



PUBLIC

SAP Data Services

Version du document : 4.2 Support Package 11 (14.2.11.0) – 2019-03-08

Guide de Designer

Contenu

1	Introduction.	19
2	Introduction à SAP Data Services.	20
3	Protection et confidentialité des données.	21
4	Présentation de ce guide.	22
4.1	A propos de ce guide.	22
4.2	Qui doit lire ce guide ?.	22
5	Interface utilisateur du Designer.	23
5.1	Zone de projet.	23
5.2	Bibliothèque d'objets locale.	24
5.3	Bibliothèque d'objets centrale.	25
	Central Object Library layout.	26
5.4	Viewing, hiding, and docking project area and object library.	27
5.5	Palette d'outils.	28
5.6	Espace de travail.	29
	Déplacement d'objets dans l'espace de travail.	30
	Mise à l'échelle de l'espace de travail.	30
	Organisation des fenêtres dans l'espace de travail.	30
	Fermeture des fenêtres dans l'espace de travail.	31
5.7	Options générales et de l'environnement.	31
	Designer : environnement.	32
	Designer : généralités.	34
	Designer : éléments graphiques.	35
	Designer Fonts options.	35
	Designer : valeurs d'attributs.	36
	Designer : connexions au référentiel central.	37
	Designer : langue.	37
	Designer : SSL.	37
	Données : généralités.	38
	Job Server : environnement.	39
	Job Server : généralités.	39
	.	39
6	Connexion au Designer.	40
6.1	Restrictions de version.	40

6.2	Réinitialisation des utilisateurs.	41
7	Mots réservés.	42
8	Objets.	43
8.1	Objets réutilisables.	44
8.2	Objets à usage unique.	46
8.3	Save reusable objects.	47
8.4	Métadonnées de l'objet.	48
8.5	Descriptions des objets.	49
8.6	Hiérarchie d'objets.	53
		55
		55
		56
8.7		57
8.8	Editeurs d'objets.	58
8.9	Création d'un objet réutilisable dans la bibliothèque d'objets.	59
8.10	Création d'un objet réutilisable à l'aide de la palette d'outils.	60
8.11	Ajout d'un objet existant.	61
8.12	Modification des noms des objets.	62
8.13	Ajout, modification et affichage de propriétés d'objet.	63
	Onglet Général.	64
	Onglet Attributs.	64
	Onglet Attributs de classe.	65
8.14	Annotations.	65
	Description des objets.	66
8.15	Création de descriptions.	66
	Ajout d'une description à un objet.	68
	Afficher une description dans l'espace de travail.	68
	Masquer la description d'un objet.	68
	Modifier les descriptions d'objets.	68
8.16	Copie d'objets.	69
8.17	Replicating objects.	70
8.18	Enregistrement et suppression des objets.	71
	Enregistrer les modifications apportées à un seul réutilisable.	72
	Enregistrer tous les objets modifiés dans le référentiel.	72
		73
	Supprimer une définition d'objet du référentiel.	73
	Supprimer un appel d'objet.	73
8.19	Recherche d'objets.	74
9	Projets et Jobs.	77

9.1	Projets.	77
	Objets constituant un projet.	77
	Création d'un projet.	78
	Ouvrir un projet existant.	78
	Enregistrer toutes les modifications apportées à un projet.	79
9.2	Jobs.	79
	Créer un job dans la zone de projet.	80
	Créer un job dans la bibliothèque d'objets.	80
	Conventions d'appellation pour objets dans jobs.	81
10	Qu'est-ce qu'une banque de données.	82
10.1	Banques de données de bases de données.	83
	Définition d'une banque de données d'une base de données.	85
	Interface de l'ordinateur central.	88
	Banques de données Amazon Redshift.	91
	Apache Impala.	92
	Banques de données Hive.	92
	HP Vertica datastore.	93
	Création d'une banque de données SAP HANA avec cryptage SSL.	95
	Création d'une banque de données Microsoft SQL Server avec cryptage SSL.	96
	Banque de données SAP Vora.	97
	Parcourir les métadonnées depuis une banque de données de base de données.	99
	Importation de métadonnées par une banque de données de base de données.	101
10.2	Banques de données en mémoire.	108
	Définir une banque de données en mémoire.	109
	Création d'une table en mémoire.	109
	Utiliser une table en mémoire comme une source ou une cible.	110
	Option Mettre à jour le schéma.	110
	Options de cible de la table en mémoire.	111
	Option Créer un ID de ligne.	111
	Dépannage des tables en mémoire.	112
10.3	Banques de données de cache persistant.	112
	Création de banques de données de cache persistant.	113
	Création de tables de cache persistant.	114
	Utilisation de tables de cache persistant comme des sources.	116
10.4	Banques de données liées.	117
	Relations entre les liens de base de données et les banques de données.	118
	Liaison d'une banque de données cible à une banque de données source à l'aide d'un lien de base de données.	119
10.5	Banques de données de l'adaptateur.	120
10.6	Banques de données de l'application.	121
10.7	Banques de données de services Web.	123

	Définition d'une banque de données de services Web.	124
	Exploration de métadonnées WSDL et WADL par une banque de données de services Web	124
10.8	Modification d'une définition de banque de données.	125
	Modification des options de banque de données.	125
	Modifier les propriétés de la banque de données.	126
10.9	Création et gestion de plusieurs configurations de banques de données.	126
	Définitions.	127
	Pourquoi utiliser plusieurs configurations de banque de données ?	128
	Création d'une nouvelle configuration de banque de données.	129
	Utilisation des alias.	130
	Création d'un alias.	130
	Fonctions pour identifier la configuration.	131
	Configuration des banques de données dans les flux de données.	132
	Solutions de portabilité.	136
	Rename Owner option.	141
	Définition d'une configuration système.	146
11	Formats de fichiers.	149
11.1	Fonctionnalités des formats de fichiers.	149
	Lecture de plusieurs fichiers en une fois.	150
	Identification des noms de fichiers sources.	151
	Ignorer les lignes avec des marqueurs spécifiés.	151
	About special markers.	152
	Threads de traitement en parallèle.	152
	Erreur de traitement pour les sources de fichier plat.	153
11.2	Création d'un nouveau format de fichier plat.	154
	Définition des attributs de colonne lors de la création d'un format de fichier.	155
	Formats des nombres.	156
	Formats de date au niveau du champ.	157
	Use a sample file to define column attributes.	158
	Modélisation d'un format de fichier sur un exemple de fichier.	159
11.3	Réplication d'un format de fichier pour la création d'un format de fichier.	160
11.4	Création d'un format de fichier à partir d'un schéma de table plate existant.	161
11.5	Créer un fichier source ou cible spécifique.	162
11.6	Edition de formats de fichier.	163
	Modifier un fichier source ou cible.	164
	Modification de plusieurs propriétés de colonne.	164
11.7	Protocoles de transfert de fichier.	165
	Transferts personnalisés de fichiers.	166
	Utilisation d'un programme de transfert personnalisé.	171

12	Format de fichier HDFS.	172
12.1	Configuration des résultats du script Pig personnalisé en tant que source.	172
13	Utilisation de format de fichier du copybook COBOL.	174
13.1	Créer un format de fichier de copybook COBOL.	174
13.2	Créer un format de fichier de copybook COBOL et un fichier de données.	175
13.3	Créer des règles pour identifier quels enregistrements représentent quels schémas.	176
13.4	Identifier le champ qui contient la longueur de l'enregistrement du schéma.	176
14	Création des formats de classeur Microsoft Excel sur les plateformes UNIX.	177
14.1	Créer un format de classeur Microsoft Excel sur UNIX.	177
15	Création de formats de fichier de journal Web.	179
15.1	Fonction word_ext.	180
15.2	Fonction concat_date_time.	180
15.3	Fonction WL_GetKeyValue.	181
16	Formats de fichier non structurés.	182
17	Objets d'emplacement de fichier.	183
17.1	Gestion des fichiers de données locales ou à distance.	184
17.2	Création d'un objet d'emplacement de fichier.	188
17.3	Obtenir une autorisation SSH.	189
17.4	À propos de l'empreinte de la clé d'hôte.	190
17.5	Modification d'un objet d'emplacement de fichier existant.	190
17.6	Créer des configurations multiples pour un objet d'emplacement de fichier.	191
	Création de configurations multiples d'objet d'emplacement de fichier.	191
17.7	Associer des objets d'emplacement de fichier à des formats de fichier.	192
	Association d'un format de fichier plat à un objet d'emplacement de fichier.	193
	Association des fichiers de schéma XML, DTD ou JSON.	194
	Association des copybooks COBOL à un objet d'emplacement de fichier.	195
	Association de classeurs Excel à un objet d'emplacement de fichier.	196
17.8	Ajout d'un objet d'emplacement de fichier à un flux de données.	197
17.9	Utilisation de scripts pour déplacer des fichiers depuis ou vers un serveur à distance.	197
	Déplacement de fichiers vers et à partir des conteneurs Azure.	198
18	Flux de données.	200
18.1	Qu'est-ce qu'un flux de données ?.	200
	Dénomination des flux de données.	200
	Exemple d'un flux de données.	201
	Étapes dans un flux de données.	201
	Les flux de données comme des étapes dans les workflows.	202
	Jeux de données intermédiaires dans un flux de données.	202
	Codes d'opération.	203

	Transfert des paramètres vers les flux de données.	203
18.2	Création et définition des flux de données.	204
	Définir un nouveau flux de données à l'aide de la bibliothèque d'objets.	204
	Définir un nouveau flux de données à l'aide de la palette d'outils.	205
	Modification des propriétés d'un flux de données.	205
18.3	Objets sources et cibles.	206
	Objets sources.	207
	Objets cibles.	207
	Ajout d'objets sources ou cibles aux flux de données.	211
	Tables de modèle.	212
	Conversion des tables de modèle en tables régulières.	213
18.4	Comprendre la propagation de colonnes.	214
	Ajout de colonnes dans un flux de données.	215
	Propager les colonnes dans un flux de données contenant une transformation Merge.	216
18.5	Tables de recherche et fonction lookup_ext.	217
	Accès à l'éditeur lookup_ext.	218
	Exemple : Définition d'une fonction lookup_ext simple.	218
	Exemple : Définition d'une fonction lookup_ext complexe.	221
18.6	Exécution d'un flux de données.	223
	Poussée vers le bas des opérations vers le serveur de base de données.	224
	Exécution de flux de données distribués.	224
	Equilibrage de charge.	225
	Caches.	226
18.7	Vue d'ensemble d'un audit de flux de données.. . . .	226
19	Transformations.	227
19.1	Ajout d'une transformation à un flux de données.	230
19.2	Editeurs de transformations.	230
19.3	Aide intégrée pour Data Services.	231
19.4	Configurations de transformation.	232
	Création d'une configuration de transformation.	233
	Ajout d'un champ défini par l'utilisateur.	234
	Editeur d'options classées.	234
	Editeurs des transformations Associate, Match et User-Defined.	235
20	Workflows.	237
20.1	Etapes d'un workflow.	237
20.2	Ordre d'exécution dans les workflows.	238
20.3	Exemple de workflow.	239
20.4	Création de workflows.	239
	Créer un workflow à l'aide de la bibliothèque d'objets.	239
	Créer un workflow à l'aide de la palette d'outils.	240

	Spécifier qu'un job exécute le workflow une seule fois.	240
	Qu'est-ce qu'un workflow unique ?.	240
	Qu'est-ce qu'un workflow continu ?.	241
20.5	Structures conditionnelles.	242
	Définir un conditionnel.	243
20.6	Boucles While.	244
	Définition d'une boucle WHILE.	245
	Utilisation d'une boucle WHILE avec la visualisation des données.	246
20.7	Blocs Try/Catch.	246
	Définition d'un bloc Try/Catch.	247
	Catégories d'exceptions disponibles.	249
	Exemple : détermination des détails d'une erreur.	249
	Meilleures pratiques Catch.	250
20.8	Scripts.	250
	Création d'un script.	251
	Débugage de scripts à l'aide de la fonction d'impression.	252
20.9	Éditeur rapide et outils fonctionnels.	252
	Accès à l'éditeur rapide.	253
	Utilisation de la liste de sélection et d'une info-bulle.	253
	Navigation vers une fonction.	254
	Recherche d'une fonction.	254
	Validation.	254
21	Données imbriquées.	256
21.1	Représentation de données hiérarchiques.	256
21.2	Mise en forme des documents XML.	259
	Spécification schéma XML.	260
	À propos de l'importation des schémas XML.	261
	Types de données abstraits.	269
	Importation de groupes de substitution.	271
	Limiter le nombre de groupes de substitution à importer.	271
	Spécification d'options de source pour les fichiers XML.	272
	Mappage de schémas facultatifs.	274
	Utilisation de Document Type Definitions (DTD).	275
	Génération de schémas DTD et XML à partir d'un schéma NRDM.	280
21.3	Opérations sur les données imbriquées.	280
	Vue d'ensemble des données imbriquées et de la transformation Query.	280
	Construction de la clause FROM.	281
	Imbrication de colonnes.	284
	Utilisation de colonnes corrélées dans des données imbriquées.	286
	Lignes d'éléments uniques et données imbriquées.	287
	Regroupement de valeurs de schémas imbriqués.	287

	Suppression de l'imbrication de données imbriquées.	287
	Transformation de niveaux inférieurs de données imbriquées.	290
21.4	Extraction et analyse XML pour les colonnes.	291
	Exemples de scénarios.	292
21.5	Extraction JSON.	298
	Extraction des données de la chaîne JSON à l'aide de la fonction <code>extract_from_json</code>	299
22	Flux de données intégrés.	300
22.1	Vue d'ensemble des flux de données intégrés.	300
22.2	Exemples de flux de données imbriquées.	300
22.3	Création de flux de données intégrés.	301
	Utilisation de l'option Créer un flux de données intégré.	302
	Création de flux de données intégrés à partir de flux existants.	303
	Utilisation des flux de données intégrés.	304
	Test séparé d'un flux de données intégré.	306
	Dépannage des flux de données intégrés.	307
23	Variables et paramètres.	308
23.1	Affichage de variables et paramètres.	310
23.2	Fenêtre Variables et paramètres.	311
23.3	Onglets Définitions et Appels.	311
23.4	Types de variables et paramètres.	313
23.5	Accès à Afficher utilisation depuis la fenêtre Variables et paramètres.	314
23.6	Utiliser des variables locales et des paramètres.	315
	Paramètres.	315
	Ajouter un paramètre à un workflow ou à un flux de données.	316
	Transférer les valeurs dans le flux de données.	316
	Définir la valeur du paramètre dans l'appel de flux.	317
	Définir une variable locale.	318
	Définir une variable locale.	318
23.7	Utiliser des variables globales	319
	Définition d'une variable globale.	319
	Répliquer une variable globale.	320
	Afficher des variables globales à partir de la boîte de dialogue Propriétés.	321
	Valeurs pour les variables globales.	321
	Définir une valeur de variable globale comme propriété de job.	322
	Définir une valeur de variable globale comme propriété d'exécution.	322
	Classement automatique de valeurs de variables globales dans un job.	323
23.8	Règles de variables locales et globales.	325
23.9	Variables d'environnement.	326
23.10	Définir des noms de fichier lors de l'exécution à l'aide de variables.	327
	Utilisation d'une variable dans un nom de fichier plat.	327

23.11	Paramètres de substitution.	328
	Présentation des paramètres de substitution.	328
	Tâches dans l'Éditeur de paramètres de substitution.	330
	Configurations des paramètres de substitution et configurations système.	335
	Association de configurations des paramètres de substitution avec des configurations système.	336
	Remplacement d'un paramètre de substitution dans l'Administrateur.	337
	Exécution d'un job avec des paramètres de substitution	338
	Exportation et importation de paramètres de substitution.	338
24	Jobs en temps réel.	340
24.1	Traitement d'un message de demande/réponse.	340
24.2	Qu'est-ce qu'un job en temps réel ?.	341
	Contenu d'un job en temps réel.	341
	Temps réel par rapport à job.	342
	Messages.	342
	Traitement des messages.	344
	Exemples de jobs en temps réel.	344
24.3	Création de jobs en temps réel.	346
	Réorganisation des métadonnées	346
	Modèles de job en temps réel.	347
	Using real-time job models.	349
	Créer un job en temps réel avec un flux de données unique.	351
24.4	Objets source et cible en temps réel.	352
	Pour afficher un schéma source ou cible de message XML.	353
	Chargement des tables cible comme transaction unique.	353
	Sources et cibles secondaires.	353
	Chargement opérationnel de tables.	354
	Astuces relatives à la conception pour flux de données dans des jobs en temps réel.	355
24.5	Testing real-time jobs.	355
	Exécution d'un job en temps réel en mode test.	355
	Utilisation de la fonction d'affichage des données.	356
	Utilisation d'une cible de fichier XML.	356
24.6	Building blocks for real-time jobs.	357
	Complément aux données de message.	357
	Création de branche de flux de données en fonction d'une valeur de cache de données.	360
	Appel de fonctions d'application.	361
24.7	Conception d'applications en temps réel.	361
	Baisse des requêtes nécessitant un accès à une application back-office.	362
	Messages de jobs en temps réel aux instances d'adaptateur.	362
	Service en temps réel appelé par une instance d'adaptateur.	362
	Démarrage et arrêt des services en temps réel.	363

25	Exécution de jobs.	364
25.1	Vue d'ensemble d'une exécution de job.	364
25.2	Preparing for job execution.	364
	Validation des jobs et des composants de jobs.	364
	S'assurer que le Job Server est en cours d'exécution.	365
	Paramétrage des options d'exécution du job.	366
25.3	Exécution des jobs comme des tâches immédiates.	370
	Exécuter un job comme une tâche immédiate.	370
	Onglet Surveillance	371
	Onglet Journal	371
25.4	Débogage des erreurs d'exécution.	372
	Utilisation des journaux.	372
	Propriétés de suivi.	375
	Examen des données cibles.	378
	Création d'un package de débogage.	378
25.5	Modification des options du Job Server.	379
	Modifier les valeurs d'option pour un Job Server individuel.	381
	Utilisation des noms de lecteurs mappés dans un chemin d'accès.	383
26	Évaluation des données.	384
26.1	Utilisation du Profileur de données.	385
	Sources de données à profiler.	386
	Connexion au serveur du profileur.	386
	Profiler statistics.	387
	Exécution d'une tâche du profileur.	390
	Surveillance des tâches du profileur à l'aide de Designer.	395
	Affichage des résultats du profileur.	396
26.2	Utilisation de la fonction d'affichage des données pour déterminer la qualité des données.	402
	Onglet Données.	402
	Onglet Profil.	403
	Onglet Profil de relation ou Profil de colonne.	404
26.3	Utilisation de la transformation Validation.	404
	Analyse du profil de colonne.	404
	Définition d'une règle de validation en fonction du profil de colonne.	406
	Editeur de règles.	407
26.4	Utilisation de l'Audit	412
	Audit d'objets dans un flux de données.	412
	Accès à la fenêtre Audit.	416
	Définition de points de vérification, de règles et des actions en cas d'échec.	417
	Indications pour sélectionner les points de vérification	420
	Audit du flux de données intégré.	420
	Résolution des étiquettes d'audit non valides.	423

	Affichage des résultats d'audit.	423
27	Qualité des données.	426
27.1	Présentation de la qualité des données.	426
27.2	Data Cleanse.	426
	A propos du nettoyage de données.	426
	Cycle de vie du package de nettoyage des données : développement, déploiement et gestion	428
	Configuration de la transformation Data Cleanse.	432
	Classement et définition de priorité des moteurs d'analyse.	433
	Données d'analyse.	433
	A propos de la normalisation des données.	447
	Affectation de descriptions de sexes et de titres de civilité.	447
	Préparation d'enregistrements pour correspondance.	448
	Données propres à une région.	449
27.3	Géocodage.	457
	Géocodage d'adresse	458
	Géocodage inversé.	462
	Recherche textuelle de POI	472
	Compréhension de la sortie.	475
	Standardisation des données de sortie Geocoder.	477
	Utilisation d'autres transformations.	478
27.4	Match.	479
	Correspondance et consolidation à l'aide de Data Services et Information Steward.	479
	Stratégies de correspondance.	480
	Composants de la correspondance.	481
	Match Wizard.	483
	Transformations pour flux de données de correspondance.	490
	Utilisation de l'éditeur de correspondance et de l'éditeur associé.	492
	Sources physiques et logiques.	493
	Match preparation.	497
	Match criteria.	520
	Post-match processing.	538
	Correspondance par association.	557
	Correspondance Unicode.	558
	Critères de correspondance phonétiques.	560
	Configuration des rapports de correspondance.	563
27.5	Address Cleanse.	565
	Fonctionnement d'Address Cleanse.	566
	Rapports Address cleanse.	568
	Préparation des données d'entrée.	569
	Détermination des transformations à utiliser.	570

	Identification du pays de destination.	573
	Configuration des fichiers de référence.	574
	Définition des options de standardisation.	576
	Traitement des adresses japonaises.	577
	Traitement des adresses chinoises.	587
	Pays pris en charge (Global Address Cleanse).	591
	Certification Nouvelle-Zélande.	594
	Option de la liste de propositions de Global Address Cleanse.	597
	Transformation Global Suggestion List.	598
27.6	Au-delà du nettoyage d'adresse de base.	599
	USPS DPV®.	599
	LACSLink.	609
	SuiteLink™.	618
	USPS DSF2®.	621
	Présentation de NCOALink®.	631
	USPS eLOT®.	647
	Early Warning System (EWS).	648
	USPS RDI®.	649
	GeoCensus (USA Regulatory Address Cleanse).	652
	Z4Change (USA Regulatory Address Cleanse).	656
	Présentation des listes de propositions.	657
	Rapport de statistiques de sources de données multiples.	660
27.7	Tables de référentiel pour les statistiques de qualité de données.	665
27.8	Prise en charge de Data Quality pour les types de données natives.	665
	Définitions du type de données Qualité des données.	666
27.9	Data Quality support pour valeurs nulles.	666
28	Query transform.	667
28.1	Entrée de schéma et Sortie de schéma.	667
28.2	Schéma d'entrée de transformation Query.	669
28.3	Schéma de sortie de transformation Query.	671
28.4	Commandes de schémas de sortie.	672
28.5	Recherche dans un schéma d'entrée ou de sortie.	674
29	Fonctions et procédures.	676
29.1	A propos des fonctions.	676
	Fonctions comparées aux transformations.	677
	Opération d'une fonction.	677
	Arithmétique dans les fonctions de dates.	678
	Insertion de fonctions dans les expressions.	678
	Fonctions personnalisées.	681
	Fonctions intégrées.	685

29.2	À propos des procédures.	690
	Exigences.	691
	Création de procédures stockées dans une base de données.	692
	Importation de métadonnées pour des procédures stockées.	697
	Structure d'une procédure stockée.	698
	Appel de procédures stockées.	699
	Vérification du statut d'exécution.	704
30	Langage de script.	706
30.1	Utilisation de la langue de script.	706
30.2	Syntaxe du langage.	707
	Syntaxe des instructions de scripts.	707
	Syntaxe des références de colonnes et de tables dans les expressions.	708
	Chaînes.	708
	Variables.	710
	Interpolation de variables.	710
	Fonctions et procédures stockées.	711
	Valeurs nulles.	712
	Debogage et validation.	714
30.3	Modèles de scripts.	715
	Fonction Square.	715
	Fonction RepeatString.	715
31	Fonctionnement de l'extraction.	717
31.1	Utilisation de la transformation Entity Extraction.	718
31.2	Entités, types, sous-types et faits.	718
31.3	Différences entre traitement des données de texte et transformations Data Cleanse.	720
32	Conception et débogage.	722
32.1	Utilisation de Afficher les cas d'emploi.	722
	Accès à Afficher utilisation depuis la bibliothèque d'objets.	723
	Accès à Afficher les cas d'emploi depuis l'espace de travail.	726
	Limitations.	726
32.2	Utilisation de l'affichage des données.	727
	Accessing View Data.	728
	Affichage des données dans l'espace de travail.	728
	Propriétés de l'affichage des données.	730
	Options de la barre d'outils d'affichage des données.	734
	Onglets d'affichage des données.	735
32.3	Utilisation de l'affichage des données lors de la conception.	738
	Affichage des données lors de la conception.	739
	Configuration de l'affichage des données lors de la conception.	740

	Définition de variables pour les expressions.	740
32.4	Utilisation du débogueur interactif.	741
	Avant de lancer le débogueur interactif.	741
	Démarrage et arrêt du débogueur interactif.	744
	Volets.	746
	Options et barre d'outils du menu Débogage.	751
	Affichage des données passées par les transformations.	753
	Optimiseur de poussée vers le bas.	754
	Restrictions.	754
32.5	Comparaison d'objets.	755
	Comparer deux objets différents.	755
	Comparaison de deux versions du même objet.	756
	Présentation de la fenêtre Affichage des différences.	756
	Navigation dans les différences.	760
32.6	Calcul des mappages de colonnes.	761
	Calculer automatiquement les mappages de colonnes.	761
	Calculer automatiquement les mappages de colonnes.	761
32.7	Contournement de workflows et flux de données spécifiques.	762
	Contournement d'un flux de données simple ou d'un workflow unique.	763
	Contournement de flux de données ou workflows multiples.	763
	Désactivation du contournement.	763
33	Mécanismes de récupération.	765
33.1	Récupération suite à un échec d'exécution de job.	765
33.2	Récupération automatique de jobs.	766
	Activation de la récupération automatique.	766
	Marquage des unités de récupération.	766
	Exécution en mode de récupération.	767
	Garantie d'un chemin d'exécution correct.	768
	Utilisation de blocs Try/Catch avec récupération automatique.	769
	Garantie que les données ne sont pas dupliquées dans des cibles.	771
	Utilisation de SQL de préchargement pour autoriser des flux de données ré-exécutables	772
33.3	Récupération manuelle de jobs à l'aide de tables de statuts.	774
33.4	Traitement de données avec problèmes.	774
	Utilisation de fichiers de dépassement.	775
	Filtrage pour valeurs manquantes ou erronées	775
	Traitement de faits avec dimensions manquantes.	776
33.5	Echange de métadonnées.	777
	Echange de métadonnées.	777
33.6	Chargement du fichier de données volumineux avec option de récupération.	779
	Activation de l'option de récupération pour le chargement de données volumineuses.	779
	Limitations.	780

34	Capture de données modifiées.	782
34.1	Actualisation complète.	782
34.2	Capture des modifications uniquement.	782
34.3	CDM basée sur la source et basée sur la cible.	783
	CDM basée sur la source.	783
	CDM basée sur la cible.	784
34.4	Utilisation de CDC avec des sources Oracle.	784
	Présentation de la CDM pour les bases de données Oracle.	784
	Configuration de la CDM Oracle.	788
	Création d'une banque de données CDM pour Oracle.	789
	Importation de données CDC dans des tables	790
	Affichage d'une table CDM importée.	792
	Configuration d'une table source CDM Oracle.	794
	Création d'un flux de données avec une source CDM Oracle.	796
	Maintaining CDC tables and subscriptions.	797
	Restrictions.	798
34.5	Utilisation de la CDM avec les sources d'un ordinateur central Attunity.	798
	Configuration de la CDM Attunity.	799
	Configuration du progiciel pour la CDM sur les sources d'ordinateur central.	800
	Importation des données CDM de l'ordinateur central.	802
	Configuration d'une source CDM d'ordinateur central.	803
	Utilisation des points de contrôle de l'ordinateur central.	804
	Restrictions.	805
34.6	Utilisation de la CDM avec SAP Replication Server.	805
	Aperçu de l'utilisation d'un workflow continu et de fonctions	806
	Présentation de la méthode de modélisation SAP PowerAMC.	811
34.7	Utilisation de la CDM avec les bases de données Microsoft SQL Server.	821
	Restrictions.	823
	Colonnes de SAP BusinessObjects Data Services.	823
	Méthode de capture des données modifiées (CDM).	824
	Méthode de suivi des modifications.	826
	Méthode du serveur de réplication.	828
34.8	Utilisation de la CDM avec des sources basées sur l'horodatage.	835
	Traitement des horodatages.	836
	Chevauchements.	838
	Types d'horodatages.	843
	Timestamp-based CDC examples.	845
	Autres astuces de conception de job.	850
34.9	Utilisation de la CDM pour les cibles.	851
35	Data Services data types.	852

36	Conversion de types de données.	855
36.1	Calcul de date.	855
36.2	Types de données non pris en charge.	856
36.3	Conversion de/en types de données internes.	856
	Flux Attunity.	857
	Copybook Cobol.	858
	Data Federator.	860
	IBM DB2.	860
	Informix.	863
	Microsoft Excel.	864
	Microsoft SQL Server.	865
	MySQL.	868
	Netezza.	870
	ODBC.	871
	Oracle.	873
	SQL Anywhere.	875
	SAP ASE.	877
	SAP Sybase IQ.	879
	Teradata.	881
36.4	Conversion de types de données dans des expressions.	883
36.5	Conversion parmi des types de données Number.	884
	Division algorithm.	886
	Conversion of decimals with different scale and precision.	886
	Conversion between string and number data types.	887
	Conversion between Oracle string and date data types.	887
36.6	Conversions entre types de données explicites.	888
36.7	Conversion entre des types de données natifs.	888
	date.	889
	datetime.	889
	decimal.	890
	double.	891
	int (entier).	892
	caractères (varchar).	892
37	Surveillance des Jobs.	894
37.1	Administrateur.	894
38	Développement multi-utilisateurs.	895
38.1	Référentiel central et référentiel local.	895
38.2	Utilisateurs multiples.	897
38.3	Sécurité et référentiel central.	900
38.4	Configuration d'environnement multi-utilisateurs.	900

	Création d'un référentiel central non sécurisé.	901
	Définition d'une connexion à un référentiel central non sécurisé.	901
	Activation d'un référentiel central.	902
38.5	Implémentation de la sécurité du référentiel central.	904
	Présentation.	904
	Création d'un référentiel central sécurisé.	906
	Ajout d'un administrateur multi-utilisateurs (facultatif).	907
	Configuration des groupes et utilisateurs.	907
	Définition d'une connexion vers un référentiel central sécurisé.	908
	Utilisation d'objets dans un référentiel central sécurisé.	908
38.6	Travail dans un environnement multi-utilisateurs.	909
	Filtrage.	910
	Ajout d'objets au référentiel central.	911
	Extraction d'objets.	913
	Annulation de l'extraction.	916
	Rechargement d'objets.	917
	Etiquetage d'objets.	919
	Obtention d'objets.	922
	Comparaison d'objets.	924
	Affichage de l'historique de l'objet.	924
	Suppression d'objets.	927
38.7	Migration de jobs multi-utilisateurs.	927
	Gestion des phases des applications.	928
	Copie de contenu entre référentiels centraux.	929
	Migration du référentiel central.	930

1 Introduction

2 Introduction à SAP Data Services

SAP Data Services offre une solution d'entreprise unique pour l'intégration des données, la qualité des données, le profilage des données et le traitement des données texte.

Les entreprises peuvent utiliser Data Services pour intégrer, transformer, améliorer et fournir des données sécurisées aux processus de gestion critiques. Les entreprises du secteur informatique peuvent dépendre de Data Services pour une efficacité opérationnelle optimale, pour améliorer la qualité des données et accéder à des sources et des applications hétérogènes. Data Services fournit toutes ces fonctionnalités grâce à :

- Une interface utilisateur de développement unique
- Un référentiel des métadonnées
- Une couche de connectivité de données
- Un environnement d'exécution
- Une console de gestion

3 Protection et confidentialité des données

SAP fournit des fonctionnalités et fonctions spécifiques pour assurer la conformité par rapport à la législation en vigueur, notamment la protection des données.

La protection des données est associée à de nombreuses obligations légales et préoccupations de confidentialité. Outre la conformité avec les réglementations applicables en matière de confidentialité des données, SAP doit prendre en compte la conformité avec la législation propre à un secteur dans différents pays. SAP n'indique pas si les fonctionnalités et fonctions fournies constituent la meilleure méthode pour prendre en charge les exigences de l'entreprise, du secteur, d'une région ou d'un pays. En outre, les informations fournies par SAP ne prodiguent aucun conseil ni aucune recommandation quant aux fonctionnalités supplémentaires requises dans des environnements informatiques particuliers. Prenez vos décisions relatives à la protection des données au cas par cas, en tenant compte de votre infrastructure système et de la législation en vigueur.

i Remarque

Dans la plupart des cas, la conformité de SAP par rapport aux lois applicables en matière de confidentialité et de protection des données n'est pas couverte par une fonctionnalité du produit. Le logiciel SAP prend en charge la conformité de la protection des données en fournissant des fonctionnalités de sécurité et des fonctions spécifiques pertinentes en matière de protection des données, comme le blocage et la suppression simplifiés des données personnelles. SAP ne fournit aucun conseil juridique, sous aucune forme. Dans ce document, SAP utilise des définitions et autres termes ne provenant d'aucune source légale.

Notez que le logiciel SAP place les données que vous indiquez dans des journaux de suivi, exemples de rapports et référentiels (données aux conséquences néfastes), etc. En d'autres termes, vos données sont stockées dans des emplacements autres que les fichiers de sortie. Il est de votre responsabilité de supprimer ces données.

Voici quelques exemples indiquant où Data Services utilise les données client :

- Lorsque vous activez l'option [Suivre toutes les données](#), Data services imprime les données dans certains fichiers journaux.
- Si un job utilisant la fonctionnalité de chargement par lots échoue, Data Services enregistre des fichiers contenant des données client dans le répertoire de chargement par lots, afin que vous puissiez vérifier et analyser les données. L'emplacement de chargement par lots par défaut est %DS_COMMON_DIR%/log/BulkLoader.
- Les fonctions smtp_to et d'expéditeur utilisent des numéros d'ID d'e-mail comme paramètres d'entrée. Data Services enregistre les numéros d'ID d'e-mail dans le référentiel Data Services.
- Data Services place un échantillon de données client dans le référentiel Data Services lorsque l'option "Complet" est activée lors de la génération de données aux conséquences néfastes pour Global Address Cleanse (GAC) et Universal Data Cleanse (UDC).
- L'exemple de rapport GAC contient un échantillon de données d'adresse de l'utilisateur.

4 Présentation de ce guide

SAP Data Services Designer fournit un environnement de développement d'interface utilisateur graphique dans lequel vous définissez une logique d'application des données permettant d'extraire, de transformer et de charger des données de la source à la cible.

Bienvenue dans le *Guide Designer*. Ce guide décrit toutes les fonctionnalités du GUI Designer et comment vous pouvez utiliser les nombreuses fonctionnalités de manipulation de différents types de données à partir de systèmes traditionnels de gestion de base de données, Big Data et données du Cloud.

[A propos de ce guide \[page 22\]](#)

[Qui doit lire ce guide ? \[page 22\]](#)

4.1 A propos de ce guide

Le guide contient deux types d'informations :

- Des informations conceptuelles qui aident à comprendre Data Services Designer et son fonctionnement
- Des informations procédurales qui expliquent étape par étape comment accomplir une tâche

Vous trouverez ce guide très utile :

- En étudiant le produit
- En accomplissant les tâches de la phase de conception et de tests précoces de vos projets de déplacement de données.
- Comme source générale d'information durant toutes les phases de vos projets

4.2 Qui doit lire ce guide ?

Cette documentation ainsi que celles d'autres produits Data Services repose sur les principes suivants :

- Vous êtes développeur d'applications, consultant ou administrateur de base de données travaillant sur l'extraction, l'entreposage, l'intégration ou la qualité des données.
- Vous comprenez les concepts de vos systèmes de sources de données, de RDBMS, de Business Intelligence et de messagerie.
- Vous comprenez les besoins en matière de données de votre organisation.
- Vous connaissez bien SQL (Structured Query Language).
- Si vous désirez utiliser ce produit pour concevoir un traitement en temps réel, vous devez bien connaître :
 - les formats de schéma DTD et XML pour les fichiers XML ;
 - les services Web de publication (protocoles WSDL, REST, HTTP, HTTPS et SOAP, etc.).
- Vous connaissez bien les environnements d'installation de Data Services : Microsoft Windows ou UNIX.

5 Interface utilisateur du Designer

Cette section présente l'interface graphique utilisateur de Designer.

[Zone de projet \[page 23\]](#)

[Bibliothèque d'objets locale \[page 24\]](#)

La bibliothèque d'objets locale donne accès aux objets réutilisables.

[Bibliothèque d'objets centrale \[page 25\]](#)

La bibliothèque d'objets centrale fait office de contrôle de source pour gérer les modifications apportées aux objets dans un environnement avec plusieurs utilisateurs.

[Viewing, hiding, and docking project area and object library \[page 27\]](#)

View, hide, or auto hide the project area and object library, or move the panes to new locations on the Designer screen.

[Palette d'outils \[page 28\]](#)

[Espace de travail \[page 29\]](#)

L'espace de travail comporte un emplacement qui permet de manipuler les objets système et d'assembler les processus de déplacement de données en mode graphique.

[Options générales et de l'environnement \[page 31\]](#)

Affiche les groupes d'options relatifs à Designer, aux données et aux options du Job Server.

5.1 Zone de projet

La zone de projet affiche les objets utilisés dans chaque projet de façon hiérarchique. Les onglets disponibles dans la partie inférieure permettent d'effectuer différentes tâches. Ces onglets sont les suivants :



Créer, afficher et gérer des projets. Propose une vue hiérarchique de tous les objets utilisés dans chaque projet.



Afficher le statut des jobs en cours d'exécution. Lorsque vous sélectionnez une exécution de job spécifique, son statut s'affiche, indiquant les étapes qui sont terminées et celles qui sont en cours d'exécution. Ces tâches peuvent également être réalisées à l'aide de l'Administrateur.



Afficher l'historique des jobs terminés. Les journaux peuvent également être consultés à l'aide de l'Administrateur.

Pour gérer l'emplacement de la zone de projet, cliquez avec le bouton droit de la souris sur sa bordure grise et sélectionnez/désélectionnez [Ancrage](#) ou bien sélectionnez [Masquer](#) dans le menu.

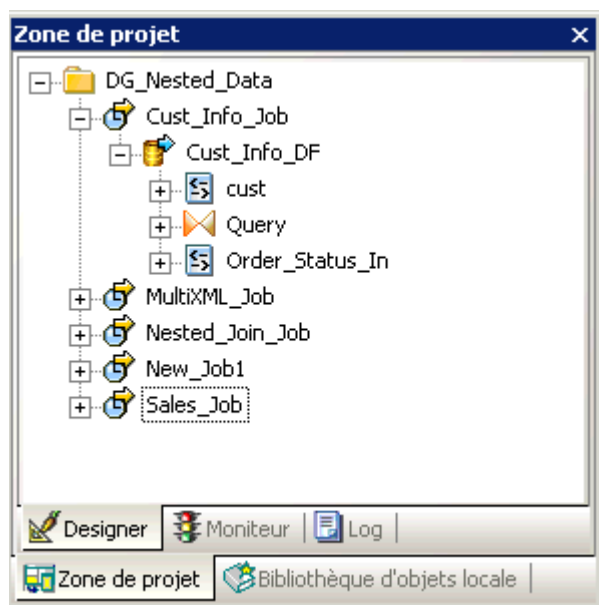
- Lorsque vous sélectionnez [Ancrage](#), vous pouvez cliquer sur la zone de projet et la faire glisser pour l'ancrer à un bord de la fenêtre de Designer ou l'en détacher. Lorsque vous faites glisser la zone de projet pour l'éloigner d'un bord de la fenêtre de Designer, elle reste détachée. Pour basculer rapidement entre vos

derniers emplacements avec ancrage et sans ancrage, il vous suffit de cliquer deux fois sur la bordure grise.

Lorsque vous désélectionnez **Ancrage**, vous pouvez cliquer sur la zone de projet et la faire glisser vers n'importe quel emplacement de l'écran sans qu'elle ne soit ancrée dans la fenêtre de Designer.

- Lorsque vous sélectionnez **Masquer**, la zone de projet n'est plus visible dans la fenêtre de Designer. Pour afficher à nouveau la zone de projet, cliquez sur son icône dans la barre d'outils.

Voici un exemple de l'onglet **Designer** de la fenêtre de projet, qui affiche la hiérarchie du projet :



Lorsque vous descendez dans la hiérarchie des objets au niveau de l'espace de travail de Designer, votre position dans la hiérarchie du projet est mise en surbrillance dans la fenêtre.

5.2 Bibliothèque d'objets locale

La bibliothèque d'objets locale donne accès aux objets réutilisables.

Il peut aussi bien s'agir d'objets système intégrés, comme des transformations, que des objets que vous créez et enregistrez, comme les banques de données, jobs, flux de données et workflows.

La bibliothèque d'objets locale se présente sous la forme d'une fenêtre dans votre référentiel local. Elle vous évite d'accéder directement au référentiel. Les mises à jour du référentiel interviennent dans le cadre du fonctionnement normal du progiciel. Lorsque vous enregistrez les objets que vous créez, ils sont ajoutés au référentiel. Les objets enregistrés sont accessibles par le biais de la bibliothèque d'objets locale.

Pour gérer l'emplacement de la bibliothèque d'objets, cliquez avec le bouton droit sur sa bordure grise et sélectionnez/désélectionnez **Ancrage** ou bien sélectionnez **Masquer** dans le menu.

- Lorsque vous sélectionnez **Ancrage**, vous pouvez cliquer sur la bibliothèque d'objets et la faire glisser pour l'ancrer à un bord de la fenêtre de Designer ou l'en détacher. Lorsque vous faites glisser la bibliothèque d'objets pour l'éloigner d'un bord de la fenêtre de Designer, elle reste détachée. Pour basculer rapidement entre vos derniers emplacements avec ancrage et sans ancrage, il vous suffit de cliquer deux fois sur la bordure grise.

Lorsque vous désélectionnez [Ancrage](#), vous pouvez cliquer sur la bibliothèque d'objets et la faire glisser vers n'importe quel emplacement de l'écran sans qu'elle soit ancrée dans la fenêtre de Designer.

- Lorsque vous sélectionnez [Masquer](#), la bibliothèque d'objets n'est plus visible dans la fenêtre de Designer. Pour afficher à nouveau la bibliothèque d'objets, cliquez sur son icône dans la barre d'outils.

Pour plus d'informations sur le référentiel central et référentiel local lors d'un développement multi-utilisateurs, voir le *Guide de Designer*.

5.3 Bibliothèque d'objets centrale

La bibliothèque d'objets centrale fait office de contrôle de source pour gérer les modifications apportées aux objets dans un environnement avec plusieurs utilisateurs.

Par le biais de la bibliothèque d'objets centrale, les utilisateurs autorisés accèdent à un référentiel de bibliothèque qui contient des versions d'objets enregistrés, dans leurs référentiels locaux. La bibliothèque d'objets centrale permet de gérer les personnes qui peuvent consulter les objets stockés dans le référentiel central et qui peuvent modifier ces objets.

Les utilisateurs doivent faire partie d'un groupe d'utilisateurs autorisé à exécuter des tâches dans la bibliothèque d'objets centrale. Les administrateurs peuvent affecter des autorisations à un groupe entier d'utilisateurs et affecter différents niveaux d'autorisations aux utilisateurs d'un groupe.

❖ Exemple

Tous les utilisateurs du groupe A disposent d'un accès en lecture à la bibliothèque d'objets centrale. Dans le groupe B, un utilisateur peut copier des objets dans la bibliothèque d'objets locale à partir de la bibliothèque d'objets centrale, mais ne peut pas extraire et archiver des objets. En d'autres termes, il ne peut pas mettre à jour des objets dans la bibliothèque d'objets centrale, avec des modifications apportées localement. Un autre utilisateur du groupe B dispose de tous les droits d'accès et peut donc extraire des objets, modifier ces objets, archiver les modifications apportées à la bibliothèque d'objets centrale et télécharger de nouveaux objets dans la bibliothèque d'objets centrale.

Dans Designer, les utilisateurs vérifient un objet du référentiel central à l'aide de la bibliothèque d'objets centrale. Après la vérification, aucun autre utilisateur ne peut travailler sur cet objet jusqu'à ce qu'il soit à nouveau contrôlé dans le référentiel central. Les autres utilisateurs peuvent modifier leur copie locale de l'objet, sans incidence sur la version dans le référentiel central.

Pour contrôler l'emplacement de la bibliothèque d'objets centrale dans la fenêtre de [Designer](#), cliquez avec le bouton droit sur la bordure et sélectionnez [Ancrage](#), [Variable](#) ou [Masquer](#) dans le menu déroulant.

- Si vous sélectionnez [Variable](#), faites glisser la bibliothèque d'objets centrale à un endroit quelconque de la fenêtre de Designer.
- Lorsque vous sélectionnez [Ancrage](#), faites glisser la bibliothèque d'objets centrale sur l'une des icônes de position qui s'affiche. Chaque icône affiche un poste : en haut, en bas, à droite et à gauche. Lorsque vous disposez la bibliothèque d'objets centrale sur une icône, la bibliothèque est ancrée à la position que vous avez sélectionnée. Faites par exemple glisser la bibliothèque sur l'icône à gauche pour l'ancrer côté gauche de la fenêtre de [Designer](#).
- Lorsque vous faites glisser la bibliothèque d'objets centrale hors de la bordure de la fenêtre de [Designer](#), la bibliothèque reste détachée.

- Pour basculer rapidement entre vos derniers emplacements avec ancrage et sans ancrage, cliquez deux fois sur la bordure grise.
- Sélectionnez *Masquer* pour que la bibliothèque ne s'affiche pas dans la fenêtre de Designer.

Mise en forme de la bibliothèque d'objets centrale

Les informations de la bibliothèque d'objets centrale diffèrent de celles d'une bibliothèque d'objets locale. Il existe plusieurs icônes situées en haut de la boîte de dialogue pour réaliser les tâches suivantes :

- Sélectionner divers types d'extraction d'objet
- Afficher l'historique
- Modifier la connexion au référentiel central
- Actualiser le contenu de la bibliothèque d'objets centrale

La partie supérieure affiche également le groupe d'utilisateurs actuel auquel vous appartenez et le nom du référentiel central.

Le volet principal de la bibliothèque d'objets centrale contient une liste des types d'objet en fonction de l'onglet inférieur que vous sélectionnez.

Lorsqu'un utilisateur extrait un objet, une coche rouge apparaît sur le nom de l'objet dans la colonne de gauche.

Le volet principal contient plusieurs colonnes avec des informations pour chaque objet :

- *Utilisateur à l'origine de l'extraction* : Le nom de l'utilisateur qui a actuellement extrait l'objet de la bibliothèque, ou champ non renseigné si l'objet n'est pas extrait.
- *Référentiel d'extraction* : Le nom du référentiel dans lequel l'objet est extrait, ou champ non renseigné si l'objet n'est pas extrait.
- *Autorisation* : Le type d'autorisation pour le groupe qui s'affiche en haut dans la zone *Autorisation du groupe*. Si vous ajoutez un nouvel objet à la bibliothèque centrale, le groupe actuel obtient les autorisations COMPLÈTES et tous les autres groupes obtiennent l'autorisation de LECTURE.
- *Version la plus récente* : Un numéro de version et un horodatage indiquant quand le logiciel a enregistré cette version de l'objet.
- *Description* : Informations sur l'objet qui ont été saisies lorsque l'objet a été ajouté à la bibliothèque.

Informations associées

[Bibliothèque d'objets locale \[page 24\]](#)

[Développement multi-utilisateurs \[page 895\]](#)

[Référentiel central et référentiel local \[page 895\]](#)

[Travail dans un environnement multi-utilisateurs \[page 909\]](#)

5.3.1 Central Object Library layout

The information in the central object library is different from the local object library.

In the central object library, there are several icons located at the top of the pane to perform the following tasks:

- Select various object checkout types
- Show history
- Edit central repository connection
- Refresh the content of the central object library

The top of the pane also displays the current user group to which you belong and the name of the central repository.

The main area of the central object library contains:

- A list of the object types based on the lower tab that you choose.
- When you check out an object, a red check mark appears over the object name in the left column.

The main area also contains several columns with information for each object:

- *Check out user*: The name of the user who currently has the object checked out of the library, or blank when the object is not checked out.
- *Check out repository*: The name of the repository that contains the checked-out object, or blank when the object is not checked out.
- *Permission*: The authorization type for the group that appears in the *Group Permission* box at the top. When you add a new object to the central object library, the current group gets FULL permission to the object and all other groups get READ permission.
- *Latest version*: A version number and a timestamp that indicate when the software saved this version of the object.
- *Description*: Information about the object that you entered when you added the object to the library.

5.4 Viewing, hiding, and docking project area and object library

View, hide, or auto hide the project area and object library, or move the panes to new locations on the Designer screen.

The project area and object library panes appear in the left side of Designer by default. Perform the following steps to hide, auto hide, view, or dock and undock the panes:

1. To hide a pane:
 - a. Right-click the gray border around the pane.
 - b. Select *Hide* from the popup menu.
2. To auto hide a pane:
 - a. Right-click the gray border around the pane.
 - b. Select *Auto Hide* from the popup menu.

The software moves the pane header to the nearest border of Designer. Click the header to unhide the pane. Click the push pin icon in the upper right corner of the pane to rehide.

3. To view the panes:
 - a. Select  *Tools* > *Project Area*. 
 - b. Select  *Tools* > *Object Library*. 

4. To undock, move, and dock the panes:
 - a. Click the gray border around the pane and drag the pane.
 - b. Drag the pane until your mouse pointer reaches one of the positional arrow icons that appear as you drag the pane.

Positional arrows appear on screen while you drag the pane. Each arrow indicates a position on the Designer screen to automatically dock the pane to that position.
 - c. Double-click the pane border to move to the last position whether it is a docked or undocked position.

Play around with the panes to learn the various ways to hide, unhide, and dock the panes. To go back to the default layout, select ► [View](#) ► [Apply Default Layout](#). ►

5.5 Palette d'outils

La palette d'outils est une fenêtre distincte qui apparaît par défaut sur le bord droit de l'espace de travail de Designer. Vous pouvez déplacer la palette d'outils à l'emplacement de votre choix sur l'écran ou l'ancrer à un bord de la fenêtre de Designer.

Les icônes disponibles dans la palette d'outils permettent de créer des objets dans l'espace de travail. Celles qui ne peuvent pas être utilisées sur le diagramme ouvert dans l'espace de travail sont désactivées.

Pour afficher le nom d'une icône, placez le curseur dessus jusqu'à ce que l'info-bulle de l'icône apparaisse.

Lorsque vous créez un objet à l'aide de la palette d'outils, vous créez une définition d'un objet. Si un nouvel objet est réutilisable, il sera automatiquement disponible dans la bibliothèque d'objets une fois que l'aurez créé.

Si vous sélectionnez l'icône de flux de données dans la palette d'outils et que vous définissez un nouveau flux de données, vous pourrez ultérieurement extraire ce flux de données de la bibliothèque d'objets et ajouter un appel à la définition existante, par exemple.

La palette d'outils contient les icônes suivantes :

Icône	Outil	Description (classe)	Disponible
	Pointeur	Modifie le pointeur d'outil en pointeur de sélection permettant de sélectionner et déplacer des objets dans un diagramme.	Partout
	Workflow	Crée un workflow. (réutilisable)	Jobs et workflows
	Flux de données	Crée un flux de données. (réutilisable)	Jobs et workflows
	Flux de données ABAP	Uniquement utilisé avec l'application SAP.	
	Transformation Query	Crée un modèle pour une requête. A utiliser pour définir les mappages de colonnes et les sélections de lignes (usage unique)	Flux de données

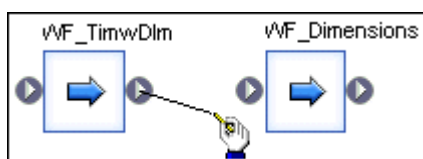
Icône	Outil	Description (classe)	Disponible
	Modèle de table	Crée une table pour une cible. (usage unique)	Flux de données
	Modèle de schémas imbriqués	Crée un modèle JSON ou XML. (usage unique)	Flux de données
	Transport de données	Uniquement utilisé avec l'application SAP.	
	Script	Crée un objet de script. (usage unique)	Jobs et workflows
	Structure conditionnelle	Crée un objet conditionnel. (usage unique)	Jobs et workflows
	Try	Crée un objet try. (usage unique)	Jobs et workflows
	Catch	Crée un objet Catch. (usage unique)	Jobs et workflows
	Annotation	Crée une annotation. (usage unique)	Jobs, workflows et flux de données

5.6 Espace de travail

L'espace de travail comporte un emplacement qui permet de manipuler les objets système et d'assembler les processus de déplacement de données en mode graphique.

Lorsque vous ouvrez ou sélectionnez un job ou un flux quelconque dans une hiérarchie de jobs, l'espace de travail devient "actif" avec votre sélection.

Ces processus sont représentés par des icônes que vous pouvez glisser-déposer dans un espace de travail pour créer un diagramme d'espace de travail. Ce diagramme offre une représentation visuelle d'une application de déplacement de données complète ou partielle.



[Déplacement d'objets dans l'espace de travail \[page 30\]](#)

Pour déplacer les objets dans l'espace de travail, utilisez les commandes de souris standard.

[Mise à l'échelle de l'espace de travail \[page 30\]](#)

La mise à l'échelle de l'espace de travail vous permet de modifier la vue d'un job, d'un workflow ou d'un flux de données.

[Organisation des fenêtres dans l'espace de travail \[page 30\]](#)

Le menu Fenêtre permet d'organiser les différentes fenêtres ouvertes dans l'espace de travail.

[Fermeture des fenêtres dans l'espace de travail \[page 31\]](#)

Lorsque vous explorez un objet dans la zone de projet ou l'espace de travail, une vue de la définition d'objet s'affiche dans l'espace de travail. Ces vues utilisent des ressources système et doivent être fermées lorsqu'elles ne sont pas utilisées.

5.6.1 Déplacement d'objets dans l'espace de travail

Pour déplacer les objets dans l'espace de travail, utilisez les commandes de souris standard.

Pour déplacer un objet vers un autre emplacement de l'espace de travail, procédez comme suit :

1. Cliquez sur l'objet pour le sélectionner.
2. Faites glisser l'objet jusqu'à l'emplacement souhaité de l'espace de travail.

5.6.2 Mise à l'échelle de l'espace de travail

La mise à l'échelle de l'espace de travail vous permet de modifier la vue d'un job, d'un workflow ou d'un flux de données.

Vous pouvez, par exemple, augmenter l'échelle pour examiner une partie définie d'un workflow ou réduire l'échelle pour pouvoir examiner l'ensemble d'un workflow sans devoir opérer de défilement.

Il existe plusieurs manières de modifier l'échelle de l'espace de travail :

- Dans la liste déroulante au niveau de la barre d'outils, sélectionnez une échelle prédéfinie ou saisissez une valeur personnalisée (100 %, par exemple).
- Cliquez avec le bouton droit dans l'espace de travail et sélectionnez l'échelle souhaitée.
- Sélectionnez [Mettre à l'échelle](#). Le Designer calcule l'échelle pertinente pour le projet complet dans la zone d'affichage en cours.
- Si vous sélectionnez [Mise à l'échelle globale](#), l'espace de travail est visible en intégralité dans la zone d'affichage en cours.

5.6.3 Organisation des fenêtres dans l'espace de travail

Le menu Fenêtre permet d'organiser les différentes fenêtres ouvertes dans l'espace de travail.

Vous pouvez organiser les fenêtres des façons suivantes :

- cascade
- mosaïque horizontale
- mosaïque verticale

5.6.4 Fermeture des fenêtres dans l'espace de travail

Lorsque vous explorez un objet dans la zone de projet ou l'espace de travail, une vue de la définition d'objet s'affiche dans l'espace de travail. Ces vues utilisent des ressources système et doivent être fermées lorsqu'elles ne sont pas utilisées.

Cette vue est repérable grâce à un onglet en bas de l'espace de travail ; autrement dit, plus vous ouvrez d'objets dans l'espace de travail, plus les onglets sont nombreux (vous pouvez afficher ou masquer ces onglets via le menu ► *Outils* ► *Options* ►. Accédez aux options ► *Designer* ► *Général* ► et sélectionnez/désélectionnez *Afficher les onglets dans l'espace de travail*).

i Remarque

Ces vues utilisent des ressources système. Si vous avez un grand nombre de vues ouvert, vous risquez de constater une dégradation des performances.

Fermez les vues une par une en cliquant sur la case de fermeture en haut à droite de l'espace de travail. Pour fermer toutes les vues ouvertes, sélectionnez ► *Fenêtre* ► *Fermer toutes les fenêtres* ► ou cliquez sur l'icône *Fermer toutes les fenêtres* dans la barre d'outils.

Informations associées

[Options générales et de l'environnement \[page 31\]](#)

5.7 Options générales et de l'environnement

Affiche les groupes d'options relatifs à Designer, aux données et aux options du Job Server.

Pour ouvrir la fenêtre *Options*, sélectionnez ► *Outils* ► *Options* ►.

Pour développer les options, cliquez sur le signe plus. Lorsque vous sélectionnez une option ou un groupe d'options, une description s'affiche sur la droite.

[Designer : environnement \[page 32\]](#)

Vous pouvez modifier les paramètres par défaut pour le reporting de métadonnées et les job servers, ainsi que les paramètres de port de communication pour le Designer.

[Designer : généralités \[page 34\]](#)

Vous pouvez définir les paramètres par défaut pour les options couramment utilisées dans le Designer.

[Designer : éléments graphiques \[page 35\]](#)

Sélectionnez et prévisualisez les éléments de style afin de personnaliser vos espaces de travail. Ces options permettent d'établir une distinction claire entre l'espace de travail de conception de jobs/workflows et l'espace de travail de conception de flux de données.

[Designer Fonts options \[page 35\]](#)

Use the *Font* options to set font type, style, and size for text in the workspace pane.

[Designer : valeurs d'attributs \[page 36\]](#)

Lors d'une réimportation, Data Services conserve ou efface l'ancienne valeur pour certains attributs de table et de colonne. Vous pouvez contrôler les actions effectuées par Data Services lors d'une réimportation pour certains attributs.

[Designer : connexions au référentiel central \[page 37\]](#)

Le référentiel central fournit une bibliothèque d'objets partagée permettant aux développeurs d'enregistrer des objets dans leurs référentiels locaux et de les en extraire. Utilisez ces options pour définir les paramètres de connexion par défaut.

[Designer : langue \[page 37\]](#)

Ces options fournissent des paramètres régionaux dans une langue autre que l'anglais (paramètre par défaut) pour l'affichage de l'interface utilisateur Data Services et de tout texte généré dans d'autres langues par cette interface. Vous pouvez sélectionner les paramètres régionaux à la fois pour l'interface utilisateur et les données affichées.

[Designer : SSL \[page 37\]](#)

Par défaut, les chemins d'accès aux certificats SSL et aux fichiers de clé sont automatiquement configurés pendant l'installation. Il n'est pas nécessaire de les modifier à moins que vous ne souhaitiez utiliser vos propres certificats.

[Données : généralités \[page 38\]](#)

Utilisez les options de la fenêtre Données - Généralités pour définir les options de données par défaut dans le logiciel.

[Job Server : environnement \[page 39\]](#)

Data Services utilise des processus et des threads pour exécuter des jobs qui extraient les données de sources, transforment les données et les chargent dans un entrepôt de données. Le nombre de processus et de threads en cours d'exécution simultanément influence la performance des jobs Data Services.

[Job Server : généralités \[page 39\]](#)

Utilisez cette fenêtre pour modifier les valeurs d'option par défaut pour un Job Server individuel.

[Page](#)

5.7.1 Designer : environnement

Vous pouvez modifier les paramètres par défaut pour le reporting de métadonnées et les job servers, ainsi que les paramètres de port de communication pour le Designer.

Administrateur par défaut pour le reporting des métadonnées

Option	Description
Administrateur	Sélectionnez l'Administrateur utilisé par l'outil de reporting des métadonnées. Un Administrateur est défini par un nom d'hôte et un port.

Job Server par défaut

Option	Description
Actuel	Affiche la valeur actuelle du Job Server par défaut.

Option	Description
<i>Nouveau</i>	Permet d'indiquer une nouvelle valeur pour le Job Server par défaut à partir d'une liste déroulante des Job Servers associés au référentiel. Les modifications sont appliquées immédiatement.

Si un référentiel est associé à plusieurs Job Servers, un seul Job Server doit être défini comme Job Server par défaut à utiliser lors de la connexion.

i Remarque

Les options spécifiques au job et les noms de chemin indiqués dans Designer se rapportent au Job Server par défaut actuel. Si vous modifiez le Job Server par défaut, modifiez ces options et noms de chemin.

Ports de communication Designer

Option	Description
<i>Autoriser Designer à définir le port de communication du job server</i>	Si la case est cochée, Designer définit automatiquement un port disponible pour la réception des messages du Job Server actuel. Elle est cochée par défaut. Décochez-la pour indiquer un port d'écoute ou une plage de ports.
<i>De</i> <i>A</i>	Saisissez les numéros de port dans les zones de texte correspondantes. Pour indiquer un port d'écoute spécifique, saisissez le même numéro de port dans les deux zones de texte de port <i>De</i> et <i>A</i>. Les modifications seront seulement appliquées au redémarrage du progiciel. Uniquement activé lorsque vous désélectionnez le paramètre précédent. Permet d'indiquer une plage de ports dans laquelle Designer peut choisir un port d'écoute. Vous pouvez décider de restreindre le port utilisé pour la communication entre Designer et le Job Server lorsque les deux composants sont séparés par un pare-feu.
<i>Débogueur interactif</i>	Permet de définir un port de communication avec un Job Server pour Designer, à utiliser lors de l'exécution en mode débogage.
<i>Groupe de serveurs pour le référentiel local</i>	Si le référentiel local auquel vous vous êtes connecté à l'ouverture de Designer est associé à un groupe de serveurs, le nom du groupe de serveurs apparaît.

Informations associées

[Modification du port du débogueur interactif \[page 744\]](#)

5.7.2 Designer : généralités

Vous pouvez définir les paramètres par défaut pour les options couramment utilisées dans le Designer.

Option	Description
<i>Afficher la taille de l'échantillonnage des données (lignes)</i>	Contrôle la taille de l'échantillon utilisé pour afficher les données dans les sources et cibles des flux de données en cours dans l'espace de travail. Pour visualiser les données, cliquez sur l'icône de loupe dans les objets sources et cibles.
<i>Nombre de caractères dans le nom de l'icône d'espace de travail</i>	Contrôle la longueur des noms d'objets affichés dans l'espace de travail. Les noms d'objets peuvent dépasser cette valeur mais Designer affiche seulement le nombre saisi ici. La valeur par défaut est de 17 caractères.
<i>Nombre maximal d'éléments de l'arborescence de schéma à développer automatiquement</i>	Nombre d'éléments affichés dans l'arborescence de schéma. Les noms d'éléments ne doivent pas dépasser cette valeur. Saisissez un nombre pour le <i>Schéma d'entrée</i> et le <i>Schéma de sortie</i> . La valeur par défaut est 100.
<i>Paramètres par défaut aux variables portant le même nom</i>	Lorsque vous déclarez une variable au niveau du workflow, le progiciel transmet automatiquement la valeur sous forme de paramètre du même nom à un flux de données appelé par un workflow.
<i>Importer les domaines automatiquement</i>	Cochez cette case pour importer automatiquement les domaines lors de l'importation d'une table faisant référence à un domaine.
<i>Procéder à une validation complète avant l'exécution du job</i>	Cette option permet d'exécuter une validation complète du job préalablement à son exécution. Elle est désactivée par défaut. Si vous conservez ce paramétrage par défaut, vous devez valider la conception manuellement avant l'exécution du job.
<i>Ouvrir le moniteur à l'exécution du job</i>	Détermine le comportement de Designer lors de l'exécution d'un job. Si cette option est activée, Designer bascule de l'espace de travail à la vue du moniteur lors de l'exécution d'un job. Dans le cas contraire, l'espace de travail reste tel quel. L'option est activée par défaut.
<i>Calculer automatiquement les mappages de colonnes</i>	Calcule les informations relatives aux tables et colonnes cible et aux sources utilisées pour les compléter. Le progiciel utilise ces informations pour les rapports de métadonnées (sur l'impact et le lignage des données, par exemple), l'auto-documentation ou les rapports personnalisés. Les informations sur le mappage des colonnes sont sauvegardées dans la table AL_COLMAP (vue ALVW_MAPPING) une fois que vous avez enregistré un flux de données, importé des objets dans un référentiel ou exporté des objets depuis un référentiel. Si l'option est sélectionnée, veillez à valider le job complet avant de l'enregistrer car le calcul du mappage des colonnes est sensible aux erreurs et les flux de données présentant des problèmes de validation seront ignorés.
<i>Afficher la boîte de dialogue une fois le job terminé</i>	Permet d'indiquer si vous voulez obtenir une alerte ou simplement lire les messages de suivi.
<i>Afficher les onglets dans l'espace de travail</i>	Permet d'indiquer si vous voulez naviguer à l'aide des onglets en bas de l'espace de travail.
<i>Fenêtre simple dans l'espace de travail</i>	Permet d'afficher une seule fenêtre dans la zone d'espace de travail (par exemple un job ou une transformation). Si vous ouvrez une autre fenêtre, la fenêtre précédente se ferme et la nouvelle s'ouvre.
<i>Afficher la page d'accueil au démarrage</i>	Permet d'afficher la page d'accueil Designer à l'ouverture de Designer.
<i>Activer la description de l'objet lors de l'instanciation</i>	Active la description d'éléments pour les éléments placés dans votre flux (par exemple un lecteur ou une transformation mis dans un job, un workflow ou un flux de données).

Option	Description
Exclure les éléments non exécutables de l'exportation vers le document XML	Exclut les éléments non traités lors de l'exécution du job des documents XML exportés. Par exemple, les coordonnées d'affichage de l'espace de travail de Designer ne sont pas exportées.

Pour en savoir plus sur l'actualisation de l'onglet des données d'utilisation, voir le *Guide de la Console de gestion*.

Informations associées

[Utilisation de l'affichage des données \[page 727\]](#)

5.7.3 Designer : éléments graphiques

Sélectionnez et prévisualisez les éléments de style afin de personnaliser vos espaces de travail. Ces options permettent d'établir une distinction claire entre l'espace de travail de conception de jobs/workflows et l'espace de travail de conception de flux de données.

Option	Description
Type de flux de l'espace de travail	Passez de l'un à l'autre des types de flux d'espace de travail (Job/Workflow et Flux de données) pour visualiser les paramètres par défaut. Modifiez les paramètres de chaque type à l'aide des options restantes.
Type de ligne	Choisissez un style pour les lignes de connecteurs d'objets.
Épaisseur de ligne	Définit l'épaisseur de ligne du connecteur.
Style d'arrière-plan	Choisissez un motif d'arrière-plan simple ou en mosaïque pour le type de flux sélectionné.
Jeu de couleurs	Définit la couleur d'arrière-plan en bleu, gris ou blanc.
Utiliser le filigrane de navigation	Ajoute un filigrane à l'arrière-plan du type de flux sélectionné. Cette option est uniquement disponible avec un style d'arrière-plan simple.

5.7.4 Designer Fonts options

Use the *Font* options to set font type, style, and size for text in the workspace pane.

For each option, click *Choose Font* to open the *Font* dialog box. Options include *Font*, *Font style*, and *Size*.

Option	Description
<i>Object label in workspace</i>	Specifies the font options for the text that appears above each object in the workspace, such as the object name.
<i>Object description and annotation in workspace</i>	Specifies the font options for the text of the object description and annotations.
<i>Custom controls</i>	Specifies the font options for text other than object labels, descriptions, and annotations.

Informations associées

[Options générales et de l'environnement \[page 31\]](#)

5.7.5 Designer : valeurs d'attributs

Lors d'une réimportation, Data Services conserve ou efface l'ancienne valeur pour certains attributs de table et de colonne. Vous pouvez contrôler les actions effectuées par Data Services lors d'une réimportation pour certains attributs.

Attribut	Type d'objet	Description
<i>Description professionnelle</i>	Colonne et table	Description commerciale d'une table ou d'une colonne.
<i>Nom commercial</i>	Colonne et table	Champ logique. L'attribut définit et exécute les jobs qui extraient, transforment et chargent des données physiques pendant que les données Nom commercial restent intactes.
<i>Type de contenu</i>	Colonne	Définit le type de données dans une colonne.
<i>Description</i>	Colonne et table	Description de la colonne ou table.
<i>Nombre de lignes estimé</i>	Table	Estimation de la taille de la table utilisée lors du calcul de l'ordre de lecture des tables pour l'exécution des opérations de jointure.
<i>Etiquette de l'objet</i>	Colonne et table	Etiquette utilisée pour décrire un objet.
<i>Utilisation de la table</i>	Table	Champ d'étiquette utilisé par exemple pour marquer une table comme un fait ou une dimension.

5.7.6 Designer : connexions au référentiel central

Le référentiel central fournit une bibliothèque d'objets partagée permettant aux développeurs d'enregistrer des objets dans leurs référentiels locaux et de les en extraire. Utilisez ces options pour définir les paramètres de connexion par défaut.

Options	Description
Connexions au référentiel central	Affiche les connexions au référentiel central et le référentiel central actif. Pour activer un référentiel central, cliquez avec le bouton droit sur l'une des connexions de la liste et sélectionnez Activer .
Réactiver automatiquement	Indiquez si vous voulez que le référentiel central actif soit ré-activé chaque fois que vous vous connectez au progiciel à l'aide du référentiel local en cours.

5.7.7 Designer : langue

Ces options fournissent des paramètres régionaux dans une langue autre que l'anglais (paramètre par défaut) pour l'affichage de l'interface utilisateur Data Services et de tout texte généré dans d'autres langues par cette interface. Vous pouvez sélectionner les paramètres régionaux à la fois pour l'interface utilisateur et les données affichées.

Option	Description
Paramètres régionaux du produit	Spécifie la langue de l'interface utilisateur et de tous les messages du produit.
Paramètres régionaux de visualisation préférés	Spécifie les paramètres régionaux à utiliser pour la présentation des données utilisateur. Par exemple, le format de la date doit correspondre aux paramètres régionaux de visualisation préférés.

5.7.8 Designer : SSL

Par défaut, les chemins d'accès aux certificats SSL et aux fichiers de clé sont automatiquement configurés pendant l'installation. Il n'est pas nécessaire de les modifier à moins que vous ne souhaitiez utiliser vos propres certificats.

i Remarque

Si vous modifiez des options SSL autres que [Utiliser le protocole SSL pour le profileur](#), vous devez redémarrer Designer et tous les serveurs Data Services.

Option	Description
<i>Fichier de certificat du serveur</i>	Chemin d'accès au fichier de certificat du serveur. Le certificat du serveur doit être au format PEM.
<i>Fichier de clé privée du serveur</i>	Chemin d'accès au fichier de clé privée du serveur.
<i>Utiliser le fichier de mots de passe de la clé privée du serveur</i>	Sélectionnez cette option et indiquez l'emplacement du fichier de mots de passe si vous souhaitez utiliser un fichier de mots de passe de la clé privée.
<i>Dossier de certificats approuvés</i>	Indiquez l'emplacement du dossier de certificats approuvés. Les extensions valides pour les certificats dans le dossier des certificats approuvés comprennent .pem, .crt et .cer. Quelle que soit l'extension du fichier, tous les contenus du fichier de certificats doivent être au format PEM.
<i>Utiliser le protocole SSL pour le profileur</i>	Sélectionnez cette option pour utiliser le protocole SSL pour les communications entre Designer et le serveur du profileur.

5.7.9 Données : généralités

Utilisez les options de la fenêtre Données - Généralités pour définir les options de données par défaut dans le logiciel.

Les options suivantes sont disponibles :

Option	Description										
<i>Changement de siècle</i>	Indique la façon dont le progiciel interprète le siècle pour les années à deux chiffres. Les années à deux chiffres supérieures ou égales à cette valeur sont interprétées comme 19##. Les années à deux chiffres inférieures à cette valeur sont interprétées comme 20##. La valeur par défaut est 15. Si, par exemple, vous donnez la valeur 15 à l'option <i>Changement de siècle</i> : <table> <tr> <th>Année à deux chiffres</th><th>Interprétation</th></tr> <tr> <td>99</td><td>1999</td></tr> <tr> <td>16</td><td>1916</td></tr> <tr> <td>15</td><td>1915</td></tr> <tr> <td>14</td><td>2014</td></tr> </table>	Année à deux chiffres	Interprétation	99	1999	16	1916	15	1915	14	2014
Année à deux chiffres	Interprétation										
99	1999										
16	1916										
15	1915										
14	2014										
<i>Convertir les valeurs vides en valeurs nulles pour le Bulk Loader Oracle</i>	Convertit les valeurs vides en valeurs nulles lors du chargement de données via l'utilitaire Bulk Loader Oracle : - si la colonne ne fait pas partie de la clé primaire - et si la colonne accepte la valeur Null.										

5.7.10 Job Server : environnement

Data Services utilise des processus et des threads pour exécuter des jobs qui extraient les données de sources, transforment les données et les chargent dans un entrepôt de données. Le nombre de processus et de threads en cours d'exécution simultanément influence la performance des jobs Data Services.

Option	Description
<i>Nombre maximal de traitements de moteur</i>	Définit le nombre limite de traitements parallèles du moteur pour le Job Server.

5.7.11 Job Server : généralités

Utilisez cette fenêtre pour modifier les valeurs d'option par défaut pour un Job Server individuel.

Une fois le Job Server sélectionné, vous pouvez modifier la valeur par défaut pour un certain nombre d'éléments. Par exemple, vous pouvez modifier le nombre de fois qu'un Job Serveur tente d'établir une connexion au FTP s'il a initialement échoué.

Informations associées

[Modifier les valeurs d'option pour un Job Server individuel \[page 381\]](#)

5.7.12



Option	Description

6 Connexion au Designer

Pour vous connecter au progiciel, vous devez avoir accès à un référentiel local. En règle générale, vous créez un référentiel lors de l'installation. Vous pouvez néanmoins créer un référentiel à tout moment à l'aide du Gestionnaire de référentiel et configurer les droits d'accès dans le CMS (Central Management Server).

Par ailleurs, chaque référentiel doit être associé à au moins un Job Server pour que vous puissiez exécuter des jobs liés aux référentiels dans Designer. Vous définissez habituellement un Job Server que vous associez à un référentiel lors de l'installation. En revanche, vous pouvez définir ou modifier des Job Server ou les liens entre les référentiels et les Job Server à tout moment à l'aide du Gestionnaire de serveurs.

Lorsque vous vous connectez à Designer, vous devez vous connecter en tant qu'utilisateur défini dans le CMS (Central Management Server).

1. Saisissez vos références d'utilisateur pour le CMS.

Option	Description
<i>Système</i>	Spécifiez le nom de serveur et, facultativement, le port du CMS.
<i>Nom d'utilisateur</i>	Spécifiez le nom d'utilisateur à utiliser pour la connexion au CMS.
<i>Mot de passe</i>	Spécifiez le mot de passe à utiliser pour la connexion au CMS.
<i>Authentification</i>	Spécifiez le type d'authentification utilisé par le CMS.

2. Cliquez sur [Se connecter](#).
Le logiciel tente de se connecter au CMS à l'aide des informations spécifiées. Lorsque vous êtes connecté, la liste des référentiels locaux disponibles s'affiche.
3. Sélectionnez le référentiel que vous souhaitez utiliser.
4. Pour vous connecter à l'aide du référentiel sélectionné, cliquez sur [OK](#).
Lorsque vous cliquez sur [OK](#), vous devez saisir le mot de passe pour le référentiel Data Services. Ce comportement par défaut peut être modifié en ajoutant les droits nécessaires au référentiel dans le CMC. Pour en savoir plus, voir le *Guide d'administration*.

Pour changer les référentiels dans Designer, faites un clic droit dans la bibliothèque d'objets locale et sélectionnez [Changer de bibliothèque d'objets](#).

6.1 Restrictions de version

Les restrictions de version doivent être prises en compte lors de l'utilisateur de Designer.

La version du référentiel doit être associée à la même **version principale** que Designer et être inférieure ou égale à la version de Designer.

Comme illustré dans l'exemple suivant des versions de Data Services, les chiffres de la troisième position à gauche indiquent la désignation de la version majeure.

Par exemple, dans SAP Data Services version 14.2.10.1748, les numéros ont la signification suivante :

- **14.2** est une version de lancement.
- **10** est la version principale, également connue en tant que service pack.
- **1748** est la version mineure, également connue sous le nom de la version du correctif.

Lors de la connexion, le logiciel vous alerte en cas de non-concordance entre la version de Designer et celle de votre référentiel. Une fois connecté, vous pouvez afficher la version du progiciel et la version du référentiel en sélectionnant [Aide > A propos de Data Services](#).

Certaines fonctions de la version actuelle de Designer peuvent ne pas être prises en charge si vous n'êtes pas connecté à la dernière version du référentiel.

6.2 Réinitialisation des utilisateurs

Si plusieurs personnes essayent de se connecter à un même référentiel, il est possible que vous deviez réinitialiser ou déconnecter un utilisateur pour continuer.

La fenêtre Réinitialiser les utilisateurs affiche les utilisateurs et l'heure à laquelle ils se sont connectés au référentiel.

Cette fenêtre vous offre plusieurs possibilités. Vous pouvez :

- [Réinitialiser les utilisateurs](#) dans le référentiel et définir votre propre utilisateur comme l'utilisateur connecté actuellement.
- [Continuer](#) et vous connecter au système indépendamment du fait que d'autres utilisateurs soient connectés ou non.
- [Quitter](#) pour interrompre la tentative de connexion et fermer la session.

i Remarque

Utilisez [Réinitialiser les utilisateurs](#) ou [Continuer](#) uniquement si vous êtes sûr d'être le seul utilisateur connecté au référentiel. Les modifications a posteriori risquent d'endommager le référentiel.

7 Mots réservés

8 Objets

Toutes les « entités » que vous définissez, modifiez ou utilisez dans SAP Data Services Designer sont appelées objets.

Chaque objet s'inscrit dans l'une de ses deux classes :

- Usage unique
- Réutilisable

La classe d'objet détermine la méthode de création et d'extraction de l'objet.

i Remarque

Pour plus d'informations sur les objets spécifiques à une source, consultez le document en supplément applicable pour cette source. Par exemple, pour obtenir des informations sur les applications SAP utilisées en tant que source, voir le *Supplément pour SAP*.

[Objets réutilisables \[page 44\]](#)

Vous pouvez réutiliser et répliquer la plupart des objets définis dans le progiciel.

[Objets à usage unique \[page 46\]](#)

Les objets à usage unique apparaissent uniquement en tant que composants d'autres objets.

[Save reusable objects \[page 47\]](#)

When you save a reusable object, the software stores the language that describes the object in the repository.

[Métadonnées de l'objet \[page 48\]](#)

Lorsque vous créez un objet, vous associez l'objet à un jeu d'options, de propriétés et d'attributs.

[Descriptions des objets \[page 49\]](#)

Data Services fonctionne à l'aide d'objets, et un grand nombre sont combinés pour générer les processus Data Services.

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

L'éditeur affiche les schémas d'entrée et de sortie de l'objet ainsi que les options pour l'objet.

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

Vous pouvez créer un objet réutilisable dans la bibliothèque d'objets plutôt que de le créer à l'aide de la palette d'outils.

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

Vous pouvez réutiliser et répliquer la plupart des objets définis dans le progiciel à l'aide la palette d'outils.

[Ajout d'un objet existant \[page 61\]](#)

Ajouter un objet existant ou créer un appel pour un objet existant.

[Modification des noms des objets \[page 62\]](#)

Vous pouvez modifier le nom d'un objet à partir de l'espace de travail ou de la bibliothèque d'objets.

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

Vous pouvez ajouter, afficher et dans certains cas, modifier les propriétés d'un objet via sa page de propriétés.

[Annotations \[page 65\]](#)

Use annotations like stick on notes to describe aspects of a flow of data such as in jobs, work flows, or data flows.

[Création de descriptions \[page 66\]](#)

Les descriptions permettent de donner des informations sur les objets.

[Copie d'objets \[page 69\]](#)

Vous pouvez couper ou copier des objets, puis les coller à l'emplacement souhaité dans l'espace de travail.

[Replicating objects \[page 70\]](#)

Replicating an object creates a copy of the object.

[Enregistrement et suppression des objets \[page 71\]](#)

Enregistrer un objet dans le logiciel équivaut à sauvegarder la langue qui décrit cet objet dans le référentiel.

[Recherche d'objets \[page 74\]](#)

Dans la bibliothèque d'objets, vous pouvez rechercher des objets définis dans un référentiel ou des objets disponibles par le biais d'une banque de données.

8.1 Objets réutilisables

Vous pouvez réutiliser et répliquer la plupart des objets définis dans le logiciel.

Après avoir défini et enregistré un objet réutilisable, SAP Data Services stocke la définition dans le référentiel. Vous pouvez ensuite réutiliser la définition aussi souvent que vous le voulez en créant des appels. Les objets réutilisables sont accessibles par le biais de la bibliothèque d'objets.

Un objet réutilisable possède une définition unique ; tous les appels de l'objet font référence à cette définition. Si vous modifiez la définition de l'objet à un endroit, puis enregistrez cet objet, la modification s'applique à tous les autres appels à l'objet.

Un flux de données est, par exemple, un objet réutilisable. Plusieurs jobs, comme un job de chargement hebdomadaire et un job de chargement quotidien, peuvent appeler le même flux de données. Si le flux de données change, les deux jobs appellent la nouvelle version du flux de données.

Utiliser la fonction Glisser-déposer sur un objet de la bibliothèque revient à créer une référence (ou un `<appel>`) pour la définition d'objet existante.

Il est possible de modifier des objets réutilisables à tout moment, quel que soit le projet ouvert en cours. Par exemple, si vous ouvrez un nouveau projet, vous pouvez accéder à la bibliothèque d'objets, ouvrir un flux de données, puis le modifier. L'objet restera inutilisable jusqu'à ce que vous l'enregistriez de manière explicite.

Les fonctions sont des objets réutilisables qui ne sont pas disponibles dans la bibliothèque d'objets. Le logiciel fournit l'accès à ces objets via l'Assistant Fonction, quel que soit leur emplacement d'utilisation.

Certains objets de la bibliothèque d'objets ne sont pas réutilisables dans toutes les instances :

- Les banques de données se trouvent dans la bibliothèque d'objets car elles représentent une méthode de catégorisation et d'accès aux métadonnées externes.
- Les transformations intégrées sont "réutilisables" à chaque suppression de transformation ou à chaque fois qu'une nouvelle instance de la transformation est créée.

L'enregistrement d'un objet réutilisable revient à stocker la langue qui décrit l'objet dans le référentiel. La description d'un objet réutilisable inclut ces composants :

- Propriétés de l'objet
- Options de l'objet
- Appels que cet objet effectue vers d'autres objets
- Définition des objets à usage unique appelés par cet objet

Si un objet contient un appel à un autre objet réutilisable, seul l'appel au second objet est enregistré ; les modifications apportées à ce dernier ne le sont pas.

La description est stockée même si la validation de l'objet n'aboutit pas.

Les objets sont enregistrés sans invite :

- Lorsque vous importez un objet dans le référentiel.
- Lorsque vous terminez la modification des éléments suivants :
 - Banques de données
 - Formats de fichiers plats
 - Schémas imbriqués, tels que le format DTD, le format JSON ou le schéma XML

Vous pouvez, de manière explicite, enregistrer l'objet réutilisable actuellement ouvert dans l'espace de travail en sélectionnant *Enregistrer* dans le menu *Projet*. Si un objet à usage unique est ouvert dans l'espace de travail, la commande *Enregistrer* n'est pas disponible.

Pour enregistrer tous les objets modifiés dans le référentiel, sélectionnez *Tout enregistrer* dans le menu *Projet*.

Vous êtes invité à enregistrer tous les objets possédant des modifications lorsque vous exécutez un job, puis quittez l'application Designer.

Rubrique parent : [Objets \[page 43\]](#)

Informations associées

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.2 Objets à usage unique

Les objets à usage unique apparaissent uniquement en tant que composants d'autres objets.

Ils fonctionnent uniquement dans le contexte dans lequel ils ont été créés.

L'enregistrement d'un objet à usage unique revient à stocker la langue qui décrit l'objet dans le référentiel. La description d'un objet à usage unique peut uniquement être enregistrée comme élément de l'objet réutilisable qui appelle l'objet à usage unique.

La description est stockée même si l'objet n'est pas validé.

Rubrique parent : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.3 Save reusable objects

When you save a reusable object, the software stores the language that describes the object in the repository.

The description of a reusable object includes these components:

- Properties of the object
- Options for the object
- Calls this object makes to other objects
- Definition of single-use objects called by this object

The transform stores the object description even if the software cannot successfully validate the object.

The software saves some reusable objects without prompting you to save them:

- When you import an object into the repository, the software saves the object.
- When you finish editing the following objects and close the editor, the software saves the object:
 - Datastores
 - Flat file formats
 - Nested schemas, such as DTD format, JSON format, or XML Schema

Save other reusable objects by using the [Save](#) or [Save All](#) options:

- Save the reusable object currently open in the workspace by choosing [Save](#) from the [Project](#) menu.

Remarque

If a single-use object is open in the workspace, the [Save](#) command is not available.

- Save all objects in the repository that have changes by choosing [Save All](#) from the [Project](#) menu.

Additionally, the software prompts you to save all objects that have changes when you execute a job or when you exit Designer.

Parent topic: [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.4 Métadonnées de l'objet

Lorsque vous créez un objet, vous associez l'objet à un jeu d'options, de propriétés et d'attributs.

Après la création et l'enregistrement d'un objet, le logiciel enregistre les options avec des métadonnées qui identifient l'objet.

Métadonnées de l'objet

Informations d'objets	Description
Options	Définit les opérations de l'objet telles que les informations de connexion à la base de données, les actions, les mots de passe et les noms d'utilisateur. Par exemple, dans une banque de données, une option représente le nom de la base de données à laquelle la banque de données se connecte. Affichez et modifiez les options de l'objet en ouvrant l'éditeur d'objets.
Propriétés	Documente les informations de l'objet telles que le nom et la description de l'objet. Les propriétés décrivent un objet ; elles n'ont aucune incidence sur le fonctionnement de l'objet. Affichez les propriétés d'un objet en cliquant avec le bouton droit sur le nom de l'objet et en sélectionnant <i>Propriétés</i>.
Attributs	Fournit des informations supplémentaires sur un objet comme la date de création, la méthode d'enregistrement, et si l'objet est activé pour le service Web. Les valeurs d'attributs peuvent avoir une incidence sur le comportement de l'objet. Affichez les attributs de l'objet en ouvrant les propriétés de l'objet et en sélectionnant l'onglet <i>Attributs</i>.

Rubrique parent : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.5 Descriptions des objets

Data Services fonctionne à l'aide d'objets, et un grand nombre sont combinés pour générer les processus Data Services.

Le tableau suivant répertorie tous les objets présents dans Data Services, avec leur classe et leur description. Pour en savoir plus sur chaque objet, voir les rubriques propres à chaque objet.

Descriptions des objets

Objet	Classe d'objet	Description
Annotation	Usage unique	Note qui contient des informations sur les aspects d'un flux, une partie d'un flux ou un diagramme dans l'espace de travail. Vous créez l'annotation et la joignez à l'élément correspondant dans l'espace de travail.
Job batch	Réutilisable	Définit les activités que le logiciel exécute à un moment donné, y compris les messages de suivi, de surveillance et d'erreur. Ajoutez des jobs uniquement à des projets. Une fois un job batch enregistré, le logiciel le place dans la bibliothèque d'objets. Le job batch dans la bibliothèque d'objets est une référence directe à l'objet du job batch. Les projets peuvent contenir des références à plusieurs objets de job batch, mais seulement une référence par job batch.
Catch	Usage unique	Spécifie les étapes à exécuter si une erreur survient dans un groupe d'exceptions donné lors de l'exécution d'un job.
Format de fichier copy-book COBOL	Réutilisable	Définit le format d'une source de fichier de copybook COBOL.

Objet	Classe d'objet	Description
Structure conditionnelle	Usage unique	Spécifie les étapes à exécuter selon les résultats d'une condition.
Flux de données	Réutilisable	Spécifie les exigences en termes d'extraction, de transformation et de chargement des données des sources vers les cibles. Un flux de données peut être une partie d'un job d'arrière-plan ou un job en temps réel.
Banque de données	Réutilisable	Spécifie les informations de connexion pour accéder à une base de données ou à une autre source de données. Vous ne pouvez pas ajouter l'objet de banque de données lui-même dans l'espace de travail.
Document	Réutilisable	Décrit la structure de données des schémas imbriqués compliqués. Disponible uniquement pour certains objets de banque de données d'adaptateur.
DTD	Réutilisable	Décrit le format qu'un fichier ou message XML lit ou écrit. DTD est l'acronyme de « Document Type Definition », c'est-à-dire définition du type de document.
Format de classeur Excel	Réutilisable	Définit le format d'une source de classeur Excel.
Format de fichier	Réutilisable	Décrit l'emplacement et le nom d'un fichier plat ainsi que l'organisation des données du fichier plat dans un fichier source ou cible.
Emplacement du fichier	Réutilisable	Définit le protocole de transfert de fichier à utiliser pour transférer les fichiers de données. Inclut les informations concernant les serveurs distants et locaux. Associez un objet d'emplacement de fichier à un format de fichier afin de l'utiliser en tant que source ou cible dans le flux de données. • En tant que source, le logiciel utilise les informations de l'objet d'emplacement de fichier afin de transférer les données source du serveur distant au serveur local. • En tant que cible, le logiciel utilise les informations de l'objet d'emplacement de fichier afin de transférer le fichier de sortie du serveur local au serveur distant.
Fonction	Réutilisable	Renvoie une valeur.
Format de fichier HDFS	Réutilisable	Décrit la structure d'un système de fichiers distribué Hadoop.
Fichier JSON	Usage unique	Cible ou source batch ou en temps réel. • En tant que source, un fichier JSON convertit les données JSON entrantes en données que le logiciel peut traiter. • En tant que cible, un fichier JSON convertit les données produites par un flux de données, y compris les données imbriquées en fichier au format JSON.

Objet	Classe d'objet	Description
Message JSON	Usage unique	Source ou cible en temps réel. - En tant que source, les messages JSON convertissent les requêtes JSON entrantes en données que le job en temps réel pourra traiter. - En tant que cible, les messages JSON convertissent les résultats du job en temps réel (y compris les données hiérarchiques) en réponse JSON, puis transmettent les messages au serveur d'accès.
Journal	Usage unique	Enregistre les informations relatives à l'exécution particulière d'un job donné.
Fonction de message	Réutilisable	S'applique aux messages XML lorsque la configuration est appropriée. Disponible uniquement pour certaines banques de données d'adaptateur.
Modèle de schémas imbriqués	Usage unique	Cible qui crée un fichier JSON qui correspond à un schéma d'entrée donné. Ne requiert pas de DTD, schéma JSON ni schéma XML.
Message sortant	Réutilisable	Communications hiérarchiques XML que les jobs en temps réel peuvent publier vers les adaptateurs. Disponible uniquement pour certaines banques de données d'adaptateur.
Projet	Usage unique	Contient des groupes d'un ou plusieurs jobs pour faciliter l'accès.
Transformation Query	Usage unique	Définit les conditions pour extraire un jeu de données spécifié.
Job en temps réel	Réutilisable	Définit les activités que le logiciel exécute à la demande en temps réel. Créez des jobs en temps réel dans Designer. Puis configurez-les et exécutez-les en tant que services associés à un serveur d'accès dans l'Administrateur. - Concevez les jobs en temps réel conformément aux règles de modèles de flux de données. - Exécutez les jobs en temps réel selon un système "demande-réponse".
Script	Usage unique	Évalue les expressions, les fonctions d'appels et affecte des valeurs aux variables.
Source	Usage unique	Contient les données source que le logiciel lit et traite dans un flux de données.

Objet	Classe d'objet	Description
Table	Réutilisable	Définit une table SGBD externe pour laquelle les métadonnées ont été importées, ou une table cible dans laquelle les données sont ou ont été placées. Une table est associée à une banque de données spécifique. Une table n'existe pas indépendamment d'une connexion à la banque de données. Une table extrait ou stocke les données basées sur le schéma de la définition de table à partir duquel elle a été créée.
Cible	Usage unique	Accepte les données extraites et transformées dans un flux de données.
Modèle de table	Réutilisable	Représente une nouvelle table que vous ajoutez à une base de données. À l'exception des banques de données SAP, toutes les banques de données possèdent un modèle par défaut qui vous permet de créer des tables dans la banque de données. Le logiciel crée le schéma pour chaque instance d'un modèle de table lors de l'exécution. Le schéma créé est basé sur les données chargées dans le modèle de table.
Transformation	Réutilisable	Contient les paramètres qui effectuent des opérations spécifiques sur les jeux de données. Requiert zéro ou plusieurs jeux de données ; produit zéro ou un jeu de données, qui peut être divisé en plusieurs jeux de données.
Try	Usage unique	Introduit un bloc Try/Catch dans un flux de données. Vous pouvez avoir plusieurs instructions Catch pour une seule instruction Try.
Boucle While	Usage unique	Entraîne la répétition en boucle d'une séquence d'étapes tant qu'une condition définie renvoie la valeur Vrai.
Workflow	Réutilisable	Contient un ordre spécifique de flux de données et d'opérations qui prennent en charge un flux de données.
Fichier XML	Usage unique	Convertit les données XML vers les formats applicables selon que le fichier soit une source ou une cible pour les jobs batch et en temps réel. • En tant que source, convertit les données XML entrantes en données que le logiciel peut traiter. • En tant que cible, convertit les données produites par un flux de données, y compris les données imbriquées en fichier au format XML.

Objet	Classe d'objet	Description
Message XML	Usage unique	Convertit les messages XML selon qu'il soit une source ou une cible dans les jobs en temps réel. • En tant que source, convertit les requêtes XML entrantes en données que le job en temps réel pourra traiter. • En tant que cible, convertit les résultats du job en temps réel (y compris les données hiérarchiques) en réponse XML, puis transmettent les messages au serveur d'accès.
Schéma XML	Réutilisable	Décrit la structure des données d'un fichier ou message XML, pour qu'un document XML puisse lire ou écrire les données.

Rubrique parent : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

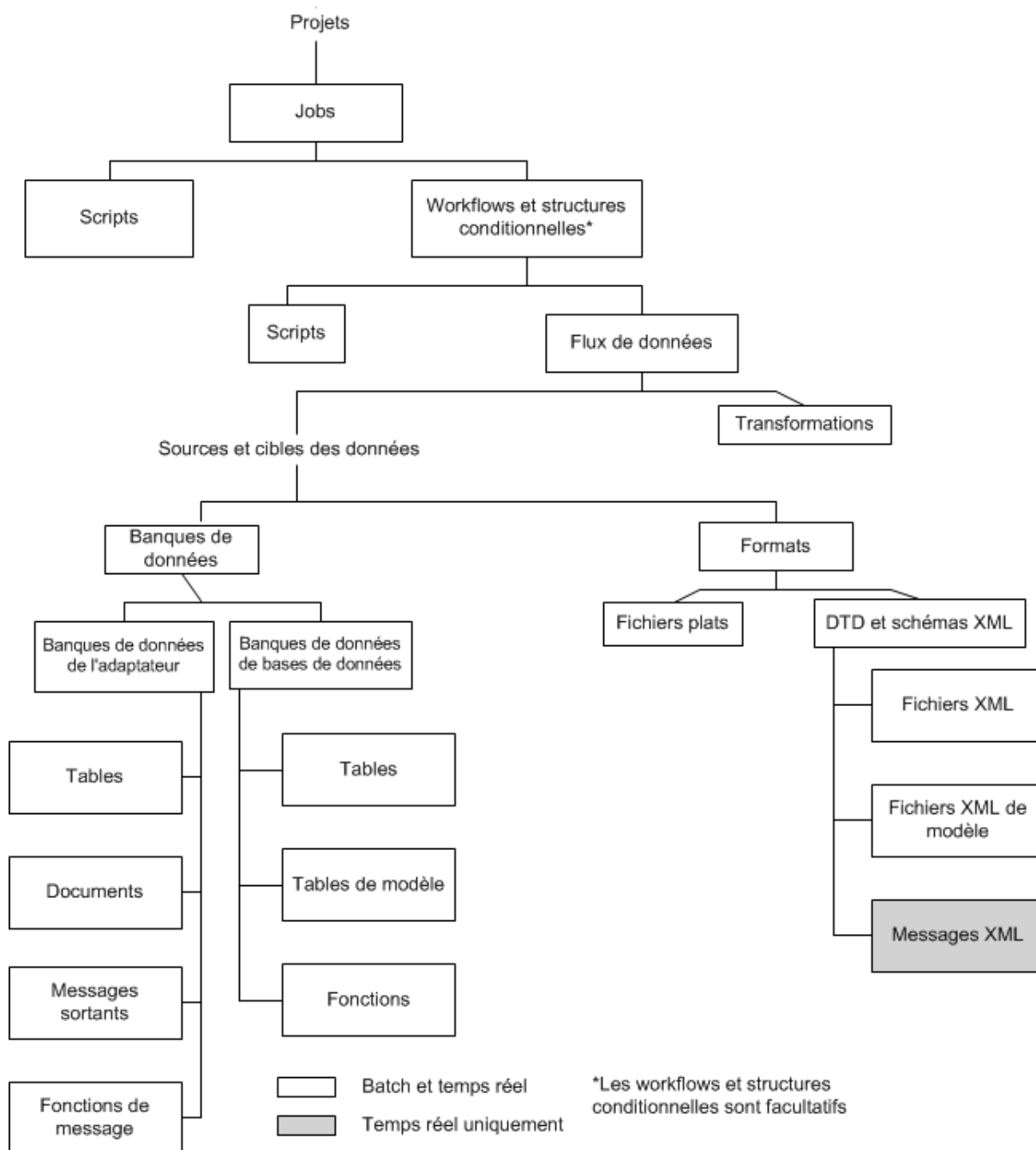
[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.6 Hiérarchie d'objets



-
-

[Page](#)

[Page](#)

[Page](#)

Rubrique parent : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.6.1

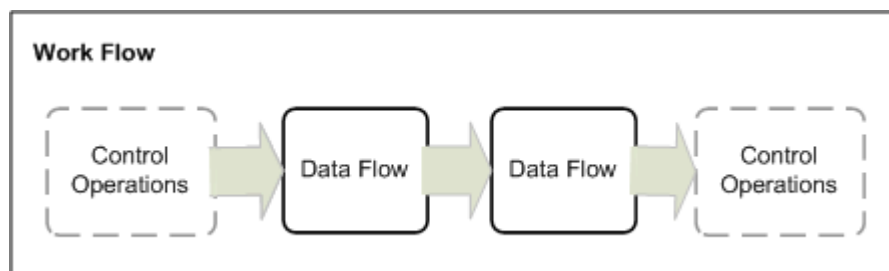
8.6.2

-
-
-
- Transmettre les paramètres vers et depuis les flux de données
-
-

Bouton	Composant
	Workflow Procédure
	Flux de données
	Script
	Structure conditionnelle
	
	Instruction Try
	Catch
	Annotation

8.6.3

-
-



Modèles

Plutôt que de créer votre propre job à partir de zéro, servez-vous des modèles. Si l'un d'eux est lié à votre problème métier, utilisez-le et modifiez les options dans les transformations en fonction de vos besoins spécifiques.

Pour chaque scénario, un modèle déjà configuré peut être utilisé pour résoudre le problème métier posé.

Vous pouvez le consulter régulièrement pour accéder à des informations actualisées et utiliser les forums pour poser des questions ou formuler une demande.

<http://scn.sap.com/docs/DOC-8820> 

8.7

Voici quelques exemples :

Préfixe	Suffixe	Objet	Exemple
JOB		Job	JOB_SalesOrg
WF		Workflow	WF_SalesOrg
		Flux de données	DF_Currency
		Banque de données	

Rubrique parent : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.8 Editeurs d'objets

L'éditeur affiche les schémas d'entrée et de sortie de l'objet ainsi que les options pour l'objet.

Cliquez sur le nom de l'objet dans l'espace de travail pour ouvrir son éditeur. S'il existe plusieurs options, elles sont regroupées dans des onglets au niveau de l'éditeur.

Un schéma est une structure de données qui peut contenir des colonnes, d'autres schémas imbriqués et des fonctions (le contenu correspondant aux éléments du schéma). Une table est un schéma qui contient uniquement des colonnes.

Un éditeur permet d'effectuer les opérations suivantes :

- Annuler ou rétablir les actions effectuées dans la fenêtre (clic droit puis *Annuler* ou *Rétablir*).
- Rechercher une chaîne dans l'éditeur (clic droit puis *Rechercher*).
- Glisser-déposer des noms de colonnes dans les cases d'option pertinentes à partir du schéma d'entrée.
- Utiliser des couleurs pour identifier les chaînes et les commentaires dans les zones de texte autorisant la modification des expressions (mots-clés affichés en bleu ; chaînes affichées en rose entre guillemets ; commentaires débutant par le signe livre et affichés en vert).

i Remarque

Il est impossible d'ajouter des commentaires à une clause de mappage dans une transformation Query. Si vous ajoutez des commentaires, le job ne s'exécute pas et il est impossible de l'exporter. Utilisez plutôt la fonction de description d'objet ou d'annotation de l'espace de travail. Par exemple, la syntaxe suivante n'est **pas** prise en charge dans l'onglet Mappage :

```
table.column # comment
```

Rubrique parent : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.9 Création d'un objet réutilisable dans la bibliothèque d'objets

Vous pouvez créer un objet réutilisable dans la bibliothèque d'objets plutôt que de le créer à l'aide de la palette d'outils.

1. Pour ouvrir la bibliothèque d'objets, sélectionnez ► [Outils](#) ► [Bibliothèque d'objets](#) ►.
2. Cliquez sur l'onglet correspondant au type d'objet.
3. Cliquez avec le bouton droit n'importe où, hormis sur les objets existants, et sélectionnez [Nouveau](#).
4. Cliquez avec le bouton droit sur le nouvel objet et sélectionnez [Propriétés](#). Saisissez des options comme le nom et la description pour définir l'objet.

Présentation des tâches : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.10 Création d'un objet réutilisable à l'aide de la palette d'outils

Vous pouvez réutiliser et répliquer la plupart des objets définis dans le progiciel à l'aide la palette d'outils.

1. Dans la palette d'outils, cliquez sur l'icône de l'objet que vous voulez créer.
2. Déplacez le curseur vers l'espace de travail et cliquez à nouveau.

L'icône de l'objet apparaît à l'endroit où vous avez cliqué dans l'espace de travail.

Présentation des tâches : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.11 Ajout d'un objet existant

Ajouter un objet existant ou créer un appel pour un objet existant.

1. Pour ouvrir la bibliothèque d'objets, sélectionnez ► [Outils](#) ► [Bibliothèque d'objets](#) ►.
2. Cliquez sur l'onglet correspondant à un type d'objet.
3. Sélectionnez un objet.
4. Faites glisser l'objet vers l'espace de travail.

i Remarque

Les objets déposés dans l'espace de travail doivent obéir à la logique de la hiérarchie. Vous pouvez, par exemple, déposer un flux de données dans un job, mais vous ne pouvez pas déposer un workflow dans un flux de données.

Présentation des tâches : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.12 Modification des noms des objets

Vous pouvez modifier le nom d'un objet à partir de l'espace de travail ou de la bibliothèque d'objets.

Vous pouvez également créer une copie d'un objet existant.

i Remarque

Il n'est pas possible de modifier le nom des objets intégrés.

1. Pour modifier le nom d'un objet dans l'espace de travail, procédez comme suit :
 - a. Cliquez pour sélectionner l'objet dans l'espace de travail.
 - b. Cliquez avec le bouton droit et sélectionnez *Renommer*.
 - c. Modifiez le texte affiché dans la zone de texte de nom.
 - d. Pour enregistrer le nouveau nom, cliquez en dehors de la zone de texte ou appuyez sur Entrée.
2. Pour modifier le nom d'un objet dans la bibliothèque d'objets, procédez comme suit :
 - a. Sélectionnez l'objet dans la bibliothèque d'objets.
 - b. Cliquez avec le bouton droit et sélectionnez *Propriétés*.
 - c. Modifiez le texte dans la première zone de texte.
 - d. Cliquez sur *OK*.
3. Pour copier un objet, procédez comme suit :
 - a. Sélectionnez l'objet dans la bibliothèque d'objets.
 - b. Cliquez avec le bouton droit et sélectionnez *Répliquer*.
 - c. Le progiciel effectue une copie de l'objet supérieur (mais pas des objets qu'il appelle) et lui donne un nouveau nom que vous pouvez modifier.

Présentation des tâches : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)
[Copie d'objets \[page 69\]](#)
[Replicating objects \[page 70\]](#)
[Enregistrement et suppression des objets \[page 71\]](#)
[Recherche d'objets \[page 74\]](#)

8.13 Ajout, modification et affichage de propriétés d'objet

Vous pouvez ajouter, afficher et dans certains cas, modifier les propriétés d'un objet via sa page de propriétés.

1. Sélectionnez l'objet dans la bibliothèque d'objets.
2. Cliquez avec le bouton droit et sélectionnez *Propriétés*. L'onglet *Général* de la fenêtre *Propriétés* s'ouvre.
3. Renseignez les feuilles de propriétés. Les feuilles de propriétés diffèrent selon le type d'objet mais les onglets Général, Attributs et Attributs de classe, décrits dans les sections qui suivent, sont les plus communs.
4. Lorsque vous avez terminé, cliquez sur *OK* pour enregistrer les modifications apportées aux propriétés de l'objet et fermer la fenêtre.

Vous pouvez également cliquer sur *Appliquer* pour enregistrer les modifications sans fermer la fenêtre.

[Onglet Général \[page 64\]](#)

Permet de modifier le nom de l'objet ainsi que de saisir ou modifier la description correspondante.

[Onglet Attributs \[page 64\]](#)

Permet d'affecter des valeurs aux attributs de l'objet en cours.

[Onglet Attributs de classe \[page 65\]](#)

Affiche les attributs disponibles pour le type d'objet sélectionné.

Présentation des tâches : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)
[Objets à usage unique \[page 46\]](#)
[Save reusable objects \[page 47\]](#)
[Métadonnées de l'objet \[page 48\]](#)
[Descriptions des objets \[page 49\]](#)
[Hiérarchie d'objets \[page 53\]](#)
[Page](#)
[Editeurs d'objets \[page 58\]](#)
[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)
[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)
[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)
[Annotations \[page 65\]](#)
[Création de descriptions \[page 66\]](#)
[Copie d'objets \[page 69\]](#)
[Replicating objects \[page 70\]](#)
[Enregistrement et suppression des objets \[page 71\]](#)
[Recherche d'objets \[page 74\]](#)

8.13.1 Onglet Général

Permet de modifier le nom de l'objet ainsi que de saisir ou modifier la description correspondante.

L'onglet *Général* comporte deux propriétés d'objet principales : le nom et la description.

Vous pouvez ajouter des descriptions d'objets à des objets à usage unique ou des objets réutilisables. Vous pouvez activer et désactiver la description d'un objet ; pour cela, cliquez dessus avec le bouton droit dans l'espace de travail et sélectionnez ou désélectionnez *Afficher les descriptions activées*.

En fonction de l'objet, d'autres propriétés peuvent être affichées dans l'onglet *Général*, par exemple :

- *Exécuter une seule fois*
- *Restaurer en tant qu'unité*
- *Degré de parallélisme*
- *Utiliser les liens de base de données*
- *Type de cache*
- *Contournement*

Informations associées

[Banques de données liées \[page 117\]](#)
[Mécanismes de récupération \[page 765\]](#)
[Création et définition des flux de données \[page 204\]](#)

8.13.2 Onglet Attributs

Permet d'affecter des valeurs aux attributs de l'objet en cours.

Pour affecter une valeur à un attribut, sélectionnez l'attribut et saisissez la valeur dans la zone *Valeur* en bas de la fenêtre.

Certaines valeurs d'attribut sont définies par le progiciel et ne peuvent pas être modifiées. Lorsque vous sélectionnez un attribut qui possède une valeur système, le champ *Valeur* n'est pas accessible.

8.13.3 Onglet Attributs de classe

Affiche les attributs disponibles pour le type d'objet sélectionné.

Tous les objets de flux de données possèdent, par exemple, les mêmes attributs de classe.

Pour créer un attribut pour une classe d'objets, cliquez avec le bouton droit dans la liste d'attributs et sélectionnez [Ajouter](#). Le nouvel attribut est maintenant disponible pour l'ensemble des objets de la classe.

Pour supprimer un attribut, sélectionnez-le, cliquez avec le bouton droit et sélectionnez [Supprimer](#). Vous ne pouvez pas supprimer les attributs de classe prédéfinis par Data Services.

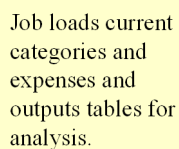
8.14 Annotations

Use annotations like stick on notes to describe aspects of a flow of data such as in jobs, work flows, or data flows.

Together, object descriptions and annotations allow you to document an SAP Data Services application. For example, you can describe the incremental behavior of individual jobs with annotations and include a description for each object.

❖ Exemple

For example, the following annotation describes the purpose of a job:



Job loads current categories and expenses and outputs tables for analysis.

Annotations are different than object descriptions because annotations are related to a process or application and descriptions are related to individual objects.

When you import or export a job, work flow, or data flow, you import or export associated annotations.

Parent topic: [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Création de descriptions \[page 66\]](#)

[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.14.1 Description des objets

Pour ajouter des commentaires sur les objets, vous pouvez utiliser les descriptions.

Les annotations permettent d'expliquer un job, un workflow ou un flux de données. Les descriptions et les annotations des objets sont visibles dans l'espace de travail. Utilisées conjointement, les descriptions et les annotations permettent de documenter une application SAP Data Services . Vous pouvez décrire le comportement incrémental de différents jobs à l'aide de plusieurs annotations et affecter à chaque objet une description de base, par exemple.

❖ Exemple

Ce job charge les catégories et dépenses actuelles puis génère des tables à des fins d'analyse.

Informations associées

[Création de descriptions \[page 66\]](#)

8.15 Création de descriptions

Les descriptions permettent de donner des informations sur les objets.

Elles sont visibles dans les diagrammes de l'espace de travail. De ce fait, les descriptions sont un moyen pratique d'ajouter des commentaires aux objets dans l'espace de travail.

Une description est associée à un objet spécifique. Lorsque vous importez ou exportez cet objet du référentiel (lors de la migration entre environnements de développement, de test et de production, par exemple), vous importez ou exportez également sa description.

Designer détermine quand afficher les descriptions des objets en fonction d'un paramètre au niveau système et d'un paramètre au niveau objet. Les deux paramètres doivent être activés pour que la description d'un objet particulier s'affiche.

Le paramètre au niveau système est spécifique à votre configuration. Il est désactivé par défaut. Pour l'activer, sélectionnez *AffichageDescriptions activées* ou cliquez sur le bouton *Afficher les descriptions activées* dans la barre d'outils.

Le paramètre au niveau objet est enregistré avec l'objet dans le référentiel. Il est également désactivé par défaut, sauf si vous ajoutez ou modifiez une description à partir de l'espace de travail. Pour activer le paramètre au niveau objet, cliquez avec le bouton droit sur l'objet et sélectionnez *Activer la description de l'objet*.

Une ellipse après le texte d'une description indique la présence de texte supplémentaire. Pour voir tout le texte, redimensionnez la description par cliquer-déplacer. Lorsque vous déplacez un objet, sa description est elle aussi déplacée. Pour connaître l'objet associé à une description donnée, affichez le nom de l'objet dans la barre d'état.

[Ajout d'une description à un objet \[page 68\]](#)

[Afficher une description dans l'espace de travail \[page 68\]](#)

[Masquer la description d'un objet \[page 68\]](#)

[Modifier les descriptions d'objets \[page 68\]](#)

Rubrique parent : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

8.15.1 Ajout d'une description à un objet

1. Dans la zone de projet ou la bibliothèque d'objets, cliquez avec le bouton droit sur un objet et sélectionnez [Propriétés](#).
2. Saisissez vos commentaires dans la zone de texte [Description](#).
3. Cliquez sur [OK](#).

La description de l'objet s'affiche dans la bibliothèque d'objets.

8.15.2 Afficher une description dans l'espace de travail

1. Dans la zone de projet, sélectionnez un objet existant (tel qu'un job) contenant un objet auquel vous avez ajouté une description (tel qu'un workflow).
2. Dans le menu [Affichage](#), sélectionnez [Descriptions activées](#).

Vous pouvez également cliquer sur le bouton Afficher les descriptions activées dans la barre d'outils.

3. Cliquez avec le bouton droit sur le workflow et sélectionnez [Activer la description d'objets](#).

La description s'affiche dans l'espace de travail au-dessous de l'objet.

8.15.3 Masquer la description d'un objet

1. Dans le diagramme de l'espace de travail, cliquez avec le bouton droit sur un objet.

Vous pouvez également sélectionner plusieurs objets en procédant comme suit :

- Appuyez sur la touche Ctrl et maintenez-la enfoncée tout en sélectionnant des objets dans le diagramme de l'espace de travail, puis cliquez avec le bouton droit sur l'un des objets sélectionnés.
- Faites glisser une zone de sélection autour de tous les objets à sélectionner, puis cliquez avec le bouton droit sur l'un des objets sélectionnés.

2. Dans le menu contextuel, désélectionnez [Activer la description d'objets](#).

La description de l'objet sélectionné est masquée même si l'option [Afficher les descriptions activées](#) est sélectionnée, car le basculement au niveau objet est prioritaire sur le basculement au niveau système.

8.15.4 Modifier les descriptions d'objets

1. Dans l'espace de travail, double-cliquez sur la description d'un objet.

2. Saisissez, coupez, copiez ou collez du texte dans la description.
3. Dans le menu *Projet*, sélectionnez *Enregistrer*.

Vous pouvez également cliquer avec le bouton droit sur un objet et sélectionner Propriétés pour ouvrir la fenêtre *Propriétés* de l'objet et ajouter ou modifier sa description.

i Remarque

Si vous tentez de modifier la description d'un objet réutilisable, le progiciel vous informe que la description sera mise à jour dans chaque occurrence de l'objet au niveau de tous les jobs. Vous pouvez cocher la case *Ne plus afficher ce message* pour éviter cette alerte. Néanmoins, une fois désactivée, l'alerte ne peut être réactivée qu'avec l'intervention du support technique.

8.16 Copie d'objets

Vous pouvez couper ou copier des objets, puis les coller à l'emplacement souhaité dans l'espace de travail.

Il est possible de copier et coller plusieurs objets dans un même ou dans d'autres flux de données, workflows ou jobs. Par ailleurs, vous pouvez couper ou copier des appels de flux de données et de workflows, puis les coller dans des objets valides au niveau de l'espace de travail.

Les références à des variables globales, à des variables locales, à des paramètres et à des paramètres de substitution sont copiées. Néanmoins, vous devez définir chacun de ces éléments dans son nouveau contexte.

i Remarque

L'opération de collage duplique les objets sélectionnés dans un flux mais continue à appeler les objets d'origine. Autrement dit, l'opération de collage utilise l'objet d'origine dans un autre emplacement. L'opération de réplication crée un nouvel objet dans la bibliothèque d'objets.

Pour couper ou copier des objets puis les coller, procédez comme suit :

1. Dans l'espace de travail, sélectionnez les objets à couper ou à copier.
Pour sélectionner plusieurs objets, utilisez les raccourcis clavier Ctrl-clic, Maj-clic ou Ctrl+A
2. Cliquez avec le bouton droit et sélectionnez *Couper* ou *Copier*.
3. Cliquez dans le même flux ou sélectionnez-en un autre. Cliquez avec le bouton droit et sélectionnez *Coller*.
Si cela s'avère nécessaire pour éviter un conflit d'appellation, un nouveau nom est généré automatiquement.

i Remarque

Si vous cliquez avec le bouton droit et sélectionnez *Coller*, les objets sont collés à l'emplacement sélectionné.

Les objets sont collés dans l'angle supérieur gauche de l'espace de travail si vous utilisez l'une des méthodes suivantes :

- Clic sur l'icône *Coller*.
- Clic sur  *Édition*  *Coller* .
- Utilisation du raccourci clavier Ctrl+V.

Si vous utilisez une méthode qui colle les objets dans l'angle supérieur gauche, les objets collés par la suite sont placés les uns au-dessus des autres.

Présentation des tâches : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

[Replicating objects \[page 70\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.17 Replicating objects

Replicating an object creates a copy of the object.

Replicate objects from the object library. The software generates a different name for the copy. The copy is separate object from the original. You can edit the copy without affecting the original object or any occurrences of the original object.

1. Select the applicable tab in the object library.
2. Right-click the object to be replicated.
3. Select *Replicate* from the dropdown menu.

The software copies the object and adds it to the list of objects. The software prefixes the original object name with `copy_x_transform_name`. For example, the first replicated copy is `copy_1_transform_name`, the second is `copy_2_transform_name`, and so on.

Task overview: [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

[Copie d'objets \[page 69\]](#)

[Enregistrement et suppression des objets \[page 71\]](#)

[Recherche d'objets \[page 74\]](#)

8.18 Enregistrement et suppression des objets

Enregistrer un objet dans le logiciel équivaut à sauvegarder la langue qui décrit cet objet dans le référentiel.

Vous pouvez enregistrer les objets réutilisables ; les objets à usage unique sont seulement enregistrés dans le cadre de la définition d'objet réutilisable qui les appelle.

Il est possible d'enregistrer les modifications de l'objet réutilisable actif dans l'espace de travail. Lorsque vous enregistrez l'objet, ses propriétés, les définitions des objets à usage unique qu'il appelle et les appels éventuels d'autres objets réutilisables sont sauvegardés dans le référentiel. Seul l'appel est enregistré, et non le contenu des objets réutilisables pris en compte.

Le logiciel sauvegarde la description même si l'objet n'est pas complet ou comporte une erreur (validation non effectuée).

Présentation des tâches : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)

[Hiérarchie d'objets \[page 53\]](#)

[Page](#)

[Editeurs d'objets \[page 58\]](#)

[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)

[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)

[Ajout d'un objet existant \[page 61\]](#)

[Modification des noms des objets \[page 62\]](#)

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Annotations \[page 65\]](#)

[Création de descriptions \[page 66\]](#)

[Copie d'objets \[page 69\]](#)

[Replicating objects \[page 70\]](#)

[Recherche d'objets \[page 74\]](#)

8.18.1 Enregistrer les modifications apportées à un seul réutilisable

1. Ouvrez le projet dans lequel figure votre objet.
2. Sélectionnez **Projet > Enregistrer**.

Cette commande enregistre tous les objets ouverts dans l'espace de travail.

Répétez ces étapes pour les autres objets à enregistrer.

8.18.2 Enregistrer tous les objets modifiés dans le référentiel

1. Sélectionnez **Projet > Tout enregistrer**.

Le logiciel liste les objets réutilisables qui ont été modifiés depuis la dernière opération d'enregistrement.

2. (facultatif) Désélectionnez tout objet listé pour éviter de l'enregistrer.
3. Cliquez sur **OK**.

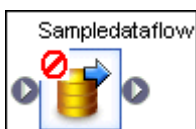
i Remarque

Le logiciel vous invite à enregistrer tous les objets modifiés lorsque vous exécutez un job et que vous quittez Designer. L'enregistrement d'un objet réutilisable entraîne l'enregistrement des objets à usage unique qu'il contient le cas échéant.

8.18.3

8.18.4 Supprimer une définition d'objet du référentiel

1. Dans la bibliothèque d'objets, sélectionnez l'objet.
2. Cliquez avec le bouton droit et sélectionnez *Supprimer*.
 - Si vous tentez de supprimer un objet en cours d'utilisation, le progiciel émet un message d'avertissement et vous propose d'utiliser la fonction *Afficher les cas d'emploi*.
 - Si vous cliquez sur *Oui*, tous les appels de l'objet sont marqués par une icône rouge « supprimé » pour indiquer que les appels sont non valides. Pour produire un job exécutable, vous devez supprimer ou remplacer ces appels.



i Remarque

Les objets intégrés comme les transformations ne peuvent pas être supprimés de la bibliothèque d'objets.

Informations associées

[Utilisation de Afficher les cas d'emploi \[page 722\]](#)

8.18.5 Supprimer un appel d'objet

1. Ouvrez l'objet qui contient l'appel à supprimer.
2. Cliquez avec le bouton droit sur l'appel d'objet et sélectionnez *Supprimer*.

Si vous supprimez un objet réutilisable de l'espace de travail ou de la zone de projet, seul l'appel de l'objet est supprimé. La définition d'objet figure toujours dans la bibliothèque d'objets.

8.19 Recherche d'objets

Dans la bibliothèque d'objets, vous pouvez rechercher des objets définis dans un référentiel ou des objets disponibles par le biais d'une banque de données.

1. Cliquez avec le bouton droit dans la bibliothèque d'objets et sélectionnez [Rechercher](#).

La fenêtre [Rechercher](#) s'affiche.

2. Saisissez les valeurs pertinentes pour la recherche.

Les options disponibles dans la fenêtre [Rechercher](#) sont décrites en détail après cette procédure.

3. Cliquez sur [Rechercher](#).

Les objets correspondant à vos entrées sont affichés dans la fenêtre. Une ligne de statut située en bas de la fenêtre [Rechercher](#) indique où la recherche a été effectuée (localement ou centralement), le nombre d'éléments trouvés et la durée de la recherche.

Dans la fenêtre des résultats de la recherche, le menu contextuel affiche les options suivantes :

- [Ouvrir](#) un élément
- [Importer](#) des tables externes en tant que métadonnées du référentiel
- [Enregistrer sous](#) pour exporter les résultats de la recherche dans un fichier CSV
- [Afficher les cas d'emploi](#) pour afficher les objets parent dans la fenêtre de [sortie](#)
- [Localiser dans la bibliothèque](#) pour sélectionner l'objet dans la bibliothèque locale ou centrale
- [Propriétés](#) pour afficher les attributs

Vous pouvez également faire glisser des objets depuis la fenêtre de résultats de la recherche et les déposer à l'emplacement de votre choix.

La fenêtre [Rechercher](#) permet d'accéder aux options suivantes :

Option	Description
Rechercher dans	Emplacement où effectuer la recherche. Sélectionnez un référentiel ou une banque de données spécifique. Lorsque vous sélectionnez une banque de données, vous pouvez également choisir d'effectuer la recherche dans les données importées (Données internes) ou dans la banque de données complète (Données externes). Vous pouvez également choisir d'effectuer la recherche dans le Référentiel local ou le Référentiel central.
Type d'objet	Type d'objet à rechercher. Dans le cas d'une recherche dans le référentiel, vous disposez des valeurs Tables, Fichiers, Flux de données, Workflows, Jobs, Hiérarchies, IDocs et Domaines. Dans le cas d'une recherche dans une banque de données ou une application, effectuez une sélection parmi les types d'objet disponibles par l'intermédiaire de la banque de données.

Option	Description
<i>Nom</i>	Recherche une chaîne de texte dans le nom de l'objet. Si la recherche porte sur le référentiel, le nom n'est pas sensible à la casse. Si la recherche porte sur une banque de données et que le nom est sensible à la casse dans cette banque de données, saisissez le nom tel qu'il figure dans la banque de données ou l'application et mettez-le entre guillemets doubles (") afin de conserver la casse. Vous pouvez indiquer si l'information à rechercher <i>Contient</i> le nom spécifié ou est <i>Egale</i> au nom spécifié à l'aide de la liste déroulante en regard du champ <i>Nom</i>.
<i>Description</i>	Recherche une chaîne de texte dans la description de l'objet. Les objets importés dans le référentiel ont une définition provenant de leur source. Par défaut, les objets que vous créez dans Designer ne possèdent aucune description, sauf si vous en ajoutez une. La recherche renvoie les objets dont l'attribut de description contient la valeur saisie.
<i>Tout rechercher</i>	Recherche une chaîne de texte dans toutes les parties de l'objet. Pour les jobs, la recherche est effectuée dans le job lui-même et dans tous les éléments du job. (<i>Contient</i> est la seule option de recherche.)

La fenêtre de recherche comporte également un bouton *Avancé* qui permet de rechercher les objets d'après leurs valeurs d'attribut. Vous pouvez uniquement effectuer une recherche par valeurs d'attribut dans le référentiel.

Le bouton *Avancé* permet d'accéder aux options suivantes :

Option	Description
Attribut	Attribut de l'objet dans lequel effectuer la recherche.
Valeur	Valeur de l'attribut à rechercher.
Correspondance	Type de recherche exécuté. Pour rechercher un attribut contenant la valeur indiquée, sélectionnez <i>Contient</i>. Pour rechercher un attribut contenant uniquement la valeur indiquée, sélectionnez <i>Egal à</i>.

Présentation des tâches : [Objets \[page 43\]](#)

Informations associées

[Objets réutilisables \[page 44\]](#)

[Objets à usage unique \[page 46\]](#)

[Save reusable objects \[page 47\]](#)

[Métadonnées de l'objet \[page 48\]](#)

[Descriptions des objets \[page 49\]](#)
[Hiérarchie d'objets \[page 53\]](#)
[Page](#)
[Editeurs d'objets \[page 58\]](#)
[Création d'un objet réutilisable dans la bibliothèque d'objets \[page 59\]](#)
[Création d'un objet réutilisable à l'aide de la palette d'outils \[page 60\]](#)
[Ajout d'un objet existant \[page 61\]](#)
[Modification des noms des objets \[page 62\]](#)
[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)
[Annotations \[page 65\]](#)
[Création de descriptions \[page 66\]](#)
[Copie d'objets \[page 69\]](#)
[Replicating objects \[page 70\]](#)
[Enregistrement et suppression des objets \[page 71\]](#)
[Utilisation de Afficher les cas d'emploi \[page 722\]](#)

9 Projets et Jobs

Les objets Projet et Job représentent les deux niveaux supérieurs d'organisation pour les flux d'application que vous créez à l'aide de Designer.

[Projets \[page 77\]](#)

Un projet désigne un objet réutilisable qui vous permet de regrouper des jobs.

[Jobs \[page 79\]](#)

Un job est le seul objet que vous pouvez exécuter.

9.1 Projets

Un projet désigne un objet réutilisable qui vous permet de regrouper des jobs.

Un projet désigne le niveau le plus élevé d'organisation proposé par le logiciel. L'ouverture d'un projet permet d'accéder facilement à un groupe d'objets dans l'interface utilisateur.

Vous pouvez utiliser un projet pour regrouper des jobs ayant des planifications dépendantes les uns des autres ou pour lesquelles vous souhaitez effectuer un suivi commun..

Les projets ont des caractéristiques communes :

- Les projets sont listés dans la bibliothèque d'objets.
- Vous pouvez ouvrir un seul projet à la fois.
- Les projets ne peuvent pas être partagés entre plusieurs utilisateurs.

[Objets constituant un projet \[page 77\]](#)

Dans un projet, les objets s'affichent hiérarchiquement dans la zone de projet.

[Création d'un projet \[page 78\]](#)

[Ouvrir un projet existant \[page 78\]](#)

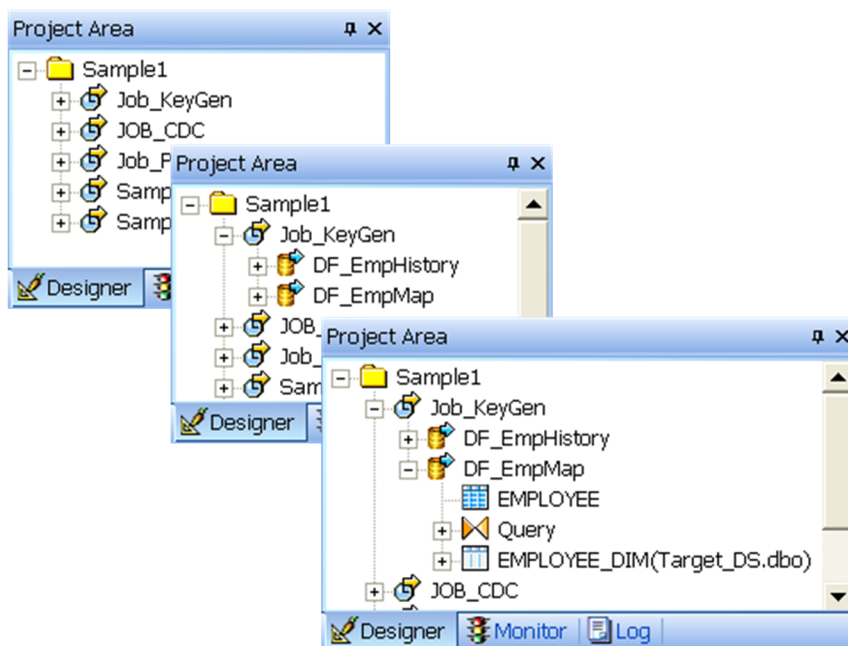
[Enregistrer toutes les modifications apportées à un projet \[page 79\]](#)

9.1.1 Objets constituant un projet

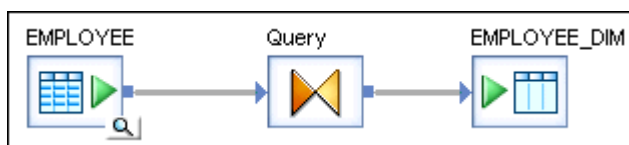
Dans un projet, les objets s'affichent hiérarchiquement dans la zone de projet.

Si un signe plus (+) s'affiche en regard d'un objet, développez-le pour afficher les objets de niveau inférieur figurant dans l'objet. Le logiciel indique le contenu comme noms dans la hiérarchie de zone de projet et des icônes dans l'espace de travail.

Dans l'exemple suivant, le job Job_KeyGen contient deux flux de données et le flux de données DF_EmpMap contient plusieurs objets.



Chaque élément sélectionné dans la zone de projet s'affiche également dans l'espace de travail :



9.1.2 Création d'un projet

1. Sélectionnez **Projet > Nouveau > Projet**.
2. Entrez le nom de votre nouveau projet.

Le nom peut comporter des caractères alphanumériques et des traits de soulignement (_). Il ne peut pas contenir d'espaces.

3. Cliquez sur **Créer**.

Le nouveau projet s'affiche dans la zone de projet. Si vous ajoutez des jobs et d'autres objets de niveau inférieur, ils apparaissent également dans la zone de projet.

9.1.3 Ouvrir un projet existant

1. Sélectionnez **Projet > Ouvrir**.
2. Sélectionnez le nom d'un projet existant dans la liste.
3. Cliquez sur **Ouvrir**.

i Remarque

Si un autre projet a déjà été ouvert, le logiciel le ferme et ouvre le nouveau.

9.1.4 Enregistrer toutes les modifications apportées à un projet

1. Sélectionnez **Projet** > **Tout enregistrer**.

Le logiciel liste les jobs, les workflows et les flux de données modifiés depuis le dernier enregistrement.

2. (facultatif) Désélectionnez tout objet listé pour éviter de l'enregistrer.
3. Cliquez sur **OK**.

i Remarque

Le logiciel vous invite à enregistrer tous les objets modifiés lorsque vous exécutez un job et que vous quittez Designer. L'enregistrement d'un objet réutilisable entraîne l'enregistrement des objets à usage unique qu'il contient le cas échéant.

9.2 Jobs

Un job est le seul objet que vous pouvez exécuter.

Vous pouvez exécuter manuellement des jobs et les tester en mode développement. En mode production, vous pouvez planifier des jobs batch et configurer des jobs en temps réel comme services qui exécutent un processus lorsque le logiciel reçoit une demande de message.

Un job se compose d'étapes que vous souhaitez exécuter ensemble. Chaque étape est représentée par une icône objet que vous placez dans l'espace de travail pour créer un diagramme de jobs. Un diagramme de jobs se compose de deux objets ou plus reliés ensemble. Vous pouvez inclure l'un des objets suivants dans une définition de job :

- Flux de données
 - Sources
 - Cibles
 - Transformations
- Workflows
 - Scripts
 - Structures conditionnelles
 - Boucles WHILE
 - Blocs Try/Catch

Si un job devient complexe, organisez son contenu en workflows individuels puis créez un job individuel qui appelle ces workflows.

Les jobs en temps réel utilisent les mêmes composants que les jobs batch. Vous pouvez ajouter des workflows et des flux de données aux jobs batch et aux jobs en temps réel. Lorsque vous faites glisser une icône de workflow ou de flux de données dans un job, vous demandez au logiciel de valider ces objets selon les exigences du type de job (batch ou temps réel).

Il existe certaines restrictions en ce qui concerne l'utilisation de certaines fonctions du logiciel avec les jobs en temps réel.

i Remarque

Si un job utilisant la fonctionnalité de chargement par lot échoue, Data Services enregistre des fichiers de données contenant des données client dans le répertoire de chargement par lots. Cette opération est effectuée pour que vous puissiez examiner et analyser les données. L'emplacement de chargement par lots par défaut est %DS_COMMON_DIR%/log/BulkLoader. Il est de votre responsabilité de supprimer les fichiers après leur analyse.

[Créer un job dans la zone de projet \[page 80\]](#)

[Créer un job dans la bibliothèque d'objets \[page 80\]](#)

[Conventions d'appellation pour objets dans jobs \[page 81\]](#)

Informations associées

[Workflows \[page 237\]](#)

[Jobs en temps réel \[page 340\]](#)

9.2.1 Créer un job dans la zone de projet

1. Dans la zone de projet, sélectionnez le nom du projet.
2. Cliquez avec le bouton droit de la souris et sélectionnez *Nouveau job* ou *Job en temps réel*.
3. Modifiez le nom.

Le nom peut comporter des caractères alphanumériques et des traits de soulignement (_). Il ne peut pas contenir d'espaces.

Le logiciel ouvre un nouvel espace de travail où vous définissez le job.

9.2.2 Créer un job dans la bibliothèque d'objets

1. Accédez à l'onglet *Jobs*.
2. Cliquez avec le bouton droit de la souris sur *Jobs batch* ou *Jobs en temps réel* et sélectionnez *Nouveau*.
3. Un nouveau job s'affiche avec un nom par défaut.
4. Cliquez avec le bouton droit de la souris et sélectionnez *Propriétés* pour modifier le nom des objets et ajouter une description.

Le nom peut comporter des caractères alphanumériques et des traits de soulignement (_). Il ne peut pas contenir d'espaces.

5. Pour ajouter le job au projet en cours, faites-le glisser dans la zone de projet.

9.2.3 Conventions d'appellation pour objets dans jobs

Nous vous recommandons de suivre les conventions d'appellation cohérentes pour simplifier l'identification d'objet dans tous les systèmes de votre entreprise. Vous pourrez ainsi travailler plus facilement avec des métadonnées dans toutes les applications, par exemple :

- Applications de modélisation des données
- Applications ETL
- Applications de reporting
- Kits de développement de logiciel Adapter

Des exemples de conventions recommandées pour l'utilisation de jobs et autres objets sont indiqués dans le tableau suivant.

Préfixe	Suffixe	Objet	Exemple
DF_	n/a	Flux de données	DF_Currency
EDF_	_Input	Flux de données imbriquées	EDF_Example_Input
EDF_	_Output	Flux de données imbriquées	EDF_Example_Output
RTJob_	n/a	Job en temps réel	RTJob_OrderStatus
WF_	n/a	Workflow	WF_SalesOrg
JOB_	n/a	Job	JOB_SalesOrg
n/a	_DS	Banque de données	ORA_DS
DC_	n/a	Configuration de banque de données	DC_DB2_production
SC_	n/a	Configuration du système	SC_ORA_test
n/a	_Memory_DS	Banque de données en mémoire	Catalog_Memory_DS
PROC_	n/a	Procédure stockée	PROC_SalesStatus

Bien que Designer soit une interface utilisateur graphique avec des icônes représentant des objets dans sa fenêtre, d'autres interfaces peuvent vous amener à identifier les types d'objets par le texte seul. En utilisant un préfixe ou un suffixe vous pouvez plus facilement identifier votre type d'objet.

Outre les préfixes et les suffixes, vous pouvez être amené à fournir des noms standardisés pour des objets identifiant une action spécifique dans tous les types d'objets. Par exemple : DF_OrderStatus, RTJob_OrderStatus.

Outre les préfixes et les suffixes, les conventions d'appellation peuvent également inclure des identificateurs de nom du chemin. Par exemple, la convention d'appellation de procédure stockée peut être la suivante :

<datastore>.<owner>.<PROC_Name>

<datastore>.<owner>.<package>.<PROC_Name>

10 Qu'est-ce qu'une banque de données

Les banques de données représentent des configurations de connexion entre le logiciel et les bases de données ou les applications.

Ces configurations peuvent être directes ou via un adaptateur. Les configurations de banque de données permettent au logiciel d'accéder aux métadonnées d'une base de données ou d'une application, et de lire depuis ou d'écrire dans cette base de données ou application pendant que le logiciel exécute un job.

Les banques de données SAP Data Services peuvent se connecter à :

- Des bases de données et des systèmes de fichier de l'ordinateur principal.
- Des applications qui ont des adaptateurs pré-packagés ou écrits par l'utilisateur.
- J.D. Edwards One World et J.D. Edwards World, applications Oracle, PeopleSoft, applications SAP, SAP Data Quality Management, microservices pour les données d'emplacement, SAP NetWeaver BW, applications Siebel et Google BigQuery. Pour plus d'informations, voir les guides suppléments appropriés.
- Serveurs à distance utilisant FTP, SFTP et SCP.

i Remarque

Le logiciel lit et écrit des données stockées dans des fichiers plats à l'aide de formats de fichiers plats. Le logiciel lit et écrit des données stockées dans des documents XML à l'aide de schémas DTD et XML.

Les informations spécifiques auxquelles un objet de banque de données peut accéder dépendent de la configuration de connexion. Lorsque la base de données ou l'application est modifiée, apportez les modifications correspondantes dans les informations de banque de données dans le logiciel. Le logiciel ne détecte pas automatiquement les nouvelles informations.

i Remarque

Les objets supprimés d'une connexion à la banque de données sont identifiés dans la zone du projet et dans l'espace de travail par une icône rouge "supprimé".  Cet indicateur visuel permet de trouver et de mettre à jour les flux de données affectés par les modifications de la banque de données.

Il est possible de créer plusieurs configurations pour une banque de données. Cela vous permet de prévoir en avance les différents environnements dans lesquels la banque de données peut être utilisée et limite le travail impliqué par la migration de jobs. Par exemple, il est possible d'ajouter un ensemble de configurations (DEV, TEST et PROD) au même nom de banque de données. Ces paramètres de connexion restent avec la banque de données lors de l'exportation ou de l'importation. Il est possible de grouper tous les ensembles de configurations de la banque de données dans une configuration système. Lors de l'exécution ou de la planification d'un job, sélectionnez une configuration système et ainsi l'ensemble des configurations de la banque de données pour l'environnement actuel.

[Banques de données de bases de données \[page 83\]](#)

Permet à Data Services d'effectuer des opérations de lecture et d'écriture dans les types de base de données pris en charge.

[Banques de données en mémoire \[page 108\]](#)

Une banque de données en mémoire est un conteneur pour des tables en mémoire.

[Banques de données de cache persistant \[page 112\]](#)

Une banque de données de cache persistant est un conteneur pour ces tables de cache.

[Banques de données liées \[page 117\]](#)

Plusieurs fournisseurs de base de données prennent en charge des chemins de communication à sens unique entre un serveur de base de données à un autre. Oracle appelle ces chemins liens de base de données.

[Banques de données de l'adaptateur \[page 120\]](#)

Les banques de données d'adaptateur permettent d'accéder aux données d'application, accessibles via l'adaptateur.

[Banques de données de l'application \[page 121\]](#)

Configurez une banque de données de l'application avec les informations de connexion à l'application.

[Banques de données de services Web \[page 123\]](#)

Les banques de données de services Web représentent une connexion depuis Data Services vers une source de données externe basée sur les services Web.

[Modification d'une définition de banque de données \[page 125\]](#)

Comme tous les objets, les banques de données sont définies par des options et des propriétés.

[Création et gestion de plusieurs configurations de banques de données \[page 126\]](#)

La création de plusieurs configurations pour une seule banque de données permet de consolider des connexions aux banques de données distinctes pour des sources ou des cibles similaires dans une banque de données source ou cible avec plusieurs configurations.

Informations associées

[Banques de données de bases de données \[page 83\]](#)

[Formats de fichiers \[page 149\]](#)

[Mise en forme des documents XML \[page 259\]](#)

[Création et gestion de plusieurs configurations de banques de données \[page 126\]](#)

10.1 Banques de données de bases de données

Permet à Data Services d'effectuer des opérations de lecture et d'écriture dans les types de base de données pris en charge.

Les banques de données de bases de données représentent des connexions uniques ou multiples avec :

- Des systèmes hérités à l'aide de Attunity Connect
- Bases de données HP Vertica, IBM DB2, Informix, Microsoft SQL Server, MySQL, Netezza, Oracle, SAP ASE, SAP Data Federator, SAP SQL Anywhere, SAP HANA, SAP Vora, Sybase IQ et Teradata (utilisant des connexions natives)
- D'autres bases de données (par ODBC)
- Un référentiel, à l'aide d'une banque de données en mémoire ou d'une banque de données de cache persistant

Vous créez une connexion à la plupart des sources de données à l'aide du nom de serveur au lieu du nom du DSN (nom de source de données) ou du nom du TNS (Transparent Network Substrate). Les connexions avec nom de serveur (également appelées connexions sans DSN et connexions sans TNS) suppriment le besoin de configurer les mêmes entrées DSN ou TNS sur chaque ordinateur dans un environnement distribué.

Pour en savoir plus sur les connexions sans DSN et sans TNS, voir le *Guide d'administration*.

[Définition d'une banque de données d'une base de données \[page 85\]](#)

Définissez au moins une banque de données d'une base de données pour chaque base de données ou système de fichier central avec lesquels vous échangez des données.

[Interface de l'ordinateur central \[page 88\]](#)

Le logiciel fournit la banque de données du connecteur Attunity qui accède aux sources de données de l'ordinateur central par Attunity Connect.

[Banques de données Amazon Redshift \[page 91\]](#)

Utiliser une banque de données Amazon Redshift pour importer et charger les tables, charger des fichiers de données Amazon S3, et plus encore.

[Apache Impala \[page 92\]](#)

Créez une banque de données de base de données pour Apache Impala, qui est une base de données source pour Apache Hadoop.

[Banques de données Hive \[page 92\]](#)

Utilisez une banque de données Hive pour accéder aux données depuis votre entrepôt Hive.

[HP Vertica datastore \[page 93\]](#)

Use the HP Vertica datastore to import HP Vertica tables to use as source or targets in a data flow.

[Création d'une banque de données SAP HANA avec cryptage SSL \[page 95\]](#)

Création d'une banque de données SAP HANA avec cryptage SSL. Le cryptage SSL protège les données lors du transfert entre le serveur de base de données et Data Services.

[Création d'une banque de données Microsoft SQL Server avec cryptage SSL \[page 96\]](#)

Création d'une banque de données Microsoft SQL Server avec cryptage SSL. Le cryptage SSL protège les données lors du transfert entre le serveur de base de données et Data Services. Data Services utilise la connexion automatique par défaut de SQL Server. Un certificat SSL n'est donc pas nécessaire, et son utilisation n'est d'ailleurs pas prise en charge par Data Services.

[Banque de données SAP Vora \[page 97\]](#)

Avec une banque de données SAP Vora, accédez aux tables Vora à l'aide du pilote ODBC SAP HANA et du protocole de connexion unidirectionnelle SAP HANA.

[Parcourir les métadonnées depuis une banque de données de base de données \[page 99\]](#)

Vous pouvez afficher les métadonnées pour les objets importés ou non importés et pour contrôler si les métadonnées ont changé pour les objets déjà importés.

[Importation de métadonnées par une banque de données de base de données \[page 101\]](#)

Pour les banques de données de base de données, il est possible d'importer les métadonnées pour des tables et des fonctions.

Rubrique parent : [Qu'est-ce qu'une banque de données \[page 82\]](#)

Informations associées

[Banques de données en mémoire \[page 108\]](#)

[Banques de données de cache persistant \[page 112\]](#)

[Banques de données liées \[page 117\]](#)

[Banques de données de l'adaptateur \[page 120\]](#)

[Banques de données de l'application \[page 121\]](#)

[Banques de données de services Web \[page 123\]](#)

[Modification d'une définition de banque de données \[page 125\]](#)

[Création et gestion de plusieurs configurations de banques de données \[page 126\]](#)

10.1.1 Définition d'une banque de données d'une base de données

Définissez au moins une banque de données d'une base de données pour chaque base de données ou système de fichier central avec lesquels vous échangez des données.

Pour définir une banque de données de base de données, obtenez les privilèges d'accès à la base de données ou au système de fichier décrits par la banque de données.

Par exemple, le fait de permettre au logiciel d'utiliser le SQL paramétré lors de la lecture ou de l'écriture sur des bases de données DB2 autorise l'utilisateur (de la banque de données/base de données) à créer, exécuter et déposer les procédures stockées. Si l'utilisateur n'est pas autorisé à créer, exécuter et déposer les procédures stockées, les jobs continuent d'exécuter. Cependant, les jobs produiront un message d'avertissement et exécuteront moins efficacement.

1. Dans l'onglet *Banques de données* de la bibliothèque d'objets, cliquez avec le bouton droit de la souris et sélectionnez *Nouveau*.
2. Saisissez le nom de la nouvelle banque de données dans le champ *Nom de la banque de données*.

Le nom peut contenir tous les caractères alphabétiques ou numériques, ou des tirets de soulignement (_). Il ne peut pas contenir d'espace.

3. Dans la zone *Type de banque de données*, sélectionnez *Base de données*.

Le logiciel affiche les options pertinentes pour ce type. Vous trouverez de plus amples informations sur ces options et sur les types de base de données qui ne sont pas traités ici dans la section "Banques de données de base de donnée" du *Guide de référence*.

La table suivante contient des informations sur les situations spéciales liées aux types de base de données :

Type de base de données	Informations
Oracle	Le type de base de données Oracle prend en charge les connexions sans TNS. Pour utiliser une connexion sans TNS, désélectionnez <i>Utiliser le nom du TNS</i> et saisissez le nom de l'hôte, le SID et le port.
Data Federator	Lorsque vous utilisez ce type de base de données, vous devez spécifier le nom du catalogue et le nom du schéma dans l'URL. Si vous ne le faites pas, il est possible que toutes les tables de chaque catalogue s'affichent. 1. Sélectionnez Admin ODBC puis l'onglet DSN système. 2. Mettez en surbrillance <i>Data Federator</i> puis cliquez sur <i>Configurer</i>. 3. Dans l'option URL, saisissez le nom du catalogue et le nom du schéma. Par exemple : <pre>jdbc:leselect://localhost/ <NomCatalogue>;schema=<NomSchéma></pre>
SAP HANA	Pour les versions antérieures à SAP HANA 2.0 SPS 01 MDC, les banques de données de base de données accèdent à une base de données unique. Pour les versions SAP HANA 2.0 SPS 01 MDC et versions ultérieures, les banques de données de base de données accèdent à une base de données mutualisée, définie.

De plus, si vous sélectionnez les connexions sans DSN (vous désélectionnez l'option *Utiliser le nom de source de données*), les informations suivantes s'appliquent :

- Applicable uniquement aux types de base de données suivants : DB2, Informix, MySQL, Netezza, ou SAP HANA. Pour obtenir la dernière liste des bases de données prises en charge pour les connexions sans DSN, voir les *Notes de version*.
 - Saisissez les informations dans les options suivantes : Nom du serveur de base de données, Nom de base de données (pour DB2 et MySQL), Informations de ports, Nom d'utilisateur et Mot de passe.
 - Si vous sélectionnez Informix et que vous souhaitez utiliser des connexions sans DSN lorsque Data Services se trouve sur un autre ordinateur que le serveur Informix, vous devez identifier l'hôte Informix comme suit :
 1. Accédez au dossier d'installation du client Informix (par exemple : C:\Program Files\IBM\Informix\Client-SDK\bin Run setnet32.exe).
 2. Dans l'onglet d'informations sur le serveur, saisissez le nom du serveur Informix IBM, le nom d'hôte et les autres informations requises.
 3. Faites du serveur Informix IBM le serveur par défaut.
 - Pour les connexions sans DSN à une base de données Informix, le Designer peut obtenir le nom de l'hôte Informix pour le nom de serveur Informix indiqué.
4. Sélectionnez *Activer le transfert automatique des données* pour permettre à la transformation Data_Transfer d'utiliser les tables de transfert dans cette banque de données pour pousser vers le bas les

opérations de base de données ultérieures. Cette case à cocher s'affiche pour toutes les bases de données, sauf le connecteur Attunity, Data Federator, le cache mémoire et le cache persistant.

5. Pour ajouter plus d'informations, sélectionnez [Avancé](#).

Cliquez sur la cellule sous chaque option de configuration et saisissez ou sélectionnez une valeur.

6. Si vous souhaitez que le logiciel convertisse un type de données de votre source qu'il ne prend normalement pas en charge, sélectionnez [Importer les types de données non pris en charge comme VARCHAR de taille](#) et saisissez le nombre de caractères que vous autorisez. Vous trouverez de plus amples informations sur les types de données dans le *Guide de référence*.
7. Cliquez sur [OK](#) pour enregistrer la banque de données de base de données.

i Remarque

Sur les versions de Data Integrator antérieures à la version 11.7.0, le type de base de données approprié à utiliser lors de la création d'une banque de données sur Netezza était ODBC. SAP Data Services 11.7.1 fournit une option Netezza spécifique comme type de base de données à la place d'ODBC. Lors de l'utilisation de Netezza comme base de données avec le logiciel, nous recommandons de sélectionner l'option Netezza du logiciel comme type de base de données plutôt qu'ODBC.

Informations associées

[Création et gestion de plusieurs configurations de banques de données \[page 126\]](#)

10.1.1.1 Limiting imported Data Federator catalogs

Limit Data Federator tables so that you do not import tables from all catalogs.

To limit the tables to tables in a specific catalog, specify the catalog name and the schema name in the URL. In the datastore editor for Data Federator:

1. Select [ODBC Admin](#) and then the [System DSN](#) tab.
2. Highlight [Data Federator](#) and then click [Configure](#).
3. In the URL option, enter the catalog name and the schema name. For example: `jdbc:lesselect://localhost/<catalogname>;schema=<schemaname>`

10.1.1.2 Configuring DSN-less connection for Informix that uses a different server

When your Data Services installation is on a different computer than your Informix server, follow a different procedure to configure a DSN-less connection.

Identify the Informix host as follows:

1. Go to your Informix client installation folder. For example: `C:\Program Files\IBM\Informix\Client-SDK\bin` Run `setnet32.exe`.

2. In the [Server Information](#) tab, enter the name of the IBM Informix server, the host name, and other required information.
3. Make the IBM Informix server the default server.

Designer obtains the Informix host name for the Informix server name you provided.

10.1.2 Interface de l'ordinateur central

Le logiciel fournit la banque de données du connecteur Attunity qui accède aux sources de données de l'ordinateur central par Attunity Connect.

Les sources de données auxquelles Attunity Connect accède sont dans la liste suivante : Pour une liste complète des sources, consultez la documentation d'Attunity.

- Adabas
- DB2 UDB pour OS/390 et DB2 UDB pour OS/400
- IMS/DB
- VSAM
- Fichiers plats sur OS/390 et fichiers plats sur OS/400

10.1.2.1 Conditions préalables pour une banque de données Attunity

Le connecteur Attunity accède aux données de l'ordinateur central à l'aide du logiciel que vous devez installer manuellement sur le serveur de l'ordinateur central et sur l'ordinateur du client local (Job Server). Le logiciel se connecte au connecteur Attunity à l'aide de son interface ODBC.

Il n'est pas nécessaire d'acheter un gestionnaire de pilotes ODBC distinct pour UNIX et pour les plateformes Windows.

Serveurs

Installez et configurez le produit Attunity Connect sur le serveur (par exemple, un ordinateur z Series).

Clients

Pour accéder aux données de l'ordinateur central à l'aide du connecteur Attunity, installez le produit Attunity Connect. Le pilote ODBC est requis. Attunity offre également un outil facultatif appelé Attunity Studio, qu'il est possible d'utiliser pour la configuration et l'administration.

Configurez les sources de données ODBC sur le client (Job Server de SAPData Services).

Lorsque vous installez un Job Server sur UNIX, l'installateur vous invite à fournir un chemin du répertoire d'installation pour le logiciel du connecteur Attunity. De plus, il n'est pas nécessaire d'installer un gestionnaire de pilotes parce que le logiciel charge les pilotes ODBC directement sur les plateformes UNIX.

Pour plus d'informations sur la manière d'installer et de configurer ces produits, consultez leur documentation.

10.1.2.2 Création et configuration d'une banque de données Attunity

Pour utiliser l'option de la banque de données du connecteur Attunity, mettez à niveau le référentiel vers la version 6.5.1 ou ultérieure de SAP Data Services.

Pour créer et configurer une banque de données du connecteur Attunity :

1. Dans l'onglet *Banques de données* de la bibliothèque d'objets, faites un clic droit et sélectionnez *Nouveau*.
2. Saisissez un nom pour la banque de données.
3. Dans la zone *Type de banque de données*, sélectionnez *Base de données*.
4. Dans la case *Type de la base de données*, sélectionnez *Connecteur Attunity*.
5. Saisissez le nom de la source de données Attunity, l'emplacement du démon Attunity (*Emplacement de l'hôte*), le numéro du port du démon Attunity, et un nom unique de l'espace de travail du serveur Attunity.
6. Pour modifier les options par défaut (telles que *Lignes par validation* ou *Langue*), cliquez sur le bouton *Avancé*.
7. Cliquez sur *OK*.

Vous pouvez maintenant utiliser la nouvelle connexion de banque de données pour importer les tables de métadonnées dans le référentiel actuel.

10.1.2.3 Spécification de plusieurs sources de données dans une banque de données Attunity

La banque de données du connecteur Attunity permet d'accéder à plusieurs sources de données Attunity sur le même emplacement de démon Attunity.

Si plusieurs types de données existent sur le même ordinateur, par exemple une base de données DB2 et VSAM, il est possible que vous souhaitiez accéder aux deux types de données à l'aide d'une seule connexion. Par exemple, vous pouvez utiliser une connexion unique pour relier les tables (et pousser vers le bas l'opération de jointure vers un serveur distant), ce qui réduit la quantité de données transmises par le réseau.

Pour spécifier plusieurs sources dans l'éditeur de banque de données :

1. Séparez les noms des sources de données par des points virgules dans la case de la source de données Attunity à l'aide du format suivant :

```
AttunityDataSourceName;AttunityDataSourceName
```

Par exemple, si vous possédez une source de données DB2 appelée DSN4 et une source de données VSAM appelée Navdemo, saisissez les valeurs suivantes dans la case Source de données :

```
DSN4 ; Navdemo
```

2. Si vous faites une liste de plusieurs noms de source de données pour une banque de données du connecteur Attunity, assurez-vous de respecter les conditions suivantes :
 - Toutes les sources de données Attunity doivent être accessibles avec le même nom d'utilisateur et mot de passe.
 - Toutes les sources de données Attunity doivent utiliser le même espace de travail. Lorsque vous configurez l'accès aux sources de données dans Attunity Studio, utilisez le même nom d'espace de travail pour chaque source de données.

10.1.2.4 Convention d'appellation Data Services pour les tables Attunity

Le format Data Services pour l'accès aux tables Attunity est particulier à Data Services.

Puisqu'une seule banque de données peut accéder à plusieurs systèmes logiciels qui ne partagent pas le même espace de noms, le nom de la source de données Attunity doit être indiqué lorsqu'elle fait référence à une table. Avec un connecteur Attunity, faites précéder le nom de la table avec les noms de la source de données et du propriétaire, séparés par deux points. Le format est le suivant :

```
AttunityDataSource:OwnerName.TableName
```

Lors de l'utilisation de Designer pour créer vos jobs avec des tables Attunity importées, Data Services génère automatiquement le SQL correct pour ce format. Cependant, lorsque vous créez SQL, assurez-vous d'utiliser ce format. Il est possible de créer SQL dans les constructions suivantes :

- fonction SQL ;
- transformation SQL ;
- fonction pushdown_sql ;
- commandes de pré-chargement dans un chargeur de table ;
- commandes de post-chargement dans un chargeur de table.

i Remarque

Pour les tables de Data Services, la longueur maximum du nom du propriétaire pour la plupart des types de référentiels est de 256 caractères (64 pour MySQL et 128 pour MS SQL server). Dans le cas des tables Attunity, la longueur maximum du nom de la source de données Attunity et du nom réel du propriétaire est de 63 caractères (les deux-points comptent pour 1 caractère).

10.1.2.5 Limitations pour la banque de données Attunity

Toutes les fonctionnalités Data Services sont disponibles lorsque vous utilisez une banque de données du connecteur Attunity, excepté pour quelques fonctionnalités.

- Chargement par lots
- Fonctions importées (importe les métadonnées pour les tables uniquement)
- Tables de modèle (création de tables)
- Le type de données date-heure prend en charge jusqu'à 2 multiples décimaux de secondes uniquement
- Data Services ne peut pas charger des données d'horodatage dans une colonne d'horodatage dans une table car Attunity tronque les données varchar à 8 caractères, ce qui n'est pas assez pour représenter correctement une valeur d'horodatage.
- Lors de l'exécution d'un job sur UNIX, le job peut échouer avec l'erreur suivante :

```
[D000] Cannot open file /usr1/attun/navroot/def/sys System error 13: The file access permissions do not allow the specified action.; (OPEN)
```

Cette erreur se produit à cause des autorisations de fichier insuffisantes pour certains des fichiers dans le répertoire d'installation Attunity. Pour éviter cette erreur, remplacez les autorisations de fichier pour tous les fichiers dans le répertoire Attunity par 777 en exécutant la commande suivante depuis le répertoire d'installation Attunity :

```
$ chmod -R 777 *
```

10.1.3 Banques de données Amazon Redshift

Utiliser une banque de données Amazon Redshift pour importer et charger les tables, charger des fichiers de données Amazon S3, et plus encore.

Utilisez une banque de données de base de données Redshift pour les tâches suivantes :

- Importation de tables
- Lecture ou chargement de tables Redshift dans un flux de données
- Aperçu des données
- Création et importation de modèles de tables
- Chargement des fichiers de données Amazon S3 dans une table Redshift à l'aide de la fonction intégrée `load_from_s3_to_redshift`.

Le tableau suivant décrit les options spécifiques à Redshift lors de la création ou de la modification d'une banque de données.

Options de la fenêtre principale

Option	Description
Activer le transfert automatique des données	Sélectionnez cette option pour activer les tables de transfert dans cette banque de données. La transformation Data_Transfer utilise les tables de transfert pour pousser vers le bas les opérations de base de données ultérieures. Cette option est activée par défaut.

Utilisez un pilote ODBC Amazon Redshift pour vous connecter à la base de données du cluster Redshift. Le pilote ODBC Redshift se connecte à Redshift uniquement sur les plateformes Windows et Linux.

Pour plus d'informations concernant le téléchargement et l'installation du pilote ODBC Amazon Redshift, consultez la documentation Amazon Redshift sur le site Internet d'Amazon.

i Remarque

Activez les paramètres de SSL (Secure Socket Layer) dans le pilote ODBC Amazon Redshift. Dans la fenêtre [Configuration du DSN du pilote ODBC Amazon Redshift](#), définissez l'option [Authentification SSL](#) sur [autoriser](#).

Pour en savoir plus sur la prise en charge d'Amazon Redshift, voir le *Supplément pour le Big Data*.

10.1.4 Apache Impala

Créez une banque de données de base de données pour Apache Impala, qui est une base de données source pour Apache Hadoop.

Avant de créer une banque de données Impala Apache, importez le pilote Cloudera et créez un nom de source de données (DSN). Pour créer une banque de données Apache Impala, ouvrez l'éditeur de banque de données et sélectionnez ODBC pour le [Type de données](#). Sélectionnez ensuite le DSN que vous avez créé pour cette banque de données et renseignez les options restantes, le cas échéant.

Utilisez la banque de données pour importer les tables Impala et utiliser les données d'Impala en tant que source ou cible dans un flux de données.

Avant de travailler avec Apache Impala, tenez compte des restrictions suivantes :

- Data Services prend en charge Impala 2.5 et versions ultérieures.
- Data Services prend uniquement en charge les types de données scalaires Impala. Data Services ne prend pas en charge les types complexes tels qu'ARRAY, STRUCT ou MAP.

10.1.5 Banques de données Hive

Utilisez une banque de données Hive pour accéder aux données depuis votre entrepôt Hive.

Hive est un entrepôt de données qui facilite la lecture, l'écriture et la gestion de grands jeux de données, résidant dans l'archivage réparti, comme HDFS. Pour demander Hive, utilisez HiveQL, un type de syntaxe SQL.

Les banques de données Hive utilisent l'adaptateur Hive pour se connecter à votre entrepôt Hive. Data Services prend en charge deux types de banques de données Hive :

- Banque de données de l'adaptateur Hive : Data Services est installé au sein du cluster Hadoop. Configurez la banque de données sans utiliser le nom de source de données (DSN) ou un pilote ODBC de Hive.
- Banque de données de la base de données Hive : Data Services est installé sur un ordinateur. Configurez la banque de données via DSN et un pilote ODBC Hive pris en charge.

Utilisez l'un des types de banque de données Hive pour les tâches suivantes :

- Lire et écrire dans le serveur Hive à l'aide des tables Hive comme tables source ou cible.
- Utiliser un modèle de table Hive dans votre flux de données.
- Afficher un aperçu des données de table Hive.

Pilotes ODBC Hive pris en charge

La banque de données de base de données Hive prend en charge les pilotes ODBC suivants :

- Cloudera
- Hortonworks
- MapR

Pour plus d'informations sur les versions de pilote actuellement prises en charge, reportez-vous à la *Matrice de disponibilité des produits (PAM)*.

Restrictions

- Les opérations telles que DELETE, UPDATE et UPINSERT ne sont pas prises en charge nativement par le serveur Hive.
- Le SQL paramétré n'est pas pris en charge car le serveur Hive ne prend pas en charge le marqueur de paramètre.

10.1.6 HP Vertica datastore

Use the HP Vertica datastore to import HP Vertica tables to use as source or targets in a data flow.

After you create the HP Vertica database datastore, you can import HP Vertica tables into Data Services. Use the tables as sources or targets in dataflows, and create HP Vertica template tables.

Configure an HP Vertica database datastore using a supported driver and a DSN (data source name). See the *Product Availability Matrix (PAM)* for supported drivers.

SSL protocol is available for HP Vertica database datastores. Before you can create an SSL-enabled HP Vertica datastore, ensure the following:

- You have installed and configured MIT Kerberos 5

- For Windows operating systems, you have created a DSN in the ODBC Data Source Administrator
- For UNIX operating systems, you have created a DSN in the Data Services Connection Manager for UNIX

For descriptions of HP Vertica datastore options, see the *Reference Guide*.

For instruction to enable MIT Kerberos and to create a DSN, see the *Administrator Guide*.

10.1.6.1 Création d'une banque de données HP Vertica avec cryptage SSL

Le cryptage SSL protège les données lors du transfert entre le serveur de base de données et Data Services.

Un administrateur doit installer MIT Kerberos 5 et activer Kerberos pour le protocole SSL HP Vertica. De plus, un administrateur doit créer un nom de source de données (DSN) SSL à l'aide de l'administrateur de source de données ODBC afin qu'il soit disponible lors de la création de la banque de données. Voir le *Guide d'administration* pour plus d'informations sur la configuration de MIT Kerberos.

Le cryptage SSL pour HP Vertica est disponible dans SAP Data Services version 4.2 Support Package 7 Patch 1 (14.2.7.1) or versions ultérieures.

i Remarque

L'activation du cryptage SSL ralentit la performance des jobs.

i Remarque

Une banque de données de base de données HP Vertica requiert de sélectionner DSN comme méthode de connexion. Les connexions sans DSN ne sont pas autorisées pour la banque de données HP Vertica avec cryptage SSL.

1. Dans Designer sélectionnez ► *Projet* ► *Nouveau* ► *Banque de données* ►.
2. Renseignez les options comme vous le feriez pour une banque de données de base de données HP Vertica. Renseignez les options suivantes spécifiquement pour le cryptage SSL :

Options spécifiques à SS

Option	Valeur
<i>Utiliser le nom de source de données (DSN)</i>	Sélectionnez cette option.
<i>Nom de la source de données</i>	Sélectionnez le fichier de source de données DSN SSL HP Vertica précédemment créé dans l'administrateur de source de données ODBC.

3. Renseignez les options avancées applicables restantes puis enregistrez votre banque de données.

10.1.6.2 Chargement par lots pour HP Vertica

Configurez une banque de données de base de données HP Vertica pour le chargement par lots en augmentant la taille de validation dans le chargeur et en utilisant l'option d'équilibrage de charge de connexion native, lorsque vous configurez le pilote ODBC.

Consultez le Guide *Connexion à HP Vertica* à l'adresse https://my.vertica.com/docs/7.1.x/PDF/HP_Vertica_7.1.x_ConnectingToHPVertica.pdf .

Aucune option spécifique de chargement par lots définie lorsque vous créez une banque de données de base de données. Cependant, lorsque vous chargez des données dans une cible HP Vertica, dans un flux de données, le logiciel exécute automatiquement une instruction HP Vertica contenant une instruction `COPY LOCAL`. Via cette instruction le pilote ODBC est en lecture et le flux de données de fichier s'effectue du client au serveur.

Vous pouvez augmenter encore davantage la vitesse de chargement en effectuant les paramétrages suivants dans Designer :

1. Augmenter les lignes par validation en source :
 - a. Dans Designer, ouvrez le flux de données applicable.
 - b. Dans l'espace de travail, cliquez deux fois sur l'objet cible de banque de données HP Vertica pour l'ouvrir.
 - c. Ouvrez l'onglet *Options* dans le volet inférieur.
 - d. Augmentez le nombre de lignes dans l'option *Lignes par validation*.
2. Dans HP Vertica, activez l'option permettant d'utiliser l'équilibrage de chargement de connexion native lorsque vous configurez le pilote ODBC pour HP Vertica.

10.1.7 Création d'une banque de données SAP HANA avec cryptage SSL

Création d'une banque de données SAP HANA avec cryptage SSL. Le cryptage SSL protège les données lors du transfert entre le serveur de base de données et Data Services.

Un administrateur doit importer et configurer le certificat de base de données SAP HANA. De plus, vous devez créer une source de données (DSN) SSL afin qu'elle soit disponible lors de la création de la banque de données. Des informations sur l'importation et la configuration du certificat de base de données SAP HANA sont disponibles dans le *Guide d'administration*.

Le cryptage SSL est disponible dans SAP Data Services, version 4.2 SP7 (14.2.7.0) ou version ultérieure.

i Remarque

L'activation du cryptage SSL ralentira la performance des jobs.

i Remarque

Une banque de données de base de données SAP HANA requiert de sélectionner DSN comme méthode de connexion. Les connexions sans DSN ne sont pas autorisées lors de l'activation du cryptage SSL.

i Remarque

Si vous utilisez le conteneur de base de données mutualisée (MDC) de la version 2.0 SPS 01 SAP HANA ou version ultérieure, spécifiez le numéro de port et le nom du serveur de base de données spécifique à la base de données mutualisée à laquelle vous accédez.

1. Dans Designer sélectionnez ► *Projet* ► *Nouveau* ► *Banque de données* ►.
2. Renseignez les options comme vous le feriez pour une banque de données de base de données SAP HANA. Renseignez les options suivantes spécifiquement pour le cryptage SSL :

Options spécifiques à SS

Option	Valeur
<i>Utiliser le nom de source de données (DSN)</i>	Sélectionner
<i>Nom de la source de données</i>	Sélectionnez le fichier de source de données DSN SSL SAP HANA qui a été précédemment créé (voir Conditions préalables ci-dessus).

Trouvez des descriptions pour toutes les options des banques de données de base de données SAP HANA dans le *Guide de référence*.

3. Renseignez les options Avancées applicables restantes puis enregistrez votre banque de données.

10.1.8 Création d'une banque de données Microsoft SQL Server avec cryptage SSL

Création d'une banque de données Microsoft SQL Server avec cryptage SSL. Le cryptage SSL protège les données lors du transfert entre le serveur de base de données et Data Services. Data Services utilise la connexion automatique par défaut de SQL Server. Un certificat SSL n'est donc pas nécessaire, et son utilisation n'est d'ailleurs pas prise en charge par Data Services.

Un administrateur doit importer et configurer le certificat de base de données Microsoft SQL Server. Des informations sur l'importation et la configuration du certificat de base de données Microsoft SQL Server sont disponibles dans le *Guide d'administration*.

Le cryptage SSL pour les banques de données de base de données Microsoft SQL Server est disponible dans SAP Data Services, version 4.2 SP7 (14.2.7.0) ou version ultérieure.

i Remarque

L'activation du cryptage SSL ralentira la performance des jobs.

1. Dans Designer sélectionnez ► *Projet* ► *Nouveau* ► *Banque de données* ►.
2. Renseignez les options dans l'éditeur de banque de données comme vous le feriez pour toute autre banque de données de base de données Microsoft SQL Server. Assurez-vous de renseigner les options spécifiques à SSL suivantes.

Options spécifiques à SS

Option	Description
<i>Sous-type de base de données</i>	Sélectionnez Sur site .
<i>Utiliser le cryptage SSL</i>	Sélectionnez Oui . La valeur par défaut est Non . Uniquement disponible lorsque vous sélectionnez Sur site pour le <i>Sous-type de base de données</i> .

3. Renseignez les options applicables restantes puis enregistrez la banque de données.
Trouvez des descriptions pour toutes les options des banques de données de base de données Microsoft SQL Server dans le *Guide de référence*.

10.1.9 Banque de données SAP Vora

Avec une banque de données SAP Vora, accédez aux tables Vora à l'aide du pilote ODBC SAP HANA et du protocole de connexion unidirectionnelle SAP HANA.

Utilisez la banque de données SAP Vora comme source dans un flux de données et un modèle de table pour la cible. Pour le chargement des données, le logiciel charge les données du modèle de table vers un fichier intermédiaire CSV dans un fichier configuré local, un système de fichiers HDFS local ou un système de fichiers HDFS Amazon S3 local. Le logiciel charge ensuite la table à partir du fichier local et ajoute les données dans la table existante dans SAP Vora.

Effectuer les tâches suivantes avec la banque de données SAP Vora :

- Importez les tables Vora.
- Ajoutez les données aux tables Vora existantes à l'aide de INSERT.
- Utilisez le chargement par lots.
- Affichez les données des tables Vora dans SAP Data Services.
- Parcourez les métadonnées

Prenez en compte les limitations suivantes lorsque vous utilisez une banque de données SAP Vora.

- La banque de données ne fonctionne pas pour les vues et les partitions de SAP Vora.
- La banque de données utilise le moteur disque relationnel SAP Vora. Cela ne s'applique pas aux autres moteurs, comme le moteur graphique ou le moteur de collection SAP Vora.
- La banque de données n'autorise pas de mappage de colonne partiel.

Exigences de la banque de données SAP Vora :

- Utilisation avec SAP Vora version 2.0 et ultérieures. Pour accéder à SAP Vora avec les versions antérieures à 2.0, utilisez la banque de données ODBC.
- Utilisez SAP HANA version 2.0, Support Package 2 du pilote ODBC, pour le protocole de connexion unidirectionnelle SAP HANA.
- L'utilisateur de la banque de données doit être enregistré comme un « utilisateur Vora » de SAP Vora. Pour plus de détails sur les types d'utilisateur, voir le *Guide du développeur* de SAP Vora.

10.1.9.1 Configuration DSN pour SAP Vora sur Unix et Linux

SAP Vora requiert une connexion de type DSN.

Téléchargez et installez le pilote ODBC SAP HANA version 2.0 SP2 ou ultérieure. Le fichier s'intitule `libodbcHDB.so`.

Pour utiliser l'interface utilisateur graphique du Gestionnaire de connexions, vous devez installer la bibliothèque GTK+2. Le GTK+2 est une boîte à outils multi-plateformes libre qui crée des interfaces utilisateur. Pour d'autres informations sur l'obtention et l'installation de GTK+2, voir <https://www.gtk.org/> . Les instructions suivantes partent du principe que vous disposez d'une interface utilisateur pour le Gestionnaire de connexions.

1. Dans une invite de commande, ouvrez le Gestionnaire de connexion comme suit :

Exemple de code

```
$ cd $LINK_DIR/bin/  
$ ./DSConnectionManager
```

La boîte de dialogue *Gestionnaire de connexions SAP Data Services* s'ouvre.

2. Dans l'onglet *Sources de données*, sélectionnez SAP Vora et cliquez sur *Ajouter*.

La boîte de dialogue *Configuration pour SAP Vora* s'ouvre. Si vous avez téléchargé et installé le pilote ODBC SAP HANA, le logiciel renseigne automatiquement la première option, *Fichier Ini ODBC*. Si vous n'avez pas installé le pilote, cette option est vide. Revenez en arrière, installez le pilote et suivez ces étapes à nouveau.

3. Renseignez les options restantes comme le décrit le tableau suivant.

Options du pilote

Option	Description
<i>Nom de la source de données</i>	Sélectionnez le nom de la source de données à partir de la flèche déroulante.
<i>Nom d'utilisateur</i>	Saisissez le nom d'utilisateur pour accéder à la table SAP Vora.
<i>Mot de passe</i>	Saisissez le mot de passe pour accéder à la table SAP Vora.
<i>Pilote</i>	Saisissez l'emplacement et le nom du pilote ODBC SAP HANA. Nom : <code>libodbcHDB.so</code> .
<i>Nom de l'hôte</i>	Saisissez le nom du serveur.
<i>Port</i>	Saisissez le numéro de port.

4. Facultatif. Cliquez sur *Tester la connexion*. Lorsque la connexion est réussie, cliquez sur *OK*.
5. Cliquez sur *Fermer* pour fermer le Gestionnaire de connexions.

10.1.9.2 Configuration DSN pour SAP Vora sur Windows

Les banques de données SAP Vora requièrent une connexion de type DSN.

Téléchargez et installez le pilote ODBC SAP HANA version 2.0 SP2 ou ultérieure. Le fichier s'intitule `libodbcHDB.exe`.

Ouvrez l'administrateur ODBC. Ouvrez l'administrateur à partir de l'éditeur de banque de données, ou bien ouvrez-le dans le menu Démarrer.

1. Ouvrez l'onglet [Système DSN](#) puis sélectionnez [vora2](#).
2. Cliquez sur [Configurer](#).

La boîte de dialogue [SAP HDB \(FULL_BUILD_VERSION\)](#) s'ouvre.

3. Saisissez un nom dans [Nom de source de données](#). Vous pouvez saisir une description dans [Description](#).
4. Entrez le nom du serveur et le numéro de port distinct avec deux-points dans [Port du serveur](#).

❖ Exemple

[vora:30115](#)

5. Cliquez sur [Se connecter](#) pour tester la connexion. Quand vous avez terminé, cliquez sur [OK](#).

10.1.10 Parcourir les métadonnées depuis une banque de données de base de données

Vous pouvez afficher les métadonnées pour les objets importés ou non importés et pour contrôler si les métadonnées ont changé pour les objets déjà importés.

Le logiciel enregistre les informations de métadonnées pour tous les objets importés dans une banque de données.

[Afficher les métadonnées de banque de données \[page 99\]](#)

[Affichage des objets importés \[page 100\]](#)

[Déterminer si un schéma a changé depuis son importation \[page 101\]](#)

[Afficher les informations d'index secondaires pour les tables \[page 101\]](#)

10.1.10.1 Afficher les métadonnées de banque de données

1. Sélectionnez l'onglet [Banques de données](#) dans la bibliothèque d'objets.
2. Sélectionnez une banque de données, faites un clic droit et sélectionnez [Ouvrir](#). (Il est également possible de double-cliquer sur l'icône de la banque de données.)

Le logiciel ouvre l'explorateur de banque de données dans l'espace de travail. L'explorateur de banque de données liste les tables dans la banque de données. Il est possible de voir des tables dans la base de données externe ou dans le référentiel interne. Il est également possible de faire une recherche dans ces composants.

3. Sélectionnez [Métadonnées externes](#) pour afficher les tables dans la base de données externe.

Si vous sélectionnez une table ou plus, il est possible de faire un clic droit pour plus d'options.

Commande	Description
Ouvrir (uniquement disponible si vous sélectionnez une seule table).	Ouvre l'éditeur pour les métadonnées de table.
Importer	Importe (ou réimporte) des métadonnées de la base de données vers le référentiel.
Rapprocher	Contrôle les différences entre les métadonnées de la base de données et celles dans le référentiel.

4. Sélectionnez [Métadonnées du référentiel](#) pour afficher les tables importées.

Si vous sélectionnez une table ou plus, il est possible de faire un clic droit pour plus d'options.

Commande	Description
Ouvrir (uniquement disponible si vous sélectionnez une seule table).	Ouvre l'éditeur pour les métadonnées de table.
Rapprocher	Contrôle les différences entre les métadonnées du référentiel et celles de la base de données.
Réimporter	Réimporte les métadonnées de la base de données vers le référentiel.
Supprimer	Supprime la table ou les tables du référentiel.
Propriétés (uniquement disponible si vous sélectionnez une seule table).	Affiche les propriétés de la table sélectionnée.
Afficher les données	Ouvre la fenêtre Afficher les données qui permet de voir les données qui sont actuellement dans la table.

Informations associées

[Importer par recherche \[page 106\]](#)

10.1.10.2 Affichage des objets importés

1. Accédez à l'onglet [Banques de données](#) dans la bibliothèque d'objets.

2. Cliquez sur le signe plus (+) à côté du nom de la banque de données pour afficher les types d'objet dans la banque de données. Par exemple, les banques de données de base de données ont des fonctions, des tables, et des tables de modèle.
3. Cliquez sur le signe plus (+) à côté d'un type d'objet pour afficher les objets de ce type importés depuis la banque de données. Par exemple, cliquez sur le signe plus (+) à côté des tables pour afficher les tables importées.

10.1.10.3 Déterminer si un schéma a changé depuis son importation

1. Dans la fenêtre du navigateur qui affiche la liste des tables du référentiel, sélectionnez [Métadonnées externes](#).
2. Sélectionnez la table ou les tables dont vous souhaitez contrôler les modifications.
3. Faites un clic droit et sélectionnez [Rapprocher](#).

La colonne Modifiée affiche OUI pour indiquer que les tables de la base de données sont différentes des métadonnées importées dans le logiciel. Pour utiliser les métadonnées du logiciel les plus récentes, réimportez la table.

La colonne Importée affiche OUI pour indiquer que la table a été importée dans le référentiel.

10.1.10.4 Afficher les informations d'index secondaires pour les tables

Les informations d'index secondaires peuvent aider à comprendre le schéma d'une table importée.

1. Dans l'onglet des banques de données de Designer, cliquez avec le bouton droit de la souris sur la table pour ouvrir le menu contextuel.
2. Dans le menu contextuel, cliquez sur [Propriétés](#) pour ouvrir la fenêtre Propriétés.
3. Dans la fenêtre Propriétés, cliquez sur l'onglet [Index](#). La partie gauche de la fenêtre affiche la liste d'index.
4. Cliquez sur un index pour voir le contenu.

10.1.11 Importation de métadonnées par une banque de données de base de données

Pour les banques de données de base de données, il est possible d'importer les métadonnées pour des tables et des fonctions.

[Informations de table importée \[page 102\]](#)

Le logiciel détermine et enregistre un set spécifique d'informations de métadonnées pour les tables. Après avoir importé les métadonnées, il est possible de modifier les noms de colonne, les descriptions et les types de données.

[Informations sur varchar et colonne à partir de SAP Data Federator \[page 103\]](#)

Toutes les colonnes de décimales importées dans Data Services depuis une source de données SAP Data Federator sont converties à la précision et à l'échelle décimale (28,6).

[Informations de procédures et de fonctions stockées importées \[page 103\]](#)

Le logiciel peut importer des fonctions et des procédures stockées à partir d'un certain nombre de bases de données.

[Importer par navigation \[page 104\]](#)

[Comparer par nom \[page 105\]](#)

[Importer par recherche \[page 106\]](#)

[Réimportation d'objets \[page 107\]](#)

Si vous avez déjà importé un objet tel qu'une banque de données, une fonction ou une table, il est possible de le réimporter, ce qui met à jour les métadonnées de l'objet depuis la base de données (la réimportation remplace toutes les modifications que vous avez apportées à l'objet dans le logiciel).

10.1.11.1 Informations de table importée

Le logiciel détermine et enregistre un set spécifique d'informations de métadonnées pour les tables. Après avoir importé les métadonnées, il est possible de modifier les noms de colonne, les descriptions et les types de données.

Les modifications sont propagées vers tous les objets qui appellent ces objets.

Métadonnées	Description
Nom de la table	Le nom de la table tel qu'il apparaît dans la base de données. i Remarque La longueur maximale dépend du type de référentiel Data Services. Pour la plupart des types de référentiels, la longueur maximale est de 256 caractères ; elle est de 64 caractères pour MySQL et de 128 caractères pour MS SQL server.
Description de la table	La description de la table.
Nom de colonne	Le nom de la colonne.
Description de la colonne	La description de la colonne.
Type de données de la colonne	Le type de données pour la colonne. Si une colonne est définie comme un type de données non pris en charge, le logiciel convertit le type de données en un type pris en charge. Dans certains cas, si le logiciel ne peut pas convertir le type de données, il ignore la colonne entière.

Métadonnées	Description
Type de contenu de la colonne	Le type de contenu identifie le type de données dans le champ.
Colonne de la clé primaire	La/Les colonne(s) qui contient/contiennent la clé primaire pour la table. Après qu'une table ait été ajoutée à un diagramme de flux de données, ces colonnes sont indiquées dans la liste de colonne par une icône de clé à côté du nom de colonne.
Attribut de la table	Les informations enregistrées par le logiciel sur la table telles que la date de création et la date de modification si ces valeurs sont disponibles.
Nom du propriétaire	Nom du propriétaire de la table.

i Remarque

Le nom du propriétaire pour les sources de données MySQL et Netezza correspond au nom de la base de données ou du schéma dans lequel la table apparaît.

10.1.11.2 Informations sur varchar et colonne à partir de SAP Data Federator

Toutes les colonnes de décimales importées dans Data Services depuis une source de données SAP Data Federator sont converties à la précision et à l'échelle décimale (28,6).

Toutes les colonnes varchar importées dans le logiciel à partir d'une source de données SAP Data Federator sont de type varchar (1024).

Il est possible de modifier la précision ou l'échelle décimale et la taille varchar dans le logiciel après l'importation depuis la source de données SAP Data Federator.

10.1.11.3 Informations de procédures et de fonctions stockées importées

Le logiciel peut importer des fonctions et des procédures stockées à partir d'un certain nombre de bases de données.

Vous pouvez importer des procédures stockées depuis les bases de données DB2, MS SQL Server, Oracle, SAP HANA, SQL Anywhere, SAP ASE, Sybase IQ et Teradata.

Il est également possible d'importer des fonctions stockées et des packages depuis Oracle. Il est possible d'utiliser ces fonctions et procédures dans les spécifications d'extraction que vous avez données à Data Services.

Les informations importées pour les fonctions incluent :

- Les paramètres de fonction
- Le type de renvoi
- Le nom, le propriétaire

Les fonctions et les procédures importées apparaissent sur l'onglet *Banques de données* de la bibliothèque d'objets. Les fonctions et les procédures apparaissent dans la branche *Fonction* de chaque arborescence de banque de données.

Il est possible de configurer les fonctions et les procédures importées grâce à l'assistant de fonction et l'éditeur rapide dans une catégorie identifiée par le nom de la banque de données.

Informations associées

[Guide de référence : à propos des procédures \[page 690\]](#)

10.1.11.4 Importer par navigation

i Remarque

Les fonctions ne peuvent pas être importées par navigation.

1. Ouvrez la bibliothèque d'objets.
2. Accédez à l'onglet *Banques de données*.
3. Sélectionnez la banque de données que vous souhaitez utiliser.
4. Faites un clic droit et sélectionnez *Ouvrir*.

Les éléments disponibles pour importer via la banque de données apparaissent dans l'espace de travail.

Dans certains environnements, les tables sont organisées et affichées sous forme d'arborescence. Si c'est le cas, un signe plus (+) se trouve à gauche du nom. Cliquez sur le signe plus pour naviguer dans l'arborescence.

L'espace de travail contient des colonnes qui indiquent si la table a déjà été importée dans le logiciel (Importée) et si le schéma de table a été modifié depuis qu'il a été importé (Modifié). Pour vérifier si le référentiel contient les métadonnées les plus récentes pour un objet, faites un clic droit sur l'objet et sélectionnez *Rapprocher*.

5. Sélectionnez les éléments pour lesquels vous souhaitez importer des métadonnées.

Par exemple, pour importer une table, vous devez sélectionner une table plutôt qu'un dossier qui contient des tables.

6. Faites un clic droit et sélectionnez *Importer*.
7. Dans la bibliothèque d'objets, accédez à l'onglet *Banques de données* pour afficher la liste des objets importés.

10.1.11.5 Comparer par nom

1. Ouvrez la bibliothèque d'objets.
2. Cliquez sur l'onglet *Banques de données*.
3. Sélectionnez la banque de données que vous souhaitez utiliser.
4. Faites un clic droit et sélectionnez *Importer par nom*
5. Dans la fenêtre Importer par nom, sélectionnez le type d'élément que vous souhaitez importer de la liste *Type*.

Si vous importez une procédure stockée, sélectionnez *Fonction*.

6. Pour importer des tables :
 - a. Saisissez un nom de table dans la case *Nom* pour indiquer une table particulière, ou sélectionnez la case à cocher *Tout*, si disponible, pour spécifier toutes les tables.

Si le nom respecte la casse dans la base de données (et non pas tout en majuscules), saisissez le nom tel qu'il apparaît dans la base de données et utilisez les guillemets doubles (") autour du nom pour conserver la casse.

Si vous utilisez Netezza 7.x, vous pouvez entrer un nom de schéma dans la zone de schémas pour limiter les tableaux indiqués à un schéma particulier. Si vous ne renseignez pas le nom de schéma, les tables spécifiées sont limitées au schéma par défaut.
 - b. Saisissez un nom de propriétaire dans la case *Propriétaire* pour limiter les tables indiquées à un propriétaire particulier. Si vous laissez le nom de propriétaire vide, vous indiquez des tables correspondantes qui ne tiennent pas compte du propriétaire (c'est-à-dire, toutes les tables avec le nom de table indiqué).
7. Importer des fonctions et des procédures :
 - Dans la case *Nom*, saisissez le nom de la fonction ou de la procédure stockée.

Si le nom respecte la casse dans la base de données (et non pas tout en majuscules), saisissez le nom tel qu'il apparaît dans la base de données et utilisez les guillemets doubles (") autour du nom pour conserver la casse. Autrement, le logiciel convertit les noms en des caractères majuscules.

Il est également possible de saisir le nom d'un package. Un package Oracle est une collection encapsulée d'objets de programme reliés (par exemple des procédures, fonctions, variables, curseurs et exceptions) stockés ensemble dans une base de données. Le logiciel permet d'importer des procédures ou des fonctions créées dans des packages et de les utiliser comme des procédures ou des fonctions de niveau supérieur.

Si vous saisissez un nom de package, le logiciel importe toutes les fonctions et les procédures stockées définies dans le package Oracle. Il est impossible d'importer une fonction ou une procédure individuelle définie dans un package.
 - Saisissez un nom de propriétaire dans la case *Propriétaire* pour limiter les fonctions indiquées à un propriétaire particulier. Si vous laissez le nom de propriétaire vide, vous indiquez des fonctions correspondantes qui ne tiennent pas compte du propriétaire (c'est-à-dire, toutes les fonctions avec le nom indiqué).
 - Si vous importez une fonction Oracle ou une procédure stockée et que l'une des conditions suivantes s'applique, décochez la case *Joignable depuis l'expression SQL*. Une procédure stockée ne peut pas être poussée vers le bas vers une base de données à l'intérieur d'une autre instruction SQL lorsque la procédure stockée contient une instruction DDL, qu'elle termine la transaction actuelle avec COMMIT ou ROLLBACK, ou qu'elle émet une commande ALTER SESSION ou ALTER SYSTEM.
8. Cliquez sur *OK*.

10.1.11.6 Importer par recherche

i Remarque

Les fonctions ne peuvent pas être importées par recherche

1. Ouvrez la bibliothèque d'objets.
2. Cliquez sur l'onglet *Banques de données*.
3. Sélectionnez le nom de la banque de données que vous souhaitez utiliser.
4. Faites un clic droit et sélectionnez *Rechercher*.

La fenêtre Rechercher apparaît.

5. Saisissez le nom entier de l'élément ou une partie dans la zone de texte *Nom*.

Si le nom respecte la casse dans la base de données (et non pas tout en majuscules), saisissez le nom tel qu'il apparaît dans la base de données et utilisez les guillemets doubles (") autour du nom pour conserver la casse.

6. Sélectionnez *Contient* ou *Egal à* depuis la liste déroulante à droite en fonction de si vous avez fourni une valeur de recherche complète ou partielle.

Egal à sélectionne uniquement la chaîne de recherche complète. C'est à dire que vous devez rechercher nom_table.propriétaire plutôt que nom_table.

7. (Facultatif) Saisissez une description dans la zone de texte *Description*.
8. Sélectionnez le type d'objet dans la case *Type*.
9. Sélectionnez la banque de données dans laquelle vous souhaitez faire la recherche à partir de la case *Rechercher dans*.
10. Sélectionnez *Externe* dans la boîte déroulante à droite de la boîte *Rechercher dans*.

Externe indique que le logiciel recherche l'élément dans la base de données entière définie par la banque de données.

Interne indique que le logiciel recherche uniquement les éléments qui ont été importés.

11. Accédez à l'onglet *Avancé* pour faire une recherche à l'aide des valeurs d'attribut du logiciel.

Les options avancées s'appliquent uniquement aux recherches d'éléments importés.

12. Cliquez sur *Rechercher*.

Le logiciel liste les tables qui correspondent à vos critères de recherche.

13. Pour importer une table à partir de la liste retournée, sélectionnez la table, faites un clic droit et sélectionnez *Importer*.

10.1.11.7 Réimportation d'objets

Si vous avez déjà importé un objet tel qu'une banque de données, une fonction ou une table, il est possible de le réimporter, ce qui met à jour les métadonnées de l'objet depuis la base de données (la réimportation remplace toutes les modifications que vous avez apportées à l'objet dans le logiciel).

Dans les anciennes versions du logiciel, pour réimporter les objets, vous ouvriez la banque de données, affichez les métadonnées du référentiel et sélectionnez les objets à réimporter. Dans cette version du logiciel, il est possible de réimporter les objets à l'aide de la bibliothèque d'objets à plusieurs niveaux :

- Objets individuels : réimporte les métadonnées pour un objet individuel tel qu'une table ou une fonction
- Niveau de nœud de catégorie : réimporte les définitions de tous les objets de ce type dans cette banque de données, par exemple, toutes les tables dans la banque de données
- Niveau de banque de données : réimporte l'ensemble de la banque de données et tous les objets dépendants, dont les tables, les fonctions, les IDOC et les hiérarchies

Réimportation de tables existantes

Si vous utilisez Netezza 6.x ou 7.x, les tables mises à niveau à partir de SP9 ou version inférieure seront réimportées normalement sans nom de domaine de schéma. Pour Netezza 7.x, les tables dont le nouveau nom du schéma est importé dans SP10, seront réimportées de la même façon, en affichant le nom du schéma.

10.1.11.7.1 Réimporter les objets depuis la bibliothèque d'objets

1. Dans la bibliothèque d'objets, cliquez sur l'onglet *Banques de données*.
2. Faites un clic droit sur un objet individuel et cliquez sur *Réimporter*, ou faites un clic droit sur un nœud de catégorie ou sur un nom de banque de données et cliquez sur *Tout réimporter*.

Il est également possible de sélectionner plusieurs objets individuels en appuyant sur Ctrl ou Shift et en cliquant.

3. Cliquez sur *Oui* pour réimporter les métadonnées.
4. Si vous avez sélectionné plusieurs objets à réimporter (par exemple avec *Tout réimporter*), le logiciel demande confirmation pour chaque objet à moins que vous cochiez la case *Ne plus me poser la question pour les objets restants*.

Il est possible d'ignorer des objets à réimporter en cliquant sur Non pour cet objet.

Si vous n'êtes pas sûr de vouloir réimporter (et donc remplacer) l'objet, cliquez sur Afficher utilisation pour afficher l'emplacement où l'objet est actuellement utilisé dans les jobs.

10.2 Banques de données en mémoire

Une banque de données en mémoire est un conteneur pour des tables en mémoire.

Le logiciel permet de créer une banque de données de base de données à l'aide de la *Mémoire* comme *Type de la base de données*. Les banques de données en mémoire sont conçues pour améliorer la performance de traitement des flux de données qui exécutent des jobs en temps réel. Les données (en petite quantité en général dans un job en temps réel) sont stockées en mémoire pour offrir un accès immédiat plutôt que d'accéder aux données sources d'origine.

Une banque de données offre normalement une connexion à une base de données, à une application ou à un adaptateur. En revanche, une banque de données en mémoire contient des schémas de tables en mémoire enregistrés dans le référentiel.

Les tables en mémoire sont des schémas qui permettent de cacher les données intermédiaires. Les tables en mémoire peuvent cacher les données des tables de base de données relationnelles ou des fichiers de données hiérarchiques tels que les messages XML et les IDocs SAP (qui contiennent tous les deux des schémas imbriqués).

Les tables en mémoire peuvent être utilisées pour :

- Déplacer les données entre des flux de données dans des jobs en temps réel. En cachant les données intermédiaires, la performance des jobs en temps réel avec plusieurs flux de données est largement meilleure qu'elle le serait si des fichiers ou des tables ordinaires étaient utilisés pour stocker les données intermédiaires. Pour obtenir la meilleure performance, utilisez uniquement les tables en mémoire lors du traitement de petites quantités de données.
- Stockez les données de table en mémoire pour la durée du job. En enregistrant les données de table en mémoire, la fonction LOOKUP_ECT et les autres transformations et fonctions qui ne nécessitent pas d'opérations de base de données peuvent accéder aux données sans avoir à les lire depuis une base de données distante.

La durée de vie des données de table en mémoire correspond à la durée du job. Les données des tables en mémoire ne peuvent pas être partagées entre plusieurs jobs en temps réel. La prise en charge de l'utilisation des tables en mémoire pour les jobs batch n'est pas disponible.

[Définir une banque de données en mémoire \[page 109\]](#)

[Création d'une table en mémoire \[page 109\]](#)

[Utiliser une table en mémoire comme une source ou une cible \[page 110\]](#)

[Option Mettre à jour le schéma \[page 110\]](#)

Il est possible que vous souhaitiez mettre à jour rapidement un schéma de table cible en mémoire si le schéma précédent a été modifié.

[Options de cible de la table en mémoire \[page 111\]](#)

[Option Créer un ID de ligne \[page 111\]](#)

[Dépannage des tables en mémoire \[page 112\]](#)

Rubrique parent : [Qu'est-ce qu'une banque de données \[page 82\]](#)

Informations associées

Banques de données de bases de données [page 83]

Banques de données de cache persistant [page 112]

Banques de données liées [page 117]

Banques de données de l'adaptateur [page 120]

Banques de données de l'application [page 121]

Banques de données de services Web [page 123]

Modification d'une définition de banque de données [page 125]

Création et gestion de plusieurs configurations de banques de données [page 126]

10.2.1 Définir une banque de données en mémoire

1. A partir du menu *Projet*, sélectionnez ► *Nouveau* ► *Banque de données* ►.
2. Dans la case *Nom*, saisissez le nom de la nouvelle banque de données.

Assurez-vous d'utiliser la convention d'appellation "Memory_DS". Les noms des banques de données sont ajoutés aux noms des tables lorsque les icônes de table apparaissent dans l'espace de travail. Les tables en mémoire sont représentées dans l'espace de travail avec des icônes de table ordinaires. Par conséquent, étiquetez une banque de données en mémoire pour distinguer ses tables en mémoire des tables de base de données ordinaires dans l'espace de travail.

3. Dans la case *Type de la banque de données*, laissez le paramètre par défaut *Base de données*.
4. Dans la case *Type de la base de données*, sélectionnez *Mémoire*.

Aucun attribut supplémentaire n'est requis pour la banque de données en mémoire.

5. Cliquez sur *OK*.

10.2.2 Création d'une table en mémoire

1. A partir de la palette d'outils, cliquez sur l'icône de modèle de table. 
2. Cliquez dans un flux de données pour placer la table de modèle.

La fenêtre *Créer une table* s'ouvre.

3. A partir de la fenêtre *Créer une table*, sélectionnez la banque de données en mémoire.
4. Saisissez un nom de table.
5. Si vous souhaitez une colonne d'ID de ligne générée par le système dans la table, cliquez sur la case à cocher *Créer un ID de ligne*.
6. Cliquez sur *OK*.

La table en mémoire apparaît dans l'espace de travail comme une icône de table de modèle.

7. Connectez la table en mémoire au flux de données comme une cible.

8. Depuis le menu [Projet](#), sélectionnez [Enregistrer](#).

Dans l'espace de travail, l'icône de la table en mémoire est remplacée par une icône de table cible, et la table apparaît dans la bibliothèque d'objets sous la liste des tables de la banque de données en mémoire.

Informations associées

[Option Créer un ID de ligne \[page 111\]](#)

10.2.3 Utiliser une table en mémoire comme une source ou une cible

1. Dans la bibliothèque d'objets, cliquez sur l'onglet [Banques de données](#).
2. Développez la banque de données en mémoire qui contient la table en mémoire que vous souhaitez utiliser.
3. Développez [Tables](#).

Une liste des tables apparaît.

4. Sélectionnez la table en mémoire que vous souhaitez utiliser comme une source ou une cible, et faites-la glisser dans un flux de données ouvert.
5. Connectez la table en mémoire comme une source ou une cible dans le flux de données.

Si vous utilisez une table en mémoire comme une cible, ouvrez l'éditeur de table cible de la table en mémoire pour définir les options de table.

6. Enregistrez le job

Informations associées

[Options de cible de la table en mémoire \[page 111\]](#)

10.2.4 Option Mettre à jour le schéma

Il est possible que vous souhaitiez mettre à jour rapidement un schéma de table cible en mémoire si le schéma précédent a été modifié.

Pour ce faire, utilisez l'option [Mettre à jour le schéma](#). Autrement, vous devriez ajouter une nouvelle table en mémoire pour mettre à jour un schéma.

1. Faites un clic droit sur l'icône de la table cible en mémoire dans l'espace de travail.
2. Sélectionnez [Mettre à jour le schéma](#).

Le schéma de l'objet précédent est utilisé pour mettre à jour le schéma de la table cible en mémoire. La table en mémoire actuelle est mise à jour dans le référentiel. Toutes les occurrences de la table en mémoire actuelle sont mises à jour avec le nouveau schéma.

10.2.5 Options de cible de la table en mémoire

L'option *Supprimer les données de la table avant le chargement* est disponible pour les cibles de la table en mémoire. L'option est activée par défaut (la case est sélectionnée). Pour définir cette option, ouvrez l'éditeur de table cible en mémoire. Si vous décochez cette option, de nouvelles données sont ajoutées aux données de table existantes.

10.2.6 Option Créer un ID de ligne

Si l'option *Créer un ID de ligne* est cochée dans la fenêtre Créer table en mémoire, le logiciel génère une colonne d'entier appelée *DI_Row_ID* dans laquelle la première ligne insérée obtient la valeur 1, la deuxième ligne insérée obtient la valeur 2, etc. Cette nouvelle colonne vous permet d'utiliser une expression LOOKUP_EXT en tant qu'itérateur dans un script.

i Remarque

La même fonctionnalité est disponible pour les autres types de banque de données qui utilisent la fonction SQL.

Utilisez la colonne *DI_Row_ID* pour itérer sur une table à l'aide d'une fonction `lookup_ext` dans un script. Exemple :

```
$NumOfRows = total_rows (memory_DS..table1)
$I = 1;
$count=0
while ($count < $NumOfRows)
begin
    $data =
        lookup_ext([memory_DS..table1, 'NO_CACHE','MAX'],[A],[O],[DI_Row_ID,'=',
$I]);
    $I = $I + 1;
    if ($data != NULL)
    begin
        $count = $count + 1;
    end
end
end
```

Dans le script précédent, `table1` est une table en mémoire. Le nom de la table est précédé par son nom de banque de données (`memory_DS`), d'un point, d'un espace vide (où se trouverait le propriétaire de la table pour une table ordinaire), puis d'un second point. Il n'existe pas de propriétaires pour les banques de données en mémoire, donc les tables sont uniquement identifiées par le nom de la banque de données et le nom de la table comme indiqué.

Sélectionnez les arguments de la fonction LOOKUP_EXT (ligne 7) depuis l'éditeur de fonctions lorsque vous définissez une fonction LOOKUP_EXT.

La fonction `TOTAL_ROWS(DatastoreName.Owner.TableName)` renvoie le nombre de lignes d'une table particulière dans une banque de données. Cette fonction peut être utilisée avec tout type de banque de données. Dans le cas d'une utilisation avec une banque de données en mémoire, utilisez la syntaxe suivante : `TOTAL_ROWS( <DatastoreName..TableName> )`

Le logiciel fournit également une fonction intégrée qu'il est possible d'utiliser pour effacer explicitement les données d'une table en mémoire. Cette fonction offre un meilleur contrôle que celui qu'offre le job actif sur les données et l'utilisation de la mémoire. La fonction `TRUNCATE_TABLE( <DatastoreName..TableName> )` peut uniquement être utilisée avec des tables en mémoire.

10.2.7 Dépannage des tables en mémoire

- Une erreur possible, particulièrement avec l'utilisation des tables en mémoire, est que le logiciel exécute en dehors de la zone de mémoire virtuelle. Le logiciel quitte s'il exécute en dehors de la mémoire en exécutant une opération.
- Une erreur de validation et d'exécution se produit si le schéma d'une table en mémoire ne correspond pas au schéma de l'objet précédent dans le flux de données.
Pour corriger cette erreur, utilisez l'option Mettre à jour le schéma ou créez une nouvelle table en mémoire qui correspond au schéma de l'objet précédent dans le flux de données.
- Deux fichiers journaux contiennent les informations spécifiques aux tables en mémoire : le journal `trace_memory_reader` et le journal `trace_memory_loader`.

10.3 Banques de données de cache persistant

Une banque de données de cache persistant est un conteneur pour ces tables de cache.

Le logiciel permet de créer une banque de données de base de données à l'aide d'un *Cache persistant* comme *Type de la base de données*. Les banques de données de cache persistant offrent les avantages ci-dessous pour les flux de données qui traitent de larges volumes de données.

- Il est possible de stocker une grande quantité de données dans un cache persistant que le logiciel charge rapidement dans la mémoire pour fournir un accès immédiat durant un job. Par exemple, il est possible d'accéder à une table de recherche ou à une table de comparaison localement (plutôt que de les lire depuis une base de données distante).
- Il est possible de créer des tables de cache qui peuvent être partagées par les flux de données (à la différence d'une table en mémoire qui ne peut pas être partagée entre plusieurs jobs en temps réel). Par exemple, si une grande table de recherche utilisée dans une fonction `lookup_ext` est rarement modifiée, il est possible de créer un cache une seule fois, et les jobs ultérieurs peuvent utiliser ce cache plutôt que d'en créer un à chaque fois.

Une banque de données offre normalement une connexion à une base de données, à une application ou à un adaptateur. En revanche, une banque de données de cache persistant contient des schémas de tables de cache enregistrés dans le référentiel.

Les tables de cache persistant permettent de cacher de grandes quantités de données. Les tables de cache persistant peuvent cacher des données provenant de tables de base de données relationnelles et de fichiers.

i Remarque

Il est impossible de cacher les données des fichiers de données hiérarchiques tels que les messages XML et les IDocs SAP (qui contiennent tous les deux des schémas imbriqués). Il est impossible d'effectuer des insertions, des suppressions ou des mises à jour incrémentielles sur une table de cache persistant.

Vous créez une table de cache persistant en chargeant les données dans la table cible de cache persistant qui utilise un flux de données. Vous pouvez ensuite lire depuis la table de cache dans un autre flux de données. Lorsque vous chargez des données dans une table de cache persistant, le logiciel tronque et recrée systématiquement la table.

[Création de banques de données de cache persistant \[page 113\]](#)

Il est possible de créer des banques de données de cache persistant à l'aide de la fenêtre *Editeur de banque de données*.

[Création de tables de cache persistant \[page 114\]](#)

Lors de la création d'une table de cache persistant, il n'est pas nécessaire d'indiquer le schéma de la table ou d'importer les métadonnées de la table.

[Utilisation de tables de cache persistant comme des sources \[page 116\]](#)

Rubrique parent : [Qu'est-ce qu'une banque de données \[page 82\]](#)

Informations associées

[Banques de données de bases de données \[page 83\]](#)

[Banques de données en mémoire \[page 108\]](#)

[Banques de données liées \[page 117\]](#)

[Banques de données de l'adaptateur \[page 120\]](#)

[Banques de données de l'application \[page 121\]](#)

[Banques de données de services Web \[page 123\]](#)

[Modification d'une définition de banque de données \[page 125\]](#)

[Création et gestion de plusieurs configurations de banques de données \[page 126\]](#)

10.3.1 Création de banques de données de cache persistant

Il est possible de créer des banques de données de cache persistant à l'aide de la fenêtre *Editeur de banque de données*.

1. À partir du menu *Projet*, sélectionnez ► *Nouveau* ► *Banque de données* .
2. Dans la case *Nom*, saisissez le nom de la nouvelle banque de données.

Assurez-vous d'utiliser une convention d'appellation telle que "Persist_DS". Les noms des banques de données sont ajoutés aux noms des tables lorsque les icônes de table apparaissent dans l'espace de travail. Les tables de cache persistant sont représentées dans l'espace de travail avec des icônes de table

ordinaires. Par conséquent, étiquetez une banque de données de cache persistant pour distinguer ses tables de cache persistant des tables de base de données ordinaires dans l'espace de travail.

3. Dans la case *Type de la banque de données*, laissez le paramètre par défaut *Base de données*.
4. Dans la case *Type de la base de données*, sélectionnez *Cache persistant*.
5. Dans la case *Répertoire du cache*, vous pouvez soit saisir soit naviguer vers un répertoire dans lequel vous souhaitez stocker le cache persistant.
6. Cliquez sur *OK*.

10.3.2 Création de tables de cache persistant

Lors de la création d'une table de cache persistant, il n'est pas nécessaire d'indiquer le schéma de la table ou d'importer les métadonnées de la table.

A la place, le logiciel crée automatiquement le schéma pour chaque table de cache persistant en fonction du schéma précédent. Lors de la première sauvegarde du job, le logiciel définit le schéma de la table de cache persistant et enregistre la table. Par la suite, la table apparaît avec une icône de table dans l'espace de travail et dans la bibliothèque d'objets sous la banque de données de cache persistant.

Vous créez une table de cache persistant sous l'une des formes suivantes :

- Comme une table de modèle cible dans un flux de données
- Comme une partie de la transformation Data_Transfer lors de l'exécution d'un job

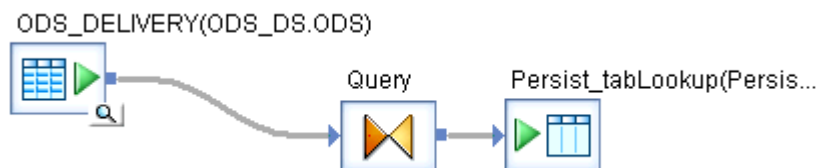
10.3.2.1 Créer une table de cache persistant comme cible dans un flux de données

Créez une table de cache persistant à l'aide d'un modèle de table puis utilisez ce dernier comme cible dans le flux de données.

1. Utilisez l'une des méthodes suivantes pour ouvrir la fenêtre *Créer un modèle* :
 - Depuis la palette d'outils :
 1. Cliquez sur l'icône de modèle de table. 
 2. Cliquez dans un flux de données pour placer la table de modèle dans l'espace de travail.
 3. Dans la fenêtre *Créer un modèle*, sélectionnez la banque de données de cache persistant.
 - Depuis la bibliothèque d'objets :
 1. Développez une banque de données de cache persistant.
 2. Cliquez sur l'icône de table de modèle et faites-la glisser dans l'espace de travail.
2. Sur la fenêtre *Créer un modèle*, saisissez un nom de table.
3. Cliquez sur *OK*.

La table de cache persistant apparaît dans l'espace de travail comme une icône de table de modèle.

4. Connectez la table de cache persistant au flux de données comme cible (généralement une transformation Query).



5. Dans la transformation Query, mappez les colonnes Entrée de schéma que vous souhaitez inclure dans la table de cache persistant.
6. Ouvrez l'éditeur de table cible de la table de cache persistant pour définir les options de table.
7. Sur l'onglet Options de l'éditeur de table cible de cache persistant, il est possible de modifier les options suivantes pour la table de cache persistant.
 - *Comparaison des colonnes* : indique la manière dont les colonnes d'entrée sont mappées aux colonnes de table de cache persistant. Il existe deux options :
 - Compare_by_position : le logiciel ne prend pas en compte les noms de colonnes et mappe les colonnes sources aux colonnes cibles par position.
 - Comparer_par_nom : le logiciel mappe les colonnes sources aux colonnes cibles par le nom. Cette option est celle par défaut
 - *Inclure les clés en double* : sélectionnez cette case à cocher pour cacher les clés en double. Cette option est sélectionnée par défaut.
8. Sur l'onglet Clés, indiquez la/les colonne(s) clé(s) à utiliser comme clé dans la table de cache persistant.
9. Depuis le menu *Projet*, sélectionnez *Enregistrer*. Dans l'espace de travail, l'icône de la table de modèle est remplacée par une icône de table cible, et la table apparaît dans la bibliothèque d'objets sous la liste des tables de la banque de données de cache persistant.

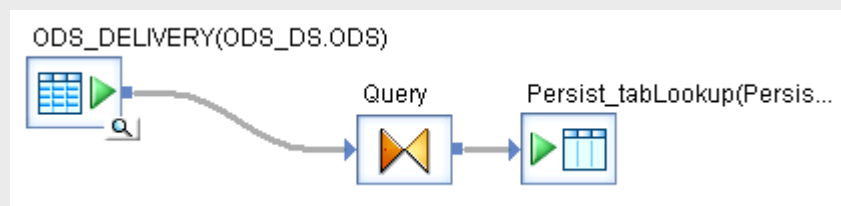
10.3.2.2 Creating a persistent cache table schema in a data flow

Create the persistent cache table schema in the work flow starting with the upstream object.

Follow the steps in [Créer une table de cache persistant comme cible dans un flux de données \[page 114\]](#). Use the following diagram as an example:

❖ Exemple

The object that precedes the persistent cache table is a Query transform in the following diagram:



After you connect the persistent cache target table to the data flow, perform the following steps:

1. Open the object editor for the object that directly precedes the persistent cache table.

In the example, it is the Query transform editor.

2. Select the columns to include in the persistent cache table.
3. Drag the columns from the *Schema In* pane to the *Schema Out* pane.
4. Click the name of the persistent cache table to open the editor.
5. Set options as described in the following table.

Option	Description
<i>Column comparison</i>	Specifies how the software maps the input columns to persistent cache table columns. ◦ <i>Compare by position</i>: The software disregards the column names and maps source columns to target columns by position. ◦ <i>Compare by name</i>: The software maps source columns to target columns by name. <i>Compare by name</i> is the default setting.
<i>Include duplicate keys</i>	Specifies whether the software caches duplicate keys. Selected: The software caches duplicate keys. Selected is the default setting. Not selected: The software does not cache duplicate keys.

6. Open the *Keys* tab to specify the key column or columns to use as the key or keys in the persistent cache table.
7. Select ► *Project* ► *Save*. ►

The persistent cache template table icon changes to a target table icon. The software lists the table in the object library in the Table node under the applicable persistent cache datastore.

10.3.3 Utilisation de tables de cache persistant comme des sources

Après avoir créé une table de cache persistant comme une cible dans un flux de données, il est possible d'utiliser une table de cache persistant comme une source dans n'importe quel flux de données. Il est également possible de l'utiliser comme une table de recherche ou une table de comparaison.

10.4 Banques de données liées

Plusieurs fournisseurs de base de données prennent en charge des chemins de communication à sens unique entre un serveur de base de données à un autre. Oracle appelle ces chemins liens de base de données.

Dans DB2, le chemin de communication à sens unique d'un serveur de base de données à un autre est fourni par un serveur d'informations qui permet à un set de serveurs d'obtenir des données des sources de données distantes. Dans Microsoft SQL Server, les serveurs liés fournissent le chemin de communication à sens unique d'un serveur de base de données à un autre. Ces solutions permettent aux utilisateurs locaux d'accéder à des données sur une base de données distante, qui peut se trouver sur l'ordinateur local ou distant, du même type de base de données ou d'un type différent.

Par exemple, un serveur de base de données local Oracle, appelé Ordres, peut stocker un lien de base de données pour accéder aux informations dans une base de données Oracle distante, appelée Clients. Les utilisateurs connectés à Clients ne peuvent cependant pas utiliser le même lien pour accéder aux données dans Ordres. Les utilisateurs connectés à la base de données Client doivent définir un lien différent, stocké dans le dictionnaire de données de la base de données Client, pour accéder aux données sur Ordres.

Le logiciel fait référence aux chemins de communication entre les bases de données sous le nom de liens de base de données. Les banques de données dans une relation de lien de base de données sont appelées banques de données liées. Le logiciel utilise les banques de données liées pour améliorer sa performance en poussant vers le bas des opérations vers une base de données cible à l'aide d'une banque de données cibles.

[Relations entre les liens de base de données et les banques de données \[page 118\]](#)

[Liaison d'une banque de données cible à une banque de données source à l'aide d'un lien de base de données \[page 119\]](#)

Un lien de base de données enregistre les informations sur la manière de se connecter à une source de données distante.

Rubrique parent : [Qu'est-ce qu'une banque de données \[page 82\]](#)

Informations associées

[Banques de données de bases de données \[page 83\]](#)

[Banques de données en mémoire \[page 108\]](#)

[Banques de données de cache persistant \[page 112\]](#)

[Banques de données de l'adaptateur \[page 120\]](#)

[Banques de données de l'application \[page 121\]](#)

[Banques de données de services Web \[page 123\]](#)

[Modification d'une définition de banque de données \[page 125\]](#)

[Création et gestion de plusieurs configurations de banques de données \[page 126\]](#)

10.4.1 Relations entre les liens de base de données et les banques de données

Un lien de base de données enregistre les informations sur la manière de se connecter à une source de données distante, telles que son nom d'hôte, le nom de la base de données, le nom de l'utilisateur, le mot de passe et le type de base de données. Les mêmes informations sont stockées dans une banque de données de base de données SAP Data Services. Il est possible d'associer la banque de données à une autre banque de données et ensuite d'importer un lien de base de données externe comme une option de banque de données. Les banques de données doivent se connecter aux bases de données définies dans le lien de base de données.

Les conditions supplémentaires sont les suivantes :

- Un serveur local pour des liens de base de données doit être un serveur cible dans le logiciel
- Un serveur distant pour des liens de base de données doit être un serveur source dans le logiciel
- Un lien de base de données externe (qui existe en premier dans une base de données) établit la relation entre une banque de données cible et une banque de données source
- Une banque de données locale peut être reliée à zéro ou plusieurs banques de données à l'aide d'un lien de base de données pour chaque base de données distante
- Deux banques de données peuvent être reliées ensemble à l'aide d'un seul lien

Le diagramme ci-dessous affiche les relations possibles entre les liens de base de données et les banques de données liées :



Quatre liens de base de données, les DBLink 1 à 4, sont sur la base de données DB1 et le logiciel les lit par la banque de données DS1.

- Le DBLink 1 relie la banque de données DS1 à la banque de données DS2. Cette relation est appelée banque de données liée DBLink1 (la banque de données liée porte le même nom que le lien de base de données externe).
- DBLink2 n'est mappé à aucune banque de données dans le logiciel car il relie DS1 à DS2, qui sont également reliées à DBLink1. Bien qu'il ne s'agisse pas d'un cas ordinaire, il est possible de créer plusieurs liens de base de données externe qui connectent à la même source distante. Cependant, le logiciel n'autorise qu'un seul lien de base de données entre une banque de données cible et une paire de banques de données source. Par exemple, si vous sélectionnez DBLink1 pour lier la banque de données cible DS1 à la banque de données source DS2, vous ne pouvez pas importer DBLink2 pour réaliser la même chose.
- DBLink3 n'est mappé à aucune banque de données dans le logiciel car il n'existe aucune banque de données définie pour la source de données distante à laquelle le lien de base de données externe fait référence.
- DBLink4 relie DS1 avec DS3.

10.4.2 Liaison d'une banque de données cible à une banque de données source à l'aide d'un lien de base de données

Un lien de base de données enregistre les informations sur la manière de se connecter à une source de données distante.

Exécutez les étapes suivantes dans Designer.

1. Dans l'onglet *Banques de données* de la bibliothèque d'objets, cliquez avec le bouton droit sur une banque de données cible et sélectionnez *Modifier*.

Si le type de base de données prend en charge les liens de base de données, la liste des options de configuration comprend l'option *Banques de données liées*, dans la section *Avancé*.

i Remarque

L'éditeur de banque de données permet de modifier les liens de base de données vers les banques de données cible uniquement pour la configuration par défaut. Ainsi, si votre banque de données source contient plusieurs configurations (par exemple : Config1, Config2 et Config3), modifiez la configuration par défaut avant d'essayer d'importer ou de modifier les liens qui la concernent (par exemple, déclarez Config2 comme configuration par défaut pour pouvoir la modifier).

2. Cliquez sur l'étiquette *Banques de données liées*.

La fenêtre *Ajouter une banque de données liée* s'ouvre.

3. Dans la fenêtre *Ajouter une banque de données liée*, sélectionnez une banque de données à laquelle sera liée votre banque de données cible en fonction des paramètres du lien de base de données à importer.

Par exemple, si la banque de données cible est DS_Emp (informations sur les employés) et que le lien de base de données à importer doit associer les informations sur les employés aux informations sur les ventes, sélectionnez DS_Sales (informations sur les ventes).

Les banques de données de la zone de liste ont des types de base de données pris en charge pour les banques de données liées.

i Remarque

L'éditeur de banque de données n'autorise qu'un seul lien de base de données entre une paire banque de données cible/banque de données source. Par conséquent, si la banque de données cible B comporte déjà un lien vers la banque de données source A, vous ne pouvez pas importer un autre lien de base de données associant la banque de données A et la banque de données B.

4. Cliquez sur *OK*.

La fenêtre *Editeur de banque de données* affiche la banque de données que vous avez sélectionnée.

5. Sélectionnez le bouton Liste à droite de *Non lié* ou double-cliquez sur la cellule.

La fenêtre *Lien de base de données* s'ouvre.

6. Pour créer un lien vers une banque de données ou modifier le lien existant, sélectionnez *Utiliser le lien de base de données*.

i Remarque

Pour supprimer un lien existant, sélectionnez *Ne pas lier*.

7. Sélectionnez un lien de base de données dans la liste lue dans la connexion de la configuration par défaut de la banque de données cible que vous modifiez.
Cette zone de liste contient les liens préalablement définis dans le SGBD.
8. Sélectionnez la configuration de banque de données source à utiliser avec ce lien de base de données.
9. (Facultatif) Sélectionnez [Détails](#) pour afficher d'autres informations sur les liens ou les tester.
La coche indique le lien à utiliser. Si vous utilisez la fenêtre [Détails](#), cliquez sur OK quand vous avez terminé.
10. Dans la boîte de dialogue [Lien de base de données](#), cliquez sur [OK](#).

10.5 Banques de données de l'adaptateur

Les banques de données d'adaptateur permettent d'accéder aux données d'application, accessibles via l'adaptateur.

En fonction de l'implémentation de l'adaptateur, les adaptateurs permettent de :

- Parcourir les métadonnées d'application
- Importer les métadonnées d'application dans un référentiel
- Déplacer les données batch et en temps réel entre le logiciel et les applications

SAP offre un kit de développement de logiciel (SDK) pour adaptateur pour développer vos propres adaptateurs personnalisés. Vous pouvez également acheter les adaptateurs pré-packagés du logiciel pour accéder aux métadonnées et aux données d'applications dans n'importe quelle application. Pour plus d'informations sur ces produits, contactez votre représentant commercial SAP.

Les adaptateurs sont représentés dans Designer par des banques de données d'adaptateurs. Les jobs fournissent un mouvement de données batch et en temps réel entre le logiciel et les applications par des objets subordonnés d'une banque de données d'adaptateur :

Objets subordonnés	Utilisés comme	Utilisés pour
Tables	Source ou cible	Mouvement de données batch
Documents	Source ou cible	
Fonctions	Appel de fonction dans une requête	
Fonctions de message	Appel de fonction dans une requête	Mouvement de données en temps réel
Messages sortants	Cible uniquement	

Les adaptateurs peuvent donner accès aux données et aux métadonnées d'une application, ou simplement aux métadonnées. Par exemple, si la source de données est compatible avec SQL, l'adaptateur peut être conçu pour accéder aux métadonnées. Le logiciel extrait les données de l'application ou charge les données directement dans l'application.

Pour en savoir plus sur l'installation, la configuration et l'utilisation des adaptateurs, voir *Supplément pour adaptateurs*.

Rubrique parent : [Qu'est-ce qu'une banque de données \[page 82\]](#)

Informations associées

[Banques de données de bases de données \[page 83\]](#)

[Banques de données en mémoire \[page 108\]](#)

[Banques de données de cache persistant \[page 112\]](#)

[Banques de données liées \[page 117\]](#)

[Banques de données de l'application \[page 121\]](#)

[Banques de données de services Web \[page 123\]](#)

[Modification d'une définition de banque de données \[page 125\]](#)

[Création et gestion de plusieurs configurations de banques de données \[page 126\]](#)

[Objets sources et cibles \[page 206\]](#)

[Objets source et cible en temps réel \[page 352\]](#)

10.6 Banques de données de l'application

Configurez une banque de données de l'application avec les informations de connexion à l'application.

Le tableau suivant décrit les applications disponibles lorsque vous créez une banque de données de l'application.

Descriptions de banques de données de l'application

Application	Banque de données
Google BigQuery	Se connecte à votre projet BigQuery afin de charger les données aux tables Google existantes et extraire (lire) les données des tables existantes à traiter dans Data Services. • Téléchargez les données générées par Data Services dans les tables existantes dans votre compte Google BigQuery pour lancer des requêtes. • Extrayez des données de vos tables Google BigQuery afin de les utiliser comme source dans les flux de données Data Services. Servez-vous d'un modèle Google BigQuery comme cible pour créer automatiquement la table dans votre compte Google et chargez les données traitées dans la table Google. Pour en savoir plus, voir le <i>Supplément pour Google BigQuery</i>.

Application	Banque de données
JDE OneWorld	Se connecte à la base de données dans votre application JDE OneWorld. Les options de configuration de la banque de données sont basées sur le type de base de données que vous sélectionnez. Le type de banque de données JDE OneWorld utilise les types de banques de données suivants : • DB2 • Microsoft SQL Server • ODBC • Oracle Pour en savoir plus sur la configuration des options de banque de données pour des applications JD Edwards, voir le document <i>Supplément pour J.D. Edwards</i>.
JD Edwards World	Se connecte à la base de données ODBC dans votre application JDE World. Pour en savoir plus sur la configuration des options de banque de données pour des applications JD Edwards, voir le document <i>Supplément pour J.D. Edwards</i>.
Applications Oracle	Se connecte aux données dans vos applications Oracle. Pour en savoir plus sur la configuration des options de banque de données pour des applications Oracle, voir le document <i>Supplément pour Oracle Applications</i>.
PeopleSoft	Se connecte aux types de base de données suivantes dans votre application PeopleSoft. Les options varient en fonction du type de base de données choisi. Les bases de données Peoplesoft applicables sont : • Microsoft SQL Server • Oracle Pour en savoir plus sur les applications PeopleSoft, voir le document <i>Supplément pour PeopleSoft</i>.
Serveur de réplication CDM	Se connecte aux jobs de capture des données modifiées (CDM) à l'aide du modèle de mouvements de données SAP PowerAMC.
Applications SAP	Se connecte à l'application SAP applicable. Trouvez des descriptions pour les banques de données de l'application SAP dans le <i>Supplément pour SAP</i>.
Source SAP BW	
Cible SAP BW	
SAP DQM Microservices	Se connecte à SAP Data Quality Management, microservices pour les données d'emplacement pour utiliser la transformation DQM Microservices dans Data Services. Recherchez les descriptions des options de banque de données pour la banque de données DQM microservices dans le <i>Supplément pour SAP</i>.

Application	Banque de données
Siebel	Se connecte aux types de base de données suivantes dans votre application Siebel. • DB2 • Microsoft SQL Server • Oracle Pour en savoir plus sur les applications Siebel, voir le document <i>Supplément pour Siebel</i>.
Service Web REST	Retrouvez les descriptions complètes des options dans le <i>Guide d'intégration</i> .
Service Web SOAP	Retrouvez les descriptions complètes des options dans le <i>Guide d'intégration</i> .

Une fois la banque de données créée, vous pouvez y importer des métadonnées relatives aux objets, tels que des tables et fonctions dans la bibliothèque d'objets.

Pour des informations complètes sur les banques de données d'adaptateur, consultez le *Supplément pour adaptateurs*.

Rubrique parent : [Qu'est-ce qu'une banque de données \[page 82\]](#)

Informations associées

[Banques de données de bases de données \[page 83\]](#)

[Banques de données en mémoire \[page 108\]](#)

[Banques de données de cache persistant \[page 112\]](#)

[Banques de données liées \[page 117\]](#)

[Banques de données de l'adaptateur \[page 120\]](#)

[Banques de données de services Web \[page 123\]](#)

[Modification d'une définition de banque de données \[page 125\]](#)

[Création et gestion de plusieurs configurations de banques de données \[page 126\]](#)

10.7 Banques de données de services Web

Les banques de données de services Web représentent une connexion depuis Data Services vers une source de données externe basée sur les services Web.

Pour plus d'informations, voir « Accès à un service Web à l'aide de Designer » dans le *Guide d'intégration*.

Rubrique parent : [Qu'est-ce qu'une banque de données \[page 82\]](#)

Informations associées

[Banques de données de bases de données \[page 83\]](#)

[Banques de données en mémoire \[page 108\]](#)

[Banques de données de cache persistant \[page 112\]](#)

[Banques de données liées \[page 117\]](#)

[Banques de données de l'adaptateur \[page 120\]](#)

[Banques de données de l'application \[page 121\]](#)

[Modification d'une définition de banque de données \[page 125\]](#)

[Création et gestion de plusieurs configurations de banques de données \[page 126\]](#)

10.7.1 Définition d'une banque de données de services Web

Vous devez définir au moins une banque de données pour chaque service Web avec lequel vous échangez des données.

Pour définir une banque de données, vous devez avoir les privilèges d'accès aux services Web appropriés décrits par la banque de données.

1. Dans l'onglet *Banques de données* de la bibliothèque d'objets, faites un clic droit et sélectionnez *Nouveau*.
2. Saisissez le nom de la nouvelle banque de données dans le champ *Nom de la banque de données*.
Le nom peut contenir tous les caractères alphabétiques ou numériques, ou des tirets de soulignement (_). Il ne peut pas contenir d'espace.
3. Sélectionnez le *Type de banque de données*.
Choisissez une option de service Web. Lorsque vous sélectionnez un type de banque de données, Data Services affiche les autres options pertinentes pour ce type.
4. Indiquez l'*URL du service Web*.
L'URL doit accepter les connexions et renvoyer le WSDL. Pour les services Web REST, vous pouvez saisir une URL ou le chemin d'accès au fichier local WADL. Voir "Technologies de services Web" dans le *Guide d'intégration* pour en savoir plus sur les fichiers WSDL et WADL.
5. Cliquez sur *OK*.
La configuration de banque de données est enregistrée dans le référentiel de métadonnées et la nouvelle banque de données apparaît dans la bibliothèque d'objets.

10.7.2 Exploration de métadonnées WSDL et WADL par une banque de données de services Web

Data Services enregistre les informations de métadonnées pour tous les objets importés dans une banque de données. Vous pouvez utiliser Data Services pour afficher les métadonnées pour les objets importés ou non importés et pour contrôler si les métadonnées ont changé pour les objets déjà importés.

Voir "Technologies de services Web" dans le *Guide d'intégration* pour en savoir plus sur les fichiers WSDL et WADL.

10.8 Modification d'une définition de banque de données

Comme tous les objets, les banques de données sont définies par des options et des propriétés.

Les options contrôlent l'opération des objets. Par exemple, le nom de la base de données à laquelle se connecter est une option de banque de données.

Les propriétés documentent l'objet. Par exemple, le nom de la banque de données et la date à laquelle elle a été créée sont des propriétés de banque de données. Les propriétés sont simplement des descriptions de l'objet et n'affectent pas son opération.

[Modification des options de banque de données \[page 125\]](#)

[Modifier les propriétés de la banque de données \[page 126\]](#)

Rubrique parent : [Qu'est-ce qu'une banque de données \[page 82\]](#)

Informations associées

[Banques de données de bases de données \[page 83\]](#)

[Banques de données en mémoire \[page 108\]](#)

[Banques de données de cache persistant \[page 112\]](#)

[Banques de données liées \[page 117\]](#)

[Banques de données de l'adaptateur \[page 120\]](#)

[Banques de données de l'application \[page 121\]](#)

[Banques de données de services Web \[page 123\]](#)

[Création et gestion de plusieurs configurations de banques de données \[page 126\]](#)

10.8.1 Modification des options de banque de données

1. Accédez à l'onglet *Banques de données* dans la bibliothèque d'objets.
2. Faites un clic droit sur le nom de la banque de données et sélectionnez *Modifier*.

L'éditeur de banque de données apparaît (la barre de titre pour cette boîte de dialogue affiche Modifier la banque de données). Il est possible de réaliser les tâches suivantes :

- Modifier les informations de connexion pour la configuration de banque de données actuelle.
- Cliquez sur *Avancé* et modifiez les propriétés pour la configuration actuelle.
- Cliquez sur *Modifier* pour ajouter, modifier ou supprimer des configurations supplémentaires. La boîte de dialogue *Configurations pour la banque de données* s'ouvre lorsque vous sélectionnez *Modifier* dans l'éditeur de banque de données. Après avoir ajouté une nouvelle configuration à une banque de données existante, vous pouvez utiliser les champs dans la grille pour modifier les valeurs et les propriétés de connexion pour la nouvelle configuration.

3. Cliquez sur *OK*.

Les options prennent immédiatement effet.

10.8.2 Modifier les propriétés de la banque de données

1. Accédez à l'onglet Banque de données dans la bibliothèque d'objets.
2. Faites un clic droit sur le nom de la banque de données et sélectionnez *Propriétés*.

La fenêtre Propriétés s'ouvre.

3. Modifiez les propriétés de la banque de données.
4. Cliquez sur *OK*.

10.9 Création et gestion de plusieurs configurations de banques de données

La création de plusieurs configurations pour une seule banque de données permet de consolider des connexions aux banques de données distinctes pour des sources ou des cibles similaires dans une banque de données source ou cible avec plusieurs configurations.

Il est alors possible de sélectionner un ensemble de configurations qui inclut les sources et les cibles que vous souhaitez en sélectionnant une configuration système lorsque vous exécutez ou planifiez le job. La possibilité de créer plusieurs configurations de banque de données offre une plus grande facilité d'utilisation pour les scénarios de portabilité de job, tels que :

- OEM (différentes bases de données pour la conception et la distribution)
- Migration (différentes connexions pour DEV, TEST et PROD)
- Multi-instance (bases de données avec des versions différentes ou des paramètres régionaux)
- Multi-utilisateurs (bases de données pour les référentiels centraux et locaux)

Définitions [page 127]

Termes que vous devez bien connaître lors de la création et de la gestion de plusieurs configurations de banque de données

Pourquoi utiliser plusieurs configurations de banque de données ? [page 128]

En créant plusieurs configurations de banque de données, il est possible de réduire le temps de développement de bout en bout dans un environnement d'entrepôt des données d'entreprise multi-source, 24h/24, 7j/7, car il est possible de transporter facilement les jobs dans différents types, versions et instances de base de données.

Création d'une nouvelle configuration de banque de données [page 129]

Utilisation des alias [page 130]

Un alias est un nom de propriétaire logique que vous pouvez créer pour les objets que vous utilisez dans différents environnements de base de données.

Création d'un alias [page 130]

Fonctions pour identifier la configuration [page 131]

Le logiciel fournit des fonctions qui sont utiles lors de l'utilisation de plusieurs configurations de banque de données cible et source.

[Configuration des banques de données dans les flux de données \[page 132\]](#)

Lorsque vous utilisez une banque de données de configuration multiple dans un flux de données, soyez prudent lorsque vous modifiez les configurations.

[Solutions de portabilité \[page 136\]](#)

Définissez plusieurs configurations sources ou cibles pour une seule banque de données si vous souhaitez modifier rapidement les connexions à une base de données source ou cible différente.

[Rename Owner option \[page 141\]](#)

A shared alias name in a central repository environment makes tracking the object easier when it has been checked out by multiple users.

[Définition d'une configuration système \[page 146\]](#)

Lors de la conception de jobs, déterminez et créez des configurations de banque de données et des configurations système en fonction de votre environnement et de vos règles d'entreprise.

Rubrique parent : [Qu'est-ce qu'une banque de données \[page 82\]](#)

Informations associées

[Banques de données de bases de données \[page 83\]](#)

[Banques de données en mémoire \[page 108\]](#)

[Banques de données de cache persistant \[page 112\]](#)

[Banques de données liées \[page 117\]](#)

[Banques de données de l'adaptateur \[page 120\]](#)

[Banques de données de l'application \[page 121\]](#)

[Banques de données de services Web \[page 123\]](#)

[Modification d'une définition de banque de données \[page 125\]](#)

[Solutions de portabilité \[page 136\]](#)

10.9.1 Définitions

Termes que vous devez bien connaître lors de la création et de la gestion de plusieurs configurations de banque de données

Terme	Définition
Configuration de banque de données	Permet de fournir plusieurs sources ou cibles de métadonnées pour les banques de données. Chaque configuration est une propriété d'une banque de données qui fait référence à un ensemble d'options configurables (telles que le nom de connexion à la base de données, le type de base de données, le nom d'utilisateur, le mot de passe, et le paramètre régional) et à leurs valeurs.

Terme	Définition
Configuration de banque de données par défaut	La configuration de banque de données utilisée par le logiciel pour parcourir et importer les objets de la base de données (tables et fonctions), et pour exécuter les jobs si aucune configuration système n'est indiquée. Si une banque de données a plus d'une configuration, sélectionnez une configuration par défaut au besoin. Si la banque de données a une seule configuration, le logiciel l'utilise comme configuration par défaut.
Configuration de banque de données actuelle	La configuration de banque de données utilisée par le logiciel pour exécuter un job. Si vous définissez une configuration système, le logiciel exécute le job à l'aide de la configuration de système. Indiquez une configuration actuelle pour chaque configuration système. Si vous ne créez pas une configuration de système, ou que la configuration système ne spécifie pas de configuration pour une banque de données, le logiciel utilise la configuration de banque de données par défaut comme la configuration actuelle lors de l'exécution du job.
Objets de base de données	Les tables et les fonctions qui sont importées d'une banque de données. Les objets de base de données possèdent normalement des propriétaires. Certains objets de base de données n'ont pas de propriétaire. Par exemple, les objets de base de données dans une banque de données ODBC connectée à une base de données Access n'ont pas de propriétaire.
Nom du propriétaire	Le nom du propriétaire pour un objet de base de données (par exemple, une table) dans une base de données sous-jacente. Également connu comme nom du propriétaire de la base de données ou nom du propriétaire physique.
Alias	Un nom de propriétaire logique. Créez un alias pour les objets qui sont dans des environnements de base de données différents s'il existe des noms de propriétaire différents dans ces environnements. Il est possible de créer un alias depuis un éditeur de banque de données pour n'importe quelle configuration de banque de données.
Objets dépendants	Les objets dépendants sont les jobs, les workflows, les flux de données et les fonctions personnalisées dans lesquels un objet de base de données est utilisé. Les informations d'objets dépendants sont générées par l'utilitaire cas d'emploi.

10.9.2 Pourquoi utiliser plusieurs configurations de banque de données ?

En créant plusieurs configurations de banque de données, il est possible de réduire le temps de développement de bout en bout dans un environnement d'entrepôt des données d'entreprise multi-source, 24h/24, 7j/7, car il est possible de transporter facilement les jobs dans différents types, versions et instances de base de données.

Par exemple, le déplacement peut être aussi simple que :

1. Créer une nouvelle configuration dans une banque de données source ou cible existante.
2. Ajouter un alias de banque de données puis de mapper les configurations avec des noms de propriétaire d'objets différents à cet alias.
3. Définir une configuration système puis d'ajouter des configurations de banque de données requises pour un environnement particulier. Sélectionner une configuration système lors de l'exécution d'un job.

10.9.3 Création d'une nouvelle configuration de banque de données

1. A partir de l'onglet Banques de données de la bibliothèque d'objets, faites un clic droit sur une banque de données existante et sélectionnez [Modifier](#).
2. Cliquez sur [Avancé](#) pour afficher les informations de la configuration existante.

Chaque banque de données doit avoir au moins une configuration. Si une seule configuration existe, c'est la configuration par défaut.

3. Cliquez sur [Modifier](#) pour ouvrir la fenêtre Configurations pour la banque de données.

4.  Cliquez sur l'icône [Créer une configuration](#) dans la barre d'outils.

La fenêtre Créer une configuration s'ouvre.

5. Dans la fenêtre Créer une configuration :
 - a. Saisissez un [Nom](#) de configuration unique et logique.
 - b. Sélectionnez un [Type de la base de données](#) dans le menu déroulant.
 - c. Sélectionnez une [Version de la base de données](#) dans le menu déroulant.
 - d. Dans la section Valeurs des tables cibles et des transformations SQL, le logiciel présélectionne la valeur [Utiliser les valeurs de](#) en fonction du type et de la version de la base de données existante. Designer utilise automatiquement la transformation SQL et les valeurs cible existantes pour le même type et version de base de données.

De plus, si la base de données que vous souhaitez associer à une nouvelle configuration est une version postérieure à celle associée aux autres configurations existantes, Designer renseigne automatiquement l'option [Utiliser les valeurs de](#) avec l'ancienne version.

Cependant, si un type et une version de base de données ne sont pas déjà indiqués dans une configuration existante, ou si la version de la base de données est plus ancienne que la configuration existante, vous pouvez choisir d'utiliser les valeurs d'une autre configuration existante ou de la configuration par défaut pour le type et la version de base de données.

- e. Cochez ou décochez l'option [Restaurer des valeurs si elles existent déjà](#).

Lorsque vous supprimez les configurations de banque de données, le logiciel enregistre toutes les valeurs de cible et les transformations SQL associées. Si vous créez une nouvelle configuration de banque de données avec le même type et la même version de base de données que celle précédemment supprimée, l'option Restaurer des valeurs si elles existent déjà permet d'accéder et de profiter des paramètres de valeur enregistrés.

- Si vous conservez cette option (sélectionnée par défaut), le logiciel utilise les valeurs personnalisées cibles et de transformation SQL des anciennes configurations de banque de données supprimées.
 - Si vous désélectionnez [Restaurer des valeurs si elles existent déjà](#), le logiciel n'essaye pas de restaurer les valeurs de cible et de transformation SQL, vous permettant de fournir de nouvelles valeurs.
- f. Cliquez sur [OK](#) pour enregistrer la nouvelle configuration.

Si la banque de données contient des flux de données préexistants avec des transformations SQL ou des objets cibles, le logiciel doit ajouter de nouvelles valeur de type et de version de base de données pour ces transformations et objets cibles. Dans ce cas, lorsque vous ajoutez une nouvelle configuration de banque de données, le logiciel affiche la fenêtre [Nouvelles valeurs ajoutées - Objets modifiés](#) qui

fournit des informations détaillées sur les flux de données affectés et les objets modifiés. Ces mêmes résultats s'affichent également dans la fenêtre Sortie de Designer.

Pour chaque banque de données, le logiciel requiert qu'une configuration soit désignée comme la configuration par défaut. Le logiciel utilise la configuration par défaut pour importer les métadonnées et conserve également la configuration par défaut lors de l'exportation et des opérations multi-utilisateurs. La première configuration de banque de données est automatiquement désignée comme la configuration par défaut ; cependant, après avoir ajouté une ou plusieurs configurations de banque de données supplémentaires, il est possible d'utiliser l'éditeur de banque de données pour marquer une configuration différente comme configuration par défaut.

Lorsque vous exportez un référentiel, le logiciel conserve toutes les configurations de toutes les banques de données, y compris les paramètres de texte de transformation SQL et de l'éditeur de table cible. Si la banque de données que vous exportez existe déjà dans le référentiel cible, le logiciel remplace les configurations dans la cible par les configurations sources. Le logiciel exporte les configurations système séparément des autres objets de jobs reliés.

10.9.4 Utilisation des alias

Un alias est un nom de propriétaire logique que vous pouvez créer pour les objets que vous utilisez dans différents environnements de base de données.

Il est possible de créer un alias depuis un éditeur de banque de données pour n'importe quelle configuration de banque de données. Après la création d'un alias (par exemple, ALIAS1, ALIAS2), naviguez horizontalement vers chaque colonne de configuration. Saisissez le nom du propriétaire auquel ce nom d'alias est mappé ou sélectionnez un nom de propriétaire dans la liste déroulante.

i Remarque

Data Services n'étiquette pas spécifiquement le nom du propriétaire. Le nom du propriétaire s'affiche dans la grille, mais c'est le nom de l'alias qui se trouve dans la ligne. Il n'y a aucune ligne intitulée Nom de propriétaire.

Quand vous supprimez un nom d'alias, l'opération de suppression supprime également chaque nom de propriétaire appliqué pour chaque configuration. Le logiciel supprime la ligne sélectionnée, qui comprend l'alias et tous les noms de propriétaires affectés.

Pour obtenir des instructions pour créer des alias, voir le *Guide de Designer*.

10.9.5 Création d'un alias

1. Dans l'éditeur de banque de données, cliquez sur *Avancé*, puis sur *Alias (cliquez ici pour les créer)*.

La fenêtre Créer un alias s'ouvre.

2. Sous *Nom de l'alias dans Designer*, utilisez uniquement des caractères alphanumériques et le symbole de tiret de soulignement (_) pour saisir un nom d'alias.
3. Cliquez sur *OK*.

La fenêtre *Créer un alias* se ferme et le nouvel alias apparaît sous la catégorie Alias.

Lorsque vous définissez un alias de banque de données, le logiciel remplace l'alias de la configuration de banque de données indiqué pour le nom du propriétaire réel lors de l'importation de métadonnées pour les objets de base de données. Il est également possible de renommer les tables et les fonctions après les avoir importées.

Informations associées

[Modification du nom du propriétaire de table et de fonction \[page 142\]](#)

10.9.6 Fonctions pour identifier la configuration

Le logiciel fournit des fonctions qui sont utiles lors de l'utilisation de plusieurs configurations de banque de données cible et source.

SAP Data Services	Catégorie SAP Data Services	Description
Fonction		
type_de_BdD	Divers	Renvoie le type de base de données de la configuration de banque de données actuelle.
db_version	Divers	Renvoie la version de base de données de la configuration de banque de données actuelle.
db_database_name	Divers	Renvoie le nom de la base de données de la configuration de banque de données actuelle si le type de base de données est MS SQL Server ou SAP ASE.
db_owner	Divers	Renvoie le nom du propriétaire réel qui correspond au nom d'alias donné sous la configuration de banque de données actuelle.
current_configuration	Divers	Renvoie le nom de la configuration de banque de données qui est utilisée lors de l'exécution.
current_system_configuration	Divers	Renvoie le nom de la configuration système actuelle. Si aucune configuration système n'est définie, renvoie une valeur nulle.

Le logiciel lie les paramètres de transformation SQL et d'éditeur de table cible utilisés dans un flux de données aux configurations de banque de données. Il est également possible d'utiliser une interpolation de variable dans un texte SQL avec ces fonctions pour permettre à la transformation SQL de s'effectuer correctement sans tenir compte de la configuration utilisée par Job server lors de l'exécution du job.

Utilisez l'Administrateur pour sélectionner une configuration système ainsi que pour afficher la configuration de banque de données sous-jacente associée lorsque vous :

- Exécutez des jobs batch
- Planifiez des jobs batch
- Affichez l'historique des jobs batch
- Créez des services pour les jobs en temps réel

Pour utiliser plusieurs configurations correctement, concevez les jobs pour ne pas avoir à modifier les schémas, types de données, fonctions, variables, etc. lorsque vous passez d'une configuration de banque de données à une autre. Par exemple, si vous avez une banque de données avec une configuration pour des sources Oracle et des sources SQL, assurez-vous que les schémas de métadonnées de table correspondent exactement. Utilisez les mêmes noms de table, noms d'alias, numéros et ordre de colonnes, ainsi que les mêmes noms de colonne, types de données et types de contenu.

Informations associées

[Astuces de portabilité des jobs \[page 138\]](#)

10.9.7 Configuration des banques de données dans les flux de données

Lorsque vous utilisez une banque de données de configuration multiple dans un flux de données, soyez prudent lorsque vous modifiez les configurations.

Lorsque le flux de données contient les objets suivants, la modification de la configuration de la banque de données peut entraîner une modification de la langue du flux de données :

- Cibles de table
- Type de transfert de table utilisé dans la transformation Data_Transfer comme une cible
- transformations SQL

Le logiciel ajoute les options de cible et le texte de la transformation SQL aux configurations de banque de données supplémentaires en fonction de leurs définitions dans la configuration existante. Le tableau suivant décrit les options de cible que vous pouvez utiliser lorsque vous ajoutez une nouvelle configuration à une banque de données existante.

Nouvelle configuration	Options cible
Le même type de base de données et la même version ou une version plus récente qu'une configuration existante	Le logiciel utilise automatiquement les valeurs d'une transformation SQL, d'un éditeur de table cible ou d'un éditeur de transformation Data_Transfer existants, le cas échéant, y compris les options de Bulk Loader.

Nouvelle configuration	Options cible
Autre type de base de données ou le même type de base de données, mais une version plus ancienne que la configuration existante	Vous pouvez sélectionner des valeurs dans un type et une version de base de données existants en sélectionnant l'option dans la liste Utiliser les valeurs de .

La liste [Utiliser les valeurs de](#) contient toujours les options suivantes :

- Valeurs par défaut
- Type et version de base de données pour chaque configuration actuellement associée à la banque de données

❖ Exemple

La banque de données contient deux configurations : Oracle 9i et Microsoft SQL Server 2000. La banque de données est connectée à un flux de données. Puis, vous ajoutez une troisième configuration à la banque de données pour DB2. Les valeurs par défaut suivantes s'affichent dans la liste [Utiliser les valeurs de](#) :

- Oracle 9i
- Microsoft SQL Server 2000

Les valeurs par défaut sont les mêmes que celles qui apparaissent pour toutes les cibles de base de données, les tables cibles Data_Transfer et les transformations SQL. Le texte SQL par défaut est toujours vide. La liste suivante contient une partie des valeurs d'options cible par défaut :

- [Taille de la validation de la ligne](#) = 1000
- [Comparaison de colonne](#) = Comparer par nom
- [Supprimer les données de la table avant le chargement](#) = non sélectionnée
- [Supprimer et recréer la table](#) = non sélectionnée pour les tables régulières (sélectionnée pour les tables de modèle)

Si vous sélectionnez la case à cocher [Restaurer les valeurs si elles existent déjà](#), qui est sélectionnée par défaut, le logiciel crée la nouvelle configuration. Le logiciel détermine si les valeurs de la transformation SQL, de l'éditeur de table cible ou de l'éditeur de transformation Data_Transfer existent déjà pour la nouvelle base de données. Si les valeurs de la base de données existent déjà, le logiciel restaure l'option de chargement par lot. Cependant, si aucune valeur n'existe pour la base de données, le logiciel définit l'option de chargement par lot sur [Aucun](#), la valeur par défaut.

De la même façon, si vous décochez [Restaurer des valeurs si elles existent déjà](#), le logiciel définit l'option de chargement par lot sur [Aucun](#), la valeur par défaut.

❖ Exemple

Vous travaillez dans un environnement multi-utilisateurs et vous avez une banque de données locale qui dispose de configurations pour Oracle 9i et SQL Server 2000. Vous avez également des flux de données existants qui utilisent des tables cible de cette banque de données. Vous supprimez ensuite Oracle 9i parce que vous avez extrait une version différente de la banque de données du référentiel central. Par la suite, vous souhaitez ajouter une configuration Oracle 9i à cette banque de données.

Le fait de supprimer une version conduit le logiciel à supprimer la configuration, mais pas les valeurs de la table cible, de la table cible Data_Transfer ou de la transformation SQL. Si vous sélectionnez [Restaurer des valeurs si elles existent déjà](#) lors de la création d'une nouvelle configuration, le logiciel détermine si les valeurs existent

déjà pour la base de données. Si le logiciel ne peut pas trouver ces valeurs, le logiciel utilise des valeurs spécifiées dans la case [Utiliser les valeurs de](#).

Après avoir cliqué sur [Appliquer](#) pour enregistrer la nouvelle configuration, le logiciel :

- copie les valeurs existantes de la transformation SQL, de l'éditeur de table cible et de l'éditeur de table cible Data_Transfer.
- Affiche un rapport des objets modifiés dans une fenêtre pop-up.
- Affiche un rapport des objets modifiés dans la boîte de dialogue de sortie.

[About the Use Values From option \[page 134\]](#)

SAP Data Services populates values in the data flow based on the setting in the [Use Values From](#) option in the [Create New Configuration](#) dialog box.

[Contenus du rapport pour les objets modifiés \[page 135\]](#)

Le logiciel affiche un rapport pour les objets modifiés une fois que vous avez ajouté une nouvelle configuration à une banque de données existante et cliqué sur [Appliquer](#).

Informations associées

[Contenus du rapport pour les objets modifiés \[page 135\]](#)

10.9.7.1 About the Use Values From option

SAP Data Services populates values in the data flow based on the setting in the [Use Values From](#) option in the [Create New Configuration](#) dialog box.

The values that you set in the [Use values from](#) dropdown list affect the values that Data Services uses to populate options in the target table, SQL transform, and—or the Data_Transfer target table. For information about the values available in [Use values from](#), see [Configuration des banques de données dans les flux de données \[page 132\]](#).

The [Use values from](#) list always contains the following options:

- [Default values](#)
- Database type and version for each configuration currently associated with the datastore

❖ Exemple

Your datastore contains two configurations: Oracle 9i and Microsoft SQL Server 2000. The data flow uses a table from the datastore. When you add a new configuration for DB2 to the datastore, the following values appear in the [Use values from](#) dropdown menu:

- [Default values](#)
- Oracle 9i
- Microsoft SQL Server 2000

When you select [Default values](#), Data Services uses the same defaults that appear for all database targets, Data_Transfer targets, and—or SQL transforms. In SQL transforms, the default SQL text is always blank. The following list contains some of the default target option values:

- *Row commit size* = 1000
- *Column comparison* = Compare by name
- *Delete data from table before loading* = not selected
- *Drop and re-create table* = not selected for regular tables (Selected for template tables)

The *Restore values if they already exist* checkbox in the *Create New Configuration* dialog box is selected by default. If the configuration is for the same database type and version, the software uses values from a deleted configuration. If you deselect the checkbox, you supply the values.

After you click *OK* to save a new configuration, the software performs the following tasks:

- Copies any existing SQL transform, target table, and Data_Transfer target table values.
- Displays a report of the modified objects in a *Modified Object Report* dialog box.
- Displays a report of the modified objects in the *Output* dialog box.

10.9.7.2 Contenus du rapport pour les objets modifiés

Le logiciel affiche un rapport pour les objets modifiés une fois que vous avez ajouté une nouvelle configuration à une banque de données existante et cliqué sur *Appliquer*.

Description du rapport pour les objets modifiés

Colonne du rapport	Description
Flux de données	Noms des flux de données où la langue a été modifiée.
Objets modifiés	Objets dans les flux de données qui ont été affectés.
Type d'objet	Types d'objets concernés par la modification. Par exemple, la table cible ou la transformation SQL.
Utilisation	Utilisation des objets, tels que la source ou la cible.
Possède un Bulk Loader.	Si les objets ont un Bulk Loader.
Bulk Loader copié	Si l'option du Bulk Loader est copiée.
Anciennes valeurs	S'il existe des anciennes valeurs.
Valeurs restaurées	Si les anciennes valeurs sont restaurées.

Il est possible d'utiliser ce rapport comme guide pour modifier manuellement les valeurs pour les options de cibles, de tables cible Data_Transfer et de transformations SQL, si nécessaire. Dans la boîte de dialogue qui contient le rapport des objets modifiés, vous pouvez trier les résultats en cliquant sur les en-têtes de colonne. Il est également possible d'enregistrer la sortie vers un fichier. La fenêtre pop-up apparaît une fois que vous avez cliqué sur *Appliquer* pour chaque nouvelle configuration ajoutée.

Le logiciel efface et affiche également les résultats dans la fenêtre Sortie après chaque configuration ajoutée. Puisque les fenêtres d'éditeur de banque de données sont modales, il est impossible de voir la fenêtre Sortie en entier ou de la manipuler. Cependant, il est possible de cliquer deux fois sur l'un des objets dans le rapport pour afficher le flux de données.

Informations associées

[Configuration des banques de données dans les flux de données \[page 132\]](#)

10.9.8 Solutions de portabilité

Définissez plusieurs configurations sources ou cibles pour une seule banque de données si vous souhaitez modifier rapidement les connexions à une base de données source ou cible différente.

Le logiciel offre plusieurs solutions différentes pour les job de déplacement.

[Migration entre environnements \[page 136\]](#)

Le déplacement des métadonnées du référentiel vers un autre environnement (par exemple du développement au test, ou du test à la production) utilise des bases de données sources et cibles différentes.

[Chargement de plusieurs instances \[page 137\]](#)

Le scénario de migration pour le chargement de plusieurs instances d'une source de données dans un entrepôt de données cible, est identique au scénario de migration, sauf que vous utilisez un seul référentiel.

[Déploiement d'OEM \[page 137\]](#)

Vous pouvez concevoir des jobs pour un type de base de données et déployez ces jobs vers d'autres types de base de données comme partenaire OEM.

[Astuces de portabilité des jobs \[page 138\]](#)

[Développement multi-utilisateurs \[page 140\]](#)

Lorsque vous utilisez le système de gestion du référentiel central, plusieurs développeurs, chacun avec leur propre référentiel local, peuvent recharger et extraire les jobs.

Informations associées

[Travail dans un environnement multi-utilisateurs \[page 909\]](#)

10.9.8.1 Migration entre environnements

Le déplacement des métadonnées du référentiel vers un autre environnement (par exemple du développement au test, ou du test à la production) utilise des bases de données sources et cibles différentes.

Le processus inclut généralement les caractéristiques suivantes :

- Les environnements utilisent le même type de base de données mais peuvent avoir des versions ou des paramètres régionaux de base de données uniques.
- Les objets de base de données (tables et fonctions) peuvent appartenir à des propriétaires différents.
- Chaque environnement a un nom de connexion à la base de données, un nom d'utilisateur, un mot de passe, d'autres propriétés de connexion et un mappage du propriétaire uniques.

- Vous utilisez une procédure de migration du référentiel typique. Soit vous exportez les jobs vers un fichier ATL puis importez le fichier ATL dans un autre référentiel, soit vous exportez directement les jobs d'un référentiel à un autre.

Puisque le logiciel remplace les configurations de banques de données lors de l'exportation, il est recommandé d'ajouter des configurations pour l'environnement cible (par exemple, ajouter des configurations pour l'environnement test lors de la migration du développement au test) au référentiel source (par exemple, ajouter au référentiel de développement avant la migration vers l'environnement test). L'utilitaire Exportation enregistre les configurations supplémentaires dans l'environnement cible, ce qui signifie que vous n'avez pas besoin de modifier les banques de données avant d'exécuter des jobs déplacés dans l'environnement cible.

Cette solution offre les avantages suivants :

- Temps de d'arrêt de production minimum : vous pouvez démarrer les jobs dès que vous les avez exportés.
- Problèmes de sécurité minimaux : les testeurs et les opérateurs en production n'ont pas besoin d'autorisation pour modifier les objets du référentiel.

10.9.8.2 Chargement de plusieurs instances

Le scénario de migration pour le chargement de plusieurs instances d'une source de données dans un entrepôt de données cible, est identique au scénario de migration, sauf que vous utilisez un seul référentiel.

10.9.8.3 Déploiement d'OEM

Vous pouvez concevoir des jobs pour un type de base de données et déployez ces jobs vers d'autres types de base de données comme partenaire OEM.

Le déploiement inclut généralement les caractéristiques suivantes :

- L'instance nécessite plusieurs types et versions de base de données source.
- Puisqu'une banque de données peut seulement accéder à une instance à la fois, il est possible que vous deviez déclencher des fonctions lors de l'exécution pour correspondre aux différentes instances. Si c'est le cas, le logiciel nécessite un texte SQL différent pour les fonctions (telles que lookup_ext et sql) et les transformations (telles que la transformation SQL). Le logiciel nécessite également des paramètres différents pour la table cible (configurables dans l'éditeur de table cible).
- Les instances peuvent utiliser des paramètres régionaux différents.
- Les tables de base de données dans différentes bases de données appartiennent à des propriétaires différents.
- Chaque instance a un nom de connexion à la base de données, un nom d'utilisateur, un mot de passe, d'autres propriétés de connexion et des mappages du propriétaire uniques.
- Vous exportez les jobs vers des fichiers ATL pour le déploiement.

[Déployer des jobs vers d'autres types de base de données en tant que partenaires OEM. \[page 138\]](#)

10.9.8.3.1 Déployer des jobs vers d'autres types de base de données en tant que partenaires OEM.

1. Développez des jobs pour un type particulier de base de données en suivant les étapes décrites dans le scénario [Chargement de plusieurs instances \[page 137\]](#).

Pour prendre en charge une nouvelle instance sous un nouveau type de base de données, le logiciel copie les propriétés de base de données de la table cible et de la transformation SQL à partir de la configuration précédente vers chaque configuration supplémentaire lorsque vous l'enregistrez.

Si vous avez sélectionné une méthode de chargeur par lots pour une ou plusieurs table(s) cible dans les flux de données du job, et que de nouvelles configurations s'appliquent à des types de base de données différents, ouvrez les cibles et définissez manuellement l'option de chargeur par lots (si vous souhaitez toujours utiliser la méthode de chargeur par lots avec le nouveau type de base de données). Le logiciel ne copie pas les options de chargeur par lot pour les cibles d'un type de base de données à un autre.

Lorsque le logiciel enregistre une nouvelle configuration, il génère également un rapport qui fournit une liste des cibles automatiquement définies pour le chargement par lots. Faites référence à ce rapport pour apporter des modifications manuelles si besoin.

2. Si le texte SQL dans une transformation SQL ne s'applique pas au nouveau type de base de données, modifiez le texte SQL pour le nouveau type de base de données.

Si le texte SQL contient des noms de propriétaire ou de base de données codés en dur, pensez à remplacer ces noms avec des variables pour fournir des noms de propriétaire et des noms de base de données pour plusieurs types de base de données. Ainsi, vous ne devez pas modifier le texte SQL pour chaque environnement.

3. Puisque le logiciel ne prend pas en charge un texte SQL unique pour chaque type ou version de base de données des fonctions `sql()`, `lookup_ext()`, et `pushdown_sql()`, utilisez la fonction `db_type` et les fonctions similaires pour obtenir le type et la version de base de données de la configuration de banque de données actuelle et fournir le texte SQL correct pour ce type et cette version de base de données à l'aide de la technique de substitution (interpolation) de variables.

10.9.8.4 Astuces de portabilité des jobs

- Le logiciel suppose que les métadonnées d'une table ou d'une fonction sont les mêmes dans les différents types et versions de base de données indiqués dans les différentes configurations d'une même banque de données. Par exemple, si vous importez une table lorsque la configuration par défaut de la banque de données est Oracle, puisque vous utilisez la table dans un job pour extraire depuis DB2, le job exécutera.
- Importez les métadonnées pour un objet de base de données à l'aide de la configuration par défaut et utilisez les mêmes métadonnées avec toutes les configurations définies dans la même banque de données.
- Le logiciel prend en charge ses options dans certains types ou versions de base de données qu'il ne prend pas en charge dans d'autres. Par exemple, le logiciel prend en charge la lecture parallèle sur les tables partitionnées par hachage Oracle, mais pas sur les tables partitionnées par hachage de DB2 ou d'autres bases de données. Si vous importez une table partitionnée par hachage Oracle et définissez le flux de données pour exécuter en parallèle, le logiciel lit à partir de chaque partition en parallèle. Cependant, lorsque vous exécutez un job à l'aide des sources de l'environnement DB2, la lecture parallèle ne se produit pas.

- Les fonctionnalités suivantes prennent en charge la portabilité de job :
 - Transformation SQL améliorée
Avec la transformation SQL améliorée, il est possible de saisir un texte SQL différent pour les différents types/version de base de données, et d'utiliser une substitution de variables dans le texte SQL pour permettre au logiciel de lire le bon texte pour la configuration de banque de données associée.
 - Editeur de table cible amélioré
A l'aide des options de l'éditeur de table cible amélioré, il est possible de configurer les cibles de table de base de données pour les différents types/version de base de données pour correspondre à leurs configurations de banque de données.
 - Editeur de banque de données amélioré
A l'aide de l'éditeur de banque de données amélioré, il est possible de choisir lors de la création d'une nouvelle configuration de banque de données de copier les propriétés de la base de données (y compris les options de cible de table et de banque de données, ainsi que le texte de la transformation SQL) à partir d'une configuration existante, ou d'utiliser les valeurs actuelles.
- Lors de la conception d'un job qui est exécuté à partir de types ou versions de base de données différentes, nommez les tables de base de données, les fonctions, et les procédures stockées de la même façon pour toutes les sources. Si vous créez des configurations pour des bases de données non sensibles à la casse et pour des bases de données sensibles à la casse dans la même banque de données, il est recommandé de nommer les tables, fonctions, et procédures stockées avec des caractères tout en majuscules.
- Les schémas de table doivent correspondre dans les bases de données de la banque de données. Cela signifie que le nombre de colonnes, les noms des colonnes, et les positions des colonnes doit être exactement les mêmes. Les types de données de la colonne doivent être les mêmes ou compatibles. Par exemple, si vous avez une colonne VARCHAR dans une source Oracle, utilisez également une colonne VARCHAR dans la source Microsoft SQL Server. Si vous avez une colonne DATE dans une source Oracle, utilisez une colonne DATETIME dans la source Microsoft SQL Server. Définissez les clés primaires et étrangères de la même manière.
- Les schémas de procédures stockées doivent correspondre. Lorsque vous importez une procédure stockées depuis une configuration de banque de données, et que vous l'utilisez pour une autre configuration de banque de données, le logiciel suppose que la signature de la procédure stockées est exactement la même pour les deux bases de données. Par exemple, si une procédure stockée est une fonction stockée (uniquement les fonctions stockées prises en charge par Oracle), alors vous devez l'utiliser comme une fonction avec toutes les autres configurations dans une banque de données (en d'autres termes, toutes les bases de données doivent être Oracle). Si la procédure stockée a trois paramètres dans une base de données, elle doit avoir exactement trois paramètres dans les autres bases de données. De plus, les noms, positions, types de données et les types in/out des paramètres doivent correspondre exactement.

Informations associées

[Développement multi-utilisateurs \[page 895\]](#)

[Configuration d'environnement multi-utilisateurs \[page 900\]](#)

10.9.8.5 Développement multi-utilisateurs

Lorsque vous utilisez le système de gestion du référentiel central, plusieurs développeurs, chacun avec leur propre référentiel local, peuvent recharger et extraire les jobs.

L'environnement de développement inclut généralement les caractéristiques suivantes :

- Il possède un référentiel central et un nombre de référentiels locaux.
- Plusieurs environnements de développement sont parfois fusionnés (via des opérations de référentiel central telles que les rechargements et les extractions). Lorsque la fusion se produit, les noms de propriétaire réels (utilisés au départ pour l'importation d'objets) doivent être ensuite mappés à un ensemble d'alias partagés entre tous les utilisateurs.
- Le logiciel conserve l'historique des objets (versions et étiquettes).
- Les instances partagent le même type de base de données mais peuvent avoir des versions ou des paramètres régionaux de base de données différents.
- Les objets de base de données peuvent appartenir à des propriétaires différents.
- Chaque instance a un nom de connexion à la base de données, un nom d'utilisateur, un mot de passe, d'autres propriétés de connexion et un mappage du propriétaire uniques.

Dans le scénario de développement multi-utilisateurs, vous devez définir des alias pour que le logiciel puisse conserver correctement l'historique pour tous les objets dans l'environnement partagé.

10.9.8.5.1 Déplacement de jobs dans un environnement multi-utilisateurs

Lors du déplacement de jobs dans un environnement multi-utilisateurs, prenez en compte ces points :

- Renommez les propriétaires de table et les propriétaires de fonction pour consolider les noms de propriétaire des objets de base de données d'objets dans des alias
 - La modification du nom se produit dans les référentiels locaux. Pour renommer les objets de base de données stockés dans le référentiel central, extrayez la banque de données vers un référentiel local et appliquez l'outil de renommage dans le référentiel local.
 - Si les objets à renommer ont des objets dépendants, le logiciel vous demande d'extraire les objets dépendants.
 - Si tous les objets dépendants peuvent être extraits, le renommage crée un nouvel objet qui possède l'alias et supprime l'objet d'origine qui possède le nom du propriétaire d'origine.
 - Si tous les objets dépendants ne peuvent pas être extraits (les flux de données sont extraits par un autre utilisateur), le logiciel affiche un message qui vous laisse l'option de continuer ou d'annuler l'opération. Si vous ne pouvez pas extraire certains objets dépendants, l'outil de renommage affecte uniquement les flux que vous pouvez extraire. Après le renommage, l'objet d'origine coexiste avec le nouvel objet. Le nombre de flux touchés par le processus de renommage affecte les informations sur l'utilisation et les cas d'emploi dans Designer pour l'objet d'origine et le nouvel objet.
- Vous êtes responsable du rechargement de tous les objets dépendants qui ont été extraits durant le processus de renommage du propriétaire. Le rechargement de nouveaux objets n'entraîne pas automatiquement le rechargement des objets dépendants qui ont été extraits.
 - Le logiciel ne supprime pas les objets d'origine du référentiel central lorsque vous rechargez les nouveaux objets.

- Soyez prudents car le fait de recharger dans les banques de données et de les extraire comme des opérations multi-utilisateur peut remplacer les configurations de banque de données.
- Conservez les configurations de banque de données de tous les utilisateurs en ne remplaçant pas les configurations qu'ils ont créées. Ajoutez plutôt une configuration et définissez-la comme configuration par défaut lorsque vous travaillez dans votre propre environnement.
 - Lorsque votre groupe termine la phase de développement, il est recommandé que le dernier développeur supprime les configurations qui s'appliquent aux environnements de développement et ajoutez les configurations qui s'appliquent aux environnements de test ou de production.

10.9.9 Rename Owner option

A shared alias name in a central repository environment makes tracking the object easier when it has been checked out by multiple users.

If all users in a multiple user environment keep the alias name when they work with the object in their local repositories, Data Services tracks dependencies for objects that your team checks in and out of the central repository.

You can rename the owner of imported tables, template tables, and functions. When you rename an owner name with an alias name, the instances of a table or function in a data flow are affected. The datastore from which they were imported is not affected.

The software supports both case-sensitive and case-insensitive owner renaming in the following ways:

- If the objects to be renamed are from a case-sensitive database, Data Services preserves case sensitivity.
- If the objects to be renamed are from a datastore that contains both case-sensitive and case-insensitive databases, Data Services bases the casing of the new owner name on the case sensitivity of the default configuration.

To ensure that all objects are portable across all configurations, enter all owner names and object names using uppercase characters.

During the owner renaming process, Data Services:

- Updates the dependent objects to use the new owner name. Dependent objects include jobs, work flows, and data flows that use the renamed object.
- Lists the object in the object library with the new owner name. The object library displays usage or where used information that reflects the number of updated dependent objects.
- Deletes the metadata for the object with the original owner name from the object library and the repository only after it successfully updates all the dependent objects.

[Modification du nom du propriétaire de table et de fonction \[page 142\]](#)

Utilisez le renommage de propriétaire pour affecter un alias de métadonnées unique au lieu du nom du propriétaire réel pour les objets de la base de données dans la banque de données.

[Utilisation de la fenêtre Renommer dans un scénario multi-utilisateurs \[page 143\]](#)

L'utilisation d'un alias pour tous les objets stockés dans un référentiel central permet au logiciel de suivre tous les objets rechargés par plusieurs utilisateurs.

10.9.9.1 Modification du nom du propriétaire de table et de fonction

Utilisez le renommage de propriétaire pour affecter un alias de métadonnées unique au lieu du nom du propriétaire réel pour les objets de la base de données dans la banque de données.

Vous pouvez renommer le propriétaire des tables importées, tables de modèle ou des fonctions. La consolidation de métadonnées sous un nom d'alias unique permet d'accéder à des informations cohérentes et précises sur les dépendances à tout moment tout en permettant également de passer plus facilement d'une configuration à une autre lorsque vous déplacez les jobs dans différents environnements.

Lors de l'utilisation d'objets stockés dans un référentiel central, un alias partagé permet de suivre facilement les objets rechargés par plusieurs utilisateurs. Si tous les utilisateurs des référentiels locaux utilisent le même alias, le logiciel peut suivre les dépendances pour les objets rechargés et extraits du référentiel central par votre équipe.

Lorsque vous renommez un propriétaire, les instances d'une table ou d'une fonction dans un flux de données sont affectées, et non pas la banque de données depuis laquelle elles ont été importées.

1. À partir de l'onglet Banque de données de la bibliothèque d'objets locale, développez une table, une table de modèle ou une catégorie de fonction.
2. Faites un clic droit sur la table ou sur la fonction et sélectionnez *Renommer le propriétaire*.
3. Saisissez un Nouveau nom du propriétaire puis cliquez sur *Renommer*.

Lorsque vous saisissez un Nouveau nom du propriétaire, le logiciel l'utilise comme un alias de métadonnées pour la table ou la fonction.

i Remarque

Si l'objet que vous renommez existe déjà dans la banque de données, le logiciel détermine si les deux objets ont le même schéma. Si tel est le cas, le logiciel poursuit. Si les schémas sont différents, le logiciel affiche un message à cet effet. Il est possible que vous deviez choisir un nom d'objet différent.

Le logiciel prend en charge les renommages de propriétaire sensibles à la casse ou non sensibles à la casse.

- Si les objets que vous souhaitez renommer proviennent d'une base de données sensible à la casse, le mécanisme de renommage du propriétaire conserve la sensibilité à la casse.
- Si les objets que vous souhaitez renommer proviennent d'une banque de données qui contient des bases de données sensibles à la casse et non sensibles à la casse, le logiciel reposera la sensibilité à la casse des nouveaux noms de propriétaire sur la sensibilité à la casse de la configuration par défaut. Pour s'assurer que tous les objets sont portables dans toutes les configurations de ce scénario, saisissez tous les noms de propriétaire et les noms d'objet avec des caractères tout en majuscules.

Lors du processus de renommage du propriétaire :

- Le logiciel met à jour les objets dépendants (jobs, workflows et flux de données qui utilisent l'objet renommé) pour utiliser le nouveau nom du propriétaire.
- La bibliothèque d'objets affiche l'entrée de l'objet avec le nouveau nom du propriétaire. Les informations d'utilisation et de cas d'emploi affichées reflètent le nombre d'objets dépendants mis à jour.
- Si le logiciel met à jour avec succès tous les objets dépendants, il supprime les métadonnées pour l'objet avec le nom du propriétaire d'origine de la bibliothèque d'objets et du référentiel.

10.9.9.2 Utilisation de la fenêtre Renommer dans un scénario multi-utilisateurs

L'utilisation d'un alias pour tous les objets stockés dans un référentiel central permet au logiciel de suivre tous les objets rechargés par plusieurs utilisateurs.

Si tous les utilisateurs des référentiels locaux utilisent le même alias, le logiciel peut suivre les dépendances pour les objets rechargés et extraits du référentiel central par votre équipe.

Lorsque vous rechargez et extrayez les objets d'un référentiel central, en fonction du statut d'extraction d'un objet renommé et de l'éventuelle association de l'objet à des objets dépendants, il existe plusieurs comportements possibles lorsque vous sélectionnez le bouton [Renommer](#).

Scénario	Comportement lors de l'utilisation de la fonctionnalité Renommer.
1 - L'objet n'est pas extrait et n'a aucun objet dépendant dans le référentiel local ou central.	Lorsque vous cliquez sur Renommer , le logiciel renomme le propriétaire de l'objet.
2 - L'objet est extrait et n'a aucun objet dépendant dans le référentiel local ou central.	Lorsque vous cliquez sur Renommer , le logiciel renomme le propriétaire de l'objet.
3 - L'objet n'est pas extrait et a un ou plusieurs objet(s) dépendant(s) (dans le référentiel local).	Lorsque vous cliquez sur Renommer, le logiciel affiche une seconde fenêtre avec une liste des objets dépendants (qui utilisent ou font référence à l'objet renommé). Si vous cliquez sur Continuer, le logiciel renomme les objets et modifie les objets dépendants qui font référence à l'objet renommé à l'aide du nouveau nom du propriétaire. Si vous cliquez sur Annuler, Designer renvoie à la fenêtre Renommer le propriétaire. i Remarque Il est possible qu'un objet ait toujours un ou plusieurs objet(s) dépendant(s) dans le référentiel central. Cependant, si l'objet à renommer n'est pas extrait, le mécanisme Renommer le propriétaire (par dessin) n'affecte pas les objets dépendants dans le référentiel central.

Scénario

4 - L'objet est extrait et a un ou plusieurs objet(s) dépendant(s).

Comportement lors de l'utilisation de la fonctionnalité Renommer.

Ce scénario est relativement complexe.

- Si vous n'êtes pas connecté au référentiel central, le message de statut est :

```
This object is checked out from
central repository <X>. Please
select Tools | Central Repository...
to activate that repository before
renaming.
```

- Si vous êtes connecté au référentiel central, la fenêtre Renommer le propriétaire s'ouvre.
Lorsque vous cliquez sur [Renommer](#), une deuxième fenêtre s'ouvre pour afficher les objets dépendants et un statut qui indique leur état d'extraction et leur emplacement. Si un objet dépendant se trouve dans le référentiel local uniquement, le message de statut vous indiquera qu'il est uniquement utilisé dans un référentiel local et que l'extraction n'est pas nécessaire.

- Si l'objet dépendant se trouve dans le référentiel central et qu'il n'est pas extrait, le message de statut vous indiquera qu'il n'est pas extrait.

- Si l'objet dépendant est extrait ou extrait par un autre utilisateur, le message de statut affiche le nom du référentiel extrait. Par exemple,
`Oracle.production.user1.`

Comme dans le scénario 2, le but de cette deuxième fenêtre est d'afficher les objets dépendants. De plus, cette fenêtre permet d'extraire les objets dépendants nécessaires du référentiel central, sans avoir à accéder à la fenêtre Bibliothèque d'objets centrale.

Cliquez sur le bouton [Actualiser la liste](#) pour mettre à jour le statut d'extraction dans la liste. Cette mise à jour est utile lorsque le logiciel identifie un objet dépendant dans le référentiel central mais qu'un autre utilisateur l'a extrait.

Lorsque cet utilisateur recharge l'objet dépendant, cliquez sur [Actualiser la liste](#) pour mettre à jour le statut et vérifier que l'objet dépendant n'est plus extrait.

→ Conseil

Pour tirer le meilleur parti de la fonctionnalité Renommer le propriétaire, extrayez les objets dépendants associés du référentiel central. Cette extraction aide à éviter d'avoir des objets dépendants qui font référence à des objets avec des noms de pro-

Scénario	Comportement lors de l'utilisation de la fonctionnalité Renommer.
	propriétaire qui n'existent pas. A partir du référentiel central, sélectionnez un ou plusieurs objet(s), puis faites un clic droit et sélectionnez <i>Extraire</i>. Après avoir extrait l'objet dépendant, Designer met à jour le statut. Si l'extraction était réussie, le statut affiche le nom du référentiel local.
4a - Vous cliquez sur <i>Continuer</i>, mais un ou plusieurs objet(s) dépendant(s) ne sont pas extraits du référentiel central.	Dans cette situation, le logiciel affiche une autre boîte de dialogue pour vous prévenir des objets pas encore traités et pour confirmer votre désir de continuer. Vous devez : • cliquer sur <i>Non</i> pour revenir à la boîte de dialogue précédente qui affiche les objets dépendants ; • cliquer sur <i>Oui</i> pour continuer à renommer l'objet sélectionné et modifier ses objets dépendants. Le logiciel modifie les objets qui ne sont pas extraits dans le référentiel local pour faire référence au nouveau nom de propriétaire. Il est de votre responsabilité de maintenir une cohérence avec les objets dans le référentiel central.
4b - Vous cliquez sur <i>Continuer</i> et tous les objets dépendants sont extraits du référentiel central.	Le logiciel renomme le propriétaire de l'objet sélectionné et modifie tous les objets dépendants pour faire référence au nouveau nom du propriétaire. Bien qu'il semble que l'objet d'origine ait un nouveau nom du propriétaire, le logiciel n'a en réalité pas modifié l'objet d'origine ; il a créé un nouvel objet identique à l'original mais utilise le nouveau nom du propriétaire. L'objet d'origine avec l'ancien nom du propriétaire existe toujours. Le logiciel effectue alors une "annulation de l'extraction" sur l'objet d'origine. C'est alors à vous de devoir recharger l'objet renommé. Lorsque l'opération de renommage est réussie, dans l'onglet <i>Banque de données</i> de la bibliothèque d'objets locale, le logiciel met à jour la table ou la fonction avec le nouveau nom du propriétaire et la fenêtre Sortie affiche le message suivant : <pre>Object <Object_Name>: Owner name <Old_Owner> successfully renamed to <New_Owner>, including references from dependent objects.</pre> Si le logiciel n'a pas renommé avec succès le propriétaire, la fenêtre Sortie affiche le message suivant : <pre>Object <Object_Name>: Owner name <Old_Owner> could not be renamed to <New_Owner>.</pre>

10.9.10 Définition d'une configuration système

Lors de la conception de jobs, déterminez et créez des configurations de banque de données et des configurations système en fonction de votre environnement et de vos règles d'entreprise.

Pour ce faire, vous devez connaître la différence entre les configurations de banque de données et les configurations système.

<i>Configurations de banque de données</i>	Chaque configuration de banque de données définit une connexion à une base de données particulière depuis une banque de données unique.
<i>Configurations système</i>	Chaque configuration système définit un ensemble de configurations de banque de données qu'il est préférable d'utiliser ensemble lors de l'exécution d'un job. Il est possible de définir une configuration de système si le référentiel contient au moins une banque de données avec plusieurs configurations. Il est également possible d'associer des configurations de paramètres de substitution à des configurations système.

Créez des configurations de banque de données pour les banques de données du référentiel avant de créer des configurations système pour les organiser et les associer.

Sélectionnez une configuration de système à utiliser au moment de l'exécution. Dans de nombreuses entreprises, un concepteur de jobs définit les configurations de banque de données et de système requises, puis un administrateur système détermine quelle configuration de système utiliser lors de la planification et du lancement d'un job.

Le logiciel conserve les configurations système séparément des jobs. Il est impossible de recharger ou d'extraire les configurations système dans un environnement multi-utilisateurs. Cependant, il est possible d'exporter des configurations système vers un fichier plat distinct que vous pouvez ensuite importer.

[Création d'une configuration système \[page 146\]](#)

[Creating a system configuration by duplication \[page 147\]](#)

Use a duplicate system configuration as a template to create a new configuration.

[Renaming and deleting system configurations \[page 148\]](#)

The *Edit System Configurations* dialog box has toolbar icons for renaming and deleting system configurations.

[Exportation d'une configuration système \[page 148\]](#)

10.9.10.1 Création d'une configuration système

1. Dans la barre de menu de Designer, sélectionnez ► *Outils* ► *Configurations système* ►.

La fenêtre *Modifier les configurations système* s'affiche.

2. Pour ajouter une nouvelle configuration système, réalisez l'une des opérations suivantes :
 - Cliquez sur l'icône *Créer une configuration* pour ajouter une configuration qui référence la configuration par défaut des paramètres de substitution et chaque connexion à la banque de données.
 - Sélectionnez une configuration existante et cliquez sur l'icône *Dupliquer la configuration* pour créer une copie de la configuration sélectionnée.

Il est possible d'utiliser la copie comme modèle et de modifier les sélections de paramètres de substitution ou de configuration de banque de données pour qu'elles s'adaptent à vos besoins.

3. Renommez la nouvelle configuration système si vous le souhaitez.
 - a. Sélectionnez la configuration système que vous souhaitez renommer.
 - b. Cliquez sur l'icône [Renommer la configuration](#) pour activer le mode de modification pour le champ du nom de la configuration.
 - c. Saisissez un nouveau nom unique et cliquez en dehors du champ du nom pour accepter votre choix.

Il est recommandé de suivre une convention d'affectation de nom cohérente et d'utiliser le préfixe **sc_** dans chaque nom de configuration système pour pouvoir identifier facilement ce fichier comme étant une configuration système. Cette pratique est particulièrement utile lorsque vous exportez la configuration système.
4. A partir de la liste, sélectionnez une configuration de paramètres de substitution à associer avec la configuration système.
5. Pour chaque banque de données, sélectionnez la configuration de banque de données que vous souhaitez utiliser lors de l'exécution d'un job à l'aide de la configuration système.

Si vous ne mappez pas une configuration de banque de données à une configuration système, le Job Server utilise la configuration de banque de données par défaut au moment de l'exécution.
6. Cliquez sur **OK** pour enregistrer les paramètres de la configuration système.

Informations associées

[Association de configurations des paramètres de substitution avec des configurations système \[page 336\]](#)

10.9.10.2 Creating a system configuration by duplication

Use a duplicate system configuration as a template to create a new configuration.

1. In Designer, select **Tools > System Configurations**.

The *Edit System Configurations* dialog box displays. Each existing system configuration appears in a column, with the system configuration name as the column heading.

2. Select the column of the applicable system configuration and click the [Duplicate Configuration](#) icon  from the tool bar.

The software adds a new column next to the column that you duplicated. The software assigns a name such as `<Existing_SC_Name>_Copy1`.

3. Rename the new system configuration to a name that complies with your naming conventions.

We recommend that you follow a consistent naming convention and use the prefix **sc_** to indicate that the object is a system configuration. Using a naming convention is helpful especially when you export your system configurations.

4. Continue creating the system configuration as in [Création d'une configuration système \[page 146\]](#).

10.9.10.3 Renaming and deleting system configurations

The *Edit System Configurations* dialog box has toolbar icons for renaming and deleting system configurations.

Each existing system configuration appears as a column in the *Edit System Configurations* dialog box. The system configuration name appears as the column heading.

1. Select the system configuration column to rename.

The software highlights the entire system configuration column.

2. Click the *Rename Configuration* icon  on the toolbar menu.

The software highlights the name of the system configuration.

3. Enter a new name for the system configuration based on your naming conventions.

We recommend using **sc_** as a prefix for naming system configurations.

Data Services saves the new name after you click *OK*.

4. To delete a system configuration, select the column of the system configuration to delete.

5. Click the *Delete Configuration* icon  from the toolbar menu

The software removes the system configuration column.

6. Click *OK* to save your changes and close the dialog box.

10.9.10.4 Exportation d'une configuration système

1. Dans la bibliothèque d'objets, sélectionnez l'onglet Banques de données et faites un clic droit sur une banque de données.

2. Sélectionnez ► *Référentiel* ► *Exporter les configurations système* ►.

Il est recommandé d'ajouter le préfixe CS_ à chaque fichier .atl de configuration système exportée pour identifier facilement ce fichier comme étant une configuration système.

3. Cliquez sur *OK*.

11 Formats de fichiers

Ensemble de propriétés qui décrivent la structure de métadonnées d'un fichier plat (ASCII).

Les formats de fichier décrivent la structure de métadonnées d'un fichier spécifique. Un modèle de format de fichier est une description générique qui peut être utilisée pour plusieurs fichiers de données.

Cette section évoque les formats de fichiers, la manière d'utiliser l'éditeur de format de fichier et la manière de créer un format de fichier dans le logiciel.

[Fonctionnalités des formats de fichiers \[page 149\]](#)

Le logiciel offre plusieurs possibilités pour le traitement des fichiers.

[Création d'un nouveau format de fichier plat \[page 154\]](#)

Étapes de création d'un nouveau format de fichier plat.

[Réplication d'un format de fichier pour la création d'un format de fichier \[page 160\]](#)

Étapes pour créer un format de fichier à l'aide de l'option *Répliquer*.

[Création d'un format de fichier à partir d'un schéma de table plate existant \[page 161\]](#)

Étapes pour utiliser un schéma de fichier plat existant dans un flux de données afin de créer rapidement un format de fichier plat au lieu de créer un format de fichier à partir de zéro.

[Créer un fichier source ou cible spécifique \[page 162\]](#)

Étapes pour utiliser un format de fichier comme source ou comme cible dans un flux de données.

[Edition de formats de fichier \[page 163\]](#)

Il est possible de modifier des modèles de format de fichier existants pour correspondre aux modifications apportées dans le format ou la structure d'un fichier.

[Protocoles de transfert de fichier \[page 165\]](#)

Description des deux méthodes de configuration des protocoles de transfert de fichier dans un format de fichier.

11.1 Fonctionnalités des formats de fichiers

Le logiciel offre plusieurs possibilités pour le traitement des fichiers.

[Lecture de plusieurs fichiers en une fois \[page 150\]](#)

Le logiciel peut lire plusieurs fichiers avec le même format à partir d'un seul répertoire à l'aide d'un seul objet source.

[Identification des noms de fichiers sources \[page 151\]](#)

Vous pouvez identifier le fichier source pour chaque ligne dans la cible

[Ignorer les lignes avec des marqueurs spécifiés \[page 151\]](#)

L'éditeur de format de fichier offre une façon d'ignorer les lignes qui contiennent un marqueur (ou des marqueurs) spécifié(s) lors de la lecture des fichiers.

[About special markers \[page 152\]](#)

Configure a file format so that SAP Data Services ignores rows based on specific markers.

[Threads de traitement en parallèle \[page 152\]](#)

Data Services peut utiliser des threads parallèles pour lire et charger les fichiers pour maximiser la performance.

[Erreur de traitement pour les sources de fichier plat \[page 153\]](#)

Vous pouvez configurer l'Éditeur de format de fichier pour identifier les lignes dans des sources de fichiers plats qui contiennent des erreurs.

Rubrique parent : [Formats de fichiers \[page 149\]](#)

Informations associées

[Création d'un nouveau format de fichier plat \[page 154\]](#)

[Réplication d'un format de fichier pour la création d'un format de fichier \[page 160\]](#)

[Création d'un format de fichier à partir d'un schéma de table plate existant \[page 161\]](#)

[Créer un fichier source ou cible spécifique \[page 162\]](#)

[Edition de formats de fichier \[page 163\]](#)

[Protocoles de transfert de fichier \[page 165\]](#)

11.1.1 Lecture de plusieurs fichiers en une fois

Le logiciel peut lire plusieurs fichiers avec le même format à partir d'un seul répertoire à l'aide d'un seul objet source.

1. Ouvrez l'éditeur pour le format de fichier source
2. Sous *Fichier(s) de données* dans l'éditeur de format de fichier, définissez l'*Emplacement* des fichiers sources sur *Local* ou *Job Server*.
3. Définissez le répertoire racine dans *Répertoire racine*.

i Remarque

Si le Job Server se trouve sur un autre ordinateur que Designer, il est impossible d'utiliser la fonction Parcourir pour indiquer le répertoire racine. Vous devez saisir le chemin. Vous pouvez saisir un chemin d'accès absolu ou un chemin d'accès relatif, mais le Job Server doit pouvoir y accéder.

4. Sous *Nom(s) de fichier*, saisissez soit :
 - Une liste des noms de fichier séparés par des virgules, soit
 - Un nom de fichier contenant un caractère générique (* ou ?).Exemple :
1999?????.txt lit les fichiers de l'année 1999
*.txt lit tous les fichiers avec l'extension txt à partir du Répertoire racine spécifié

11.1.2 Identification des noms de fichiers sources

Vous pouvez identifier le fichier source pour chaque ligne dans la cible

Il est possible que vous souhaitiez effectuer cette opération dans les situations suivantes :

- Vous avez spécifié un caractère générique pour lire plusieurs fichiers sources en même temps
- Vous chargez depuis différents fichiers sources sur différentes exécutions

Pour identifier un fichier source, procédez de la façon suivante :

1. Sous *Informations source* dans l'éditeur de format de fichier, définissez *Inclure la colonne Nom du fichier* sur *Oui*. Cette option génère une colonne appelée DI_FILENAME qui contient le nom du fichier source.
2. Dans l'éditeur Query, mappez la colonne DI_FILENAME de l'Entrée de schéma à la Sortie de schéma.
3. Lors de l'exécution du job, la colonne DI_FILENAME pour chaque ligne dans la cible contient le nom du fichier source.

11.1.3 Ignorer les lignes avec des marqueurs spécifiés

L'éditeur de format de fichier offre une façon d'ignorer les lignes qui contiennent un marqueur (ou des marqueurs) spécifié(s) lors de la lecture des fichiers.

Par exemple, il est possible que vous souhaitiez ignorer les marqueurs des lignes de commentaires tels que # et //.

Associés avec cette fonction, deux caractères spéciaux (le point-virgule (;) et la barre oblique inverse (\)) rendent possible la définition de plusieurs marqueurs dans la chaîne pour ignorer les marqueurs de lignes. Utilisez le point-virgule pour délimiter chaque marqueur, et la barre oblique inverse pour indiquer des caractères spéciaux comme des marqueurs (tels que la barre oblique inverse et le point-virgule).

La valeur par défaut de marqueur est une chaîne vide. Lorsque vous spécifiez la valeur par défaut, aucune ligne n'est ignorée.

Pour spécifier des marqueurs pour les lignes à ignorer, procédez de la façon suivante :

1. Ouvrez l'éditeur de format de fichier depuis la Bibliothèque d'objets ou en ouvrant un objet source dans l'espace de travail.
2. Recherchez *Ignorer les marqueurs de ligne* sous la Propriété *Format*.
3. Cliquez dans la zone de texte associée et saisissez une chaîne pour indiquer un ou plusieurs marqueurs représentant les lignes que le logiciel doit ignorer durant la lecture du fichier et/ou la création de métadonnées.

La table suivante donne quelques exemples de marqueur(s) de lignes à ignorer. (Chaque valeur est délimitée par un point-virgule à moins que le point-virgule soit précédé d'une barre oblique inverse.)

Valeur(s) de marqueur	Ligne(s) ignorée(s)
	Aucune (valeur par défaut)
abc	Toutes celles qui commencent par la chaîne abc

Valeur(s) de marqueur	Ligne(s) ignorée(s)
abc;def;hi	Toutes celles qui commencent par abc, def ou hi
abc;\;	Toutes celles qui commencent par abc ou ;
abc;\;\;	Toutes celles qui commencent par abc, \ ou ;

11.1.4 About special markers

Configure a file format so that SAP Data Services ignores rows based on specific markers.

Your data may contain rows prefaced with a marker to indicate a row that contains informational data and not record data. Or your data may have rows that start with a specific string that you want to ignore.

Specify special markers to ignore in either the [File Format Editor](#) or the [Source File Editor](#). To add more than one marker to ignore, enter markers delimited by a semicolon (;). When a marker is also a special character, use the backslash (\) escape character.

The following table provides some examples of how to enter special markers in either the [File Format Editor](#) or the [Source File Editor](#).

Marker value or values	Rows ignored
	None.
	Blank is the default value. Data Services processes all rows in your data.
abc	Rows that begin with the string abc
abc;def;hi	Rows that begin with abc, def and hi
abc;\;	Rows that begin with abc and ;
	The backslash character acts as an escape character for the semicolon.
abc;\;\;	Rows that begin with abc and \ and ;
	The backslash character acts as an escape character for the backslash and the semicolon.

11.1.5 Threads de traitement en parallèle

Data Services peut utiliser des threads parallèles pour lire et charger les fichiers pour maximiser la performance.

Pour spécifier des threads parallèles pour traiter les formats de fichier :

1. Ouvrez l'éditeur de format de fichier de l'une des manières suivantes :
 - Dans l'onglet Formats de la bibliothèque d'objets, faites un clic droit sur un nom de format de fichier et cliquez sur [Modifier](#).

- Dans l'espace de travail, double-cliquez sur l'objet source ou cible.
- 2. Trouvez les [Threads de traitement en parallèle](#) sous la Propriété [Général](#).
- 3. Indiquez le nombre de threads pour lire ou charger ce format de fichier.
Par exemple, si vous avez 4 processeurs sur l'ordinateur du Job Server, saisissez le nombre 4 dans la case [Threads de traitement en parallèle](#).

11.1.6 Erreur de traitement pour les sources de fichier plat

Vous pouvez configurer l'Éditeur de format de fichier pour identifier les lignes dans des sources de fichiers plats qui contiennent des erreurs.

Lors de l'exécution du job, le logiciel traite les lignes des sources de fichiers plats une par une. Vous pouvez afficher des informations relatives aux erreurs suivantes :

Type d'erreur	Exemple
Erreurs de conversion de type de données	Un champ peut être défini dans l'Éditeur de format de fichier comme contenant un type de données de nombre entier, alors que les données rencontrées sont de type varchar.
Erreurs de format de lignes	En cas d'un fichier de largeur fixe, le logiciel identifie une ligne qui ne correspond pas à la valeur de largeur attendue.

Ces propriétés d'erreur de traitement s'appliquent uniquement aux sources de fichiers plats.

[Options d'erreur de traitement \[page 153\]](#)

Dans l'Éditeur de format de fichier, l'ensemble de propriétés [Erreur de traitement](#) permet de choisir si le logiciel doit effectuer les actions de traitement d'erreur.

[A propos du fichier d'erreur \[page 154\]](#)

S'il est activé, le fichier d'erreur inclut les informations relatives aux erreurs.

11.1.6.1 Options d'erreur de traitement

Dans l'Éditeur de format de fichier, l'ensemble de propriétés [Erreur de traitement](#) permet de choisir si le logiciel doit effectuer les actions de traitement d'erreur.

Vous pouvez laisser le logiciel effectuer les tâches suivantes :

- vérifier l'un ou l'autre des deux types d'erreur de la source de fichier plat
- écrire la/les ligne(s) non valide(s) dans un fichier d'erreur spécifique
- arrêter le traitement du fichier source après avoir atteint un nombre spécifié de lignes non valides
- enregistrer les avertissements de conversion des types de données ou de format de lignes dans le journal des erreurs ; si c'est le cas, il est possible de limiter le nombre d'avertissements à enregistrer sans arrêter le job

11.1.6.2 A propos du fichier d'erreur

S'il est activé, le fichier d'erreur inclut les informations relatives aux erreurs.

Le format est un fichier texte délimité par un point-virgule. Il est possible d'avoir plusieurs fichiers d'entrée sources pour le fichier d'erreur. Le fichier reste sur le même ordinateur que le Job Server.

Les entrées dans un fichier d'erreur a la syntaxe suivante :

```
source file path and name; row number in source file; Data Services error;  
column number where the error occurred; all columns from the invalid row
```

L'entrée suivante illustre une erreur de format de lignes :

```
d:/acl_work/in_test.txt;2;-80104: 1-3-A column delimiter was seen after column  
number <3> for row number <2> in file <d:/acl_work/in_test.txt>. The total  
number of columns defined is <3>, so a row delimiter should be seen after column  
number <3>. Please check the file for bad data, or redefine the input schema for  
the file by editing the file format in the UI.;3;defg;234;def
```

où 3 indique une erreur qui s'est produite après la troisième colonne, et defg;234;def sont les trois colonnes de données à partir de la ligne non valide.

i Remarque

Si vous définissez l'option *Threads de traitement en parallèle* du format de fichier sur n'importe quelle valeur supérieure à 0 ou sur {aucun}, le numéro de la ligne dans la valeur du fichier source est -1.

11.2 Création d'un nouveau format de fichier plat

Étapes de création d'un nouveau format de fichier plat.

1. Cliquez avec le bouton droit sur *Fichiers plats* dans l'onglet *Formats* de la bibliothèque d'objets et sélectionnez *Nouveau*.

L'*Éditeur de format de fichier* s'ouvre.

2. Dans la colonne propriétés-valeurs à gauche, sélectionnez le type de fichier plat dans la liste déroulante *Type*.

i Remarque

Le logiciel propose les options restantes en fonction du type choisi.

3. Saisissez un *nom* descriptif pour le modèle de format de fichier.
Après avoir enregistré le modèle de format de fichier, il est impossible de modifier le nom mais vous pouvez ajouter ou modifier une description.
4. Renseignez le reste des options dans la colonne propriétés le cas échéant (voir le *Guide de référence* pour les descriptions d'options).
5. Le cas échéant, renseignez les informations dans la section attributs de colonne en haut à droite de la fenêtre de l'*Éditeur de format de fichier*.

6. Cliquez sur [Enregistrer et fermer](#) pour enregistrer le modèle de format de fichier et fermer l'éditeur de format de fichier.

[Définition des attributs de colonne lors de la création d'un format de fichier \[page 155\]](#)

Modifiez et définissez les colonnes ou les champs d'un fichier source dans la section attributs de colonne.

[Formats des nombres \[page 156\]](#)

Le point (.) et la virgule (,) sont les deux formats les plus utilisés pour déterminer les séparateurs de décimales et de milliers pour les types de données numériques.

[Formats de date au niveau du champ \[page 157\]](#)

Il est possible de spécifier un format de date au niveau du champ pour remplacer les formats de date, d'heure ou de date-heure définis dans la zone Propriétés-Valeurs.

[Use a sample file to define column attributes \[page 158\]](#)

Designate a file location and name in the *File Format Editor* to display sample data on which to base column attributes.

[Modélisation d'un format de fichier sur un exemple de fichier. \[page 159\]](#)

Présentation des tâches : [Formats de fichiers \[page 149\]](#)

Informations associées

[Fonctionnalités des formats de fichiers \[page 149\]](#)

[Réplication d'un format de fichier pour la création d'un format de fichier \[page 160\]](#)

[Création d'un format de fichier à partir d'un schéma de table plate existant \[page 161\]](#)

[Créer un fichier source ou cible spécifique \[page 162\]](#)

[Edition de formats de fichier \[page 163\]](#)

[Protocoles de transfert de fichier \[page 165\]](#)

[Définition des attributs de colonne lors de la création d'un format de fichier \[page 155\]](#)

[Transferts personnalisés de fichiers \[page 166\]](#)

[Objets d'emplacement de fichier \[page 183\]](#)

11.2.1 Définition des attributs de colonne lors de la création d'un format de fichier

Modifiez et définissez les colonnes ou les champs d'un fichier source dans la section attributs de colonne.

Attention

Les formats spécifiques au champ que vous définissez dans la section attributs de colonne écrasent le format par défaut défini dans la colonne propriétés-valeurs à gauche.

Lors de la création d'un format de fichier pour certains fichiers source, il est possible de spécifier la structure des colonnes dans la zone de travail Attributs de colonne (volet en haut à droite) de l'[Éditeur de format de fichier](#). Il est possible de modéliser les attributs de fichier à l'aide d'un exemple de fichier.

1. Cliquez sur [Champ1](#) sous la colonne [Nom de champ](#) et saisissez un nom de champ.
2. Cliquez sur la cellule [Type de données](#) et choisissez un type de données dans la liste déroulante.
3. Poursuivez la configuration du champ en renseignant les autres attributs applicables dans la ligne.

Informations supplémentaires concernant les colonnes :

- Les colonnes [Échelle](#) et [Précision](#) sont destinées à des types de données décimales et numériques.
 - Pour le [Type de contenu](#), si vous ajoutez une colonne lors de la création d'un format, le logiciel peut renseigner le type de contenu en fonction du nom du champ. Si un type de contenu approprié n'est pas disponible, le logiciel laisse le [type de contenu](#) vide par défaut.
 - Vous pouvez également saisir les informations dans le champ [Format](#) pour les types de données appropriés (type de données horodateur par exemple). Ces informations remplacent le format par défaut défini dans la zone Propriétés-Valeurs pour ce type de données.
4. Ajoutez un nouveau champ en cliquant sur la cellule vide sous la première ligne de champ.
 5. Cliquez sur [Enregistrer et fermer](#) pour enregistrer le modèle de format de fichier et fermer l'éditeur de format de fichier.

i Remarque

Vous n'avez pas besoin d'indiquer les colonnes pour les fichiers utilisés comme des cibles. Si vous indiquez des colonnes et qu'elles ne correspondent pas au schéma de sortie de la transformation précédente, le logiciel écrit dans le fichier cible à l'aide du schéma de sortie de la transformation.

i Remarque

Pour le type de données décimale ou réelle, si vous indiquez seulement un format de colonne source et que les noms de colonnes et les types de données dans le schéma cible ne correspondent pas à ceux dans le schéma source, le logiciel ne peut pas utiliser le format de colonne source spécifié. À la place, il utilise par défaut le format utilisé par la page de code sur l'ordinateur où le Job Server est installé.

Informations associées

[Création d'un nouveau format de fichier plat \[page 154\]](#)

11.2.2 Formats des nombres

Le point (.) et la virgule (,) sont les deux formats les plus utilisés pour déterminer les séparateurs de décimales et de milliers pour les types de données numériques.

Lors de la mise en forme des fichiers dans le logiciel, les types de données dans lesquels ces symboles peuvent être utilisés incluent Décimal, Numérique, Flottant et Double. Il est possible d'utiliser n'importe quel symbole pour l'indicateur des milliers, et n'importe quel symbole pour le séparateur de décimales. Exemple : 2,098.65 ou 2.089,65.

	Field Name	Data Type	Field Size	Precision	Scale	Format
	Sales Office	int				{none}
	Region	varchar	2			{none}
	Date open	date				#,##0.0
	Country	varchar	7			#,##0,0
*						

Format	Description
{aucun}	Le logiciel s'attend à ce que le nombre contienne uniquement le séparateur de décimales. La lecture des données numériques et de ce séparateur de décimales est déterminée par les paramètres régionaux du Job Server de Data Service. La virgule (,) est le séparateur de décimales lorsque les Paramètres régionaux de Data Service sont définis sur un pays qui utilise des virgules (par exemple, l'Allemagne ou la France). Le point (.) est le séparateur de décimales lorsque les Paramètres régionaux sont définis sur un pays qui utilise des points (par exemple, les États-Unis, l'Inde et le RU). Dans ce format, le logiciel renvoie une erreur si un nombre contient un séparateur des milliers. Lorsque le logiciel écrit les données, il utilise uniquement le séparateur de décimales des paramètres régionaux du Job Server. Il n'utilise pas de séparateur de milliers.
#,##0.0	Le logiciel s'attend à ce que le séparateur de décimales d'un nombre soit un point (.) et que le séparateur des milliers soit une virgule (,). Lorsque le logiciel charge les données dans un fichier plat, il utilise une virgule (,) comme séparateur des milliers et un point (.) comme séparateur de décimales.
#,##0,0	Le logiciel s'attend à ce que le séparateur de décimales d'un nombre soit une virgule (,) et que le séparateur des milliers soit un point (,). Lorsque le logiciel charge les données dans un fichier plat, il utilise un point (.) comme séparateur des milliers et une virgule (,) comme séparateur de décimales.

Les signes décimaux de début et de fin sont également pris en charge. Exemple : +12,000.00 ou 32.32-.

11.2.3 Formats de date au niveau du champ

Il est possible de spécifier un format de date au niveau du champ pour remplacer les formats de date, d'heure ou de date-heure définis dans la zone Propriétés-Valeurs.

Par exemple, lorsque le [Type de données](#) est défini sur Date, il est possible de modifier la valeur dans le champ [Format](#) correspondant par un format de date différent tel que :

- aaaa.mm.jj
- mm/jj/aa
- jj.mm.aa

11.2.4 Use a sample file to define column attributes

Designate a file location and name in the *File Format Editor* to display sample data on which to base column attributes.

Keep in mind the following information when you specify a file in the *File Format Editor*:

Purpose	File format settings	Sample file location
You do not intend to use the file format for job execution.	- Set <i>Location</i> to <i>Local</i>. - Browse to set the <i>Root directory</i>. - Browse to set the <i>File name(s)</i> for the sample file.	On the same computer as SAP Data Services.
You intend to use the file format for job execution.	- Set <i>Location</i> to <i>Job Server</i>. i Remarque Data Services disables the <i>Browse</i> icon. - Type the root directory for <i>Root directory</i>. - Type the file name in <i>File name(s)</i> Use either an absolute path or a relative path to the file, as long as the Job Server can access it. ❖ Exemple For example, a path on UNIX might be <code>/usr/data/abc.txt</code>. A path on Windows might be <code>C:\DATA\abc.txt</code>.	On the same computer as the Job Server.

Use Telnet to eliminate typing errors

When you manually enter the root directory and file name, you open yourself to errors because of typing mistakes. Also, capitalization is different between Windows and Unix environments.

For Windows, files are not case-sensitive, but they are in UNIX environments. For example, `abc.txt` and `aBc.txt` are different files in the same UNIX directory.

You cannot use Windows Explorer to determine the exact file location on Windows. To reduce the risk of typing errors, establish a connection between the Telnet protocol to the Job Server (UNIX or Windows) computer. Find the full path name of the sample file. Then, copy and paste the path name from the Telnet application directly into the *Root directory* text box in the *File Format Editor*.

11.2.5 Modélisation d'un format de fichier sur un exemple de fichier.

1. Depuis l'onglet *Formats* dans la bibliothèque d'objets locale, créez un nouveau modèle de format de fichier ou modifiez un modèle de format de fichier plat existant.
2. Sous *Fichier(s) de données* :
 - Si l'exemple de fichier se trouve sur le même ordinateur que Designer Data Services, définissez *Emplacement* sur *Local*. Naviguez pour définir les options *Répertoire racine* et *Fichier(s)* pour spécifier l'exemple de fichier.

i Remarque

Lors de la conception, il est possible de spécifier un fichier qui se trouve sur l'ordinateur où Designer s'exécute ou sur l'ordinateur où le Job Server s'exécute. Indiquez l'emplacement du fichier dans la propriété *Emplacement*. Lors de l'exécution, vous devez spécifier un fichier qui se trouve sur l'ordinateur du Job Server qui exécute le job.

- Si l'exemple de fichier se trouve sur le même ordinateur que le Job Server actuel, définissez *Emplacement* sur *Job Server*. Saisissez les options *Répertoire racine* et *Fichier(s)* pour spécifier l'exemple de fichier. Lorsque vous sélectionnez *Job Server*, l'icône *Parcourir* est désactivée ; vous devez donc saisir le chemin d'accès au fichier. Vous pouvez saisir un chemin d'accès absolu ou un chemin d'accès relatif, mais le Job Server doit pouvoir y accéder. Par exemple, un chemin sur UNIX devrait être `/usr/data/abc.txt`. Un chemin sur Windows devrait être `C:\DATA\abc.txt`.

i Remarque

Dans le système d'exploitation Windows, les fichiers ne sont pas sensibles à la casse ; cependant, les noms de fichier sont sensibles à la casse dans l'environnement UNIX. (Par exemple, `abc.txt` et `aBc.txt` sont deux fichiers différents dans le même répertoire UNIX.)

Pour réduire le risque de saisir des erreurs, il est possible d'effectuer un telnet (établir une connexion au protocole Telnet) sur l'ordinateur du Job Server (UNIX ou Windows) et de trouver le nom complet du chemin du fichier que vous souhaitez utiliser. Ensuite, copiez-collez le nom du chemin depuis l'application de telnet directement dans la zone de texte *Répertoire racine* dans l'éditeur de format de fichier. Il est impossible d'utiliser Windows Explorer pour déterminer l'emplacement exact du fichier sur Windows.

3. Si le type de fichier est délimité, définissez le séparateur de colonne approprié pour l'exemple de fichier. Il est possible de choisir dans la liste déroulante ou de spécifier des séparateurs Unicode en saisissant directement le code de caractère Unicode sous la forme `/XXXX` où XXXX est un code de caractère Unicode décimal. Par exemple, `/44` est le caractère Unicode pour la virgule (,).
4. Sous *Entrée/Sortie*, définissez *Ignorer l'en-tête de la ligne* sur *Oui* si vous souhaitez utiliser la première ligne dans le fichier pour désigner les noms de champ.

L'éditeur de format de fichier affiche les noms de colonne dans la zone *Aperçu des données* et crée automatiquement la structure des métadonnées.

5. Facultatif. Modifiez la structure des métadonnées si nécessaires dans la zone des attributs de colonne en haut à droite de l'*Éditeur de format de fichier*.

Pour les fichiers délimités et de largeur fixe, il est possible de modifier la structure des métadonnées dans la zone de travail Attributs de la colonne :

- a. Faites un clic droit pour insérer ou supprimer des champs.
- b. Renommez les champs.
- c. Définissez les types de données
- d. Saisissez des longueurs de champ pour les types de données *Blob* et *Varchar*.
- e. Saisissez les informations d'échelle et de précision pour les types de données *Numérique* et *Décimale*.
- f. Saisissez les informations du champ *Format* pour les types de données appropriés si vous le souhaitez. Ces informations de format remplacent le format par défaut défini dans la zone Valeurs dans les Propriétés pour ce type de données.
- g. Saisissez les informations de *Type de contenu*. Vous n'avez pas besoin d'indiquer les colonnes pour les fichiers utilisés comme des cibles. Si vous avez ajouté une colonne lors de la création d'un nouveau format, le type de contenu peut automatiquement se renseigner en fonction du nom de champ. Si un type de contenu approprié ne peut pas être automatiquement renseigné, il est vide par défaut.

Pour les fichiers de largeur fixe, il est possible de modifier la structure des métadonnées dans la zone de travail Attributs de la colonne :

- a. Cliquez pour sélectionner et mettre en surbrillance les colonnes
- b. Faites un clic droit pour insérer ou supprimer des champs.

i Remarque

La zone Aperçu des données ne peut pas afficher les données Blob.

6. Cliquez sur *Enregistrer et fermer* pour enregistrer le modèle de format de fichier et fermer l'éditeur de format de fichier.

11.3 Réplication d'un format de fichier pour la création d'un format de fichier

Étapes pour créer un format de fichier à l'aide de l'option *Répliquer*.

1. Dans l'onglet *Formats* de la bibliothèque d'objets, cliquez avec le bouton droit sur un format de fichier existant et sélectionnez *Répliquer* dans le menu déroulant.

L'*Éditeur de format de fichier* s'ouvre et affiche le schéma du format de fichier copié.

2. Saisissez un nouveau nom unique dans l'option *Nom*.

i Remarque

Vous devez saisir un nouveau nom pour le fichier répliqué. Le logiciel ne vous autorise pas à enregistrer le fichier répliqué avec le même nom que celui d'origine (ou que celui d'un autre objet de format de fichier existant). De plus, c'est votre seule chance de modifier la valeur de la propriété *Nom*. Une fois enregistré, il n'est plus possible de modifier à nouveau le nom.

3. Modifiez les autres propriétés le cas échéant.

Recherchez les propriétés disponibles lorsque l'éditeur de format de fichier est en mode source ou en mode cible.

4. Pour enregistrer et afficher le nouveau schéma de format de fichier, cliquez sur [Enregistrer](#).

Pour terminer le processus de réplication (même après avoir modifié le nom et cliqué sur Enregistrer), cliquez sur [Annuler](#) ou appuyez sur le bouton  de votre clavier.

5. Cliquez sur [Enregistrer et fermer](#).

Présentation des tâches : [Formats de fichiers \[page 149\]](#)

Informations associées

[Fonctionnalités des formats de fichiers \[page 149\]](#)

[Création d'un nouveau format de fichier plat \[page 154\]](#)

[Création d'un format de fichier à partir d'un schéma de table plate existant \[page 161\]](#)

[Créer un fichier source ou cible spécifique \[page 162\]](#)

[Edition de formats de fichier \[page 163\]](#)

[Protocoles de transfert de fichier \[page 165\]](#)

[Formats de fichiers \[page 149\]](#)

11.4 Création d'un format de fichier à partir d'un schéma de table plate existant

Étapes pour utiliser un schéma de fichier plat existant dans un flux de données afin de créer rapidement un format de fichier plat au lieu de créer un format de fichier à partir de zéro.

Avant de suivre la procédure ci-dessous, ouvrez le flux de données applicable contenant l'objet avec le schéma à copier.

1. Cliquez sur le nom de l'objet.

L'éditeur d'objet s'ouvre en affichant le schéma (par exemple [Entrée de schéma](#) situé en haut à gauche de l'éditeur)

2. Cliquez avec le bouton droit sur le nom du schéma et sélectionnez [Créer format de fichier](#) dans le menu déroulant.

L'[Éditeur de format de fichier](#) s'ouvre et affiche les paramètres actuels du schéma.

3. Saisissez un nom pour le nouveau schéma dans l'option [Nom](#).
4. Définissez les propriétés applicables pour le nouveau schéma, le cas échéant, et cliquez sur [Enregistrer et fermer](#).

Le logiciel enregistre le nouveau modèle de format de fichier dans le référentiel. Vous pouvez y accéder dans l'onglet [Formats](#) de la bibliothèque d'objets dans le dossier [Fichiers plats](#).

Présentation des tâches : [Formats de fichiers \[page 149\]](#)

Informations associées

[Fonctionnalités des formats de fichiers \[page 149\]](#)

[Création d'un nouveau format de fichier plat \[page 154\]](#)

[Réplication d'un format de fichier pour la création d'un format de fichier \[page 160\]](#)

[Créer un fichier source ou cible spécifique \[page 162\]](#)

[Edition de formats de fichier \[page 163\]](#)

[Protocoles de transfert de fichier \[page 165\]](#)

11.5 Créer un fichier source ou cible spécifique

Étapes pour utiliser un format de fichier comme source ou comme cible dans un flux de données.

Créez un flux de données dans l'espace de travail et choisissez le fichier plat que vous utiliserez comme source ou cible.

1. Cliquez sur un nom de format de fichier plat dans le dossier *Fichiers plats* de l'onglet *Formats* dans la bibliothèque d'objets locale.
2. Faites glisser le modèle de format de fichier vers l'espace de travail du flux de données.

Un menu pop-up s'affiche.

3. Sélectionnez *Créer la source* pour définir un format de fichier source, ou sélectionnez *Créer la cible* pour définir un format de fichier cible.

Le fichier plat source ou cible apparaît dans votre espace de travail.

4. Cliquez sur le nom de l'objet de format de fichier dans l'espace de travail pour ouvrir l'éditeur.

L'éditeur d'objet s'ouvre. Les propriétés disponibles varient si vous avez choisi *Créer la source* ou *Créer la cible*.

5. Renseignez les propriétés *Nom(s) de fichier(s)* et *Emplacement* dans *Fichier(s) de données*.

i Remarque

Il est possible d'utiliser des variables comme des noms de fichier.

6. Connectez l'objet de format de fichier à l'objets applicable dans le flux de données.

Présentation des tâches : [Formats de fichiers \[page 149\]](#)

Informations associées

[Fonctionnalités des formats de fichiers \[page 149\]](#)

[Création d'un nouveau format de fichier plat \[page 154\]](#)

[Réplication d'un format de fichier pour la création d'un format de fichier \[page 160\]](#)

[Création d'un format de fichier à partir d'un schéma de table plate existant \[page 161\]](#)

[Edition de formats de fichier \[page 163\]](#)

[Protocoles de transfert de fichier \[page 165\]](#)

[Définir des noms de fichier lors de l'exécution à l'aide de variables \[page 327\]](#)

11.6 Edition de formats de fichier

Il est possible de modifier des modèles de format de fichier existants pour correspondre aux modifications apportées dans le format ou la structure d'un fichier.

Il est impossible de modifier le nom d'un modèle de format de fichier.

Par exemple, si vous avez un champ de date dans un fichier source ou cible dont le format est mm/jj/aa et que les données pour ce champ passent au format jj-mm-aa à cause de changements dans le programme qui génère le fichier source, vous pouvez modifier le modèle de format de fichier correspondant et les informations sur le format de la date.

Pour les formats de fichier source ou cible spécifiques, il est possible de modifier les propriétés qui définissent uniquement cette source ou cette cible telles que le nom et l'emplacement du fichier.

⚠ Attention

Si le modèle est utilisé dans d'autres jobs (l'utilisation est supérieure à 0), les modifications que vous apportez au modèle sont également apportées aux fichiers qui utilisent le modèle.

Pour modifier un modèle de format de fichier, procédez de la façon suivante :

1. Dans l'onglet [Formats](#) de la bibliothèque d'objets, cliquez deux fois sur un format de fichier plat existant (ou faites un clic droit et sélectionnez [Modifier](#)).
L'éditeur de format de fichier s'ouvre avec les valeurs du format existant.
2. Modifiez les valeurs selon vos besoins.
Recherchez les propriétés disponibles lorsque l'éditeur de format de fichier est en mode source ou en mode cible.
3. Cliquez sur [Enregistrer](#).

Présentation des tâches : [Formats de fichiers \[page 149\]](#)

Informations associées

[Fonctionnalités des formats de fichiers \[page 149\]](#)

[Création d'un nouveau format de fichier plat \[page 154\]](#)

[Réplication d'un format de fichier pour la création d'un format de fichier \[page 160\]](#)

[Création d'un format de fichier à partir d'un schéma de table plate existant \[page 161\]](#)

[Créer un fichier source ou cible spécifique \[page 162\]](#)

11.6.1 Modifier un fichier source ou cible

1. Dans l'espace de travail, cliquez sur le nom d'un fichier source ou cible.

L'Éditeur de format de fichier s'ouvre : L'*Éditeur de fichiers source* ou l'*Éditeur de fichiers cible* s'ouvre, affichant les propriétés du fichier source ou cible sélectionné.

2. Modifiez les propriétés souhaitées.

Il existe des propriétés communes aux modes source et cible. Toutefois, des options sont spécifiques à chaque mode. Retrouvez les descriptions complètes des options dans le *Guide de référence*.

! Restriction

Toutes les modifications de valeurs que vous apportez dans les modes source ou cible remplacent celles du format de fichier d'origine.

Pour modifier les propriétés qui ne sont pas disponibles en mode source ou cible, vous devez modifier le modèle de format de fichier du fichier.

3. Cliquez sur *Enregistrer*.

11.6.2 Modification de plusieurs propriétés de colonne

Étapes pour définir les mêmes propriétés de colonne, telles que le type de données et le type de contenu, pour plusieurs colonnes d'un format de fichier dans la zone des attributs de colonne de l'*Éditeur de format de fichier*.

Utilisez ces étapes lors de la création d'un nouveau format de fichier ou de la modification d'un format de fichier existant.

i Remarque

Pour plus de clarté, nous faisons référence à chaque champ comme colonne séparée. Par exemple, le champ Prénom et Nom sont deux colonnes. Nous ne faisons pas référence aux colonnes à proprement parler (*Nom de champ*, *Type de données*, etc.) de la zone attributs de la colonne.

1. Cliquez avec le bouton droit sur un format de fichier existant sous le dossier *Fichiers plats* de l'onglet *Formats* de la bibliothèque d'objets et sélectionnez *Modifier*.
L'*Éditeur de format de fichier* s'ouvre.
2. Dans la zone des attributs de colonne (volet en haut à droite), sélectionnez les colonnes que vous souhaitez modifier.
 - Pour sélectionner des colonnes consécutives, sélectionnez la première colonne et appuyez sur la touche **Maj** puis sélectionnez la dernière colonne.
 - Pour sélectionner des colonnes non consécutives, appuyez sur la touche **Contrôle** et cliquez sur chaque colonne applicable.

3. Faites un clic droit et sélectionnez *Propriétés*.
La fenêtre *Propriétés de colonnes multiples* s'ouvre. L'option *Nom* affiche le nom de chaque colonne sélectionnée et ne peut être modifié.
4. Modifiez les options applicables, *Type de données* ou *Type de contenu* par exemple, puis cliquez sur *OK*.
Les valeurs que vous définissez dans la boîte de dialogue *Propriétés de colonnes multiples* apparaissent dans l'*Éditeur de format de fichier* comme les valeurs de la zone attributs de colonne.

Informations associées

[Formats de fichiers \[page 149\]](#)

11.7 Protocoles de transfert de fichier

Description des deux méthodes de configuration des protocoles de transfert de fichier dans un format de fichier.

Pour utiliser les paramètres du protocole de transfert de fichiers, il est possible de choisir de définir un programme de transfert personnalisé ou de créer un objet d'emplacement de fichier. Ces deux méthodes doivent être définies avec un format de fichier spécifique dans l'Éditeur de format de fichier. La table suivante décrit les méthodes de protocole de transfert de fichiers.

Méthode de transfert de fichier	Description
Programme de transfert personnalisé	Uniquement pour les types de fichiers de largeur fixe et délimités. Lit et charge les fichiers à l'aide des paramètres du programme de transfert personnalisé dans le groupe d'options du <i>Transfert personnalisé</i> dans l' <i>Éditeur de format de fichier</i> .
Objet d'emplacement de fichier	Pour divers types de fichiers. Contient des informations de protocole de transfert de fichiers pour les protocoles tels que FTP, SFTP ou SCP et inclut les informations de serveurs locaux et à distance. Créez un objet d'emplacement de fichier et associez-le à un format de fichier dans l' <i>Éditeur de format de fichier</i> .

Rubrique parent : [Formats de fichiers \[page 149\]](#)

Informations associées

[Fonctionnalités des formats de fichiers \[page 149\]](#)

[Création d'un nouveau format de fichier plat \[page 154\]](#)

[Réplication d'un format de fichier pour la création d'un format de fichier \[page 160\]](#)

[Création d'un format de fichier à partir d'un schéma de table plate existant \[page 161\]](#)

[Créer un fichier source ou cible spécifique \[page 162\]](#)

[Edition de formats de fichier \[page 163\]](#)

[Transferts personnalisés de fichiers \[page 166\]](#)

[Objets d'emplacement de fichier \[page 183\]](#)

11.7.1 Transferts personnalisés de fichiers

Le logiciel peut lire et charger les fichiers à l'aide d'un programme de transfert de fichier tiers pour les fichiers plats.

Définissez un programme de transfert de fichiers personnalisé lorsque vous créez ou modifiez un format de fichier pour un fichier plat délimité ou de largeur fixe. Il est possible d'utiliser des programmes de transfert tiers (personnalisés) pour :

- Incorporer des applications de transfert de fichiers de standards de société comme une partie de l'exécution d'un job de logiciel.
- Fournir une grande flexibilité et une sécurité élevée pour les fichiers transférés par un pare-feu.

L'option de programme de transfert personnalisé permet de spécifier :

- Un programme de transfert personnalisé (appelé durant l'exécution d'un job)
- Des arguments supplémentaires, en fonction de ce qui est disponible dans votre programme, tels que :
 - Des données de connexion
 - Des mécanismes de cryptage/décryptage
 - Des mécanismes de compression

Informations associées

[Utilisation d'un programme de transfert personnalisé \[page 171\]](#)

11.7.1.1 Variables de système de transfert personnalisé pour les fichiers plats

L'utilisation de variables comme arguments de programme de transfert personnalisé, permet de collecter les informations de connexion saisies dans le logiciel et d'utiliser ces données au moment de l'exécution avec votre programme de transfert personnalisé.

i Remarque

Si vous avez détaillé des informations telles qu'un serveur local ou à distance à ajouter aux variables de transfert personnalisé, envisagez plutôt de créer un objet d'emplacement de fichier. La boîte de dialogue Emplacement du fichier vous fournit les options de saisie des emplacements de serveur.

Lorsque vous définissez des options de transfert personnalisé pour des cibles et des sources de fichiers externes, certaines informations de transfert, comme le nom du serveur distant à partir duquel ou vers lequel le fichier est transféré, peuvent avoir besoin d'être saisies littéralement comme argument de programme de transfert. Vous pouvez saisir d'autres informations à l'aide des variables de système suivantes :

Données saisies pour :	Sont remplacées par cette variable si cette option a été définie dans le champ Argument
Nom d'utilisateur	\$AW_USER
Mot de passe	\$AW_PASSWORD
Répertoire local	\$AW_LOCAL_DIR
Fichier(s)	\$AW_FILE_NAME

Les options de transfert personnalisé suivantes utilisent un fichier de commande Windows (Myftp.cmd) avec cinq arguments. Les arguments 1 à 4 sont des variables de système :

- Les variables Utilisateur et Mot de passe sont pour le serveur externe
- La variable Répertoire local est pour l'emplacement dans lequel les fichiers transférés sont stockés dans le logiciel
- La variable Nom du fichier est pour les noms des fichiers à transférer

L'argument 5 fournit le nom littéral du serveur externe.

i Remarque

Si vous ne spécifiez aucun fichier de sortie standard (comme `ftp.out` dans l'exemple ci-dessous), le logiciel écrit la sortie standard dans le journal de suivi du job.

```
@echo off
set USER=%1
set PASSWORD=%2
set LOCAL_DIR=%3
set FILE_NAME=%4
set LITERAL_HOST_NAME=%5
set INP_FILE=ftp.inp
echo %USER%>%INP_FILE%
echo %PASSWORD%>>%INP_FILE%
echo lcd %LOCAL_DIR%>>%INP_FILE%
echo get %FILE_NAME%>>%INP_FILE%
echo bye>>%INP_FILE%
ftp -s%INPT_FILE% %LITERAL_HOST_NAME%>ftp.out
```

Informations associées

[Objets d'emplacement de fichier \[page 183\]](#)

11.7.1.2 Options de transfert personnalisé pour les fichiers plats

i Remarque

Si vous avez détaillé des informations comme les noms d'utilisateurs et les mots de passe à ajouter aux variables de transfert personnalisé, envisagez plutôt de créer un objet d'emplacement de fichier. La boîte de dialogue Emplacement de fichier vous fournit les options de saisie des emplacements de serveur.

Dans les options de programme de transfert personnalisé, seule l'option *Exécutable de programme* est obligatoire.

La saisie des valeurs *Nom d'utilisateur*, *Mot de passe* et *Arguments* est facultative. Ces options sont fournies pour que vous indiquiez des arguments qui peuvent être traités par le programme de transfert personnalisé (comme les données de connexion).

Il est également possible d'utiliser *Arguments* pour activer ou désactiver les fonctionnalités intégrées dans le programme telles que les mécanismes de cryptage/décryptage et de compression. Par exemple, vous pouvez concevoir votre programme de transfert pour que, lorsque vous saisissez `-sSecureTransportOn` ou `-CCompressionYES`, la sécurité ou la compression soit activée.

i Remarque

Les arguments disponibles dépendent de ce qui est inclus dans le programme de transfert personnalisé. Consultez la documentation du programme de transfert personnalisé pour une liste des arguments valides.

Vous pouvez utiliser la case *Arguments* pour saisir un nom d'utilisateur et un mot de passe. Cependant, le logiciel fournit également des cases *Nom d'utilisateur* et *Mot de passe* distinctes. En saisissant les variables `<AW_USER>` et `<AW_PASSWORD>` comme *Arguments* puis en utilisant les cases *Utilisateur* et *Mot de passe* pour saisir des chaînes littérales, ces cases supplémentaires sont utiles de deux manières :

- Vous pouvez mettre plus facilement à jour les utilisateurs et les mots de passe dans le logiciel lorsque vous configurez le logiciel pour utiliser un programme de transfert, et lorsque vous exportez par la suite le job. Par exemple, lorsque vous migrez le job vers un autre environnement, il est possible que vous souhaitiez changer les informations de connexion sans faire défiler les autres arguments.
- Vous pouvez utiliser les propriétés de masque et de cryptage de la case *Mot de passe*. Les données saisies dans la case *Mot de passe* sont masquées dans les fichiers journaux et sur l'écran, stockées dans le référentiel et cryptées par Data Services.

i Remarque

Le logiciel envoie un mot de passe au programme de transfert personnalisé en texte en clair. Si vous n'autorisez pas que les mots de passe en clair soient exposés comme des arguments dans les exécutables de ligne de commande, configurez le programme personnalisé pour :

- récupérer le mot de passe depuis un emplacement approuvé ;
- hériter des privilèges de sécurité à partir du programme appelant (dans ce cas, le logiciel).

Informations associées

[Objets d'emplacement de fichier \[page 183\]](#)

11.7.1.3 Options de transfert personnalisé

L'option de transfert personnalisé permet d'utiliser un programme tiers pour transférer les sources et les cibles de fichiers plats.

Il est possible de configurer le programme de transfert personnalisé dans la fenêtre Editeur de format de fichier. Comme d'autres paramètres de format de fichier, il est possible de remplacer les paramètres de programme de transfert personnalisé s'ils sont modifiés pour une source ou une cible dans un flux de données en particulier. Il est également possible de modifier l'option de transfert personnalisé lors de l'exportation d'un format de fichier.

Informations associées

[Configurer un programme de transfert personnalisé dans l'éditeur de format de fichier \[page 169\]](#)

[Protocoles de transfert de fichier \[page 165\]](#)

11.7.1.3.1 Configurer un programme de transfert personnalisé dans l'éditeur de format de fichier

1. Sélectionnez l'onglet *Formats* dans la bibliothèque d'objets.
2. Faites un clic droit sur *Fichiers plats* dans l'onglet et sélectionnez *Nouveau*.

L'Editeur de format de fichier s'ouvre.

3. Sélectionnez soit le type de fichier *Délimité* soit *Largeur fixe*.

i Remarque

Bien que l'option de programme de transfert personnalisé n'est pas prise en charge par les types de fichier des applications SAP, il est possible de l'utiliser comme une méthode de transport de données pour un flux de donnée SAP ABAP.

4. Saisissez un nom de format
5. Sélectionnez *Oui* pour l'option *Programme de transfert personnalisé*.
6. Développez *Transfert personnalisé* et saisissez le nom du programme de transfert personnalisé et les arguments.
7. Remplissez les autres cases dans la fenêtre de l'éditeur de format de fichier.

Dans la section Fichier(s) de données, indiquez l'emplacement du fichier dans le logiciel.

Pour indiquer des variables de système pour Répertoire racine et Fichier(s) dans la case Arguments :

- Associer la variable de système `$<AW_LOCAL_DIR>` avec l'argument du répertoire local du programme de transfert personnalisé.
- Associer la variable de système `$<AW_FILE_NAME>` avec l'argument du nom de fichier du programme de transfert personnalisé.

Par exemple, saisissez : `-$AW_LOCAL_DIR\AW_FILE_NAME`

Lorsque le programme exécute, les paramètres Répertoire racine et Fichier(s) sont remplacés par ces variables et lus par le programme de transfert personnalisé.

i Remarque

L'indicateur -1 utilisé dans l'exemple ci-dessus est un indicateur de programme personnalisé. Les arguments que vous pouvez utiliser comme des arguments de programme personnalisé dans le logiciel dépendent de ce qu'attend le programme de transfert personnalisé.

8. Cliquez sur [Enregistrer](#).

11.7.1.4 Astuces de conception

Utiliser ces astuces de conception avec les options de transfert personnalisé.

- Les variables ne sont pas prises en charge dans les noms de fichier lors de l'appel d'un programme de transfert personnalisé pour le fichier.
- Il est uniquement possible de modifier les options du transfert personnalisé dans la fenêtre Éditeur de format de fichier (ou Éditeur de banque de données dans le cas de l'application SAP) avant qu'elles soient exportées. Il est impossible de modifier les mises à jour des sources et cibles de fichier au niveau du flux de données une fois exportées. Une fois importés, il est possible d'ajuster les paramètres de l'option de transfert personnalisé au niveau du flux de données. Ils remplacent les paramètres au niveau du format de fichier.

Lors de la conception d'un programme de transfert personnalisé pour travailler avec le logiciel, souvenez-vous que :

- Le logiciel s'attend à ce que le programme de transfert appelé renvoie 0 en cas de réussite, et une valeur différente de zéro en cas d'échec.
- Le logiciel fournit des informations de suivi avant et après que le programme de transfert personnalisé exécute. Le programme de transfert complet et ses arguments avec le mot de passe masqué (le cas échéant) sont écrits dans le journal de suivi. Lorsque "Transfert personnalisé exécuté" apparaît dans le journal de suivi, le programme de transfert personnalisé est terminé.
- Si le programme de transfert personnalisé termine avec succès (code de retour = 0), le logiciel contrôle les éléments suivants :
 - Pour un flux de données ABAP, si le fichier de transport n'existe pas dans le répertoire local, il affiche une erreur et le logiciel s'arrête.
 - Pour une source de fichier, si le fichier ou les fichiers à lire par le logiciel n'existent pas dans le répertoire local, le logiciel écrit un message d'avertissement dans le journal de suivi.
- Si le programme de transfert personnalisé affiche une erreur ou que l'exécution échoue (le code de retour n'est pas 0), alors le logiciel produit une erreur avec le code de retour et la sortie `stdout/stderr`.

- Si le programme de transfert personnalisé réussit mais produit une sortie standard, le logiciel émet un avertissement, enregistre les premiers 1 000 octets de la sortie produite, et poursuit le traitement.
- Le concepteur du programme de transfert personnalisé doit fournir des arguments d'option valides pour garantir que les fichiers sont transférés vers et depuis le répertoire local (spécifié dans le logiciel). Cette opération peut exiger que le fichier distant et le nom du répertoire soient spécifiés comme des arguments puis envoyés à l'interface Designer à l'aide de variables de système.

11.7.2 Utilisation d'un programme de transfert personnalisé

Un programme de transfert personnalisé fait référence à un programme de protocole de transfert de fichiers tiers que vous utilisez pour lire et charger les fichiers plats. Cette option s'applique aux fichiers plats délimités ou de largeur fixe.

Pour les autres types de fichiers plats et pour les schémas imbriqués, il est possible de configurer une définition de protocole de transfert de fichiers en créant un objet d'emplacement de fichier plutôt qu'en utilisant les paramètres du programme de transfert personnalisé.

Si vous sélectionnez *Oui* à l'option *Programme de transfert personnalisé*, vous devez renseigner les options dans la section *Transfert personnalisé* ultérieurement dans la section propriétés-valeurs.

1. Définissez l'option *Programme de transfert personnalisé* sur *Oui*.
2. Développez la section *Transfert personnalisé* et saisissez le nom du programme de transfert personnalisé et les arguments.

Informations associées

[Création d'un nouveau format de fichier plat \[page 154\]](#)

[Protocoles de transfert de fichier \[page 165\]](#)

[Objets d'emplacement de fichier \[page 183\]](#)

12 Format de fichier HDFS

Le format de fichier HDFS (système de fichiers distribué Hadoop) décrit la structure du système de fichiers.

Caractéristique	Description
Classe	Réutilisable
Accès	Dans la bibliothèque d'objets, cliquez sur l'onglet Formats .
Description	Un format de fichier HDFS décrit la structure d'un système de fichiers distribué Hadoop. Il stocke des modèles de format de fichier HDFS dans la bibliothèque d'objets. Le format se compose de plusieurs propriétés qui peuvent être définies dans l'éditeur de format de fichier. Les propriétés disponibles varient selon le mode de l'éditeur. L'éditeur de format de fichier HDFS comprend la plupart des options que proposent habituellement les éditeurs de format de fichier, mais aussi les options qui sont spécifiques au système de fichiers HDFS.

Informations associées

[Formats de fichiers \[page 149\]](#)

12.1 Configuration des résultats du script Pig personnalisé en tant que source

Utilisez un format de fichier HDFS et un script Pig personnalisé pour utiliser les résultats du script PIG en tant que source dans un flux de données.

Créez un format de fichier HDFS ou modifiez-en un existant. Utilisez la section Pig du format de fichier HDFS pour créer ou rechercher un script Pig personnalisé qui édite des données.

Suivez les étapes ci-dessous pour utiliser les résultats d'un script Pig personnalisé dans votre format de fichier HDFS en tant que source :

1. Dans l'éditeur de format de fichier HDFS, sélectionnez [Délimité](#) comme [Type](#) dans la section [Général](#).
2. Saisissez l'emplacement pour le fichier de sortie des résultats du script Pig personnalisé dans [Répertoire racine](#) dans la section [Fichier\(s\) de données](#).
3. Saisissez le nom du fichier qui contiendra les résultats du script Pig personnalisé dans [Fichier\(s\) de données](#).
4. Dans la section [Pig](#), définissez [Script Pig personnalisé](#) sur le chemin d'accès au script Pig personnalisé. L'emplacement doit se trouver sur l'ordinateur où est configuré le Job Server Data Services.

5. Renseignez les options de schéma de sortie applicables pour le script Pig personnalisé.
6. Définissez les séparateurs pour le fichier de sortie dans la section *Séparateurs*.
7. Sauvegardez le format de fichier.

Utilisez le format de fichier comme source dans un flux de données. Lorsque le logiciel exécute le script Pig personnalisé au format de fichier HDFS, il utilise les résultats du script comme données source dans le job.

13 Utilisation de format de fichier du copybook COBOL

Un format de fichier du copybook COBOL décrit la structure définie dans un fichier copybook COBOL (indiqué généralement par une extension .cpy).

Lors de la création d'un format de copybook COBOL, il est possible de :

- créer juste le format, puis configurer la source après avoir ajouté le format à un flux de données ;
- créer le format et l'associer à un fichier de données en même temps.

Cette section décrit également comment :

- créer des règles pour identifier quels enregistrements représentent quels schémas à l'aide d'une option d'ID de champ ;
- identifier le champ qui contient la longueur de l'enregistrement du schéma à l'aide d'une option de champ de longueur d'enregistrement.

Informations associées

Guide de référence : [Types de données](#), [Conversion de/en types de données internes](#) [page 856]

[Objets d'emplacement de fichier](#) [page 183]

13.1 Créer un format de fichier de copybook COBOL

1. Dans la bibliothèque d'objets locale, cliquez sur l'onglet [Formats](#), faites un clic droit sur [Copybook COBOL](#) et cliquez sur [Nouveau](#).

La fenêtre Importer le copybook COBOL s'ouvre.

2. Nommez le format en saisissant un nom dans le champ [Nom du format](#).
3. Sur l'onglet [Format](#) pour l'option [Nom du fichier](#), spécifiez le format de fichier du copybook COBOL à importer, qui porte généralement l'extension .cpy.

Lors de la conception, il est possible de spécifier un fichier de l'une des façons suivantes :

- Pour un fichier situé sur l'ordinateur où Designer s'exécute, vous pouvez utiliser le bouton [Parcourir](#).
 - Pour un fichier localisé sur l'ordinateur où le Job Server est exécuté, vous devez saisir le chemin vers le fichier. Vous pouvez saisir un chemin d'accès absolu ou un chemin d'accès relatif, mais le Job Server doit pouvoir y accéder.
4. Cliquez sur [OK](#).

Le logiciel ajoute le copybook COBOL à la bibliothèque d'objets.

5. La boîte de dialogue *Nom(s) de schéma du copybook COBOL* s'affiche. Si vous le souhaitez, sélectionnez ou double-cliquez sur un nom de schéma pour le renommer.
6. Cliquez sur *OK*.

Lorsque vous ajoutez ultérieurement le format à un flux de données, vous pouvez utiliser les options dans l'éditeur de source pour définir les sources.

13.2 Créer un format de fichier de copybook COBOL et un fichier de données

1. Dans la bibliothèque d'objets locale, cliquez sur l'onglet *Formats*, faites un clic droit sur *Copybook COBOL* et cliquez sur *Nouveau*.

La fenêtre Importer le copybook COBOL s'ouvre.

2. Nommez le format en saisissant un nom dans le champ *Nom du format*.
3. Sur l'onglet *Format* pour l'option *Nom du fichier*, spécifiez le format de fichier du copybook COBOL à importer, qui porte généralement l'extension .cpy.

Lors de la conception, il est possible de spécifier un fichier de l'une des façons suivantes :

- Pour un fichier localisé sur l'ordinateur où le Designer est exécuté, vous pouvez utiliser le bouton Parcourir.
- Pour un fichier localisé sur l'ordinateur où le Job Server est exécuté, vous devez saisir le chemin vers le fichier. Vous pouvez saisir un chemin d'accès absolu ou un chemin d'accès relatif, mais le Job Server doit pouvoir y accéder.

4. Cliquez sur l'onglet *Fichier de données*.
5. Pour l'option *Répertoire*, saisissez ou naviguez vers le répertoire qui contient le fichier de données du copybook COBOL à importer.

Si vous incluez un chemin d'accès à un répertoire à cet endroit, saisissez alors uniquement le nom du fichier dans le champ Nom.

6. Indiquez le *Nom* du fichier de données du copybook COBOL.

Si vous laissez le champ Répertoire vide, saisissez alors un chemin complet et le nom du fichier à cet endroit.

Lors de la conception, il est possible de spécifier un fichier de l'une des façons suivantes :

- Pour un fichier localisé sur l'ordinateur où le Designer est exécuté, vous pouvez utiliser le bouton Parcourir.
- Pour un fichier localisé sur l'ordinateur où le Job Server est exécuté, vous devez saisir le chemin vers le fichier. Vous pouvez saisir un chemin d'accès absolu ou un chemin d'accès relatif, mais le Job Server doit pouvoir y accéder.

7. Si un fichier de données n'est pas sur le même ordinateur que le Job Server, cliquez sur l'onglet *Accès aux données*, sélectionnez *FTP* ou *Personnalisé* puis saisissez les critères pour l'accès au fichier de données.
8. Cliquez sur *OK*.
9. La boîte de dialogue *Nom(s) de schéma du copybook COBOL* s'affiche. Si vous le souhaitez, sélectionnez ou double-cliquez sur un nom de schéma pour le renommer.

10. Cliquez sur [OK](#).

L'onglet ID de champ permet de créer des règles pour identifier quels enregistrements représentent quel schéma.

13.3 Créer des règles pour identifier quels enregistrements représentent quels schémas

1. Dans la bibliothèque d'objets locale, cliquez sur l'onglet [Formats](#), faites un clic droit sur [Copybook COBOL](#) et cliquez sur [Modifier](#).

La fenêtre Modifier le copybook COBOL s'ouvre.
2. Dans le volet supérieur, sélectionnez un champ pour représenter le schéma.
3. Cliquez sur l'onglet [ID de champ](#).
4. Sur l'onglet ID de champ, sélectionnez la case à cocher [Utiliser le champ <nom du schéma. nom du fichier> comme ID](#).
5. Cliquez sur [Insérer en dessous](#) pour ajouter une valeur éditable à la liste Valeurs.
6. Saisissez une valeur pour le champ.
7. Continuez à insérer (ajouter) les valeurs si nécessaire.
8. Sélectionnez les champs supplémentaires et insérez des valeurs si nécessaire.
9. Cliquez sur [OK](#).

13.4 Identifier le champ qui contient la longueur de l'enregistrement du schéma

1. Dans la bibliothèque d'objets locale, cliquez sur l'onglet [Formats](#), faites un clic droit sur [Copybook COBOL](#) et cliquez sur [Modifier](#).

La fenêtre Modifier le copybook COBOL s'ouvre.
2. Cliquez sur l'onglet [Champ de longueur d'enregistrement](#).
3. Pour le schéma à modifier, cliquez sur sa colonne Champ de longueur d'enregistrement pour activer un menu déroulant.
4. Sélectionnez le champ (un par schéma) qui contient la longueur d'enregistrement.

La valeur de décalage est automatiquement remplacée par la valeur par défaut, 4 ; cependant, vous pouvez la modifier avec une autre valeur numérique. Le décalage est la valeur obtenue dans la longueur d'enregistrement totale lorsqu'elle est ajoutée à la valeur dans le champ Longueur d'enregistrement.

5. Cliquez sur [OK](#).

14 Création des formats de classeur Microsoft Excel sur les plateformes UNIX

Décrit comment utiliser un classeur Microsoft Excel comme une source avec un Job Server sur une plateforme UNIX.

Pour créer des formats de classeur Microsoft Excel sur Windows, consultez le *Guide de référence*.

Pour accéder au classeur, vous devez créer et configurer une instance d'adaptateur dans l'Administrateur. La procédure suivante offre une vue d'ensemble du processus de configuration. Pour plus de détails sur la création d'adaptateurs, consultez le *Guide de la Console de gestion*.

Pensez également aux conditions suivantes :

- Pour importer le classeur, il doit être disponible sur un système de fichier Windows. Il est possible de modifier par la suite l'emplacement du fichier actuel à utiliser pour le traitement dans l'éditeur de source du format de classeur Excel. Consultez le *Guide de référence*.
- Pour réimporter ou afficher les données dans Designer, le fichier doit être disponible sous Windows.
- Les entrées dans le fichier du journal des erreurs peuvent être représentées numériquement pour les champs de date et d'heure.
De plus, Data Services écrit les enregistrements avec les erreurs vers la sortie (dans Windows, ces enregistrements sont ignorés).

Informations associées

[Objets d'emplacement de fichier \[page 183\]](#)

14.1 Créer un format de classeur Microsoft Excel sur UNIX

1. À l'aide du Gestionnaire de serveurs (<[LINK_DIR](#)>/bin/svrcfg), vérifiez que le Job Server UNIX peut prendre en charge les adaptateurs. Voir le *Guide d'installation pour Unix*.
2. Assurez-vous qu'un référentiel associé au Job Server est enregistré dans la CMC (Central Management Console). Pour enregistrer un référentiel dans la CMC, consultez le *Guide d'administration*.
3. Dans l'Administrateur, ajoutez un adaptateur pour accéder aux classeurs Excel. Voir le guide *Management Console Guide*.

Il est possible de configurer seulement un adaptateur Excel par Job Server. Utilisez les options suivantes :

- Dans l'onglet Statut, cliquez sur l'adaptateur du Job Server à droite pour le configurer.
- Sur l'onglet Configuration de l'adaptateur de la page Instances de l'adaptateur, cliquez sur [Ajouter](#).
- Dans l'onglet Configuration de l'adaptateur, saisissez le *Nom de l'instance de l'adaptateur*. Saisissez `BOExcelAdapter` (obligatoire et sensible à la casse).

Vous pouvez laisser toutes les options sur leur valeur par défaut sauf s'il s'agit d'un traitement de fichiers supérieur à 1 MB. Dans ce cas, modifiez la valeur des Options supplémentaires du programme de lancement Java `-Xms64m -Xmx512` ou `-Xms128m -Xmx1024m` (la valeur par défaut est `-Xms64m -Xmx256m`). Notez que la gestion de la mémoire Java peut empêcher le traitement de fichiers très volumineux (ou de plusieurs fichiers plus petits).

i Remarque

Lors du démarrage de Data Services 4.2 SP5, l'adaptateur Excel utilise un mécanisme basé sur les événements SAX pour analyser les fichiers Excel. Si vous souhaitez le rétablir comme dans la version 4.2 SP4 ou version précédente, définissez l'indicateur suivant sur Faux et redémarrez le Job Server. L'indicateur est spécifié dans les options du programme de lancement Java pour les paramètres de configuration BOExcelAdapter dans la Console de gestion.

`DuseSAXEventModelForXlsx=false`

Si l'indicateur `DuseSAXEventModelForXlsx` n'est pas spécifié, il est défini par défaut sur vrai et utilise le modèle d'événement SAX.

Lors du traitement de fichiers Excel volumineux, il est recommandé d'augmenter la taille heap de Java dans les options du lanceur Java. Exemple :

`-Xms256m -Xmx4096m`

4. Depuis l'onglet ► [Administrateur](#) ► [Adaptateur](#) ► [Statut de l'instance de l'adaptateur](#) ►, lancez l'adaptateur.
5. Dans Designer, onglet [Formats](#) de la bibliothèque d'objets, créez le format de fichier en important le classeur Excel. Pour plus de détails, consultez le *Guide de référence*.

15 Création de formats de fichier de journal Web

Les journaux Web sont des fichiers plats générés par les serveurs Web et qui sont utilisés pour Business Intelligence.

Les journaux Web suivent généralement les détails des accès au site Web tels que :

- Les noms de domaine du client ou les adresses IP
- Les noms d'utilisateur
- Les horodatages
- Les actions demandées (peut inclure la chaîne de recherche)
- Les octets transférés
- L'adresse en référence
- L'ID de témoin de connexion en charge

Les journaux Web utilisent un format de fichier commun et un format de fichier commun étendu.

Format de journal Web commun :

```
151.99.190.27 - - [01/Jan/1997:13:06:51 -0600]
"GET /~bacuslab HTTP/1.0" 301 -4
```

Format de journal Web commun étendu :

```
saturn5.cun.com - - [25/JUN/1998:11:19:58 -0500]
"GET /wew/js/mouseover.html HTTP/1.0" 200 1936
"http://av.yahoo.com/bin/query?p=mouseover+javascript+source+code&hc=0"
"Mozilla/4.02 [en] (x11; U; SunOS 5.6 sun4m)"
```

Le logiciel prend en charge les formats de journal Web commun et commun étendu comme des sources.

L'éditeur de format de fichier prend également en charge les éléments suivants :

- Un tiret comme un indicateur de valeur nulle
- Le fuseau horaire dans le format date-heure, par exemple 01/Jan/1997:13:06:51 -0600

Le logiciel inclut plusieurs fonctions pour le traitement des données du journal Web :

- Fonction word_ext
- Fonction concat_data_time
- Fonction WL_GetKeyValue

Informations associées

[Fonction word_ext \[page 180\]](#)

[Fonction concat_date_time \[page 180\]](#)

[Fonction WL_GetKeyValue \[page 181\]](#)

15.1 Fonction word_ext

`word_ext` est une fonction de chaîne qui étend la fonction de mot en renvoyant le mot identifié par sa position dans une chaîne délimitée.

Cette fonction est utile pour l'analyse des noms de fichier ou d'URL.

Format

```
word_ext(string, word_number, separator(s))
```

Un nombre de mot négatif signifie qu'il faut compter de droite à gauche

Exemples

`word_ext('www.bodi.com', 2, '.')` renvoie 'bodi'.

`word_ext('www.cs.wisc.edu', -2, '.')` renvoie 'wisc'.

`word_ext('www.cs.wisc.edu', 5, '.')` renvoie NULL.

`word_ext('aaa+=bbb+=ccc+zz=dd', 4, '+=')` renvoie 'zz'. Si deux séparateurs sont spécifiés (+=), la fonction recherche l'un ou l'autre.

`word_ext(',,,,,aaa,,,,bb,,,c ', 2, '.')` renvoie 'bb'. Cette fonction ignore les séparateurs consécutifs.

15.2 Fonction concat_date_time

`concat_date_time` est une fonction de date qui renvoie l'horodatage à partir d'entrées distinctes de date et d'heure.

Format

```
concat_date_time(date, time)
```

Exemple

```
concat_date_time(MS40."date",MS40."time")
```

15.3 Fonction WL_GetKeyValue

WL_GetKeyValue est une fonction personnalisée (écrite en langage de script) qui renvoie les valeurs d'un mot clé donné.

Elle est utile pour l'analyse des chaînes de recherche.

Format

```
WL_GetKeyValue(string, keyword)
```

Exemple

Une recherche sur Google pour bodi B2B est enregistrée dans un journal Web comme :

```
GET "http://www.google.com/search?hl=en&lr=&safe=off&q=bodi+B2B&btnG=Google+Search"
WL_GetKeyValue('http://www.google.com/search?hl=en&lr=&safe=off&q=bodi+B2B&btnG=Google+Search','q') returns 'bodi+B2B'.
```

16 Formats de fichier non structurés

Les formats de fichier non structurés sont un type de format de fichier plat.

Pour lire les fichiers qui contiennent du contenu non structuré, créez un format de fichier comme une source qui lit un ou plusieurs fichier(s) à partir d'un référentiel. Lors de l'exécution, l'objet source dans le flux de données produit une ligne par fichier et contient une référence à chaque fichier pour accéder à son contenu. Dans le flux de données, vous pouvez utiliser une transformation de traitement des données de texte telle que Entity Extraction pour traiter le texte non structuré ou utiliser une autre transformation pour manipuler les données.

Les types de format de fichier non structurés incluent :

Types de format de fichier non structurés	Description
Texte non structuré	Utilisez ce format pour traiter un répertoire de fichiers texte incluant : • Texte • HTML • XML Data Services enregistre le contenu de chaque fichier à l'aide du type de données Long.
Binaire non structuré	utilisez ce format pour lire les documents binaires. Data Services enregistre le contenu de chaque fichier à l'aide du type de données Blob. • Il est possible de traiter une variété de formats de documents en obtenant votre entrée à partir d'une variété de fichiers au format binaire, puis en transmettant cet objet BLOB à la transformation Traitement des données de texte. De cette manière, les formats suivants sont acceptés : ◦ Microsoft Word : 2003, 2007 et 2010 (Office Open XML) ◦ Microsoft PowerPoint : 2003, 2007 et 2010 ◦ Microsoft Excel : 2003, 2007 et 2010 ◦ Adobe PDF : 1.3 – 1.7 ◦ Microsoft RTF : 1.8 et 1.9.1 ◦ Microsoft Outlook E-mail Message : 2003, 2007, 2010 ◦ Message électronique généré : fichiers ".eml" ◦ Texte Open Document, feuille de calcul et présentation : 1.0, 1.1, 1.2 ◦ Corel WordPerfect : 6.0 (1993) – X5 (2010) • Vous pouvez également utiliser le format de fichier binaire non structuré pour déplacer un répertoire de fichiers graphiques sur le disque vers une table de base de données. Supposons que vous souhaitez associer des photos d'employés avec les données d'employés correspondantes qui sont stockées dans une base de données. Le flux de données inclurait la source du format de fichier binaire non structuré, une transformation Query qui associe la photo de l'employé avec les données de l'employé à l'aide du numéro d'ID de l'employée par exemple, et la table cible de base de données.

17 Objets d'emplacement de fichier

Les objets d'emplacement de fichier sont des objets réutilisables permettant de définir un protocole spécifique de transfert de fichiers.

Le logiciel prend en charge les protocoles de transfert de fichiers suivants :

- FTP
- SFTP
- SCP
- Local
- Stockage Cloud Google
- Stockage Cloud Azure
- Azure Data Lake Store
- Stockage Cloud S3 Amazon
- Système de fichiers distribué Hadoop (HDFS)

Pour créer un objet d'emplacement de fichier :

- Entrez les spécificités du protocole de transfert de fichiers pour le type de protocole que vous choisissez.
- Définissez un serveur local et distant pour lesquels vous disposez d'une autorisation d'accès.

Le logiciel stocke les objets d'emplacement de fichier sous *Emplacements de fichier* de l'onglet *Format* de la bibliothèque d'objets locale Designer Data Services.

Une fois l'objet d'emplacement de fichier créé, associez-le à un objet de format lorsque vous créez un objet ou modifiez un objet existant. Pour utiliser les informations du protocole de transfert de fichiers dans un flux de données, associez un objet d'emplacement de fichier à une source ou une cible dans un flux de données.

Les objets de format applicables incluent : fichiers plats, fichiers de schémas imbriqués (XML, DTD et JSON), copybooks COBOL et classeurs Excel.

i Remarque

Les classeurs Excel et les copybooks COBOL associés à un objet d'emplacement de fichier ne peuvent être utilisés qu'en tant que source dans le flux de données.

Le logiciel utilise les informations du serveur local et distant et les protocoles de transfert de fichier dans le flux de données pour transférer des données entre les serveurs locaux et distants.

❖ Exemple

- Utilisez les données du serveur local comme source dans un flux de données.
- Chargez les données de sortie générées depuis le flux de données vers le serveur local puis vers le serveur distant.

Le logiciel fournit des options de contrôle des actions sur les données existantes lorsqu'il transfère les données dans un fichier, à partir de fichiers distants ou locaux. Sélectionnez Options pour ajouter, écraser, supprimer ou enregistrer les données dans les serveurs locaux ou distants. Ces options « Append » sont disponibles en

fonction du type d'objet de format, du type de protocole et du type d'utilisation (source ou cible). Si un fichier de données n'existe pas encore dans le serveur local ou distant, le logiciel crée le fichier.

Informations associées

[Création d'un objet d'emplacement de fichier \[page 188\]](#)

[Gestion des fichiers de données locales ou à distance \[page 184\]](#)

17.1 Gestion des fichiers de données locales ou à distance

Les options d'ajout se trouvent dans les boîtes de dialogue des différents éditeurs qui dépendent du fait que vous travaillez avec une source ou une cible, avec un fichier plat, un schéma imbriqué, un copybook COBOL ou un classeur Excel.

Même si certaines options enregistrent ou suppriment des données existantes, nous y faisons toujours référence comme options d'« ajout ». Le tableau suivant répertorie les options d'ajout et indique comment elles s'appliquent aux fichiers locaux et distants pour les sources et les cibles.

Options d'ajout de fichier

Option	Objet d'emplacement de fichier utilisé comme	Emplacement de l'option	Protocole	Description
<i>Supprimer et recréer le fichier</i>	Cible (schémas imbriqués)	<i>Éditeur de fichier cible de schéma imbriqué</i>	FTP, SFTP	Activé : écrase le fichier à distance existant avec les données de sortie générées. Désactivé : ajoute les données de sortie générées du fichier local vers le fichier à distance. i Remarque Activé ou désactivé : lorsqu'un fichier à distance n'existe pas, le logiciel crée un fichier à distance et le renseigne avec les sorties générées.
			SCP	Activé ou Désactivé : Écrase le fichier à distance avec le contenu du fichier de sortie généré.
			Local, AWS S3	Activé ou Désactivé : Ajoute, écrase ou crée un fichier local.
<i>Supprimer et recréer le fichier</i>	Cible (fichiers plats, XML, modèle XML, JSON, modèle JSON et Copybook)	<i>Éditeur de fichiers cibles</i>	Azure Data Lake Store	Activé : Crée ou écrase le fichier existant dans Azure Data Lake Store avec les données générées. Désactivé : Crée ou ajoute des données générées dans le fichier existant.

Option	Objet d'emplacement de fichier utilisé comme	Emplacement de l'option	Protocole	Description
<i>Supprimer le fichier après le transfert</i>	Source (tous les objets de format applicables)	Éditeur de format de schéma imbriqué (XML, DTD, ou JSON)	Local, AWS S3	Activé : supprime le fichier local après la lecture de données dans le flux de données. Désactivé : enregistre le fichier local après la lecture de données dans le flux de données.
<i>Supprimer le fichier après le transfert</i>	Source (fichiers plats, XML, modèle XML, JSON, modèle JSON, classeur Excel et Copybook)	Éditeur de fichiers source	Azure Data Lake Store	Activé : supprime le fichier local après la lecture de données dans le flux de données. Désactivé : enregistre le fichier local après la lecture de données dans le flux de données.
<i>Supprimer le fichier après le transfert</i>	Cible (tous les objets de format applicables)	Éditeur de fichiers cibles	Azure Data Lake Store	Activé : Supprime le fichier local après le chargement des données dans Azure Data Lake Store cible. Désactivé : Enregistre le fichier local après le chargement des données dans Azure Data Lake Store cible.

Option	Objet d'emplacement de fichier utilisé comme	Emplacement de l'option	Protocole	Description
<i>Supprimer le fichier</i>	Cible (fichiers plats uniquement)	<i>Éditeur de fichiers cibles</i>	FTP, SFTP	Activé : écrase le contenu du fichier à distance avec les sorties générées. Désactivé : ajoute les sorties générées au contenu existant du fichier distant.
				i Remarque Activé ou Désactivé : lorsqu'un fichier distant n'existe pas, le logiciel crée le fichier distant et le renseigne avec les sorties générées.
			SCP	Activé ou Désactivé : écrase un fichier distant existant ou crée un fichier distant s'il n'en existe pas.
			Local, AWS S3	Activé ou Désactivé : écrase ou ajoute des données au fichier local.

Informations associées

[Associer des objets d'emplacement de fichier à des formats de fichier \[page 192\]](#)

17.2 Création d'un objet d'emplacement de fichier

Créez un objet d'emplacement de fichier à utiliser avec les formats de fichier pour définir le protocole de transfert de fichiers entre les serveurs locaux et à distance.

Suivez cette procédure pour créer un objet d'emplacement de fichier :

1. Cliquez avec le bouton droit sur le dossier *Emplacements de fichiers* dans l'onglet *Formats* de la bibliothèque d'objets et sélectionnez *Nouveau*.
2. Saisissez un *Nom* pour le nouvel objet dans la boîte de dialogue *Créer un emplacement de fichier*.
3. Sélectionnez un protocole depuis la liste déroulante *Protocole*. Le protocole par défaut est Local.

La définition de l'option *Protocole* détermine les options restantes. La table suivante répertorie de brèves descriptions des options. Pour les descriptions complètes et pour en savoir plus, voir les descriptions dans le *Guide de référence*.

Option	Brève description
<i>Hôte</i>	Nom du serveur à distance
<i>Port</i>	Numéro de port du serveur à distance
<i>Empreinte de la clé d'hôte</i>	Code de l'ordinateur hôte (SFTP et SCP uniquement)
<i>Type d'autorisation</i>	Mot de passe ou clé publique
<i>Utilisateur</i>	Nom d'utilisateur du serveur à distance
<i>Mot de passe</i>	Mot de passe du serveur à distance
<i>Chemin d'accès du fichier de clé privée pour l'autorisation SSH</i>	Chemin d'accès au fichier de clé privée (SFTP et SCP uniquement)
<i>Phrase secrète de clé privée pour l'autorisation SSH</i>	Phrase secrète du fichier de clé privée (SFTP et SCP uniquement)
<i>Chemin d'accès du fichier de clé publique pour l'autorisation SSH</i>	Chemin d'accès au fichier de clé publique (SFTP et SCP uniquement)
<i>Nombre de tentatives de connexion</i>	Nombre de nouvelles tentatives de connexion
<i>Intervalle de tentatives de connexion</i>	Délai en secondes entre chaque nouvelle tentative de connexion
<i>Répertoire à distance</i>	(facultatif) chemin d'accès du fichier pour le serveur à distance
<i>Répertoire local</i>	Chemin d'accès du fichier du serveur local pour cet objet d'emplacement de fichier

4. Cliquez sur *OK*.

Le logiciel crée l'objet d'emplacement de fichier. Il est possible de créer des configurations multiples de l'objet d'emplacement de fichier en cliquant sur [Modifier](#) plutôt que sur [OK](#). Autrement, vous pouvez modifier un objet d'emplacement de fichier existant pour créer des configurations supplémentaires.

Informations associées

[Associer des objets d'emplacement de fichier à des formats de fichier \[page 192\]](#)

[Créer des configurations multiples pour un objet d'emplacement de fichier \[page 191\]](#)

17.3 Obtenir une autorisation SSH

Le processus d'obtention des informations d'autorisation SSH varie en fonction du système hôte.

1. Exécutez une commande qui génère un fichier de clé privée et un fichier de clé publique.

Commande pour un hôte à distance Linux : `~/ .ssh`

2. Faites de cette commande votre répertoire actuel.

Pour l'hôte à distance Linux, faites de `~/ .ssh` votre répertoire actuel.

3. Ajoutez le contenu du fichier de la clé publique au fichier `authorized_keys`.

Si vous utilisez RSA; la commande Linux génère le fichier de clé privée `id_rsa` et le fichier de clé publique `id_rsa.pub`.

4. Copiez le fichier de clé privée et le fichier de clé publique dans l'hôte Job server.

Si vous utilisez RSA, copiez le fichier de clé privée `id_rsa` et le fichier de clé publique `id_rsa.pub` dans l'hôte Job server.

5. Saisissez le nom du fichier de clé privée et celui de clé publique dans les options de l'objet d'emplacement de fichier.

Si vous utilisez RSA pour Linux (avec l'exemple ci-dessus), saisissez les valeurs suivantes pour renseigner les options de l'objet d'emplacement de fichier :

- *Chemin d'accès du fichier de clé privée SSH autorisée* = `id_rsa`
- *Chemin d'accès du fichier de clé publique SSH autorisée* = `id_rsa.pub`
- *Phrase secrète du fichier de clé privée SSH autorisée* = saisissez une phrase secrète seulement si vous en avez définie une lors de la génération de la clé privée SSH.

17.4 À propos de l'empreinte de la clé d'hôte

L'empreinte de la clé d'hôte générée doit utiliser l'algorithme MD5 (message digest) ou SHA1 (secure hash). Le format doit être codé en valeurs de chaînes hexadécimales séparées par des deux-points.

Exemple de valeur de chaîne hexadécimale : `xx:xx: . . . . . :xx`. Pour MD5, la longueur d'empreinte au format hexadécimal est égale à 47 (séparateurs compris). Pour SHA1, la longueur d'empreinte au format hexadécimal est égale à 59 (séparateurs compris).

❖ Exemple

Pour un hôte Linux, obtenez les informations relatives à l'empreinte de la clé d'hôte en vous connectant à l'hôte de l'objet d'emplacement de fichier que vous avez défini dans l'éditeur d'objets d'emplacement de fichier. Exécutez ensuite une commande de génération de clé SSH telle que :

```
user@host> sudo ssh-keygen -lf /etc/ssh/ssh-host_rsa_key.pub  
user@host> sudo ssh-keygen -lf /etc/ssh/ssh-host_rsa_key.pub
```

Le résultat généré peut être soit au format SHA1, soit MD5. Il correspond à la valeur que vous avez saisie dans l'option *Empreinte de la clé d'hôte* :

Pour la sortie SHA1, vous saisissez l'*Empreinte de la clé d'hôte* suivante : `2048 49:fc:79:ef:dd:6c:d3:1b:90:e6:67:9a:d5:93:3a:ac`

Pour la sortie MD5, vous saisissez l'*Empreinte de la clé d'hôte* suivante : `root@linux-sles-sp3 (RSA)`

17.5 Modification d'un objet d'emplacement de fichier existant

Modifiez un objet d'emplacement de fichier pour modifier les paramètres ou ajouter des configurations supplémentaires.

1. Cliquez avec le bouton droit sur le dossier *Emplacements de fichiers* pour modifier dans l'onglet *Formats* de la bibliothèque d'objets et sélectionnez *Modifier*.

La boîte de dialogue *Modifier l'emplacement de fichier*<nomdufichier> s'ouvre.

2. Modifiez les autres options applicables et cliquez sur *OK*.

Vous donnez un nom à l'objet d'emplacement de fichier lorsque vous le créez. Il est impossible de modifier le nom de l'objet d'emplacement de fichier lorsque vous le modifiez.

Informations associées

[Création de configurations multiples d'objet d'emplacement de fichier \[page 191\]](#)

17.6 Créer des configurations multiples pour un objet d'emplacement de fichier

Créez des configurations multiples d'objet d'emplacement de fichier pour chaque type de protocole de transfert de fichier utilisé.

Chaque objet d'emplacement de fichier possède au moins une configuration. Si une seule configuration existe, c'est la configuration par défaut. Lorsque vous créez des configurations multiples dans un objet d'emplacement de fichier, vous pouvez modifier la configuration par défaut de n'importe quelle configuration dans l'objet.

L'utilisation de configurations multiples est pratique. Par exemple, lorsque vous utilisez différents protocoles en fonction de votre environnement, il est possible de créer un objet d'emplacement de fichier qui contient les configurations pour vos environnements de test et de production. Dans votre environnement de test, vous utilisez un protocole local car les tests ont lieu en interne. Lorsque vous migrez à l'environnement de production, vous utilisez SFTP afin de transférer en toute sécurité les fichiers entre le serveur à distance et le serveur local. Basculez entre les protocoles en modifiant l'objet d'emplacement de fichier utilisé dans votre job pour utiliser une configuration par défaut différente.

17.6.1 Création de configurations multiples d'objet d'emplacement de fichier

Étapes pour créer des configurations d'emplacement de fichier supplémentaires dans un objet d'emplacement de fichier unique et pour sélectionner une configuration par défaut.

1. Cliquez avec le bouton droit sur le nom d'un objet d'emplacement de fichier dans *Emplacements de fichiers* de l'onglet *Format* dans la bibliothèque d'objets et cliquez sur *Modifier*.

La boîte de dialogue *Modifier l'emplacement de fichier <nomdufichier>* s'ouvre.

2. Cliquez sur *Modifier...*

La boîte de dialogue *Configurations pour l'emplacement de fichier <nomdufichier>* s'ouvre.

3. Cliquez sur l'icône Créer une configuration ().

La boîte de dialogue *Créer une configuration* s'ouvre.

4. Saisissez un nom unique pour la configuration et cliquez sur *OK*.

Une nouvelle colonne Configurations apparaît avec le nouveau nom en en-tête.

5. Si cette configuration est votre configuration par défaut, sélectionnez *Oui* dans *Configuration par défaut*.

i Remarque

Si vous avez des configurations multiples dans un objet d'emplacement de fichier, le logiciel bascule automatiquement les autres configurations sur *Non* dans *Configuration par défaut*.

6. Sélectionnez le protocole depuis le menu déroulant *Protocole*.

Les options restantes apparaissant après l'option *Protocoles* dépendent du type de protocole que vous choisissez. Toutefois, si vous avez des configurations multiples (plus d'une colonne de paramètres), il peut

y avoir des options répertoriées qui ne s'appliquent pas au protocole que vous définissez. Les options ne s'appliquant pas à une configuration apparaissent dans la colonne « Sans objet ».

7. Renseignez les autres options applicables et cliquez sur [OK](#).

Pour les descriptions d'options, consultez le *Guide de référence*.

17.7 Associer des objets d'emplacement de fichier à des formats de fichier

Pour utiliser les informations de protocole de transfert de fichier que vous saisissez dans un objet d'emplacement de fichier, vous associez l'objet d'emplacement de fichier à une source de fichier spécifique ou à une cible dans le flux de données.

Vous pouvez également inclure un objet d'emplacement de fichier lorsque vous créez ou modifiez un nouveau format.

i Remarque

Certains formats ne s'appliquent pas aux objets d'emplacement de fichier. Dans ce cas, les options de l'objet d'emplacement de fichier ne s'appliquent pas. Voir la table suivante pour les types de format.

La table suivante vous explique synthétiquement comment associer un objet spécifique d'emplacement de fichier à un format de fichier spécifique.

Type de format	Associer un objet d'emplacement de fichier
Fichier plat	Créez un nouveau format de fichier plat ou modifiez-en un et ajoutez les informations de l'objet d'emplacement de fichier dans la section <i>Fichiers de données</i> de l' <i>Éditeur de format de fichier</i> .
Fichier de schéma DTD, XML, ou JSON	Modifiez un fichier de schéma et ajoutez les informations de l'objet d'emplacement de fichier dans l'onglet <i>Format</i> de l' <i>Éditeur de format</i> .
Copybook COBOL	Modifiez un copybook COBOL existant et saisissez les informations d'emplacement de fichier dans l'onglet <i>Fichier de données</i> de la boîte de dialogue <i>Modifier le copybook COBOL</i> .
Classeur Excel	Modifiez un classeur Excel existant et saisissez les informations d'emplacement de fichier dans l'onglet <i>Format</i> de la boîte de dialogue <i>Importer le classeur Excel</i> .

Informations associées

Guide de Designer : Objets d'emplacement de fichier, Gestion des fichiers de données locales ou à distance [page 184]

Guide de Designer : Emplacements de fichier, Association d'un format de fichier plat à un objet d'emplacement de fichier [page 193]

Guide de Designer : Emplacements de fichier, Association des fichiers de schéma XML, DTD ou JSON [page 194]

Guide de Designer : Emplacements de fichier, Association des copybooks COBOL à un objet d'emplacement de fichier [page 195]

Guide de Designer : Emplacements de fichier, Association de classeurs Excel à un objet d'emplacement de fichier [page 196]

17.7.1 Association d'un format de fichier plat à un objet d'emplacement de fichier

Associer un format de fichier plat existant à un objet d'emplacement de fichier pour utiliser le protocole de transfert spécifié dans l'objet d'emplacement de fichier pour accéder aux données.

i Remarque

Il est possible d'associer un objet d'emplacement de fichier existant à un modèle existant de format de fichier plat ou à un nouveau modèle. Les instructions suivantes fonctionnent avec un modèle existant de format de fichier plat.

1. Cliquez avec le bouton droit sur le format de fichier plat applicable dans la bibliothèque d'objets sous *Fichiers plats* et cliquez sur *Modifier*.

La boîte de dialogue *Éditeur de format de fichier*<filename> s'ouvre.

2. Renseignez les options suivantes dans la section *Fichiers de données* :

Option	Description
<i>Emplacement</i>	Nom de l'objet d'emplacement de fichier.
<i>Supprimer le fichier après le transfert</i>	L'option apparaît uniquement lorsque vous choisissez un objet d'emplacement de fichier pour l'option <i>Emplacement</i>. ○ Oui : le logiciel supprime la copie du fichier du serveur local après son utilisation en tant que source ou cible dans le flux de données. ○ Non : le logiciel enregistre la copie du fichier du serveur local après son utilisation en tant que source ou cible dans le flux de données.

Option	Description
	Il est possible de modifier ce paramètre dans l'Éditeur source ou cible lorsque vous définissez la paire format de fichier/objet d'emplacement de fichier en tant que source ou cible dans le flux de données.

3. Renseignez les autres options applicables liées au format de fichier puis cliquez sur [Enregistrer et fermer](#).

Informations associées

[Objets sources et cibles \[page 206\]](#)

[Gestion des fichiers de données locales ou à distance \[page 184\]](#)

17.7.2 Association des fichiers de schéma XML, DTD ou JSON

Étapes d'association d'un schéma imbriqué à un objet d'emplacement de fichier.

Créez un schéma imbriqué pour DTD, XML ou JSON.

1. Cliquez avec le bouton droit sur le modèle de schéma imbriqué applicable dans la bibliothèque d'objets sous [Schémas imbriqués](#) et cliquez sur [Modifier](#).

Le schéma s'ouvre et affiche l'onglet [Général](#), de lecture seule.

2. Ouvrez l'onglet [Format](#) et renseignez les options suivantes :

Option	Description
Emplacement du fichier	Nom de l'objet d'emplacement de fichier.
Supprimer le fichier après le transfert	L'option apparaît uniquement lorsque vous choisissez un objet d'emplacement de fichier pour l'option Emplacement du fichier. ○ Oui : le logiciel supprime la copie du fichier du serveur local après son utilisation en tant que source ou cible dans le flux de données. ○ Non : le logiciel enregistre la copie du fichier du serveur local après son utilisation en tant que source ou cible dans le flux de données. Il est possible de modifier ce paramètre dans l'éditeur de source ou de cible lorsque vous utilisez cet objet dans un flux de données.

3. Renseignez les autres options applicables liées au format de fichier puis cliquez sur [Enregistrer et fermer](#).

Lorsque vous incluez le schéma imbriqué que vous venez de créer dans un flux de données, l'objet d'emplacement de fichier apparaît dans l'onglet Source.

Informations associées

[Gestion des fichiers de données locales ou à distance \[page 184\]](#)

17.7.3 Association des copybooks COBOL à un objet d'emplacement de fichier

Étapes d'association d'un copybook COBOL à un objet d'emplacement de fichier.

i Remarque

Un objet de copybook COBOL ne peut être utilisé qu'en tant que source dans un flux de données.

1. Cliquez avec le bouton droit sur le copybook COBOL applicable dans la bibliothèque d'objets sous [Schémas imbriqués](#) et cliquez sur [Modifier](#).

La boîte de dialogue [Modifier le copybook COBOL](#) s'ouvre.

2. Renseignez les options suivantes dans l'onglet [Fichier de données](#) :

Option	Description
Emplacement du fichier	Nom de l'objet d'emplacement de fichier.
Supprimer le fichier après le transfert	L'option apparaît uniquement lorsque vous choisissez un objet d'emplacement de fichier pour l'option Emplacement du fichier. ○ Oui : le logiciel supprime la copie du fichier du serveur local après son utilisation en tant que source dans le flux de données. ○ Non : le logiciel enregistre la copie du fichier du serveur local après son utilisation en tant que source dans le flux de données. Il est possible de modifier ce paramètre dans l'éditeur de source ou de cible lorsque vous utilisez cet objet dans un flux de données. Non applicable aux objets cible.

3. Renseignez les options restantes le cas échéant. Les descriptions des options se trouvent dans le [Guide de référence](#).

L'objet d'emplacement de fichier est associé au copybook COBOL.

Informations associées

[Créer un format de fichier de copybook COBOL \[page 174\]](#)

[Gestion des fichiers de données locales ou à distance \[page 184\]](#)

17.7.4 Association de classeurs Excel à un objet d'emplacement de fichier

Étapes d'association d'un classeur Excel à un objet d'emplacement de fichier.

i Remarque

Les classeurs Excel ne peuvent être utilisés qu'en tant que source dans un flux de données.

1. Cliquez avec le bouton droit sur un classeur Excel existant dans *Classeurs Excel* de l'onglet *Formats* de la bibliothèque d'objets puis cliquez sur *Modifier*.

La boîte de dialogue *Importer le classeur Excel* s'ouvre.

2. Renseignez les options suivantes dans l'onglet *Format* :

Option	Description
<i>Emplacement du fichier</i>	Nom de l'objet d'emplacement de fichier.
<i>Supprimer le fichier après le transfert</i>	L'option apparaît uniquement lorsque vous choisissez un objet d'emplacement de fichier pour l'option <i>Emplacement du fichier</i>. ○ Oui : le logiciel supprime la copie du fichier du serveur local après son utilisation en tant que source dans le flux de données. ○ Non : le logiciel enregistre la copie du fichier du serveur local après son utilisation en tant que source dans le flux de données. Il est possible de modifier ce paramètre dans l'éditeur de source ou de cible lorsque vous utilisez cet objet dans un flux de données. Non applicable aux objets cible.

3. Renseignez les options restantes le cas échéant. Les descriptions des options se trouvent dans le *Guide de référence*.

L'objet d'emplacement de fichier est associé au classeur Excel.

Informations associées

[Créer un format de classeur Microsoft Excel sur UNIX \[page 177\]](#)

17.8 Ajout d'un objet d'emplacement de fichier à un flux de données

Étapes pour l'ajout d'un objet d'emplacement de fichier à un flux de données en tant que source ou cible.

Vous pouvez effectuer ces étapes avec un objet de format auquel est associé ou non un objet d'emplacement de fichier. Si aucun objet d'emplacement de fichier n'est associé au format d'objet, ou si vous souhaitez modifier l'objet d'emplacement de fichier associé, il est possible d'ajouter ou de modifier l'association en effectuant la procédure ci-dessous.

1. Glissez l'objet de format applicable dans un flux de données dans votre espace de travail.
2. Dans la boîte de dialogue contextuelle, sélectionnez *Créer la source* ou *Créer la cible*, le cas échéant.
3. Cliquez sur le nom de l'objet que vous venez d'ajouter. S'il s'agit d'une source, l'*Éditeur de fichiers source* s'ouvre. S'il s'agit d'une cible, l'*Éditeur de fichiers cible* s'ouvre.

Les options qui apparaissent varient en fonction du type de d'objet de format choisi (fichier plat, JSON, XML par exemple).

4. Dans l'*Éditeur de fichiers source* ou dans l'*Éditeur de fichiers cible*, sélectionnez un objet d'emplacement de fichier existant dans le menu déroulant *Emplacement*. Il est également possible de modifier une association existante à un objet d'emplacement de fichier le cas échéant.
5. Fermez la fenêtre de l'*Éditeur de fichiers source* ou de l'*Éditeur de fichiers cible*.
6. Continuez la configuration du flux de données, le cas échéant.

Informations associées

[Création et définition des flux de données \[page 204\]](#)

[Objets d'emplacement de fichier \[page 183\]](#)

[Associer des objets d'emplacement de fichier à des formats de fichier \[page 192\]](#)

[Gestion des fichiers de données locales ou à distance \[page 184\]](#)

17.9 Utilisation de scripts pour déplacer des fichiers depuis ou vers un serveur à distance

Exemples des fonctions intégrées d'objet d'emplacement de fichier dans les scripts dans un workflow.

Utilisez les deux fonctions intégrées `copy_from_remote_server` et `copy_to_remote_server` dans les scripts de flux de données afin d'identifier un objet d'emplacement de fichier et un format de fichier au lieu d'utiliser un format associé à un objet d'emplacement de fichier. Les scripts identifient l'objet d'emplacement de fichier contenant vos paramètres de protocole de transfert de fichier, afin que le logiciel utilise votre protocole pour copier les fichiers depuis ou vers un serveur à distance. Vous pouvez également saisir un nom de fichier en utilisant des caractères génériques (*) et (?) pour déplacer un groupe de fichiers qui correspondent au critère du caractère générique.

❖ Exemple

Exemples de fichier de script

Fonction	Résultat
↳ Exemple de code <pre>copy_from_remote_system('FTP_Datastore', '*.txt');</pre>	Téléchargez les fichiers qui correspondent au critère du caractère générique du nom de fichier depuis le serveur à distance.
↳ Exemple de code <pre>copy_from_remote_system('FTP_Datastore', 'mydata.txt');</pre>	Téléchargez le fichier individuel depuis le serveur à distance.
↳ Exemple de code <pre>copy_to_remote_system('FTP_Datastore', 'myremotedata.txt');</pre>	Téléchargez le fichier individuel vers le serveur à distance.
↳ Exemple de code <pre>copy_to_remote_system('FTP_Datastore', '*.txt');</pre>	Téléchargez les fichiers qui correspondent au critère du caractère générique du nom de fichier vers le serveur à distance.

17.9.1 Déplacement de fichiers vers et à partir des conteneurs Azure

Data Services permet de déplacer les fichiers d'un stockage local, comme un lecteur local ou un dossier, vers un conteneur Azure.

Data Services permet de déplacer les fichiers d'un stockage local, comme un lecteur local ou un dossier, vers un conteneur Azure. Vous pouvez utiliser un conteneur existant ou en créer un, s'il n'existe pas. Vous pouvez également importer des fichiers (appelés "blobs" dans un conteneur) du conteneur Azure vers un lecteur local ou un dossier. Les fichiers peuvent être de tout type et ne sont pas manipulés en interne par Data Services. Actuellement, Data Services prend en charge les blobs de blocs dans le type de stockage Conteneur.

Vous utilisez un format de fichier pour décrire un fichier Blob et l'utilisez dans un flux de données pour exécuter des opérations supplémentaires sur un fichier. Le format de fichier peut également être utilisé dans un script pour automatiser le téléchargement et la suppression d'un fichier local.

Les étapes suivantes sont des étapes de niveau supérieur pour le téléchargement de fichier dans un blob de stockage Conteneur dans Microsoft Azure.

1. Créez un compte de stockage dans Azure et prenez note de la clé partagée primaire. Pour plus d'informations, voir la documentation Microsoft ou l'assistance technique Microsoft.
2. Créez un objet d'emplacement de fichier avec le protocole Stockage cloud Azure. Pour des détails sur la configuration des options d'objets d'emplacement de fichier dans Azure, voir le *Guide de référence*.
3. Créez un job dans le Data Services Designer.
4. Ajoutez un script contenant la fonction appropriée dans le job.

- `copy_to_remote_system`
- `copy_from_remote_system`

Exemple

```
copy_to_remote_system('New_FileLocation', '*')
```

Un script contenant cette fonction copie tous les fichiers du répertoire local spécifié dans l'objet d'emplacement de fichier vers le conteneur spécifié dans le même objet.

5. Enregistrez et exécutez le job.

18 Flux de données

Décrit les fondements des flux de données, y compris les objets de flux de données, l'utilisation des recherches, l'exécution de flux de données et les audits.

18.1 Qu'est-ce qu'un flux de données ?

Les flux de données extraient, transforment et chargent les données.

Tout ce qui a trait aux données, y compris la lecture des sources, la transformation des données et le chargement des cibles, se produit dans un flux de données. Les lignes qui connectent les objets dans un flux de données représentent le flux des données durant les étapes de transformation des données.

Après avoir défini un flux de données, il est possible de l'ajouter à un job ou à un workflow. À l'intérieur d'un workflow, un flux de données peut envoyer et recevoir des informations vers et depuis d'autres objets grâce aux paramètres d'entrée et de sortie.



→ N'oubliez pas

Sachez que les données fournies sont placées dans les journaux de suivi, exemples de rapports et référentiels (aux conséquences néfastes sur les données), etc. En d'autres termes, vos données pourront être stockées dans des emplacements autres que les fichiers de sortie.

18.1.1 Dénomination des flux de données

Les noms des flux de données peuvent inclure des caractères alphanumériques et des tirets de soulignement (_). Ils ne peuvent pas contenir d'espace vide.

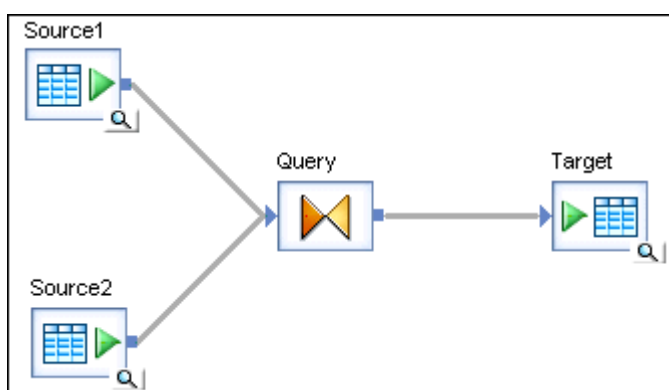
18.1.2 Exemple d'un flux de données

Supposons que vous souhaitez renseigner la table des faits de votre entrepôt de données avec de nouvelles données à partir de deux tables contenues dans la base de données de transaction source.

Le flux de données comporte les éléments suivants :

- Deux tables sources
- Une jointure entre ces tables, définie dans une transformation Query
- Une table cible où les nouvelles lignes sont placées

Vous indiquez le flux des données par ces composants en les connectant dans l'ordre dans lequel les données se déplacent dans ces derniers. Le flux de données obtenu ressemble à l'image ci-dessous :



18.1.3 Etapes dans un flux de données

Chaque icône que vous placez dans le diagramme de flux de données devient une étape dans le flux de données.

Il est possible d'utiliser les objets suivants comme des étapes dans un flux de données :

- source
- cible
- transformations

Les connexions que vous établissez entre les icônes déterminent l'ordre dans lequel le logiciel achève les étapes.

Informations associées

[Objets sources et cibles \[page 206\]](#)

[Transformations \[page 227\]](#)

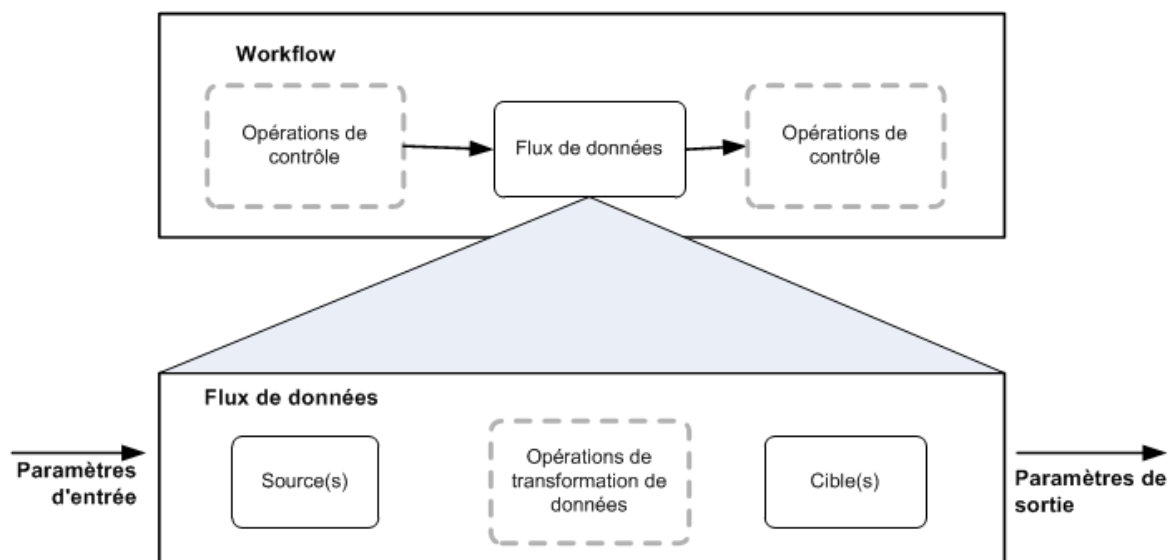
18.1.4 Les flux de données comme des étapes dans les workflows

Les flux de données sont des opérations fermées, même lorsqu'ils sont des étapes dans un workflow.

Les jeux de données créés dans un flux de données ne sont pas disponibles pour les autres étapes dans le workflow.

Un workflow n'opère pas sur les jeux de données et ne peut pas fournir plus de données à un flux de données ; cependant, un workflow peut réaliser les opérations suivantes :

- Appeler les flux de données pour effectuer les opérations de mouvement des données
- Définir les conditions appropriées pour exécuter les flux de données
- Transmettre les paramètres vers et depuis les flux de données



18.1.5 Jeux de données intermédiaires dans un flux de données

Chaque étape dans un flux de données, jusqu'à la définition de cible, produit un résultat intermédiaire (par exemple, les résultats d'une instruction SQL contenant une clause WHERE), qui se déplace vers l'étape suivante d'un flux de données.

Le résultat intermédiaire comporte un ensemble de lignes de l'opération précédente et le schéma dans lequel les lignes sont organisées. Ce résultat est appelé un jeu de données. Ce jeu de données peut, à son tour, être davantage "filtré" et dirigé dans un autre jeu de données.

18.1.6 Codes d'opération

Chaque ligne dans un jeu de données est marquée d'un code d'opération qui identifie le statut de la ligne.

Les codes d'opération sont comme suit :

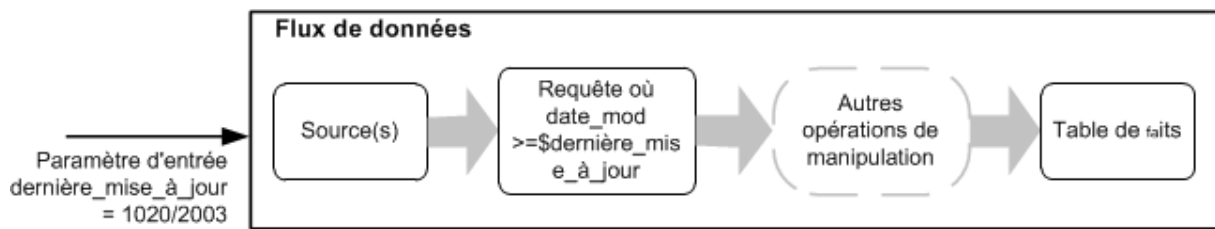
Code d'opération	Description
NORMAL	Crée une nouvelle ligne dans la cible. Toutes les lignes dans un jeu de données sont marquées d'un indicateur NORMAL lorsqu'elles sont extraites d'une source. Si une ligne est marquée d'un indicateur NORMAL lorsqu'elle est chargée dans une cible, elle est insérée comme une nouvelle ligne dans la cible.
INSERT	Crée une ligne dans la cible. Les lignes peuvent être marquées d'un indicateur INSERT par les transformations dans le flux de données pour indiquer qu'une modification a eu lieu dans un jeu de données, par rapport à une image plus ancienne du même jeu de données. La modification est enregistrée dans la cible séparément à partir des données existantes.
DELETE	Ignoré par la cible. Les lignes marquées d'un indicateur DELETE ne sont pas chargées. Les lignes peuvent être marquées d'un indicateur DELETE uniquement par la transformation Map_Operation.
UPDATE	Remplace une ligne existante dans la cible. Les lignes peuvent être marquées d'un indicateur UPDATE par les transformations dans le flux de données pour indiquer qu'une modification a eu lieu dans un jeu de données, par rapport à une image plus ancienne du même jeu de données. La modification est enregistrée dans la cible dans la même ligne que les données existantes.

18.1.7 Transfert des paramètres vers les flux de données

Les données ne circulent pas en dehors d'un flux de données, même lorsque vous ajoutez un flux de données à un workflow. Néanmoins, il est possible de transmettre des paramètres vers et depuis un flux de données.

Les paramètres évaluent les valeurs individuelles plutôt que les ensembles de valeurs. Lorsqu'un flux de données reçoit des paramètres, les étapes à l'intérieur du flux de données peuvent faire référence à ces paramètres comme des variables.

Les paramètres rendent les définitions de flux de données plus flexibles. Par exemple, un paramètre peut indiquer la dernière fois qu'une table des faits a été mise à jour. Il est possible d'utiliser cette valeur dans un flux de données pour extraire uniquement les lignes modifiées depuis la dernière mise à jour. Le schéma ci-dessous indique le paramètre `last_update` utilisé dans une requête pour déterminer le jeu de données utilisé pour charger la table des faits.



Informations associées

[Variables et paramètres \[page 308\]](#)

18.2 Création et définition des flux de données

Vous pouvez créer des flux de données à l'aide d'objets de la bibliothèque d'objets ou de la palette d'outils. Après avoir créé un flux de données, il est possible de modifier ses propriétés.

Informations associées

[Modification des propriétés d'un flux de données \[page 205\]](#)

18.2.1 Définir un nouveau flux de données à l'aide de la bibliothèque d'objets

1. Dans la bibliothèque d'objets, accédez à l'onglet *Flux de données*.
2. Sélectionnez la catégorie de flux de données, faites un clic droit et sélectionnez *Nouveau*.
3. Sélectionnez le nouveau flux de données.
4. Faites glisser le flux de données dans l'espace de travail pour un job ou un workflow.
5. Ajoutez les sources, les transformations et les cibles dont vous avez besoin.

18.2.2 Définir un nouveau flux de données à l'aide de la palette d'outils

1. Sélectionnez l'icône du flux de données dans la palette d'outils.
2. Cliquez sur l'espace de travail pour un job ou un workflow pour placer le flux de données.

Il est possible d'ajouter des flux de données à un job batch ou à un job en temps réel. Lorsque vous faites glisser une icône de flux de données dans un job, vous dites au logiciel de valider ces objets en fonction des conditions du type de job (batch ou en temps réel).

3. Ajoutez les sources, les transformations et les cibles dont vous avez besoin.

18.2.3 Modification des propriétés d'un flux de données

Après avoir créé un flux de données, il est possible de modifier ses propriétés.

1. Faites un clic droit sur le flux de données et sélectionnez *Propriétés*.

La fenêtre *Propriétés* s'ouvre pour le flux de données.

2. Modifiez les propriétés souhaitées d'un flux de données.
3. Cliquez sur *OK*.

Cette table décrit les différentes propriétés qu'il est possible de définir pour le flux de données.

Option	Description
<i>Exécuter une seule fois</i>	Lorsque vous indiquez qu'un flux de données doit être exécuté une seule fois, un job batch ne ré-exécutera jamais ce flux de données après qu'il ait terminé avec succès, sauf si le flux de données est contenu dans un workflow qui est une unité de restauration qui ré-exécute et qui n'a pas terminé avec succès ailleurs en dehors de l'unité de restauration. Il est recommandé de ne pas marquer un flux de données à Exécuter une seule fois si un workflow parent est une unité de restauration.
<i>Utiliser les liens de base de données</i>	Les liens de base de données sont des chemins de communication entre un serveur de base de données et un autre. Les liens de base de données permettent aux utilisateurs locaux d'accéder à des données sur une base de données distante, qui peut se trouver sur l'ordinateur local ou distant, du même type de base de données ou d'un type différent.
<i>Degré de parallélisme</i>	Le degré de parallélisme (DOP) est une propriété d'un flux de données qui définit le nombre de fois que chaque transformation dans un flux de données réplique pour traiter un sous-ensemble parallèle de données.
<i>Type de cache</i>	Il est possible de cacher les données pour améliorer la performance des opérations telles que les jointures, les groupements, les tris, les filtrages, les recherches et les comparaisons de table. Il est possible de sélectionner l'une des valeurs suivantes pour l'option Type de cache dans la fenêtre de Propriétés du flux de données : ◦ <i>En mémoire</i> : sélectionnez cette valeur si le flux de données traite une petite quantité de données qui peuvent être contenues dans la mémoire disponible. ◦ <i>Paginable</i> : cette valeur est celle par défaut.

Option	Description
<i>Contournement</i>	Permet de contourner l'exécution d'un flux de données au moment de la création. Cette option est uniquement disponible au niveau de l'appel du flux de données (lorsque le flux de données se trouve dans l'espace de travail Designer, par exemple). ! Restriction Vous devez créer des paramètres de substitution de contournement à utiliser avec l'option <i>Contournement</i>. Par exemple, dans la fenêtre <i>Editeur de paramètres de substitution</i>, vous pouvez créer \$\$BYPASSEnable avec la valeur Oui et \$\$BYPASSDisable avec la valeur Non (ou toute autre valeur différente de Oui). Une fois la création de votre job terminée, vous pouvez désactiver le contournement avant de le passer en mode de production. Pour en savoir plus, voir Contournement de workflows et flux de données spécifiques [page 762].

Informations associées

[Création d'un paramètre de substitution \[page 333\]](#)

18.3 Objets sources et cibles

Un flux de données lit et charge directement les données à l'aide d'objets source et cible.

Les objets source définissent les sources à partir desquelles vous lisez les données.

Les objets cible définissent les cibles vers lesquelles vous écrivez (ou chargez) les données.

Informations associées

[Objets sources \[page 207\]](#)

[Objets cibles \[page 207\]](#)

18.3.1 Objets sources

Les objets source contiennent les données à lire par le flux de données pour traitement.

Objet source	Description	Accès au logiciel
Table	Fichier formaté avec des colonnes et des lignes tel qu'utilisé dans les bases de données relationnelles.	Direct ou par adaptateur
Modèle de table	Une table de modèle créée et enregistrée dans un autre flux de données (utilisée pour le développement).	Direct
Fichier	Fichier plat délimité ou de largeur fixe.	Direct
Document	Fichier avec un format spécifique à l'application (non accessible en lecture par les analyseurs SQL ou XML).	Par adaptateur
Fichier JSON	Fichier mis en forme avec des données JSON.	Direct
Message JSON	Utilisé comme une source dans les jobs en temps réel.	Direct
Fichier XML	Fichier mis en forme avec des balises XML.	Direct
Message XML	Utilisé comme une source dans les jobs en temps réel.	Direct

Il est également possible d'utiliser des messages IDoc comme sources en temps réel pour les applications SAP.

Informations associées

[Tables de modèle \[page 212\]](#)

[Objets source et cible en temps réel \[page 352\]](#)

18.3.2 Objets cibles

Les objets cible reçoivent les données traitées qui sont chargées depuis le flux de données.

Objet cible	Description	Accès au logiciel
Document	Fichier avec un format spécifique à l'application (non accessible en lecture par les analyseurs SQL ou XML).	Par adaptateur
Fichier	Fichier plat délimité ou de largeur fixe.	Direct
Fichier JSON	Un fichier formaté dans le format JSON.	Direct
Message JSON	Voir Objets source et cible en temps réel [page 352] .	
Modèle de fichier imbriqué	Fichier JSON ou XML dont le format est basé sur la sortie de la transformation précédente (utilisé pour le développement, essentiellement pour déboguer les flux de données).	Direct
Message sortant	Voir Objets source et cible en temps réel [page 352] .	

Objet cible	Description	Accès au logiciel
Table	Fichier formaté avec des colonnes et des lignes tel qu'utilisé dans les bases de données relationnelles.	Direct ou par adaptateur
Modèle de table	Table dont le format est basé sur la sortie de la transformation précédente (utilisé pour le développement).	Direct
Fichier XML	Fichier mis en forme avec des balises XML.	Direct
Message XML	Voir Objets source et cible en temps réel [page 352] .	

Il est également possible d'utiliser des messages IDoc comme sources en temps réel pour les applications SAP.

18.3.2.1 Load Triggers tab in target editor

A load trigger operation is a template SQL statement that has placeholders for column and variable values.

Specify load trigger commands in the [Load Triggers](#) tab of the target editor in a data flow.

A load trigger specifies a SQL command that is performed by the database on an INSERT, UPDATE, or DELETE operation. Set the special operation in the load trigger to occur before, after, or instead of normal operations.

The software sets the placeholders at execution time based on the fields in the transform input schema. For each row, the software fills out the template and applies the operation against the target.

i Remarque

The software does not support load trigger options for Microsoft APS or Azure DW.

Use load triggers in situations such as when you archive updates to a data warehouse or archive incremental updates of aggregated values.

The software does not parse load triggers. Thus, when you specify a load trigger, the software does not parameterize SQL statements. As a result, load times might be higher when you use load triggers.

The software does not validate load triggers.

i Remarque

If you use an override, you cannot specify auto correct load.

❖ Exemple

For example, instead of applying an insert of a new sales order row, you use a load trigger that applies inserts and updates of aggregated values of `<sales_per_customer>` and `<sales_per_region>`.

The templates give you a row with `<customer_id>`, `<order_amount>`, `<region_id>`, and so on.

The INSERT and UPDATE statements are:

```
INSERT into order_fact
values ([customer_id], [order_amount]);
UPDATE region_fact
SET order_amount =
order_amount + [order_amount]
```

```
WHERE region_id = [region_id];
```

Enter your load triggers manually in the text box in the [Load Triggers](#) tab or drag column names from the [Schema In](#) pane. Enclose column names in curly braces or square brackets. For example, {SalesOffice} or [SalesOffice].

With curly braces, the software encloses the value in quotation marks, if needed. The software does not provide quotes around field names in square brackets. To avoid unintended results, use curly braces for varchar or char column names.

If you insert column names into the SQL statement by dragging the column names, the software inserts square brackets for you. If you require curly braces, manually change the square brackets to curly braces.

The default operations that you select from the [On operation](#) dropdown list are [#insert], [#update], and [#delete].

❖ Example

To delimit a SQL statement, use [#new]. For example:

```
[#insert] [#new]
insert into foo values ([col1], {col2}, ...)
```

Other guidelines for building the SQL statement include the following:

- To specify "before" images, add the suffix .before to the column name. To specify "after" images, add the suffix .after to the column name.
 - The default suffix for UPDATE and INSERT operations is .after.
 - The default suffix for DELETE operations is .before.
- For UPDATE operations, specify both the "before" and the "after" image values. You can specify both images for INSERT and DELETE operations, also, but it is not required.
- Include variables in the SQL statements, but not expressions.
- You can map a batch of SQL statements. Each SQL statement is separated by a new separator ([#new]). The following statement is an example for mapping insert SQL:

```
INSERT into log_table values ({col1}, {col2})
[#new]
[#insert] [#new]
delete from alt_junk where . . .
```

18.3.2.2 Setting DB2 10.5 for column based target table

When you select [Column store](#) for the DB2 option [Table Type](#), also complete specific prerequisite tasks.

A specific target set up option is [Table Type](#) where you can select either [Row store](#) or [Column store](#). This task takes you through the prerequisite steps for DB2 when you select [Column store](#). The software supports column organized tables only on Linux and AIX:

- Linux: x86-x64, Intel, and AMD processors
- AIX: POWER processors

There are two ways to enable column store in Linux and AIX:

- Default table organization: Set the registry variable DB2_WORKLOAD to ANALYTICS.
- Configure database for analytics workload: If your DB2_WORKLOAD cannot be set to ANALYTICS, create and optimally configure the database for analytics workload by following these steps.

1. Set the default table organization setting for user tables `dft_table_org` to COLUMN.

Set `dft_table_org` to column so that new tables are organized by column by default. If you do not set `dft_table_org` to column, you have to set the ORGANIZE BY COLUMN clause for each CREATE TABLE statement.

2. Set the default degree database configuration parameter to `dft_degree` to ANY.
3. Set the default extent size database configuration parameter `dft_extent_sz` to 4.
4. Ensure that the database configuration parameters for sort heap `sorheap` and sort heap threshold for shared sorts `sheapthres_shr` are **not** set to *AUTOMATIC*.

Consider increasing `sorheap` and `sheapthres_shr` values significantly for analytics workloads. A reasonable starting point is to set `sheapthres_shr` to the size of the buffer pool across all buffer pools. Set `sorheap` to a fraction, for example 1/20, of `sheapthres_shr` to enable concurrent sort operations.

5. Set the utility heap size database configuration parameter `util_heap_sz` to 1,000,000 pages and AUTOMATIC to address the resource needs of the LOAD command.
 - If the database server has at least 128 GB of memory, set `util_heap_sz` to 4,000,000 pages.
 - If concurrent load operations are running, increase the value of `util_heap_sz` to accommodate higher memory requirements.
6. Set the automatic reorganization database configuration parameter `auto_reorg` to ON.
7. Ensure that the database manager configuration parameter `sheapthres` to ON.

i Remarque

This setting applies to all databases in the instance.

8. Ensure that intraquery parallelism, which is required to access column-organized tables, is enabled.

Enable intraquery parallelism at the instance level database level. For details, see Intraquery parallelism and intrapartition parallelism in your DB2 documentation.

9. Enable concurrency control on the SYSDEFAULTMANAGEDSUBCLASS service subclass by issuing the following statement: `ALTER THRESHOLD SYSDEFAULTCONCURRENT ENABLE`.

Additional restrictions

Schemas that include column-organized tables cannot be transported.

For data replication with column-organized tables as either source or target, there are some sql statements that do not support column-organized tables as source and target tables:

- SET INTEGRITY
- CREATE TRIGGER
- CREATE EVENT MONITOR
- CREATE INDEX AND ALTER INDEX

18.3.3 Ajout d'objets sources ou cibles aux flux de données

Remplissez les conditions préalables suivantes avant d'utiliser un objet source ou cible dans un flux de données :

Pour	Connaissances requises
Tables accédées directement à partir d'une base de données	Définir une banque de données d'une base de données et importer les métadonnées de table.
Tables de modèle	Définir une banque de données d'une base de données.
Fichiers	Définir un format de fichier et importer le fichier.
Fichiers XML et messages	Importer un format de fichier XML.
Objets accédés par un adaptateur	Définir une banque de données d'adaptateur et importer des métadonnées d'objet.

1. Ouvrez le flux de données dans lequel vous souhaitez placer l'objet.
2. Si la bibliothèque d'objets n'est pas déjà ouverte, sélectionnez ► **Outils** ► **Bibliothèque d'objets** ► pour l'ouvrir.
3. Sélectionnez l'onglet de la bibliothèque d'objets approprié : sélectionnez l'onglet **Formats** pour les fichiers plats, les DTD, JSON ou les schémas XML, ou sélectionnez l'onglet **Banques de données** pour les bases de données et les objets adaptateurs.
4. Sélectionnez l'objet que vous souhaitez ajouter comme une cible ou une source. (Développez les listes réduites en cliquant sur le signe plus à côté d'une icône conteneur.)

Pour une nouvelle table de modèle, sélectionnez l'icône **Table de modèle** dans la palette d'outils.

Pour un nouveau fichier de modèle JSON ou XML, sélectionnez l'icône **Modèle de schémas imbriqués** dans la palette d'outils.

5. Déplacez l'objet dans l'espace de travail.
6. Pour les objets qui peuvent être soit sources soit cibles, un menu pop-up apparaît lorsque vous relâchez le curseur. Sélectionnez le type d'objet à réaliser.

Pour les nouvelles tables de modèle et les nouveaux fichiers de modèle XML, une deuxième fenêtre apparaît lorsque vous relâchez le curseur. Saisissez les informations requises pour le nouvel objet modèle. Les noms peuvent inclure des caractères alphanumériques et des tirets de soulignement (_). Les tables de modèle ne peuvent pas avoir le même nom qu'une table existante dans une banque de données.

7. L'objet source ou cible apparaît dans l'espace de travail.
8. Cliquez sur le nom de l'objet dans l'espace de travail

Le logiciel ouvre l'éditeur pour l'objet. Définissez les options requises pour l'objet.

i Remarque

Assurez-vous que les fichiers qui font référence aux formats des fichiers plats, DTD, JSON ou schémas XML sont accessibles depuis le Job Server où le job est exécuté et spécifiez l'emplacement du fichier relatif à cet ordinateur.

Informations associées

[Banques de données de bases de données \[page 83\]](#)

[Tables de modèle \[page 212\]](#)

[Formats de fichiers \[page 149\]](#)

[Importer un format de schéma XML ou DTD \[page 277\]](#)

[Ajout d'un objet d'emplacement de fichier à un flux de données \[page 197\]](#)

18.3.4 Tables de modèle

Lors de la conception initiale d'une application, il est possible que vous trouviez pratique d'utiliser des tables de modèle pour représenter les tables de base de données.

Avec les tables de modèle, vous ne devez pas créer au départ une nouvelle table dans votre système de gestion de base de données ni importer les métadonnées dans le logiciel. A la place, le logiciel crée automatiquement la table dans la base de données avec le schéma défini par le flux de données lors de l'exécution d'un job.

Après la création d'une table de modèle comme cible dans un flux de données, il est possible de l'utiliser comme une source dans d'autres flux de données. Bien qu'une table de modèle puisse être utilisée comme une table source dans plusieurs flux de données, elle peut être utilisée comme table cible dans un seul flux de données.

Les tables de modèle sont particulièrement utiles au début du développement de l'application lorsque vous concevez et testez un projet. Si vous modifiez et enregistrez l'opération de transformation des données dans le flux de données dans lequel la table de modèle est une cible, le schéma de la table de modèle change automatiquement. Toutes les mises à jour du schéma s'appliquent automatiquement aux autres instances de la table de modèle. Lors du processus de validation, le logiciel vous avertit de toutes les erreurs telles que celles qui découlent de la modification du schéma.

18.3.4.1 Créer une table de modèle cible

1. Utilisez l'une des méthodes suivantes pour ouvrir la fenêtre [Créer un modèle](#) :
 - Depuis la palette d'outils :
 - Cliquez sur l'icône de table de modèle. 
 - Cliquez dans un flux de données pour placer la table de modèle dans l'espace de travail.
 - Dans la fenêtre [Créer un modèle](#), sélectionnez une banque de données.
 - Depuis la bibliothèque d'objets :
 - Développez une banque de données.
 - Cliquez sur l'icône de table de modèle et faites-la glisser dans l'espace de travail.
2. Dans la fenêtre [Créer un modèle](#), saisissez un nom de table.

Si vous utilisez Netezza 7.x, vous pouvez entrer un nom de schéma dans la zone de schéma pour limiter les modèles de table à un schéma spécifique. Si vous ne renseignez pas le nom de schéma, les modèles de table sont limités au schéma par défaut.

3. Cliquez sur [OK](#).

La table apparaît dans l'espace de travail comme une icône de table de modèle.

4. Connectez la table de modèle au flux de données comme cible (généralement une transformation Query).
5. Dans la transformation Query, mappez les colonnes Entrée de schéma que vous souhaitez inclure dans la table cible.
6. Depuis le menu [Projet](#), sélectionnez [Enregistrer](#).

Dans l'espace de travail, l'icône de la table de modèle est remplacée par une icône de table cible et la table apparaît dans la bibliothèque d'objets sous la liste de la banque de données des tables.

Une fois que vous êtes satisfait de la conception du flux de données, enregistrez-le. Lorsque le job est exécuté, le logiciel utilise la le modèle de table pour créer une table dans la base de données que vous avez indiquée lors de la création du modèle de table. Une fois qu'une table de modèle est créée dans la base de données, il est possible de convertir la table de modèle du référentiel en une table régulière.

18.3.5 Conversion des tables de modèle en tables régulières

Vous devez convertir les tables de modèle en tables régulières pour profiter de certaines fonctionnalités comme le chargement par lots.

D'autres fonctionnalités, telles que l'exportation d'un objet, sont disponibles pour les tables de modèle.

i Remarque

Une fois qu'une table de modèle est convertie, il n'est plus possible de modifier le schéma.

18.3.5.1 Convertir un modèle de table en une table régulière depuis la bibliothèque d'objets

1. Ouvrez la bibliothèque d'objets et accédez à l'onglet [Banques de données](#).
2. Cliquez sur le signe plus (+) à côté de la banque de données qui contient la table de modèle que vous souhaitez convertir.

Une liste d'objets apparaît.

3. Cliquez sur le signe plus (+) à coté de [Table de modèle](#).

La liste des tables de modèle apparaît.

4. Faites un clic droit sur la table de modèle que vous souhaitez convertir et sélectionnez [Importer la table](#).

Le logiciel convertit la table de données dans le référentiel en une table régulière en l'important de la base de données. Pour mettre à jour l'icône dans tous les flux de données, sélectionnez ► [Afficher](#) ► [Actualiser](#) ►. Dans la bibliothèque d'objets de la banque de données, la table est désormais listée sous Tables au lieu de Tables de modèle.

18.3.5.2 Convertir un modèle de table en une table régulière depuis un flux de données

1. Ouvrez le flux de données qui contient la table de modèle
2. Faites un clic droit sur la table de modèle que vous souhaitez convertir et sélectionnez *Importer la table*.

Une fois qu'une table de modèle est convertie en une table régulière, il n'est plus possible de modifier le schéma de la table.

18.4 Comprendre la propagation de colonnes

Vous pouvez utiliser *Propager la colonne à partir de* dans un flux de données pour ajouter une colonne existante à partir d'une source ou d'une transformation en amont par le biais d'objets intermédiaires vers un point d'arrivée sélectionné.

Les colonnes sont ajoutées dans chaque objet sans aucune modification sur le type de données ou les autres attributs. Lorsqu'il existe plus d'un chemin possible entre le point de départ et le point d'arrivée, il est possible d'indiquer le chemin pour les colonnes ajoutées.

La propagation de colonnes est une opération de tirage. La commande *Propager la colonne à partir de* est émise à partir de l'objet dans lequel la colonne est nécessaire. La colonne est extraite de la source ou transformation en amont sélectionnée et ajoutée à chaque objet intermédiaire ainsi qu'à l'objet du point d'arrivée sélectionné.

Par exemple, dans le flux de données ci-dessous, la table source `Employee` contient les informations de nom des employés ainsi que l'ID d'employé, l'information sur le travail et les dates d'embauche. La transformation `Name_Cleanse` est utilisée pour normaliser les noms d'employés. Enfin, les données sont sorties dans un fichier XML appelé `Noms_Employés`.



Après avoir vu la sortie dans la table `Employee_Names`, vous vous rendez compte que l'initiale du milieu (colonne `minit`) doit être incluse dans la sortie. Faites un clic droit sur le schéma de niveau supérieur de la table `Employee_Names` et sélectionnez *Propager la colonne à partir de*. La fenêtre *Propager la colonne vers Noms_Employés* apparaît.

Dans le volet gauche de la fenêtre *Propager la colonne vers Employee_Names*, sélectionnez la table source `Employee` depuis la liste d'objets. La liste des colonnes de sortie affichées dans le volet droit change pour afficher les colonnes dans le schéma de l'objet sélectionné. Sélectionnez la colonne `MINIT` comme la colonne que vous souhaitez tirer à partir de la source, puis cliquez sur *Propager*.

Le schéma de la colonne `minit` est transporté par les transformations `Query` et `Name_Cleanse` vers la table `Employee_Names`.

Les caractéristiques des colonnes propagées sont les suivantes :

- La commande [Propager colonne à partir de](#) peut être émise à partir du schéma de niveau supérieur d'une transformation ou d'une cible.
- Les colonnes sont ajoutées dans chaque objet sans aucune modification sur le type de données ou les autres attributs. Une fois qu'une colonne est ajoutée au schéma d'un objet, la colonne fonctionne exactement de la même façon que si elle avait été créée manuellement.
- La colonne propagée est ajoutée à la fin de la liste de schéma dans chaque objet.
- Le nom de la colonne de sortie est généré automatiquement pour éviter des conflits de nom avec les colonnes existantes. Il est possible de modifier le nom de colonne.
- Seules les colonnes incluses dans les schémas de niveau supérieur peuvent être propagées. Les colonnes dans les schémas imbriquées ne peuvent pas être propagées.
- Une colonne peut être propagée plus d'une fois. Toutes les colonnes existantes sont affichées dans le volet droit de la fenêtre [Propager la colonne vers](#) dans le champ [Existe déjà dans](#). Chaque colonne supplémentaire a un nom unique.
- Plusieurs colonnes peuvent être sélectionnées et propagées en même temps.

i Remarque

Il est impossible de propager une colonne avec une transformation `Hierarchy_Flattening` ou une transformation `Table_Comparison`.

18.4.1 Ajout de colonnes dans un flux de données

Vous pouvez ajouter une colonne à partir d'une source ou d'une transformation en amont par le biais d'objets intermédiaires vers un point d'arrivée sélectionné à l'aide de la commande Propager.

Les colonnes sont ajoutées dans chaque objet sans aucune modification sur le type de données ou les autres attributs.

Ajouter des colonnes dans un flux de données :

1. Dans l'objet en aval dans lequel vous souhaitez ajouter la colonne (le point d'arrivée), faites un clic droit sur le schéma de niveau supérieur et cliquez sur [Propager la colonne à partir de](#).
[Propager la colonne à partir de](#) peut être émis à partir du schéma de niveau supérieur dans une transformation ou un objet cible.
2. Dans le volet gauche de la fenêtre [Propager la colonne vers](#), sélectionnez l'objet en amont qui contient la colonne que vous souhaitez mapper.
Les colonnes disponibles dans cet objet sont affichées dans le volet droit avec une liste des mappages existants depuis cette colonne.
3. Dans le volet droit, sélectionnez la colonne que vous souhaitez ajouter et cliquez soit sur [Propager](#) soit sur [Propager et fermer](#).
L'une des opérations suivantes se produit :
 - S'il n'existe qu'un seul chemin possible, la colonne sélectionnée est ajoutée par les transformations intermédiaires à l'objet en aval.
 - S'il existe plus d'un chemin possible par les objets intermédiaires, la boîte de dialogue [Choisir le routage vers](#) apparaît. Cela peut se produire lorsque le flux de données contient une transformation Query avec plusieurs objets d'entrée. Sélectionnez le chemin que vous préférez et cliquez sur [OK](#).

18.4.2 Propager les colonnes dans un flux de données contenant une transformation Merge

La propagation permet d'apporter des modifications à la colonne en aval.

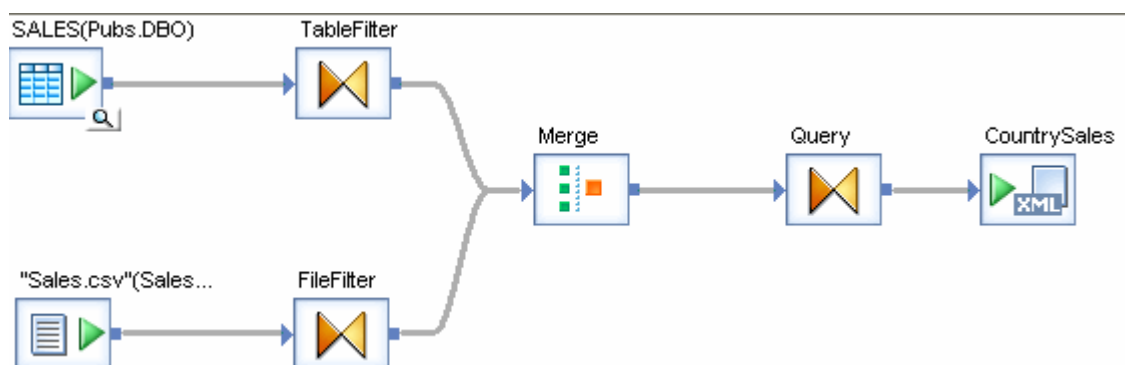
Dans un flux de données valide qui contient deux sources ou plus qui sont fusionnées à l'aide d'une transformation Merge, le schéma des entrées dans la transformation Merge doit être identique. Toutes les sources doivent avoir le même schéma, y compris :

- le même nombre de colonnes
- les mêmes noms des colonnes
- les colonnes semblables doivent avoir le même type de données

Afin de conserver un flux de données valide lors de la propagation d'une colonne par une transformation Merge, vous devez vous assurer de respecter cette restriction.

Lorsque vous propagez une colonne et qu'une transformation Merge se trouve entre le point de départ et le point d'arrivée, un message vous alerte qu'une fois l'opération de propagation terminée, le flux de données sera non valide puisque les schémas d'entrée dans la transformation Merge ne seront pas identiques. Si vous décidez de poursuivre l'opération de propagation de la colonne, vous devrez ajouter plus tard des colonnes aux schémas d'entrée dans la transformation Merge pour que le flux de données soit valide.

Par exemple, dans le flux de données ci-dessous, les données de chaque table source sont filtrées, puis les résultats sont fusionnés dans la transformation Merge.



Si vous choisissez de propager une colonne à partir de la source `SALES (Pubs . DBO)` vers la cible `CountrySales`, la colonne est ajoutée au schéma `TableFilter` mais pas au schéma `FileFilter`, donnant lieu à des schémas d'entrée différents dans la transformation `Merge` et un flux de données non valide.

Afin de conserver un flux de données valide lors de la propagation d'une colonne par une transformation `Merge`, suivez un processus à plusieurs étapes :

1. Assurez-vous que la colonne que vous souhaitez propager est disponible dans les schémas de tous les objets qui mènent dans la transformation `Merge` sur le côté en aval. Cela garantit que toutes les entrées dans la transformation `Merge` sont identiques et que le flux de données est valide.
2. Propagez la colonne sur le côté en amont de la transformation `Merge` vers le point d'arrivée souhaité.

18.5 Tables de recherche et fonction lookup_ext

Les tables de recherche contiennent des données auxquelles d'autres tables font référence.

En règle générale, les tables de recherche peuvent comporter les types de colonnes suivants :

- Colonne de recherche : permet d'établir une correspondance avec une ou plusieurs lignes en fonction des valeurs d'entrée. Pour identifier une correspondance dans une ligne, vous appliquez des opérateurs tels que =, >, <, ~. Une table de recherche peut contenir plusieurs colonnes de recherche.
- Colonne de sortie : colonne renvoyée par la ligne, qui correspond à la condition de recherche définie pour la colonne de recherche. Une table de recherche peut contenir plusieurs colonnes de sortie.
- Colonne de politique de retour : permet d'indiquer les données à renvoyer lorsque plusieurs lignes correspondent à la ou aux conditions de recherche.

Utilisez la fonction lookup_ext pour extraire des données d'une table de recherche en fonction de conditions de recherche définies par l'utilisateur pour établir une correspondance entre les données d'entrée et les données de la table de recherche. La fonction lookup_ext permet non seulement d'extraire une valeur dans une table ou un fichier en fonction des valeurs d'une table source ou d'un fichier différent, mais elle propose aussi des fonctionnalités étendues permettant d'effectuer les opérations suivantes :

- Retourner plusieurs colonnes à partir d'une même recherche.
- Effectuer une sélection parmi d'autres opérateurs, y compris les critères spéciaux, pour définir une condition de recherche.
- Indiquer une politique de retour pour la recherche.
- Appeler lookup_ext dans des scripts et des fonctions personnalisées (ce qui vous permet également de réutiliser la ou les recherches regroupées dans des scripts).
- Définir un SQL personnalisé à l'aide du paramètre SQL_override pour alimenter le cache de recherche et parvenir à affiner d'importantes quantités de données aux seules sections pertinentes pour la recherche.
- Appeler lookup_ext à l'aide de l'assistant de fonction dans le mappage de sortie de requête pour retourner plusieurs colonnes dans une transformation Query.
- Sélectionner une stratégie de mise en cache et, par exemple, décider de placer dans la mémoire cache l'intégralité de la table de recherche ou de générer des instructions SQL de façon dynamique pour chaque enregistrement d'entrée.
- Utiliser lookup_ext avec les tables de banque de données en mémoire ou les tables du cache persistant. En ce qui concerne les tables de recherche, l'utilisation du cache persistant présente les avantages suivants par rapport aux tables en mémoire :
 - Plusieurs flux de données peuvent utiliser la même table de recherche disponible dans le cache persistant.
 - Le progiciel n'a pas besoin de construire la table de recherche chaque fois qu'un flux de données l'utilise.
 - Le cache persistant n'a aucune contrainte de mémoire car il est stocké sur disque et le progiciel exécute rapidement la pagination dans la mémoire.
- Utiliser le cache paginable (qui n'est pas disponible pour les fonctions lookup et lookup_seq).
- Utiliser des expressions dans les tables de recherche et retourner les valeurs obtenues.

Pour une description des fonctions lookup et lookup_seq associées, voir le *Guide de référence*.

18.5.1 Accès à l'éditeur lookup_ext

La fonction lookup_ext possède son propre éditeur graphique.

Il existe deux façons d'appeler l'éditeur :

- en ajoutant un appel de fonction à l'intérieur d'une transformation Query (utilisez cette option si vous souhaitez que la table de recherche retourne plusieurs colonnes) ;
- à partir de l'onglet Mappage dans une requête ou une fonction de script.

18.5.1.1 Ajout d'un appel de fonction

1. Dans la transformation Query, au niveau du volet *Schéma de sortie*, sans sélectionner de colonne de sortie spécifique, cliquez avec le bouton droit sur le volet et sélectionnez *Nouvel appel de fonction*.
2. Sélectionnez la catégorie de fonction *Fonctions de recherche* et le nom de fonction *lookup_ext*.
3. Pour appeler l'éditeur, cliquez sur *Suivant*.

La section Sortie vous permet d'ajouter plusieurs colonnes au schéma de sortie.

Une fois que vous avez fermé la fenêtre lookup_ext, vous pouvez rouvrir l'éditeur graphique pour apporter des modifications (cliquez avec le bouton droit sur le nom de la fonction dans le schéma et sélectionnez *Modifier l'appel de fonction*), ce qui constitue un avantage du nouvel appel de fonction.

18.5.1.2 Appel de l'éditeur lookup_ext depuis l'onglet Mappage

1. Sélectionnez le nom de la colonne de sortie.
2. Dans l'onglet *Mappage*, cliquez sur *Fonctions*.
3. Sélectionnez la *catégorie de fonction* *Fonctions de recherche* et le *nom de fonction* *lookup_ext*.
4. Pour appeler l'éditeur, cliquez sur *Suivant*.

Dans la section Sortie, le champ *Variable* remplace le champ *Nom de la colonne de sortie*. Vous pouvez définir une colonne de sortie destinée à alimenter la colonne sélectionnée dans le schéma de sortie. Lorsque lookup_ext retourne plusieurs colonnes de sortie, utilisez des variables pour stocker les valeurs de sortie ou utilisez lookup_ext comme nouvel appel de fonction comme indiqué précédemment dans cette section.

L'éditeur graphique n'est pas disponible avec les fonctions utilisées dans les mappages. En revanche, vous pouvez traiter le texte manuellement dans l'onglet *Mappage*.

18.5.2 Exemple : Définition d'une fonction lookup_ext simple

Cette procédure décrit la définition d'une fonction lookup_ext simple à l'aide d'un nouvel appel de fonction. L'exemple associé illustre l'utilisation d'une table de recherche pour extraire les noms des services des salariés.

Pour en savoir plus sur les options disponibles pour la fonction lookup_ext, voir le *Guide de référence*.

1. Dans un flux de données, ouvrez l'éditeur de requêtes.
2. Dans le volet *Entrée de schéma*, faites glisser la colonne d'ID vers le volet *Sortie de schéma*.
3. Sélectionnez la colonne ID dans le volet *Sortie de schéma*, cliquez avec le bouton droit et sélectionnez *Nouvel appel de fonction*. Cliquez sur *Insérer en dessous*.
4. Sélectionnez la *catégorie de fonction Fonctions de recherche* et le *nom de fonction lookup_ext*, puis cliquez sur *Suivant*.

L'éditeur lookup_ext apparaît.

5. Dans la fenêtre *Lookup_ext - Sélectionner les paramètres*, sélectionnez une table de recherche :
 - a. En regard de la zone de texte *Table de recherche*, cliquez sur la flèche de liste déroulante puis cliquez deux fois sur la banque de données, le format de fichier ou le schéma actuel comprenant la table.
 - b. Sélectionnez la table de recherche et cliquez sur *OK*.

Dans l'exemple, la table de recherche est un fichier au format txt appelé ID_lookup.txt, qui réside dans D:\Data.

6. Pour la *spécification du cache*, la valeur par défaut de CACHE_DE_PRECHARGEMENT est utile lorsque le nombre de lignes de la table est peu élevé ou que vous pensez accéder à un fort pourcentage des valeurs de la table.

AUCUN_CACHE lit les valeurs de la table de recherche pour chaque ligne sans valeurs mises en cache.

Sélectionnez CACHE_DE_CHARGEMENT_A_LA_DEMANDE lorsque la table comporte un grand nombre de lignes et que vous pensez accéder fréquemment à un faible pourcentage de ses valeurs ou lorsque vous utilisez la table dans plusieurs recherches avec des conditions de comparaison extrêmement sélectives, qui donnent un petit sous-ensemble de données.

7. Pour disposer de ressources supplémentaires en vue de l'exécution de la fonction lookup_ext, sélectionnez *Exécuter en tant que processus séparé*. Cette option crée un processus de flux de données enfant séparé pour la fonction lookup_ext lorsque le progiciel exécute le flux de données.
8. Définissez une ou plusieurs conditions. Pour chacune, ajoutez un nom de colonne de table de recherche (effectuez une sélection dans la liste déroulante ou faites glisser des éléments depuis le volet *Paramètre*), sélectionnez l'opérateur pertinent et entrez une expression en la saisissant, la faisant glisser, la collant ou à l'aide de l'éditeur rapide (cliquez sur l'icône dans la colonne de droite).

Dans l'exemple, la condition correspond à `ID_DEPT = Employees.ID_DEPT`.

9. Définissez la sortie. Pour chaque colonne de sortie :
 - a. Ajoutez un nom de colonne de la table de recherche.
 - b. Le cas échéant, modifiez la valeur par défaut (valeur nulle).
 - c. Indiquez le *nom de la colonne de sortie* en la saisissant, la faisant glisser, la collant ou à l'aide de l'éditeur rapide (cliquez sur l'icône dans la colonne de droite).

Dans l'exemple, la colonne de sortie est ID_DEPT_NAME.

10. Si plusieurs correspondances sont possibles, indiquez l'ordre et définissez une politique de retour (*MAX* par défaut) pour la sélection d'une correspondance. Pour ordonner la sortie, saisissez le ou les noms de colonne dans la liste *Organiser par*.

❖ Exemple

L'exemple suivant illustre le mode d'utilisation de la table de recherche ID_lookup.txt pour extraire les noms des services des salariés.

La table des salariés (Employees) se présente comme suit :

ID	NAME	ID_DEPT
SSN111111111	Employee1	10
SSN222222222	Employee2	10
TAXID333333333	Employee3	20

La table de recherche ID_lookup.txt se présente comme suit :

ID_DEPT	ID_PATTERN	ID_RETURN	ID_DEPT_NAME
10	ms(SSN*)	=substr(ID_Pattern,4,20)	Payroll
20	ms(TAXID*)	=substr(ID_Pattern,6,30)	Comptabilité

L'éditeur lookup_ext est alors configuré comme suit.

Lookup table: Cache spec: ☐ Run as a separate process

Available parameters:

- Parameter
 - Lookup Table
 - "D:\\Data\\ID_lookup.txt"(Lookup
 - ID_DEPT
 - ID_PATTERN
 - ID_RETURN
 - ID_DEPT_NAME
 - Input Schema
 - Employees
 - ID
 - NAME
 - ID_DEPT
 - Variables

Condition:

Column in Lookup table	Op.(%)	Expression
ID_DEPT	=	Employees.ID_DEPT
*		

Output:

Column in Lookup table	Expression?	Default	Output column name
ID_DEPT_NAME	☒	NULL	ID_DEPT_NAME
*	☐	NULL	

Order by:

Column in Lookup table
*

Return policy:

(Optional) If multiple rows satisfy the lookup conditions, they are ordered by the list of columns here, and the row that satisfies the "Return policy" is returned.

Custom SQL

< Back Finish Cancel

Informations associées

Exemple : Définition d'une fonction lookup_ext complexe [page 221]

18.5.3 Exemple : Définition d'une fonction lookup_ext complexe

Cette procédure décrit la définition d'une fonction lookup_ext complexe à l'aide d'un nouvel appel de fonction. L'exemple associé utilise les mêmes tables de recherche et d'entrée que [Exemple : Définition d'une fonction lookup_ext simple \[page 218\]](#). Cet exemple illustre l'extraction et la normalisation des numéros d'identification des salariés.

Pour en savoir plus sur les options disponibles pour la fonction lookup_ext, voir le *Guide de référence*.

1. Dans un flux de données, ouvrez l'éditeur de requêtes.
2. Dans le volet *Entrée de schéma*, faites glisser la colonne d'ID vers le volet *Sortie de schéma*. Procédez de même pour la colonne Nom.
3. Dans le volet *Sortie de schéma*, cliquez avec le bouton droit sur la colonne Nom et cliquez sur *Nouvel appel de fonction*. Cliquez sur *Insérer en dessous*.
4. Sélectionnez la *Catégorie de fonction Fonctions de recherche* et le *Nom de la fonction lookup_ext*, puis cliquez sur *Suivant*.
5. Dans la fenêtre *Lookup_ext - Sélectionner les paramètres*, sélectionnez une table de recherche :
Dans l'exemple, la table de recherche (au format txt) est nommée ID_lookup.txt et réside dans D:\Data.
6. Définissez une ou plusieurs conditions.
Dans l'exemple, la condition correspond à `ID_PATTERN ~ Employees.ID`.
7. Définissez la sortie. Pour chaque colonne de sortie :
 - a. Ajoutez un nom de colonne de la table de recherche.
 - b. Si vous voulez que le progiciel interprète la colonne dans la table de recherche comme une expression et retourne la valeur calculée, cochez la case *Expression*.
 - c. Le cas échéant, modifiez la valeur par défaut (valeur nulle).
 - d. Indiquez le *nom de la ou des colonnes de sortie* en le saisissant, en le faisant glisser ou en le collant ou à l'aide de l'éditeur rapide (cliquez sur l'icône dans la colonne de droite).

Dans l'exemple, les colonnes de sortie sont ID_RETURN et ID_DEPT_NAME.

❖ Exemple

Extraction et normalisation des numéros d'identification des salariés

Dans cet exemple, vous voulez extraire et normaliser des numéros de sécurité sociale et d'identification fiscale de salariés, avec des préfixes différents. Vous allez supprimer les préfixes afin de normaliser les numéros. Vous voulez également identifier le service d'origine du numéro. Le flux de données comporte une table source Employees, une requête configurée avec look_ext et une table cible.

Configurez l'éditeur lookup_ext comme dans le graphique suivant.

Lookup_ext - Sélectionner les paramètres

Table de recherche : Spéc. du cache : ☐ Exécuter en tant que processus séparé

Paramètres disponibles :

- Table de recherche
 - "D:\Data\ID_lookup.txt"(Lookup)
 - ID_DEPT
 - ID_PATTERN
 - ID_RETURN
 - ID_DEPT_NAME
 - Schéma d'entrée
 - Employees
 - ID
 - NAME
 - ID_DEPT
- Variables

Condition :

Colonne de table de recherche	Op.(&)	Expression
ID_PATTERN	~	Employees.ID
*		

Pour démarrer, sélectionnez une table de recherche, puis créez une ou plusieurs conditions à rechercher dans la table de recherche. Par exemple : "Table recherche IDemployé = Schéma entrée IDemployé"

Sortie :

Colonne de table de recherche	Expression ?	Par défaut	Nom de la colonne
ID_RETURN	☒	NULL	ID_RETURN
ID_DEPT_NAME	☐	NULL	ID_DEPT_NAME

Spécifiez d'autres colonnes dans la table de recherche à renvoyer lorsque les conditions sont remplies. Exemple : "Table de recherche ?" si

Trier par :

Colonne de table de recherche
*

Stratégie de renvoi :

(Facultatif) Si plusieurs lignes remplissent les conditions de recherche, elles sont classées par la liste de colonnes ici et la ligne qui

SQL personnalisé

< Précédent Terminer Annuler

La condition de recherche correspond à `ID_PATTERN ~ Employees.ID`.

Le progiciel lit les différentes lignes de la table source Employees, puis passe à la table de recherche ID_lookup.txt pour identifier toutes les lignes qui répondent à la condition de recherche.

L'opérateur ~ indique que le logiciel appliquera une comparaison de modèles à Employees.ID. Lorsqu'il rencontre un modèle correspondant à Employees.ID dans ID_lookup.ID_PATTERN, le logiciel applique l'expression dans ID_lookup.ID_RETURN. Dans cet exemple, Employee1 et Employee2 ont tous deux des ID qui correspondent au modèle ms(SSN*) dans la table de recherche. Le logiciel applique alors l'expression `=substr(ID_PATTERN, 4, 20)` aux données, ce qui permet d'extraire de la chaîne de correspondance (Employees.ID) une sous-chaîne de 20 caractères maximum en partant de la quatrième position. Les résultats pour Employee1 et Employee2 sont respectivement de 11111111 et 22222222.

Pour la sortie de la colonne de recherche ID_RETURN, le progiciel évalue ID_RETURN comme une expression car la case **Expression** est cochée. Dans la table de recherche, la colonne ID_RETURN contient l'expression `=substr(ID_PATTERN, 4, 20)`. Dans cette expression, ID_PATTERN se rapporte à la colonne ID_PATTERN de la table de recherche. Lorsque la condition de recherche `ID_PATTERN ~ Employees.ID` est vraie, le progiciel évalue l'expression. Dans le cas présent, le progiciel remplace le caractère générique ID_PATTERN par la véritable valeur de Employees.ID.

La sortie comprend également la colonne ID_DEPT_NAME, que le logiciel retourne comme une valeur littérale (du fait que la case Expression n'est pas cochée). La table cible obtenue se présente comme suit :

ID	NAME	ID_RETURN	ID_DEPT_NAME
SSN11111111	Employee1	11111111	Payroll
SSN22222222	Employee2	22222222	Payroll
TAXID33333333	Employee3	33333333	Accounting

Informations associées

[Exemple : Définition d'une fonction lookup_ext simple \[page 218\]](#)

[Accès à l'éditeur lookup_ext \[page 218\]](#)

18.6 Exécution d'un flux de données

Un flux de données est une spécification déclarative à partir de laquelle le logiciel détermine les bonnes données à traiter.

Par exemple, dans des flux de données placés dans des jobs batch, l'ordre de la transaction est d'extraire, de transformer, puis de charger les données dans une cible. Les flux de données sont similaires aux instructions SQL. La spécification déclare la sortie souhaitée.

Le logiciel exécute un flux de données à chaque fois qu'un flux de données se produit dans un job. Cependant, il est possible de spécifier qu'un job batch exécute un flux de données particulier une seule fois. Dans ce cas, le logiciel exécute uniquement la première occurrence du flux de données ; le logiciel ignore les occurrences suivantes dans le job.

Vous pouvez utiliser cette fonctionnalité lorsque vous développez des jobs batch complexes avec plusieurs chemins, tels que les jobs avec des blocs Try/Catch ou des structures conditionnelles, et que vous souhaitez vous assurer que le logiciel n'exécute qu'une seule fois un flux de données particulier.

Informations associées

[Création et définition des flux de données \[page 204\]](#)

18.6.1 Poussée vers le bas des opérations vers le serveur de base de données

À partir des informations contenues dans la spécification du flux de données, le logiciel produit une sortie tout en optimisant la performance.

Par exemple, pour les sources et les cibles SQL, le logiciel crée des instructions SQL spécifiques à la base de données en fonction des diagrammes du flux de données du job. Pour optimiser la performance, le logiciel pousse vers le bas autant d'opérations de transformation que possibles vers la base de donnée source ou cible, et combine autant d'opérations que possible dans une requête vers la base de données. Par exemple, le logiciel essaye de pousser vers le bas les jointures et les évaluations de fonction. En poussant les opérations vers la base de données, le logiciel réduit le nombre de lignes et d'opérations que le moteur doit traiter.

La conception du flux de données influence le nombre d'opérations que le logiciel peut pousser vers la base de données sources ou cibles. Avant d'exécuter un job, il est possible d'examiner le SQL généré par le logiciel et de modifier la conception pour aboutir aux résultats les plus efficaces.

Il est possible d'utiliser la transformation Data_Transfer pour pousser vers le bas des opérations consommatrices de ressources n'importe où dans un flux de données vers la base de données. Les opérations consommatrices de ressources incluent les jointures, GROUP BY, ORDER BY et DISTINCT.

18.6.2 Exécution de flux de données distribués

Le logiciel offre la possibilité de distribuer les travaux de traitement des données consommatrices de mémoire et de processeurs (tels que les jointures, les regroupements, les comparaisons de table et les recherches) dans les multiples processus et ordinateurs.

Ce travail de distribution offre les bénéfices potentiels suivants :

- Une meilleure gestion de la mémoire en tirant profit de plus de ressources processeurs et de mémoire physique
- Une meilleure performance et extensibilité de job en utilisant une exécution de flux de données secondaire simultanée pour tirer profit du grid computing

Il est possible de créer un flux de données secondaire pour que le logiciel n'ait pas besoin de traiter l'ensemble du flux de données en mémoire en une seule fois. Il est également possible de distribuer les flux de données secondaires aux différents job servers d'un groupe de serveurs pour utiliser plus de mémoire et des ressources processeurs supplémentaires.

Utilisez les fonctionnalités suivantes pour diviser un flux de données en plusieurs flux de données secondaires

- L'option *Exécuter en tant que processus séparé* sur les opérations consommatrices de ressources qui incluent :
 - Transformation Hierarchy_Flattening
 - Transformation Associate
 - Transformation Country ID
 - Transformation Global Address Cleanse
 - Transformation Global Suggestion Lists
 - Match Transform

- Transformation United States Regulatory Address Cleanse
- Transformation User-Defined
- Les opérations de requête qui sont consommatrices de processeurs et de mémoire :
 - Jointures
 - GROUP BY
 - ORDER BY
 - DISTINCT
- Transformation Table_Comparison
- Fonction Lookup_ext
- Fonction Count_distinct
- Fonction Search_replace

Si vous sélectionnez l'option Exécuter en tant que processus séparé pour plusieurs opérations dans un flux de données, le logiciel divise le flux de données en des flux de données secondaires plus petits qui utilisent des ressources différentes (mémoire et ordinateur) les uns des autres. Lorsque vous indiquez plusieurs options Exécuter en tant que processus séparé, les processus de flux de données secondaire exécutent en parallèle.

- Transformation Data_Transfer
Avec cette transformation, le logiciel n'a pas besoin de traiter l'ensemble du flux de données sur l'ordinateur du Job Server. À la place, la transformation Data_Transfer peut pousser vers le bas le traitement d'une opération consommatrice de ressource vers le serveur de base de données. Cette transformation divise le flux de données en deux flux de données secondaires et transfère les données vers une table dans le serveur de base de données pour permettre au logiciel de pousser vers le bas l'opération.

18.6.3 Equilibrage de charge

Il est possible de distribuer l'exécution d'un job ou une partie d'un job dans de multiples Job Servers dans un groupe de serveurs pour mieux équilibrer les opérations consommatrices de ressources.

Vous pouvez spécifier les valeurs suivantes dans l'option *Niveau de distribution* lors de l'exécution d'un job :

- Niveau de job : un job peut s'exécuter sur un Job Server disponible.
- Niveau de flux de données : chaque flux de données dans un job peut s'exécuter sur un Job Server disponible.
- Sous-niveau de flux de données : une opération consommatrice de ressources (comme le tri, la comparaison de table ou la recherche de table) dans un flux de données peut s'exécuter sur un Job Serveur disponible.

18.6.4 Caches

Le logiciel offre l'option de cacher les données dans la mémoire pour améliorer les opérations dans les flux de données telles que :

Opération	Description
Jointures	Puisqu'une source interne d'une jointure doit être lue pour chaque ligne d'une source externe, il est possible que vous souhaitiez mettre en cache une source lorsqu'elle est utilisée comme source interne dans une jointure.
Comparaisons de tables	Puisqu'une table de comparaison doit être lue pour chaque ligne d'une source, il est possible que vous souhaitiez mettre en cache la table de comparaison.
Recherches	Puisqu'une table de recherche peut exister dans une base de données distante, il est possible que vous souhaitiez la mettre en cache dans la mémoire pour réduire les temps d'accès.

Le logiciel offre les types de caches suivants qui peuvent être utilisés par le flux de données pour toutes les opérations qu'il contient :

Cache	Description
En mémoire	Utilisez le cache en mémoire lorsque le flux de données traite une petite quantité de données qui peuvent être contenues dans la mémoire.
Cache paginable	Utilisez le cache paginable lorsque le flux de données traite une très grande quantité de données qui ne peuvent pas être contenues dans la mémoire.

Si vous divisez le flux de données en flux de données secondaires qui exécutent chacun sur un Job Server différent, chaque flux de données secondaires peut utiliser son propre type de cache.

18.7 Vue d'ensemble d'un audit de flux de données.

Il est possible d'auditer des objets dans un flux de données pour collecter les statistiques d'audit d'exécution.

Il est possible d'effectuer les tâches suivantes avec cette fonctionnalité d'audit :

- Collecter des statistiques d'audit sur les données lues dans un job, traitées par plusieurs transformations et chargées dans des cibles.
- Définir des règles à propos des statistiques d'audit pour déterminer si les données correctes sont traitées.
- Générer une notification en cas d'échec d'audit.
- Faire une requête pour les statistiques d'audit qui persistent dans le référentiel.

Informations associées

[Utilisation de l'Audit \[page 412\]](#)

19 Transformations

Les transformations opèrent sur les jeux de données en manipulant les jeux d'entrées et en produisant un ou plusieurs jeux de sortie.

A l'inverse, les fonctions opèrent sur les valeurs uniques dans des colonnes spécifiques dans un jeu de données. De nombreuses transformations intégrées sont disponibles depuis la bibliothèque d'objets sur l'onglet [Transformation](#).

Les transformations que vous pouvez utiliser dépend du package logiciel que vous avez acheté. (Si une transformation appartient à un package que vous n'avez pas acheté, elle est désactivée et ne peut pas être utilisée dans un job.)

Le logiciel présente les types de transformation suivants :

- **Data Integrator** : Transformations qui extraient, transforment et chargent les données. Ces transformations contribuent à assurer l'intégrité des données et à optimiser la productivité du développeur pour le chargement et la mise à jour des données dans un environnement d'entrepôt.
- **Data Quality** : Transformations permettant d'améliorer la qualité de vos données. Ces transformations peuvent analyser, corriger, enrichir, mettre en correspondance et consolider vos biens personnalisés et d'informations opérationnelles.
- **Plateformes** : Transformations qui effectuent des opérations de mouvements de données générales. Ces transformations vous permettent de générer, mapper et fusionner des lignes à partir de deux sources ou plus, de créer des opérations de requête SQL (expressions, recherches, jointures et filtres), d'effectuer des répartitions conditionnelles et de masquer des données personnelles pour préserver la pertinence, l'anonymat et la sécurité des données sensibles.
- **Text Data Processing** : Transformations aidant à extraire des informations spécifiques de votre texte. Elles analysent d'importants volumes de texte, ce qui vous permet d'identifier et d'extraire des entités et faits. Exemple : des clients, des produits, des emplacements et des informations financières pertinentes pour votre organisation.

Catégorie de transformations	Transformation	Description
Data Integrator	Data Transfer	Permet à un flux de données de diviser son traitement en deux sous-flux de données et de pousser vers le bas des opérations consommatrices de ressources vers le serveur de base de données.
	Date Generation	Génère une colonne contenant des valeurs de dates en fonction des dates de début et de fin, et de l'incrément que vous avez fournis.
	Effective Date	Génère une autre colonne « Date de fin de validité » en fonction de la « Date de validité » de la clé primaire.
	Hierarchy Flattening	Aplanit les données hiérarchiques sous forme de tables relationnelles pour qu'elles puissent participer à un schéma en étoile. La mise à plat hiérarchique peut être verticale ou horizontale.

Catégorie de transformations

	Transformation	Description
	History Preserving	Convertit les lignes marquées d'un indicateur Mettre à jour en Mettre à jour plus Insérer, pour que les valeurs d'origine soient conservées dans la cible. Indiquez dans quelle colonne rechercher les données mises à jour.
	Key Generation	Génère de nouvelles clés pour les données sources à partir d'une valeur basée sur les clés existantes dans la table que vous avez indiquée.
	Map CDC Operation	Trie les données d'entrée, mappe les données de sortie et résout les pré-images et les post-images pour les lignes Mettre à jour. Bien qu'elle soit habituellement utilisée pour prendre en charge la capture des données modifiées d'Oracle, cette transformation prend en charge tous les flux de données si leurs conditions d'entrée sont respectées.
	Pivot (Columns to Rows)	Fait pivoter les valeurs des colonnes spécifiques en lignes. (Voir aussi Reverse Pivot.)
	Reverse Pivot (Lignes en colonnes)	Fait pivoter les valeurs des lignes spécifiques en colonnes.
	Table Comparison	Compare deux jeux de données et produit la différence entre eux sous forme de jeu de données avec des lignes marquées d'un indicateur Insérer et Mettre à jour.
	XML Pipeline	Traite des entrées XML volumineuses dans des petits lots.
Data Quality	Associate	Combine les résultats de deux ou plus transformations Match, ou deux ou plus transformations Associate, ou une combinaison des deux, pour trouver les correspondances dans les sets de correspondance.
	Country ID	Analyse les données d'entrée puis identifie le pays de destination pour chaque enregistrement.
	Data Cleanse	Identifie et analyse les données de nom, de fonction et de société, les numéros de téléphone, les numéros de sécurité sociale, les dates et les adresses électroniques. Elle peut affecter un sexe, ajouter des titres de civilité, générer des standards Match et convertir des sources d'entrée à un format standard. Elle peut également analyser et manipuler plusieurs formes de données internationales, de données opérationnelles et de produit.
	DSF2 Walk Sequencer	Ajoute des informations de séquence de distribution aux données qu'il est possible d'utiliser avec un logiciel de tri préalable pour remplir les conditions pour des remises de séquence de tour.
	Geocoder	Utilise les coordonnées géographiques, les adresses et les données de points d'intérêt (POI) pour ajouter une adresse, une latitude et une longitude, un recensement et d'autres informations aux enregistrements.
	Global Address Cleanse	Identifie, analyse, valide et corrige les données d'adresses globales, telles que le numéro principal, le nom principal, le type principal, la direction, l'identificateur secondaire et le numéro secondaire.
	Global Suggestion Lists	Complète et renseigne les adresses avec des données minimales et peut fournir des propositions pour les correspondances possibles.
	Match	Identifie les enregistrements de correspondance en fonction des règles d'entreprise. Elle effectue également la sélection de candidat, ID unique, le meilleur enregistrement et d'autres opérations.

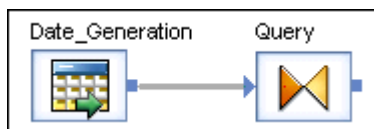
Catégorie de transformations	Transformation	Description
	USA Regulatory Address Cleanse	Identifie, analyse, valide et corrige les données d'adresses des Etats-Unis en fonction du CASS (Coding Accuracy Support System).
Plateforme	Case	Simplifie la logique de branche dans les flux de données en consolidant la logique de case ou de prise de décisions dans une transformation. Les chemins sont définis dans une table d'expressions.
	Data Mask	Utilise des techniques de masquage des données pour camoufler ou masquer les informations personnelles contenues dans vos bases de données (les numéros de compte bancaire, les numéros de carte de crédit et le revenu, par exemple). Vous pouvez utiliser les techniques suivantes dans Data Mask : • Remplacement de caractère • Variance numérique • Variance de dates • Variance de modèles • Généralisation de nombre • Généralisation de date Le masquage des données préserve la pertinence et les relations des données tout en assurant la confidentialité et l'anonymat des informations confidentielles du client et contribue à respecter les directives de protection des données.
	Map Operation	Modifie les données basées sur les codes d'opération actuels et les expressions de mappage. Les codes d'opération peuvent alors être convertis entre opérations de manipulation des données.
	Merge	Unifie les lignes à partir de deux sources ou plus dans une seule cible.
	Query	Extrait un jeu de données qui satisfait les conditions que vous avez indiquées. Une transformation Query est semblable à une instruction SQL SELECT.
	Row Generation	Génère une colonne contenant des valeurs de nombre entier démarrant à zéro par incrément de 1 pour la valeur de fin spécifiée.
	SQL	Effectue les opérations de requête SQL indiquées.
	User Defined	Effectue tout ce pourquoi il est possible d'écrire un code Python. Par exemple, il est possible d'utiliser la transformation pour créer des enregistrements et des jeux de données, ou pour renseigner un champ avec une valeur spécifique.
	Validation	Assure que les données dans le flux de données répondent à vos critères durant toutes les phases. Il est possible de filtrer ou de remplacer les données qui ne répondent pas aux critères.
Text Data Processing	Entity Extraction	Extrait des informations (entités et faits) d'un contenu au format de type texte, HTML, XML ou binaire (par exemple, PDF).

19.1 Ajout d'une transformation à un flux de données

Il est possible d'utiliser Designer pour ajouter des transformations aux flux de données.

1. Ouvrez un objet de flux de données.
2. Ouvrez la bibliothèque d'objets si elle n'est pas déjà ouverte et cliquez sur l'onglet [Transformations](#).
3. Sélectionnez la transformation ou la configuration de transformation que vous souhaitez ajouter au flux de données.
4. Faites glisser l'icône de transformation ou de configuration de transformation dans l'espace de travail du flux de données. Si vous avez sélectionné une transformation qui possède plusieurs configurations de transformation disponibles, un menu déroulant vous invite à sélectionner une configuration de transformation.
5. Dessinez les connexions du flux de données.

Pour connecter une source à une transformation, cliquez sur le carré sur le côté droit de la source et déplacez le curseur vers la flèche sur le côté gauche de la transformation.



Continuez à connecter les entrées et les sorties tel que requis pour la transformation.

- L'entrée pour la transformation peut être la sortie depuis une autre transformation, ou la sortie depuis une source ; ou bien la transformation peut ne pas exiger des données sources.
 - Vous pouvez connecter la sortie de la transformation à l'entrée d'une autre transformation ou cible.
6. Double-cliquez sur le nom de la transformation.
Cette opération ouvre l'éditeur de transformation, ce qui permet de terminer la définition de la transformation.
 7. Saisissez des valeurs d'options.
Pour qu'une colonne de données soit marquée comme une option de transformation, saisissez le nom de la colonne tel qu'il apparaît dans le schéma d'entrée ou faites glisser le nom de la colonne depuis le schéma d'entrée vers la boîte d'option.

19.2 Editeurs de transformations

Après avoir ajouté une transformation à un flux de données, configurez-la à l'aide de l'éditeur de transformation.

Les présentations de l'éditeur de transformation varient suivant le type de transformation.

❖ Exemple

La transformation la plus utilisée est la transformation Query. L'éditeur de transformation Query possède deux volets :

- Soit une zone *Entrée de schéma*, soit une zone *Sortie de schéma*, ou les deux.
- Une zone d'options à onglets où vous définissez les valeurs en fonction des besoins de la transformation et de la tâche à exécuter.

Les transformations Data Quality, telles que Match et Data Cleanse, utilisent un éditeur de transformations qui permet de définir les options et de mapper les champs d'entrée et de sortie.

L'éditeur de la transformation Entity Extraction dispose d'options d'extraction et d'une zone pour mapper les champs d'entrée et de sortie.

Référez-vous au *Guide de référence* pour des informations sur l'éditeur de transformation.

19.3 Aide intégrée pour Data Services

L'aide intégrée est disponible pour certaines transformations dans Designer.

L'aide intégrée sert à rechercher des informations supplémentaires sur les transformations et les options de SAP BusinessObjects Data Services. La rubrique change pour fournir une aide sur le contexte dans lequel vous évoluez. Lorsque vous sélectionnez une nouvelle transformation ou un nouveau groupe d'options, la rubrique est actualisée afin de refléter cette sélection.

Vous pouvez également accéder aux autres rubriques à l'aide des liens hypertexte figurant dans la rubrique ouverte, ou en vous servant de la table des matières.

Les transformations suivantes comprennent une aide intégrée :

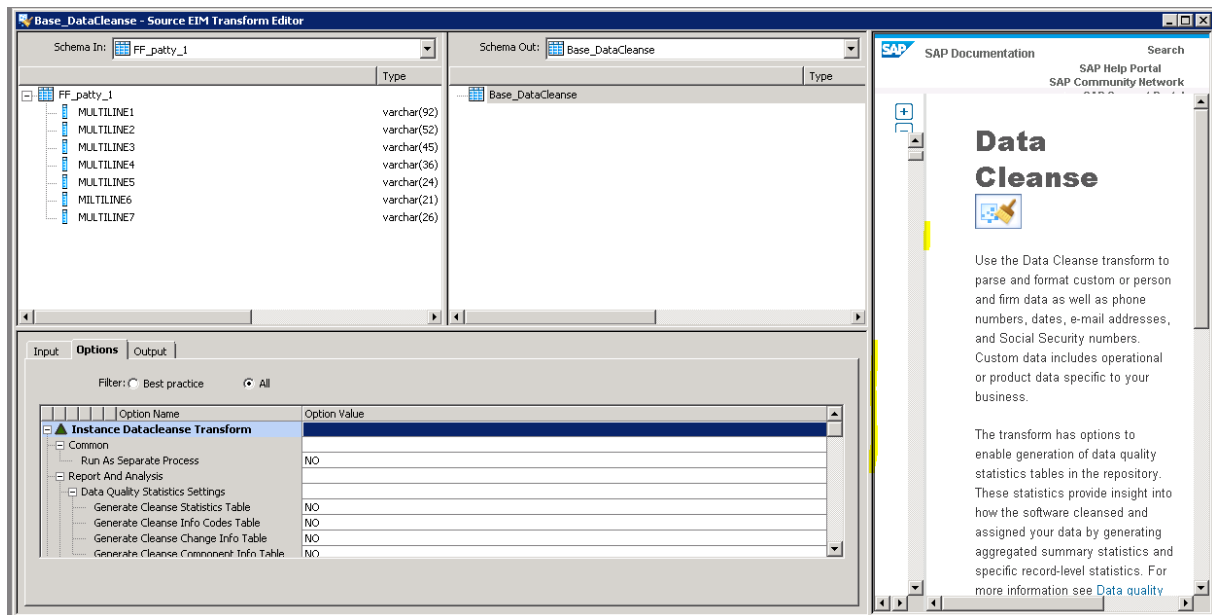
- Country ID
- Data Cleanse
- DQM Microservices
- Masque de données
- DSF2 Walk Sequencer
- Extraction
- Geocoder
- Global Address Cleanse
- Global Suggestion List
- Assistant Match
- USA Regulatory Address Cleanse
- Commun

i Remarque

Pour afficher les informations des options pour les transformations Associate, Match et User-Defined :

1. Sélectionnez la transformation dans votre flux de données.
2. Sélectionnez  *Outils* > <transformation> *Éditeur* .

Si le volet d'aide ne s'affiche pas à droite du volet de l'éditeur de transformation, faites glisser la bordure droite de l'éditeur à gauche. (voir Mise en surbrillance dans la capture d'écran suivante). Pour afficher ou masquer la table des matières d'aide, faites glisser la bordure entre le volet d'aide et le volet de la table des matières (aussi affiché en surbrillance).



19.4 Configurations de transformation

Une configuration de transformation désigne une transformation avec des champs de saisie et de sortie préconfigurés. Meilleure pratique et des options pouvant être utilisées dans plusieurs flux de données.

Ces configurations sont utiles lorsque vous utilisez de manière répétitive une transformation avec des options spécifiques et des champs d'entrée et de sortie.

Certaines transformations, par exemple, Data Quality ont des configurations en lecture seule que le logiciel installe lorsque vous installez Data Services. Vous pouvez créer vos propres configurations de transformation par la réplcation d'une configuration de transformation existante ou par la création d'une autre. Vous ne pouvez pas exécuter une exportation ou des opérations multi-utilisateurs sur des configurations de transformation en lecture seule.

Dans la boîte de dialogue *Éditeur de configuration de transformation*, vous configurez les options par défaut, les champs de saisie et de sortie. Meilleure pratique pour votre configuration de transformation. Après avoir placé une instance de la configuration de transformation dans un flux de données, vous pouvez remplacer celles qui existent par défaut.

Si vous modifiez une configuration de transformation, cette modification est héritée par chaque instance de la configuration de transformation utilisée dans les flux de données, sauf si un utilisateur a explicitement remplacé la même valeur d'option dans une instance.

Informations associées

[Création d'une configuration de transformation \[page 233\]](#)

[Ajout d'un champ défini par l'utilisateur \[page 234\]](#)

19.4.1 Création d'une configuration de transformation

Créez votre propre configuration de transformation en fonction des champs d'entrée et de sortie des meilleures pratiques communes et des options communes que vous utilisez dans vos flux de données.

1. Sélectionnez l'une des options suivantes en fonction de la manière dont vous voulez créer une configuration :
 - Cliquez avec le bouton droit de la souris sur une configuration de transformation dans l'onglet [Transformations](#) de la bibliothèque d'objets locale et sélectionnez [Nouveau](#).
 - Cliquez avec le bouton droit de la souris sur une configuration de transformation dans l'onglet [Transformations](#) de la bibliothèque d'objets locale et sélectionnez [Répliquer](#).

Vous ne pouvez pas créer une configuration de transformation lorsque les options [Nouveau](#) ou [Répliquer](#) ne sont pas disponibles dans le menu.

La fenêtre [Éditeur de configuration de la transformation](#) s'ouvre.

2. Saisissez un nom pour la nouvelle transformation dans [Nom de configuration de la transformation](#).
3. Ouvrez l'onglet [Options](#) et définissez les valeurs d'option pour déterminer la manière dont la transformation traitera vos données.

Si vous définissez les options pour les transformations Associate, Match ou User-Defined, cliquez sur [Modifier les options](#). La boîte de dialogue Éditeur associé, Éditeur de correspondance ou Éditeur défini par l'utilisateur, apparaît.

Si vous modifiez la valeur par défaut d'une option, un triangle vert s'affiche près du nom de l'option pour indiquer qu'un remplacement a été effectué.

4. Pour désigner une option comme meilleure pratique, cochez la case [Meilleure pratique](#) située près de la valeur.

Un indicateur des meilleures pratiques indique aux autres utilisateurs les options typiques à définir pour ce type de transformation.

Utilisez le filtre pour afficher toutes les options ou simplement les options qui sont désignées comme meilleure pratique.

5. Cliquez sur [Vérifier](#) pour contrôler si les valeurs d'option sélectionnées sont valides.
Des erreurs s'affichent en bas de la fenêtre, le cas échéant.

6. Dans l'onglet [Meilleures pratiques d'entrée](#), sélectionnez les champs d'entrée que vous désirez désigner comme meilleure pratique pour la configuration de transformation.

Les configurations de transformation fournies ne disposent pas des champs d'entrée de meilleures pratiques. Vous sélectionnez les champs d'entrée de meilleures pratiques en fonction du type du schéma d'entrée que vous préférez utiliser. Par exemple, vous pouvez mapper les champs de votre flux de données qui contiennent des données d'adresses si celles-ci se trouvent dans les types de champs suivants :

- Discret
- Multiligne
- Champs Discret et Multiligne combinés

Lorsque vous utilisez cette configuration de transformation dans un flux de données, sélectionnez le filtre Meilleure pratique dans l'onglet [Entrée](#) de l'éditeur de transformation. Le logiciel liste uniquement les champs que vous avez désignés comme meilleures pratiques dans la configuration de transformation.

7. Créez des champs d'entrée définis par l'utilisateur pour les configurations de transformation Associate, Match et User-Defined. Créez des champs d'entrée définis par l'utilisateur en cliquant sur [Créer](#).

8. Dans l'onglet *Meilleures pratiques de sortie*, sélectionnez les champs de sortie que vous désirez désigner comme meilleure pratique pour la configuration de transformation.
Lorsque vous utilisez cette configuration de transformation dans un flux de données, sélectionnez le filtre Meilleure pratique dans l'onglet *Sortie* de l'éditeur de transformation. Le logiciel liste uniquement les champs que vous avez désignés comme meilleures pratiques dans la configuration de transformation.
9. Cliquez sur *OK* pour enregistrer la configuration de transformation.
Le logiciel répertorie la nouvelle configuration de transformation dans la Bibliothèque d'objets locale sous la transformation de base du même type.

Utilisez la nouvelle configuration de transformation dans le flux de données.

19.4.2 Ajout d'un champ défini par l'utilisateur

Les transformations comme Associate, Match et User-Defined peuvent avoir des champs définis par l'utilisateur, car elles n'ont pas d'ensemble prédéfini de champs d'entrée.

Ajoutez un champ défini par l'utilisateur à une seule instance d'une transformation dans un flux de données ou à une configuration de transformation afin qu'il puisse être utilisé dans toutes les instances.

Dans la transformation User-Defined, vous pouvez également ajouter des champs de sortie définis par l'utilisateur.

1. Dans la Bibliothèque d'objets locale, ouvrez l'onglet *Transformations* et cliquez avec le bouton droit de la souris sur une configuration de transformation Associate, Match, ou User-Defined et sélectionnez *Modifier*. La boîte de dialogue *Éditeur de configuration de transformation* s'ouvre.
2. Ouvrez l'onglet *Meilleures pratiques d'entrée*, cliquez sur *Créer* et saisissez un nom pour le champ d'entrée.
3. Cliquez sur *OK* pour enregistrer la configuration de transformation.

Lorsque vous créez un champ défini par l'utilisateur dans la configuration de transformation, Data Services est affiché comme champ disponible dans chaque instance de la transformation dans un flux de données. Vous pouvez également créer des champs définis par l'utilisateur dans chaque instance de transformation.

19.4.3 Editeur d'options classées

Certaines transformations permettent de sélectionner et d'indiquer l'ordre de plusieurs valeurs pour une seule option.

L'option de séquence d'analyse de la transformation Data Cleanse constitue un exemple.

Pour configurer une option classée :

1. Cliquez sur les boutons de commande *Ajouter* et *Supprimer* pour supprimer des valeurs d'option entre des listes de valeurs Disponible et Sélectionné.

i Remarque

Supprimer toutes les valeurs. Pour supprimer la liste des valeurs Sélectionné et déplacer toutes les valeurs d'option à la liste de valeurs Disponible, cliquez sur *Tout supprimer*.

2. Sélectionnez une valeur dans la liste de valeurs Disponible et cliquez sur les boutons de commande Flèche vers le haut et le bas pour modifier la position de la valeur dans la liste.
3. Cliquez sur **OK** pour sauvegarder les modifications apportées à la configuration de l'option. Les valeurs sont répertoriées dans Designer et séparées par des caractères tiret vertical (|).

19.4.4 Editeurs des transformations Associate, Match et User-Defined

Les éditeurs de transformation Associate, Match et User-Defined contiennent chacun un éditeur supplémentaire pour définir des informations détaillées sur les options.

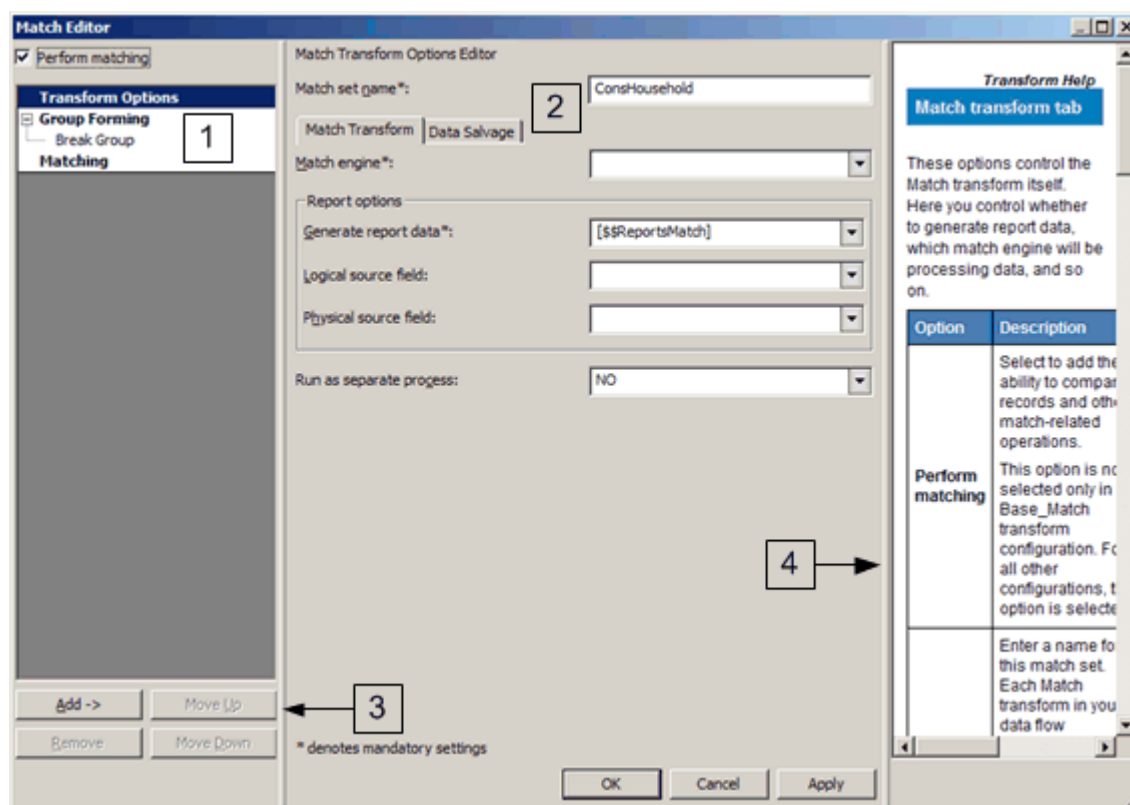
Il existe deux façons d'accéder à l'éditeur supplémentaire :

- Double-cliquez sur la transformation dans l'espace de travail pour ouvrir son éditeur, puis cliquez sur [Modifier les options](#).
- Sélectionnez la transformation dans le flux de données et sélectionnez ► **Outils** ► **Éditeur** `<nom_transformation>` ►. Par exemple, ► **Outils** ► **Éditeur Associate** ►.

i Remarque

Lorsque vous utilisez des flux de données créés dans l'Assistant de nettoyage Information Steward, vous ne pouvez pas accéder aux options de l'éditeur Match. Les options de la transformation Match ne peuvent pas être modifiées. Par conséquent, les commandes d'accès à l'éditeur Match sont inactives.

Les éditeurs des transformations Associate, Match et User-Defined ont une apparence et des fonctions similaires, et dans certains cas, ils partagent les mêmes groupes d'options.



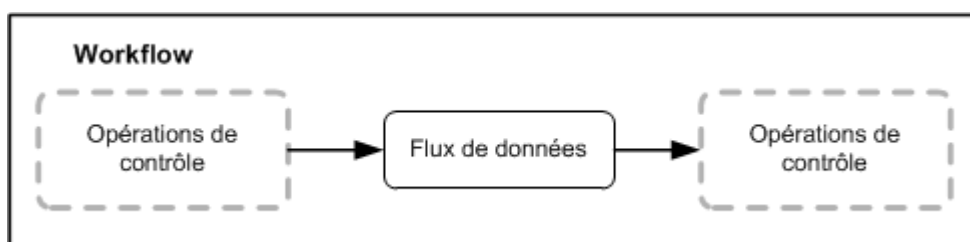
La fenêtre de l'éditeur est partagée en quatre zones :

1. *Explorateur d'options* — Dans cette zone, vous sélectionnez les groupes d'options ou les opérations disponibles pour la transformation. Pour afficher un groupe d'options masqué, cliquez avec le bouton droit sur le groupe d'options auquel il appartient et sélectionnez son nom dans le menu.
2. *Editeur d'options* — Dans cette zone, vous indiquez la valeur de l'option.
3. *Boutons* — Utilisez-les pour ajouter, supprimer et ordonner les groupes d'options.
4. *Aide intégrée* — L'aide intégrée affiche des informations supplémentaires sur l'utilisation de l'écran actif de l'éditeur.

20 Workflows

Un workflow définit le processus de prise de décision pour l'exécution de flux de données.

Par exemple, les éléments d'un workflow peuvent déterminer le chemin d'exécution en fonction d'une valeur définie par un job précédent ou peut indiquer un autre chemin si le chemin principal est incorrect. Finalement, un workflow a pour but de préparer l'exécution de flux de données et de définir l'état du système à la fin de l'exécution des flux de données.



Les jobs sont des workflows spéciaux. Les jobs ont une particularité en ce sens où vous pouvez les exécuter. Presque toutes les fonctionnalités documentées pour les workflows s'appliquent également aux jobs, à une exception : les jobs n'ont pas de paramètre.

Informations associées

[Projets \[page 77\]](#)

20.1 Etapes d'un workflow

Les étapes de workflow prennent la forme d'icône que vous placez dans l'espace de travail pour créer un diagramme de workflows.

Les objets suivants peuvent être des éléments de workflows :

- Workflows
- Flux de données
- Scripts
- Structures conditionnelles
- Boucles While
- Blocs Try/Catch

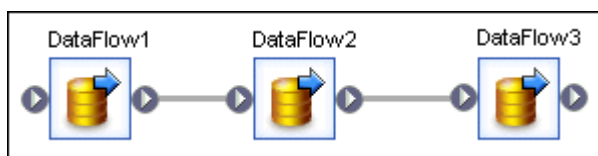
Les workflows peuvent appeler d'autres workflows et vous pouvez imbriquer des appels dans n'importe quelle profondeur. Un workflow peut également s'appeler lui-même.

Les connexions que vous établissez entre les icônes dans l'espace de travail déterminent l'ordre dans lequel des workflows sont exécutés, sauf si les jobs contenant ces workflows sont exécutés en parallèle.

20.2 Ordre d'exécution dans les workflows

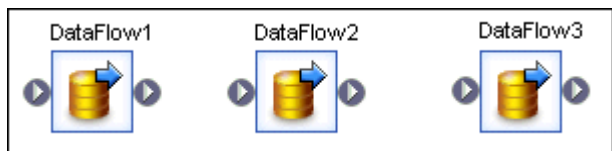
Les étapes d'un workflow sont exécutées dans une séquence de gauche à droite, indiquée par les lignes reliant les étapes.

Vous trouverez ci-dessous le diagramme d'un workflow appelant trois flux de données :

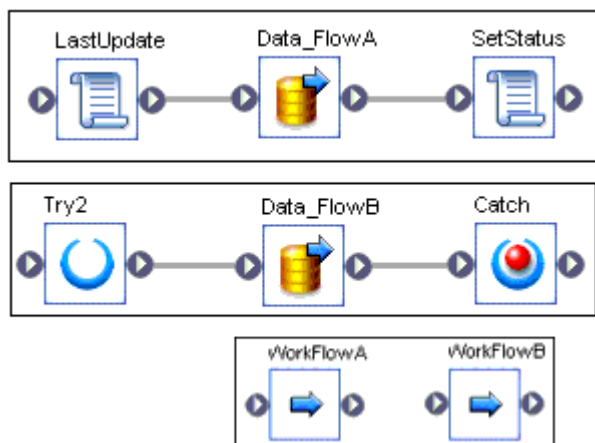


Notez que Data_Flow1 n'a aucune connexion de la gauche mais est relié de la l'arête droite à l'arête gauche de Data_Flow2 et que ce Data_Flow2 est relié à Data_Flow3. Un seul thread de contrôle relie les trois étapes. L'exécution commence par Data_Flow1 et se poursuit à travers les trois flux de données.

Reliez les étapes d'un workflow lorsqu'il existe une dépendance entre les étapes. S'il n'existe aucune dépendance, les étapes ne doivent pas être reliées. Dans ce cas, le logiciel peut exécuter les étapes indépendantes dans le workflow en tant que processus distincts. Dans le workflow suivant, le logiciel exécute les flux de données 1 à 3 en parallèle :



Pour exécuter des workflows plus complexes en parallèle, définissez chaque séquence comme workflow distinct puis appelez chacun des workflows d'un autre workflow comme dans l'exemple suivant :



Vous pouvez indiquer qu'un job exécute un workflow spécifique ou un flux de données à la fois. Dans ce cas, le logiciel exécute la première occurrence du workflow ou de flux de données ; le logiciel ignore les occurrences suivantes dans le job. Vous pouvez utiliser cette fonctionnalité pour un développement complexe de jobs avec plusieurs chemins, par exemple des jobs avec des blocs Try/Catch ou des structures conditionnelles et vous souhaitez garantir que le logiciel exécute uniquement un workflow spécifique ou un flux de données à la fois.

20.3 Exemple de workflow

Supposez que vous souhaitiez mettre à jour une table de faits. Vous définissez un flux de données dans lequel la transformation de données actuelle a lieu. Cependant, avant de déplacer des données de la source, vous souhaitez déterminer la dernière date de mise à jour de la table de faits, de telle sorte que seules les lignes ajoutées ou modifiées depuis cette date, soient extraites.

Vous devez écrire un script pour déterminer la dernière date de mise à jour. Vous pouvez alors transférer cette date au flux de données comme paramètre.

En outre, vous voulez vérifier que les connexions de données requises pour créer la table de faits sont actives lorsque les données sont lues à partir de cette dernière. Pour ce faire, dans le logiciel, définissez un bloc Catch/Try. Si les connexions ne sont pas actives, le catch exécute un script que vous avez écrit, qui envoie automatiquement un e-mail informant un Administrateur du problème.

Les scripts et la détection d'erreur ne peuvent pas être exécutés dans le flux de données. Il existe plutôt des étapes de processus de prise de décision influençant le flux de données. Ce processus de prise de décision est défini comme workflow, tel que le suivant :



Le logiciel exécute ces étapes dans l'ordre dans lequel vous les avez reliées.

20.4 Création de workflows

Vous pouvez créer des workflows à l'aide de la bibliothèque d'objets ou de la palette d'outils.

Une fois que vous avez créé un workflow, vous pouvez indiquer qu'un seul job exécute le workflow à la fois, sous forme de processus simple ou continu, même si ce dernier apparaît plusieurs fois dans le job.

20.4.1 Créer un workflow à l'aide de la bibliothèque d'objets

1. Ouvrez la bibliothèque d'objets.

2. Accédez à l'onglet [Workflows](#).
3. Cliquez avec le bouton droit de la souris et sélectionnez [Nouveau](#).
4. Faites glisser le workflow dans le diagramme.
5. Ajoutez les flux de données, les workflows, les structures conditionnelles et les scripts dont vous avez besoin.

20.4.2 Créer un workflow à l'aide de la palette d'outils

1. Sélectionnez l'icône Workflow dans la palette d'outils.
2. Cliquez sur l'endroit où vous souhaitez placer le workflow dans le diagramme.

Si plusieurs instances d'un workflow apparaissent dans un job, vous pouvez améliorer la performance de l'exécution en lançant le workflow une seule fois.

20.4.3 Spécifier qu'un job exécute le workflow une seule fois

Lorsque vous spécifiez qu'un workflow doit être exécuté une seule fois, un job ne ré-exécutera jamais ce workflow après qu'il ait terminé avec succès.

Sauf si le flux de données est contenu dans un workflow qui est une unité de restauration qui ré-exécute et qui n'a pas terminé avec succès ailleurs en dehors de l'unité de restauration.

Il est recommandé de ne pas activer l'option [Exécuter une seule fois](#) pour un workflow, si ce dernier ou un workflow parent est une unité de restauration.

1. Cliquez avec le bouton droit de la souris sur le workflow et sélectionnez [Propriétés](#).
La fenêtre Propriétés s'ouvre pour le workflow.
2. Sélectionnez [Régulier](#) dans la liste déroulante [Type d'exécution](#).
3. Cochez la case [Exécuter une seule fois](#)
4. Cliquez sur [OK](#).

20.4.4 Qu'est-ce qu'un workflow unique ?

Un workflow unique exécute tous les flux de données subordonnés d'un processus de système d'exploitation.

Si les flux de données sont destinés à être exécutés en parallèle, ils sont exécutés dans différents threads au lieu de différents processus. L'avantage d'un traitement unique est qu'il est possible de partager des ressources telles que des connexions à la base de données entre plusieurs flux de données.

i Remarque

Les workflows uniques comportent les limitations suivantes :

- Un workflow unique ne peut pas appeler ou contenir de workflow continu.

- Un workflow unique ne peut pas utiliser de flux de données secondaire. C'est pourquoi les options de transformation Data Transfer et "Exécuter en tant que processus séparé" ne sont pas valides. Le logiciel génère une erreur de validation d'exécution.
- Un workflow unique peut être exécuté uniquement par un workflow continu. Un workflow unique ne peut pas appeler d'autre workflow unique.

20.4.4.1 Spécifier qu'un job s'exécute comme workflow unique

1. Cliquez avec le bouton droit de la souris sur un workflow et sélectionnez [Propriétés](#).
2. Sélectionnez [Unique](#) dans la liste déroulante Type d'exécution.
3. Cliquez sur [OK](#).

20.4.5 Qu'est-ce qu'un workflow continu ?

Un workflow continu exécute tous les flux de données dans une boucle mais les conserve en mémoire pour l'itération suivante.

Cela libère de la nécessité de répéter certaines étapes courantes d'exécution (telles que la connexion au référentiel, l'analyse/optimisation/compilation ATL, l'ouverture des connexions à la base de données).

i Remarque

Les workflows continus comportent les limitations suivantes :

- Un workflow continu ne peut pas appeler ou contenir un autre workflow continu. Si un workflow continu en appelle un autre qui ne se termine jamais, il ne peut jamais redémarrer les processus subordonnés.
- Un workflow continu ne peut pas utiliser de flux de données secondaire. C'est pourquoi les options de transformation Data Transfer et "Exécuter en tant que processus séparé" ne sont pas valides. Le logiciel génère une erreur de validation d'exécution.
- Un workflow unique ou normal ne peut pas appeler ou contenir de workflow continu.
- Un job en temps réel ne peut pas appeler de workflow continu.
- La transformation XML_Map de plateforme ne peut pas être utilisée dans un workflow continu.
- Les transformations Data Integrator, Data_Transfer et XML_Pipeline ne peuvent pas être utilisées dans un workflow continu.

20.4.5.1 Spécifier qu'un job exécute un workflow continu

1. Cliquez avec le bouton droit de la souris sur un workflow et sélectionnez [Propriétés](#).
2. Sélectionnez [Continu](#) dans la liste déroulante Type d'exécution.
3. Accédez à l'onglet [Options continues](#).

4. Indiquez quand vous souhaitez que le workflow valide les ressources :
 - Pour valider les ressources après un certain nombre d'exécutions, sélectionnez *Après un certain nombre d'exécutions* et saisissez le nombre d'exécutions. La valeur par défaut est 100.
 - Pour valider les ressources après un certain nombre d'heures, cochez la case *Après*, sélectionnez *Nombre d'heures* et saisissez le nombre d'heures.
 - Pour valider les ressources après un certain nombre de jours, cochez la case *Après*, sélectionnez *Nombre de jours* et saisissez le nombre de jours.
 - Pour valider les ressources lorsque le résultat d'une fonction n'est pas égal à zéro, cochez la case *Après*, sélectionnez *Le résultat de la fonction n'est pas égal à zéro* et saisissez la fonction à utiliser.
5. Pour arrêter le workflow lorsque le résultat d'une fonction personnalisée est égal à zéro, sélectionnez *Lorsque le résultat de la fonction est nul* et saisissez la fonction personnalisée à utiliser.
6. Cliquez sur *OK*.

20.5 Structures conditionnelles

Les structures conditionnelles sont des objets à utilisation unique utilisés pour implémenter une logique if/then/else dans un workflow.

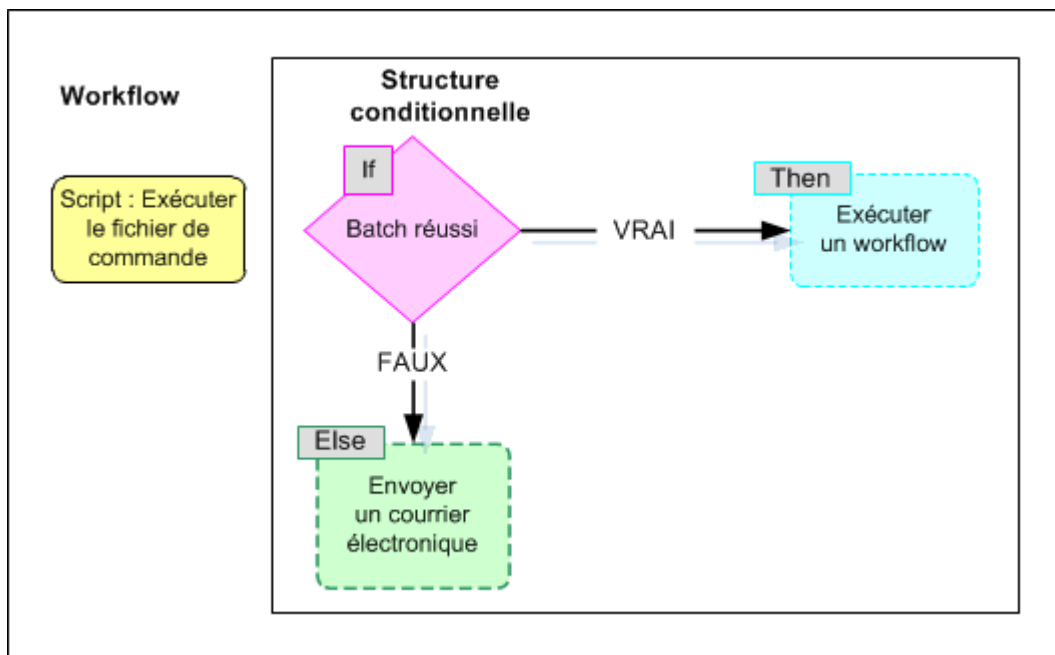
Les structures conditionnelles et leurs composants (expressions if, diagrammes then et else) sont incluses dans le périmètre des variables et des paramètres de flux de contrôle parent.

Pour définir une structure conditionnelle, indiquez une condition et deux branches logiques :

Branche conditionnelle	Description
<i>If</i>	Expression booléenne qui donne la valeur VRAI ou FAUX Vous pouvez utiliser des fonctions, des variables et des opérateurs standard pour créer l'expression.
<i>Then</i>	Éléments de workflow à exécuter si l'expression <i>If</i> donne la valeur VRAIE
<i>Else</i>	(Facultatif) Éléments de workflow à exécuter si l'expression <i>If</i> donne la valeur FALSE.

Définissez les branches *Then* et *Else* dans la structure conditionnelle.

Une structure conditionnelle peut être intégrée à un workflow. Supposez que vous utilisiez un fichier de commandes Windows pour transférer des données d'un système hérité dans le logiciel. Vous écrivez un script dans un workflow pour exécuter le fichier de commandes et renvoyer un indicateur correct. Vous définissez ensuite une structure conditionnelle qui lit l'indicateur pour déterminer si les données sont disponibles pour le reste du workflow.



Pour implémenter cette structure conditionnelle dans le logiciel, vous définissez deux workflows, un pour chaque branche de la structure conditionnelle. Si les éléments de chaque branche sont simples, vous pouvez les définir dans l'éditeur de structures conditionnelles lui-même.

Les branches *Then* et *Else* de la structure conditionnelle peuvent contenir tout objet susceptible de se trouver dans un workflow, notamment d'autres workflows, des structures conditionnelles imbriquées, des blocs Try/Catch, etc.

20.5.1 Définir un conditionnel

1. Définissez les workflows appelés par les branches *Then* et *Else* de la structure conditionnelle.

Il est recommandé de définir, tester et sauvegarder chaque workflow comme objet distinct plutôt qu'en créant ces workflows dans l'éditeur de structures conditionnelles.

2. Ouvrez le workflow dans lequel vous souhaitez placer la structure conditionnelle.
3. Cliquez sur l'icône pour une structure conditionnelle dans la palette d'outils.
4. Cliquez sur l'emplacement où vous souhaitez placer la structure conditionnelle dans le diagramme.

La structure conditionnelle apparaît dans le diagramme.

5. Cliquez sur le nom de la structure conditionnelle pour ouvrir l'éditeur de structures conditionnelles.
6. Cliquez sur *if*.
7. Saisissez l'expression booléenne contrôlant la structure conditionnelle.

Poursuivez la création de votre expression. Vous souhaitez peut-être utiliser l'assistant de fonction ou l'éditeur rapide.

8. Une fois l'expression créée, cliquez sur *OK*.
9. Ajoutez le workflow prédéfini dans la case *Then*.

Pour ajouter un workflow existant, ouvrez la bibliothèque d'objets dans l'onglet Workflows, sélectionnez le workflow souhaité puis faites-le glisser dans la case Then.

10. (Facultatif) Ajoutez le workflow prédéfini dans la case [Else](#).

Si l'expression IF donne la valeur FAUX et la case Else n'est pas cochée, le logiciel quitte la structure conditionnelle et poursuit avec le workflow.

11. Une fois la structure conditionnelle exécutée, sélectionnez [DeboggerValider](#).

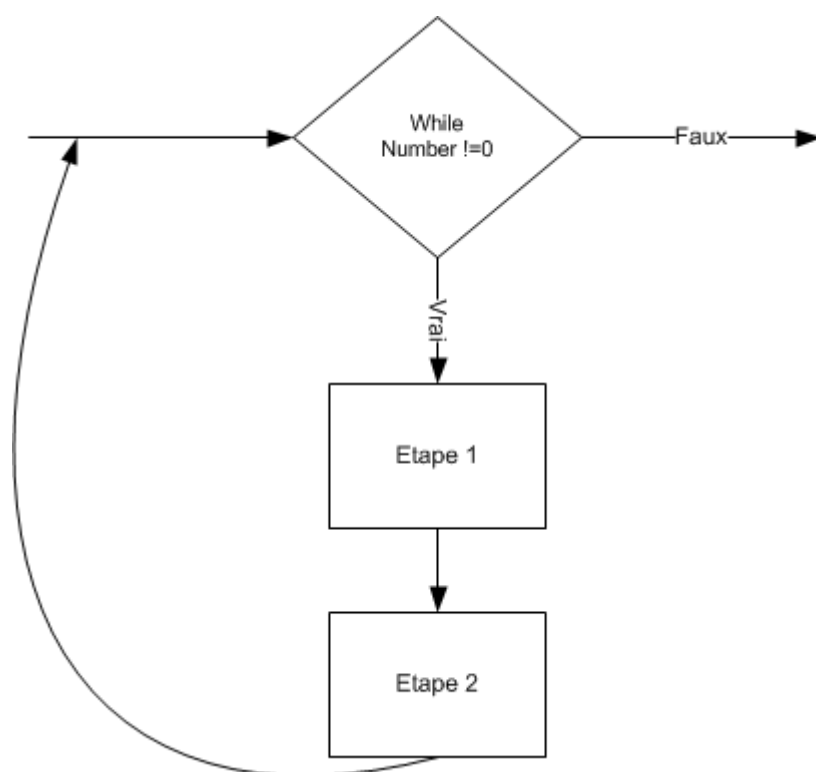
Le logiciel teste votre structure conditionnelle par rapport aux erreurs de syntaxe et affiche toutes les erreurs rencontrées.

12. La structure conditionnelle est à présent définie. Cliquez sur le bouton [Retour](#) pour revenir au workflow appelant la structure conditionnelle.

20.6 Boucles While

La boucle WHILE désigne un objet à utilisation unique que vous pouvez utiliser dans un workflow.

Utilisez une boucle WHILE pour répéter une séquence d'étapes dans un workflow tant qu'une condition est vraie.

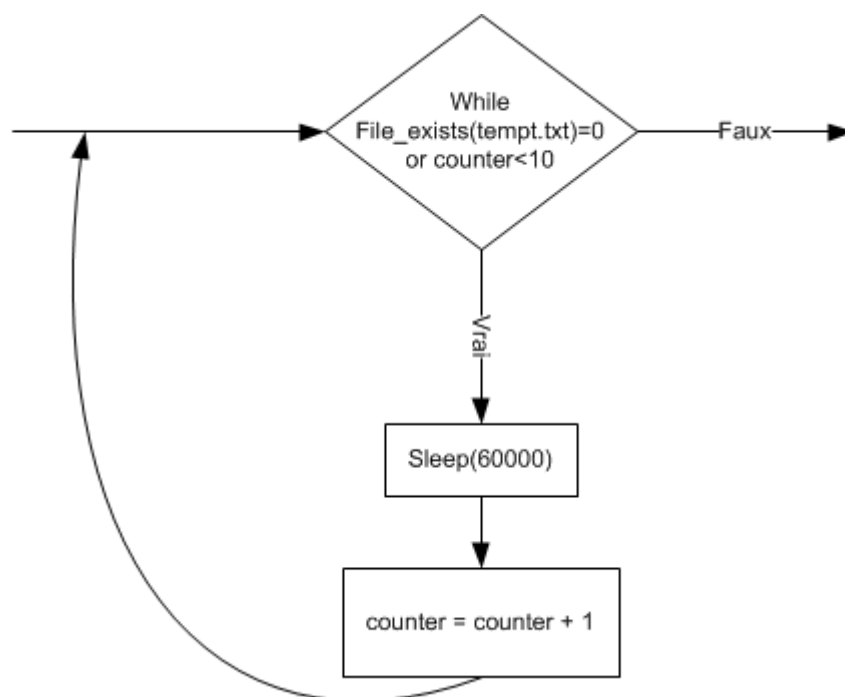


Généralement, les étapes exécutées au cours de la boucle WHILE se traduisent par une modification de la condition de telle sorte qu'éventuellement, la condition ne soit plus remplie et que le workflow existe à partir de la boucle WHILE. Si la condition n'est pas modifiée, la boucle WHILE n'est pas interrompue.

Par exemple, vous souhaitez peut-être qu'un workflow attende jusqu'à ce que le système écrive un fichier spécifique. Vous pouvez utiliser une boucle WHILE pour vérifier l'existence du fichier à l'aide de la fonction

`file_exists`. Tant que le fichier n'existe pas, vous pouvez exécuter le workflow en mode veille pour une durée spécifique, soit une minute avant que le contrôle ne reprenne.

Comme le système peut ne jamais écrire le fichier, vous devez ajouter un autre contrôle à la boucle, par exemple un compteur, pour garantir que la boucle WHILE existe éventuellement. En d'autres mots, modifiez la boucle WHILE pour vérifier l'existence du fichier et la valeur du compteur. Dès que le fichier n'existe pas et que la valeur du compteur est inférieure à une valeur spécifique, répétez la boucle WHILE. Dans chaque itération de la boucle, mettez le workflow en mode veille puis augmentez le compteur.



20.6.1 Définition d'une boucle WHILE

Vous pouvez définir une boucle WHILE dans tout workflow.

1. Ouvrez le workflow où vous souhaitez placer la boucle WHILE.
2. Cliquez sur l'icône Boucle WHILE dans la palette d'outils.
3. Cliquez sur l'emplacement où vous souhaitez placer la boucle WHILE dans le diagramme d'espace de travail.

La boucle WHILE s'affiche dans le diagramme.

4. Cliquez sur la boucle WHILE pour ouvrir l'éditeur de boucles WHILE.
5. Dans la case **WHILE** en haut de l'éditeur, saisissez la condition qui doit s'appliquer à l'initialisation et répétez les étapes dans la boucle WHILE.

Vous pouvez également ouvrir l'éditeur d'expressions qui vous permet, en agrandissant l'espace, de saisir une expression et d'accéder à l'assistant de fonction. Cliquez sur **OK** dès que vous avez saisi une expression dans l'éditeur.

6. Ajoutez les étapes que vous voulez exécuter au cours de la boucle WHILE dans l'espace de travail, dans l'éditeur de boucles WHILE.

Vous pouvez ajouter tout objet valide dans un workflow incluant des scripts, des workflows et des flux de données. Reliez ces objets pour représenter l'ordre dans lequel vous souhaitez exécuter les étapes.

i Remarque

Même si vous pouvez inclure le workflow parent dans la boucle WHILE, des appels récursifs peuvent créer une boucle infinie.

7. Une fois que vous avez défini les étapes dans la boucle WHILE, sélectionnez  **Débogage**  **Valider** .

Le logiciel teste votre définition par rapport aux erreurs de syntaxe et affiche les erreurs rencontrées.

8. Fermez l'éditeur de boucles WHILE pour revenir au workflow appelant.

20.6.2 Utilisation d'une boucle WHILE avec la visualisation des données

Selon la conception de votre job, il est possible que le logiciel ne termine pas toutes les itérations d'une boucle WHILE si vous exécutez un job en mode de visualisation des données.

Lorsque vous utilisez la visualisation des données, un job est interrompu lorsque le logiciel a récupéré le nombre de lignes indiqué pour tous les objets pouvant être analysés.

Les situations suivantes peuvent se produire lors de l'utilisation de la boucle WHILE pendant l'exécution d'un job en mode de visualisation des données :

- Si la boucle WHILE contient des objets pouvant être analysés et qu'il n'y en a pas en dehors de la boucle WHILE (par exemple si la boucle WHILE est le dernier objet d'un job), l'exécution du job prendra fin après l'analyse des objets dans la boucle WHILE, peut-être à la fin de la première itération de la boucle WHILE.
- S'il reste des objets pouvant être analysés par la suite, la boucle WHILE est exécutée normalement. Les objets analysés dans la boucle WHILE indiqueront les résultats de la dernière itération.
- S'il n'existe aucun objet pouvant être analysé suite à la boucle WHILE mais s'il existe des objets pouvant être analysés terminés parallèlement à la boucle WHILE, l'exécution du job se termine à la fin de l'analyse de tous les objets pouvant être analysés. La boucle WHILE peut exécuter n'importe quel nombre d'itérations.

20.7 Blocs Try/Catch

Un bloc Try/Catch désigne une combinaison d'un objet Try et d'un ou plusieurs objets Catch qui permettent d'indiquer d'autres workflows si des erreurs surviennent lorsque le logiciel exécute un job.

Blocs Try/Catch :

- Groupes d'exceptions "Catch" "levées" par le logiciel, le DBMS ou le système d'exploitation.
- Appliquez les solutions aux groupes d'exceptions ou aux erreurs spécifiques dans un groupe.
- Poursuivez l'exécution.

Les objets Try et Catch sont des objets à utilisation unique.

Voici une méthode générale d'implémentation de la gestion des exceptions :

1. Insérez un objet Try avant les étapes de gestion des erreurs.
2. Insérez un objet Catch dans le workflow après les étapes.
3. Dans l'objet Catch, exécutez les étapes suivantes :
 - Sélectionnez un ou plusieurs groupes d'erreurs que vous souhaitez intercepter.
 - Définissez les actions exécutées par une exception levée. Ces actions peuvent être un objet de script unique, un flux de données, un workflow ou une combinaison de ces objets.
 - Facultatif. Utilisez des fonctions Catch dans le bloc Catch pour identifier les détails de l'erreur.

Si une exception est renvoyée au cours de l'exécution d'un bloc Try/Catch et si aucun objet Catch ne recherche cette exception, l'exception est traitée par la logique d'erreur normale.

Le workflow suivant indique un bloc Try/Catch autour d'un flux de données :



Dans ce cas, si le flux de données BuildTable provoque des exceptions générées par le système indiquées dans le catch Catch_A, les actions définies dans Catch_A sont exécutées.

L'action initialisée par l'objet catch peut être simple ou complexe. Voici des exemples d'actions d'exceptions possibles :

- Envoyer le message d'erreur à une base de données de reporting en ligne ou à votre groupe de support.
- Exécuter à nouveau un workflow échoué ou un flux de données.
- Exécuter une diminution de version d'un workflow échoué ou un flux de données.

Informations associées

[Définition d'un bloc Try/Catch \[page 247\]](#)

[Catégories d'exceptions disponibles \[page 249\]](#)

[Exemple : détermination des détails d'une erreur \[page 249\]](#)

20.7.1 Définition d'un bloc Try/Catch

Pour définir des blocs Try/Catch :

1. Ouvrez le workflow qui contiendra le bloc Try/Catch.
2. Cliquez sur l'icône Try dans la palette d'outils.
3. Cliquez sur l'emplacement où vous voulez placer le try dans le diagramme.

L'icône Try s'affiche dans le diagramme.

i Remarque

Aucun éditeur n'existe pour un Try ; le Try initialise le bloc Try/Catch.

4. Cliquez sur l'icône Catch dans la palette d'outils.
5. Cliquez sur l'emplacement où vous voulez placer l'objet catch dans l'espace de travail.
L'objet catch s'affiche dans l'espace de travail.
6. Reliez les objets Try et Catch aux objets liés.
7. Cliquez sur le nom de l'objet catch pour ouvrir l'éditeur catch.
8. Sélectionnez un ou plusieurs groupes dans la liste des [Exceptions](#).
Pour sélectionner tous les groupes d'exceptions, cochez la case en haut.
9. Définissez les actions à exécuter pour chaque groupe d'exceptions et ajoutez les actions à la boîte de workflows catch. Les actions peuvent être un script individuel, un flux de données, un workflow ou toute combinaison de ces objets.
 - a. Il est recommandé de définir, de tester et de sauvegarder les actions comme objet distinct plutôt que de les créer dans l'éditeur Catch.
 - b. Pour définir des actions pour des erreurs spécifiques, utilisez les fonctions catch suivantes dans un script exécutant le workflow :
 - `error_context()`
 - `error_message()`
 - `error_number()`
 - `error_timestamp()`
 - c. Pour ajouter un workflow existant à la case de workflow catch, ouvrez la bibliothèque d'objets dans l'onglet Workflows, sélectionnez le workflow souhaité et glissez-le dans la case.
10. Une fois le catch exécuté, sélectionnez ► [Validation](#) ► [Valider](#) ► [Tous les objets de la vue](#) ►.
Le logiciel teste votre définition par rapport aux erreurs de syntaxe et affiche les erreurs rencontrées.
11. Cliquez sur le bouton [Retour](#) pour revenir au workflow appelant le catch.
12. Pour intercepter plusieurs groupes d'exceptions et affecter différentes actions à chacun d'entre eux, répétez les étapes 4 à 11 pour chaque catch dans le workflow.

i Remarque

Dans une séquence de blocs Catch, si un bloc Catch intercepte une exception, les blocs Catch suivants ne seront pas exécutés. Par exemple, si votre workflow a la séquence suivante et que Catch1 intercepte une exception, Catch2 et CatchAll ne seront pas exécutés.

```
Try > DataFlow1 > Catch1 > Catch2 > CatchAll
```

Si une erreur du groupe d'exceptions listée dans le Catch survient au cours de l'exécution de ce bloc Try/Catch, le logiciel exécute le workflow Catch.

Informations associées

[Catégories d'exceptions disponibles \[page 249\]](#)

20.7.2 Catégories d'exceptions disponibles

Les catégories d'exceptions disponibles comprennent les éléments suivant :

- Erreurs d'exécution (1001)
- Erreurs d'accès à la base de données (1002)
- Erreurs de connexion à la base de données (1003)
- Erreurs de traitement de fichier plat (1004)
- Erreurs d'accès au fichier (1005)
- Erreurs d'accès au référentiel (1006)
- Erreurs de système SAP (1007)
- Exception de ressource système (1008)
- Erreurs d'exécution BW SAP (1009)
- Erreurs de traitement XML (1010)
- Erreurs de copybook COBOL (1011)
- Erreurs de classeur Excel (1012)
- Erreurs de transformation Data Quality (1013)

20.7.3 Exemple : détermination des détails d'une erreur

Cet exemple illustre la façon d'utiliser les fonctions d'erreur dans un script de catch. Supposez que vous souhaitiez déterminer des erreurs d'accès à la base de données et envoyer les détails de l'erreur à votre groupe de support.

1. Dans l'éditeur catch, sélectionnez le groupe d'exceptions que vous souhaitez déterminer. Dans cet exemple, cochez la case devant *Erreurs d'accès à la base de données(1002)*.
2. Dans la zone de workflow de l'éditeur catch, créez un objet de script avec le script suivant :

```
mail_to('support@my.com',
        'Data Service error number' || error_number(),
        'Error message: ' || error_message(),20,20);
print('DBMS Error: ' || error_message());
```

3. Cet exemple de script de catch inclut la fonction mail_to pour effectuer les opérations suivantes :
 - Indiquer l'adresse e-mail de votre groupe de support.
 - Envoyer le numéro d'erreur renvoyé par la fonction error_number() pour l'exception détectée.
 - Envoyer le message d'erreur renvoyé par la fonction error_message() pour l'exception détectée.
4. L'exemple de script de catch inclut une commande d'impression permettant d'imprimer le message d'erreur pour l'erreur de la base de données.

20.7.4 Meilleures pratiques Catch

Utilisez ces recommandations pour configurer correctement un bloc Try/Catch dans votre configuration de job.

Pour chaque objet Catch dans le bloc Try/Catch, spécifiez les éléments suivants :

- Un ou plusieurs groupes d'exception gérés par l'objet Catch.

i Remarque

Si vous souhaitez affecter des actions différentes à des groupes d'exception différents, ajoutez un Catch pour chaque ensemble d'actions.

- Les actions à exécuter lorsqu'une exception dans les groupes d'exception indiqués se produit.
Facultatif mais recommandé : Définir, tester et sauvegarder les actions comme objet distinct plutôt que de les créer dans l'éditeur Catch. Ces actions peuvent être un objet de script unique, un flux de données, un workflow ou une combinaison de ces objets.
- Les fonctions d'erreur facultatives dans le bloc Catch pour identifier les détails de l'erreur.

Si une exception est renvoyée au cours de l'exécution d'un bloc Try/Catch et si aucun objet Catch ne recherche ce groupe d'exception, l'exception est traitée par la logique d'erreur normale.

Pour les jobs batch uniquement, ne faites référence aux variables de sortie d'un bloc Try/Catch dans aucune des étapes ultérieures si vous utilisez la fonctionnalité de récupération automatique. Le fait de référencer de telles variables peut avoir une incidence sur les résultats lors de la récupération automatique.

Vous pouvez utiliser des blocs Try/Catch dans tous les composants de job en temps réel. Cependant, les blocs Try/Catch ne peuvent pas chevaucher une boucle de traitement en temps réel ni le composant d'initialisation ou de nettoyage d'un job en temps réel.

20.8 Scripts

Les scripts sont des objets à utilisation unique, utilisés pour appeler des fonctions et affecter des valeurs aux variables dans un workflow.

Par exemple, vous pouvez utiliser la fonction SQL dans un script pour déterminer la dernière heure de mise à jour pour une table et affecter cette valeur à une variable. Vous pouvez ensuite affecter la variable à un paramètre transféré dans un flux de données et qui identifie les lignes à extraire d'une source.

Un script peut contenir les instructions suivantes :

- Appels de fonction
- Instructions IF
- Instructions WHILE
- Instructions d'affectation
- Opérateurs

Les règles de base pour la syntaxe du script sont les suivantes :

- Chaque ligne se termine par un point-virgule (;).

- Les noms de variables commencent par le symbole dollar (\$).
- Les valeurs de chaîne sont entre parenthèses simples (').
- Les commentaires commencent par un symbole dièse (#).
- Les appels de fonction indiquent toujours des paramètres même si la fonction n'utilise aucun paramètre.

Par exemple, l'instruction de script suivante détermine la date du jour et affecte la valeur à la variable \$AUJOURD'HUI :

```
$TODAY = sysdate();
```

Vous ne pouvez pas utiliser des variables à moins de les déclarer dans le workflow appelant le script.

Informations associées

[Guide de référence : langage de script \[page 706\]](#)

20.8.1 Création d'un script

1. Ouvrez le workflow.
2. Cliquez sur l'icône Script dans la palette d'outils.
3. Cliquez sur l'emplacement où vous souhaitez placer le script dans le diagramme.

L'icône Script s'affiche dans le diagramme.

4. Cliquez sur le nom du script pour ouvrir l'éditeur de scripts.
5. Saisissez les instructions de scripts, chacune suivie d'un point-virgule.

L'exemple suivant illustre un script qui détermine l'heure de début d'une fonction personnalisée.

```
AW_StartJob ('NORMAL', 'DELTA', $G_STIME, $GETIME);
$GETIME =to_date(
sql('ODS_DS', 'SELECT to_char(MAX(LAST_UPDATE) ,
\'YYYY-MM-DD HH24:MI:SS\')
FROM EMPLOYEE'),
'YYYY_MMM_DDD_HH24:MI:SS');
```

Cliquez sur le bouton Fonction pour inclure des fonctions dans votre script.

6. Une fois le script terminé, sélectionnez ► **Validation** ► **Valider** ►.

Le logiciel teste votre script par rapport aux erreurs de syntaxe et affiche toutes les erreurs rencontrées.

7. Cliquez sur le bouton ... puis **Enregistrer** pour nommer et enregistrer votre script.

Le script est enregistré par défaut sous <LINK_DIR>/Data Services/DataQuality/Samples.

20.8.2 Débogage de scripts à l'aide de la fonction d'impression

Le logiciel dispose de fonctionnalité de débogage permettant d'imprimer les éléments suivants :

- Les valeurs de variables et de paramètres lors de l'exécution
- Le chemin d'exécution suivi dans un script

Vous pouvez utiliser la fonction d'impression pour écrire les valeurs des paramètres et variables d'un workflow dans le journal de suivi. Par exemple, dans un script, la ligne suivante :

```
print('The value of parameter $<x>: [$<x>]');
```

crée la sortie suivante dans le journal de suivi :

```
The following output is being printed via the Print function in <Session  
<job_name>.  
The value of parameter $<x>: <value>
```

20.9 Éditeur rapide et outils fonctionnels

SAP Data Services met à disposition l'éditeur rapide et d'autres outils fonctionnels pour vous aider à créer des scripts, des expressions et des fonctions personnalisées.

Avec l'éditeur rapide, il n'est plus nécessaire de saisir les noms des éléments existants comme les colonnes, les fonctions et les variables. Le logiciel vous donne accès à l'éditeur rapide et aux outils fonctionnels au moment où vous en avez besoin. Par exemple, ouvrez les deux outils dans l'éditeur de transformation Query lors de la création d'une clause WHERE. Ouvrez également l'éditeur rapide depuis l'onglet *Fonctions* dans la bibliothèque d'objets lors de la création ou de la modification d'une fonction personnalisée.

i Remarque

Parfois, nous faisons référence aux outils fonctionnels sous le nom "Assistant Fonction".

Lors de l'utilisation de l'éditeur rapide pour la création de fonctions, ouvrez la boîte de dialogue *Sélectionner une fonction* pour développer des fonctions personnalisées. La boîte de dialogue *Sélectionner une fonction* référence plusieurs catégories de fonctions dans le volet de gauche. Lorsque vous sélectionnez une catégorie, les noms des fonctions apparaissent dans le volet de droite. Sélectionnez un nom de fonction, puis cliquez sur *Suivant* pour continuer à définir cette fonction.

Informations associées

[Langage de script \[page 706\]](#)

[Fonctions personnalisées \[page 681\]](#)

[Tables de recherche et fonction lookup_ext \[page 217\]](#)

20.9.1 Accès à l'éditeur rapide

Accédez à l'éditeur rapide intégré et utilisez-le comme volet dans tout éditeur d'objet de Designer ou ouvrez-le sous forme de fenêtre distincte agrandie.

Par exemple, utilisez l'onglet *Mappage* dans une requête pour accéder à l'éditeur rapide :

1. Faites glisser une colonne d'un schéma d'entrée vers un schéma de sortie pour activer l'éditeur rapide.
2. Saisissez du texte et sélectionnez des options à l'aide du menu contextuel de l'éditeur rapide ou cliquez sur les points de suspension (...) pour ouvrir la fenêtre agrandie de l'éditeur rapide.

i Remarque

Il n'est pas possible d'ajouter des commentaires à une clause de mappage dans une transformation Query. La syntaxe suivante n'est, par exemple, pas prise en charge dans l'onglet Mappage :

```
table.column # comment
```

Le job ne s'exécutera pas et vous ne parviendrez pas à l'exporter. Utilisez plutôt la fonction de description d'objet ou d'annotation de l'espace de travail.

A l'ouverture de la fenêtre de l'éditeur rapide, le contexte de l'objet à partir duquel il a été ouvert est affiché dans la barre de titre.

Il est possible d'ouvrir l'éditeur rapide depuis les emplacements suivants :

- Onglet *Mappage* de l'éditeur Query
- Onglet *FROM* de l'éditeur Query
- Onglet *WHERE* de l'éditeur Query
- Editeur de script
- Editeur conditionnel
- Editeur de boucle While
- Editeur de fonctions personnalisées
- Assistant Fonction, page "Définir les paramètres d'entrée"
- SQL- Editeur de transformation
- Editeur Case

20.9.2 Utilisation de la liste de sélection et d'une info-bulle

1. Cliquez avec le bouton droit sur l'éditeur, affichez le menu et vérifiez que les éléments *Activer la liste de sélection* et *Activer l'info-bulle* sont sélectionnés.
2. Dans l'éditeur, saisissez au moins trois caractères ou le symbole de variable du signe du dollar (\$).

La liste de sélection s'ouvre sur l'éditeur et met en surbrillance le premier élément (de cette liste par ordre alphabétique) correspondant aux caractères saisis.

Vous pouvez aussi cliquer sur l'icône Ouvrir la liste de sélection dans la barre d'outils.

3. Double-cliquez sur un élément de la liste de sélection pour l'insérer dans l'éditeur et afficher une info-bulle associée.

L'info-bulle affiche la même description, définition ou syntaxe que si l'élément était sélectionné dans la bibliothèque de l'éditeur.

4. Complétez le script ou l'expression à l'aide de ces outils.

Par exemple, si vous complétez une fonction de recherche, l'info-bulle reste à l'écran pour vous permettre de suivre la syntaxe de la fonction. Si vous saisissez une valeur d'entrée avec un type de données erroné, l'info-bulle se ferme, ce qui indique une erreur.

20.9.3 Navigation vers une fonction

1. Développez les nœuds pour rechercher la fonction requise :
 - Les fonctions intégrées sont regroupées par type
 - Les fonctions personnalisées sont listées sous le nœud Personnalisé
 - Les fonctions et procédures stockées importées sont listées sous le nom de la banque de données utilisée pour les importer.
2. Cliquez sur une fonction, puis lisez sa description et sa syntaxe dans la zone jaune sous les onglets.
3. Lorsque vous disposez de la fonction voulue, placez-la dans l'éditeur.

20.9.4 Recherche d'une fonction

1. Sélectionnez la position de la fonction dans l'éditeur.
2. Dans l'onglet *Fonctions* de la bibliothèque de l'éditeur, sélectionnez le nœud *Rechercher*.
3. Saisissez une chaîne telle que 100.
4. Appuyez sur *Entrée* ou *Tabulation*.

Toutes les fonctions qui contiennent la chaîne sont renvoyées sous le nœud.

5. Pour placer la fonction dans l'éditeur, utilisez une des méthodes suivantes :
 - double-cliquez
 - glissez-déposez
 - cliquez avec le bouton droit sur la fonction et sélectionnez *Entrée*
 - sélectionnez la fonction et appuyez sur *Entrée* ou *Tabulation*

20.9.5 Validation

Validez une expression ou validez l'objet contenant l'expression.

Si le logiciel peut valider le contexte actuel, comme une expression, il active l'option *Valider* dans la barre d'outils et le menu contextuel. Si le logiciel ne peut pas activer l'option *Valider*, validez l'expression dans le contexte de l'objet tout entier.

1. Pour valider une expression à l'aide de l'affichage d'erreur intégré de l'éditeur rapide :

- a. sélectionnez l'icône *Valider* dans la barre d'outils ou cliquez avec le bouton droit et sélectionnez *Valider*.



Le logiciel répertorie toutes les erreurs dans un volet distinct sous l'éditeur.

- b. Double-cliquez sur chaque erreur.

L'éditeur redessine pour afficher l'emplacement de l'erreur dans votre texte.

- c. Corrigez l'erreur et poursuivez la validation jusqu'à ce que toutes les erreurs de validation soient fixées et qu'il n'y ait plus d'erreur de validation.

2. Pour valider une expression dans le contexte de l'objet tout entier :

- a. Fermez la fenêtre de l'éditeur rapide.

Le logiciel affiche l'expression dans l'éditeur rapide intégré.

- b. Dans le menu Designer, sélectionnez ► *Débogage* ► *Valider* ►.

Informations associées

[Débogage et validation \[page 714\]](#)

21 Données imbriquées

Le progiciel mappe les données imbriquées avec un schéma distinct qui est implicitement lié à une seule ligne et une seule colonne du schéma parent. Ce mécanisme s'intitule le NRDM (Nested Relational Data Modelling, modélisation de données relationnelles imbriquées).

Les données réelles ont souvent des relations hiérarchiques représentées dans une base de données relationnelle avec des schémas maître-détail utilisant des clés étrangères pour créer le mappage. Toutefois, certains jeux de données, tels que les documents XML et les IDocs SAP ERP, traitent les relations hiérarchiques à l'aide de données imbriquées.

Le NRDM permet d'afficher et de manipuler les relations hiérarchiques au sein des cibles, des transformations et des sources des flux de données.

Les commandes clients sont souvent présentées en utilisant l'imbrication : les postes d'une commande client sont liés à un seul en-tête et sont affichés à l'aide d'un schéma imbriqué. Chaque ligne du jeu de données de la commande client contient un schéma imbriqué du poste.

OrderNo	CustID	ShipTo1	ShipTo2	LineItems
9999	1001	123 State St	Town, CA	

Item	Qty	ItemPrice
001	2	10
002	4	5

21.1 Représentation de données hiérarchiques

Vous pouvez représenter les mêmes données hiérarchiques de différentes manières.

En voici des exemples :

Lignes multiples dans un seul jeu de données

Jeu de données de la commande

N° commande	IDClient	AdresseLivraison1	AdresseLivraison2	Élément	Qté	PrixÉlément
9999	1001	123 State St	Town, CA	001	2	10
9999	1001	123 State St	Town, CA	002	4	5

Jeux de données multiples liés par une jointure

Jeu de données de l'en-tête de commande

N° commande	IDClient	AdresseLivraison1	AdresseLivraison2
9999	1001	123 State St	Town, CA

Jeu de données des postes

N° commande	Elément	Qté	PrixElément
9999	001	2	10
9999	002	4	5

OU En-tête de commande N° de commande = Poste N° de commande

Données imbriquées

La méthode des données imbriquées peut s'avérer plus concise (aucune information répétée) et peut se mettre à l'échelle pour présenter un niveau supérieur de complexité hiérarchique. Par exemple, les colonnes d'un schéma imbriqué peuvent également contenir des colonnes. Il existe une instance unique de chaque schéma imbriqué pour chaque ligne à chaque niveau de la relation.

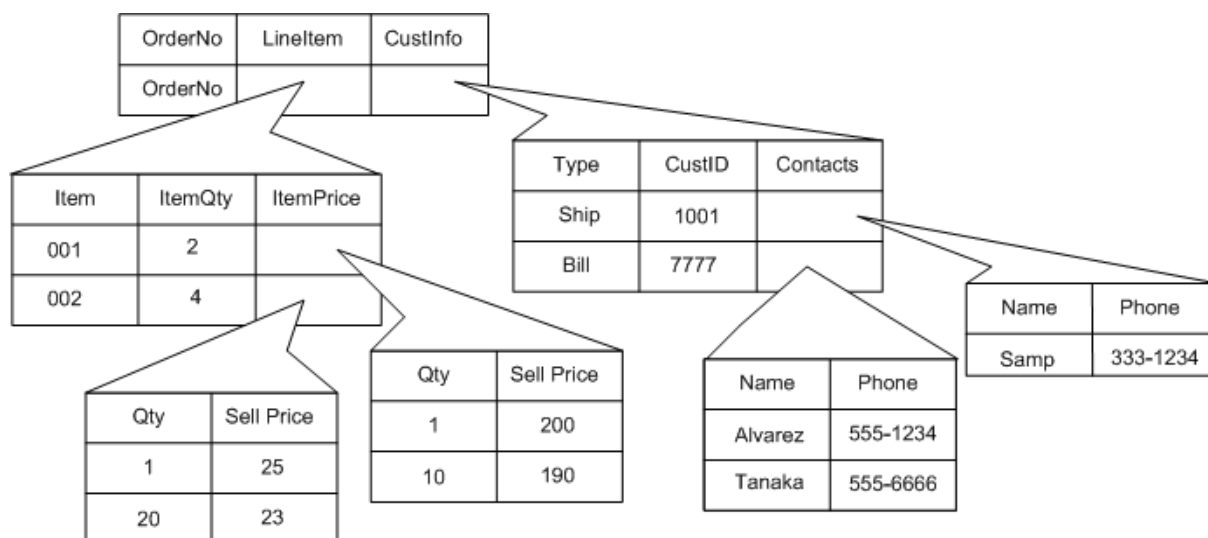
Jeu de données de la commande

OrderNo	CustID	ShipTo1	ShipTo2	LineItems
9999	1001	123 State St	Town, CA	

Item	Qty	ItemPrice
001	2	10
002	4	5

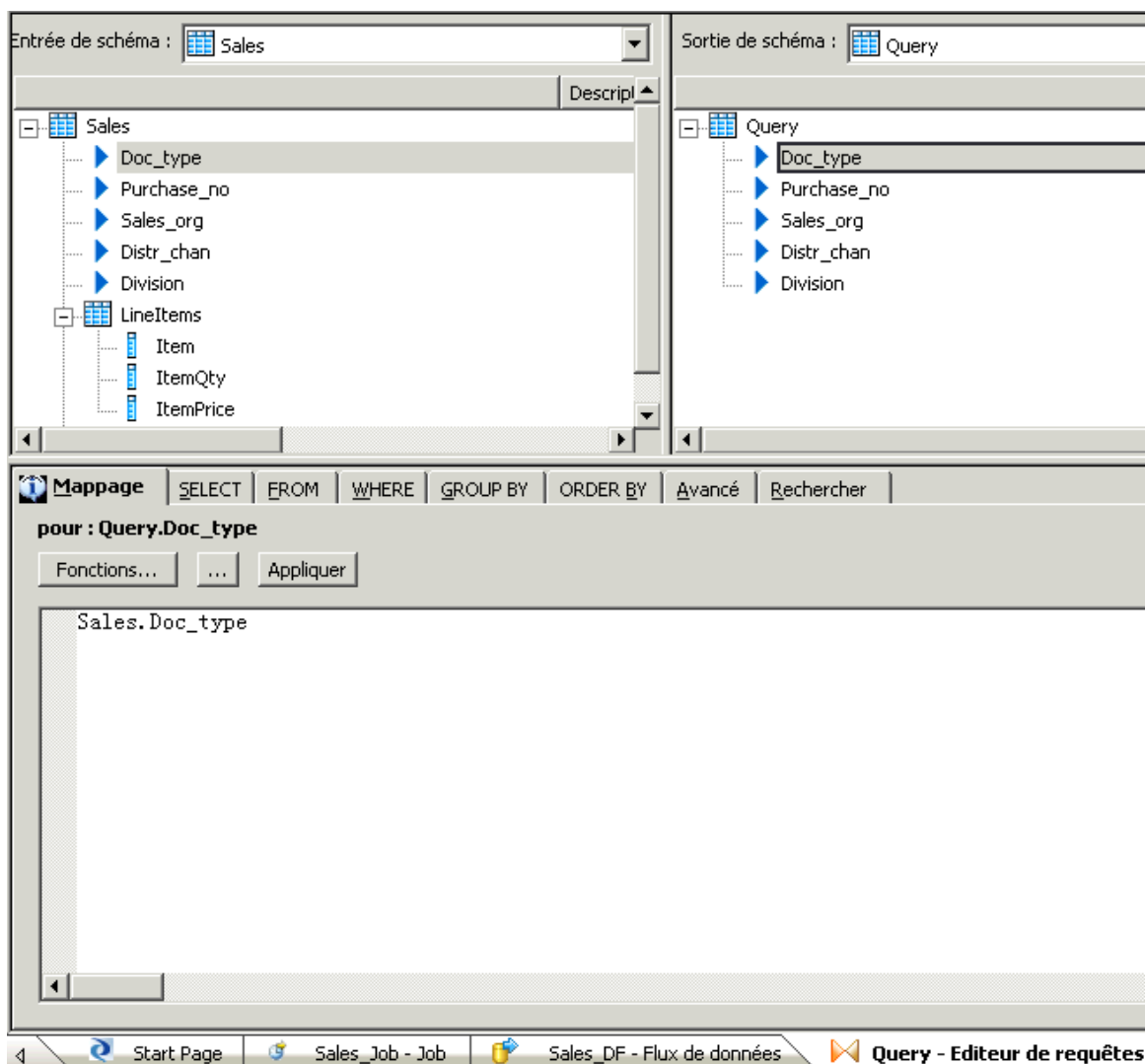
En généralisant davantage le concept des données imbriquées, chaque ligne de chaque niveau peut disposer d'un nombre quelconque de colonnes contenant des schémas imbriqués.

Jeu de données de la commande



Vous pouvez afficher la structure des données imbriquées dans les schémas d'entrée et de sortie des sources, cibles et transformations dans les flux de données. Les schémas imbriqués s'affichent avec une icône de schéma associée d'un signe plus, indiquant que l'objet contient des colonnes. La structure du schéma indique la manière dont les données sont classées.

- Ventes est le schéma de niveau supérieur.
- LinItems est un schéma imbriqué. Le signe moins précédant l'icône de schéma indique que la liste de colonnes est ouverte.
- CustInfo est un schéma imbriqué dont la liste de colonnes est fermée.



21.2 Mise en forme des documents XML

Le progiciel permet d'importer et d'exporter des métadonnées pour des documents XML (fichiers ou messages) utilisés en tant que sources ou cibles dans les jobs. Les documents XML sont des documents hiérarchiques.

Leur structure valide est stockée dans des documents aux formats distincts.

Un schéma XML (.xsd, par exemple) ou un fichier Document Type Definition (définition du type de document, .dtd) spécifient le format d'un fichier ou d'un message XML (.xml).

Lors de l'importation de métadonnées du format d'un document, celles-ci sont structurées dans le schéma interne du progiciel pour les documents hiérarchiques utilisant le modèle NRDM.

Informations associées

[Spécification schéma XML \[page 260\]](#)

[Mappage de schémas facultatifs \[page 274\]](#)

[Spécification d'options de source pour les fichiers XML \[page 272\]](#)

[Utilisation de Document Type Definitions \(DTD\) \[page 275\]](#)

[Génération de schémas DTD et XML à partir d'un schéma NRDM \[page 280\]](#)

21.2.1 Spécification schéma XML

Le progiciel prend en charge les spécifications XML Schema 1.0 du WC3.

Pour un document XML contenant des informations pour passer une commande (en-tête de commande, client et postes), le schéma XML correspondant inclut la structure de la commande et les relations entre les données.

Message avec des données

N°Commande	IDClient	AdresseLivrai- son1	AdresseLivrai- son2	Postes		
				Élément	QtéÉlément	PrixÉlément
9999	1001	123 State St	Town, CA			
				001	2	10
				002	4	5

Chaque colonne du document XML correspond à un ELEMENT ou une définition d'attribut dans le schéma XML.

Schéma XML correspondant

```
<?xml version="1.0"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="Order">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="OrderNo" type="xs:string" />
        <xs:element name="CustID" type="xs:string" />
        <xs:element name="ShipTo1" type="xs:string" />
        <xs:element name="ShipTo2" type="xs:string" />
        <xs:element maxOccurs="unbounded" name="LineItems">
          <xs:complexType>
            <xs:sequence>
              <xs:element name="Item" type="xs:string" />
              <xs:element name="ItemQty" type="xs:string" />
              <xs:element name="ItemPrice" type="xs:string" />
            </xs:sequence>
          </xs:complexType>
        </xs:element>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:schema>
```

```
</xs:sequence>
</xs:complexType>
</xs:element>
</xs:sequence>
</xs:complexType>
</xs:element>
</xs:schema>
```

21.2.2 À propos de l'importation des schémas XML

Importez les métadonnées de chaque schéma XML que vous utilisez.

La bibliothèque d'objets répertorie dans la catégorie Schémas imbriqués de l'onglet [Formats](#) les schémas XML importés.

Lors de l'importation d'un schéma XML, le logiciel lit les éléments et les attributs définis puis importe les éléments suivants :

- Structure de document
- Espace de noms
- Noms de table et de colonne
- Type de données de chaque colonne
- Type de contenu de chaque colonne
- Table imbriquée et attributs de colonne

Alors que les schémas XML effectuent une distinction entre éléments et attributs, le logiciel les importe et les convertit vers une table imbriquée et vers les attributs de colonne.

21.2.2.1 Composants Schéma XML pris en charge

SAP BusinessObjects Data Services prend en charge la quasi totalité des schémas XML valides.

Le logiciel prend en charge des fonctionnalités telles que les types abstraits et le blocage ; le logiciel importera et acceptera ces fonctions sans erreur. A l'exception des types abstraits, ce document ne traite pas de telles fonctionnalités en profondeur car elles n'ont pas d'impact direct sur les capacités du logiciel à prendre en charge les schémas XML.

Le logiciel importe les types de données Schéma XML ainsi que les noms d'élément et d'attribut et leur structure. Une fois l'importation terminée, cliquez deux fois sur un format Schéma XML depuis la bibliothèque d'objets pour afficher les noms de table et de colonne, ainsi que la structure. A partir de l'éditeur Schéma XML, cliquez avec le bouton droit sur un nom de colonne et sélectionnez Modifier les propriétés, Attributs et Types de données.

Informations associées

[Schémas XML non pris en charge \[page 262\]](#)

21.2.2.2 Schémas XML non pris en charge

Les éléments et les attributs de schéma XML suivants ne sont pas pris en charge dans SAP Data Services. Ils sont ignorés et non importés.

Composant	Description
Annotation	Le logiciel ignore les composants d'annotation appinfo et de documentation.
Attributs non natifs	Les attributs non natifs sont des attributs qui proviennent d'un espace de nom différent de celui utilisé par votre schéma XML. Le standard W3C des schémas XML permet aux utilisateurs d'ajouter des attributs non natifs aux attributs et aux éléments. Cependant, le logiciel ignore tous ces attributs.
Fichiers XDR	Le logiciel ne prend pas en charge les fichiers XDR (XML Data Reduced). Les fichiers XDR étaient utilisés comme format dans certains produits avant que le schéma XML devienne le standard. Il existe des outils tiers sur le marché qui peuvent convertir automatiquement XDR en schéma XML.

21.2.2.3 Schémas XML inclus

Un schéma XML peut être étendu en incluant des pointeurs vers d'autres fichiers de schéma XML. Pour ce faire, utilisez `<importer>`, `<inclure>` et `<redéfinir>`. Ces éléments sont définis au niveau du schéma.

La différence entre `<inclure>` et `<importer>` réside dans le fait que pour `<inclure>`, les espaces de noms doivent être identiques dans les deux schémas XML. `<Redéfinir>` est similaire à `<inclure>`, sauf que l'appelant peut redéfinir un ou plusieurs composants dans le schéma XML associé.

Lors de l'importation d'un schéma XML, SAP Data Services suit les liens vers les fichiers inclus pour définir des métadonnées supplémentaires. Les informations du schéma inclus sont enregistrées dans le référentiel afin qu'il n'y ait plus besoin d'accéder à nouveau à ces fichiers lors de l'exécution. Les inclusions peuvent être des fichiers ou des URL.

21.2.2.4 Groupes

Les schémas XML permettent de grouper des éléments, puis de faire référence au groupe. Il existe un concept similaire pour les attributs (nommé Groupe d'attributs). Dans SAP Data Services, toute référence à un groupe est remplacée par le contenu de ce groupe.

21.2.2.5 Règles d'importation des schémas XML

SAP Data Services applique les règles suivantes pour convertir un schéma XML en schéma interne du logiciel :

1. Tout élément qui contient uniquement un élément et aucun attribut devient une colonne.
2. Tout élément avec des attributs ou d'autres éléments devient une table.

3. Un attribut devient une colonne dans la table correspondant à l'élément qu'il prend en charge.
4. Toute occurrence de `<choice>`, `<sequence>` ou `<all>` utilise l'organisation donnée dans le schéma XML comme l'organisation de colonne dans le jeu de données interne.
5. Toute occurrence de `<maxOccurs>`, > de supérieure à 1 jusqu'à « non lié » devient une table avec un nom généré en interne (une table implicite).
Le nom généré en interne est le nom du parent suivi par un tiret de soulignement, puis la chaîne "nt" suivie par un numéro de séquence. Le numéro de séquence commence à 1 et augmente de 1 en 1.

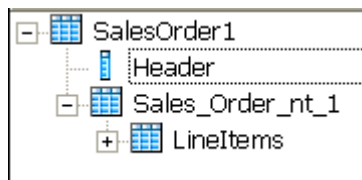
Une fois les règles appliquées, le logiciel utilise deux règles supplémentaires, sauf si cette opération affecte plusieurs lignes à un élément racine :

1. Si une table implicite contient une seule table imbriquée, alors la table implicite peut être éliminée et la table imbriquée peut être attachée directement au parent de la table implicite.

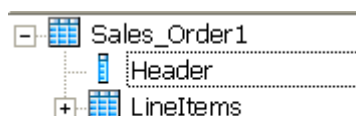
Par exemple, l'élément SalesOrder peut être défini comme suit dans un schéma XML :

```
<xs:element name="SalesOrder">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="Header"/>
      <xs:element name="LineItems" minOccurs="0"
maxOccurs="unbounded"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
```

Une fois converti dans le logiciel, l'élément LineItems avec MaxOccurs= "non lié" deviendrait une table implicite sous la table SalesOrder. L'élément LineItems serait une table imbriquée sous la table implicite.



Puisque la table implicite contient une seule table imbriquée, le format supprimerait la table implicite.



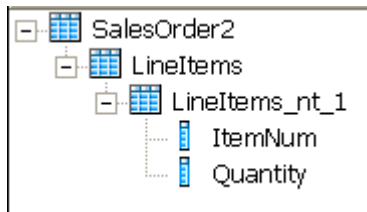
2. Si une table imbriquée contient une seule table implicite, alors la table implicite peut être éliminée et ses colonnes peuvent être placées directement sous la table imbriquée.

Par exemple, la table imbriquée LineItems peut être définie comme suit dans un schéma XML :

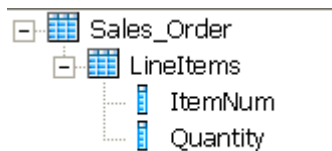
```
<xs:element name="SalesOrder">
  <xs:element name="LineItems" minOccurs="0"
maxOccurs="unbounded">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="ItemNum"/>
        <xs:element ref="Quantity"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
```

```
</xs:element>
```

Une fois converti dans le logiciel, le regroupement avec MaxOccurs= "unbounded" deviendrait une table implicite sous la table LineItems. Les éléments ItemNum et Quantity deviendrait des colonnes sous la table implicite.



Puisque la table imbriquée LineItems contient une seule table implicite, le format supprimerait la table implicite.



21.2.2.6 Limitations

Si une définition de schéma XML contient les éléments ou attributs suivants, SAP Data Services les importe avec les limitations suivantes :

- Élément Any ou anyAttribute

Il est possible d'importer un schéma XML qui contient un élément Any ou anyAttribute, ou les deux, mais le format créé par le logiciel n'affiche pas l'élément Any ou anyAttribute.

Par conséquent, le logiciel ignore le contenu de l'élément Any ou anyAttribute lorsqu'il lit un document d'instance XML. Lorsqu'un élément a le type anyType, le logiciel traite tout ce qui se trouve dedans comme une chaîne et ne reconnaît pas les sous-éléments à l'intérieur.

- Contenu mixte

La structure d'un schéma XML consiste généralement d'éléments qui contiennent des sous-éléments, et les sous-éléments au niveau le plus bas contiennent des données de caractères. Cependant, la définition d'un schéma XML permet que les données de caractères apparaissent à côté des sous-éléments, et les données de caractères ne sont pas limitées au niveau le plus bas. Pour les documents d'instance qui contiennent du contenu mixte, le logiciel ignore les données de caractères entre deux sous-colonnes, mais capture les valeurs des sous-colonnes.

21.2.2.7 Importer un schéma XML

1. Dans la bibliothèque d'objets, cliquez sur l'onglet *Format*.

2. Cliquez avec le bouton droit de la souris sur l'icône *Schémas imbriqués* puis sélectionnez ► *Nouveau* ► *Schéma XML* .
3. Saisissez les paramètres pour les schémas XML que vous importez, en fonction des descriptions d'option ci-dessous.

Option	Description
<i>Nom du format</i>	Saisissez le nom que vous voulez utiliser pour le format dans le logiciel.
<i>Nom de fichier/URL</i>	Saisissez ou naviguez vers le nom de fichier du schéma XML ou son adresse URL. i Remarque Si le Job Server se trouve sur un ordinateur autre que Designer Data Services, vous pouvez utiliser <i>Parcourir</i> pour indiquer le chemin d'accès du fichier. Vous devez saisir le chemin. Vous pouvez saisir un chemin d'accès absolu ou un chemin d'accès relatif, mais le Job Server doit pouvoir y accéder.
<i>Espace de noms</i>	Sélectionnez un schéma XML importé dans la liste déroulante uniquement si le nom de l'élément racine n'est pas unique dans le schéma XML. i Remarque Lorsque vous importez un schéma XML pour un job service Web en temps réel, utilisez un espace de noms cible unique pour le schéma. Lorsque Data Services génère le fichier WSDL pour un job en temps réel avec un schéma source ou cible qui n'a pas d'espace de noms cible, il ajoute un espace de noms cible généré automatiquement à la section de types du schéma XML. Ce processus peut réduire la performance car Data Services doit supprimer les informations de l'espace de noms de la demande de service Web au cours du traitement, puis joindre à nouveau les informations correctes de l'espace de noms avant de renvoyer la réponse au client.
<i>Nom de l'élément racine</i>	Sélectionnez le nom du noeud principal que vous souhaitez importer dans la liste déroulante. Le logiciel importe uniquement les éléments du schéma XML appartenant à ce noeud ou à tout sous-noeud.
<i>Niveau circulaire</i>	Spécifiez le nombre de niveaux uniquement si le schéma XML contient des éléments récursifs (l'élément A contient l'élément B, l'élément B contient l'élément A). Cette valeur doit correspondre au nombre de niveaux récursifs dans le contenu du schéma XML. Dans le cas contraire, le job qui utilise ce schéma XML va échouer.

Option	Description
<i>Taille varchar par défaut</i>	Configurez le logiciel pour importer des chaînes en tant que varchar de toute taille. La valeur par défaut est 1024.

4. Cliquez sur *OK*.

Après l'importation d'un schéma XML, vous pouvez modifier les propriétés de sa colonne, comme le type de données, à l'aide de l'onglet *Général* de la fenêtre Propriétés des colonnes. Il est également possible d'afficher et de modifier les attributs des colonnes et des tables imbriquées à partir de la fenêtre Propriétés des colonnes.

21.2.2.8 Mappage des types de données

SAP Data Services importe les types de données pour les éléments et les attributs des schémas XML.

Il existe deux types de types de données intégrés : les types de données Primitif et les types de données Dérivé (dérivés des primitifs). Chaque type de donnée a les valeurs suivantes définies : espace, espace lexical, et facette de contrainte.

Si la facette de contrainte `<length>` est absente lorsque les métadonnées sont importées dans le logiciel, le varchar par défaut (1024) s'applique. De la même façon, pour un nombre, les valeurs par défaut 28 et 2 s'appliquent pour la `<precision>` et l'`<échelle>`. Toutes les autres facettes comme `<minInclusive>`, `<maxInclusive>`, `<minLength>` sont importées comme attributs de colonne. Les valeurs d'énumération sont également importées comme attributs de colonne.

21.2.2.9 Types primitifs

La table suivante répertorie les types de schémas XML primitifs, des exemples et le type de données correspondant dans SAP Data Services. Les facettes de contraintes utilisées sont indiquées en gras.

Type de schéma XML	Exemple	Type de données
AnyURI	http://www.exemple.com/	Varchar(lon) : lon = <i>longueur</i> (en caractères)
Base64Binary	GpM7	Varchar(lon) : lon = <i>longueur</i> (en octets)
Booléen	{true, false, 0, 1}	Varchar(5)
Date	CCYY-MM-DD	Date/heure
DateTime	Format = CCYY-MM-DD HH:MM:SS	Date/heure
Décimal	7,999	Decimal(p, s) : p = <i>totalDigits</i> , un maximum de 28 et s = <i>fractionDigits</i> , par défaut =28,2
Double	Point flottant 64 bits	Double (dans le logiciel, il n'existe aucune différence entre réel et double)

Type de schéma XML	Exemple	Type de données
Durée	PIY2M3DT10H30M	Varchar(64)
Float	Point flottant 32 bits, 12.78e-2	Real
gDay		Varchar(12)
gMonth		Varchar(12)
gMonthDay		Varchar(12)
gYear		Varchar(12)
gYearMonth	CCYY-MM grégorien	Varchar(12)
HexBinary	0FB7	Varchar(lon) : lon = <i>longueur</i> (en octets)
Notation		N/A
Qname	po:USAddress	Varchar(lon) : lon = <i>longueur</i> (en caractères)
Chaîne	"Hello World"	Varchar(lon) : lon = <i>longueur</i> (en caractères)
Time	HH:MM:SS	Date/heure

21.2.2.10 Types dérivés

La table suivante répertorie les types de schémas XML dérivés prédéfinis, des exemples et le type de données correspondant dans SAP Data Services. Les facettes de contraintes utilisées sont indiquées en gras.

Type de schéma XML	Exemple	Type de données
NormalizedString	[aucune tabulation/CR/LF dans une chaîne]	Varchar(lon) : lon = <i>longueur</i> (en caractères)
Jeton		Varchar(lon) : lon = <i>longueur</i> (en caractères)
Langue	En-GB, en-US, fr	Varchar(lon) : lon = <i>longueur</i> (en caractères)
NMTOKEN	US, Brésil	Varchar(lon) : lon = <i>longueur</i> (en caractères)
NMTOKENS	Brésil Canada Mexique	Varchar(lon) : lon = <i>longueur</i> (en caractères)
Nom	ShipTo	Varchar(lon) : lon = <i>longueur</i> (en caractères)
NCName	USAddress	Varchar(lon) : lon = <i>longueur</i> (en caractères)
ID		Varchar(lon) : lon = <i>longueur</i> (en caractères)

Type de schéma XML	Exemple	Type de données
IDREF		Varchar(lon) : lon = <i>longueur</i> (en caractères)
IDREFS		Varchar(lon) : lon = <i>longueur</i> (en caractères)
ENTITY		Varchar(lon) : lon = <i>longueur</i> (en caractères)
ENTITIES		Varchar(lon) : lon = <i>longueur</i> (en caractères)
Integer		Int
NonPositiveInteger		Int
NegativeInteger		Int
Long		Decimal 28,8
Int		Int
Short		Int
Octet		Int
NonNegativeInteger		Int
UnsignedLong		Long
UnsignedShort		Int
UnsignedByte		Int
PositiveInteger		Int
AnyType (type ur)	type non spécifié	Varchar(255)

21.2.2.11 Types définis par l'utilisateur

Les types définis par l'utilisateur sont des attributs de schéma XML avec un espace de nom qui ne correspond pas aux schémas XML. Le standard W3C pour les schémas XML utilise un élément SimpleType pour un type défini par l'utilisateur.

Lorsque SAP Data Services trouve un type défini par l'utilisateur, il recherche le type de base et l'utilise pour affecter un type de données pour l'élément. Par exemple, si un élément X est de type NuméroTéléphone, alors son type dans le logiciel est `varchar(8)`.

Certains types simples sont basés sur d'autres types simples. Dans ce cas, le logiciel remonte jusqu'au type de base.

21.2.2.12 Types de liste

Les schémas XML ont des types de liste. Lorsqu'il rencontre une liste, SAP Data Services convertit le type de données correspondant de la liste en `varchar(1024)`. Tous les éléments de la liste sont placés dans la valeur de cette colonne comme une chaîne (exactement tel qu'elle est représentée dans le XML).

21.2.2.13 Types d'union

Un type d'union permet à une valeur d'attribut ou d'élément d'être une ou plusieurs instances d'un type issu de l'union de plusieurs types de liste et du type primitif. Lorsqu'il rencontre une union, SAP Data Services convertit le type de données correspondant de l'union en `varchar(1024)`.

21.2.2.14 Afficher et modifier les attributs des colonnes et des tables imbriquées pour le schéma XML

1. Dans la bibliothèque d'objets, cliquez sur l'onglet *Formats*.
2. Développez la catégorie *Schéma imbriqué*.
3. Cliquez deux fois sur un nom de schéma XML.

La fenêtre de l'*Éditeur de format XML* s'ouvre dans l'espace de travail. Les parties *Format du schéma XML* se trouvent dans le volet supérieur de la fenêtre.

La colonne Type affiche les types de données qu'utilise le logiciel lorsqu'il importe les métadonnées du document XML.

4. Cliquez deux fois sur une table imbriquée ou une colonne et sélectionnez l'onglet *Attributs* de la boîte de dialogue *Propriétés* pour afficher ou modifier les attributs de schéma XML.

21.2.3 Types de données abstraits

Lorsque vous créez une cible de fichier ou message XML à l'aide d'un schéma XML qui contient des éléments comportant des types de données abstraits, vous devez définir la valeur correcte de l'attribut "xsi:type" pour générer une sortie XML valide. En tant que concepteur de flux de données, vous devez savoir lequel des divers types dérivés est correct pour un élément donné.

i Remarque

Par défaut, tous les éléments ayant des types de données abstraits comportent un attribut nommé `xsi:type`.

Lors de l'utilisation de schémas XML comportant des espaces de noms, vous devez inclure l'espace de nom correct dans le nom du type. Pour obtenir la bonne balise d'espace de nom, examinez les balises d'espace de nom générées par SAP Data Services (habituellement `ns1`, `ns2`, etc.), puis utilisez la balise qui représente l'espace de nom dans lequel se trouve le type.

Exemple :

Supposons qu'un élément nommé Publication comporte un type abstrait nommé PublicationType. Lorsque le logiciel importe cet élément, il ajoute une colonne supplémentaire nommée "xsi:type" comme enfant de Publication. Vous devez alors définir l'expression pour que cette colonne soit égale au type de résultat attendu (par exemple, BookType). Pour ajouter la balise adéquate, exécutez d'abord votre job, puis notez les noms de balises générés. Dans cet exemple, ns1 pour un espace de nom nommé <http://www.bookworld.com/>. Ici, l'expression de xsi:type serait "ns1:BookType".

21.2.3.1 Import abstract data types

When you import an abstract data type element, there are specific requirements that you should know about.

Ensure that you prepare for importing abstract data types in the following ways:

- When an element is defined as abstract, ensure that a member of the element substitution group appears in the instance document.
- When a type is defined as abstract, ensure that the instance document uses a type derived from it, identified by the xsi:type attribute.

❖ Exemple

An abstract element named PublicationType has a substitution group that consists of complex types such as MagazineType, BookType, and NewspaperType. The software default behavior is to select all complex types in the substitution group or all derived types for the abstract type. However, you can choose to select a subset.

21.2.3.2 Limiter le nombre de types dérivés à importer pour un type abstrait

1. Dans la fenêtre *Importer format de schéma XML*, lorsque vous saisissez le nom du fichier ou l'adresse URL d'un schéma XML contenant un type abstrait, le bouton de commande *Type abstrait* est activé.

Par exemple, l'extrait suivant d'un xsd définit l'élément PublicationType comme étant abstrait avec des types dérivés BookType et MagazineType :

```
<xsd:complexType name="PublicationType" abstract="true">
  <xsd:sequence>
    <xsd:element name="Title" type="xsd:string"/>
    <xsd:element name="Author" type="xsd:string" minOccurs="0"
maxOccurs="unbounded"/>
    <xsd:element name="Date" type="xsd:gYear"/>
  </xsd:sequence>
</xsd:complexType>
<xsd:complexType name="BookType">
  <xsd:complexContent>
    <xsd:extension base="PublicationType">
      <xsd:sequence>
        <xsd:element name="ISBN" type="xsd:string"/>
        <xsd:element name="Publisher" type="xsd:string"/>
      </xsd:sequence>
    </xsd:extension>
  </xsd:complexContent>
</xsd:complexType>
```

```

        </xsd:extension>
      /xsd:complexContent>
    </xsd:complexType>
    <xsd:complexType name="MagazineType">
      <xsd:complexContent>
        <xsd:restriction base="PublicationType">
          <xsd:sequence>
            <xsd:element name="Title" type="xsd:string"/>
            <xsd:element name="Author" type="xsd:string"
minOccurs="0" maxOccurs="1"/>
            <xsd:element name="Date" type="xsd:gYear"/>
          </xsd:sequence>
        </xsd:restriction>
      </xsd:complexContent>
    </xsd:complexType>

```

2. Pour sélectionner un sous-ensemble de types dérivés pour un type abstrait, cliquez sur le bouton [Type abstrait](#) et effectuez les actions suivantes :
 - a. Dans la zone déroulante de la case [Type abstrait](#), sélectionnez le nom du type abstrait.
 - b. Cochez les cases en regard de chaque nom de type dérivé que vous voulez importer.
 - c. Cliquez sur [OK](#).

i Remarque

Lorsque vous modifiez votre format de schéma XML, le logiciel sélectionne tous les types dérivés pour le type abstrait par défaut. En d'autres termes, le sous-ensemble que vous avez précédemment sélectionné n'est pas conservé.

21.2.4 Importation de groupes de substitution

Un schéma XML utilise des groupes de substitution pour affecter des éléments à un groupe spécifique d'éléments qui peut être substitué pour un élément spécifique nommé appelé élément principal.

La liste des groupes de substitution peut avoir des centaines ou même des milliers de membres mais une application utilise généralement uniquement un nombre limité d'entre eux. Tous les groupes de substitution sont sélectionnés par défaut mais vous pouvez décider de sélectionner un sous-ensemble.

21.2.5 Limiter le nombre de groupes de substitution à importer

Vous pouvez sélectionner un sous-ensemble de groupes de substitution à importer.

1. Dans la fenêtre [Importer format de schéma XML](#), lorsque vous saisissez le nom du fichier ou l'adresse URL d'un schéma XML contenant des groupes de substitution, le bouton [Groupe de substitution](#) est activé.

Par exemple, l'extrait suivant d'un xsd définit l'élément `PublicationType` avec des groupes de substitution `MagazineType`, `BookType`, `AdsType` et `NewspaperType` :

```

<xsd:element name="Publication" type="PublicationType"/>
  <xsd:element name="BookStore">
    <xsd:complexType>

```

```

        <xsd:sequence>
            <xsd:element ref="Publication" maxOccurs="unbounded"/>
        </xsd:sequence>
    </xsd:complexType>
</xsd:element>
<xsd:element name="Magazine" type="MagazineType"
substitutionGroup="Publication"/>
<xsd:element name="Book" type="BookType" substitutionGroup="Publication"/>
<xsd:element name="Ads" type="AdsType" substitutionGroup="Publication"/>
<xsd:element name="Newspaper" type="NewspaperType"
substitutionGroup="Publication"/>

```

2. Cliquez sur le bouton *Groupe de substitution*.
 - a. Dans la liste déroulante, dans la case *Groupe de substitution*, cochez le nom du groupe de substitution.
 - b. Cochez les cases devant chaque nom de groupe de substitution que vous voulez importer.
 - c. Cliquez sur *OK*.

i Remarque

Lorsque vous modifiez votre format de schéma XML, le logiciel sélectionne tous les éléments pour le groupe de substitution par défaut. En d'autres termes, le sous-ensemble que vous avez précédemment sélectionné n'est pas conservé.

21.2.6 Spécification d'options de source pour les fichiers XML

Créer un flux de données avec un fichier XML comme fichier source.

Une fois les métadonnées importées pour les documents XML (fichiers ou messages), vous créez un flux de données pour utiliser les documents XML en tant que sources ou cibles dans les jobs.

Suivez ces étapes pour créer un flux de données avec un fichier XML source :

1. Dans la catégorie *Schéma imbriqué* de l'onglet *Format* de la bibliothèque d'objets, glissez le schéma XML qui définit votre fichier XML source dans votre flux de données.
2. Placez une requête dans le flux de données et connectez la source XML à l'entrée de la requête.
3. Cliquez deux fois sur la source XML dans l'espace de travail pour ouvrir l'Editeur de fichiers source XML.
4. Spécifiez le nom du fichier source XML dans la zone de texte *Fichier XML*.

Informations associées

[Lecture de plusieurs fichiers XML en une fois \[page 273\]](#)

[Identification des noms de fichiers sources \[page 151\]](#)

21.2.6.1 Lecture de plusieurs fichiers XML en une fois

Le logiciel peut lire plusieurs fichiers avec le même format à partir d'un seul répertoire à l'aide d'un seul objet source.

Configurez un flux de données selon les indications de [Spécification d'options de source pour les fichiers XML \[page 272\]](#).

1. Double-cliquez sur le fichier XML source dans le flux de données correspondant afin d'ouvrir la fenêtre de l'*Éditeur de fichiers source de schémas imbriqués* dans l'espace de travail.
2. Dans l'onglet *Source* de la partie inférieure de la fenêtre, saisissez un nom de fichier contenant un caractère générique (* ou ?) dans l'option *Fichier*.

Exemple :

D:\orders\1999?????.xml lit tous les fichiers .xml contenant « 1999 » dans leur nom à partir du répertoire spécifié.

D:\orders*.xml lit tous les fichiers avec l'extension .xml à partir du répertoire spécifié.

21.2.6.2 Identification des noms de fichiers source

Vous pouvez souhaiter identifier le fichier XML source pour chaque ligne de votre sortie de source dans les situations suivantes :

- Vous avez spécifié un caractère générique pour lire plusieurs fichiers source en une fois.
- Vous chargez à partir de plusieurs fichiers source à différents jours.

21.2.6.2.1 Identifier le fichier XML source pour chaque ligne dans la cible

Configurez un flux de données selon les indications de [Spécification d'options de source pour les fichiers XML \[page 272\]](#).

1. Double-cliquez sur le fichier XML source dans le flux de données correspondant afin d'ouvrir la fenêtre de l'*Éditeur de fichiers source de schémas imbriqués* dans l'espace de travail.
2. Dans l'onglet *Source* de la partie inférieure de la fenêtre, sélectionnez *Inclure la colonne Nom du fichier*.

Cette option génère une colonne appelée DI_FILENAME pour contenir le nom du fichier XML source.

3. Ouvrez la transformation Query dans le flux de données. Dans l'Éditeur Query, mappez la colonne DI_FILENAME de l'entrée de schéma vers la sortie de schéma.

Lors de l'exécution du job, la colonne cible DI_FILENAME contiendra le nom du fichier XML source pour chaque ligne de la cible.

21.2.7 Mappage de schémas facultatifs

Il est possible de spécifier rapidement un mappage par défaut pour les schémas facultatifs sans devoir construire manuellement une table imbriquée vide pour chaque schéma facultatif dans la transformation Query. De même, lors de l'importation de schémas XML (via des fichiers DTD ou XSD), le progiciel marque automatiquement les tables imbriquées comme facultatives si l'option correspondante a été définie dans le fichier DTD ou XSD. Le progiciel conserve cette option lorsque vous copiez et collez des schémas dans les transformations Query.

Cette fonctionnalité est particulièrement utile lorsque vous disposez de schémas XML très volumineux et de nombreux niveaux imbriqués dans les jobs. Lorsque vous rendez une colonne de schéma facultative et ne lui affectez aucun mappage, le progiciel instance automatiquement la table imbriquée lors de l'exécution du job.

Même si un élément de schéma est marqué comme facultatif, il est toujours possible de lui affecter un mappage. Il faut pour cela programmer de manière appropriée le bloc de sous-requête correspondant avec une logique d'application spécifiant la manière dont le progiciel doit générer la sortie. Toutefois, si vous modifiez une partie du bloc de sous-requête, le bloc de requête obtenu doit être complet et conforme aux règles de validation normales requises pour un bloc de requête imbriqué. Vous devez mapper tout schéma de sortie non facultatif vers un bloc de requête imbriqué valide. Le progiciel génère une valeur nulle dans l'emplacement de la liste de projets correspondant de l'ATL pour tout schéma facultatif sans bloc de sous-requête défini et associé.

21.2.7.1 Rendre une table imbriquée "facultative"

1. Cliquez avec le bouton droit de la souris sur une table imbriquée, puis sélectionnez *Facultatif* pour l'activer. Pour la désactiver, cliquez de nouveau avec le bouton droit de la souris sur la table imbriquée, puis sélectionnez à nouveau *Facultatif*.
2. Vous pouvez également cliquer avec le bouton droit de la souris sur une table imbriquée et sélectionner *Propriétés*, puis ouvrir l'onglet Attributs et définir la valeur d'attribut *Table facultative* sur *oui* ou *non*. Cliquez sur *Appliquer* puis sur *OK* pour définir la valeur.

i Remarque

Si la valeur Table facultative est différente de oui ou de non, il est impossible de marquer la table imbriquée comme facultative

Lorsque vous exécutez un job avec une table imbriquée définie comme facultative et que vous n'avez pas défini de colonne ni de table imbriquée sous cette table, le progiciel génère un ATL spécifique et n'exécute pas la validation de l'interface utilisateur pour cette table imbriquée.

Exemple :

```
CREATE NEW Query ( EMPNO int KEY ,
ENAME varchar(10),
JOB varchar (9)
NT1 al_nested_table ( DEPTNO int KEY ,
DNAME varchar (14),
NT2 al_nested_table (C1 int) ) SET("Optional
Table"= 'yes') )
AS SELECT EMP.EMPNO, EMP.ENAME, EMP.JOB,
NULL FROM EMP, DEPT;
```


i Remarque

Vous ne pouvez pas marquer comme facultatifs des schémas de niveau supérieur, des tables non imbriquées ou des tables imbriquées contenant des appels de fonction.

21.2.8 Utilisation de Document Type Definitions (DTD)

Un fichier DTD peut spécifier le format d'un document XML (fichier ou message). Le DTD décrit les données contenues dans le document XML et les relations entre les éléments dans les données.

Pour un document XML contenant des informations pour passer une commande (en-tête de commande, client et poste), le DTD correspondant inclut la structure de la commande et les relations entre les données.

Message avec des données

N°Commande	IDClient	AdresseLivrai- son1	AdresseLivrai- son2	Postes		
				Élément	QtéÉlément	PrixÉlément
9999	1001	123 State St	Town, CA			
				001	2	10
				002	4	5

Chaque colonne du document XML correspond à une définition d'ELEMENT.

Définition DTD correspondante

```
<?xml encoding="UTF-8"?>
<!ELEMENT Order (OrderNo, CustID, ShipTo1, ShipTo2, LineItems)>
<!ELEMENT OrderNo (#PCDATA)>
<!ELEMENT CustID (#PCDATA)>
<!ELEMENT ShipTo1 (#PCDATA)>
<!ELEMENT ShipTo2 (#PCDATA)>
<!ELEMENT LineItems (Item, ItemQty, ItemPrice)>
<!ELEMENT Item (#PCDATA)>
<!ELEMENT ItemQty (#PCDATA)>
<!ELEMENT ItemPrice (#PCDATA)>
```

Importez les métadonnées de chaque DTD que vous souhaitez utiliser. La bibliothèque d'objets répertorie les DTD importés dans l'onglet [Formats](#).

Vous pouvez importer des métadonnées de tout fichier XML (avec une référence à un DTD) ou DTD existant. Si vous importez les métadonnées d'un fichier XML, le progiciel extrait automatiquement le DTD pour ce fichier XML.

Lors de l'importation d'un DTD, le logiciel lit les éléments et attributs définis. Il ignore d'autres parties de la définition, telles que le texte et les commentaires. Cela permet de modifier les données XML importées ainsi que le type de données au besoin.

21.2.8.1 Règles relatives à l'importation des DTD

SAP Data Services utilise les règles de conversion pour convertir une DTD en schéma interne.

SAP Data Services applique les règles suivantes pour convertir une DTD en schéma interne :

- Tout élément contenant un PCDATA uniquement et aucun attribut devient une colonne.
- Tout élément contenant des attributs ou d'autres éléments (ou au format mixte) devient une colonne.
- Un attribut devient une colonne dans la table correspondant à l'élément qu'il prend en charge.
- Toute occurrence des opérateurs de choix (|) ou de séquence (|) utilise le classement donné dans la DTD en tant que classement de colonne dans le jeu de données interne.
- Toute occurrence d'entités multiples, telles que (*) ou (+), devient une table dont le nom est généré en interne (une table implicite).
- Le nom généré en interne est le nom du parent suivi par un trait de soulignement, puis la chaîne "nt" suivie d'un numéro de séquence. Le numéro de séquence commence à 1 et augmente de 1 en 1.

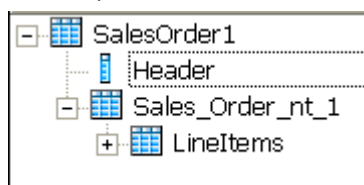
Une fois les règles appliquées, le logiciel utilise deux règles supplémentaires, sauf si cette opération affecte plusieurs lignes à un élément racine :

- Si une table implicite contient une seule table imbriquée, alors la table implicite peut être éliminée et la table imbriquée peut être attachée directement au parent de la table implicite.

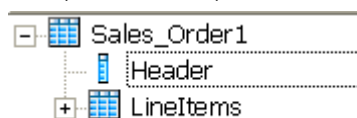
Par exemple, l'élément SalesOrder peut être défini comme suit dans une DTD :

```
<!ELEMENT SalesOrder (Header, LineItems*)>
```

Lorsqu'il est converti dans le logiciel, l'élément LineItems contenant l'opérateur Zéro ou plus devient une table implicite sous la table SalesOrder. L'élément LineItems devient alors une table imbriquée sous la table implicite.



Puisque la table implicite contient une seule table imbriquée, le format supprimerait la table implicite.

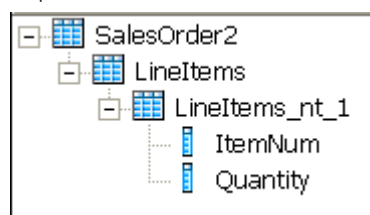


- Si une table imbriquée contient une seule table implicite, alors la table implicite peut être éliminée et ses colonnes peuvent être placées directement sous la table imbriquée.

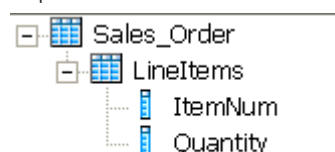
Par exemple, la table imbriquée LineItems peut être définie comme suit dans une DTD :

```
<!ELEMENT LineItems (ItemNum, Quantity)*>
```

Lorsqu'il est converti dans le logiciel, le regroupement contenant l'opérateur Zéro ou plus devient une table implicite sous la table LineItems. Les éléments ItemNum et Quantity deviennent des colonnes sous la table implicite.



Puisque la table imbriquée LineItems contient une seule table implicite, le format supprimerait la table implicite.



21.2.8.2 Importer un format de schéma XML ou DTD

1. Dans la bibliothèque d'objets, cliquez sur l'onglet *Format*.
2. Cliquez avec le bouton droit de la souris sur l'icône *Schémas imbriqués*, puis sélectionnez **Nouveau > DTD**.
3. Saisissez les paramètres dans la fenêtre Importer le format DTD :
 - Dans la zone *Nom de la définition DTD*, saisissez le nom que vous souhaitez attribuer au format DTD importé dans le logiciel.
 - Saisissez le fichier spécifiant le DTD que vous souhaitez importer.

i Remarque

Si le Job Server se trouve sur un autre ordinateur que Designer, il est impossible d'utiliser Parcourir pour indiquer le chemin du fichier. Vous devez saisir le chemin. Vous pouvez saisir un chemin d'accès absolu ou un chemin d'accès relatif, mais le Job Server doit pouvoir y accéder.

- Si vous importez un fichier XML, sélectionnez *XML* pour l'option *Type de fichier*. Si vous importez un fichier DTD, sélectionnez l'option *DTD*.
 - Dans la zone *Nom de l'élément racine*, sélectionnez le nom du nœud principal que vous souhaitez importer. Le progiciel importe uniquement les éléments du DTD appartenant à ce nœud ou à tout sous-nœud.
 - Si le DTD contient des éléments récursifs (l'élément A contient l'élément B et l'élément B contient l'élément A), spécifiez le nombre de ses niveaux en saisissant une valeur dans la zone *Niveau circulaire*. Cette valeur doit correspondre au nombre de niveaux récursifs dans le contenu du DTD. Dans le cas contraire, le job qui utilise ce DTD va échouer.
 - Vous pouvez configurer le progiciel pour importer des chaînes en tant que varchar de toute taille. Varchar 1024 est la valeur par défaut.
4. Cliquez sur *OK*.

Après l'importation d'un DTD, vous pouvez modifier les propriétés de sa colonne, comme le type de données, à l'aide de l'onglet *Général* de la fenêtre *Propriétés des colonnes*. Il est également possible d'afficher et de modifier les attributs des colonnes et des tables imbriquées du DTD à partir de la fenêtre *Propriétés des colonnes*.

21.2.8.3 Considérations en matière de conception

Utilisez des techniques de conception pour améliorer les performances et ajuster les modèles de données relationnelles imbriquées pour les DTD.

Les zones suivantes vous permettent d'améliorer les performances et d'ajuster les résultats du modèle de données relationnelles imbriquées pour une DTD donnée :

- **Récursion**
Si la DTD contient un élément qui utilise un élément ascendant dans sa définition, SAP Data Services développe la définition de l'ascendant pour un nombre de niveaux fixe. Par exemple, pour la définition suivante de l'élément "A" :
A : B, C
B : E, F
F : A, H
Le logiciel produit une table pour l'élément "F" qui inclut une extension de "A". Dans cette seconde extension de "A", "F" réapparaît, jusqu'à ce que le nombre de niveaux définis soit atteint. Dans l'extension finale de "A", l'élément "F" apparaît uniquement avec l'élément "H" dans sa définition.
- **Noms de colonnes répétés**
Le logiciel n'accepte pas que plusieurs colonnes d'un même niveau de table portent le même nom. Si le schéma interne que le logiciel produit à partir d'une DTD contient des noms dupliqués, le logiciel ajoute un suffixe à chaque instance de nom dupliqué pour garantir l'unicité des noms de colonnes. Le suffixe est un tiret de soulignement suivi d'un numéro de séquence, qui commence par 1 par incrément de 1.
Une DTD peut produire des noms dupliqués lorsqu'elle contient un élément répété à un niveau ou un élément possédant une valeur scalaire avec un attribut du même nom.
- **DTD ambiguë**
Vous pouvez créer une DTD si le logiciel ne possède pas assez d'informations pour prendre une décision lors de la génération du jeu de données internes. Le logiciel réagit à une DTD ambiguë en renvoyant une erreur pour le message XML lors de l'exécution. Exemple de définition ambiguë :

DTD	A: ((B, (C) *) (B, (D) *)) +
Schéma dans Data Services	A: (B, (C) *, B, (D) *) *
Sortie XML	A: text <D>1</D> <D>2</D>

Le logiciel utilisera les données de l'élément B pour remplir la première colonne B, puis les données de l'élément D pour remplir l'élément D. Si ces données sont reconverties en XML, elles ne seront plus valides pour la DTD.

Métadonnées

Si vous supprimez une DTD de la bibliothèque d'objets, les sources ou cibles du fichier ou message XML basées sur ce format ne sont pas valides. Le logiciel marque les objets source ou cible à l'aide d'une icône indiquant que les appels ne sont plus valides.



Pour restaurer les objets non valides, vous devez supprimer la source ou la cible et la remplacer par une source ou cible basée sur une DTD existante.

21.2.8.4 Afficher et modifier les attributs des colonnes et des tables imbriquées des DTD

1. Dans la bibliothèque d'objets, cliquez sur l'onglet *Formats*.
2. Développez la catégorie *Schémas imbriqués*.
3. Cliquez deux fois sur le nom d'un DTD.

La fenêtre *Format DTD* s'affiche dans l'espace de travail.

4. Cliquez deux fois sur une colonne ou une table imbriquée.

La fenêtre *Propriétés des colonnes* s'ouvre.

5. Sélectionnez l'onglet *Attributs* pour afficher ou modifier les attributs du DTD.

21.2.8.5 Vérification des erreurs

Contrôlez que SAP Data Services vérifie la validité de chaque fichier ou message XML entrant.

Si vous décidez de vérifier chaque fichier ou message XML, le logiciel utilise la DTD importée et stockée dans le référentiel et non celui spécifié par un fichier ou message XML donné. Si un fichier ou message n'est pas valide au regard de la DTD, le job en temps réel produit une erreur et s'arrête.

Lors de la phase de développement, vous pouvez valider tous les fichiers et messages à tester au regard des conditions d'erreur. Lors de la phase de production, si vous acceptez des messages ou fichiers non valides rares, vous risquez d'obtenir des données non valides ou ambiguës.

Tous les fichiers ou messages produits par le logiciel pour une cible de message ou de fichier XML sont validés par rapport à la DTD importée.

Vous pouvez activer ou désactiver la validation d'une source ou cible de message ou de fichier XML dans cet éditeur d'objet.

21.2.9 Génération de schémas DTD et XML à partir d'un schéma NRDM

Vous pouvez cliquer avec le bouton droit sur un schéma à partir d'un éditeur de requêtes dans Designer et générer un schéma DTD ou XML correspondant à la structure du schéma sélectionné (NRDM ou relationnel).

Cette fonctionnalité est utile pour stocker des données dans un fichier XML puis les lire dans un autre flux de données.

1. Générez un schéma DTD ou XML.
2. Utilisez le schéma DTD ou XML pour configurer un format XML.
3. Utilisez le format XML pour configurer une source XML pour le fichier stocké.

Le schéma DTD ou XML généré reposera sur les informations suivantes :

- Les colonnes deviennent des éléments ou des attributs, selon la configuration de l'attribut de type XML sur ATTRIBUT ou ELEMENT.
- Si l'attribut Requis est défini sur NON, l'élément ou l'attribut correspondant est marqué comme facultatif.
- Les tables imbriquées deviennent des éléments intermédiaires.
- L'attribut Type natif est utilisé pour définir le type de l'élément ou de l'attribut.
- Lors de la génération de schémas XML, les valeurs MinOccurs et MaxOccurs sont définies selon les attributs Occurrences minimum et Occurrences maximum de la table imbriquée correspondante.

Aucune autre information n'est prise en compte lors de la génération du schéma DTD ou XML.

21.3 Opérations sur les données imbriquées

Cette section décrit les opérations que vous pouvez effectuer sur les données imbriquées.

21.3.1 Vue d'ensemble des données imbriquées et de la transformation Query

Lors du travail avec des données imbriquées, la transformation Query fournit une interface pour exécuter des instructions SELECT à chaque niveau de la relation définie dans le schéma de sortie.

Avec les données relationnelles, une transformation Query permet d'exécuter une instruction SELECT. Le mappage entre les schémas d'entrée et de sortie définit la liste de projets pour l'instruction.

Vous utilisez la transformation Query pour manipuler les données imbriquées. Pour extraire une partie des données imbriquées uniquement, vous pouvez utiliser la transformation XML_Pipeline.

Sans données imbriquées, la transformation Query suppose que la clause FROM dans l'instruction SELECT contient les jeux de données connectés en tant qu'entrées dans l'objet de requête. Lorsque vous travaillez avec des données imbriquées, il est nécessaire de définir explicitement la clause FROM dans une requête. Le progiciel assiste en définissant les entrées de niveau supérieur en tant que valeurs de la clause FROM par défaut pour le schéma de sortie de niveau supérieur.

Les autres éléments de l'instruction SELECT définis par la requête fonctionnent de la même manière avec les données imbriquées et les données plates. Toutefois, une instruction SELECT ne pouvant inclure que des références aux jeux de données relationnelles, une requête qui inclut des données imbriquées contient également une instruction SELECT pour définir les opérations de chaque schéma parent et enfant dans la sortie.

L'Éditeur Query contient un onglet pour chaque clause de la requête :

- SELECT permet de spécifier des lignes d'éléments uniques de la sortie (sans prendre en compte les lignes identiques doublon).
- FROM répertorie tous les schémas d'entrée et permet de spécifier des paires et des conditions de jointure.

Les paramètres saisis pour les onglets suivants s'appliquent uniquement au schéma en cours (affiché dans la zone de texte Sortie de schéma située en haut à droite de l'Éditeur Query) :

- WHERE
- GROUP BY
- ORDER BY

21.3.2 Construction de la clause FROM

La clause FROM permet de spécifier les tables et les vues à utiliser dans une instruction JOIN.

La clause FROM se trouve au bas de l'onglet FROM. Elle se remplit automatiquement avec les informations incluses dans les sections Schéma(s) d'entrée en haut et Paires de jointures au milieu de l'onglet. Vous pouvez modifier la clause FROM en modifiant les schémas sélectionnés dans la zone Schéma(s) d'entrée et la section Paires de jointures.

Les schémas sélectionnés dans la section Schéma(s) d'entrée (et reflétés dans la clause FROM), notamment les colonnes contenant des schémas imbriqués, peuvent être inclus dans la sortie.

Lorsque vous incluez plus d'un schéma dans la section Schéma(s) d'entrée (en sélectionnant la case à cocher [De](#)), il est possible de spécifier des paires et des conditions de jointure et de saisir un cache et un classement de jointure pour chaque schéma d'entrée.

Les descriptions de la clause FROM et le comportement de la requête sont similaires pour les données imbriquées et les données relationnelles. Le schéma en cours permet de distinguer plusieurs instructions SELECT au sein d'une seule requête. Toutefois, les instructions SELECT étant dépendantes les unes des autres et l'interface utilisateur facilitant la construction de jeux de données arbitraires, la détermination des clauses FROM appropriées pour plusieurs niveaux d'imbrication peut s'avérer complexe.

Une clause FROM peut inclure :

- Tout schéma de niveau supérieur de l'entrée
- Tout schéma représentant une colonne d'un schéma dans la clause FROM du schéma parent
- Toute condition de jointure des paires de jointure

La clause FROM forme un chemin pouvant débuter à tout niveau de la sortie. Le premier schéma du chemin doit toujours être un schéma de niveau supérieur de l'entrée.

Les données produites par une instruction SELECT d'un schéma de niveau inférieur diffèrent en fonction de la présence d'un schéma dans la clause FROM du niveau supérieur.

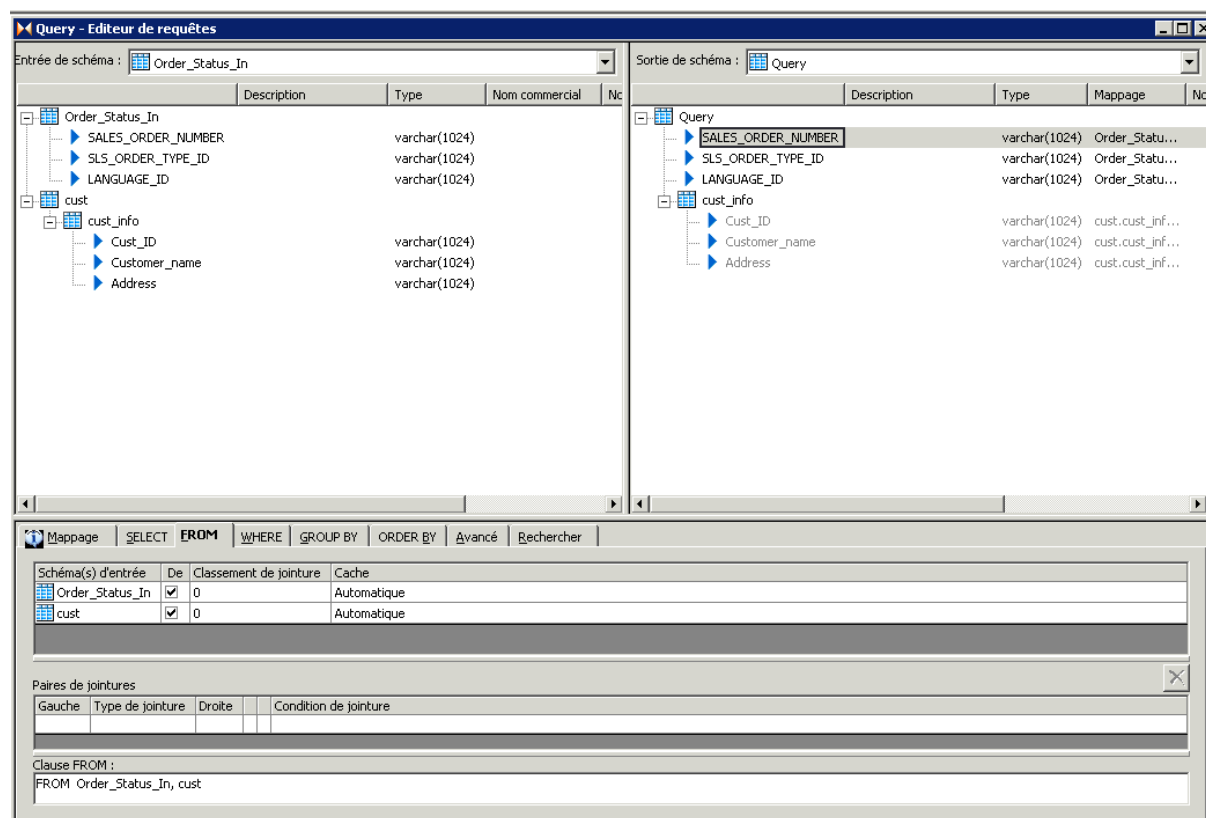
Les deux exemples suivants utilisent le jeu de données d'une commande client pour illustrer des scénarios dans lesquels les valeurs de la clause FROM modifient les données obtenues de la requête.

Informations associées

[Guide de référence : Transformation, Requête, Entrée de schéma et Sortie de schéma \[page 667\]](#)

21.3.2.1 Exemple : la clause FROM inclut toutes les entrées de niveau supérieur

Pour inclure les informations détaillées d'un client pour toutes les commandes dans la sortie, reliez le schéma Order_Status_In du niveau supérieur avec le schéma Cust. Incluez les deux schémas d'entrée du niveau supérieur dans la clause FROM pour générer les données appropriées. Lorsque vous sélectionnez les deux schémas d'entrée dans la zone Schéma(s) d'entrée de l'onglet FROM, ils apparaissent automatiquement dans la clause FROM.



Respectez les points suivants :

- La table Schéma(s) d'entrée dans l'onglet FROM inclut les deux schémas de niveau supérieur Order_Status_In et Cust (ce qui se reflète également dans la clause FROM).

- Le volet Sortie de schéma affiche le schéma imbriqué cust_info et les colonnes Cust_ID, Customer_name et Address.

21.3.2.2 Exemple : la clause FROM de niveau inférieur contient une entrée de niveau supérieur

Supposons que vous souhaitiez faire apparaître les informations détaillées d'un schéma dans chaque ligne d'un niveau inférieur d'un autre schéma. Par exemple, l'entrée inclut un schéma de niveau supérieur intitulé Materials et un schéma imbriqué intitulé LineItems. Vous souhaitez que l'entrée comprenne les informations détaillées des articles pour chaque poste. Le graphique ci-dessous illustre cette configuration dans Designer.

The image displays two screenshots of the SAP Data Designer interface, illustrating nested schema configurations in the FROM clause.

Left Screenshot: The 'Schema In' is 'LineItems'. The 'Schema Out' is 'LineItems'. The 'FROM' tab is active, showing a table with columns: Input schema(s), From, Join rank, and Cache. The table contains three rows: Materials (From: ☐, Join rank: 0, Cache: Automatic), Order (From: ☐, Join rank: 0, Cache: Automatic), and Order.LineItems (From: ☒, Join rank: 0, Cache: Automatic). The 'FROM clause' is: FROM "Order".LineItems.

Right Screenshot: The 'Schema In' is 'Materials'. The 'Schema Out' is 'LineItems'. The 'FROM' tab is active, showing a table with columns: Input schema(s), From, Join rank, and Cache. The table contains three rows: Materials (From: ☒, Join rank: 0, Cache: Automatic), Order (From: ☐, Join rank: 0, Cache: Automatic), and Order.LineItems (From: ☒, Join rank: 0, Cache: Automatic). The 'Join pairs' table shows a join between Order.LineItems (Left) and Materials (Right) with a join type of 'Left outer join' and a join condition of '"Order".LineItems.Item = Materials.Item'. The 'FROM clause' is: FROM "Order".LineItems __SAP_LEFT_OUTER_JOIN Materials ON ("Order".LineItems.Item = Materials.Item).

L'exemple sur la gauche affiche la configuration suivante :

- La zone Schéma d'entrée dans l'onglet FROM affiche le schéma imbriqué LineItems sélectionné.
- L'onglet FROM affiche la clause FROM « FROM "Order".LineItems ».

L'exemple sur la droite affiche la configuration suivante :

- Le schéma Materials.Description est mappé au schéma de sortie LineItems.Item.

- Les Schéma(s) d'entrée Materials et Order.LineItems sont sélectionnés dans la zone Schéma d'entrée dans l'onglet FROM (la colonne From contient des cases cochées).
- Une paire de jointure est créée en reliant le schéma imbriqué Order.LineItems au schéma de niveau supérieur Materials à l'aide d'un type de jointure externe gauche.
- Une condition de jointure est ajoutée lorsque le champ Item sous le schéma imbriqué LineItems est égal au champ Item du schéma de niveau supérieur Materials.

Clause FROM obtenue :

```
"Order".LineItems.Item = Materials.Item
```

21.3.3 Imbrication de colonnes

Lorsque vous imbriquez des lignes d'un schéma dans un autre schéma, le jeu de données produit dans le schéma imbriqué représente le résultat d'une requête sur le premier schéma utilisant les valeurs associées du second.

Par exemple, si vous disposez d'informations concernant une commande client dans un schéma d'en-tête et un schéma de postes, vous pouvez imbriquer les postes sous le schéma d'en-tête. Les postes d'une seule ligne du schéma d'en-tête sont égaux aux résultats d'une requête incluant le numéro de commande :

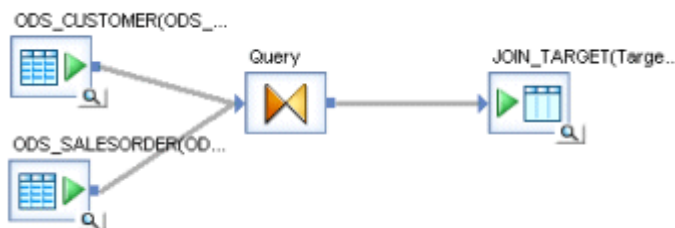
```
SELECT * FROM LineItems
WHERE Header.OrderNo = LineItems.OrderNo
```

Vous pouvez utiliser une transformation Query pour construire un jeu de données imbriquées à partir des données relationnelles. Lorsque vous indiquez les colonnes incluses dans le schéma imbriqué, spécifiez la requête utilisée pour définir le jeu de données imbriquées pour chaque ligne du schéma parent.

21.3.3.1 Construire un jeu de données imbriqué

Suivez les étapes suivantes pour configurer un jeu de données imbriquées.

1. Créez un flux de données avec les sources d'entrée que vous souhaitez inclure dans le jeu de données imbriquées.
2. Placez une transformation Query et une table cible dans le flux de données. Connectez les sources à l'entrée de la requête.



- Ouvrez la transformation Query et configurez la *liste Select*, la *clause FROM* et la *clause WHERE* afin de décrire l'instruction SELECT exécutée par la requête pour déterminer le jeu de données de niveau supérieur.

Option	Notes
<i>Liste Select</i>	mappez les éléments du schéma d'entrée au schéma de sortie en faisant glisser les colonnes du schéma d'entrée vers le schéma de sortie. Il est également possible d'inclure de nouvelles colonnes ou des expressions de mappage pour les colonnes.
<i>Clause FROM</i>	incluez les sources d'entrée dans la liste de l'onglet <i>FROM</i> , puis incluez toute jointure et condition de jointure requises pour définir les données.
<i>Clause WHERE</i>	incluez tout filtrage requis pour définir le jeu de données pour la sortie de niveau supérieur.

- Créez un nouveau schéma dans la sortie.

Cliquez avec le bouton droit de la souris dans la zone Sortie de schéma de l'Editeur Query, sélectionnez *Nouveau schéma de sortie*, puis nommez le nouveau schéma. Une nouvelle icône de schéma s'affiche dans la sortie, imbriquée sous le schéma de niveau supérieur.

Vous pouvez également faire glisser un schéma complet de l'entrée vers la sortie.

- Remplacez le schéma de sortie actuel par le schéma imbriqué en cliquant avec le bouton droit de la souris sur le schéma imbriqué puis en sélectionnant *En faire l'élément actuel*.

L'Editeur Query se modifie pour afficher le nouveau schéma actuel.

- Indiquez la clause FROM, la *liste SELECT* et la clause *WHERE* afin de décrire l'*instruction SELECT* exécutée par la requête pour déterminer le jeu de données de niveau supérieur.

Option	Notes
<i>Clause FROM</i>	si vous avez créé un nouveau schéma de sortie, il convient de faire glisser les schémas de l'entrée pour remplir la clause FROM. Si vous avez fait glisser un schéma existant de l'entrée vers la sortie de niveau supérieur, ce schéma est automatiquement mappé et répertorié dans l'onglet FROM.
<i>Liste Select</i>	seules les colonnes respectant les exigences de la clause FROM sont disponibles.
<i>Clause WHERE</i>	seules les colonnes respectant les exigences de la clause FROM sont disponibles.

- Si la sortie l'exige, imbriquez un autre schéma à ce niveau.

Répétez les étapes 4 à 6 dans ce schéma actuel pour tous les schémas imbriqués que vous souhaitez configurer.

- Si la sortie l'exige, imbriquez un autre schéma sous le niveau supérieur.

Faites du schéma de niveau supérieur le schéma actuel.

Informations associées

[Construction de la clause FROM \[page 281\]](#)

[Guide de référence : Transformation, Requête, Entrée de schéma et Sortie de schéma \[page 667\]](#)

21.3.4 Utilisation de colonnes corrélées dans des données imbriquées

La corrélation permet d'utiliser des colonnes d'un schéma de niveau supérieur pour construire un schéma imbriqué.

Dans un modèle relationnel imbriqué, les colonnes d'un schéma imbriqué sont implicitement liées aux colonnes de la ligne parente. Pour tirer parti de cette relation, il est possible d'utiliser des colonnes du schéma parent pour construire le schéma imbriqué. La colonne de niveau supérieur est une colonne corrélée.

L'inclusion d'une colonne corrélée dans un schéma imbriqué peut servir à deux fins :

- La colonne corrélée est une clé dans le schéma parent. Inclure la clé dans le schéma imbriqué permet de conserver une relation entre les deux schémas après leur conversion d'un modèle de données imbriqué à un modèle relationnel.
- La colonne corrélée est un attribut dans le schéma parent. Inclure l'attribut dans le schéma imbriqué permet d'utiliser cet attribut pour simplifier les requêtes corrélées sur les données imbriquées.

Pour inclure une colonne corrélée dans un schéma imbriqué, il n'est pas nécessaire d'inclure le schéma contenant cette colonne dans la clause FROM du schéma imbriqué.

1. Créez un flux de données avec une source incluant un schéma parent avec un schéma imbriqué.

Par exemple, la source peut être un schéma d'en-tête de commande incluant une colonne `LineItems` qui contient un schéma imbriqué.

2. Connectez une requête à la sortie de la source.
3. Dans l'éditeur de requêtes, copiez toutes les colonnes du schéma parent à la sortie.

Outre les colonnes de niveau supérieur, le progiciel crée une colonne intitulée `LineItems` qui contient un schéma imbriqué correspondant au schéma imbriqué `LineItems` dans l'entrée.

4. Remplacez le schéma actuel par le schéma `LineItems`.
5. Incluez une colonne corrélée dans le schéma imbriqué.

Les colonnes corrélées peuvent inclure des colonnes du schéma parent et tout autre schéma inclus dans la clause FROM du schéma parent.

Faites par exemple glisser la colonne `OrderNo` du schéma d'en-tête vers le schéma `LineItems`. L'inclusion de la colonne corrélée crée une nouvelle colonne de sortie dans le schéma `LineItems` intitulée `OrderNo` et la mappe à la colonne `Order.OrderNo`. Le jeu de données créé pour `LineItems` inclut toutes les colonnes `LineItems` et `OrderNo`.

Si la colonne corrélée provient d'un schéma autre que le parent immédiat, les données du schéma imbriqué contiennent uniquement les lignes correspondant aux valeurs associées dans la ligne actuelle du schéma parent et les valeurs de la colonne corrélée.

21.3.5 Lignes d'éléments uniques et données imbriquées

L'option *Lignes d'éléments uniques* dans les transformations Query supprime toute ligne doublon au niveau supérieur d'une jointure.

Cette option est particulièrement utile pour éviter des produits croisés dans des jointures générant une sortie imbriquée.

21.3.6 Regroupement de valeurs de schémas imbriqués

Lorsque vous spécifiez une clause GROUP BY pour un schéma contenant un schéma imbriqué, l'opération de regroupement combine les schémas imbriqués de chaque groupe.

Par exemple, pour assembler tous les postes inclus dans toutes les commandes pour chaque statut d'un jeu de commande, vous pouvez définir la clause GROUP BY au niveau supérieur du jeu de données sur la colonne de statut (Order.State) et créer un schéma de sortie qui inclut une colonne State (définie sur Order.State) et le schéma imbriqué LinelItems.

Jeu de données de commande

OrderNo	CustID	Etat	Postes
9999	1000	CA	
9999	1001	CA	
9777	1202	TX	

Item	ItemQty	ItemPrice
001	2	10
002	4	5

Item	ItemQty	ItemPrice
001	7	23
002	7	10

Item	ItemQty	ItemPrice
001	9	99
002	1	2

Jeu de données de commande avec état GroupBy

Etat	Postes
CA	
TX	

Item	ItemQty	ItemPrice
001	2	10
002	4	5

Item	ItemQty	ItemPrice
001	7	23
002	7	10

Item	ItemQty	ItemPrice
001	9	99
002	1	2

Il en résulte un jeu de lignes (une pour chaque statut) comprenant la colonne State et le schéma imbriqué LinelItems contenant tous les schémas LinelItems pour toutes les commandes de ce statut.

21.3.7 Suppression de l'imbrication de données imbriquées

Pour charger un jeu de données contenant des schémas imbriqués dans une cible relationnelle (non imbriquée), il est nécessaire de supprimer l'imbrication des lignes imbriquées.

Une commande client peut par exemple utiliser un schéma imbriqué pour définir la relation entre l'en-tête et les postes de la commande. Pour charger les données dans des schémas relationnels, il convient de supprimer

l'imbrication des multiniveaux. La suppression de l'imbrication génère un produit croisé du schéma de niveau supérieur (parent) et du schéma imbriqué (enfant).

Jeu de données imbriquées

OrderNo	CustID	Postes
9999	1001	

Item	ItemQty
001	2
002	4

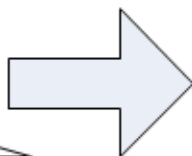


Schéma d'en-tête

OrderNo	CustID	Item	ItemQty
9999	1001	001	2
9999	1001	002	4

Il est également possible de charger différentes colonnes de différents niveaux imbriqués dans des schémas distincts. Une commande client peut par exemple être aplanie afin que le numéro de commande soit conservé séparément avec chaque poste et les informations d'en-têtes et de postes chargées dans différents schémas.

Jeu de données imbriquées

OrderNo	CustID	Postes
9999	1001	

Item	ItemQty
001	2
002	4

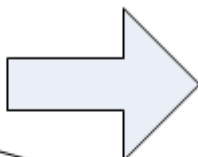


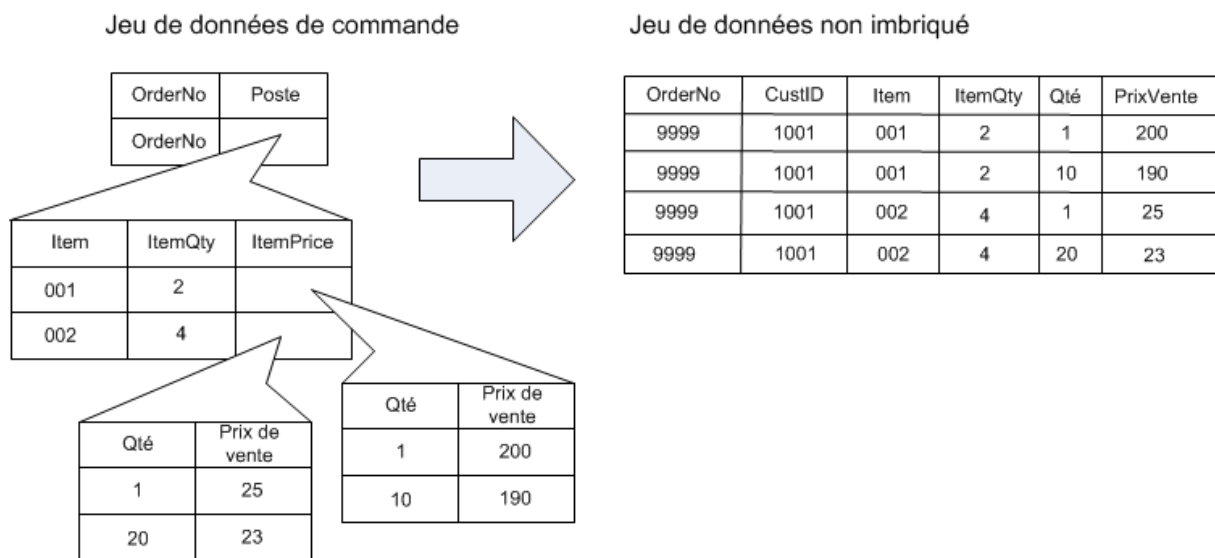
Schéma d'en-tête

OrderNo	CustID
9999	1001

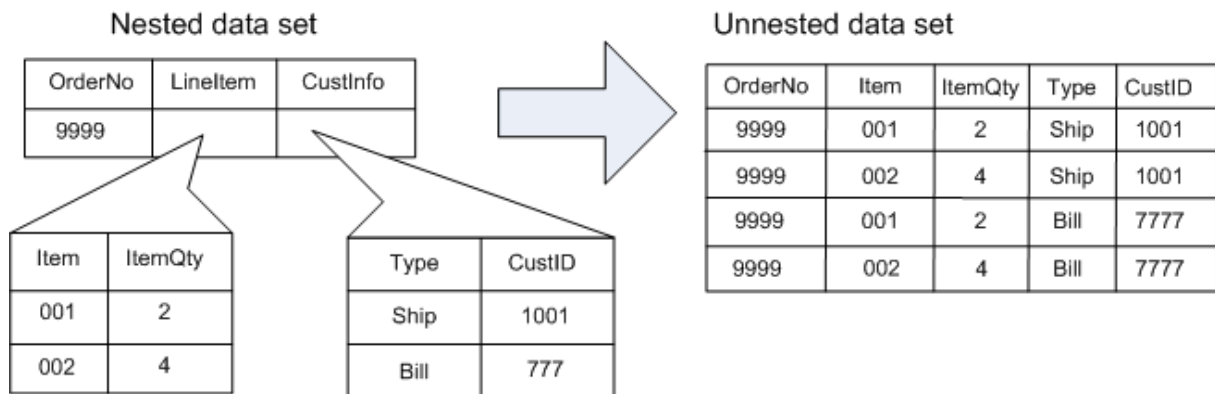
Schéma de poste

OrderNo	Item	ItemQty
9999	001	2
9999	002	4

Le progiciel permet de supprimer l'imbrication de tout nombre de schémas imbriqués à toute profondeur. Quel que soit le nombre de niveaux impliqués, le résultat des schémas desquels l'imbrication a été supprimée est un produit croisé des schémas parent et enfant. Lorsque plusieurs niveaux d'imbrications doivent être supprimés, l'imbrication du schéma enfant le plus profond est supprimée en premier, puis celle du résultat (le produit croisé du schéma parent et du schéma enfant le plus profond) et ainsi de suite jusqu'à l'imbrication du schéma de niveau supérieur.



La suppression de l'imbrication de tous les schémas (produit croisé de l'ensemble des données) peut générer un résultat inattendu. Par exemple, si une commande client inclut plusieurs valeurs client, telles que les adresses de livraison et de facturation, l'aplatissement d'une commande client par la suppression de l'imbrication des schémas client et de postes produit des lignes de données pouvant ne pas être utiles pour le traitement de la commande.



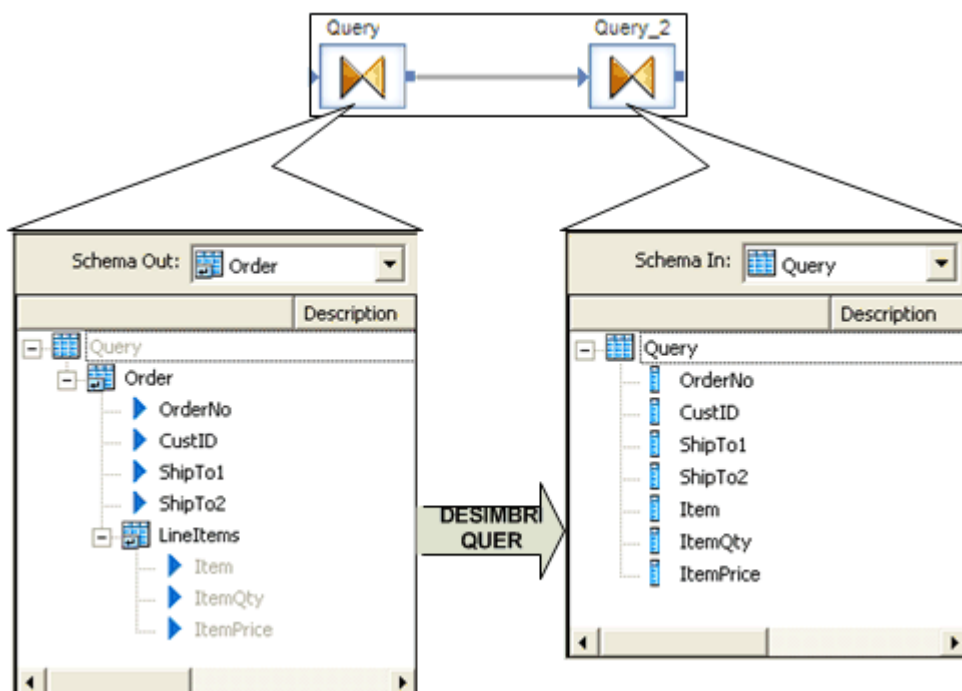
21.3.7.1 Suppression de l'imbrication de données imbriquées

1. Créez la sortie dont vous voulez supprimer l'imbrication dans le schéma de sortie d'une requête.

Les données de colonnes ou schémas non imbriqués peuvent s'avérer plus complexes à filtrer après l'opération de suppression de l'imbrication. Vous pouvez exécuter la commande Couper pour supprimer des colonnes ou schémas du niveau supérieur. Pour supprimer des schémas ou colonnes imbriqués dans des schémas imbriqués, faites du schéma imbriqué le schéma actuel puis coupez les colonnes non nécessaires ou les colonnes imbriquées.

2. Cliquez avec le bouton droit de la souris sur le nom du schéma dont vous souhaitez supprimer l'imbrication et sélectionnez *Supprimer l'imbrication*.

La sortie de la requête (l'entrée de la prochaine étape dans le flux de données) inclut les données dans la nouvelle relation, comme illustré dans le diagramme suivant.



21.3.8 Transformation de niveaux inférieurs de données imbriquées

Les données imbriquées incluses dans l'entrée de transformations (à l'exception des transformations Query ou XML_Pipeline) passent par la transformation sans être incluses dans l'opération de transformation. Seules les colonnes de niveau supérieur du jeu de données d'entrée sont disponibles pour des transformations ultérieures.

21.3.8.1 Transformer des valeurs de niveaux inférieurs de schémas imbriqués

1. Exécutez l'une des actions suivantes pour obtenir les données imbriquées :
 - Utilisez une transformation Query pour supprimer l'imbrication des données.
 - Utilisez une transformation XML_Pipeline pour sélectionner des parties des données imbriquées.
 - Exécutez la transformation.

2. Imbriguez de nouveau les données pour recréer les relations imbriquées.

Informations associées

[Suppression de l'imbrication de données imbriquées \[page 287\]](#)

21.4 Extraction et analyse XML pour les colonnes

Le progiciel permet d'extraire des données XML stockées dans une table source ou dans une colonne de fichier plat, de les transformer en données NRDM, puis de les charger dans une colonne de fichier cible ou plat.

En outre vous pouvez extraire des données de fichiers et de messages XML, leur représentation sous forme de données NRDM lors d'une transformation puis leur chargement dans un fichier ou un message XML.

De plus en plus de fournisseurs de bases de données permettent de stocker des données XML dans une seule colonne. Le champ est en général une valeur varchar, long ou clob. La fonctionnalité de gestion XML du progiciel prend également en charge la lecture et l'écriture de tels champs. Le progiciel fournit quatre fonctions prenant en charge l'extraction et le chargement de colonnes :

- `extract_from_xml`
- `load_to_xml`
- `long_to_varchar`
- `varchar_to_long`

La fonction `extract_from_xml` extrait le contenu XML stocké dans une seule colonne et crée la structure NRDM correspondante pour permettre sa transformation par le progiciel. Cette fonction prend uniquement en charge les données varchar.

Pour permettre l'extraction et l'analyse de colonnes, les données de colonnes long et clob doivent être converties en varchar avant de pouvoir être transformées par le progiciel.

- Le progiciel convertit une entrée de type de données clob en données varchar si vous sélectionnez l'option *Importer les types de données non pris en charge comme VARCHAR de taille* lors de la création d'une connexion de banque de données de base de données dans l'Editeur de banque de données.
- Si la source utilise un type de données long, appliquez la fonction `long_to_varchar` pour convertir les données en valeur varchar.

i Remarque

Le progiciel limite la taille des données XML prises en charge par ces méthodes à 100 Ko en raison de la limitation actuelle de ce type de données varchar. Il est prévu que cette restriction soit levée à l'avenir.

La fonction `load_to_xml` génère du contenu XML d'une structure NRDM donnée dans le progiciel, puis charge ce contenu dans une colonne varchar. Si vous souhaitez qu'un job convertisse la sortie en colonne long, appliquez la fonction `varchar_to_long` qui utilise la sortie de la fonction `load_to_xml` en tant qu'entrée.

21.4.1 Exemples de scénarios

Les scénarios suivants décrivent l'utilisation des fonctions permettant d'extraire des données XML d'une colonne source et de les charger dans une colonne cible.

Informations associées

[Extraction de données XML d'une colonne dans le progiciel \[page 292\]](#)

[Chargement de données XML vers une colonne du type de données long \[page 294\]](#)

[Extraction de chaînes XML de qualité des données à l'aide de la fonction `extract\_from\_xml` \[page 295\]](#)

21.4.1.1 Extraction de données XML d'une colonne dans le progiciel

Ce scénario utilise les fonctions `long_to_varchar` et `extract_from_xml` pour extraire des données XML d'une colonne avec des données de type long.

Exécutez les étapes préalables suivantes avant le début des étapes principales :

1. Importer une table Oracle contenant une colonne nommée *Contenu*, le type de données long et les données XML pour un bon de commande.
2. Importer dans le référentiel le schéma XML `PO.xsd` qui fournit le format pour les données XML.
3. Créer un projet, un job et un flux de données pour votre conception.
4. Ouvrir le flux de données et déposer la table source avec la colonne nommée *contenu* dans le flux de données.

Pour extraire des données XML d'une colonne dans le logiciel, suivez les étapes préalables ci-dessus puis suivez les étapes ci-dessous :

1. Créez une requête avec une colonne de sortie du type de données varchar et assurez-vous que sa taille est suffisante pour contenir les données XML.
2. Nommez cette colonne de sortie *contenu*.
3. Dans la section Mappage de l'éditeur de requêtes, ouvrez l'Assistant Fonction, sélectionnez le type de fonction *Conversion*, puis sélectionnez la fonction *long_to_varchar* et configurez-la en saisissant ses paramètres.

```
long_to_varchar(content, 4000)
```

Le second paramètre de cette fonction (4000 dans l'exemple ci-dessus) représente la taille maximale des données XML stockées dans la colonne de la table. Utilisez ce paramètre avec prudence. Si la taille n'est pas suffisante pour contenir le maximum de données XML pour la colonne, le progiciel tronquera les données, engendrant une erreur d'exécution. Inversement, ne saisissez pas un nombre trop important, ce qui exploiterait trop de mémoire de l'ordinateur lors de l'exécution.

4. Dans l'éditeur de requêtes, mappez la colonne de table source vers une nouvelle colonne de sortie.
5. Créez une seconde requête utilisant la fonction `extract_from_xml` pour extraire les données XML.

Pour appeler la fonction `extract_from_xml` :

- cliquez avec le bouton droit de la souris sur le contexte actuel dans la requête
- Sélectionnez *Nouvel appel de fonction*.
- Lorsque l'Assistant Fonction s'ouvre, sélectionnez *Conversion* et *extract_from_xml*.

i Remarque

Vous pouvez utiliser la fonction *extract_from_xml* dans un nouvel appel de fonction. Autrement, la fonction ne sera pas affichée dans l'Assistant Fonction.

- Entrez les valeurs pour les paramètres d'entrée et cliquez sur *Suivant*.

Paramètres d'entrée

Paramètre	Description
<i>Nom de colonne XML</i>	Saisissez <i>contenu</i> , qui correspond à la colonne de sortie dans la requête précédente contenant les données XML.
<i>Nom de colonne DTD ou XML</i>	Saisissez le nom du schéma du bon de commande (dans ce cas <i>PO</i>).
<i>Activez la validation</i>	Saisissez 1 si vous souhaitez que le progiciel valide les données XML avec le schéma spécifié. Dans le cas contraire, saisissez 0.

- Pour la fonction, sélectionnez une ou plusieurs colonnes que vous souhaitez utiliser à la sortie.

Imaginez que ce schéma de bon de commande contienne cinq éléments de niveau supérieur : DateCommande, AdresseLivraison, AdresseFacturation, Commentaire et Eléments. Vous pouvez sélectionner un nombre quelconque de colonnes de niveau supérieur d'un schéma XML qui inclut des données de colonnes de modèle de données scalaires ou relationnelles imbriquées (NRDM). Le type de renvoi de la colonne est défini dans le schéma. Si la fonction échoue en raison d'une erreur lors de la tentative de production de la sortie XML, le progiciel renvoie une valeur nulle pour les colonnes scalaires et des tables imbriquées vides pour les colonnes NRDM. La fonction `extract_from_xml` ajoute également deux colonnes :

- AL_ERROR_NUM — renvoie des codes d'erreurs : 0 pour une réussite et un entier autre que zéro pour les échecs
- AL_ERROR_MSG : renvoie un message d'erreur si la valeur de AL_ERROR_NUM est autre que 0.
Renvoie une valeur nulle si la valeur de AL_ERROR_NUM est 0

Sélectionnez une ou plusieurs de ces colonnes en tant que sortie appropriée pour la fonction `extract_from_xml`.

- Cliquez sur *Terminer*.

Le progiciel génère l'appel de fonction dans le contexte en cours et remplit le schéma de sortie de la requête avec les colonnes de sortie spécifiées.

Avec les données converties dans la structure NRDM, vous êtes prêt à lui appliquer des opérations de transformation appropriées.

Par exemple, si vous souhaitez charger la structure NRDM vers un fichier XML cible, créez un fichier XML cible puis connectez la seconde requête à ce dernier.

i Remarque

Si vous souhaitez modifier l'appel de fonction, cliquez avec le bouton droit de la souris sur l'appel de fonction dans la seconde requête et sélectionnez *Modifier l'appel de fonction*.

Dans cet exemple, pour extraire des données XML d'une colonne du type de données long, nous avons créé deux requêtes : la première pour convertir les données à l'aide de la fonction `long_to_varchar` et la seconde pour ajouter la fonction `extract_from_xml`.

Il est également possible d'utiliser une seule requête en saisissant l'expression de la fonction `long_to_varchar` directement dans le premier paramètre de la fonction `extract_from_xml`. Le premier paramètre de la fonction `extract_from_xml` peut utiliser une colonne du type de données varchar ou une expression qui renvoie les données de type varchar.

Si le type de données de la colonne source n'est pas long mais varchar, n'incluez pas la fonction `long_to_varchar` à votre flux de données.

21.4.1.2 Chargement de données XML vers une colonne du type de données long

Ce scénario utilise les fonctions `load_to_xml` et `varchar_to_long` pour convertir une structure NRDM en données scalaires de type varchar dans un format XML puis la charger vers une colonne du type de données long.

Dans cet exemple, vous souhaitez convertir en données XML une structure NRDM pour un bon de commande à l'aide de la fonction `load_to_xml`. Vous voulez ensuite charger ces données vers une colonne de table Oracle intitulée *contenu* du type de données long. La fonction `load_to_xml` renvoyant une valeur du type de données varchar, vous utilisez la fonction `varchar_to_long` pour convertir la valeur de type de données varchar en valeur de type de données long.

1. Créez une requête et connectez-la à une requête ou source précédente (présentant la structure NRDM d'un bon de commande). Dans cette requête, créez une colonne de sortie du type de données varchar intitulée *contenu*. Assurez-vous que la taille de la colonne est suffisante pour contenir les données XML.
2. Dans la zone *Mappage*, ouvrez l'Assistant Fonction, cliquez sur la catégorie *Fonctions de conversion* puis sélectionnez la fonction *load_to_xml*.
3. Cliquez sur *Suivant*.
4. Entrez les valeurs pour les paramètres d'entrée.

La fonction `load_to_xml` dispose de sept paramètres.

5. Cliquez sur *Terminer*.

Dans la zone de mappage de la fenêtre Query, notez l'expression de la fonction :

```
load_to_xml(PO, 'PO', 1, '<?xml version="1.0" encoding = "UTF-8" ?>', NULL, 1, 4000)
```

Dans cet exemple, la fonction convertit la structure NRDM du bon de commande PO en données XML et affecte la valeur au contenu de la colonne de sortie.

6. Créez une autre requête avec les colonnes de sortie correspondant aux colonnes de la table cible.

- a. Supposons que la colonne soit intitulée *contenu* et soit du type de données long.
- b. Ouvrez l'Assistant Fonction dans la section de mappage de la requête puis sélectionnez la catégorie *Fonctions de conversion*.
- c. Utilisez la fonction `varchar_to_long` pour mapper la colonne d'entrée *contenu* vers la colonne de sortie *contenu*.

La fonction `varchar_to_long` n'utilise qu'un seul paramètre d'entrée.

- d. Saisissez une valeur pour le paramètre d'entrée.

```
varchar_to_long(contenu)
```

7. Connectez cette requête à une cible de base de données.

De même que dans l'exemple utilisant la fonction `extract_from_xml`, vous avez ici utilisé deux requêtes. La première requête a converti une structure NRDM en données XML puis a affecté la valeur à une colonne du type de données `varchar`. La seconde requête a converti le type de données `varchar` en long.

Il est possible de ne mettre qu'une seule requête en œuvre en utilisant les deux fonctions dans une seule expression :

```
varchar_to_long( load_to_xml(PO, 'PO', 1, '<?xml version="1.0" encoding =  
"UTF-8" ?>', NULL, 1, 4000) )
```

Si le type de données de la colonne de la table de la base de données cible stockant les données XML est `varchar`, la fonction `varchar_to_long` n'est pas utile dans la transformation.

Informations associées

[Guide de référence : fonctions et procédures \[page 676\]](#)

21.4.1.3 Extraction de chaînes XML de qualité des données à l'aide de la fonction `extract_from_xml`

Ce scénario utilise la fonction `extract_from_xml` pour extraire des données XML des transformations Geocoder, Global Suggestion Lists, Global Address Cleanse et USA Regulatory Address Cleanse

La transformation Geocoder, la transformation Global Suggestion Lists et la fonctionnalité Liste de propositions dans les transformations Global Address Cleanse et USA Regulatory Address Cleanse peuvent éditer un champ contenant une chaîne XML. Les transformations éditent les champs suivants pouvant contenir une chaîne XML.

Transformation	Champ de sortie XML	Description de champ de sortie
Geocoder	Liste_Résultats	Contient une chaîne de sortie XML lorsque plusieurs enregistrements sont renvoyés pour une recherche. Le

Transformation	Champ de sortie XML	Description de champ de sortie
		contenu dépend des données disponibles.
Global Address Cleanse Global Suggestion List USA Regulatory Address Cleanse	Liste_Propositions	Contient une chaîne de sortie XML incluant toutes les valeurs de champ de composant de la liste de propositions indiquées dans les options de transformation. Pour éditer ces champs comme champs XML, vous devez sélectionner XML comme style de sortie dans les options de transformation.

Pour utiliser les données figurant dans les chaînes XML (par exemple, dans une application Web qui utilise le job publié comme service Web), vous devez extraire les données. Il existe deux méthodes d'extraction de données :

Méthodes d'extraction de données

Méthode	Description
Insérer une transformation Query à l'aide de la fonction <code>extract_from_xml</code> .	Cette méthode permet d'insérer une transformation Query dans le flux de données après la transformation Geocoder, Global Suggestion Lists, Global Address Cleanse ou USA Regulatory Address Cleanse Utilisez ensuite la fonction <code>extract_from_xml</code> pour analyser les données de sortie imbriquées. Cette méthode est considérée comme meilleure pratique car elle propose des données de sortie analysées, facilement accessibles dans un intégrateur.
Développez un flux de données simple qui ne supprime pas l'imbrication des données	Cette méthode permet d'éditer facilement le champ de sortie contenant la chaîne XML sans suppression de l'imbrication de données. Cette méthode permet au développeur d'application ou à l'intégrateur de sélectionner de manière dynamique les composants de sortie dans le schéma de sortie final avant de les exposer comme service Web. Le développeur d'application doit travailler étroitement avec le concepteur de flux de données pour comprendre le flux de données au-delà d'un service Web en temps réel. Le développeur d'application doit comprendre les options de transformation et indiquer les éléments à renvoyer de la liste de propositions d'adresse retour puis supprimer l'imbrication de la chaîne de sortie XML pour générer des éléments discrets d'adresse.

Informations associées

[Présentation de la qualité des données \[page 426\]](#)

21.4.1.3.1 Extraction de chaînes XML de qualité des données à l'aide de la fonction `extract_from_xml`

Ce scénario utilise la fonction `extract_from_xml` pour extraire les chaînes XML.

1. Créez un fichier XSD pour la sortie.
2. Dans l'onglet [Format](#) de la bibliothèque d'objets locale, créez un schéma XML pour la XSD de sortie.
3. Dans l'onglet [Format](#) de la bibliothèque d'objets locale, créez un schéma XML pour `gac_suggestion_list.xsd`, `global_suggestion_list.xsd`, `urac_suggestion_list.xsd` ou `result_list.xsd`.
4. Dans le flux de données, incluez le champ suivant dans la [Sortie de schéma](#) de la transformation :

Transformation	Champ
○ Global Address Cleanse ○ Global Suggestion List ○ USA Regulatory Address Cleanse	Liste_Propositions
Geocoder	Liste_Résultats

5. Ajoutez une transformation Query après la transformation Global Address Cleanse, Global Suggestion Lists, USA Regulatory Address Cleanse ou Geocoder. Terminez-la comme suit.
 - a. Passez par tous les champs sauf le champ Liste_Propositions ou le champ Liste_Résultats à partir de l'entrée de schéma jusqu'à la sortie de schéma. Pour ce faire, faites glisser les champs directement du schéma d'entrée vers le schéma de sortie.
 - b. Dans [Sortie de schéma](#), faites un clic droit sur le nœud Requête et sélectionnez [Nouveau schéma de sortie](#). Saisissez Liste_Propositions ou Liste_Résultats comme le nom du schéma (ou peu importe le nom du champ qui est dans la sortie XSD).
 - c. Dans [Sortie de schéma](#), faites un clic droit sur le champ Liste_Propositions ou Liste_Résultats et sélectionnez [En faire l'élément actuel](#)
 - d. Dans [Sortie de schéma](#), faites un clic droit sur le champ Liste_Propositions ou Liste_Résultats et sélectionnez [Nouvel appel de fonction](#).
 - e. Sélectionnez [Fonctions de conversion](#) dans la colonne [Catégories de fonction](#) et `extract_from_xml` dans la colonne [Nom de la fonction](#) et cliquez sur [Suivant](#).
 - f. Dans la fenêtre [Définir les paramètres d'entrée](#), saisissez les informations suivantes et cliquez sur [Suivant](#).

Options Définir les paramètres d'entrée

Option	Description
Nom du champ XML	Sélectionnez le champ Liste_Propositions ou Liste_Résultats de la transformation en amont.

Option	Description
DTD ou nom de schéma	Sélectionnez le schéma XML que vous avez créé pour gac_suggestion_list.xsd, urac_suggestion_list.xsd, ou result_list.xsd.
Activer la validation	Saisissez 1 pour permettre la validation.

- g. Sélectionnez LISTE ou ENREGISTREMENT dans la liste des paramètres à gauche et cliquez sur le bouton de la flèche à droite pour l'ajouter à la liste Paramètres de sortie sélectionnés.
- h. Cliquez sur [Terminer](#).

La [Sortie de schéma](#) inclut les champs de la liste de propositions/liste de résultats dans le champ Liste_Propositions ou Liste_Résultats.

6. Incluez le schéma XML pour le XML de sortie après la Requête. Ouvrez le schéma XML pour valider que les champs sont identiques dans l'[Entrée de schéma](#) et dans la [Sortie de schéma](#).
7. Si vous extrayez des données depuis une transformation Global Address Cleanse, Global Suggestion Lists ou USA Regulatory Address Cleanse et que vous avez choisi de sortir uniquement un sous-ensemble des champs de sortie de la liste de propositions disponibles dans l'onglet [Options](#), insérez une deuxième transformation Query pour spécifier les champs que vous souhaitez sortir. Cette opération permet de sélectionner les composants de sortie dans le schéma de sortie final avant qu'il soit exposé comme un service Web.

Informations associées

[Présentation de la qualité des données \[page 426\]](#)

21.5 Extraction JSON

Outre l'extraction des données de fichiers et de messages JSON, leur représentation sous forme de données NRDM lors d'une transformation puis leur chargement dans un fichier ou un message JSON, le logiciel permet également d'extraire des données JSON stockées dans une colonne de fichier plat ou table source.

Le logiciel fournit la fonction `extract_from_json` pour prendre en charge l'extraction à partir des colonnes.

La fonction `extract_from_json` stocke le contenu JSON dans une seule colonne et crée la structure NRDM correspondante pour permettre sa transformation par le logiciel. Cette fonction prend uniquement en charge les données varchar.

Pour permettre l'extraction et l'analyse de colonnes, les données de colonnes long et clob doivent être converties en varchar avant de pouvoir être transformées par le logiciel.

- Le logiciel convertit une entrée de type de données clob en données varchar si vous sélectionnez l'option [Importer les types de données non pris en charge en tant que VARCHAR de taille](#) lors de la création d'une connexion à la banque de données de base de données dans l'Editeur de banque de données.
- Si la source utilise un type de données long, appliquez la fonction `long_to_varchar` pour convertir les données en valeur varchar.

21.5.1 Extraction des données de la chaîne JSON à l'aide de la fonction `extract_from_json`

1. Créez un fichier de schéma ou un fichier de données JSON pour la sortie.
2. Dans l'onglet *Format* de la Bibliothèque d'objets locale, cliquez sur ► *Nouveau* ► *Schéma JSON* ► pour créer un schéma pour votre sortie à l'aide du fichier de schéma ou du fichier de données JSON ci-dessus.
3. Dans le flux de données, dans la sortie de schéma, cliquez avec le bouton droit sur une colonne et sélectionnez *Nouvel appel de fonction*.
4. Dans la catégorie *Fonctions de conversion*, sélectionnez *extract_from_json* et cliquez *Suivant*.
5. Dans la fenêtre Définir les paramètres d'entrée, saisissez les informations suivantes et cliquez sur *Suivant* :

Option	Description
Nom de champ JSON	Sélectionnez le champ Liste_Propositions ou Liste_Résultats de la transformation en amont.
Nom du schéma	Sélectionnez le schéma que vous avez créé
Activez la validation	Saisissez 1 pour permettre la validation.

6. Sélectionnez *LISTE* ou *ENREGISTREMENT* dans la liste des paramètres à gauche et cliquez sur le bouton de droite en forme de flèche pour l'ajouter à la liste des paramètres de sortie sélectionnés.
7. Cliquez sur *Terminer*.
La sortie de schéma inclut les champs de la liste de propositions/liste de résultats dans le champ Liste_Propositions ou Liste_Résultats.

22 Flux de données intégrés

Un flux de données intégré est un flux de données qui est appelé depuis un autre flux de données.

Le logiciel fournit des options faciles à utiliser pour créer des flux de données intégrés.

22.1 Vue d'ensemble des flux de données intégrés

Les données sont transmises dans ou en dehors du flux de données intégré à partir du flux parent par une seule source ou cible.

Un flux de données intégré peut contenir autant de sources ou cibles que nécessaire, mais uniquement une seule entrée ou une seule sortie peut transmettre les données vers ou depuis le flux de données parent.

Il est possible de créer les types de flux de données intégrés suivants :

Type	A utiliser lorsque vous souhaitez...
Une entrée	Ajouter un flux de données intégré à la fin d'un flux de données.
Une sortie	Ajouter un flux de données intégré au début d'un flux de données.
Aucune entrée ni sortie	Répliquer un flux de données existant.

Un flux de données intégré est une aide à la conception qui n'a aucun effet sur l'exécution des jobs. Lorsque le logiciel exécute le flux de données parent, il développe les flux de données intégrés, optimise le flux de données parent puis l'exécute.

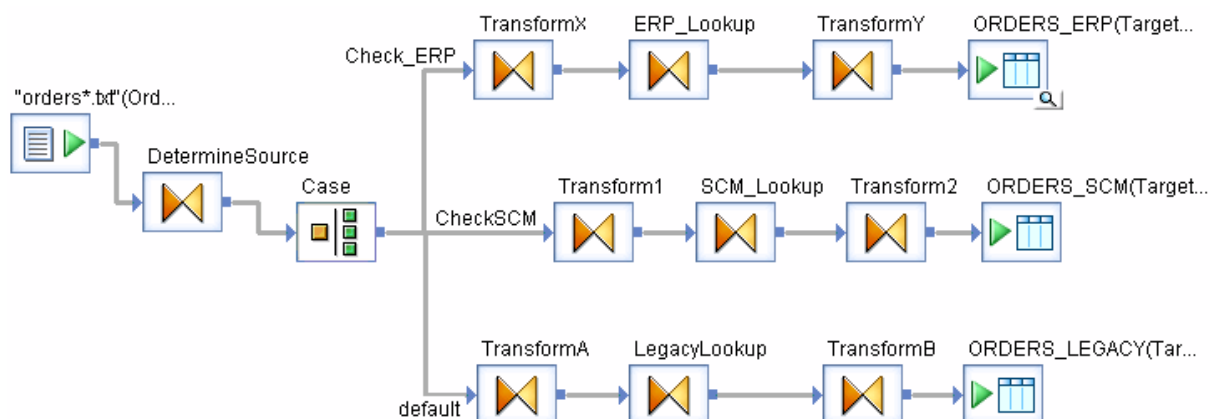
Utilisez les flux de données intégrés pour :

- Simplifier l'affichage du flux de données. Regroupez des sections d'un flux de données dans les flux de données intégrés pour permettre une documentation et une présentation plus claires.
- Réutiliser la logique du flux de données. Enregistrez les sections logiques d'un flux de données pour pouvoir utiliser la logique exacte dans d'autres flux de données, ou fournir une manière simple de répliquer la logique et de la modifier pour d'autres flux.
- Déboguer la logique du flux de données. Répliquez les sections d'un flux de données comme des flux de données intégrés pour pouvoir les exécuter indépendamment.

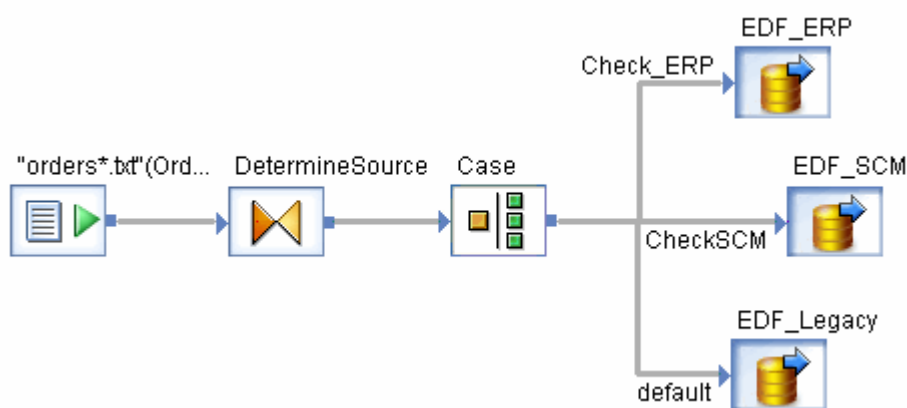
22.2 Exemples de flux de données imbriqués

Exemples de cas présentant où et comment utiliser les flux de données intégrés.

Dans cet exemple, un flux de données utilise une source unique pour charger trois systèmes cibles différents. La transformation Case envoie chaque ligne de la source aux différentes transformations qui les traitent pour obtenir une sortie cible unique.



Il est possible de simplifier le flux de données parent en utilisant des flux de données intégrés pour les trois cas différents.



22.3 Création de flux de données intégrés

Vous pouvez créer un flux de données intégré ou intégrer un flux de données existant.

Pour créer des flux de données intégrés, vous pouvez utiliser l'une des deux méthodes suivantes :

- Sélectionnez les objets dans un flux de donnée, faites un clic droit et sélectionnez *Créer un flux de données intégré*.
- Faites glisser un flux de données complet et validé de la bibliothèque d'objets vers un flux de données ouvert dans l'espace de travail. Puis :
 - Ouvrez le flux de données que vous venez à l'instant d'ajouter.
 - Faites un clic droit sur un objet que vous souhaitez utiliser comme un port d'entrée ou de sortie et sélectionnez *Créer le port* pour cet objet.

Le logiciel marque l'objet que vous sélectionnez comme le point de connexion pour ce flux de données intégré.

i Remarque

Il est possible d'indiquer un seul port, ce qui signifie que le flux de données intégré peut apparaître uniquement au début ou à la fin du flux de données parent.

Data Services ignore certains des fichiers physiques requis pour les sources et cibles auxquelles des ports sont affectés.

- Lors de l'utilisation d'un flux de données intégré, certaines données seront directement transmises de l'appelant vers la ou les transformations en regard du port source.
- Lors de l'utilisation directe d'un flux de données, Data Services utilisera les fichiers physiques dans les sources et cibles, mais ignorera les ports.

22.3.1 Utilisation de l'option Créer un flux de données intégré

Utilisez cette option pour créer un flux de données intégré.

1. Sélectionnez les objets dans un flux de données ouvert à l'aide d'une des méthodes suivantes :

- Cliquez sur l'espace vide et faites glisser le rectangle autour des objets.
- Appuyez sur Ctrl et cliquez sur chaque objet.

Assurez-vous que les objets que vous sélectionnez sont :

- tous connectés les uns aux autres ;
- connectés à d'autres objets en fonction du type de flux de données intégré que vous souhaitez créer, tel qu'une entrée, une sortie, ou aucune entrée ni sortie.

2. Faites un clic droit et sélectionnez *Créer un flux de données intégré*.

La fenêtre Créer un flux de données intégré s'ouvre, avec le flux de données intégré connecté au flux parent par un objet d'entrée.

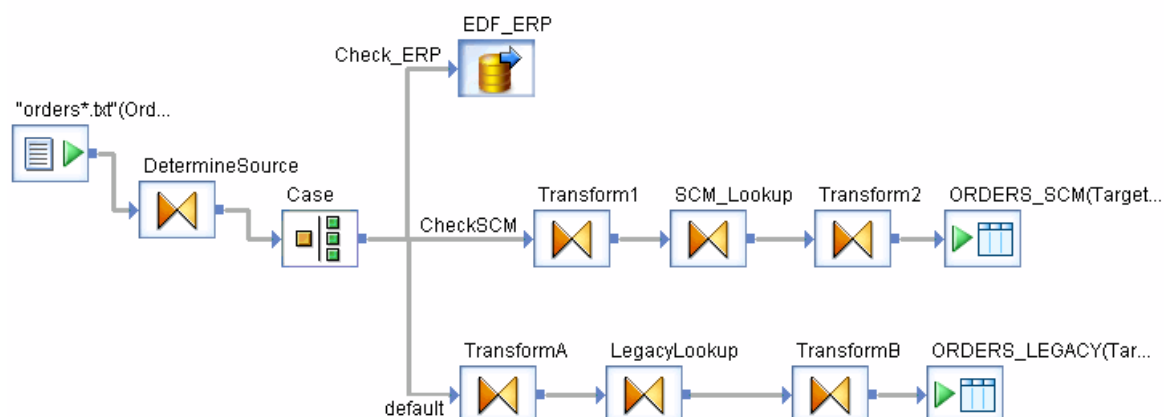
3. Nommez le flux de données intégré à l'aide de la convention `FDI_<NomFDI>` ; par exemple `FDI_ERP`.

Si vous décochez la case Remplacer les objets dans le flux de données d'origine, le logiciel n'apporte aucune modification dans le flux de données d'origine. Le logiciel enregistre le nouvel objet du flux de données intégré dans le référentiel et l'affiche dans la bibliothèque d'objets sous l'onglet Flux de données.

Vous pouvez utiliser un flux de données intégré créé sans remplacement comme un flux de données autonome pour le dépannage.

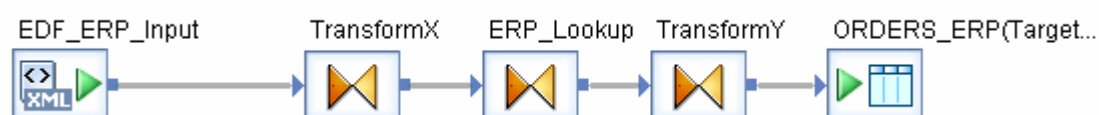
Si l'option Remplacer les objets dans le flux de données d'origine est sélectionnée, le flux de données d'origine devient un flux de données parent, qui a un appel vers le nouveau flux de données intégré.

4. Cliquez sur *OK*.



Le flux de données intégré apparaît dans le nouveau flux de données parent.

5. Cliquez sur le nom du flux de données intégré pour l'ouvrir.



6. Notez que le logiciel a créé un nouvel objet, `FDI_ERP_Entrée`, qui correspond au port d'entrée qui connecte ce flux de données intégré au flux de données parent.

Lorsque vous utilisez l'option *Créer un flux de données intégré*, le logiciel crée automatiquement un objet d'entrée ou de sortie en fonction de l'objet qui est connecté au flux de données intégré lors de sa création.

Par exemple, si un flux de données intégré a une connexion de sortie, le flux de données intégré inclut alors un objet de fichier XML cible étiqueté `<NomFDI>_Sortie`.

Les conventions d'affectation de nom pour chaque type de flux de données intégré sont :

Type	Conventions d'affectation de nom
Une entrée	<code><NomFDI>_Entrée</code>
Une sortie	<code><NomFDI>_Sortie</code>
Aucune entrée ni sortie	Le logiciel crée un flux de données intégré sans objet d'entrée ou de sortie.

22.3.2 Création de flux de données intégrés à partir de flux existants

Vous pouvez appeler un flux de données existant depuis un autre flux de données.

Vous devrez placer le flux de données dans le flux de données parent, puis marquer quelle source ou quelle cible utiliser pour transmettre les données entre les flux de données parent et intégré.

1. Faites glisser un flux de données valide existant de la bibliothèque d'objets dans un flux de données qui est ouvert dans l'espace de travail.
2. Pensez à renommer le flux à l'aide de la convention d'appellation `FDI_<NomFDI>`.

Le flux de données intégré apparaît sans aucune pointe de flèche (ports) dans l'espace de travail.

3. Ouvrez le flux de données intégré.
4. Faites un clic droit sur un objet cible ou source (fichier ou table) et sélectionnez *Créer le port*.

i Remarque

Assurez-vous de spécifier uniquement un port de sortie ou d'entrée.

De la même façon qu'un flux de données normal, les différents types de ports de flux de données intégrés sont indiqués par des marquages directionnels sur l'icône du flux de données intégré.

22.3.3 Utilisation des flux de données intégrés

Lorsque vous créez et configurez un flux de données intégré à l'aide de l'option *Créer un flux de données intégré*, le logiciel crée un nouveau fichier XML d'entrée ou de sortie et enregistre le schéma dans le référentiel comme un schéma XML.

Il est possible de réutiliser un flux de données intégré en le faisant glisser de l'onglet *Flux de données* de la bibliothèque d'objets vers d'autres flux de données. Pour enregistrer l'heure de mappage, il est possible que vous souhaitiez utiliser l'option Mettre à jour le schéma ou l'option Schéma de correspondance.

Le scénario d'exemple suivant utilise les deux options :

- Créez le flux de données 1.
- Sélectionnez les objets dans le flux de données 1 et créez le flux de données intégré 1 pour que le flux de données parent 1 appelle le flux de données intégré 1.
- Créez le flux de données 2 et le flux de données 3 et ajoutez le flux de données intégré 1 aux deux flux.
- Retournez au flux de données 1. Modifiez le schéma de l'objet précédant le flux de données intégré 1 et utilisez l'option Mettre à jour le schéma avec le flux de données intégré 1. Cette option met à jour le schéma du flux de données intégré 1 dans le référentiel.
- Désormais, les schémas dans les flux de données 2 et 3 qui sont alimentés dans le flux de données intégré 1 seront différents du schéma attendu par le flux de données intégré.
- Utilisez l'option Schéma de correspondance pour le flux de données intégré 1 dans les flux de données 2 et 3 pour résoudre les incohérences au moment de l'exécution. L'option Schéma de correspondance affecte uniquement les paramètres dans le flux de données actuel.

Les sections suivantes décrivent plus en détail l'utilisation des options Mettre à jour le schéma et Schéma de correspondance.

22.3.3.1 Mise à jour des schémas

Le logiciel fournit une option pour mettre à jour un schéma d'entrée d'un flux de données intégré.

Cette option met à jour le schéma d'un objet d'entrée d'un flux de données intégré avec le schéma de l'objet précédent dans le flux de données parent. Toutes les occurrences du flux de données intégré se mettent à jour lorsque vous utilisez cette option.

1. Ouvrez le flux de données parent du flux de données intégré.
2. Faites un clic droit sur l'objet du flux de données intégré et sélectionnez [Mettre à jour le schéma](#).

22.3.3.2 Correspondance des données entre les flux de données parent et intégré

Le schéma d'un objet d'entrée d'un flux de données intégré peut correspondre à un schéma de l'objet précédent dans le flux de données parent par le nom ou par le poste. Une correspondance par poste est celle par défaut.

22.3.3.2.1 Spécifier la manière dont les schémas doivent correspondre

1. Ouvrez le flux de données parent du flux de données intégré.
2. Faites un clic droit sur l'objet du flux de données intégré et sélectionnez [Schéma de correspondancePar nom](#) ou [Schéma de correspondancePar poste](#).

L'option Schéma de correspondance affecte uniquement les paramètres pour le flux de données actuel.

Data Services permet également au schéma de l'objet précédent dans le flux de données parent d'avoir plus ou moins de colonnes que le flux de données intégré. Le flux de données intégré ignore les colonnes supplémentaires et lit les colonnes manquantes comme des valeurs nulles.

Les colonnes dans les deux schémas doivent avoir des types de données identiques ou convertibles. Consultez la section sur la "Conversion de type" dans le *Guide de référence* pour plus d'informations.

22.3.3.3 Suppression d'objets de flux de données intégré

Il est possible de supprimer les ports du flux de données intégré ou de supprimer l'intégralité des flux de données intégrés.

22.3.3.3.1 Supprimer un port

Faites un clic droit sur l'objet d'entrée ou de sortie dans le flux de données intégré et décochez *Créer le port*. Data Services supprime la connexion à l'objet parent.

i Remarque

Il est impossible de supprimer un port en supprimant simplement la connexion dans le flux parent.

22.3.3.3.2 Suppression d'un flux de données intégré

Sélectionnez-le dans le flux de données parent ouvert et sélectionnez *Supprimer* dans le menu clic droit ou dans le menu Edition.

Si vous supprimez des flux de données intégrés de la bibliothèque d'objets, l'icône du flux de données intégré apparaît avec un cercle rouge barré dans le flux de données parent.

EDF_Example



Supprimez ces objets de flux de données intégré qui n'existent plus des flux de données parent.

22.3.4 Test séparé d'un flux de données intégré

Les flux de données intégrés peuvent être testés en les exécutant séparément comme des flux de données réguliers.

1. Indiquez un fichier XML pour le port d'entrée ou le port de sortie.

Lorsque vous utilisez l'option Créer un flux de données intégré, un objet du fichier XML d'entrée ou de sortie est créé puis connecté (facultatif) à l'objet précédent ou suivant dans le flux de données parents.

Pour tester le fichier XML sans le flux de données parent, cliquez sur le nom du fichier XML pour ouvrir son éditeur source ou cible pour indiquer un nom de fichier.

2. Placez le flux de données intégré dans un job.
3. Exécutez le job.

Il est également possible d'utiliser les fonctionnalités suivantes pour tester des flux de données intégrés :

- Afficher les données pour les exemples de données passées dans un flux de données intégré.
- Statistiques d'audit sur les données lues depuis les sources, transformées, et chargées dans les cibles, et règles sur les statistiques d'audit pour vérifier que les données attendues sont traitées.

Informations associées

[Conception et débogage \[page 722\]](#)

22.3.5 Dépannage des flux de données intégrés

Les situations suivantes produisent des erreurs :

- Un port d'entrée et un port de sortie sont indiqués dans un flux de données intégré.
- Flux de données qui n'existent plus piégés.
- Connexion au flux de données parent supprimée alors que l'option [Créer le port](#) dans le flux de données intégré reste sélectionnée.
- Les transformations avec des séparateurs (comme la transformation Case) indiquées comme l'objet du port de sortie car un séparateur produit plusieurs sorties, et les flux de données intégrés peuvent seulement en avoir une.
- Variables et paramètres déclarés dans le flux de données intégré qui ne sont pas déclarés dans le flux de données parent.
- Intégration du même flux de données à tous les niveaux de ce flux de données.
Il est cependant possible d'avoir des niveaux d'intégration illimités. Par exemple, le flux de données FDI1 appelle le flux de données intégré FDI1 qui appelle le FDI2.

Informations associées

[Suppression d'un flux de données intégré \[page 306\]](#)

[Supprimer un port \[page 306\]](#)

23 Variables et paramètres

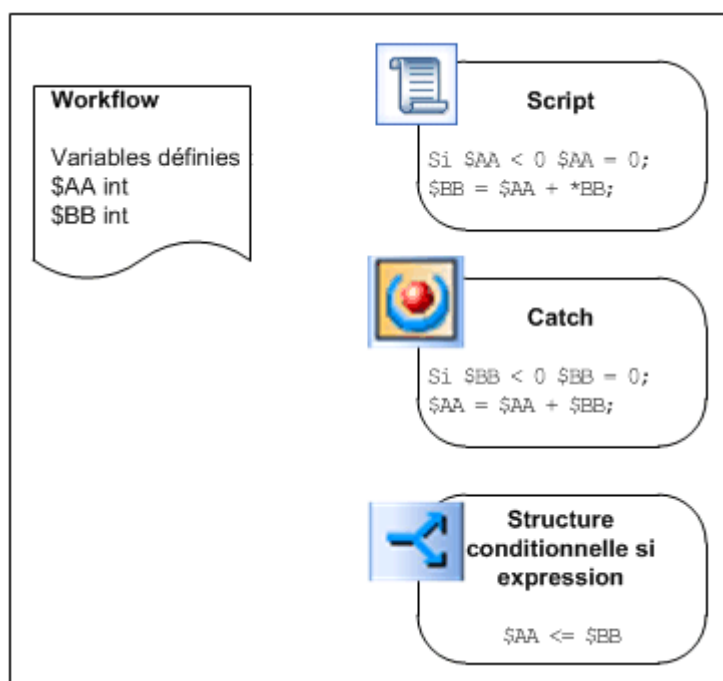
Vous pouvez améliorer la flexibilité et la réutilisabilité des workflows et des flux de données en utilisant des variables locales et globales lorsque vous créez des jobs.

Les variables sont des caractères génériques symboliques pour les valeurs. Les types de données d'une variable peuvent être de n'importe quel type pris en charge par le logiciel comme un entier, une décimale, une date ou une chaîne texte.

Vous pouvez utiliser des variables dans des expressions pour simplifier la prise de décision ou l'utilisation de données (utilisation de l'arithmétique ou substitution de caractère). Par exemple, une variable peut être utilisée dans une instruction `LOOP` ou `IF` pour vérifier une valeur de variable afin de définir l'étape à exécuter :

```
If $<amount_owed> > 0 print('$<invoice.doc>');
```

Si vous définissez des variables dans un job ou un workflow, le logiciel les utilise généralement dans un script, un catch ou un processus conditionnel.



Vous pouvez utiliser des variables dans des flux de données. Par exemple, utilisez-les dans une fonction personnalisée ou dans la clause `WHERE` d'une transformation Query.

Il existe deux types de variables, chacune ayant ses propres restrictions.

Type de variable	Restrictions
Locale	Limitées à l'objet dans lequel elles sont créées (job ou workflow). Vous devez utiliser des paramètres pour transférer des variables locales aux objets enfant (workflows et flux de données).
Globale	Limitées au job dans lequel elles sont créées ; cependant, les paramètres ne doivent pas être nécessairement transférés aux workflows et aux flux de données.

i Remarque

Si vous avez des workflows exécutés en parallèle, les variables globales ne sont pas affectées.

i Remarque

Les paramètres sont des expressions transférées dans un workflow ou un flux de données lorsqu'ils sont appelés dans un job.

Vous créez des variables locales, des paramètres et des variables globales à l'aide de la fenêtre Variables et paramètres dans Designer.

Vous pouvez définir des valeurs pour des variables locales ou globales dans des objets de script. Vous pouvez également définir des valeurs de variables globales à l'aide d'un job externe, d'une exécution ou de propriétés de planification.

Vous pouvez utiliser des variables comme noms de fichier pour les sources et cibles suivantes :

- Sources et cibles de fichier plat
- Sources et cibles de fichier XML
- Cibles de message XML (exécutées dans Designer en mode test)
- Sources et cibles de fichier IDoc (dans un environnement d'application SAP)
- Sources et cibles de message IDoc (dans un environnement d'application SAP)

Affichage de variables et paramètres [page 310]

Utilisez la fenêtre *Variables et paramètres* pour afficher, modifier et créer des variables locales et globales et des paramètres pour les objets sélectionnés dans Designer.

Fenêtre Variables et paramètres [page 311]

La boîte de dialogue *Variables et paramètres* se maintient au-dessus de ce que vous avez ouvert dans la zone de projet et l'espace de travail de Designer.

Onglets Définitions et Appels [page 311]

La boîte de dialogue *Variables et paramètres* est constituée des onglets *Définitions* et *Appels*. La valeur *Contexte* reflète le nom de l'objet sélectionné.

Types de variables et paramètres [page 313]

Les variables et paramètres sont utilisés différemment en fonction du type d'objet et de si la variable est locale ou globale.

Accès à Afficher utilisation depuis la fenêtre Variables et paramètres [page 314]

Utilisez *Afficher utilisation* pour les objets dans la boîte de dialogue *Variables et paramètres*.

[Utiliser des variables locales et des paramètres \[page 315\]](#)

Pour transférer une variable locale vers un autre objet, définissez la variable locale, puis à partir de l'objet appelant, créez un paramètre et mappez le paramètre à la variable locale en saisissant une valeur de paramètre.

[Utiliser des variables globales \[page 319\]](#)

Les variables globales sont globales uniquement dans un job. Il n'est pas nécessaire de définir des paramètres lorsque vous utilisez des variables globales.

[Règles de variables locales et globales \[page 325\]](#)

Il existe des règles à suivre lors de la définition de valeurs pour des variables locales et globales.

[Variables d'environnement \[page 326\]](#)

[Définir des noms de fichier lors de l'exécution à l'aide de variables \[page 327\]](#)

Vous pouvez définir des noms de fichier lors de l'exécution en indiquant une variable comme nom de fichier.

Informations associées

[Paramètres \[page 315\]](#)

23.1 Affichage de variables et paramètres

Utilisez la fenêtre *Variables et paramètres* pour afficher, modifier et créer des variables locales et globales et des paramètres pour les objets sélectionnés dans Designer.

1. Dans Designer sélectionnez ► *Outils* ► *Variables* ►.

La fenêtre *Variables et paramètres* s'ouvre.

2. Double-cliquez sur l'objet applicable dans la bibliothèque d'objets ou cliquez sur un objet dans la zone de projet ou l'espace de travail.

L'étiquette *Contexte* de la fenêtre *Variables et paramètres* affiche le nom de l'objet sélectionné ou de l'objet affiché dans l'espace de travail. Si aucun objet n'est sélectionné ou si l'objet n'est pas applicable aux variables et paramètres, le contenu de la fenêtre ne change pas ou reste vide.

3. Le cas échéant, ouvrez un autre objet.

Le contenu de *Variables et paramètres* change pour refléter les variables et paramètres de l'objet nouvellement ouvert.

Informations associées

[Fenêtre Variables et paramètres \[page 311\]](#)

[Types de variables et paramètres \[page 313\]](#)

[Onglets Définitions et Appels \[page 311\]](#)

23.2 Fenêtre Variables et paramètres

La boîte de dialogue *Variables et paramètres* se maintient au-dessus de ce que vous avez ouvert dans la zone de projet et l'espace de travail de Designer.

Le *Contexte*

Lorsque vous ouvrez d'autres objets, cette boîte ne se ferme pas, elle ne se minimise pas non plus, mais il est possible de la glisser à une autre position si nécessaire. Lorsque vous cliquez sur d'autres objets applicables, le logiciel modifie le contenu de l'étiquette (en affichant le nom de l'objet ouvert) en haut de la boîte de dialogue *Variables et paramètres* pour afficher les variables et les paramètres de l'objet sélectionné. L'étiquette *Contexte* reflète le nom de l'objet nouvellement sélectionné. Si l'objet n'est pas applicable aux variables et paramètres, ou s'il n'a pas de variable ni de paramètre définis, le contenu de la boîte ne se modifie pas ou reste vide.

Deux onglets composent la partie principale de la boîte : l'étiquette *Définitions* (en haut de la boîte de la dialogue) affiche le nom de l'objet sélectionné et *Appels*.

- *Définitions* : afficher et créer des variables et paramètres.
- *Appels* : afficher le nom de chaque paramètre défini pour tous les objets dans une définition d'objet parent.

De plus, si vous cliquez droit avec le bouton sur une variable ou un paramètre dans les onglets et sélectionnez *Afficher utilisation* dans le menu contextuel, vous obtiendrez des informations concernant les autres objets utilisant la variable et le paramètre en question.

Informations associées

[Onglets Définitions et Appels \[page 311\]](#)

[Types de variables et paramètres \[page 313\]](#)

[Accès à Afficher utilisation depuis la fenêtre Variables et paramètres \[page 314\]](#)

23.3 Onglets Définitions et Appels

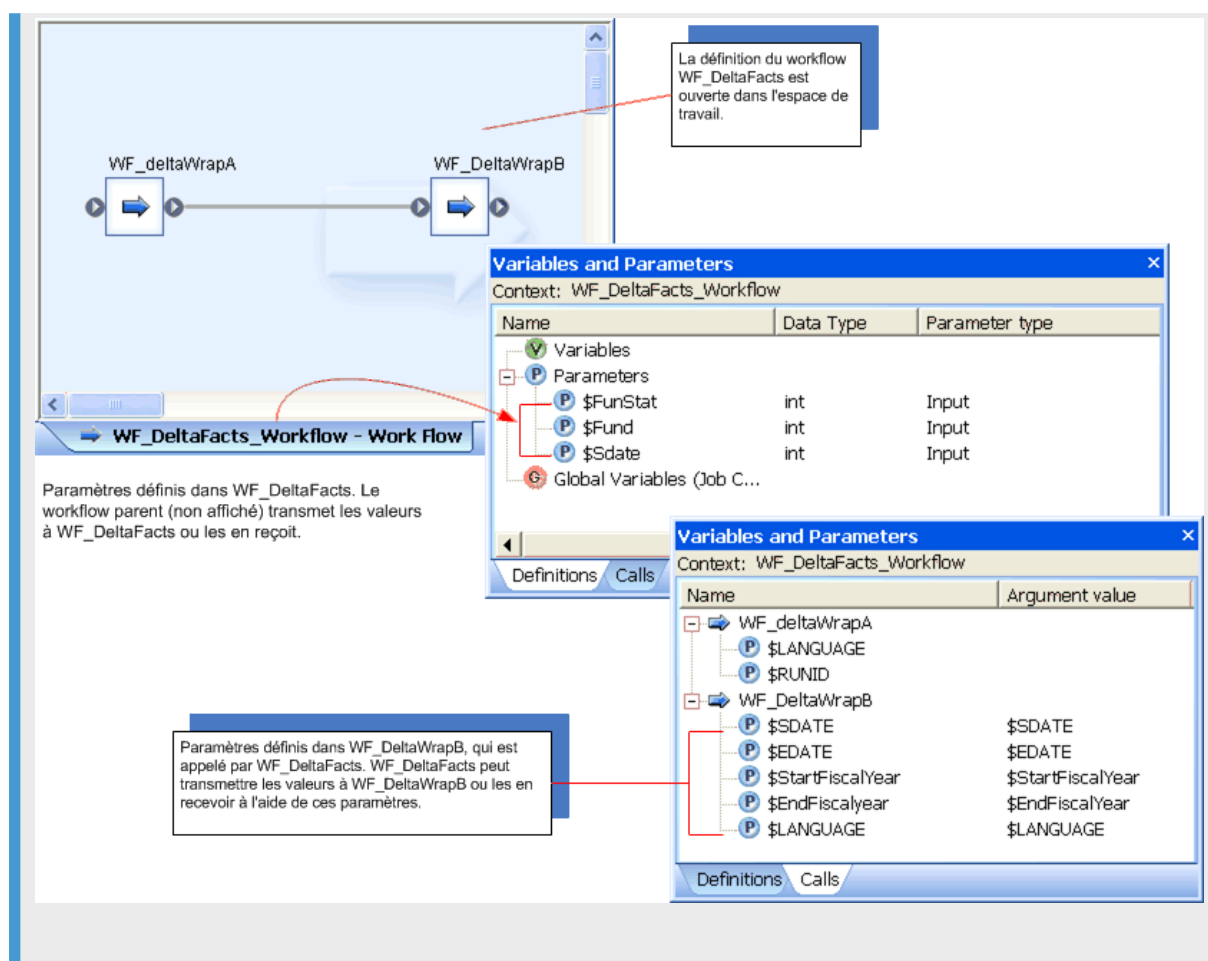
La boîte de dialogue *Variables et paramètres* est constituée des onglets *Définitions* et *Appels*. La valeur *Contexte* reflète le nom de l'objet sélectionné.

Dans les onglets *Définitions* et *Appels*, il est possible de cliquer avec le bouton droit sur une variable ou un nom de paramètre et sélectionner *Afficher utilisation* depuis le menu contextuel. Une boîte de dialogue *Rechercher* s'ouvre, répertoriant les autres objets utilisant la variable ou le paramètre mis en surbrillance.

Onglet	Objectif
<i>Définitions</i>	Afficher les variables (nom et type de données) et les paramètres (nom, type de données et type de paramètre) pour l'objet sélectionné et créer de nouvelles variables et de nouveaux paramètres. • Définir la variable locale et les paramètres au niveau du workflow et du flux de données. • Définir les variables globales au niveau de job.
<i>Appels</i>	Afficher le nom de chaque paramètre défini pour les objets dans une définition d'objet parent et saisir les valeurs de chaque paramètre. • Type de paramètre d'entrée : les valeurs peuvent être des constantes, des variables ou autre paramètre. • Sortie ou type de paramètre d'entrée/sortie : les valeurs peuvent être des variables ou des paramètres. En outre, les valeurs de l'onglet <i>Appels</i> doivent utiliser : • Le même type de données que la variable, car elles se situent dans une entrée ou un type de paramètre entrée/sortie. • Un type de données compatible si elles se situent dans un type de paramètre de sortie. • Des règles et la syntaxe de la langue de script.

❖ Exemple

L'illustration suivante présente le contenu des onglets *Définitions* et *Appels* dans la fenêtre *Variables et paramètres* pour le workflow en cours appelé `WF_DeltaFacts`. Notez que le nom du workflow s'affiche dans l'étiquette *Contexte*, en haut de la fenêtre *Variables et paramètres*.



Informations associées

[Fenêtre Variables et paramètres \[page 311\]](#)

[Accès à Afficher utilisation depuis la fenêtre Variables et paramètres \[page 314\]](#)

[Affichage de variables et paramètres \[page 310\]](#)

23.4 Types de variables et paramètres

Les variables et paramètres sont utilisés différemment en fonction du type d'objet et de si la variable est locale ou globale.

La table suivante répertorie les variables et paramètres que vous pouvez créer en fonction du type d'objet et la façon de les utiliser.

Type d'objet	Variables et paramètres applicables	Utilisé par
Job	Variables locales	Script ou structure conditionnelle dans le job
	Variables globales	Tout objet dans le job
Workflow	Variables locales	Workflow actuel transféré vers d'autres workflows ou flux de données à l'aide d'un paramètre. Objets parents pour transférer des variables locales. Les workflows peuvent également renvoyer des variables ou des paramètres aux objets parents.
	Paramètres	
Flux de données	Paramètres	Clause WHERE, mappage de colonne ou fonction dans le flux de données. Les flux de données ne peuvent pas renvoyer des valeurs de sortie.

23.5 Accès à Afficher utilisation depuis la fenêtre Variables et paramètres

Utilisez *Afficher utilisation* pour les objets dans la boîte de dialogue *Variables et paramètres*.

1. Depuis la bibliothèque d'objets, ouvrez un job, un workflow ou un flux de données.
2. Sélectionnez **Outils** > **Variables**.
3. Soit dans l'onglet *Définitions* soit dans l'onglet *Appels*, cliquez avec le bouton droit sur le nom d'une variable applicable puis sélectionnez *Afficher utilisation* du menu contextuel.

La boîte de dialogue *Rechercher* s'ouvre et affiche une liste de tous les objets utilisant la variable.

i Remarque

Il n'est pas nécessaire de modifier les paramètres dans la boîte de dialogue *Recherche*. Le logiciel renseigne automatiquement les critères de recherche pour trouver les informations « *afficher utilisation* » pour la variable sélectionnée.

Sans fermer la boîte de dialogue *Recherche* ou celle de *Variables et paramètres*, vous pouvez sélectionner un autre objet de la bibliothèque d'objets et effectuer la même procédure pour obtenir les informations d'une autre variable.

Informations associées

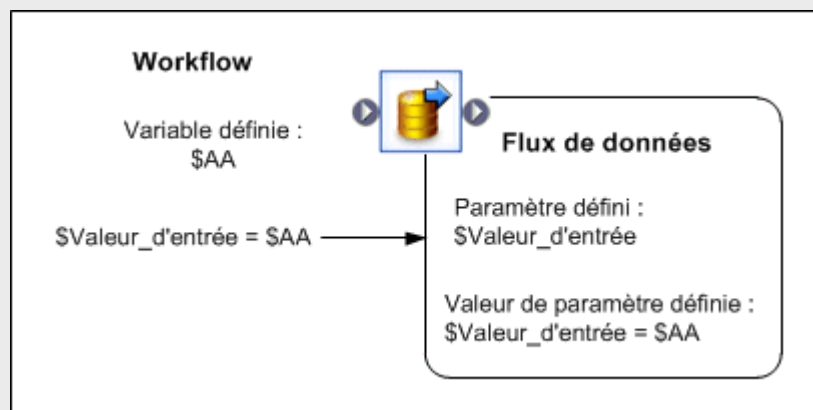
[Fenêtre Variables et paramètres \[page 311\]](#)

23.6 Utiliser des variables locales et des paramètres

Pour transférer une variable locale vers un autre objet, définissez la variable locale, puis à partir de l'objet appelant, créez un paramètre et mappez le paramètre à la variable locale en saisissant une valeur de paramètre.

❖ Exemple

Pour utiliser une variable locale dans un flux de données, définissez la variable dans un workflow parent puis transférez la valeur de la variable comme paramètre du flux de données.



23.6.1 Paramètres

Des paramètres peuvent être définis pour :

- Transférer leurs valeurs aux workflows ou à l'extérieur de workflows
- Transférer leurs valeurs aux flux de données

Un type : entrée, sortie ou entrée/sortie est affecté à chaque paramètre. La valeur transférée par le paramètre peut être utilisée par tout objet appelé par le workflow ou flux de données.

i Remarque

Vous pouvez également créer des variables locales et des paramètres à utiliser dans des fonctions personnalisées.

Informations associées

[Guide de référence : fonctions et procédures, fonctions personnalisées \[page 681\]](#)

23.6.2 Ajouter un paramètre à un workflow ou à un flux de données

1. Cliquez sur le nom du workflow ou du flux de données.
2. Cliquez sur **Outils** > **Variables** .
La fenêtre *Variables et paramètres* s'affiche.
3. Accédez à l'onglet *Définition*.
4. Sélectionnez *Paramètres*.
5. Cliquez avec le bouton droit de la souris et sélectionnez *Insérer*. Ou, pour insérer un paramètre à une position spécifique dans la liste, sélectionnez un paramètre existant, cliquez avec le bouton droit de la souris et sélectionnez *Insérer avant* ou *Insérer après*.
Un nouveau paramètre s'affiche (par exemple, `$NewParameter0`). Un cadre de contexte s'affiche et le curseur prend la forme d'une flèche avec un crayon jaune.
6. Pour modifier le nom d'une variable, cliquez sur la cellule du nom.
Le nom peut contenir des caractères alphanumériques ou des traits de soulignement, mais il ne peut pas contenir d'espace. Le nom doit toujours commencer par un symbole dollar (\$).
7. Cliquez sur la cellule du type de données pour le nouveau paramètre et sélectionnez le type de données approprié dans la liste déroulante.
Si le paramètre est une entrée ou un paramètre d'entrée/de sortie, il doit avoir le même type de données que la variable ; si le paramètre est un type de paramètre de sortie, il doit avoir un type de données compatible.
8. Cliquez sur la cellule du type de paramètre et sélectionnez le type de paramètre (entrée, sortie ou entrée/sortie).
9. Pour spécifier la position du paramètre dans la liste, cliquez avec le bouton droit de la souris sur le paramètre et sélectionnez *Déplacer vers le haut* ou *Déplacer vers le bas*.
10. Fermez la fenêtre *Variables et paramètres*.

23.6.3 Transférer les valeurs dans le flux de données

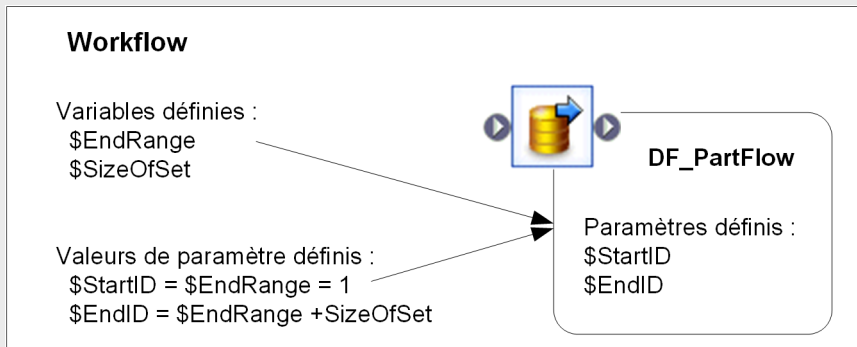
Vous pouvez utiliser une valeur transférée comme paramètre dans un flux de données pour contrôler les données transformées dans le flux de données.

❖ Exemple

Le flux de données DF_PartFlow traite quotidiennement des valeurs en stock. Il peut traiter tous les numéros de pièce fabricant utilisés ou une gamme de numéros de pièce fabricant en fonction d'exigences externes, par exemple la tranche de numéros traitée en dernier.

Si le workflow qui appelle DF_PartFlow enregistre la tranche de numéros traitée, il peut transférer la valeur de fin de la tranche \$EndRange comme paramètre aux flux de données pour indiquer la valeur de début de la tranche suivante à traiter.

Le logiciel peut calculer une nouvelle valeur de fin en fonction d'un numéro de pièce fabricant archivé à traiter chaque fois, par exemple \$SizeOfSet puis transférer cette valeur au flux de données comme valeur de fin. Une transformation Query dans le flux de données utilise les paramètres transférés pour filtrer les numéros de pièce fabricant extraits de la source.



Le flux de données pourrait être utilisé par plusieurs appels figurant dans un ou plusieurs workflows pour exécuter la même tâche sur différentes tranches de numéros en indiquant différents paramètres pour les appels spécifiques.

23.6.4 Définir la valeur du paramètre dans l'appel de flux

1. Ouvrez le job appelant, le workflow ou le flux de données.
2. Cliquez sur ► **Outils** ► **Variables** ► pour ouvrir la fenêtre *Variables et paramètres*.
3. Sélectionnez l'onglet *Appels*.

L'onglet Appels affiche tous les objets appelés à partir du job en cours, workflow ou flux de données.

4. Cliquez sur la cellule *Valeur d'argument*.
Un cadre de contexte s'affiche et le curseur prend la forme d'une flèche avec un crayon jaune.
5. Saisissez l'expression que le paramètre transférera dans la cellule.

Si le type de paramètre est une entrée, sa valeur peut être une expression contenant une constante (par exemple, 0, 3 ou 'string1'), une variable ou un autre paramètre (par exemple, \$startID ou \$parm1).

Si le type de paramètre est une sortie ou une entrée/sortie, la valeur doit être une variable ou un paramètre. La valeur ne peut pas être une constante car par définition, la valeur d'un paramètre sortie ou entrée/sortie peut être modifiée par tout objet dans le flux.

Pour indiquer des valeurs spécifiques, utilisez la syntaxe suivante :

Type de valeur	Syntaxe spécifique
Variable	<\$variable_name>

Type de valeur	Syntaxe spécifique
Chaîne	<code><'string>'</code>

23.6.5 Définir une variable locale

Étapes pour utiliser la fenêtre Variables et paramètres afin de définir une variable locale pour un job ou un flux de données.

1. Dans Designer, cliquez sur le nom du job ou du workflow dans la zone de projet ou l'espace de travail ou cliquez deux fois sur la bibliothèque d'objets.
2. Cliquez sur  **Outils** > **Variables** .
La fenêtre *Variables et paramètres* s'ouvre. La fenêtre affiche toutes les variables actuellement définies pour cet objet.
3. En ayant ouvert l'onglet Définition, cliquez avec le bouton droit sur **Variables** et sélectionnez *Insérer*. S'il existe des variables répertoriées, insérez une variable à une position spécifique dans la liste en cliquant avec le bouton droit sur une variable puis sélectionnez *Insérer avant* ou *Insérer après*.
Une nouvelle variable apparaît dans la liste sous le nom `$NewVariable0`. Une boîte de contexte s'affiche autour de la cellule du nom et le curseur prend la forme d'une flèche avec un crayon jaune pour modification.
4. Cliquez sur `$NewVariable0` dans la cellule du nom et saisissez un nouveau nom pour la variable : le nom doit commencer avec le symbole dollar (\$).
Le nom peut contenir des caractères alphanumériques ou des traits de soulignement, mais il ne peut pas contenir d'espace. Le nom doit toujours commencer par un symbole dollar (\$).
5. Cliquez sur la cellule du type de données pour la nouvelle variable et sélectionnez le type de données approprié dans la liste déroulante.
6. Facultatif. Si plus d'une variable est créée, indiquez l'ordre de la nouvelle variable dans la liste à l'aide du menu contextuel (cliquez avec le bouton droit) et sélectionnez *Déplacer vers le haut* ou *Déplacer vers le bas*.
7. Une fois les variables créées et modifiées, fermez la fenêtre Variables et paramètres en cliquant sur le « x » en haut à droite.

23.6.6 Définir une variable locale

Étapes pour copier une variable existante afin d'en créer une nouvelle.

1. Dans Designer, cliquez sur le nom du job ou du workflow dans la zone de projet ou l'espace de travail ou cliquez deux fois sur la bibliothèque d'objets.
2. Cliquez sur  **Outils** > **Variables** .
La fenêtre *Variables et paramètres* s'affiche. La fenêtre affiche toutes les variables actuellement définies pour cet objet.

3. En ayant ouvert l'onglet *Définition*, cliquez avec le bouton droit sur la variable à copier et cliquez sur *Répliquer*.
Une fenêtre de propriétés s'ouvre et affiche une nouvelle variable avec les mêmes propriétés que la variable répliquée et un nom modifiable (par exemple, \$Copy_1_NewVariable1).
4. Saisissez le nouveau nom de la variable et effectuez toute autre modification de Type de données, le cas échéant.
Le nom peut contenir des caractères alphanumériques ou des traits de soulignement, mais il ne peut pas contenir d'espace. Le nom doit toujours commencer par un symbole dollar (\$).
5. Cliquez sur *OK*.
La nouvelle variable apparaît dans la liste de variables de l'onglet *Définition*.
6. Facultatif. Cliquez avec le bouton droit sur le nom de la variable de la liste puis sélectionnez *Déplacer vers le haut* ou *Déplacer vers le bas* pour déplacer la variable à un nouvel emplacement dans la liste.
7. Une fois les variables créées et modifiées, fermez la fenêtre Variables et paramètres en cliquant sur le « x » en haut à droite.

23.7 Utiliser des variables globales

Les variables globales sont globales uniquement dans un job. Il n'est pas nécessaire de définir des paramètres lorsque vous utilisez des variables globales.

Lorsque vous affectez un nom à une variable globale, ce nom est exclusif au job, et il est impossible d'utiliser la même variable globale ou le même nom de variable globale dans un autre job.

Les variables globales apportent un maximum de flexibilité lors de l'exécution. Par exemple, au cours de la production, vous pouvez modifier des valeurs de variables globales par défaut au cours de l'exécution à partir d'une planification de job ou d'un appel SOAP sans avoir à ouvrir un job dans Designer.

Vous pouvez définir les valeurs des variables globales dans les objets de script, ou vous pouvez utiliser un job externe, une exécution ou des propriétés de planification.

23.7.1 Définition d'une variable globale

Étapes pour utiliser la fenêtre Variables et paramètres afin de définir une variable globale pour un job.

1. Dans Designer, cliquez sur le nom du job dans la zone de projet ou l'espace de travail ou cliquez deux fois sur la bibliothèque d'objets.
2. Cliquez sur  *Outils* > *Variables* .
- La fenêtre *Variables et paramètres* s'affiche. La fenêtre affiche toutes les variables actuellement définies pour cet objet.
3. En ayant ouvert l'onglet Définition, cliquez avec le bouton droit sur *Variables globales* et sélectionnez *Insérer*. S'il existe des variables répertoriées, insérez une variable à une position spécifique dans la liste en cliquant avec le bouton droit sur une variable puis sélectionnez *Insérer avant* ou *Insérer après*.

Une nouvelle variable apparaît dans la liste sous le nom `$NewJobGlobalVariable0`. Une boîte de contexte s'affiche autour de la cellule du nom et le curseur prend la forme d'une flèche avec un crayon jaune pour modification.

4. Cliquez sur `$NewJobGlobalVariable0` dans la cellule du nom et saisissez un nouveau nom pour la variable : le nom doit commencer avec le symbole dollar (\$).

Le nom peut contenir des caractères alphanumériques ou des traits de soulignement, mais il ne peut pas contenir d'espace. Le nom doit toujours commencer par un symbole dollar (\$).

5. Cliquez sur la cellule du type de données pour la nouvelle variable et sélectionnez le type de données approprié dans la liste déroulante.
6. Facultatif. Si plus d'une variable est créée, indiquez l'ordre de la nouvelle variable dans la liste à l'aide du menu contextuel (cliquez avec le bouton droit) et sélectionnez *Déplacer vers le haut* ou *Déplacer vers le bas*.
7. Une fois les variables créées et modifiées, fermez la fenêtre Variables et paramètres en cliquant sur le « x » en haut à droite.

23.7.2 Répliquer une variable globale

Étapes pour copier une variable globale existante afin d'en créer une nouvelle.

1. Dans Designer, cliquez sur le nom du job dans la zone de projet ou l'espace de travail ou cliquez deux fois sur un nom de job dans la bibliothèque d'objets.
2. Cliquez sur **Outils** > **Variables** .
La fenêtre *Variables et paramètres* s'affiche. La fenêtre affiche toutes les variables actuellement définies pour cet objet.
3. En ayant ouvert l'onglet *Définition*, cliquez avec le bouton droit sur la variable globale à copier et cliquez sur *Répliquer*.
Une fenêtre de propriétés s'ouvre et affiche une nouvelle variable avec les mêmes propriétés que la variable répliquée et un nom modifiable (par exemple, `$Copy_1_NewJobGlobalVariable1`).
4. Modifiez le nom de la nouvelle variable globale et effectuez toute autre modification de Type de données, le cas échéant.
Le nom peut contenir des caractères alphanumériques ou des traits de soulignement, mais il ne peut pas contenir d'espace. Le nom doit toujours commencer par un symbole dollar (\$).
5. Cliquez sur *OK*.
La nouvelle variable globale apparaît dans la liste de variables globales de l'onglet *Définition*.
6. Facultatif. Cliquez avec le bouton droit sur le nom de la variable globale de la liste puis sélectionnez *Déplacer vers le haut* ou *Déplacer vers le bas* pour déplacer la variable à un nouvel emplacement dans la liste.
7. Une fois les variables globales créées et modifiées, fermez la fenêtre Variables et paramètres en cliquant sur le « x » en haut à droite.

23.7.3 Afficher des variables globales à partir de la boîte de dialogue Propriétés

Il est possible d'afficher les variables globales d'un job à l'aide de la fenêtre Variables et paramètres mais vous pouvez également les afficher depuis la boîte de dialogue Propriétés d'un job sélectionné.

Les variables globales définies dans un job sont visibles pour les objets relatifs à ce job. Une variable globale définie dans un job ne peut pas être modifiée ou affichée à partir d'un autre job.

Pour afficher des variables globales à partir de la boîte de dialogue Propriétés d'un job :

1. Dans la bibliothèque d'objets, sélectionnez l'onglet *Jobs*.
2. Cliquez avec le bouton droit sur le job applicable et sélectionnez *Propriétés*.
3. Cliquez sur l'onglet *Variable globale*.
4. Facultatif. Utilisez la zone *Filtre* pour filtrer la liste de variables globales en fonction de caractéristiques spécifiques.

23.7.4 Valeurs pour les variables globales

Outre la définition d'une valeur de variable globale dans un job à l'aide d'un script d'initialisation, vous pouvez également définir et gérer des valeurs de variable globale à l'extérieur d'un job.

Les valeurs définies à l'extérieur d'un job sont traitées de la même façon que celles définies à l'intérieur d'un job (dans un script d'initialisation par exemple). Ces deux méthodes ont leurs avantages.

Lorsque vous avez défini des valeurs à la fois internes et externes pour une variable globale dans un job, le logiciel choisit automatiquement la valeur définie à l'intérieur du job comme valeur à utiliser.

Définir des valeurs à l'extérieur d'un job a certains avantages. Par exemple, il est possible d'ouvrir la boîte de dialogue Propriétés d'un job pour afficher les valeurs de substitution globales au lieu d'ouvrir le job.

Vous définissez les valeurs de variable globale à l'extérieur d'un job des façons suivantes :

- Comme propriété de job
- Comme propriété d'exécution ou de planification

i Remarque

Il n'est pas nécessaire d'affecter une valeur à une variable globale. Toutefois, le logiciel lit comme NULL une variable globale sans valeur.

Lorsque vous définissez des valeurs pour des variables globales à l'extérieur d'un job, il est possible d'afficher les valeurs dans les boîtes de dialogue Propriétés et Propriétés d'exécution dans Designer et dans les pages Options d'exécution et Planification de l'Administrateur. De plus, vous pouvez modifier facilement les valeurs lorsque vous testez ou planifiez le job.

i Remarque

Vous ne pouvez pas transférer des variables globales comme arguments de ligne de commande pour des jobs en temps réel.

23.7.5 Définir une valeur de variable globale comme propriété de job

1. Cliquez avec le bouton droit de la souris sur la librairie d'objets ou sur la zone de projet
2. Cliquez sur [Propriétés](#).
3. Cliquez sur l'onglet [Variable globale](#).

Toutes les variables globales créées dans ce job s'affichent.

4. Pour filtrer par nom les variables globales affichées, saisissez une partie du nom dans la zone [Filtre](#).
5. Saisissez des valeurs pour les variables globales dans ce job.

Vous pouvez utiliser toute instruction utilisée dans un script avec cette option.

6. Pour définir plusieurs valeurs de variable globale, utilisez la touche Contrôle pour sélectionner les variables, cliquez avec le bouton droit de la souris et sélectionnez [Mettre à jour la valeur](#). Saisissez la valeur dans l'éditeur rapide et cliquez sur [OK](#).
7. Cliquez sur [OK](#).

Le logiciel sauvegarde les valeurs dans le référentiel comme propriétés de job.

Vous pouvez également afficher et modifier ces valeurs par défaut dans la boîte de dialogue Propriétés d'exécution de Designer et dans les pages Options d'exécution et Planification de l'Administrateur. Vous pouvez ainsi substituer les valeurs de propriété du job lors de l'exécution.

Informations associées

Guide de référence : [Langage de script \[page 706\]](#)

23.7.6 Définir une valeur de variable globale comme propriété d'exécution

1. Exécutez un job à partir de Designer ou bien exécutez ou planifiez un job batch à partir de l'Administrateur.

i Remarque

A des fins de test, vous pouvez exécuter des jobs en temps réel à partir de Designer en mode test. Vérifiez que vous avez défini les propriétés d'exécution pour un job en temps réel.

2. Affichez les variables globales dans le job et leurs valeurs par défaut (le cas échéant).
3. Modifiez les valeurs des variables globales le cas échéant.
4. Si vous utilisez Designer, cliquez sur [OK](#). Si vous utilisez l'Administrateur, cliquez sur [Exécuter](#) ou [Planifier](#).

Le job est exécuté à l'aide des valeurs que vous saisissez. Les valeurs saisies comme propriétés d'exécution ne sont pas sauvegardées. Les valeurs saisies en tant que propriétés de planification sont sauvegardées mais sont accessibles uniquement depuis l'Administrateur.

23.7.7 Classement automatique de valeurs de variables globales dans un job

Lorsque vous définissez des valeurs pour les variables globales et que vous saisissez des valeurs différentes pour la même variable globale à l'aide de différentes méthodes, le logiciel utilise le classement automatique afin de sélectionner la valeur à utiliser pour une exécution de job.

Le logiciel sélectionne la valeur de classement la plus élevée pour l'utiliser dans le job. Une valeur saisie comme propriété de job a le plus petit classement. Une valeur définie à l'intérieur d'un job a le classement le plus élevé.

23.7.7.1 Propriété d'exécution et propriété de job

Si vous définissez une valeur de variable globale à l'aide d'une propriété de job et que vous définissez une variable globale différente (pour la même variable globale) à l'aide d'une propriété d'exécution, la valeur de propriété d'exécution écrase la valeur de propriété de job et devient la valeur par défaut pour l'exécution de job en cours.

i Remarque

Vous ne pouvez pas sauvegarder de valeurs de variables globales de propriété d'exécution.

❖ Exemple

Par exemple, supposez qu'un job, JOB_Test1, ait trois variables globales déclarées :

- \$YEAR
- \$MONTH
- \$DAY

Vous définissez la variable « \$YEAR » en tant que propriété de job avec la valeur « 2003 ».

Pour l'exécution de job initiale, vous définissez les valeurs de \$MONTH et \$DAY en tant que propriétés d'exécution comme suit :

- \$MONTH = JANUARY'
- \$DAY = 31

Le logiciel exécute une liste d'instructions contenant des valeurs par défaut pour JOB_Test1 :

```
$YEAR=2003;  
$MONTH=' JANUARY ' ;  
$DAY=31 ;
```

Pour la seconde exécution de job, vous définissez les variables \$YEAR = « 2002 » et \$MONTH = « JANUARY » en tant que propriétés de job mais ne définissez pas de valeur pour \$DAY. Le logiciel exécute la liste suivante d'instructions :

```
$YEAR=2002 ;  
$MONTH=' JANUARY ' ;
```

La valeur de \$DAY n'ayant pas été définie lors de l'exécution du second job, le logiciel la lit comme NULL ; la valeur \$YEAR est 2002 et non 2003 car vous avez défini « 2002 » en tant que propriété d'exécution lors de la seconde exécution.

23.7.7.2 Propriétés de job et propriétés de planification

Si vous avez défini une valeur de variable globale à la fois pour une propriété de planification et de job, la valeur de propriété de planification se substitue à la valeur de propriété de job et devient la valeur par défaut externe pour l'exécution de job en cours.

Le logiciel sauvegarde la valeur de propriété de planification dans le référentiel. Cependant, ces valeurs sont uniquement associées à une planification de job, pas au job lui-même. Par conséquent, vous devez afficher et modifier ces valeurs dans l'Administrateur.

23.7.7.3 Propriété interne au job et propriété externe

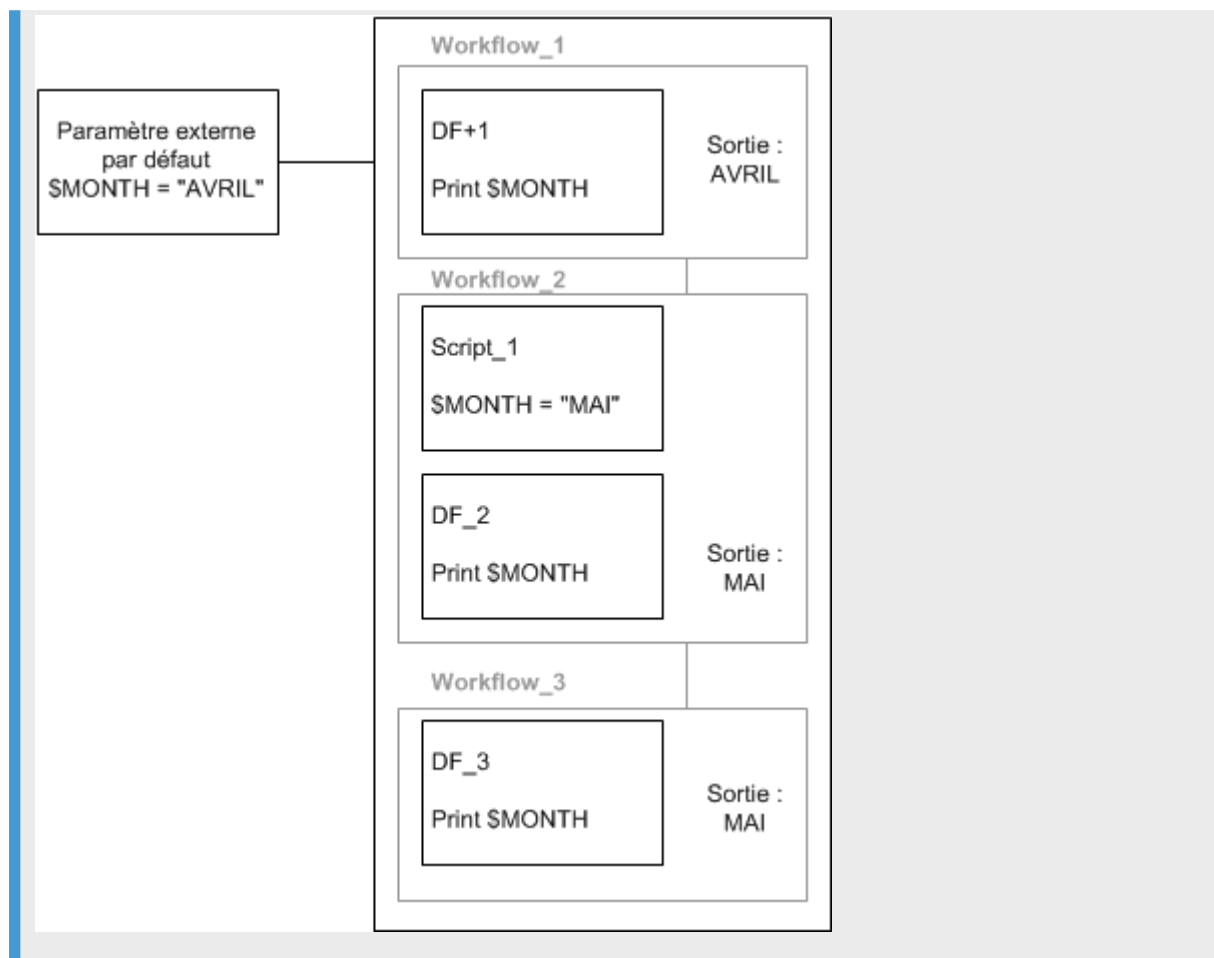
La valeur est déterminée lorsque le logiciel tente d'appliquer la valeur externe.

Une valeur de variable globale définie à l'intérieur d'un job se substitue toujours à toute valeur externe. Cependant, la substitution n'est pas exécutée tant que le logiciel tente d'appliquer les valeurs externes au job en cours d'exécution avec la valeur interne. Jusqu'à là, le logiciel traite les valeurs d'exécution, de planification ou de propriété comme valeurs par défaut.

❖ Exemple

Par exemple, supposez un job nommé JOB_Test2 contenant trois workflows, chacun comportant un flux de données. Le deuxième flux de données se trouve dans un workflow précédé d'un script dans lequel la valeur MAY est définie pour \$MONTH. Le premier et le troisième flux de données ont la même variable globale sans valeur définie. La propriété d'exécution \$MONTH = 'APRIL' constitue la valeur de variable globale.

Dans ce scénario, 'APRIL' devient la valeur par défaut pour le job. 'APRIL' reste la valeur de la variable globale tant qu'elle ne se chevauche pas avec une autre valeur pour la même variable dans le deuxième workflow. La valeur du script se trouvant dans le job, 'MAY' se substitue à 'APRIL' pour la variable \$MONTH. Le logiciel continue à traiter le job avec cette nouvelle valeur.



23.8 Règles de variables locales et globales

Il existe des règles à suivre lors de la définition de valeurs pour des variables locales et globales.

La table suivante décrit les règles spécifiques à suivre lorsque vous définissez ces variables.

Règle pour	Description
Attribution de nom	- Les variables locales et globales doivent avoir des noms uniques dans leur contexte de job. - Toute modification de nom apportée à une variable globale peut être effectuée au niveau du job uniquement.

Règle pour	Description
Réplication de jobs et de workflows	- Lorsque vous répliquez tous les objets, les variables locales et globales définies dans ce contexte de job sont également répliquées. - Lorsque vous répliquez un flux de données ou un workflow, tous les paramètres et toutes les variables locales et globales sont également répliqués. Cependant vous devez valider ces variables locales et globales dans le contexte de job dans lequel elles ont été créées. i Remarque Si vous tentez de valider un flux de données ou un workflow contenant des variables globales sans job, Data Services indique une erreur.
Importation et exportation	- Lorsque vous exportez un objet du job, vous exportez également toutes les variables locales et globales définies pour ce job. - Lorsque vous exportez un objet de niveau inférieur (par exemple, un flux de données) sans le job parent, la variable globale n'est pas exportée. Seul l'appel vers cette variable globale est exporté. i Remarque Si vous utilisez cet objet dans un autre job sans définir la variable globale dans le nouveau job, une erreur de validation se produira.

23.9 Variables d'environnement

Vous pouvez utiliser des variables d'environnement système dans des jobs, des workflows ou des flux de données. Les fonctions `get_env`, `set_env` et `is_set_env` donnent accès aux variables de système d'exploitation sous-jacent qui se comportent comme l'autorise le système d'exploitation.

Vous pouvez définir temporairement la valeur d'une variable d'environnement dans un job, un workflow ou un flux de données. Une fois définie, la valeur est visible pour tous les objets de ce job.

Utilisez les fonctions `get_env`, `set_env` et `is_set_env` pour définir, récupérer et tester les valeurs de variables d'environnement.

Informations associées

[Guide de référence : Fonctions et procédures \[page 676\]](#)

23.10 Définir des noms de fichier lors de l'exécution à l'aide de variables

Vous pouvez définir des noms de fichier lors de l'exécution en indiquant une variable comme nom de fichier.

Les variables peuvent être utilisées comme noms de fichier pour les sources et cibles suivantes :

- Fichiers plats
- Fichiers XML et messages
- Fichiers IDoc et messages (dans un environnement SAP)

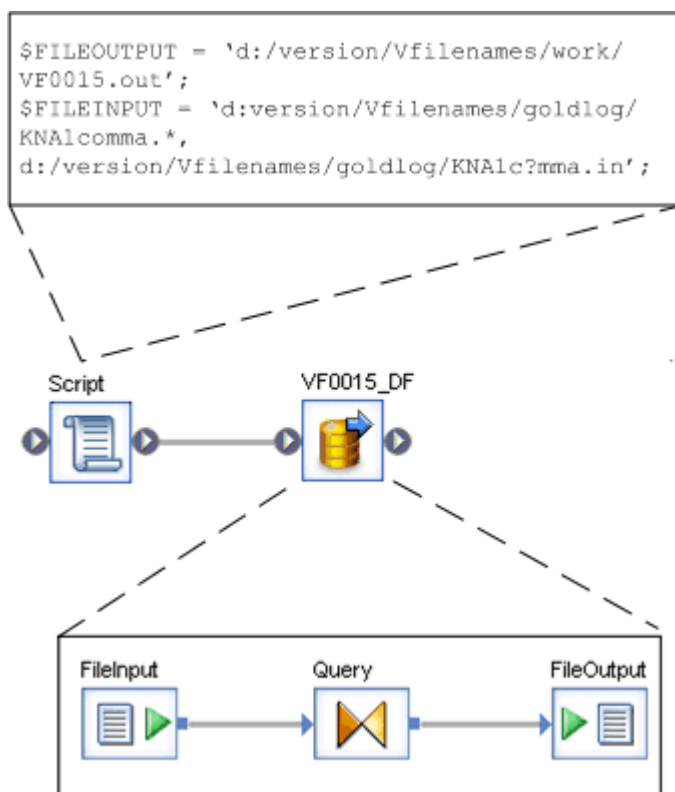
Les variables peuvent également être utilisées comme noms de fichier pour la fonction `lookup_ext` (pour un fichier plat utilisé comme paramètre de table de recherche).

23.10.1 Utilisation d'une variable dans un nom de fichier plat

Définissez un nom de fichier plat à l'aide d'une variable au cours de l'exécution.

1. Créez une variable locale ou globale à l'aide de **► Outils ► Variables ►** dans la fenêtre Variables et paramètres. Voir [Définir une variable locale \[page 318\]](#) et [Définition d'une variable globale \[page 319\]](#) pour en savoir plus.
2. Créez un script pour définir la valeur d'une variable locale ou globale en suivant les étapes dans [Création d'un script \[page 251\]](#), ou appelez une variable d'environnement système.
3. Déclarez la variable dans l'éditeur de format de fichier ou dans l'éditeur de fonctions comme paramètre `lookup_ext`.
 - Lorsque vous définissez une valeur de variable pour un fichier plat, indiquez le nom de fichier et le nom du répertoire. Saisissez la variable dans la propriété *Fichier(s)* sous *Fichier(s) de données* dans l'éditeur de format de fichier. Vous ne pouvez pas saisir une variable dans la propriété *Répertoire racine*.
 - Pour les recherches, remplacez le chemin et le nom de fichier dans le cadre *Table de recherche* de l'éditeur de fonctions `lookup_ext`, par le nom de variable.

Le schéma suivant indique comment définir des valeurs pour des variables dans les sources et cibles de fichier plat dans un script.



Lorsque vous utilisez des variables comme sources et cibles, vous pouvez également utiliser plusieurs noms de fichier et caractères génériques. Ni l'un ni l'autre n'est pris en charge dans la fonction `lookup_ext`.

Le schéma ci-dessus donne un exemple sur la façon d'utiliser plusieurs noms de variables et caractères génériques. Notez que la variable `$FILEINPUT` comprend deux noms de fichier (séparés par une virgule). Les deux noms (`KNA1comma.*` et `KNA1c?mma.in`) utilisent également les caractères génériques (`*` et `?`) pris en charge par le logiciel.

Informations associées

[Guide de référence : langage de script \[page 706\]](#)

23.11 Paramètres de substitution

23.11.1 Présentation des paramètres de substitution

Utiliser le même paramètre de substitution dans plusieurs objets pour définir de manière efficace la même valeur dans un emplacement, au lieu de modifier chaque objet séparément.

Lorsque vous utilisez des paramètres de substitution, vous devez uniquement modifier une fois la valeur dans la boîte de dialogue [Variables et Paramètres](#) pour modifier la valeur de tous les objets qui utilisent le paramètre.

Par exemple, si vous créez un job faisant référence à un répertoire unique sur votre ordinateur local et que vous exportez ce job vers un autre ordinateur, le job recherchera le seul répertoire dans le nouvel environnement. Si ce répertoire n'existe pas, le job ne sera pas exécuté. Lorsque vous utilisez des paramètres de substitution, le logiciel met automatiquement à jour cette valeur dans tous les objets qui utilisent le paramètre de substitution.

Pour plus de confort, vous pouvez configurer un groupe de paramètres de substitution pour un environnement d'exécution spécifique en associant leurs valeurs constantes sous une configuration de paramètre de substitution. Créez plusieurs configurations pour regrouper les paramètres de substitution en fonction d'un environnement d'exécution, par exemple. Vous pouvez alors contrôler de nombreuses valeurs différentes en modifiant les configurations.

Utilisez des paramètres dans les expressions, les instructions SQL, les champs d'option et les chaînes de constante. Les paramètres de substitution peuvent apparaître comme une valeur dans de nombreuses options et de nombreux éditeurs d'expressions, afin que vous puissiez choisir une variable de substitution spécifique à partir de la liste de valeurs.

❖ Exemple

Utilisez un paramètre de substitution dans un script, tel que :

```
Print('Data read in : [$$FilePath]'); or Print('[$$FilePath]');
```

Utilisez un paramètre de substitution dans un format de fichiers comme le nom de fichier : [\$FilePath]/file.txt

Informations associées

[Création ou renommage d'une configuration pour un paramètre de substitution \[page 334\]](#)

[Paramètres de substitution par défaut pour la qualité de données \[page 330\]](#)

23.11.1.1 Paramètres de substitution par rapport aux variables globales

Les paramètres de substitution diffèrent des variables globales dans leur manière d'affecter vos jobs.

Les paramètres de substitution s'appliquent au niveau du référentiel. Les variables globales s'appliquent uniquement au job dans lequel elles sont définies. Par exemple, vous utilisez une variable globale pour une valeur de paramètre qui n'est pas calculée tant que le job est en cours d'exécution, alors que vous utilisez un paramètre de substitution pour les constantes qui ne sont pas modifiées pendant l'exécution. Un paramètre de substitution défini dans un référentiel local donné est disponible pour tous les jobs de ce référentiel. Par conséquent, l'utilisation d'un paramètre de substitution signifie que vous ne devez pas définir une variable globale dans chaque job pour paramétrer une valeur constante.

Le tableau suivant décrit les principales différences entre les variables globales et les paramètres de substitution.

Variables globales	Paramètres de substitution
Définies au niveau du job	Définis au niveau du référentiel
Ne peuvent pas être partagées via des jobs	Disponibles pour tous les jobs dans un référentiel
Spécifiques au type de données	Aucun type de données (toutes les chaînes)
La valeur peut varier au cours de l'exécution de job	Ensemble de constantes avant l'exécution du job (constantes)

Vous pouvez utiliser les paramètres de substitution dans tous les emplacements où les variables globales sont prises en charge ; par exemple, utilisez les paramètres de substitution dans les zones suivantes, lesquelles prennent également en charge les variables globales :

- Clauses WHERE de transformation Query
- Mappages
- Identificateurs d'instruction SQL pour transformation SQL
- Options de fichier plat
- Transformations User-defined
- options de transformation Address cleanse
- Seuils de correspondance

23.11.1.2 Paramètres de substitution par défaut pour la qualité de données

Le logiciel installe certains paramètres de substitution par défaut utilisés par des transformations Data Quality.

Par exemple, la table suivante liste les paramètres de substitution par défaut pour la transformation USA Regulatory Address Cleanse.

Paramètre de substitution	Description
\$\$RefFilesAddressCleanse	Définit l'emplacement des répertoires address cleanse.
\$\$ReportsAddressCleanse	Active la collection de données pour la création de rapports avec les statistiques address cleanse. Les options sont Oui ou Non.

23.11.2 Tâches dans l'Éditeur de paramètres de substitution

Vous pouvez utiliser l'[Éditeur de paramètres de substitution](#) pour exécuter de nombreuses tâches liées à la création et à la modification de paramètres de substitution et de configuration.

Ouvrez l'[Éditeur de paramètres de substitution](#) de Designer en sélectionnant ► [Outils](#) ► [Configurations des paramètres de substitution](#) ►.

Ce que vous pouvez faire dans l'Éditeur de paramètres de substitution

Tâche	Méthode
Créer un paramètre	Cliquez avec le bouton droit sur un paramètre de substitution existant, près de l'emplacement où vous souhaitez placer le nouveau paramètre, puis sélectionnez Insérer .
Copier un paramètre existant	Cliquez avec le bouton droit sur un paramètre de substitution existant puis sélectionnez Copier .
	i Remarque Voir Coller un paramètre copié.
Coller un paramètre copié	Cliquez avec le bouton droit sur un paramètre de substitution existant, en dessous de l'emplacement où vous souhaitez placer le paramètre copié, cliquez avec le bouton droit et sélectionnez Coller . Le paramètre collé apparaît dans la ligne au-dessus du paramètre sélectionné.
Modifier un paramètre existant	Cliquez avec le bouton droit sur le paramètre existant. Une icône crayon apparaît à gauche du nom du paramètre. Entrez vos modifications dans la zone de texte. Lorsque vous avez terminé, cliquez sur une autre ligne pour afficher vos modifications.
Filtrer la liste des paramètres pour une recherche plus aisée	Saisissez le nom ou le nom partiel du paramètre dans la zone de texte Filtre et la liste des paramètres affiche automatiquement un sous-ensemble des paramètres de substitution qui correspondent à votre filtre.
	i Remarque Vous n'avez pas besoin d'inclure les symboles dollar dans le cadre du filtre. Par exemple, saisissez « référentiel » et non « \$\$référentiel ».
Afficher utilisation	À l'aide du bouton : sélectionnez un paramètre existant et cliquez sur le bouton Afficher utilisation À l'aide du menu contextuel : Cliquez avec le bouton droit sur le paramètre existant et sélectionnez Afficher utilisation à partir du menu contextuel. La boîte de dialogue Rechercher s'ouvre et affiche une liste de tous les objets utilisant le paramètre.
Réorganiser la liste de paramètres	Sélectionnez un paramètre et glissez-le vers une nouvelle position dans la liste.

Tâche	Méthode
Trier les configurations	<i>i</i> Remarque Vous devez avoir créé plus d'une configuration. Cliquez sur l'icône <i>Trier les noms de configuration par ordre croissant</i> ou <i>Trier les noms de configuration par ordre décroissant</i> dans la barre d'outils.
Définir un groupe de configurations par défaut	Cliquez sur la flèche déroulante à côté de l'option <i>Configuration par défaut</i> puis sélectionnez un nom de configuration.
Créer un groupe de configuration	Cliquez sur l'icône <i>Créer une configuration des paramètres de substitution</i> dans la barre d'outils.
Ajouter une configuration au paramètre de substitution	Cliquez deux fois sur la zone de texte de configuration dans la ligne de paramètre applicable. L'icône crayon apparaît à gauche du nom du paramètre. Saisissez la nouvelle configuration.
Renommer une configuration existante.	Sélectionnez l'en-tête de colonne de la configuration applicable et cliquez sur l'icône <i>Renommer la configuration des paramètres de substitution</i> dans la barre d'outils. Saisissez le nouveau nom.
Créer une copie de configuration	Sélectionnez la configuration applicable et cliquez sur l'icône <i>Créer une configuration des paramètres de substitution dupliquée</i> dans la barre d'outils.
Supprimer une configuration	Sélectionnez l'en-tête de colonne d'une configuration (il ne peut pas s'agir d'une configuration par défaut) puis cliquez sur l'icône <i>Supprimer la configuration des paramètres de substitution</i> dans la barre d'outils.
Déplacer la configuration par défaut désignée vers l'avant (la première colonne à côté de la liste de paramètres de substitution).	Cliquez sur l'icône <i>Déplacer la configuration par défaut vers l'avant</i> .

Informations associées

[Création d'un paramètre de substitution \[page 333\]](#)

23.11.2.1 Attribution de noms aux paramètres de substitution

Lorsque vous définissez des paramètres de substitution et leur attribuez des noms, utilisez les règles suivantes :

- Le préfixe de nom correspond à deux signes dollar \$\$ (les variables globales comportent un préfixe avec un signe dollar). Lorsque vous ajoutez de nouveaux paramètres de substitution dans l'éditeur de paramètres de substitution, l'éditeur ajoute automatiquement le préfixe.
- Lorsque vous saisissez des noms dans l'éditeur de paramètres de substitution, n'utilisez pas de ponctuation (ni guillemets ni crochets) sauf des traits de soulignement. Les caractères suivants ne sont pas autorisés :

```
, : / ' \ " = < > + | - * % ; \t [ ] ( ) \r \n $ ] +
```

- Vous pouvez saisir des noms directement dans des zones, des mappages de colonnes, des options de transformation etc. Cependant, vous devez les mettre entre crochets, par exemple, [\$\$SamplesInstall].
- Les noms peuvent contenir tout caractère numérique ou alphanumérique ou des traits de soulignement mais pas d'espaces.
- Les noms ne sont pas sensibles à la casse.
- La longueur maximale pour la plupart des types de référentiels est de 256 caractères (64 pour MySQL et 128 pour MS SQL server).
- Les noms doivent être uniques dans le référentiel.

23.11.2.2 Création d'un paramètre de substitution

Comment ajouter un paramètre de substitution à une configuration.

Vous créez des paramètres de substitution pour gagner du temps en saisissant la valeur une seule fois et en l'enregistrant dans une configuration de paramètre de substitution. Vous définissez une valeur constante pour le paramètre de substitution pour la raison suivante : lorsque vous utilisez le paramètre dans plusieurs objets, vous devez saisir la valeur constante une seule fois, lors de sa création.

1. Dans Designer, sélectionnez ► *Outils* ► *Configurations des paramètres de substitution* ►.

L'*Editeur de paramètres de substitution* s'ouvre. La première colonne, appelée Paramètre de substitution, liste les paramètres de substitution disponibles dans le référentiel. La seconde colonne contient la valeur constante pour le paramètre. Par défaut, l'en-tête de colonne est Configuration1.

2. Cliquez deux fois sur la ligne vide. Une icône crayon apparaît à gauche de la ligne; Saisissez le nom. (Il se peut que vous deviez faire défiler jusqu'en bas de la colonne Paramètres de substitution pour une ligne vide.)

Le logiciel ajoute automatiquement un préfixe de sigle de double dollar (\$\$) au nom lorsque vous naviguez dans la cellule.

3. Cliquez deux fois dans la cellule suivante sous la colonne de configuration applicable et saisissez la valeur constante que représente le paramètre dans cette configuration. Le logiciel applique la valeur constante au paramètre lorsque vous exécutez le job.

Informations associées

[Création ou renommage d'une configuration pour un paramètre de substitution \[page 334\]](#)

[Exportation et importation de paramètres de substitution \[page 338\]](#)

23.11.2.3 Création ou renommage d'une configuration pour un paramètre de substitution

Comment créer un groupe de configurations ou renommer un groupe de configurations existant.

Vous pouvez souhaiter créer un second groupe de configurations afin de pouvoir affecter plus d'une valeur constante à un paramètre de substitution. Par exemple, si vos fichiers de répertoires pour le test d'un job sont situés dans un emplacement différent des répertoires que vous utilisez dans la production, vous pourriez affecter deux valeurs constantes différentes pour le paramètre. Le logiciel nomme automatiquement le premier groupe de configuration Configuration1 et les groupes suivants Configuration_1, Configuration_n+1, etc.

1. Pour modifier le nom d'un groupe de configurations, sélectionnez l'en-tête de la colonne de configuration applicable pour la mettre en surbrillance, puis cliquez sur l'icône *Renommer la configuration des paramètres de substitution* dans la barre d'outils.

2. Saisissez le nouveau nom dans l'en-tête de la colonne en surbrillance.

Le nouveau nom apparaît dans l'en-tête de colonne.

3. Pour ajouter une configuration, cliquez sur l'icône *Créer une configuration des paramètres de substitution* dans la barre d'outils.

Une nouvelle colonne de configuration apparaît avec le nom affecté automatiquement en surbrillance (par exemple, Configuration_1).

4. Saisissez un nom unique dans l'en-tête de la colonne en surbrillance.

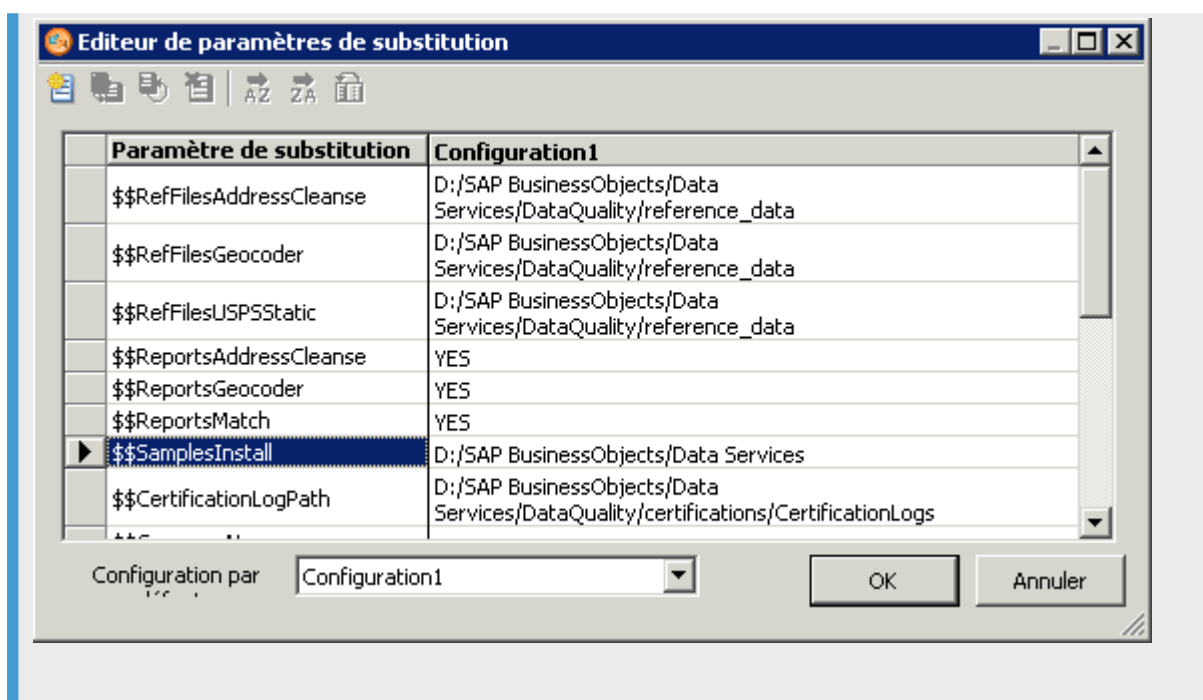
Le logiciel affiche deux colonnes de configuration.

Pour modifier la configuration par défaut qui s'appliquera lorsque vous exécutez des jobs, sélectionnez-la dans la zone de liste déroulante en haut de la fenêtre.

❖ Exemple

La capture d'écran suivante affiche deux groupes de configuration des paramètres de substitution. Le paramètre de substitution \$\$NetworkDir a la valeur D:/Data/Staging dans la configuration nommée Windows_Subst_Param_Conf et la valeur /usr/data/staging dans la configuration UNIX_Subst_Param_Conf.

Notez que chaque configuration peut contenir plusieurs paramètres de substitution.



Informations associées

[Configurations des paramètres de substitution et configurations système \[page 335\]](#)

23.11.3 Configurations des paramètres de substitution et configurations système

Une configuration système regroupe un ensemble de configurations de banque de données et une configuration des paramètres de substitution.

Une configuration des paramètres de substitution peut être associée à une ou plusieurs configurations système. Par exemple, vous pouvez créer une configuration système pour votre système local et une configuration système différente pour votre système à distance. Selon votre environnement, les deux configurations système peuvent indiquer la même configuration des paramètres de substitution ou chaque configuration système peut nécessiter une configuration des paramètres de substitution différente.

Au cours de l'exécution d'un job, vous pouvez définir la configuration système et le job sera exécuté avec les valeurs de la configuration des paramètres de substitution associée.

Informations associées

[Qu'est-ce qu'une banque de données \[page 82\]](#)

23.11.4 Association de configurations des paramètres de substitution avec des configurations système

Comment associer une configuration des paramètres de substitution avec une configuration système.

Associer une configuration des paramètres de substitution avec une configuration système :

1. Dans Designer, ouvrez l'*Éditeur de configuration système* en sélectionnant **Outils** > *Configurations système*.

L'éditeur contient au moins deux colonnes : la première colonne contient des banques de données existantes et la colonne est nommée Paramètre de substitution. La seconde colonne et les suivantes sont nommées d'après la configuration système existante. Par exemple, si vous avez une configuration système, il existe une colonne nommée d'après le nom de votre configuration système.

2. Dans la ligne pour l'objet de banque de données applicable, sélectionnez la flèche déroulante dans la cellule sous la colonne de configuration système applicable.

Une liste de configurations des paramètres de substitution existants apparaît.

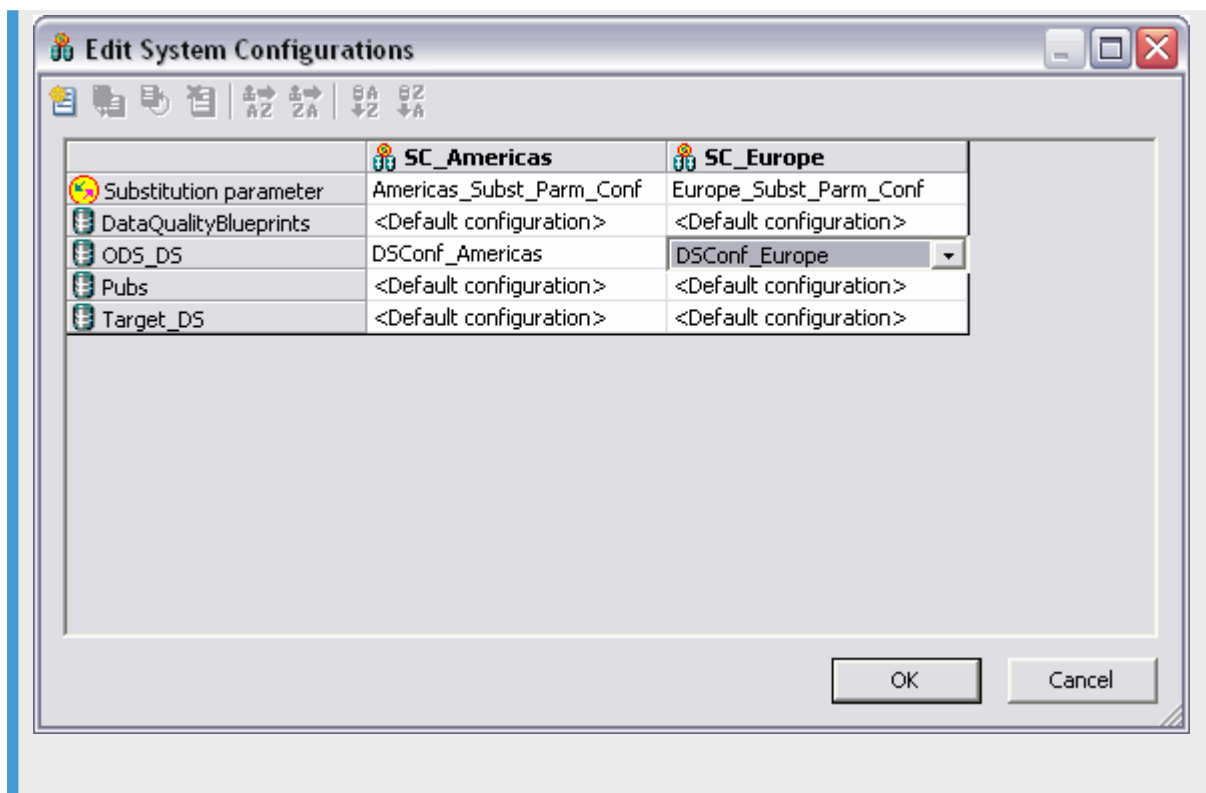
3. Sélectionnez une configuration des paramètres de substitution à partir de la liste déroulante.
4. Si vous souhaitez créer une configuration système, cliquez sur l'icône *Créer une configuration* dans la barre d'outils. Une nouvelle colonne apparaît. Saisissez le nom de la nouvelle configuration système dans l'entête de colonne.
5. Quand vous avez terminé, cliquez sur *OK*.

❖ Exemple

L'exemple suivant présente deux configurations système, **SC_Amerique** et **SC_Europe**. Il existe deux configurations des paramètres de substitution, une pour chaque région : **Americas_Subst_Parm_Conf** et **Europe_Subst_Parm_Conf**. Chaque configuration des paramètres de substitution définit où les fichiers de source de données sont localisés pour cette région. Par exemple (non affiché dans la capture d'écran ci-dessous) :

- Les fichiers de source de données pour **Americas_Subst_Parm_Conf** sont dans D:/Data/Americas.
- Les fichiers de source de données pour **Europe_Subst_Parm_Conf** sont dans D:/Data/Europe.

Sélectionnez la configuration des paramètres de substitution appropriée pour chaque banque de données et chaque configuration système applicable. Par exemple, la banque de données ODS_DS (ligne sous la première colonne, Paramètre de substitution) est associée à la configuration des paramètres de substitution DSConf_Americas et à la configuration système SC_Americas.



Informations associées

[Définition d'une configuration système \[page 146\]](#)

23.11.5 Remplacement d'un paramètre de substitution dans l'Administrateur

Dans l'Administrateur, vous pouvez remplacer les paramètres de substitution ou sélectionner une configuration système pour indiquer une configuration des paramètres de substitution sur quatre pages :

- Exécuter le job batch
- Planifier le job batch
- Commande de l'exécution de l'exportation
- Configuration du service en temps réel

Par exemple, la page Exécuter le job batch affiche le nom de la configuration système sélectionnée, la configuration des paramètres de substitution et le nom de chaque paramètre de substitution et sa valeur.

Pour remplacer un paramètre de substitution :

1. Sélectionnez la configuration système appropriée.
2. Sous *Paramètres de substitution*, cliquez sur *Ajouter un paramètre ignoré* qui affiche les paramètres de substitution disponibles.

3. Dans la liste déroulante, sélectionnez le paramètre de substitution à remplacer.
4. Dans la deuxième colonne, saisissez la valeur de remplacement. Saisissez la valeur comme une chaîne sans guillemets (contrairement aux variables globales).
5. Exécutez le job.

23.11.6 Exécution d'un job avec des paramètres de substitution

Pour voir les détails sur la façon dont les paramètres de substitution sont utilisés dans le job au cours de l'exécution dans le journal de suivi Designer :

1. Cliquez avec le bouton droit de la souris sur le nom de job et cliquez sur [Propriétés](#).
2. Cliquez sur l'onglet [Suivi](#).
3. Pour l'option [Effectuer le suivi des assembleurs](#), définissez la valeur sur [Oui](#).
4. Cliquez sur [OK](#).

Lorsque vous exécutez un job à partir de Designer, la fenêtre Propriétés d'exécution s'affiche. Vous disposez des options suivantes :

- Dans l'onglet [Options d'exécution](#) du menu déroulant [Configuration système](#), vous pouvez sélectionner la configuration système avec laquelle vous souhaitez exécuter le job. Si vous ne sélectionnez pas de configuration système, le logiciel applique la configuration des paramètres de substitution par défaut comme le définit l'éditeur de paramètres de substitution.
Vous pouvez cliquer sur [Parcourir](#) pour afficher la fenêtre [Sélectionner la configuration système](#) afin de voir la configuration des paramètres de substitution associée à chaque configuration système. La fenêtre [Sélectionner la configuration système](#) est en lecture seule. Pour modifier une configuration système, cliquez sur ► [Outils](#) ► [Configurations système](#) ►.
- Vous pouvez remplacer la valeur des paramètres de substitution spécifiques au moment de l'exécution. Cliquez sur l'onglet [Paramètres de substitution](#), sélectionnez un paramètre de substitution dans la colonne Nom et saisissez une valeur en cliquant deux fois sur la cellule Valeur.
Pour remplacer les valeurs du paramètre de substitution, lorsque vous lancez un job via un service Web, reportez-vous au [Guide d'intégration](#).

Informations associées

[Association de configurations des paramètres de substitution avec des configurations système \[page 336\]](#)

[Remplacement d'un paramètre de substitution dans l'Administrateur \[page 337\]](#)

23.11.7 Exportation et importation de paramètres de substitution

Les paramètres de substitution sont archivés dans un référentiel local avec leurs valeurs configurées. Le logiciel ne comprend pas de paramètres de substitution en temps qu'élément d'une exportation normale. Vous pouvez

cependant exporter des paramètres de substitution et des configurations vers d'autres référentiels en les exportant dans un fichier et en important le fichier dans un autre référentiel.

23.11.7.1 Exportation de paramètres de substitution

1. Cliquez avec le bouton droit de la souris sur la bibliothèque d'objets locale ► [référentiel](#) ► [Exporter les configurations des paramètres de substitution](#) .
2. Cochez la case de la colonne [Exportation](#) pour les configurations des paramètres de substitution à exporter.
3. [Sauvegardez](#) le fichier.
Le logiciel le sauvegarde comme un fichier texte avec une extension .atl.

23.11.7.2 Importation de paramètres de substitution

Les paramètres de substitution doivent d'abord avoir été exportés dans un fichier ATL.

Tenez compte des comportements suivants lors de l'importation de paramètres de substitution :

- Le logiciel ajoute tous les nouveaux paramètres de substitution et les configurations au référentiel local de destination.
 - Si le référentiel a un paramètre de substitution avec le même nom que celui du fichier exporté, l'importation écrasera la valeur du paramètre. De la même façon, si le référentiel a une configuration des paramètres de substitution avec le même nom que la configuration exportée, l'importation écrasera toutes les valeurs de paramètre pour cette configuration.
1. Dans Designer, cliquez avec le bouton droit sur la bibliothèque d'objets et sélectionnez ► [Référentiel](#) ► [Importer à partir du fichier](#) .
 2. Naviguez jusqu'au fichier à importer.
 3. Cliquez sur [OK](#).

Informations associées

[Exportation de paramètres de substitution \[page 339\]](#)

24 Jobs en temps réel

Le logiciel prend en charge une transformation de données en temps réel.

Temps réel signifie que le logiciel peut recevoir des demandes de systèmes ERP et d'applications Web et envoyer les réponses immédiatement après l'accès aux données demandées d'une mémoire cache de données ou d'une deuxième application. Vous définissez des opérations de traitement des messages à la demande en créant des jobs en temps réel dans le Designer.

24.1 Traitement d'un message de demande/réponse

Les messages transférés via un système en temps réel comprennent les informations requises pour exécuter une transaction commerciale.

Le contenu de ce message peut varier :

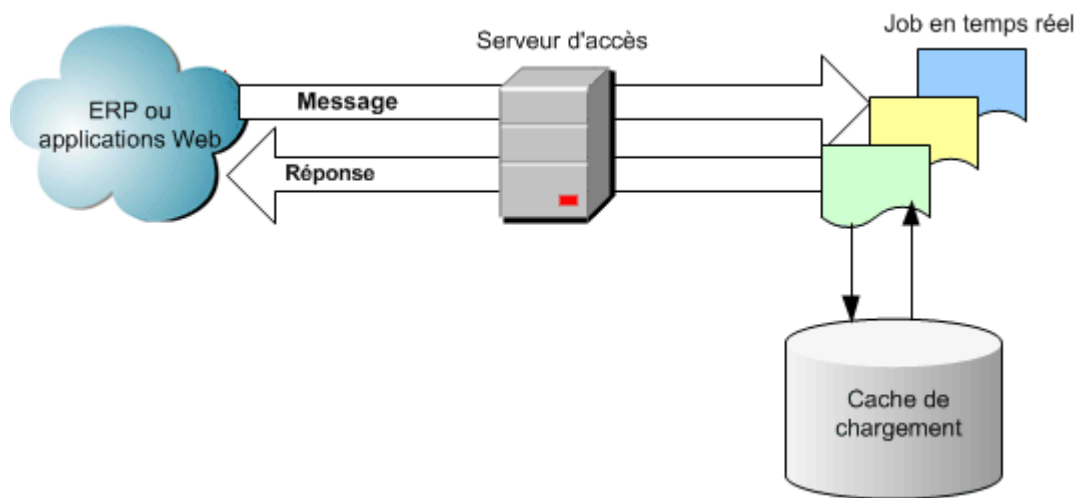
- Il pourrait être une commande client ou une facture traitée par un système ERP, destinée à un cache de données.
- Il pourrait être une demande de statut de commande créée par une application Web nécessitant une réponse de la part d'un cache de données ou d'un système back-office.

Le serveur d'accès écoute sans cesse les messages entrants. Lorsqu'il reçoit un message, le serveur d'accès dirige le message vers un processus d'attente qui effectue un ensemble prédéfini d'opérations pour le type de message. Le serveur d'accès reçoit alors une réponse pour le message et répond à l'application initiale.

Deux composants prennent en charge le traitement du message de demande-réponse :

Composant	Description
Serveur d'accès	Écoute les messages et dirige chaque message en fonction du type de message.
Job en temps réel	Effectue un ensemble prédéfini d'opérations pour ce type de message et crée une réponse.

Le traitement peut nécessiter l'ajout de données supplémentaire au message à partir d'un cache de données ou le chargement des données du message dans un cache de données. Le serveur d'accès renvoie la réponse à l'application initiale.



24.2 Qu'est-ce qu'un job en temps réel ?

Designer permet de définir le traitement de messages en temps réel à l'aide d'un job en temps réel.

Vous créez un job en temps réel différent pour chaque type de message que votre système peut créer.

24.2.1 Contenu d'un job en temps réel

Une boucle de traitement en temps réel peut contenir les objets suivants :

- Un flux de données unique pouvant contenir :
 - Une source en temps réel unique : message XML (requis)
 - Des sources : fichiers, fichiers XML et tables, y compris des tables SAP.
 - Une cible en temps réel unique : message XML (requis)
 - Des cibles : fichiers, fichiers XML, tables et modèles de tables
 - Des transformations, y compris des requêtes
- Plusieurs flux de données pouvant contenir :
 - Une source en temps réel unique dans le premier flux de données — Message XML (requis)
 - Des sources : fichiers, fichiers XML et tables, y compris des tables SAP.
 - Une source en temps réel unique dans le dernier flux de données — Message XML (requis)
 - Des cibles : fichiers, fichiers XML, tables et modèles de tables
 - Des transformations, y compris des requêtes
 - Des tables en mémoire, utilisées comme tables intermédiaires pour déplacer des jeux de données vers le flux de données suivant dans le job.
- Plusieurs workflows, scripts, expressions conditionnelles, boucles While, etc.

Il est également possible de créer des jobs en utilisant des IDocs. Pour en savoir plus, voir le document *Supplément pour SAP*.

i Remarque

Les jobs en temps réel ne peuvent pas utiliser les fonctionnalités suivantes : la transformation `Data_Transfer`, s'exécuter en tant que processus séparé, ou s'exécuter à un niveau de distribution de job inférieur à celui d'un niveau de job.

24.2.2 Temps réel par rapport à job

Les jobs en temps réel extraient des données du corps du message reçu et de toute source secondaire utilisée dans le job.

Tout comme un job batch, un job en temps réel extrait, transforme et charge des données. Chaque job en temps réel peut extraire des données d'un type de message individuel. Il peut également extraire des données d'autres sources, par exemple de tables ou de fichiers.

Les mêmes transformations que vous pouvez définir dans des batch jobs sont disponibles dans des jobs en temps réel. Cependant vous pouvez utiliser les transformations différemment dans des jobs en temps réel. Par exemple, vous pouvez utiliser des branches et des contrôles logiques plus souvent que vous ne le souhaiteriez dans les jobs batch. Si un client souhaite connaître le moment où il peut enlever sa commande au centre de distribution, vous serez amené à créer un job `CheckOrderStatus` à l'aide d'une fonction de recherche pour compter les postes de commande client puis une transformation `Case` pour indiquer le statut sous forme de chaînes : « Aucun poste n'est prêt pour l'enlèvement » ou « X poste de votre commande sont prêts pour l'enlèvement » ou « Votre commande est prête pour l'enlèvement ».

Dans les jobs en temps réel, le logiciel écrit également des données dans les cibles de message et les cibles secondaires en parallèle. Chaque message reçoit ainsi une réponse dès que possible.

Contrairement aux jobs batch, les jobs en temps réel ne sont pas exécutés en réponse à une planification ou un déclencheur interne ; les jobs en temps réel sont exécutés comme des services en temps réel lancés via l'Administrateur. Les services en temps réel attendent alors des messages du serveur d'accès. Lorsque le serveur d'accès reçoit un message, il le transfère à un service en temps réel en cours d'exécution conçu pour traiter ce type de message. Le service en temps réel traite le message et renvoie une réponse. Le service en temps réel continue à écouter et traiter les messages à la demande jusqu'à ce qu'il reçoive une instruction pour arrêter.

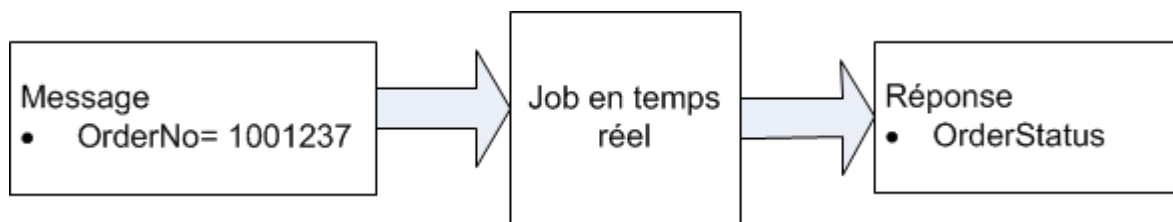
24.2.3 Messages

Les messages types comprennent des informations requises pour implémenter une transaction commerciale spécifique et pour donner une réponse appropriée.

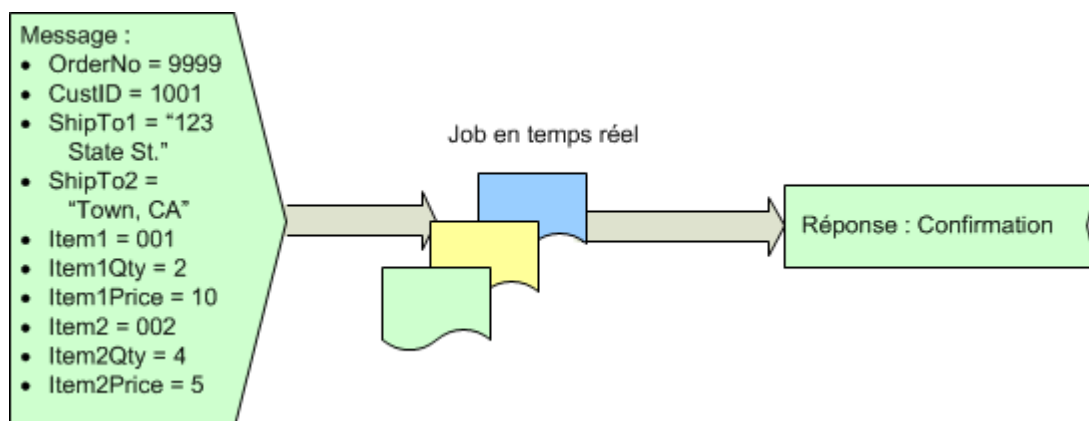
La façon dont vous créez un job en temps réel dépend du type de message que vous souhaitez qu'il traite.

Par exemple, supposez qu'un message contienne des informations requises pour déterminer le statut d'une commande spécifique. Le message peut juste contenir le numéro de la commande client. Le job en temps réel correspondant peut utiliser l'entrée pour interroger les sources correctes et renvoyer les informations produites appropriées.

Dans ce cas, le message contient des données qui peuvent être représentées comme colonne individuelle dans un tableau comportant une seule ligne.



Dans un deuxième cas, un message pourrait être une commande client à saisir dans un système ERP. Le message pourrait contenir le numéro de la commande, des informations client et les détails du poste de la commande. Le traitement des messages pourrait renvoyer une confirmation indiquant que la commande a été correctement soumise.



Dans ce cas, le message contient des données qui ne peuvent pas être représentées dans un seul tableau ; les informations de l'en-tête de la commande peuvent être représentées par un tableau et les postes de la commande par un deuxième tableau. Le logiciel représente les données d'en-tête et de poste du message dans une relation imbriquée.

Table de niveau supérieur

OrderNo	CustID	ShipTo1	ShipTo2	Postes
9999	1001	123 State St	Town, CA	

Table imbriquée

Item	ItemQty	ItemPrice
001	2	10
002	4	5

Au cours du traitement du message, le job en temps réel traite toutes les lignes de la table imbriquée pour chaque ligne de la table de niveau supérieur. Dans cette commande client, les deux postes sont traités pour la ligne individuelle des informations d'en-tête.

Les jobs en temps réel peuvent envoyer uniquement une ligne de données dans un message de réponse (cible de message). Cependant vous pouvez structurer des cibles de message de sorte que toutes les données figurent sur une seule ligne en imbriquant des tableaux dans des colonnes d'un seul tableau de niveau supérieur.

Les flux de données du logiciel prennent en charge l'imbrication de tableaux dans d'autres tableaux.

Informations associées

[Données imbriquées \[page 256\]](#)

24.2.4 Traitement des messages

Contrairement aux jobs batch, les jobs en temps réel sont conçus pour traiter plusieurs messages (pas uniquement des fichiers ou tables de données).

Pour les transformations qui requièrent toutes les données du message en même temps, telles que les requêtes incluant des fonctions d'agrégat, les données sont mises en cache temporairement. La transformation exécute l'opération spécifiée, puis vide les caches avant de recevoir le message suivant.

Pour tester un job en temps réel via Designer, la procédure recommandée consiste à tester un message et à créer un fichier test cible pour recevoir les données. Un job en temps réel supprimera les données après le traitement de chaque message si les paramètres par défaut du système sont utilisés. Vous devez donc désélectionner :

- [Supprimer les données de la table avant le chargement](#) pour une cible de table.
- [Supprimer le fichier](#) pour une cible de fichier plat
- [Supprimer et recréer le fichier](#) pour une cible XML

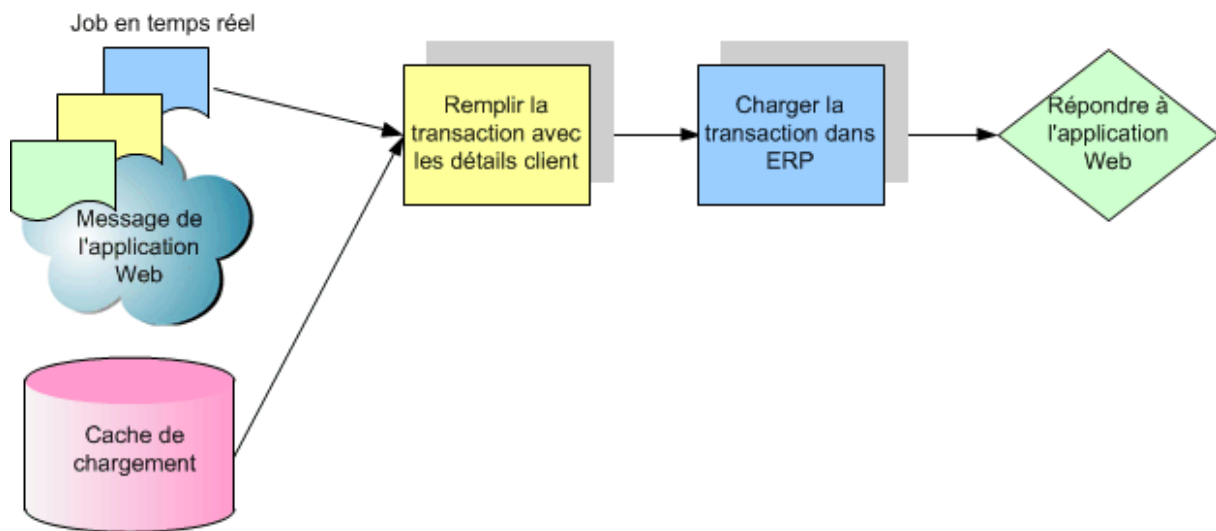
24.2.5 Exemples de jobs en temps réel

Exemple qui donne une description de niveau supérieur sur la façon dont des jobs en temps réel accèdent à des scénarios en temps réel classiques.

Les sections suivantes décrivent les objets réels que vous souhaitez utiliser pour créer la logique dans Designer.

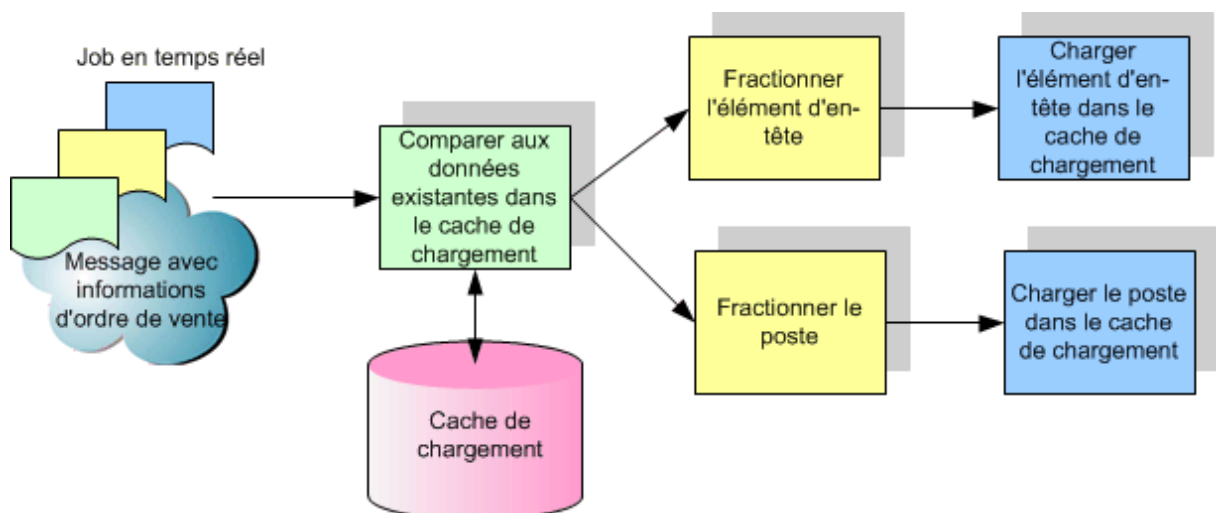
24.2.5.1 Chargement de transactions dans une application back-office

Un job en temps réel peut recevoir une transaction de la part d'une application Web et la charge dans une application back-office (ERP, SCM, système hérité). Une transformation Query vous permet d'inclure des valeurs d'un cache de données pour compléter la transaction avant de l'appliquer par rapport à l'application back-office (par exemple, un système ERP).



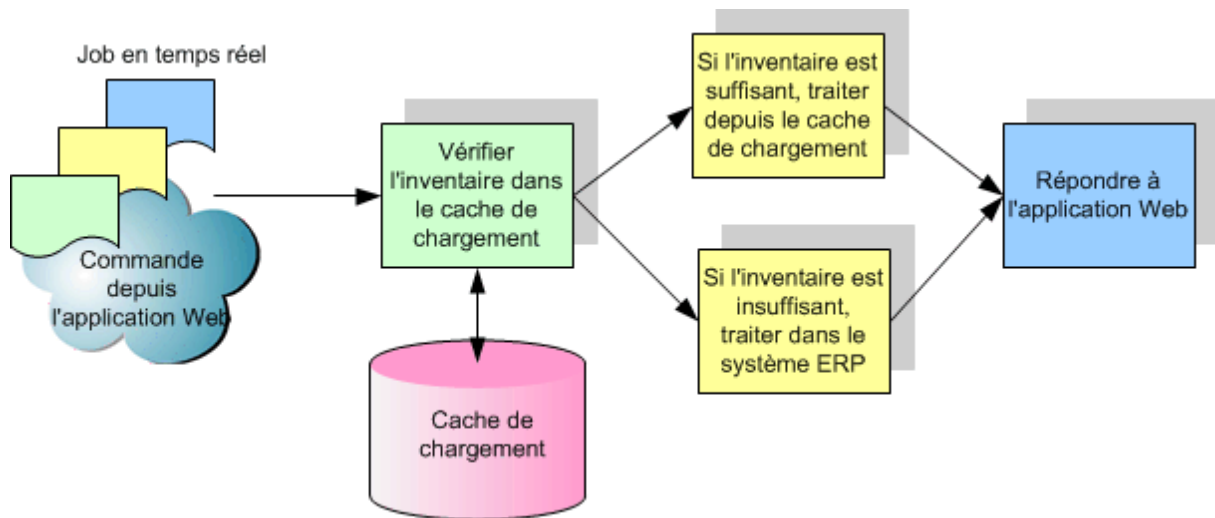
24.2.5.2 Collection de données back-office dans un cache de données

Vous pouvez utiliser des messages pour conserver le cache de données actuel. Les jobs en temps réel peuvent recevoir des messages d'une application back-office et les charger dans un cache de données ou un entrepôt de données.



24.2.5.3 Récupération de valeurs, de cache de données, applications back-office

Vous pouvez créer des jobs en temps réel qui utilisent des valeurs issues d'un cache de données pour déterminer si l'application back-office (par exemple, un système ERP) doit être interrogée directement.



24.3 Création de jobs en temps réel

Vous pouvez créer des jobs en temps réel en utilisant les mêmes objets comme jobs batch (flux de données, workflows, structures conditionnelles, boucles WHILE etc.).

Cependant, l'utilisation d'objet doit respecter un modèle de job en temps réel valide.

24.3.1 Réorganisation des métadonnées

Si vous utilisez un objet utilisé dans un job en temps réel à partir de la bibliothèque d'objets, les appels à l'objet sont remplacés par une icône indiquant que les appels ne sont plus valides.

Vous pouvez insérer n'importe quel nombre de workflows et de flux de données dans un job en temps réel tant que les modèles de flux de données associés à une boucle de traitement en temps réel sont suivis.

- Un flux de données unique dans une boucle : doit contenir la cible et la source du message.
- Plusieurs flux de données dans une boucle : doivent contenir une source du message dans le premier flux de données et une cible du message dans le dernier flux de données.

Au niveau du job, les workflows et les flux de données ne peuvent pas être désignés pour s'exécuter en parallèle. Cela est par contre possible dans un flux de données de niveau job.

Les messages d'un flux de données sont importants. Le tableau suivant indique comment les flux de données peuvent être utilisés.

Si un flux de données possède :	Il peut être utilisé comme :
Une source de message XML	Premier flux de données dans une boucle de traitement en temps réel.
Une cible de message XML	Dernier flux de données dans une boucle de traitement en temps réel.
Une source de message XML et une cible de message XML.	Uniquement un flux de données dans une boucle de traitement en temps réel.
Aucune source ou cible de message.	Ni le premier ni le dernier flux de données, ou uniquement un flux de données dans une boucle de traitement en temps réel. Flux de données dans un job batch.

Les flux de données qui ne contiennent aucun message peuvent être utilisés dans des jobs batch ou des boucles de traitement en temps réel (entre le premier et le dernier flux de données).

24.3.2 Modèles de job en temps réel

À la différence des jobs batch qui déplacent en général de grandes quantités de données à des moments planifiés, un job en temps réel, une fois démarré en tant que service en temps réel, est à l'écoute d'une demande éventuelle.

Lorsqu'un job en temps réel reçoit une demande (en général pour accéder à un nombre réduit d'enregistrements), le logiciel traite la requête, renvoie une réponse et continue à écouter. Cette logique d'écoute-traitement-écoute forme une boucle de traitement.

Un job en temps réel est constitué de trois composants de traitement : l'initialisation, une boucle de traitement en temps réel et le nettoyage.

- Le composant d'initialisation (facultatif) peut être un script, un workflow, un flux de données ou une combinaison d'objets. Il ne s'exécute qu'au démarrage d'un service en temps réel.
- La boucle de traitement en temps réel est un conteneur de logique de traitement unique du job. Vous pouvez y spécifier un nombre quelconque de workflows et de flux de données.
- Le composant de nettoyage (facultatif) peut être un script, un workflow, un flux de données ou une combinaison d'objets. Il ne s'exécute qu'à l'arrêt d'un service en temps réel.

Dans une boucle de traitement en temps réel, une source de message unique doit être incluse à la première étape et une cible de message unique doit être incluse à la dernière étape.

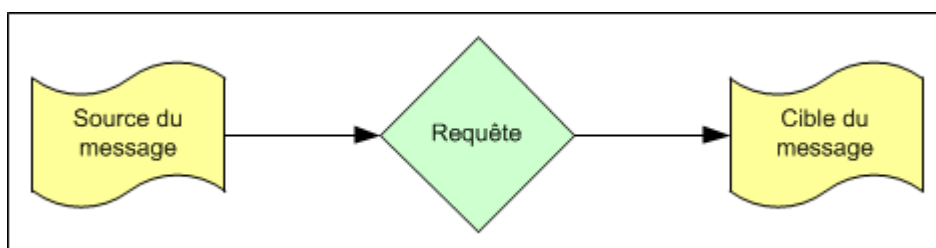
Les modèles suivants prennent en charge cette règle :

- Modèle de flux de données simple
- Modèle de flux de données multiples
- Modèle de flux de données Demande/Confirmation

Informations associées

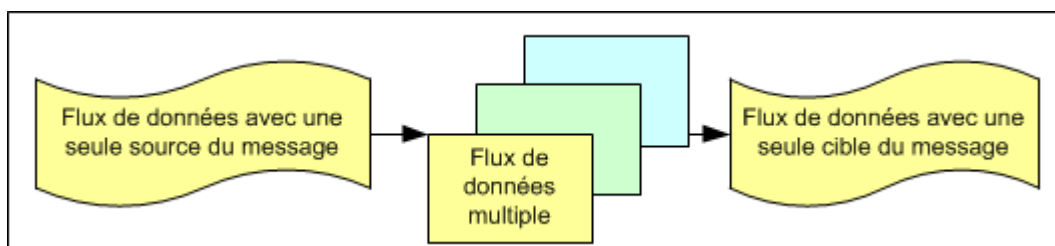
24.3.2.1 Modèle de flux de données simple

Le modèle de flux de données simple permet de créer un job en temps réel à l'aide d'un flux de données simple dans sa boucle de traitement en temps réel. Ce flux de données simple doit comprendre une source de message simple et une cible de message simple.



24.3.2.2 Modèle de flux de données multiples

Le modèle de flux de données multiples permet de créer un job en temps réel à l'aide de flux de données multiples dans sa boucle de traitement en temps réel.



L'utilisation de flux de données multiples, permet de garantir que les données de chaque message sont entièrement traitées dans un flux de données initial avant le début du traitement des flux de données suivants. Par exemple, si les données représentent 40 éléments, l'ensemble des 40 éléments doit être transféré par le premier flux de données vers une table de stockage ou une table en mémoire avant le transfert vers un deuxième flux de données. Vous pouvez ainsi contrôler et collecter toutes les données dans un message à tout point d'un job en temps réel pour la conception et la correction des erreurs.

Si vous utilisez des flux de données multiples dans une boucle en temps réel :

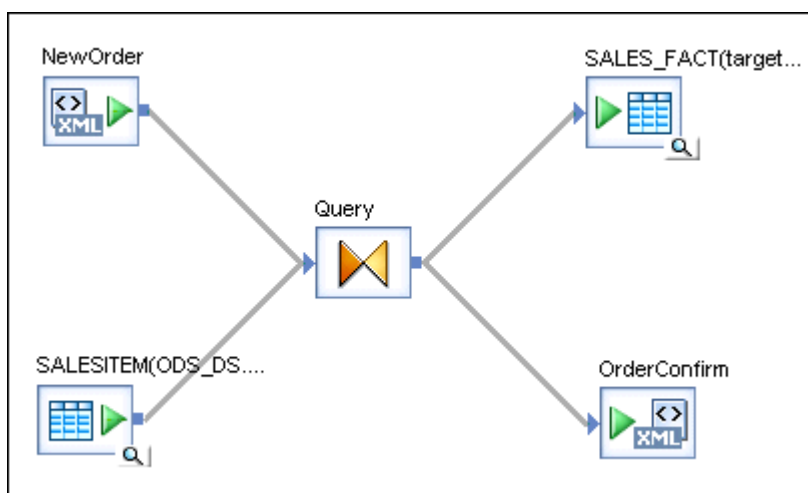
- Le premier objet de la boucle doit être un flux de données. Ce flux de données doit avoir une seule source de message.
- Le dernier objet de la boucle doit être un flux de données. Ce flux de données doit avoir une cible de message.
- Des flux de données supplémentaires ne peuvent pas avoir des sources ou des cibles de message.
- Vous pouvez ajouter n'importe quel nombre de flux de données supplémentaires à la boucle et vous pouvez les ajouter à n'importe quel nombre de workflows.

- Tous les flux de données peuvent utiliser des tables en mémoire entrantes et/ou sortantes pour transférer des jeux de données au flux de données suivant. Les tables en mémoire stockent les données dans la mémoire parallèlement à l'exécution d'une boucle. Elles améliorent la performance des jobs en temps réel avec des flux de données multiples.

24.3.3 Using real-time job models

24.3.3.1 Modèle de flux de données simple

Lorsque vous utilisez un flux de données simple dans une boucle de traitement en temps réel, l'apparence de votre diagramme de flux de données peut être la suivante :



Notez que le flux de données a une source de message et une cible de message.

24.3.3.2 Modèle de flux de données multiples

Lorsque vous utilisez des flux de données multiples dans une boucle de traitement en temps réel, vos diagrammes de flux de données peuvent ressembler à ceux de l'exemple de scénario suivant dans lequel Data Services écrit des données sur plusieurs cibles selon la conception de vos flux de données multiples.

Exigences de l'exemple de scénario :

Votre job doit effectuer les tâches suivantes, en terminant une tâche avant de passer à la suivante :

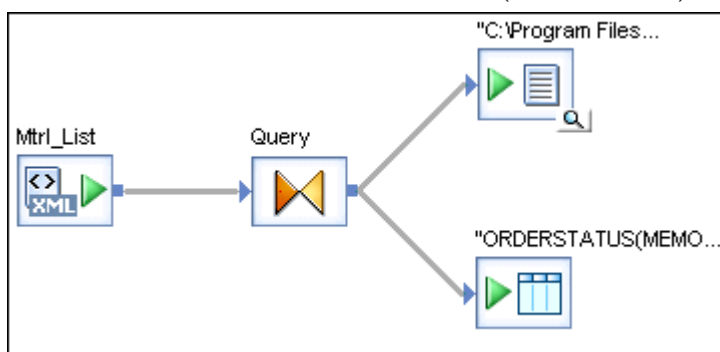
- recevoir des demandes sur le statut d'ordres individuels de la part d'un portail Web et enregistrer chaque message sur un fichier plat de sauvegarde ;

- exécuter une jointure de demande pour rechercher le statut de l'ordre et l'écrire dans une table de base de données client ;
- répondre à chaque message avec les résultats de la jointure de demande.

Solution :

Créer d'abord un job en temps réel et ajouter un flux de données, un workflow et un autre flux de données à la boucle en temps réel. Ajouter ensuite un flux de données au workflow. Puis configurer les tâches dans chaque flux de données :

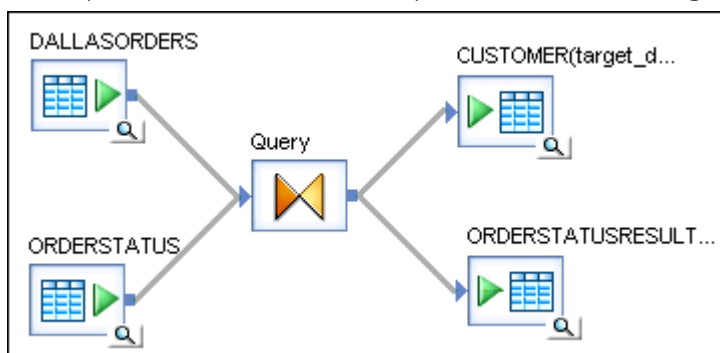
- Le premier flux de données reçoit le message XML (à l'aide d'une source de message XML) et enregistre le message dans le fichier plat (cible de format de fichier plat). Pendant ce temps, ce même flux de données écrit les données dans une table en mémoire (cible de la table).



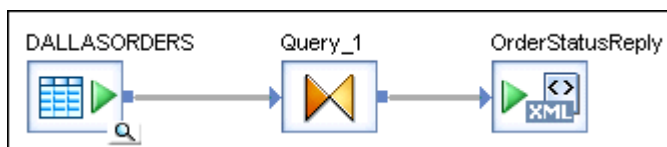
i Remarque

Vous voulez peut-être créer une table en mémoire pour déplacer les données vers des flux de données séquentiels.

- Le deuxième flux de données lit les données de message de la table en mémoire (source de la table), crée une jointure avec les données archivées (source de la table) et écrit les résultats dans une table de base de données (cible de la table) et une nouvelle table en mémoire (cible de la table).
Notez que ce flux de données ne comporte ni source de message ni cible de message.



- Le dernier flux de données envoie la réponse. Il lit le résultat de la jointure dans la table en mémoire (source de la table) et charge la réponse (cible du message XML).



Informations associées

[Conception d'applications en temps réel \[page 361\]](#)

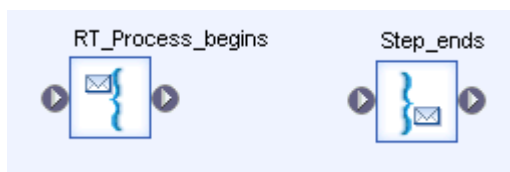
[Banques de données en mémoire \[page 108\]](#)

24.3.4 Créer un job en temps réel avec un flux de données unique

Cette rubrique décrit les étapes permettant de créer un job à l'aide d'un flux de données unique.

1. Dans Designer, créez un projet, ou ouvrez un projet existant.
2. Cliquez avec le bouton droit sur une zone vide de la zone du projet et sélectionnez *Nouveau job en temps réel* dans le menu déroulant.

Un nouvel icône de job apparaît dans la zone de projet nommée `New_RTJob1`. L'espace de travail affiche la structure du projet, qui se compose d'un marqueur de début et de fin.



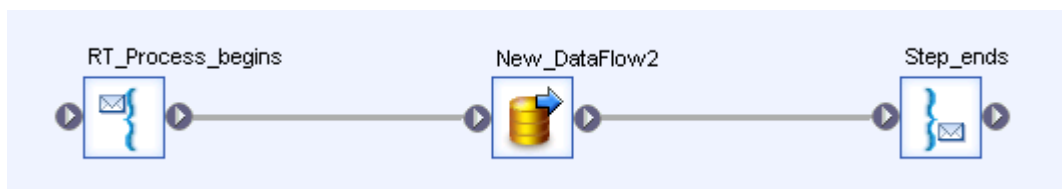
Ces marqueurs représentent le début et la fin d'une boucle en temps réel.

3. Renommez `New_RTJob1`.

Il est plus facile d'identifier vos objets en temps réel en ajoutant un préfixe au nom de l'objet. Par exemple, lorsque vous nommez ce job, le nom doit commencer par le préfixe `RT_Job_<Projectname>`.

Le logiciel répertorie tous les jobs en temps réel enregistrés sous l'onglet Job de la bibliothèque d'objets locale, par ordre alphabétique. Par conséquent, tous les jobs en temps réel sont regroupés. Toutefois, les noms de job peuvent également s'afficher dans les éditeurs de texte utilisés pour créer des appels d'adaptateur ou de services Web. Dans ces cas, un préfixe sauvegardé avec le nom du job vous aidera à l'identifier.

4. Ajouter un flux de données à l'espace de travail de job en temps réel à l'aide de la palette d'outils.
5. Reliez les objets du job dans l'espace de travail au flux de données comme objet intermédiaire.



6. Créez le flux de données en incluant une source et une cible.
7. Ajoutez, configurez et connectez un (des) objet(s) d'initialisation et de nettoyage le cas échéant.

Informations associées

[Création et définition des flux de données \[page 204\]](#)

[Objets source et cible en temps réel \[page 352\]](#)

24.4 Objets source et cible en temps réel

Les jobs en temps réel doivent contenir un objet source et/ou cible en temps réel.

Normalement les objets disponibles sont les suivants :

Objet	Description	Utilisé en tant que :	Accès au logiciel
Message XML	Message XML structuré au format de schéma DTD ou XML	Source ou cible	Directement ou via des adaptateurs
Message sortant	Message en temps réel avec un format spécifique à l'application (non lisible par un analyseur XML)	Cible	Via un adaptateur

Vous pouvez également utiliser des messages IDoc comme sources en temps réel pour des applications SAP. Pour plus d'informations, voir *Supplément pour SAP*.

La procédure d'ajout de sources et de cibles à des jobs en temps réel est identique à celle d'ajout de sources et de cibles à des jobs batch avec les éléments supplémentaires suivants :

Pour	Connaissances requises	Emplacement de la bibliothèque d'objets
Messages XML	Importer un schéma DTD ou XML pour définir un format	Onglet Formats
Message sortant	Définir une banque de données d'adaptateur et importer des métadonnées d'objet.	Onglet Banques de données sous banque de données d'adaptateur

Informations associées

[Importer un format de schéma XML ou DTD \[page 277\]](#)

24.4.1 Pour afficher un schéma source ou cible de message XML

Dans l'espace de travail d'un job en temps réel, cliquez sur le nom d'une source ou d'une cible de message XML pour ouvrir son éditeur.

Si la source ou la cible de message XML contient des données imbriquées, le schéma affiche des tables imbriquées pour représenter les relations entre les données.

24.4.2 Chargement des tables cible comme transaction unique

Dans un job en temps réel, il est possible de charger plusieurs tables à partir d'une banque de données unique dans une transaction simple. Lorsque le chargement transactionnel est activé pour les tables cible, SAP Data Services envoie des instructions `INSERT` pour toute table incluse dans la transaction vers la base de données à traiter. Il est également possible de définir les premières tables à charger en spécifiant l'ordre de transaction des tables.

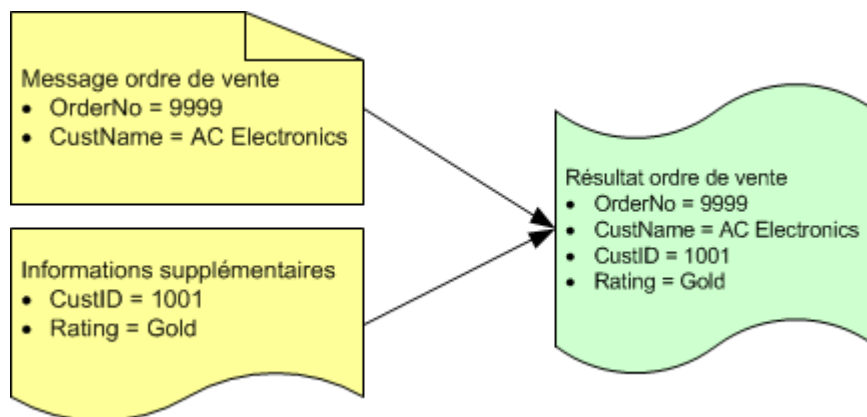
Si le flux de données inclut une cible en temps réel, il est toujours chargé parallèlement aux autres cibles pour que la durée de chargement soit la plus courte possible.

24.4.3 Sources et cibles secondaires

Les jobs en temps réel peuvent également avoir des sources ou des cibles secondaires.

Par exemple, supposez que vous traitez un message contenant une commande client à partir d'une application Web. La commande contient le nom du client mais lorsque vous appliquez la commande par rapport à votre système ERP, vous devez fournir des informations client plus détaillées.

Dans un flux de données d'un job en temps réel, vous pouvez compléter le message par les informations client pour créer le document complet à envoyer au système ERP. Les informations supplémentaires peuvent provenir du système ERP lui-même ou d'un cache de données contenant les mêmes informations.



Les tables et les fichiers (notamment fichiers XML) en tant que sources peuvent fournir ces informations supplémentaires.

Le logiciel lit des données à partir de sources secondaires conformément à la façon dont vous configurez le flux de données. Le logiciel charge des données dans les cibles secondaires parallèlement à un message cible.

Ajoutez des sources et des cibles secondaires aux flux de données des jobs en temps réel de la même façon qu'aux flux de données des jobs batch.

Informations associées

[Objets sources et cibles \[page 206\]](#)

[Ajout d'objets sources ou cibles aux flux de données \[page 211\]](#)

24.4.4 Chargement opérationnel de tables

Les tables cible dans des jobs en temps réel prennent en charge un chargement opérationnel dans lequel les données résultant du traitement d'un flux de données simple peuvent être chargées dans plusieurs tables en tant que transaction simple.

Aucun élément de la transaction ne s'applique si un élément échoue.

i Remarque

Les tables cible dans des jobs batch prennent également en charge un chargement opérationnel. Cependant soyez prudent lorsque vous activez cette option pour un job batch car elle nécessite l'utilisation d'une mémoire, ce qui peut réduire la performance lorsque vous déplacez d'importantes quantités de données.

Vous pouvez indiquer l'ordre dans lequel les tables sont incluses dans la transaction à l'aide de l'éditeur de table cible. Cette fonctionnalité prend en charge un scénario dans lequel vous avez un groupe de tables avec des clés externes qui dépendent d'un groupe contenant des clés primaires.

Vous pouvez utiliser le chargement opérationnel uniquement lorsque toutes les cibles dans un flux de données sont dans la même banque de données. Si le flux de données charge des tables dans plusieurs banques de données, les cibles de chaque banque de données sont chargées indépendamment. Alors que plusieurs cibles d'une banque de données peuvent être incluses dans une transaction, les cibles d'autres banques de données doivent être incluses dans une autre transaction.

Vous pouvez indiquer le même ordre de transaction ou des ordres de transaction distincts pour toutes les cibles à inclure dans la même transaction. Si vous indiquez le même ordre de transaction pour toutes les cibles dans la même banque de données, les tables sont toujours incluses dans la même transaction mais sont chargées ensemble. Le chargement est validé une fois que le chargement de toutes les tables est terminé dans la transaction.

Si vous indiquez des ordres de transaction distincts pour toutes les cibles dans la même banque de données, les ordres de transaction indiquent les ordres de chargement des tables. La table avec l'ordre de transaction le plus petit est chargée en premier etc., jusqu'à ce que la table avec l'ordre de transaction le plus grand soit

chargée en dernier. Deux tables ne sont pas chargées simultanément. Le chargement est validé lorsque le chargement de la dernière table est terminé.

24.4.5 Astuces relatives à la conception pour flux de données dans des jobs en temps réel

Lorsque vous concevez des flux de données en temps réel, vous devez suivre certaines règles.

- Si vous incluez une table à une jointure avec une source en temps réel, le logiciel inclut le jeu de données de la source en temps réel comme boucle externe de la jointure. Si plusieurs sources supplémentaires sont incluses dans la jointure, vous pouvez contrôler quelle table est incluse à la boucle suivante la plus externe de la jointure à l'aide des classements de jointure des tables.
- Dans des jobs en temps réel, ne mettez pas en mémoire cache les données de sources secondaires sauf si ces données sont statiques. Les données seront lues au début du job en temps réel et ne seront pas mises à jour lors de l'exécution du job.
- Si aucune ligne n'est transférée à la cible XML, le job en temps réel renvoie une réponse vide au serveur d'accès. Par exemple, si le système reçoit une demande pour un numéro de produit qui n'existe pas, votre job peut être conçu de telle sorte qu'aucune donnée ne soit transférée au message de réponse. Vous pouvez être amené à fournir des instructions appropriées à votre utilisateur (traitement d'exception dans votre job) dans le cadre de ce type de scénario.
- Si plusieurs lignes sont transférées à la cible XML, la cible lit la première ligne et rejette les autres lignes. Afin d'éviter ce problème, utilisez vos connaissances relatives au Modèle de données relationnel imbriqué et structurez vos formats de message source et cible de telle sorte qu'une "ligne" corresponde à un message. Le modèle de données relationnel imbriqué permet de structurer toute quantité de données sur une seule "ligne" car les colonnes des tableaux peuvent contenir d'autres tableaux.
- Les mécanismes de récupération ne sont pas pris en charge dans les jobs en temps réel.

Informations associées

[Données imbriquées \[page 256\]](#)

24.5 Testing real-time jobs

24.5.1 Exécution d'un job en temps réel en mode test

Vous pouvez tester des conceptions de job en temps réel sans configurer le job comme service associé à un serveur d'accès.

En mode test, vous pouvez exécuter un job en temps réel à l'aide d'un exemple de message source à partir d'un fichier pour déterminer si le logiciel crée le message cible escompté.

24.5.1.1 Spécifier un exemple de message XML et un fichier test cible

1. Dans les éditeurs de cible et de source de message XML, entrez un nom de fichier dans la case [Fichier test XML](#).

Entrez le nom de chemin complet pour le fichier source qui contient votre échantillon de données. Utilisez les chemins pour les fichiers test relatifs à l'ordinateur qui exécute le job server pour le référentiel actuel.

2. Exécutez le job.

Le mode test est toujours activé pour des jobs en temps réel. Le logiciel lit des données à partir du fichier test source et les charge dans le fichier test cible.

24.5.2 Utilisation de la fonction d'affichage des données

Utiliser la fonctionnalité d'affichage de données pour capturer un échantillon de données de sortie pour garantir que la conception fonctionne.

Ce processus permet de s'assurer que la transformation conçue renvoie les résultats escomptés.

Informations associées

[Conception et débogage \[page 722\]](#)

24.5.3 Utilisation d'une cible de fichier XML

Utilisez une cible de fichier XML pour capturer le message créé par un flux de données en autorisant le renvoi du message au serveur d'accès.

Tout comme un message XML, vous définissez un fichier XML en important un schéma DTD ou XML pour le fichier puis faites glisser le format dans la définition de flux de données. Contrairement aux messages XML, vous pouvez inclure des fichiers XML comme sources ou cibles dans des jobs batch et des jobs en temps réel.

24.5.3.1 Utiliser un fichier pour une capture de sortie à partir d'étapes de job en temps réel

1. Dans l'onglet [Formats](#) de la bibliothèque d'objets, faites glisser le schéma DTD ou XML dans un flux de données d'un job en temps réel.

Un menu vous invite à indiquer la fonction du fichier.

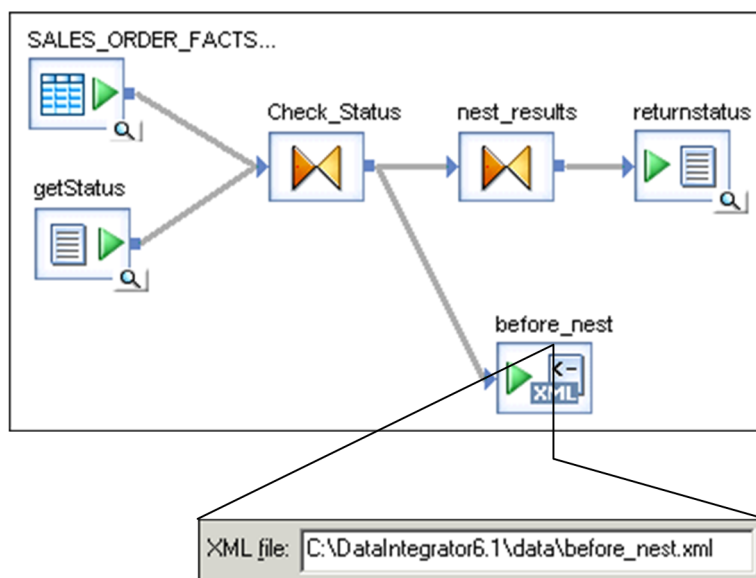
2. Sélectionnez *Créer la cible du fichier XML*.

La cible du fichier XML s'affiche dans l'espace de travail.

3. Dans l'éditeur de fichiers, indiquez l'emplacement où le logiciel écrit les données.

Saisissez un nom de fichier par rapport à l'ordinateur exécutant le job server.

4. Reliez la sortie de l'étape du flux de données de capture à l'entrée du fichier.



24.6 Building blocks for real-time jobs

24.6.1 Complément aux données de message

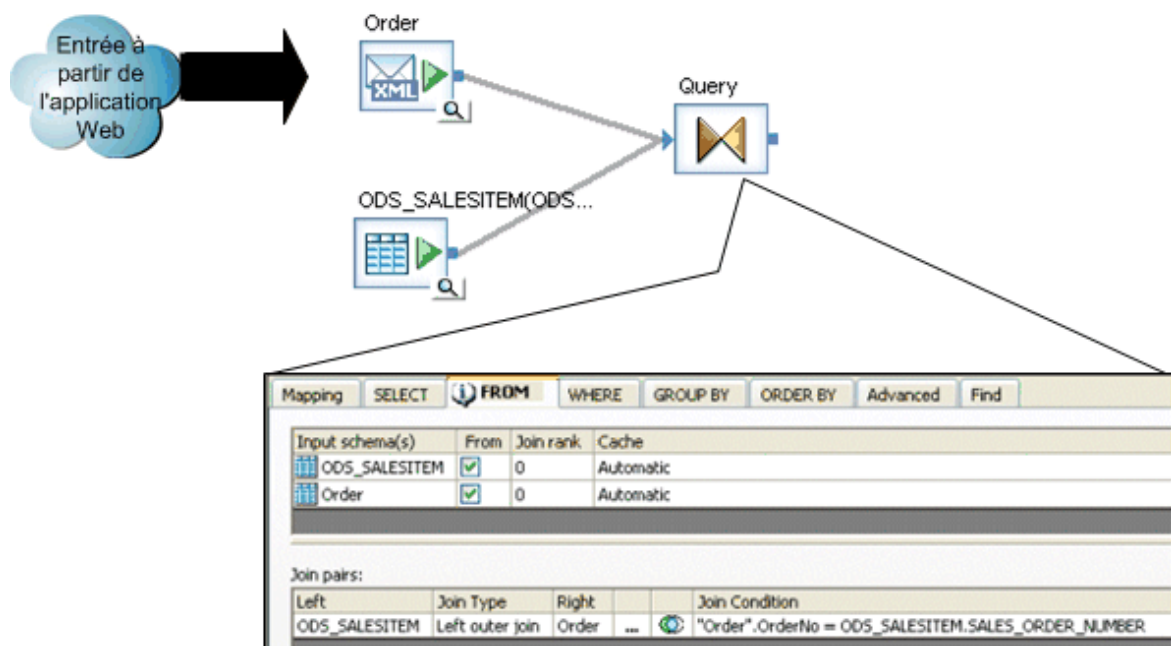
Vous pouvez définir les étapes dans le job en temps réel pour compléter les informations du message si les données incluses dans des messages de sources en temps réel ne correspondent pas exactement à vos besoins pour le traitement ou l'archivage des informations.

Une technique permettant de compléter les données dans une source en temps réel comprend les étapes suivantes :

1. Inclure une table ou un fichier en tant que source.

Outre la source en temps réel, inclure les fichiers ou les tables pour lesquels vous avez besoin d'informations complémentaires.
2. Utiliser une demande pour extraire les données requises de la table ou du fichier.
3. Utiliser les données de la source en temps réel pour rechercher les données supplémentaires nécessaires.

Vous pouvez inclure une expression JOIN dans la demande pour extraire les valeurs spécifiques requises de la source supplémentaire.



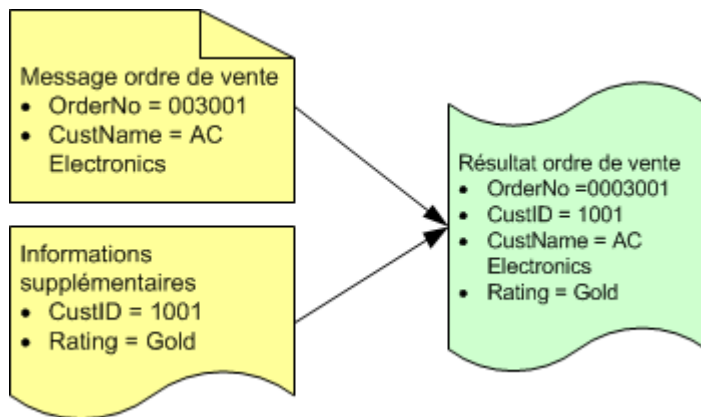
La condition de jointure relie les deux schémas d'entrée, ce qui se traduit par une sortie uniquement pour le document de poste de vente et les postes de ligne inclus dans la saisie à partir de l'application.

Utilisez prudemment les données dans la jointure pour garantir un retour de valeur. Si la jointure ne renvoie aucune donnée, la demande ne crée aucune ligne et le message retourne sans valeur au serveur d'accès. Si vous ne pouvez pas garantir le retour d'une valeur, tenez compte des alternatives suivantes :

- [Appel de la fonction de recherche](#) – renvoie une valeur par défaut si le système ne trouve aucune correspondance.
- [Jointure externe](#) – renvoie toujours une valeur, même si le système ne trouve aucune correspondance.

24.6.1.1 Complément à l'exemple de données de message

Dans cet exemple, un message de demande comprend des informations sur la commande client et son message de réponse renvoie le statut de la commande. La logique applicative utilise le numéro de client et la classification de priorité pour déterminer le niveau du statut à renvoyer. Le message contient uniquement le nom du client et le numéro de la commande. Un job en temps réel est alors défini pour récupérer le numéro de client et la classification d'autres sources avant de déterminer le statut de la commande.



1. Incluez la source en temps réel au job en temps réel.
2. Incluez la source complémentaire au job en temps réel.

Cette source pourrait être un tableau ou un fichier. Dans cet exemple, les informations complémentaires requises ne changent pas très souvent, il est donc judicieux d'extraire les données d'un cache de données plutôt que d'accéder directement à un système ERP.

3. Relier les sources.

Dans une transformation Query, créez une jointure sur le nom du client :

```
Message.CustName = Cust_Status.CustName
```

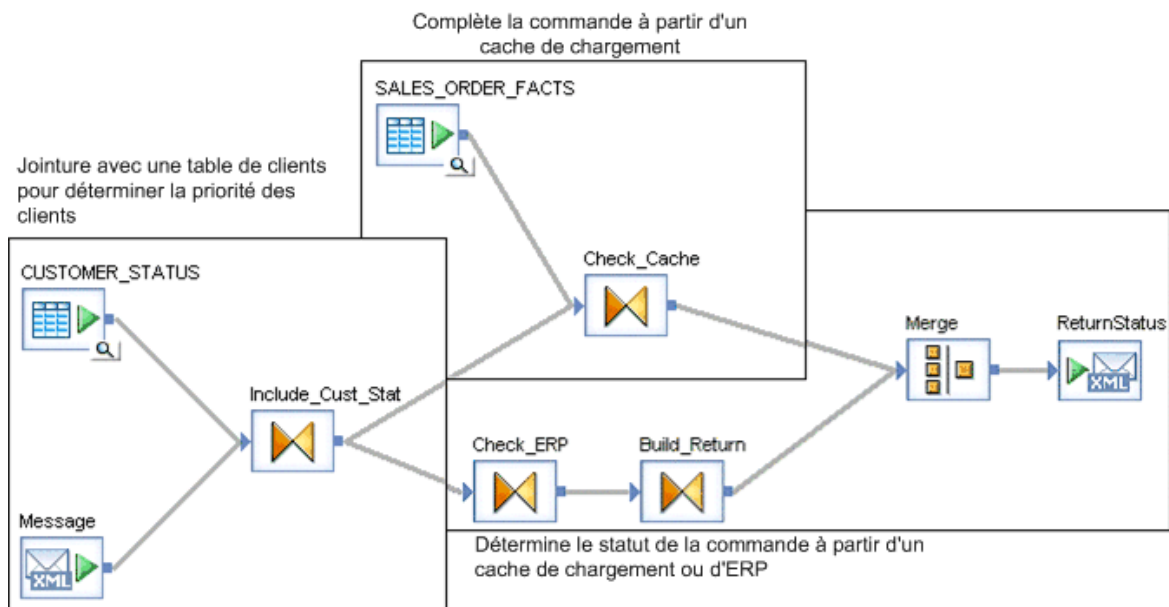
Vous pouvez créer la sortie pour inclure uniquement les colonnes requises par le job en temps réel pour déterminer le statut de la commande.

4. Complétez le job en temps réel pour déterminer le statut de la commande.

L'exemple indiqué ici détermine le statut de la commande dans l'une des deux méthodes, en fonction de la valeur du statut client. Le statut de la commande des clients les mieux classés est déterminé directement de l'ERP. Le statut de la commande pour les autres clients est déterminé à partir d'un cache de données d'informations relatives à la commande client.

La logique peut être organisée en flux de données simple ou multiples. L'illustration ci-dessous présente un modèle de flux de données simple.

Les deux branches renvoient le statut de commande pour chaque ligne de poste de la commande. Le flux de données fusionne les résultats et crée la réponse. La section suivante décrit comment concevoir des chemins de branche dans un flux de données.



24.6.2 Création de branche de flux de données en fonction d'une valeur de cache de données

L'une des choses les plus importantes que vous pouvez faire avec un job en temps réel est de concevoir une logique qui détermine si des réponses doivent être générées à partir d'un cache de données ou de données d'une application back-office (ERP, SCM, CRM).

Voici une technique permettant de créer cette logique :

1. Déterminez la règle d'accès au cache de données et à l'application back-office.
2. Comparez les données de la source en temps réel à la règle.
3. Définissez chaque chemin pouvant résulter de la sortie.

Vous pouvez être amené à prendre en compte le cas où la règle indique l'accès à l'application back-office mais le système n'est pas disponible actuellement.

4. Fusionnez les résultats de chaque chemin dans un jeu de données simple.
5. Transférez le résultat unique vers la cible en temps réel.

Vous pouvez être amené à effectuer un contrôle des erreurs et un traitement des exceptions pour être sûr qu'une valeur est transférée à la cible. Si la cible reçoit un jeu vide, le job en temps réel renvoie une réponse vide (début et fin de balise XML uniquement) au serveur d'accès.

Cet exemple décrit une section d'un job en temps réel qui traite une nouvelle commande client. La section assure le contrôle du stock disponible des produits commandés, elle répond à la question "le stock magasin est-il suffisant pour satisfaire cette commande ?"

La règle contrôlant l'accès à l'application back-office indique que le stock (Stck.) doit être supérieur à une valeur prédéfinie (IMargin), supérieure à la quantité commandée (Qté) pour considérer la valeur en stock des données mises en cache comme acceptable.

Le logiciel compare chaque poste de la commande dans l'ordre de mappage.

Commande d'achat entrante

N° commande	IDClient	Poste		
		Elément	Article	Qté
9999	1001	001	7333	300
		002	2288	1400

Cache de données de stock

Article	Stock	IMargin
7333	600	100
2288	1500	200

i Remarque

La quantité d'articles dans la commande client est comparée aux valeurs en stock du cache de données.

24.6.3 Appel de fonctions d'application

Un job en temps réel peut utiliser des fonctions d'application pour traiter des données.

Vous pouvez inclure des tables comme paramètres d'entrée ou de sortie à la fonction.

Les fonctions d'application nécessitent des valeurs de saisie pour certains paramètres et certains peuvent ne pas être renseignés. Vous devez déterminer les exigences de la fonction pour préparer les entrées appropriées.

Pour créer l'entrée, vous pouvez indiquer la table de niveau supérieur, les colonnes de niveau supérieur et toutes les tables imbriquées un niveau au-dessous des tables listées dans la clause FROM du contexte appelant la fonction. Si la fonction d'application comprend une structure comme paramètre d'entrée, vous devez indiquer les colonnes individuelles composant la structure.

Un flux de données peut contenir plusieurs étapes qui appellent une fonction, récupèrent des résultats puis les placent dans les colonnes et les tables requises pour une réponse.

24.7 Conception d'applications en temps réel

Vous pouvez créer des applications en temps réel qui répondent à vos besoins en matière de ressource et d'informations internes et externes.

Le logiciel propose une connexion à faible impact fiable entre une application Web et des applications back-office, par exemple un système ERP. Chaque implémentation d'un système ERP étant différente et le logiciel incluant une logique d'aide à la décision polyvalente, vous pouvez décider de concevoir un système.

24.7.1 Baisse des requêtes nécessitant un accès à une application back-office

Ensemble de recommandations et de considérations permettant de réduire le temps passé à tester des cycles de développement.

Les informations auxquelles vos clients sont autorisés à accéder via votre application Web peuvent influencer la performance que les clients voient sur le Web. Vous pouvez améliorer la performance par vos décisions en matière de conception d'application Web. Notamment, vous pouvez structurer votre application pour réduire le nombre de requêtes nécessitant un accès direct à l'application back-office (ERP, SCM, système hérité).

Par exemple, si votre système ERP prend en charge une structure de détermination du prix complexe qui comprend des dépendances, par exemple la priorité du client, la disponibilité du produit ou la quantité commandée, vous ne pourrez peut-être pas compter sur les valeurs d'un cache de données pour les informations prix. L'autre solution serait de demander les informations prix directement à partir du système ERP. L'accès au système ERP semble plus lent que l'accès direct à la base de données, réduisant la performance des expériences de vos clients avec l'application Web.

Pour réduire l'impact des requêtes nécessitant un accès direct au système ERP, modifiez votre application Web. Via l'exemple de la détermination du prix, personnalisez l'application pour éviter l'affichage des informations prix avec des informations produit standard et affichez la détermination du prix uniquement après le choix d'un produit ou d'une quantité spécifique par le client. Ces techniques paraissent évidentes dans le cadre des systèmes de réservations de compagnies aériennes qui fournissent des informations prix, une offre pour un vol spécifique, contrairement à d'autres sites Web qui indiquent la détermination du prix pour chaque article affiché comme élément des catalogues de produit.

24.7.2 Messages de jobs en temps réel aux instances d'adaptateur

Si un job en temps réel souhaite envoyer un message à une instance d'adaptateur, vous devez décider si vous devez créer un appel de fonction de message ou un message sortant.

- Les appels de fonction message permettent à l'instance d'adaptateur de collecter des demandes et d'envoyer des réponses.
- Les objets de message sortant peuvent uniquement envoyer des messages sortants. Ils ne peuvent pas être utilisés pour recevoir des messages.

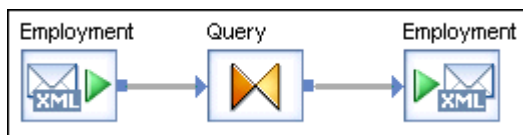
24.7.3 Service en temps réel appelé par une instance d'adaptateur

Lorsqu'une instance de l'opération (dans un adaptateur) reçoit un message d'une ressource d'information, elle le convertit en fichier XML (le cas échéant) puis envoie le message XML à un service en temps réel.

Dans le service en temps réel, le message de l'adaptateur est représenté par un objet de schéma DTD ou XML (archivé dans l'onglet [Formats](#) de la bibliothèque d'objets). Le schéma DTD ou XML représente le schéma de données pour la ressource d'information.

Le service en temps réel traite le message de la ressource de l'information (transmis par l'adaptateur) et renvoie une réponse.

Dans l'exemple de flux de données ci-dessous, le Query traite un message (représenté ici par "Activité") reçu d'une source (une instance de l'adaptateur) et renvoie la réponse à une cible (à nouveau une instance de l'adaptateur).



i Remarque

Ce document utilise des termes propres à la programmation Java. (Reportez-vous à la documentation SDK sur l'adaptateur pour en savoir plus sur les termes tels que instance de l'opération et ressource de l'information.)

24.7.4 Démarrage et arrêt des services en temps réel

Pour les phases de développement et de test, vous pouvez exécuter manuellement un job en temps réel depuis Designer. Designer exécute un job en temps réel en mode test.

Lorsque vous testez un service en temps réel ou l'exécutez dans un environnement de production, le serveur d'accès déclenche le Job Server pour qu'il traite une requête à l'aide de la logique créée dans la boucle de traitement en temps réel. Le serveur d'accès peut également déclencher le Job Server pour arrêter les boucles de traitement en temps réel. Dans un environnement de production, vous contrôlez le fonctionnement des services en temps réel à l'aide de l'Administrateur.

25 Exécution de jobs

Cette section contient une vue d'ensemble de l'exécution d'un job par le logiciel, les étapes pour exécuter les jobs, déboguer les erreurs et modifier les options du job server.

25.1 Vue d'ensemble d'une exécution de job

Il est possible d'exécuter les jobs de trois façons différentes. En fonction de vos besoins, vous pouvez configurer :

Job	Description
Jobs immédiats	Le logiciel lance les jobs batch et les jobs en temps réel, puis les exécute immédiatement depuis Designer. Pour ces jobs, Designer et le Job Server désigné (où le job s'exécute, généralement plusieurs fois sur le même ordinateur) doivent être en cours d'exécution. Vous exécuterez probablement les jobs immédiats uniquement durant le cycle de développement.
Jobs planifiés	Les jobs batch sont planifiés. Pour planifier un job, utilisez l'Administrateur ou un planificateur tiers. Lorsque les jobs sont planifiés par un logiciel tiers : • Le job démarre en dehors du logiciel. • Le job opère sur un job batch (ou script shell pour UNIX) qui a été exporté du logiciel. Lorsqu'un job est appelé par un planificateur tiers : • Le Job Server correspondant doit être en cours d'exécution. • Designer n'a pas besoin d'être en cours d'exécution.
Services	Les jobs en temps réel sont configurés comme des services qui écoutent en continu les requêtes d'un serveur d'accès et traitent les requêtes à la demande dès leur réception. Utilisez l'Administrateur pour créer un service à partir d'un job en temps réel.

25.2 Preparing for job execution

25.2.1 Validation des jobs et des composants de jobs

Il est également possible de valider explicitement les jobs et leurs composants lors de leur création en :



Cliquant sur le bouton *Tout valider* dans la barre d'outils (ou en sélectionnant **► Valider ► Tous les objets de la vue** depuis le menu *Débogage*). Cette commande contrôle la syntaxe de la définition d'objet pour l'espace de travail actif et pour tous les objets qui sont appelés de façon récursive depuis la vue de l'espace de travail actif.



En cliquant sur le bouton *Valider l'affichage en cours* dans la barre d'outils (ou en sélectionnant **► Valider ► Vue actuelle** depuis le menu *Débogage*). Cette commande contrôle la syntaxe de la définition d'objet pour l'espace de travail actif.

Vous pouvez définir les options de Designer (**► Outils ► Options ► Designer ► Général**) pour valider les jobs démarrés dans Designer avant l'exécution du job. Le paramètre par défaut est de ne pas valider.

Le logiciel valide également les jobs avant de les exporter.

Si lors de la validation le logiciel découvre une erreur dans la définition d'objet, il ouvre une boîte de dialogue indiquant qu'une erreur existe, puis il ouvre la fenêtre Sortie pour afficher l'erreur.

S'il existe des erreurs, double-cliquez sur l'erreur dans la fenêtre Sortie pour ouvrir l'éditeur de l'objet contenant l'erreur.

Si vous ne pouvez pas lire le texte d'erreur complet dans la fenêtre, vous pouvez accéder à des informations supplémentaires en faisant un clic droit sur la liste des erreurs et en sélectionnant *Afficher* depuis le menu contextuel.

Les messages d'erreur ont les niveaux de gravité suivants :

Gravité	Description
Information 	Message d'information uniquement ; n'empêche pas le job d'exécuter. Aucune action n'est requise.
Avertissement 	L'erreur n'est pas assez grave pour arrêter l'exécution du job, mais il est possible que vous obteniez des résultats imprévus. Par exemple, si le type de données d'une colonne source dans une transformation dans un flux de données ne correspond pas au type de données d'une colonne cible dans une transformation, le logiciel vous alerte avec un message d'avertissement.
Erreur 	L'erreur est assez grave pour arrêter l'exécution du job. Vous devez réparer l'erreur avant que le job exécute.

25.2.2 S'assurer que le Job Server est en cours d'exécution

Avant d'exécuter un job (comme une tâche immédiate ou planifiée), assurez-vous que le Job Server est associé avec le référentiel dans lequel le client exécute.

Lorsque Designer démarre, il affiche le statut du Job Server pour le référentiel auquel vous êtes connecté.

Icône	Description
	Le Job Server est en cours d'exécution
	Le Job Server est inactif

Le nom du Job Server actif et le numéro du port apparaissent dans la barre de statut lorsque le curseur se trouve sur l'icône.

25.2.3 Paramétrage des options d'exécution du job

Les options pour les jobs incluent le débogage et le suivi. Bien qu'il s'agisse d'options d'objet (elles affectent la fonction de l'objet), elles se trouvent soit dans la fenêtre Propriété soit dans la fenêtre Exécution associée avec le job.

Les options d'exécution pour les jobs peuvent être définies pour une instance unique ou comme une valeur par défaut.

- Le menu de clic droit [Exécuter](#) définit les options uniquement pour une exécution unique et remplace les paramètres par défaut.
- Le menu de clic droit [Propriétés](#) définit les paramètres par défaut.

25.2.3.1 Options d'exécution

Utilisez les options d'exécution pour aider à capturer et à diagnostiquer les erreurs à l'aide des options de journal, d'audit, de collection de statistiques ou de récupération.

Les informations du journal sont enregistrées dans l'un des quatre fichiers suivants dans le répertoire `<DS_COMMON_DIR>\log\<Nom_Job_Server>\<nom référentiel>`:

- Fichier journal des erreurs
- Fichier journal de surveillance
- Fichier journal de suivi
- Fichier du package de débogage

Lors de l'exécution d'un job, il est également possible de sélectionner une configuration système et un Job Server ou un groupe de serveurs depuis l'onglet [Options d'exécution](#) de la fenêtre [Propriétés d'exécution](#).

Sélectionnez l'onglet [Options d'exécution](#) pour définir les options suivantes.

Options	Description
Collecter les statistiques pour la surveillance	Sélectionnez cette option pour afficher les statistiques de cache dans le Moniteur de performance, situé dans l'Administrateur de la console de gestion. Paramètre par défaut : non sélectionné. i Remarque Utilisez cette option uniquement pour regarder la taille du cache. Pour plus d'informations, consultez la section Utilisation de caches dans le <i>Guide d'optimisation de la performance</i>.
Collecter les statistiques pour l'optimisation	Sélectionnez cette option pour collecter des statistiques que l'optimiseur Data Services utilise pour sélectionner un type de cache optimal : en mémoire et paginable. Paramètre par défaut : non sélectionné. Pour plus d'informations, consultez le <i>Guide d'optimisation de la performance</i>.
Désactiver la collecte des statistiques de validation des données	Sélectionnez cette option si ne vous souhaitez pas collecter de statistiques de validation des données pour toutes les transformations de validation du job. Paramètre par défaut : non sélectionné. Pour en savoir plus sur les statistiques de validation des données, consultez le <i>Guide Management Console Guide</i>.
Niveau de distribution	Sélectionnez le niveau avec lequel vous souhaitez distribuer le traitement de la tâche vers plusieurs Job Servers. Les options incluent : • Job : l'ensemble du job s'exécute sur un Job Server disponible. • Flux de données : chaque flux de données du job s'exécute sur un Job Server disponible. • Flux de données secondaire : chaque flux de données secondaire, une transformation ou fonction distincte, par exemple, d'un flux de données s'exécute sur un Job Server disponible. Pour en savoir plus, voir la section « Utilisation du grid computing pour distribuer l'exécution des flux de données » et « Distribution de l'exécution du flux de données » dans le <i>Guide d'optimisation de la performance</i>.
Activer l'audit	Sélectionnez cette option pour collecter les statistiques d'audit pour l'exécution de ce job spécifique. Paramètre par défaut : sélectionné. Pour plus d'informations sur l'audit, consultez le <i>Guide de Designer</i>.
Activer la restauration	S'applique à des jobs batch uniquement. Sélectionnez cette option pour activer la fonctionnalité de restauration automatique. Une fois activée, le logiciel enregistre les résultats des étapes terminées et vous permet de reprendre les jobs qui ont échoué. Paramètre par défaut : non sélectionné. Il est impossible d'activer la fonctionnalité de restauration automatique lors de l'exécution d'un job en mode d'analyse des données. De plus, cette option n'est disponible qu'en tant que propriété d'exécution. Elle n'est pas disponible en tant que propriété par défaut. Pour plus d'informations, voir le <i>Guide de Designer</i>.
Exporter les rapports Data Quality	Sélectionnez cette option pour générer et exporter tous les rapports de job spécifiés vers l'emplacement indiqué dans le nœud Configuration du serveur de rapport dans l'Administrateur de la console de gestion. Paramètre par défaut : Les rapports sont exportés dans <code><DS_COMMON_DIR>\DataQuality\reports\<repository\job></code>.

Options	Description
<i>Job Server ou groupe de serveurs</i>	Sélectionnez un Job Server ou un groupe de serveurs pour exécuter le job. La liste contient les Job Servers et les groupes de serveurs liés au référentiel du job. Définissez le Job Server avec un nom d'hôte et un port. Définissez un groupe de serveurs par le nom de groupe de serveurs. Lorsque vous sélectionnez un Job Server ou un groupe de serveurs, rappelez-vous que de nombreux objets dans le Designer ont des options définies pour l'emplacement du Job Server. Par exemple : - Noms de répertoire et de fichier pour les fichiers source et cible - Répertoires de chargement par lot Pour plus d'informations, consultez le <i>Guide de la Console de gestion</i>.
<i>Surveiller le taux d'échantillonnage</i> (nombre de secondes)	Saisissez le nombre de secondes écoulées avant que le logiciel écrive les informations dans le fichier journal de suivi et mette à jour les événements du job. Le logiciel écrit les informations sur le statut de chaque source, cible ou transformation. Paramètre par défaut : 5 secondes. Par exemple, si vous saisissez 30, les journaux se mettent à jour toutes les 30 secondes. Lors de la définition de la valeur, évaluez les améliorations de la performance obtenues en effectuant moins d'appels vers le système d'exploitation par rapport à la capacité à trouver des erreurs rapidement. Avec un taux d'échantillonnage de la surveillance plus élevé, plus de données sont collectées avant l'appel au système d'exploitation pour ouvrir le fichier : la performance est améliorée. Cependant, avec un taux d'échantillonnage de la surveillance plus élevé, il s'écoule plus de temps avant que vous puissiez voir les erreurs. La définition du taux sur 0 désactive la fonctionnalité de surveillance.
<i>i</i> Remarque Si vous utilisez un scanner anti-virus sur vos fichiers, excluez le journal de l'analyse de virus. Autrement, l'analyse de virus analyse le journal répété durant l'exécution du job, ce qui cause une dégradation de la performance.	
<i>Imprimer tous les messages de suivi</i>	Cochez cette case pour imprimer tous les messages de suivi dans le fichier journal de suivi pour le Job Server actuel. Cette option remplace les propriétés de suivi définies dans l'onglet <i>Suivi</i>.
<i>Restaurer après le dernier échec d'exécution</i>	S'applique à des jobs batch uniquement. Sélectionnez cette option pour reprendre un job échoué. Le logiciel extrait les résultats de toutes les étapes qui ont été correctement exécutées précédemment et réexécute toutes les autres étapes. Cette option est une propriété d'exécution et n'est disponible que si le job n'a pas été exécuté ou si vous avez désactivé le mode de récupération pendant une opération antérieure. Pour plus d'informations sur la récupération automatique, consultez le <i>Guide de Designer</i>.

Options	Description
Configuration du système	Sélectionnez la configuration du système à utiliser lors de l'exécution de ce job. Une configuration de système définit un ensemble de configurations de banques de données qui détermine les connexions aux banques de données. Si aucune configuration système n'est spécifiée, le logiciel utilise la configuration de banque de données par défaut pour chaque banque de données. Cette option est une propriété d'exécution. Cette option n'est disponible que si des configurations système sont définies dans le référentiel. Pour plus d'informations concernant la création et la gestion de plusieurs configurations de banque de données, consultez le <i>Guide de Designer</i>.
Utiliser les statistiques collectées	Sélectionnez cette option si vous souhaitez que l'optimisateur utilise les statistiques de cache collectées lors d'une précédente exécution du job. Paramètre par défaut : sélectionné. Pour plus d'informations, consultez le <i>Guide d'optimisation de la performance</i>.
Instance unique	Sélectionnez cette option pour éviter plusieurs instances d'un job en cours d'exécution.

25.2.3.2 Définir les options d'exécution pour chaque exécution du job

1. À partir de la zone [Projet](#), cliquez avec le bouton droit de la souris sur le nom du job et sélectionnez [Propriétés](#).
2. Sélectionnez les options dans la fenêtre Propriétés.

Informations associées

[Ajout, modification et affichage de propriétés d'objet \[page 63\]](#)

[Propriétés de suivi \[page 375\]](#)

[Définir une valeur de variable globale comme propriété de job \[page 322\]](#)

[Définir une valeur de variable globale comme propriété d'exécution \[page 322\]](#)

[Options d'exécution \[page 366\]](#)

25.2.3.3 Job d'instance unique

Assurez-vous qu'une seule instance d'un job s'exécute à la fois en activant l'option [Job d'instance unique](#).

Il est possible d'éviter des instances multiples d'exécution d'un job en cochant la case [Instance unique](#) dans la fenêtre [Propriétés](#). Une fois l'option [Instance unique](#) activée, le système abandonne toute instance supplémentaire du job lorsque l'instance en cours s'exécute. Dans la Console de gestion, vous verrez l'icône X rouge en regard de la seconde instance du job dans l'onglet [Statut du job batch](#).

Consultez le fichier journal des erreurs des jobs pour découvrir la raison pour laquelle le logiciel interrompt le job. Si le logiciel a interrompu le job car il s'agissait de la seconde instance du même job, un message d'erreur qui inclut l'heure de début du job en cours d'exécution, ainsi que le nom du journal de suivi du job s'est affiché. Vous pouvez vérifier que la première instance du job est en cours d'exécution en recherchant le job dans la fenêtre *Statut du job batch*. Il est possible que la première instance du job se soit bloquée et que le système n'ait pas été en mesure de mettre à jour le statut de job sur « échec ». Dans ce cas, interrompez le job.

Pour garantir qu'un job est une instance unique :

1. Cliquez avec le bouton droit de la souris dans la zone de projet ou dans la bibliothèque d'objets et sélectionnez *Propriétés*.
2. Ouvrez l'onglet *Options d'exécution* de la fenêtre *Propriétés* puis cochez la case *Instance unique*.
3. Cliquez sur *OK*.

25.3 Exécution des jobs comme des tâches immédiates

Les tâches immédiates ou "à la demande" sont initiées depuis Designer. Designer et le Job Server doivent être tous deux en cours d'exécution pour que le job s'exécute.

25.3.1 Exécuter un job comme une tâche immédiate

1. Dans la zone de projet, sélectionnez le nom du job.
2. Cliquez avec le bouton droit et sélectionnez *Exécuter*.

Le logiciel vous invite à enregistrer tous les objets qui comportent des modifications qui n'ont pas été enregistrées.

3. L'étape suivante est différente si la case *Procéder à une validation complète avant l'exécution du job* est cochée dans les Options de Designer :
 - Si vous n'avez pas sélectionné cette case à cocher, une fenêtre s'ouvre et affiche les propriétés d'exécution (débogage et suivi) pour le job. Passez à l'étape suivante.
 - Si vous avez sélectionné cette case à cocher, le logiciel valide le job avant son exécution. Vous devez corriger toutes les erreurs graves avant l'exécution du job. Il est possible également que des messages d'avertissement apparaissent, comme par exemple des messages qui indiquent que les valeurs de date seront converties en valeurs date-heure. Corrigez-les si vous le souhaitez (ils n'empêchent pas l'exécution du job) ou cliquez sur *OK* pour continuer. Après la validation du job, une fenêtre s'ouvre et affiche les propriétés d'exécution (débogage et suivi) pour le job.
4. Définissez les propriétés d'exécution.

Vous pouvez sélectionner le Job Server de votre choix pour traiter ce job, les profils de banque de données pour les sources et cibles le cas échéant, activer une récupération automatique, remplacer les propriétés de suivi par défaut ou sélectionner des variables globales lors de l'exécution.

i Remarque

La configuration des propriétés d'exécution affecte ici une modification temporaire pour l'exécution en cours uniquement.

5. Cliquez sur [OK](#).

Dès que le logiciel débute l'exécution, la fenêtre d'exécution s'ouvre avec le bouton du journal de suivi actif.

Utilisez les boutons en haut de la fenêtre du journal pour afficher le journal de suivi, le journal de surveillance et le journal des erreurs (s'il existe des erreurs).

Une fois le job terminé, utilisez un outil de requête RDBMS pour contrôler les contenus de la table ou du fichier cible.

Informations associées

[Designer : généralités \[page 34\]](#)

[Propriétés de suivi \[page 375\]](#)

[Débogage des erreurs d'exécution \[page 372\]](#)

[Définir une valeur de variable globale comme propriété de job \[page 322\]](#)

[Définir une valeur de variable globale comme propriété d'exécution \[page 322\]](#)

[Options d'exécution \[page 366\]](#)

25.3.2 Onglet Surveillance

L'onglet [Surveillance](#) liste les journaux de suivi de toutes les exécutions d'un job en cours ou les plus récentes.

Les icônes des feux de signalisation dans l'onglet [Surveillance](#) ont les significations suivantes :

- Un feu vert indique que le job est en cours d'exécution
Il est possible de faire un clic droit et de sélectionner Supprimer le job pour arrêter un job qui est toujours en cours d'exécution.
- Un feu rouge indique que le job est arrêté
Il est possible de faire un clic droit et de sélectionner Propriétés pour ajouter une description pour un journal de suivi spécifique. Cette description est enregistrée avec le journal auquel il est possible d'accéder ultérieurement par l'onglet Journal.
- Une croix rouge indique que le job a rencontré une erreur

25.3.3 Onglet Journal

Il est également possible de sélectionner l'onglet [Journal](#) pour afficher l'historique du journal de suivi d'un job.

Cliquez sur le journal de suivi pour l'ouvrir dans l'espace de travail.

Utilisez les icônes du journal de suivi, de surveillance et des erreurs (de gauche à droite en haut de la fenêtre d'exécution du job dans l'espace de travail) pour afficher chaque type de journal disponible pour la date et heure à laquelle le job a été exécuté.



25.4 Débogage des erreurs d'exécution

La table suivante répertorie les outils qui peuvent vous aider à comprendre les erreurs d'exécution :

Outil	Définition
Package de débogage	Crée un package de débogage de job incluant le fichier ATL ainsi que les fichiers journaux, de surveillance et de suivi. Cette action vous permet, et permet au support client SAP, de corriger les erreurs relatives à l'exécution.
Journal des erreurs	Affiche le nom de l'objet qui était exécuté au moment où une erreur s'est produite, et le texte du message d'erreur obtenu. Si le job a exécuté par rapport aux données SAP, certaines erreurs ABAP sont aussi disponibles dans le journal des erreurs.
Journal de surveillance	Affiche chaque étape de chaque flux de données dans le job, le nombre de lignes qui se déplacent dans chaque étape, et la durée de chaque étape.
Données cible	Examine toujours les données cibles pour voir si le job a produit les résultats attendus.
Journal de suivi	Enumère les étapes exécutées dans le job et l'heure à laquelle l'exécution a commencé et terminé.

Informations associées

[Utilisation des journaux \[page 372\]](#)

[Examen des journaux de suivi \[page 374\]](#)

[Examen des journaux de surveillance \[page 374\]](#)

[Examen des journaux des erreurs \[page 374\]](#)

[Examen des données cibles \[page 378\]](#)

[Création d'un package de débogage \[page 378\]](#)

25.4.1 Utilisation des journaux

Cette section décrit la manière d'utiliser les journaux dans Designer.

- Pour ouvrir le journal de suivi à l'exécution du job, sélectionnez ► **Outils** ► **Options** ► **Designer** ► **Général** ► **Ouvrir le moniteur à l'exécution du job** ►.
- Pour copier le contenu du journal à partir d'un journal ouvert, sélectionnez une ou plusieurs lignes et utilisez les commandes clés [Ctrl+C].

25.4.1.1 Accéder à un journal lors de l'exécution du job

Si vous définissez vos options de manière ouvrir le volet Moniteur lors de l'exécution de jobs, et que votre Designer s'exécute lorsque débute l'exécution de job, la fenêtre d'exécution s'ouvre automatiquement, affichant les informations du journal de suivi.

Cliquez sur les icônes du moniteur et du journal des erreurs (icônes du milieu et de droite en haut de la fenêtre d'exécution) pour afficher ces journaux. L'icône de journal des erreurs est disponible seulement si votre exécution de job génère des avertissements ou des erreurs. L'icône d'avertissement (triangle jaune avec un point d'exclamation) indique que le job contient des avertissements. L'icône d'avertissement se transforme en icône d'erreur (cercle rouge avec un X blanc au dessus) l'exécution de job contient des erreurs.



Il est possible de cliquer sur les différentes icônes lorsque le job s'exécute pour afficher chaque journal. Le volet Moniteur reste ouvert jusqu'à ce que vous le fermiez.

25.4.1.2 Accès à un journal après que la fenêtre d'exécution a été fermée

1. Dans la zone du projet, cliquez sur l'onglet *Journal*.
2. Cliquez sur un nom de job pour afficher tous les fichiers des journaux de suivi, de surveillance et des erreurs dans l'espace de travail. Ou développez le job qui vous intéresse pour afficher la liste des fichiers du journal de suivi et cliquez sur l'un d'entre eux.

Les indicateurs de journal signifient les éléments suivants :

Indicateur du journal de job	Description
 N_	Indique que le job a exécuté avec succès sur le Job Server explicitement sélectionné.
	Indique que job a été exécuté avec succès par un groupe de serveurs. Le Job Server listé a exécuté le job.
	Indique que le job a rencontré une erreur sur le Job Server explicitement sélectionné.
	Indique que le job a rencontré une erreur alors qu'il était exécuté par un groupe de serveurs. Le Job Server listé a exécuté le job.

3. Cliquez sur l'icône du journal pour l'exécution du job qui vous intéresse. (Identifiez l'exécution à partir du poste dans la séquence ou de l'horodatage.)
4. Utilisez la zone de liste pour passer d'un type de journal à un autre ou pour afficher *Aucun journal* ou *Tous les journaux*.

25.4.1.3 Supprimer un journal

Supprimez les fichiers journaux de la zone de projet dans Designer.

⚠ Attention

Avant de supprimer les fichiers journaux, assurez-vous que vous n'avez pas besoin des informations qu'ils contiennent pour des fins de reporting ou statistiques. Certains fichiers journaux contiennent d'importantes informations dont le logiciel a besoin pour créer des rapports réglementaires ou certifiés ainsi que des tables statistiques et certains fichiers journaux sont écrasés lorsque vous exécutez un autre job. Nous vous suggérons donc, si vous devez créer des données réglementaires, certifiées ou statistiques basées sur un processus spécifique, de créer des copies de sauvegarde de ces fichiers avant de les supprimer.

Pour supprimer manuellement les fichiers journaux du Designer :

1. Dans la zone du projet, cliquez sur l'onglet *Journal*.
2. Faites un clic droit sur le journal que vous souhaitez supprimer et sélectionnez *Supprimer le journal*.

25.4.1.4 Examen des journaux de suivi

Utilisez les journaux de suivi pour déterminer à quel endroit une exécution a échoué, si les étapes d'exécution se sont produites dans l'ordre attendu, et quelles parties de l'exécution prennent le plus de temps.

25.4.1.5 Examen des journaux de surveillance

Le journal de surveillance quantifie les activités des composants du job. Il fait une liste du temps passé sur un composant donné d'un job et du nombre de lignes de données qui se déplacent par ce composant.

25.4.1.6 Examen des journaux des erreurs

Le logiciel produit un journal des erreurs pour chaque exécution de job. Utilisez les journaux des erreurs pour déterminer la manière dont une exécution a échoué. Si l'exécution s'est terminée sans erreur, le journal des erreurs est vide.

25.4.1.7 Désactiver les avertissements survenant lors de l'exécution du job

Désactiver les avertissements qui sont journalisés pendant l'exécution.

Par défaut, Data Services journalise un avertissement pour chaque enregistrement d'occurrences. Ce processus peut compliquer la navigation dans le journal

Pour réduire le nombre d'avertissements apparaissant dans le journal des erreurs, procédez comme suit :

1. Ouvrez le fichier `DSConfig.txt` situé dans `%DS_COMMON_DIR%\conf`.
2. Activez le paramètre `Disable_Runtime_warning` dans la section `AI_Engine` sur la valeur `True`.
3. Enregistrez et fermez le fichier.

25.4.2 Propriétés de suivi

Les propriétés de suivi sont des types d'information que le logiciel surveille lors de l'exécution du job, et écrit dans les fichiers de suivi.

Utilisez l'onglet [Suivi](#) de la fenêtre [Propriétés](#) pour sélectionner les informations que SAP Data Services surveille et écrit dans le fichier journal de suivi au cours d'un job. Le logiciel écrit les messages de suivi dans le journal de suivi associé au Job Server actuel et des messages d'erreur dans le journal des erreurs associé au Job Server actuel.

Pour activer ou désactiver une propriété de suivi, affichez l'onglet [Suivi](#) et sélectionnez [Oui](#) ou [Non](#) dans la liste [Valeur](#) pour cette propriété.

La table suivante contient des descriptions pour chaque propriété de suivi par ordre alphabétique.

Suivi	Description
Communication du serveur d'accès	Ecrit les messages échangés entre le serveur d'accès et un fournisseur de service, dont les types de messages suivants : • Un message d'enregistrement qui annonce au serveur d'accès que le fournisseur de service est prêt. • Une demande d'exécution que le serveur d'accès envoie au fournisseur de services. • Une réponse du service au serveur d'accès. • Toutes les requêtes du serveur d'accès pour arrêter.
Assembleurs	Ecrit des messages pour les paramètres de substitution et les transformations SDK : • Paramètres de substitution : écrit des messages de suivi tels que la configuration de substitution utilisée et la valeur substituée pour chaque paramètre de substitution. • Transformations SDK : écrit des informations spécifiques à la transformation indiquées au format XML. Les informations spécifiques à la transformation peuvent être hiérarchiques. Le message de suivi des assembleurs écrit le fichier XML extrait ou assemblé au moment de l'exécution.

Suivi	Description
Données d'audit	Ecrit un message lors de l'audit : • Collecte une statistique à un point d'audit • Détermine si une règle d'audit a réussi ou échoué
Flux de données	Ecrit un message lorsque le flux de données démarre, lorsque qu'il a été correctement terminé ou quand il s'est arrêté à cause d'une erreur. Ce suivi indique également à quel moment le Bulk Loader démarre, à quel moment les avertissements de chargeur par lot ont lieu, et à quel moment le chargeur par lot se termine.
Package de débogage	Crée un package contenant le fichier ATL ainsi que les fichiers journaux, de surveillance et de suivi pour le job.
Lecteur de fichiers IDoc	Ecrit un message lors de la lecture de fichiers IDoc, dont : • début de la lecture du fichier IDoc • arrêt de la lecture du fichier IDoc • résultat de la validation du fichier IDoc
Source de la mémoire	Ecrit un message pour chaque ligne extraite de la table en mémoire.
Cible de la mémoire	Ecrit un message pour chaque ligne insérée dans la table en mémoire.
Flux de données optimisé	Pour conseil et assurance client
Fonction RFC	Ecrit un message lié aux appels RFC dont : • début de l'appel RFC • fin de l'appel RFC • un message pour chaque enregistrement reçu du logiciel pour l'appel RFC
Ligne	Ecrit un message lorsqu'une transformation importe ou exporte une ligne.
Lecteur de table SAP	Ecrit des messages à partir des lecteurs des tables du système SAP, y compris : • début de la lecture depuis la table • fin de la lecture depuis la table • début de la connexion au système SAP où se trouve la table • fin de la connexion au système SAP
Scripts et fonctions de script	Écrit un message lorsque le logiciel exécute un script ou appelle une fonction de script. Plus précisément, ce suivi lie un message lorsque : • Le script est appelé. Les scripts peuvent être démarrés à tout niveau, du niveau de job vers le niveau de flux de données. Une notation supplémentaire (et distincte) est réalisée lorsqu'un script est appelé depuis un autre script. • Une fonction est appelée par le script. • Le script a été correctement terminé.
Session	Ecrit un message lorsque la description du job est lue depuis le référentiel, lorsque le job est optimisé, et lorsque le job est exécuté.

Suivi	Description
Fonctions SQL	Ecrit les données extraites avant les fonctions SQL : • Chaque ligne extraite par la requête nommée avant le SQL est envoyée dans la fonction <code>key_generation</code> • Chaque ligne extraite par la requête nommée avant le SQL est envoyée dans la fonction de recherche (mais uniquement si <code>CACHE_DE_PRECHARGEMENT</code> n'est pas spécifié). • Lorsque le courrier est envoyé à l'aide de la fonction <code>mail_to</code>.
Chargeurs SQL	Ecrit un message lorsque le Bulk Loader : • Démarre • Envoie un message d'avertissement • Se termine correctement • Se termine sans succès, si vous sélectionnez l'option Nettoyer le répertoire du Bulk Loader après le chargement De plus, pour Microsoft SQL Server et SAP ASE, écrit lorsque le Bulk Loader du SQL Server : • Termine un envoi de ligne réussi • Rencontre une erreur Cette instance rapporte tous les SQL envoyés par le logiciel à la base de données cible, y compris : • Lorsqu'une commande <code>truncate table</code> s'exécute si l'option Supprimer les données de la table avant le chargement est sélectionnée. • Tout paramètre inclus dans les commandes SQL <code>PRECHARGEMENT</code> • Avant qu'un lot d'instructions SQL soit soumis • Lorsqu'une table de modèle est créée (et également supprimée, si vous activez l'option Supprimer/Créer) • Lorsqu'une commande <code>delete from table</code> s'exécute si vous activez la correction automatique (environnement Informix uniquement). Ce suivi écrit également toutes les lignes chargées par le logiciel dans la cible.
SQL seulement	À utiliser avec l'option Effectuer le suivi des transformations SQL, l'option Effectuer le suivi des lecteurs SQL ou l'option Effectuer le suivi des chargeurs SQL pour arrêter l'écriture des messages de suivi pour les données envoyées et reçues de la base de données.
Lecteurs SQL	Ecrit le bloc de requête SQL qu'un script, une transformation Query ou une fonction SQL soumet au système. Ecrit également les résultats SQL.
Transformations SQL	Ecrit un message (à l'aide de la transformation <code>Table_Comparison</code>) sur l'existence ou non d'une ligne dans la table cible qui correspond à une ligne d'entrée de la table source. Le message de suivi a lieu avant l'envoi de la requête à la cible et pour chaque ligne extraite lorsque la requête nommée est envoyée (mais uniquement si la mise en cache n'est pas activée).

Suivi	Description
Procédure stockée	Écrit un message lorsque le logiciel appelle une procédure stockée. Le rapport inclut : • Le moment où la procédure stockée démarre • La requête SQL soumise pour l'appel à la procédure stockée • La valeur (ou les valeurs) du paramètre d'entrée (ou des paramètres) • La valeur (ou les valeurs) du paramètre de sortie (ou des paramètres) • La valeur renvoyée (si la procédure stockée est une fonction stockée) • Le moment où la procédure stockée se termine.
Tables	Écrit un message lorsqu'une table est créée ou supprimée. Le message indique la banque de données à laquelle appartient la table créée et l'instruction SQL utilisée pour créer la table.
Suivi ABAP	Écrit un message lorsqu'un flux de données ABAP démarre ou s'arrête, et pour rapporter le statut du job ABAP.
Suivi de l'exécution parallèle	Écrit un message qui décrit la manière dont les données dans un flux de données sont traitées en parallèle.
Transformation	Écrit un message lorsqu'une transformation débute, se termine, ou s'arrête.
Workflow	Écrit un message lorsque la description du workflow est lue depuis le référentiel, lorsque le workflow est optimisé, lorsque le workflow s'exécute, et lorsque le workflow s'achève.

25.4.3 Examen des données cibles

La meilleure mesure de la réussite d'un job est l'état des données cibles. Examinez toujours les données pour vous assurer que l'opération de mouvement des données a produit les résultats que vous attendiez. Assurez-vous que :

- Les données n'aient pas été tronquées ou converties dans des types incompatibles.
- Les données n'aient pas été dupliquées dans la cible.
- Les données ne se soient pas perdues entre les mises à jour de la cible.
- Les clés générées aient été correctement incrémentées.
- Les valeurs mises à jour aient été correctement gérées.

25.4.4 Création d'un package de débogage

Créer un package de débogage de job incluant le fichier ATL ainsi que les fichiers journaux, de surveillance et de suivi.

Vous pouvez créer un package de débogage de job incluant le fichier ATL ainsi que les fichiers journaux, de surveillance et de suivi. Envoyez le fichier .zip ou .tar généré au support client SAP pour vous aider à corriger les erreurs relatives à l'exécution. Vous pouvez générer le package de débogage lors de l'exécution du job soit

depuis la fenêtre Propriétés d'exécution de Designer ou la page Options d'exécution de l'Administrateur de la console de gestion.

1. Dans la zone de projet Designer, cliquez avec le bouton droit de la souris sur le nom du job et sélectionnez [Exécuter](#).

Le logiciel vous invite à enregistrer tous les objets qui comportent des modifications qui n'ont pas été enregistrées.

2. Dans la fenêtre [Propriétés d'exécution](#), sélectionnez l'onglet [Suivi](#).
3. Sélectionnez [Créer un package de débogage](#) et remplacez la valeur par [Oui](#).
4. Cliquez sur [OK](#).

Le logiciel génère le fichier de package de débogage .zip ou .tar à l'emplacement suivant :

`<DS_COMMON_DIR>\log\<Nom_Job_Server>\<nom référentiel>` sur l'ordinateur sur lequel le Job Server est installé.

Vous pouvez télécharger le package de débogage à partir de l'Administrateur de la console de gestion. Pour en savoir plus, voir le *Guide de la Console de gestion de SAP Data Services*.

25.5 Modification des options du Job Server

Familiarisez-vous avec les aspects les plus techniques de la gestion des données effectuée par le logiciel (à l'aide du *Guide de référence*) et certaines de ses interfaces comme celles pour les adaptateurs et les applications SAP.

Il existe de nombreuses options disponibles dans le logiciel pour le dépannage et l'ajustement d'un job.

Option	Description de l'option	Valeur par défaut
Adapter Data Exchange Time-out	(Pour les adaptateurs). Définit la durée durant laquelle un appel de fonction ou un message sortant attend une réponse de l'opération de l'adaptateur.	10800000 (3 heures)
Adapter Start Time-out	(Pour les adaptateurs). Définit la durée durant laquelle l'Administrateur ou Designer attend une réponse du Job Server qui gère les adaptateurs (démarrage/arrêt/statut).	90000 (90 secondes)
AL_JobServerLoadBalance-Debug	Permet à un Job Serveur de noter les informations du groupe de serveurs si la valeur est définie sur VRAI. Les informations sont enregistrées sous : <code><LINK_DIR> /log/<JobServerName>/server_eventlog.txt</code> .	FAUX
AL_JobServerLoadOSPolling	Définit l'intervalle d'interrogation (en secondes) que le logiciel utilise pour obtenir les informations sur le statut utilisées pour calculer l'index d'équilibrage de chargement. Cet index est utilisé par des groupes de serveurs.	60

Option	Description de l'option	Valeur par défaut
Display DI Internal Jobs	Affiche la banque de données interne du logiciel, CD_DS_d0cafae2, et les jobs associés dans la bibliothèque d'objets. La banque de données CD_DS_d0cafae2 prend en charge deux jobs internes. Le premier calcule les dépendances d'utilisation sur les tables de référentiel, et le deuxième met à jour les configurations du groupe de serveurs. Si vous modifiez le mot de passe du référentiel, le nom d'utilisateur ou d'autres informations de connexion, remplacez la valeur par défaut de cette option par VRAI, fermez et rouvrez Designer, puis mettez à jour la configuration de banque de données CD_DS_d0cafae2 pour qu'elle corresponde à la nouvelle configuration du référentiel. Cela permet au job de calcul des dépendances d'utilisation (CD_JOBd0cafae2) et au job du groupe de serveurs (di_job_al_mach_info) d'exécuter sans aucune erreur de connexion.	FAUX
FTP Number of Retry	Définit le nombre de tentatives pour une connexion au FTP qui échoue initialement.	0
FTP Retry Interval	Définit l'intervalle de tentative de connexion au FTP en millisecondes.	1000
Global_DOP	Définit le degré de parallélisme pour tous les flux de données exécutés par un Job Server donné. Il est également possible de définir le <i>Degré de parallélisme</i> pour les flux de données individuels à partir de la fenêtre Propriétés de chaque flux de données. Si la valeur de <i>Degré de parallélisme</i> d'un flux de données est 0, alors le Job Server utilise la valeur Global_DOP. Le Job Server utilise la valeur de <i>Degré de parallélisme</i> du flux de données si elle est définie sur une autre valeur que 0 car elle remplace la valeur Global_DOP.	2
Ignore Reduced Msg Type	(Pour les applications SAP). Désactive le traitement du type de message réduit IDoc pour tous les types de message si la valeur est définie sur TRUE.	FAUX
Ignore Reduced Msg Type_foo	(Pour les applications SAP). Désactive le traitement du type de message réduit IDoc pour un type de message spécifique (tel que <foo>) si la valeur est définie sur TRUE.	FAUX
OCI Server Attach Retry	Le moteur appelle la fonction OCIServerAttach d'Oracle à chaque fois qu'il établit une connexion avec Oracle. Si le moteur appelle cette fonction trop rapidement (en traitant des flux de données parallèles par exemple), la fonction peut échouer. Pour corriger cette erreur, augmenter la valeur de tentative à 5.	3

Option	Description de l'option	Valeur par défaut
Splitter Optimization	Le logiciel peut se bloquer si vous créez un job dans lequel une source de fichier alimente deux requêtes. Si cette option est définie sur VRAI, le moteur crée en interne deux fichiers sources qui alimentent les deux requêtes au lieu d'un séparateur qui alimente les deux requêtes.	FAUX
Use Explicit Database Links	Les jobs avec des liens de base de données normalement importés font preuve d'une performance améliorée car le logiciel utilise ces liens pour pousser vers le bas le traitement vers une base de données. Si vous définissez cette option sur FAUX, aucun flux de données n'utilise les banques de données liées. L'utilisation des banques de données liées peut également être désactivée depuis n'importe quelle boîte de dialogue des propriétés des flux de données. L'option de niveau du flux de données a la priorité sur cette option de niveau du Job Server.	VRAI
Use Domain Name	Ajoute un nom de domaine à un nom de Job Server dans le référentiel. Cet ajout crée un nom de serveur complet et permet à Designer de localiser un Job Server sur un domaine différent. Si le nom complet de l'hôte (y compris le domaine) n'est pas correctement stocké, la connexion entre Designer ou le Job Server et le référentiel peut échouer. Si ce problème survient, attribuez à l'option la valeur FAUX avant d'ajouter les référentiels.	VRAI

25.5.1 Modifier les valeurs d'option pour un Job Server individuel

- Sélectionnez le Job Server avec lequel vous souhaitez travailler en le configurant comme Job Server par défaut.
 - Sélectionnez **Outils** > **Options** > **Designer** > **Environnement**.
 - Sélectionnez un Job Server dans la section *Job Server par défaut*.
 - Cliquez sur **OK**.
- Sélectionnez **Outils** > **Options** > **Job Server** > **Général**.
- Saisissez la section et la clé que vous souhaitez utiliser dans la liste suivante des paires de valeurs :

Section	Clé
int	AdapterDataExchangeTimeout
int	AdapterStartTimeout
AL_JobServer	AL_JobServerLoadBalanceDebug
AL_JobServer	AL_JobServerLoadOSPolling
string	DisplayDIInternalJobs
AL_Engine	FTPNumberOfRetry
AL_Engine	FTPRetryInterval
AL_Engine	Global_DOP
AL_Engine	IgnoreReducedMsgType
AL_Engine	IgnoreReducedMsgType_foo
AL_Engine	OCIServerAttach_Retry
AL_Engine	SPLITTER_OPTIMIZATION
AL_Engine	UseExplicitDatabaseLinks
Repository	UseDomainName

4. Saisissez une valeur.

Par exemple, saisissez les informations suivantes pour modifier la valeur par défaut du nombre de fois qu'un Job Serveur retente d'établir une connexion au FTP s'il a initialement échoué.

Option	Exemple de valeur
Section	AL_Engine
Clé	FTPNumberOfRetry
Valeur	2

Ces paramètres modifient la valeur par défaut pour l'option FTPNumberOfRetry passant de zéro à deux.

5. Pour enregistrer les paramètres et fermer la fenêtre Options, cliquez sur [OK](#).
6. Sélectionnez à nouveau un Job Serveur par défaut en répétant l'étape 1, si besoin.

25.5.2 Utilisation des noms de lecteurs mappés dans un chemin d'accès

Le progiciel prend uniquement en compte les chemins d'accès au format UNC (Universal Naming Convention) pour les répertoires. Si vous configurez un chemin d'accès pour un lecteur mappé, le progiciel convertit le lecteur en question dans son équivalent UNC.

Pour éviter que votre lecteur mappé ne soit reconverti en chemin d'accès UNC, vous devez ajouter les noms des lecteurs dans la fenêtre [Options](#) de Designer.

1. Sélectionnez [Outils](#) > [Options](#).
2. Dans la fenêtre [Options](#), développez [Job Server](#) et sélectionnez [Général](#).
3. Dans la zone d'édition [Section](#), saisissez **MappedNetworkDrives**.
4. Dans la zone d'édition [Clé](#), saisissez **LecteurLocal1** pour établir le mappage avec un lecteur local ou **LecteurDistant1** pour établir le mappage avec un lecteur distant.
5. Dans la zone d'édition [Valeur](#), saisissez une lettre de lecteur, comme **M:** pour un lecteur local ou **\<nom_ordinateur>\<nom_partage>** pour un lecteur distant.
6. Pour fermer la fenêtre, cliquez sur OK.

Si vous voulez ajouter un autre lecteur mappé, vous devez fermer la fenêtre [Options](#) et recommencer la saisie. Veillez à ce que chaque entrée dans la zone d'édition [Clé](#) corresponde à un nom unique.

26 Évaluation des données

Puisque les systèmes opérationnels changent fréquemment, le contrôle de la qualité des données devient critique dans vos jobs d'extraction, de transformation et de chargement (ETL).

Le logiciel propose des contrôles de qualité des données qui identifient et réparent les erreurs dans les données. Ces fonctionnalités peuvent aider à s'assurer que vous avez des informations fiables.

Vous pouvez utiliser les fonctionnalités suivantes proposées par le logiciel pour déterminer et améliorer la qualité et la structures de vos données source :

Fonctionnalité	Description
Profileur de données	Utilisez le profileur de données pour déterminer les aspects suivants de vos données : • La qualité des données source avant de les extraire. Le Profileur de données peut identifier les anomalies dans les données sources pour vous aider à mieux définir des mesures de correction pour la transformation Validation, Data Quality ou pour d'autres transformations. • La distribution, la relation et la structure des données sources pour mieux concevoir les jobs et les flux de données, ainsi que l'entrepôt de données cible. • Le contenu des données sources et cibles pour pouvoir vérifier que le job d'extraction des données renvoie les résultats attendus.
Afficher les données	La fonctionnalité d'affichage de données vous permet d'effectuer les tâches suivantes : • Afficher les données sources avant d'exécuter un job pour vous aider à créer des conceptions de job de meilleure qualité. • Comparer des échantillons de données de différentes étapes de votre job pour vérifier que le job d'extraction des données renvoie les résultats attendus.
Affichage des données lors de la conception	Utilisez la fonctionnalité Affichage des données lors de la conception pour visualiser et analyser l'entrée et la sortie en temps réel lors de la conception d'une transformation. Utilisez l'Affichage des données lors de la conception avant même que le flux de données soit terminé ou valide.
Transformation Validation	Utilisation de la transformation Validation pour : • Vérifier que les données source répondent à vos règles de gestion. • Prendre des mesures appropriées lorsque les données ne répondent pas aux règles d'entreprise.
Audit du flux de données	Utilisation de la fonctionnalité d'audit du flux de données pour : • Définir des règles qui déterminent si une source, une transformation ou un objet cible traite les données correctes. • Définir les mesures à prendre lorsqu'une règle d'audit échoue.

Fonctionnalité	Description
Transformations Data Quality	Utilisation des transformations Data Quality pour améliorer la qualité des données.
Tableaux de bord Validation des données	Utilisez les tableaux de bord de validation des données dans l'outil de rapports de métadonnées pour évaluer l'exactitude des données cibles Évaluez la fiabilité des données en fonction des règles de validation créées dans vos jobs batch. Ce retour permet aux utilisateurs professionnels de réviser, d'évaluer et d'identifier les incohérences ou les erreurs potentielles dans les données sources.

Informations associées

[Présentation de la qualité des données \[page 426\]](#)

[Utilisation du Profileur de données \[page 385\]](#)

[Utilisation de la fonction d'affichage des données pour déterminer la qualité des données \[page 402\]](#)

[Utilisation de l'affichage des données lors de la conception \[page 738\]](#)

[Utilisation de la transformation Validation \[page 404\]](#)

[Utilisation de l'Audit \[page 412\]](#)

26.1 Utilisation du Profileur de données

Le Profileur de données s'exécute sur un serveur de profileur pour fournir les informations de profileur de données suivantes, qui peuvent être affichées par plusieurs utilisateurs :

Analyse	Description
Analyse de colonne	Le Profileur de données fournit deux types de profils de colonne : • Profilage de base : cette information inclut la valeur minimale, la valeur maximale, la valeur moyenne, la longueur minimale de la chaîne et la longueur maximale de la chaîne. • Profilage détaillé : l'analyse de colonne détaillée inclut le nombre distinct, le pourcentage distinct, la médiane, la longueur médiane de la chaîne, le nombre modèle et le pourcentage modèle.
Analyse de relation	Cette information identifie les incohérences de données entre deux colonnes pour lesquelles vous définissez une relation, y compris les colonnes qui ont une relation de clé primaire et clé étrangère existante. Il est possible d'enregistrer deux niveaux de données : • Enregistrez les données uniquement dans les colonnes que vous avez sélectionnées pour la relation. • Enregistrez les valeurs dans toutes les colonnes dans chaque ligne.

26.1.1 Sources de données à profiler

Il est possible d'exécuter le Profileur de données sur les données contenues dans les sources suivantes : Voir les *Notes de version* pour la liste complète des sources prises en charge par le Profileur de données.

Source	Includes
Bases de données	• Connecteur Attunity pour les bases de données sur l'ordinateur central • DB2 • Oracle • SQL Server • SAP Sybase IQ • SAP Sybase SQL Anywhere • Teradata
Applications	• One World de JD Edwards • JD Edwards World • Applications Oracle • PeopleSoft • Applications SAP • SAP NetWeaver Business Warehouse • Siebel
Fichiers plats	Par exemple, texte brut, fichier binaire, valeurs séparées par des virgules et valeurs séparées par un séparateur.

26.1.2 Connexion au serveur du profileur

Installez et configurez le serveur du profileur avant d'utiliser le Profileur de données.

Designer doit être connecté au serveur du profileur pour exécuter le Profileur de données et afficher ses résultats. Vous fournissez cette information de connexion sur la fenêtre Connexion au serveur du profileur.

1. Utilisez l'une des méthodes suivantes pour appeler la fenêtre Connexion au serveur du profileur.
 - Depuis le menu de la barre d'outils, sélectionnez  **Outils** > **Connexion au serveur du profileur** .
 - Sur la barre de statut en bas, double-cliquez sur l'icône Serveur du profileur qui se trouve à droite de l'icône Job server.
2. Saisissez vos références d'utilisateur pour le CMS.

Option	Description
Système	Spécifiez le nom de serveur et facultativement le port du CMS.
Nom d'utilisateur	Spécifiez le nom d'utilisateur à utiliser pour la connexion au CMS.
Mot de passe	Spécifiez le mot de passe à utiliser pour la connexion au CMS.
Authentification	Spécifiez le type d'authentification utilisé par le CMS.

3. Cliquez sur [Se connecter](#).

Le logiciel essaie de se connecter au CMS à l'aide des informations indiquées. Lorsque vous vous connectez avec succès, la liste des référentiels du profileur qui sont à votre disposition est affichée.

4. Sélectionnez le référentiel que vous souhaitez utiliser.
5. Cliquez sur [OK](#) pour vous connecter à l'aide du référentiel sélectionné.

Lorsque vous vous connectez avec succès au serveur du profileur, l'icône Serveur du profileur sur la barre de statut en bas n'a plus le signe X rouge dessus. De plus, lorsque vous déplacez le pointeur sur l'icône, la barre de statut affiche l'emplacement du serveur du profileur. Cliquez sur

26.1.3 Profiler statistics

26.1.3.1 Profil de colonne

Il est possible de générer des statistiques pour une ou plusieurs colonnes.

Les colonnes peuvent toutes appartenir à une source de données ou à plusieurs sources de données. Dans le cas d'une génération de statistiques pour plusieurs sources dans une tâche du profileur, toutes les sources doivent être dans la même banque de données.

Informations associées

[Profilage de base \[page 387\]](#)

[Profilage détaillé \[page 388\]](#)

[Exemples \[page 388\]](#)

[Affichage des attributs de colonne générés par le profileur de données \[page 397\]](#)

[Soumission des tâches du profileur de colonnes \[page 390\]](#)

26.1.3.1.1 Profilage de base

Par défaut, le Profileur de données génère des attributs de profileur de base pour chaque colonne sélectionnée.

La table suivante décrit les attributs de profileur de base pour chaque attribut :

Attribut de base	Description
Min	Valeur la plus basse dans cette colonne.
Nombre min.	Nombre de lignes qui contiennent la valeur la plus basse dans cette colonne.
Max	Valeur la plus élevée dans cette colonne.
Nombre max.	Nombre de lignes qui contiennent la valeur la plus élevée dans cette colonne.
Moyenne	Pour les colonnes numériques, valeur moyenne dans cette colonne.

Attribut de base	Description
Longueur minimale de la chaîne	Pour les colonnes avec des caractères, longueur de la valeur de chaîne la plus courte dans cette colonne.
Longueur maximale de la chaîne	Pour les colonnes avec des caractères, longueur de la valeur de chaîne la plus longue dans cette colonne.
Longueur moyenne de la chaîne	Pour les colonnes avec des caractères, longueur moyenne des valeurs de chaîne dans cette colonne.
Valeurs nulles	Nombre de valeurs NULL dans cette colonne.
% de valeurs nulles	Pourcentage de lignes contenant une valeur NULL dans cette colonne.
Zéros	Nombre de valeurs 0 dans cette colonne.
% de zéros	Pourcentage de lignes contenant une valeur 0 dans cette colonne.
Vides	Pour les colonnes avec des caractères, nombre de lignes qui contiennent un vide dans cette colonne.
% de vides	Pourcentage de lignes contenant un vide dans cette colonne.

26.1.3.1.2 Profilage détaillé

La génération d'attributs détaillés consomme plus de temps et de ressources de l'ordinateur, l'utilisation doit donc être limitée à des attributs spécifiques.

Vous pouvez générer plusieurs attributs détaillés en plus des attributs de base, mais la génération d'attributs détaillés consomme plus de temps et de ressources de l'ordinateur. Par conséquent, il est recommandé de ne pas sélectionner le profil détaillé à moins d'avoir besoin des attributs suivants :

Attribut détaillé	Description
Médiane	Valeur qui se trouve dans la ligne du milieu de la table source.
Longueur médiane de la chaîne	Pour les colonnes avec des caractères, valeur qui se trouve dans la ligne du milieu de la table source.
Distinctes	Nombre de valeurs distinctes dans cette colonne.
% de distinctes	Pourcentage de lignes contenant chaque valeur distincte dans cette colonne.
Modèles	Nombre de modèles différents dans cette colonne.
% de modèles	Pourcentage de lignes contenant chaque modèle dans cette colonne.

26.1.3.1.3 Exemples

Exemples d'utilisation des statistiques de profilage de colonne pour améliorer la qualité des données

Il est possible d'utiliser les attributs de profil de colonnes pour vous aider dans les différentes tâches, y compris les tâches suivantes :

- Obtenir des statistiques, des fréquences, des plages et des valeurs hors normes de base. Par exemple, ces statistiques de profilage peuvent indiquer qu'une valeur de colonne est notamment plus élevée que les autres valeurs dans la source de données. Il est possible que vous décidiez alors de définir une transformation Validation pour activer un indicateur dans une table différente avant de charger cette valeur hors normes dans la table cible.
- Identifier des variantes du même contenu. Par exemple, le numéro de référence peut être un type de données integer dans une source de données, et un type de données varchar dans une autre source de données. Il est possible que vous décidiez alors quel type de données vous souhaitez utiliser dans votre entrepôt de données cible.
- Découvrir des modèles et des formats de données. Par exemple, les statistiques de profilage peuvent indiquer que le numéro de téléphone se trouve sous des formats différents. Avec ces informations de profil, il est possible que vous décidiez de définir une transformation Validation pour toutes les convertir afin qu'elles utilisent le même format cible.
- Analyser la plage numérique. Par exemple, le numéro de client peut avoir une plage de chiffres dans une source, et une plage différente dans une autre source. La cible doit avoir un type de données qui s'adapte à la plage maximale.
- Identifier les informations manquantes, les valeurs nulles et les vides dans le système source. Par exemple, les statistiques de profilage peuvent indiquer que les valeurs nulles apparaissent pour les numéros de fax. Il est possible que vous décidiez alors de définir une transformation Validation pour remplacer la valeur nulle par une phrase telle que "Inconnu" dans la table cible.

26.1.3.2 Profil de relation

Un profil de relation indique le pourcentage des valeurs non correspondantes dans les colonnes de deux sources. Les sources incluent :

- les tables ;
- les fichiers plats ;
- la combinaison d'une table et d'un fichier plat.

Les colonnes clés peuvent avoir une clé primaire et une relation de la clé étrangère définies, ou elles peuvent n'avoir aucun lien (si l'une d'elle provient d'une banque de données et l'autre d'un format de fichier).

Il est possible de choisir entre deux niveaux de profils de relation à enregistrer :

Option	Description
Enregistrer uniquement les données de colonnes clés	Par défaut, le Profileur de données enregistre uniquement les données des colonnes que vous avez sélectionnées pour la relation. i Remarque Le niveau Enregistrer uniquement les données de colonnes clés n'est pas disponible avec l'utilisation des banques de données Oracle.
Enregistrer toutes les données de colonnes	Il est possible d'enregistrer les valeurs des autres colonnes dans chaque ligne, mais ce processus prend plus de temps et consomme plus de ressources d'ordinateur.

Lorsque vous affichez les résultats du profil de relation, vous pouvez descendre dans la hiérarchie pour voir les données réelles qui ne correspondent pas.

Il est possible d'utiliser le profil de relation pour vous aider dans les différentes tâches, y compris les tâches suivantes :

- Identifier les données manquantes dans le système source. Par exemple, une source de données peut inclure la région, mais il est possible que ce ne soit pas le cas d'une autre source.
- Identifier les données redondantes dans les sources de données. Par exemple, il est possible qu'il existe des noms et des adresses en double entre deux sources, ou qu'aucun nom n'existe pour une adresse dans une source.
- Valider les relations dans les sources de données. Par exemple, deux systèmes de suivi des problèmes différents peuvent inclure un sous-ensemble de problèmes signalés par le client, mais certains problèmes existent uniquement dans un système ou dans l'autre.

Informations associées

[Soumission des tâches du profileur de relations \[page 392\]](#)

[Affichage des résultats du profileur \[page 396\]](#)

26.1.4 Exécution d'une tâche du profileur

Le profileur de données permet de calculer les statistiques du profileur pour tous les ensembles de colonnes sélectionnés.

i Remarque

Cette fonctionnalité facultative n'est pas disponible pour les colonnes avec des schémas imbriqués, ou les types de données LONG ou TEXT.

Il est impossible d'exécuter une tâche du profileur de la colonne avec une tâche du profileur de la relation.

26.1.4.1 Soumission des tâches du profileur de colonnes

1. Dans la bibliothèque d'objets de Designer, vous pouvez sélectionner une table ou un fichier plat.

Pour une table, accédez à l'onglet [Banques de données](#) et sélectionnez une table. Pour profiler toutes les tables d'une banque de données, sélectionnez le nom de celle-ci. Pour sélectionner un sous-ensemble de tables dans l'onglet [Banques de données](#), maintenez la touche Ctrl enfoncée en sélectionnant chaque table.

Pour un fichier plat, accédez à l'onglet [Formats](#) et sélectionnez un fichier.

2. Après avoir sélectionné la source de données, vous pouvez générer des statistiques de profil de colonne selon une des méthodes suivantes :

- Cliquez avec le bouton droit et sélectionnez *Soumettre la demande de profil de la colonne*.
Le calcul de certaines statistiques de profil peut durer longtemps. Sélectionnez cette méthode pour que la tâche de profilage s'exécute de manière asynchrone et que vous puissiez réaliser d'autres tâches de Designer pendant l'exécution de la tâche de profilage.
Cette méthode permet également de profiler plusieurs sources en une seule tâche du profileur.
- Cliquez avec le bouton droit de la souris, sélectionnez *Afficher les données*, cliquez sur l'onglet *Profil* puis sur *Mettre à jour*. Cette option soumet une tâche de profilage synchrone et vous devez attendre qu'elle soit terminée pour pouvoir réaliser d'autres tâches dans Designer.
Vous pouvez éventuellement utiliser cette option si vous êtes déjà dans la fenêtre *Afficher les données* et que vous remarquez que les statistiques de profil n'ont pas encore été générées ou que la date à laquelle elles l'ont été est antérieure à vos besoins.

3. (Facultatif) Modifiez le nom de la tâche du profileur.

Le profileur de données génère un nom par défaut pour chaque tâche du profileur. Vous pouvez modifier le nom de la tâche afin de créer un nom plus significatif, un nom unique ou supprimer les tirets qui sont autorisés dans les noms des colonnes, mais pas dans ceux des tâches.

Si vous sélectionnez une source unique, le format du nom par défaut est le suivant :

`<nomutilisateur_t_nomsources>`

Si vous sélectionnez plusieurs sources, le format du nom par défaut est le suivant :

`<nomutilisateur_t_nompremièresource_nomdernièresource>`

Colonne	Description
nomutilisateur	Nom d'utilisateur utilisé par le progiciel pour accéder aux services système.
t	Type de profil. La valeur est C pour un profil de colonne qui obtient des attributs (tels que les valeurs inférieure et supérieure) pour chaque colonne sélectionnée.
nompremièresource	Nom de la première source par ordre alphabétique.
nomdernièresource	Nom de la dernière source par ordre alphabétique si vous en sélectionnez plusieurs.

4. Si vous sélectionnez une source, la fenêtre *Soumettre la demande de profil de la colonne* liste les colonnes et les types de données.
- Conservez la coche devant chaque colonne à profiler et supprimez la coche devant chaque colonne que vous ne désirez pas profiler.
- Vous pouvez aussi cliquer dans la case à cocher située en haut devant *Nom* pour désélectionner toutes les colonnes, puis cocher les cases.
5. Si vous avez sélectionné plusieurs sources, la fenêtre *Soumettre la demande de profileur de la colonne* liste les sources à gauche.
- a. Sélectionnez une source pour afficher ses colonnes du côté droit.
 - b. Du côté droit de la fenêtre *Soumettre la demande de profil de la colonne*, conservez la coche devant chaque colonne à profiler et supprimez la coche devant chaque colonne que vous ne souhaitez pas profiler.

Autrement, vous pouvez cliquer dans la case à cocher placée en haut devant *Nom* pour désélectionner toutes les colonnes, puis cocher chaque case des colonnes à profiler.
 - c. Répétez les étapes 1 et 2 pour chaque source de données.
6. (Facultatif) Sélectionnez *Profilage détaillé* pour une colonne.

i Remarque

Le profileur de données consomme beaucoup de ressources lorsqu'il génère des statistiques de profil détaillées. Choisissez le profilage détaillé uniquement si vous avez besoin de ces attributs : comptage distinct, pourcentage distinct, valeur médiane, longueur de chaîne médiane, modèle, nombre de modèles. Si vous avez sélectionné le profilage détaillé, vérifiez que le répertoire de cache paginable indiqué contient suffisamment d'espace disque pour la taille des données que vous profilez.

Si vous voulez des attributs détaillés pour toutes les colonnes de toutes les sources listées, cliquez sur [Profilage détaillé](#) et sélectionnez [Appliquer](#) pour toutes les colonnes de toutes les sources.

Si vous voulez supprimer le profilage détaillé pour toutes les colonnes, cliquez sur [Profilage détaillé](#) et sélectionnez [Supprimer](#) pour toutes les colonnes de toutes les sources.

7. Cliquez sur [Soumettre](#) pour exécuter la tâche de profilage.

i Remarque

Si les métadonnées de la table ont été modifiées depuis que vous les avez importées (par exemple, ajout d'une colonne), vous devez réimporter la table source avant d'exécuter la tâche du profileur.

Si vous avez cliqué sur l'option [Soumettre la demande de profil de la colonne](#) pour atteindre cette fenêtre [Soumettre la demande de profileur de la colonne](#), le volet Moniteur du profileur s'affiche automatiquement lorsque vous cliquez sur [Soumettre](#).

Si vous avez cliqué sur [Mettre à jour](#) dans l'onglet [Profil](#) de la fenêtre [Afficher les données](#), la fenêtre du moniteur [Profileur](#) ne s'affiche pas lorsque vous cliquez sur [Soumettre](#). À la place, une tâche de profilage est soumise de manière asynchrone et vous devez attendre qu'elle soit terminée pour pouvoir effectuer d'autres tâches dans Designer.

Vous pouvez également surveiller votre tâche de profileur d'après son nom dans l'Administrateur.

8. Quand la tâche du profileur est terminée, vous pouvez afficher les résultats de profil dans l'option [Afficher les données](#).

Informations associées

[Profil de colonne \[page 387\]](#)

[Surveillance des tâches du profileur à l'aide de Designer \[page 395\]](#)

[Affichage des résultats du profileur \[page 396\]](#)

26.1.4.2 Soumission des tâches du profileur de relations

Un profil de relation indique le pourcentage des valeurs non correspondantes dans les colonnes de deux sources. Les sources peuvent être un des éléments suivants :

- Tables
- Fichiers plats

- Combinaison d'une table et d'un fichier plat

Une relation de clé primaire et de clé étrangère peut être définie pour les colonnes ou bien elles peuvent être sans relation (comme quand l'une vient d'une banque de données et l'autre d'un format de fichier).

Les deux colonnes ne doivent pas obligatoirement être du même type, mais elles doivent être convertibles. Par exemple, si vous exécutez une tâche du profileur de relations sur une colonne Integer et une colonne Varchar, le profileur de données convertit les valeurs des entiers en valeur de caractères (Varchar) pour effectuer la comparaison.

i Remarque

Le profileur de données consomme beaucoup de ressources lorsqu'il génère des valeurs de relations. Si vous prévoyez d'utiliser le profilage de relations, vérifiez que le répertoire de cache paginable indiqué contient suffisamment d'espace disque pour la taille des données que vous profilez.

Informations associées

[Sources de données à profiler \[page 386\]](#)

26.1.4.2.1 Génération d'un profil de relation pour les colonnes de deux sources

1. Dans la bibliothèque d'objets de Designer, sélectionnez deux sources.

Pour sélectionner deux sources dans la même banque de données ou le même format de fichier :

- a. Accédez à l'onglet *Banque de données* ou *Format* dans la bibliothèque d'objets.
- b. Maintenez la touche Ctrl enfoncée en sélectionnant la deuxième table.
- c. Cliquez avec le bouton droit et sélectionnez *Soumettre la demande de profil de relation*.

Pour sélectionner deux sources de banques de données ou formats de fichier différents :

- a. Accédez à *Banque de données* ou *Format* dans la bibliothèque d'objets.
- b. Cliquez avec le bouton droit sur la première source, sélectionnez ► *Soumettre la demande de profil de la relation* ► *Relation avec* ►
- c. Changez de banque de données ou de format dans la bibliothèque d'objets
- d. Cliquez sur la deuxième source.

La fenêtre *Soumettre la demande de profil de la relation* s'affiche.

i Remarque

Il est impossible de créer un profil de relation pour les colonnes dont les données sont de type LONG ou TEXTE.

2. (Facultatif) Modifiez le nom de la tâche du profileur.

Vous pouvez modifier le nom de la tâche afin de créer un nom plus significatif, un nom unique ou supprimer les tirets qui sont autorisés dans les noms des colonnes, mais pas dans ceux des tâches. Le nom par défaut généré par le profileur de données pour plusieurs sources a le format suivant :

`<nomutilisateur>_<t>_<nompremièresource>_<nomdernièresource>`

Colonne	Description
nomutilisateur	Nom d'utilisateur utilisé par le progiciel pour accéder aux services système.
t	Type de profil. La valeur est R pour le profil de relation qui obtient des valeurs sans correspondance dans les deux colonnes sélectionnées.
nompremièresource	Nom de la première source sélectionnée.
nomdernièresource	Nom de la dernière source sélectionnée.

3. Par défaut, s'il existe une relation, le volet supérieur de la fenêtre *Soumettre la demande de profil de la relation* affiche une ligne entre la colonne de clé primaire et la colonne de clé étrangère des deux sources. Vous pouvez changer les colonnes à profiler.

La moitié inférieure de la fenêtre *Soumettre la demande de profil de la relation* montre que la tâche du profileur va utiliser l'opération égale (=) pour comparer les deux colonnes. Le profileur de données va déterminer les valeurs qui ne sont pas égales et calculer le pourcentage de valeurs non correspondantes.

4. Pour supprimer une relation existante entre deux colonnes, sélectionnez la ligne, cliquez avec le bouton droit et sélectionnez *Supprimer la relation sélectionnée*.

Pour supprimer toutes les relations existantes entre deux sources, exécutez une des actions suivantes :

- Cliquez avec le bouton droit dans le volet supérieur et cliquez sur *Supprimer toutes les relations*.
- Cliquez sur *Supprimer toutes les relations* près du bas de la fenêtre *Soumettre la demande de profil de la relation*.

5. S'il n'existe pas de relation de clé primaire et clé étrangère entre les deux sources de données, indiquez les colonnes à profiler. Vous pouvez redimensionner chaque source de données pour afficher toutes les colonnes.

Pour indiquer ou modifier les colonnes dont vous voulez voir les valeurs de relation :

- a. Déplacez le curseur sur la première colonne à sélectionner. Maintenez le curseur enfoncé et dessinez une ligne vers l'autre colonne à sélectionner.
- b. Si vous avez supprimé toutes les relations et que vous voulez que le profileur de données sélectionne une relation de clé primaire et clé étrangère existante, cliquez avec le bouton droit dans le volet supérieur et cliquez sur *Proposer une relation*, ou bien cliquez sur *Proposer une relation* près du bas de la fenêtre *Soumettre la demande de profil de la relation*.

6. L'option *Enregistrer uniquement les données de colonnes clés* est sélectionnée par défaut. Elle indique que le profileur de données enregistre les données uniquement dans les colonnes sélectionnées pour la relation. Aucun échantillon de données n'est visible dans les autres colonnes lorsque vous affichez le profil de la relation.

Pour voir les valeurs des autres colonnes dans le profil de la relation, sélectionnez l'option *Enregistrer les données de toutes les colonnes*.

7. Cliquez sur *Soumettre* pour exécuter la tâche du profileur.

i Remarque

Si les métadonnées de la table ont été modifiées depuis que vous les avez importées (par exemple, ajout d'une colonne), vous devez réimporter la table source avant d'exécuter la tâche du profileur.

8. Le volet du moniteur du profileur s'affiche automatiquement lorsque vous cliquez sur [Soumettre](#).
Vous pouvez également surveiller votre tâche de profileur d'après son nom dans l'Administrateur.
9. Lorsque la tâche du profileur est terminée, vous pouvez afficher les résultats du profil dans l'option [Afficher les données](#) en cliquant avec le bouton droit sur une table dans la bibliothèque d'objets.

Informations associées

[Affichage des données de profil de la relation générées par le profileur de données \[page 400\]](#)

[Surveillance des tâches du profileur à l'aide de Designer \[page 395\]](#)

[Affichage des résultats du profileur \[page 396\]](#)

26.1.5 Surveillance des tâches du profileur à l'aide de Designer

La fenêtre du moniteur du [Profileur](#) s'affiche automatiquement lorsque vous soumettez une tâche du profileur si vous avez cliqué sur la barre de menu pour afficher la fenêtre du moniteur du [Profileur](#). Vous pouvez ancrer ce volet du moniteur du profileur dans Designer ou le laisser séparé.

Le volet du moniteur du profileur affiche la tâche en cours d'exécution et toutes les tâches du profileur ayant été exécutées durant un nombre de jours configuré.

- Cliquez sur l'icône [Actualiser](#)
- Cliquez sur l'icône [Informations](#) pour afficher les sources que la tâche sélectionnée est en train de profiler.
Si la tâche a échoué, la fenêtre Informations affiche également un message d'erreur.

Le moniteur du profileur affiche les colonnes suivantes : Les icônes situées dans le coin supérieur gauche du moniteur du profileur affichent les informations suivantes :

Colonne	Description
Nom	Nom de la tâche du profileur envoyée depuis Designer. Les icônes dans le coin supérieur gauche du profileur Si la tâche du profileur s'applique à une seule source, le nom par défaut a le format suivant : <pre><username_t_sourcename></pre> Si la tâche du profileur s'applique à plusieurs sources, le nom par défaut a le format suivant : <pre><username_t_firstsourcename_lastsourcename></pre>

Colonne	Description
Type	Le type de la tâche du profileur peut être : • icône pour actualiser le volet Moniteur du profileur et afficher le dernier statut des tâches du profileur. pour actualiser le volet Moniteur du profileur et afficher le dernier statut des tâches du profileur. • Relation
Statut	Le statut d'une tâche du profileur peut être : • Terminé : la tâche s'est terminée correctement. • En attente : la tâche est dans la file d'attente car le nombre maximal de tâches simultanées a été atteint ou qu'une autre tâche profile la même table. • En cours d'exécution : la tâche est en cours d'exécution. • Erreur : la tâche s'est terminée avec une erreur. Double-cliquez sur la valeur de cette colonne Statut pour afficher le message d'erreur.
Horodatage	Date et heure d'exécution de la tâche du profileur.
Sources	Noms des tables sur lesquelles s'exécute la tâche du profileur.

Informations associées

[Exécution d'une tâche du profileur \[page 390\]](#)

26.1.6 Affichage des résultats du profileur

Le profileur de données calcule et enregistre les attributs du profileur dans un référentiel du profileur pouvant être affiché par plusieurs utilisateurs.

Informations associées

[Affichage des attributs de colonne générés par le profileur de données \[page 397\]](#)

[Affichage des données de profil de la relation générées par le profileur de données \[page 400\]](#)

26.1.6.1 Affichage des attributs de colonne générés par le profileur de données

1. Dans la bibliothèque d'objets, sélectionnez la table dont vous voulez afficher les attributs du profileur.
2. Cliquez avec le bouton droit et sélectionnez [Afficher les données](#).
3. Cliquez sur l'onglet [Profil](#) (deuxième) pour afficher les attributs de profil de colonne.
 - a. L'onglet [Profil](#) affiche le nombre d'enregistrement physiques traités par le profileur de données pour générer les valeurs de la grille de profil.
 - b. La grille de profil contient les noms de colonne de la source actuelle et les attributs de profil de chaque colonne. Pour remplir la grille de profil, exécutez une tâche du profileur ou sélectionnez les noms dans cette colonne et cliquez sur [Mettre à jour](#).
 - c. Vous pouvez trier les valeurs de chaque colonne d'attribut en cliquant sur leurs en-têtes. La valeur n/a dans la grille de profil indique qu'un attribut ne s'applique pas à un type de données.

Attribut de profil de base	Description	Type de données pertinent		
		Caractère	Numérique	Date/heure
Min.	Valeur la plus basse dans cette colonne.	Oui	Oui	Oui
Nombre min.	Nombre de lignes qui contiennent la valeur la plus basse dans cette colonne.	Oui	Oui	Oui
Max.	Valeur la plus élevée dans cette colonne.	Oui	Oui	Oui
Nombre max.	Nombre de lignes qui contiennent la valeur la plus élevée dans cette colonne.	Oui	Oui	Oui
Moyenne	Pour les colonnes numériques, valeur moyenne dans cette colonne.	n/a	Oui	Sans objet
Longueur minimale de la chaîne	Pour les colonnes avec des caractères, longueur de la valeur de chaîne la plus courte dans cette colonne.	Oui	Non	Non
Longueur maximale de la chaîne	Pour les colonnes avec des caractères, longueur de la valeur de chaîne la plus longue dans cette colonne.	Oui	Non	Non

Attribut de profil de base	Description	Type de données pertinent		
		Caractère	Numérique	Date/heure
Longueur moyenne de la chaîne	Pour les colonnes avec des caractères, longueur moyenne des valeurs de chaîne dans cette colonne.	Oui	Non	Non
Valeurs nulles	Nombre de valeurs NULL dans cette colonne.	Oui	Oui	Oui
% de valeurs nulles	Pourcentage de lignes contenant une valeur NULL dans cette colonne.	Oui	Oui	Oui
Zéros	Nombre de valeurs 0 dans cette colonne.	Non	Oui	Non
% de zéros	Pourcentage de lignes contenant une valeur 0 dans cette colonne.	Non	Oui	Non
Vides	Pour les colonnes avec des caractères, nombre de lignes qui contiennent un vide dans cette colonne.	Oui	Non	Non
% de vides	Pourcentage de lignes contenant un vide dans cette colonne.	Oui	Non	Non

- d. Si vous avez sélectionné l'option *Profilage détaillé* dans la fenêtre *Soumettre la demande de profileur de la colonne*, l'onglet *Profil* affiche également les colonnes d'attribut détaillées suivantes.

Attribut de profil détaillé	Description	Type de données pertinent		
		Caractère	Numérique	Date/heure
Distinctes	Nombre de valeurs distinctes dans cette colonne.	Oui	Oui	Oui
% de distinctes	Pourcentage de lignes contenant chaque valeur distincte dans cette colonne.	Oui	Oui	Oui
Médiane	Valeur qui se trouve dans la ligne du milieu de la table source.	Oui	Oui	Oui

Attribut de profil détaillé	Description	Type de données pertinent		
		Caractère	Numérique	Date/heure
Longueur médiane de la chaîne	Pour les colonnes avec des caractères, valeur qui se trouve dans la ligne du milieu de la table source.	Oui	Non	Non
% de modèle	Pourcentage de lignes contenant chaque valeur distincte dans cette colonne. Format de chaque modèle unique de cette colonne.	Oui	Non	Non
Modèles	Nombre de modèles différents dans cette colonne.	Oui	Non	Non

4. Cliquez sur une valeur d'attribut pour afficher la ligne entière de la table source. La moitié inférieure de la fenêtre [Afficher les données](#) affiche les lignes qui contiennent la valeur d'attribut sur laquelle vous avez cliqué. Vous pouvez masquer les colonnes que vous ne désirez pas afficher en cliquant sur l'icône Afficher/masquer les colonnes.

Par exemple, votre colonne cible ADDRESS comporte seulement 45 caractères, mais les données de profilage de cette table source Client affichent que la longueur maximale de chaîne est 46. Cliquez sur la valeur 46 pour afficher les données réelles. Vous pouvez redimensionner la largeur de la colonne pour afficher la chaîne entière.

5. (Facultatif) Cliquez sur [Mettre à jour](#) pour mettre à jour les attributs de profil. Raisons d'une mise à jour à ce moment :
 - Les attributs de profil n'ont pas encore été générés.
 - La date de génération des attributs de profil est antérieure à celle voulue. La valeur Date de dernière mise à jour dans l'angle inférieur gauche de l'onglet [Profil](#) est l'horodatage de la dernière génération des profils d'attribut.

i Remarque

L'option Mettre à jour soumet une tâche synchrone du profileur et vous devez attendre qu'elle soit terminée pour pouvoir exécuter d'autres tâches dans Designer.

La fenêtre [Soumettre la demande de profil de la colonne](#) s'affiche.

Comme les calculs de mise à jour ont un impact sur les performances, sélectionnez uniquement les noms des colonnes dont vous avez besoin pour cette opération de profilage. Vous pouvez également cliquer dans la case à cocher située en haut devant Nom pour désélectionner toutes les colonnes, puis cocher la case placée devant chaque colonne à profiler.

6. Cliquez sur une statistique dans Distinctes ou Modèles pour afficher le pourcentage de chaque valeur distincte ou valeur de modèle dans une colonne. Les valeurs de modèle, le nombre d'enregistrements de chaque valeur de modèle et les pourcentages s'affichent à droite de l'onglet [Profil](#).

Par exemple, l'onglet *Profil* suivant pour la table CUSTOMERS affiche les attributs de profil pour la colonne REGION. L'attribut Distincts de la colonne REGION affiche la statistique 19, c'est-à-dire qu'il existe 19 valeurs distinctes pour REGION.

Column	Median string length	Distincts	Nulls	Nulls %	Zeros	Zeros %	Blanks	Blanks %	Profiling data (REGION_Distincts)
CUSTOMERID	5	91	0	0%	n/a	n/a	0	0%	Value
COMPANYNAME	18	91	0	0%	n/a	n/a	0	0%	<Null>
CONTACTNAME	14	91	0	0%	n/a	n/a	0	0%	SP
CONTACTTITLE	17	12	0	0%	n/a	n/a	0	0%	OR
ADDRESS	18	91	0	0%	n/a	n/a	0	0%	WA
CITY	8	69	0	0%	n/a	n/a	0	0%	BJ
REGION	2	19	60	65.93%	n/a	n/a	0	0%	BC
POSTALCODE	5	87	1	1.09%	n/a	n/a	0	0%	Táchira
COUNTRY	6	21	0	0%	n/a	n/a	0	0%	Québec
PHONE	12	91	0	0%	n/a	n/a	0	0%	Nueva Esparta
FAX	12	70	22	24.17%	n/a	n/a	0	0%	NM
									MT

Value	Records	% of total
<Null>	60	65.93%
SP	6	6.59%
OR	4	4.40%
WA	3	3.30%
BJ	3	3.30%
BC	2	2.20%
Táchira	1	1.10%
Québec	1	1.10%
Nueva Esparta	1	1.10%
NM	1	1.10%
MT	1	1.10%

CUSTOMERID	COMPANYNAME	CONTACTNAME	CONTACTTITLE	ADDRESS	CITY	REGION
ALFKI	Alfreds Futterkiste	Maria Anders	Sales Representative	Obere Str. 57	Berlin	<Null>
ANATR	Ana Trujillo Emparedados y...	Ana Trujillo	Owner	Avda. de la Constitución 2...	México D.F.	<Null>
ANTON	Antonio Moreno Taquería	Antonio Moreno	Owner	Mataderos 2312	México D.F.	<Null>

7. Cliquez sur la statistique de la colonne Distincts pour afficher chacune des 19 valeurs et le pourcentage de lignes de la table CUSTOMERS ayant cette valeur pour la colonne REGION. En outre, les barres de la colonne la plus à droite affichent la taille relative de chaque pourcentage.
8. Les données de profilage du côté droit montrent qu'un important pourcentage des valeurs de REGION est Null. Sous Valeur, cliquez sur Null ou, sous Enregistrements, cliquez sur 60 pour afficher les autres colonnes dans les lignes qui comportent une valeur Null dans la colonne REGION.
9. Vos règles de gestion peuvent indiquer que REGION ne doit pas contenir de valeur NULL dans votre entrepôt de données cible. Décidez donc quelle valeur doit remplacer les valeurs NULL lorsque vous définissez une transformation Validation.

Informations associées

[Exécution d'une tâche du profileur \[page 390\]](#)

[Définition d'une règle de validation en fonction du profil de colonne \[page 406\]](#)

26.1.6.2 Affichage des données de profil de la relation générées par le profileur de données

Les données de profil de relation affichent le pourcentage de valeurs sans correspondance dans les colonnes de deux sources. Les sources peuvent être des tables, des fichiers plats ou une combinaison d'une table et d'un fichier plat. Une relation de clé primaire et de clé étrangère peut être définie pour les colonnes ou bien elles

peuvent être sans relation (comme quand l'une vient d'une banque de données et l'autre d'un format de fichier).

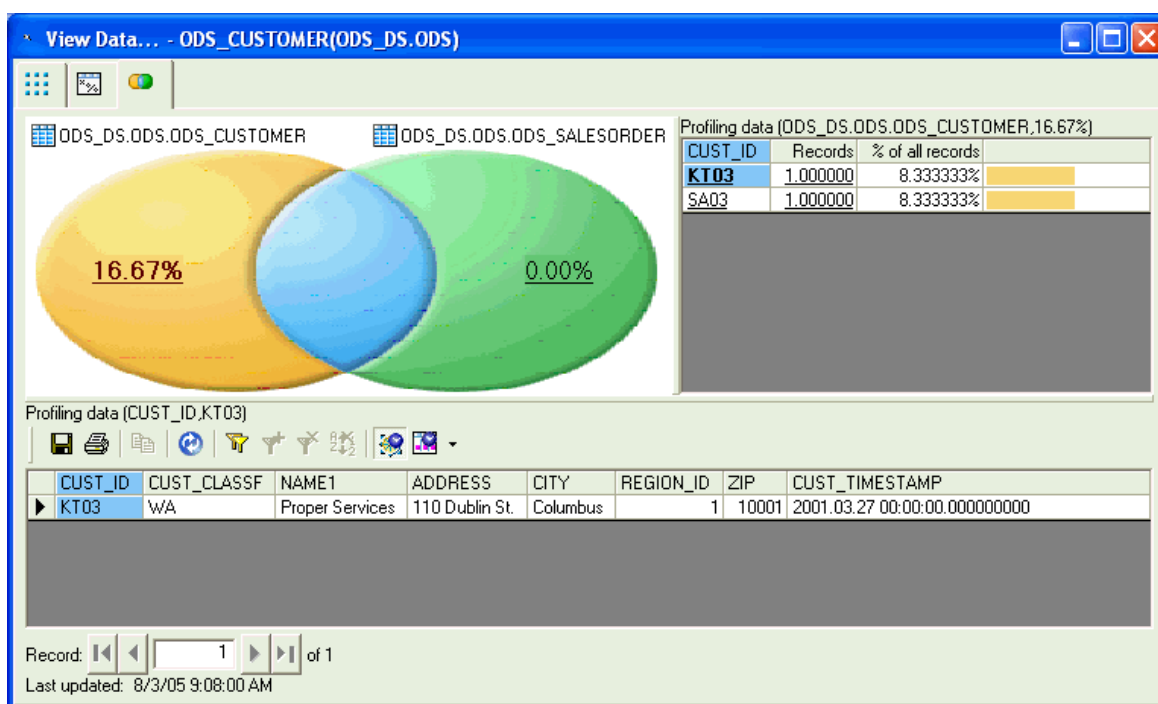
1. Dans la bibliothèque d'objets, sélectionnez la table ou le fichier dont vous voulez afficher les données de profil de relation.
2. Cliquez avec le bouton droit et sélectionnez *Afficher les données*.
3. Cliquez sur l'onglet *Relation* (troisième) pour afficher les résultats du profil de relation.

i Remarque

L'onglet *Relation* n'est visible que si vous avez exécuté une tâche du profileur de relations.

4. Cliquez sur le pourcentage non nul dans le diagramme pour afficher les valeurs de clé non contenues dans l'autre table.

Par exemple, l'onglet Relation ci-dessous de Afficher les données montre le pourcentage de clients (16,67) n'ayant pas de commande. Le profil de la relation a été défini sur la colonne CUST_ID de la table ODS_CUSTOMER et la colonne CUST_ID de la table ODS_SALESORDER. La valeur de l'ovale de gauche indique que 16,67 % des lignes de la table ODS_CUSTOMER ont des valeurs de CUST_ID qui n'existent pas dans la table ODS_SALESORDER.



Cliquez sur le pourcentage 16,67 dans l'ovale ODS_CUSTOMER pour afficher les valeurs de CUST_ID qui n'existent pas dans la table ODS_SALESORDER. Les valeurs non correspondantes KT03 et SA03 sont affichées sur le côté droit de l'onglet relation. Chaque ligne affiche une valeur de CUST_ID non correspondante, le nombre d'enregistrements comportant cette valeur de CUST_ID et le pourcentage total de clients comportant cette valeur de CUST_ID.

5. Cliquez sur l'une des valeurs du côté droit pour afficher les autres colonnes des lignes qui contiennent cette valeur.

La moitié inférieure de l'onglet *Profil de relation* affiche les valeurs des autres colonnes de la ligne comportant la valeur KT03 dans la colonne CUST_ID.

i Remarque

Si vous n'avez pas sélectionné l'option Enregistrer les données de toutes les colonnes dans la fenêtre *Soumettre la demande de profil de la relation*, vous ne pouvez pas afficher les données des autres colonnes.

Informations associées

[Soumission des tâches du profileur de relations \[page 392\]](#)

26.2 Utilisation de la fonction d'affichage des données pour déterminer la qualité des données

Utilisez l'affichage des données pour vous aider à déterminer la qualité des données source et cible. L'affichage des données permet :

- d'afficher des échantillons de données source avant d'exécuter un job afin de créer des conceptions de job de qualité supérieure ;
- de comparer des échantillons de données de différentes étapes de votre job pour vérifier que le job d'extraction des données renvoie les résultats attendus.

i Remarque

Pour utiliser la fonction d'affichage des données sur un fichier, celui-ci doit être accessible depuis Designer.

Informations associées

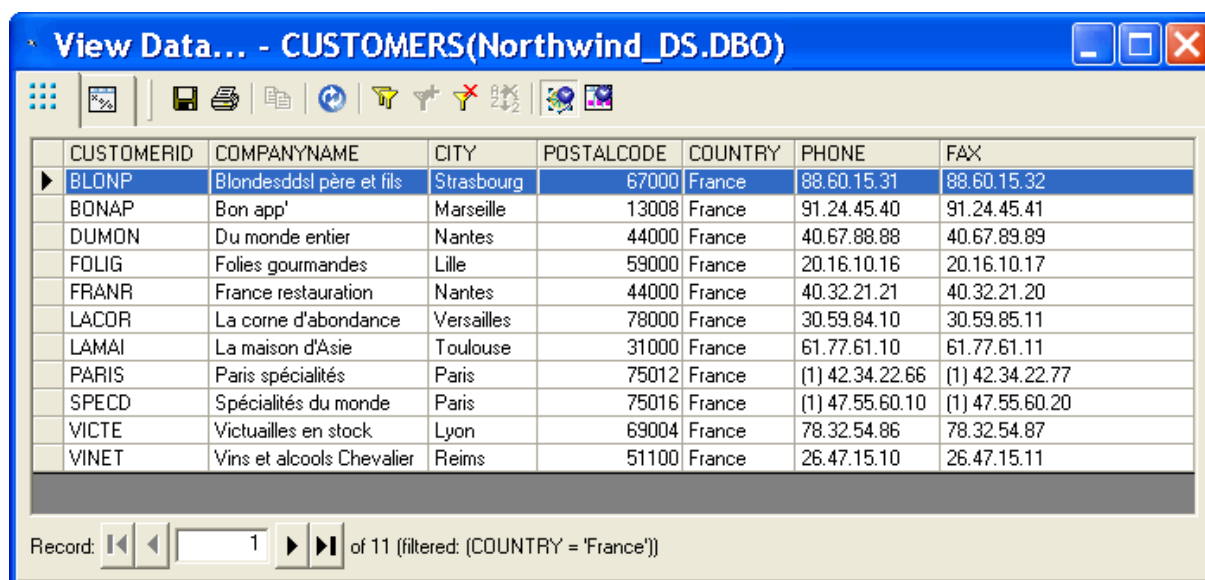
[Définition d'une règle de validation en fonction du profil de colonne \[page 406\]](#)

[Utilisation de l'affichage des données \[page 727\]](#)

26.2.1 Onglet Données

L'onglet *Données* est toujours disponible et affiche le contenu des données de modèles de lignes. Vous pouvez afficher un sous-ensemble de colonnes de chaque ligne et définir des filtres pour afficher un sous-ensemble de lignes.

Par exemple, vos règles de gestion peuvent imposer que tous les numéros de téléphone et de fax aient un seul format pour chaque pays. L'onglet *Données* ci-dessous affiche un sous-ensemble de lignes des clients résidant en France.



The screenshot shows a data viewer window titled 'View Data... - CUSTOMERS(Northwind_DS.DBO)'. It displays a table with 8 columns: CUSTOMERID, COMPANYNAME, CITY, POSTALCODE, COUNTRY, PHONE, and FAX. The data is filtered to show only customers from France. The status bar at the bottom indicates 'Record: 1 of 11 (filtered: (COUNTRY = 'France'))'.

CUSTOMERID	COMPANYNAME	CITY	POSTALCODE	COUNTRY	PHONE	FAX
BLONP	Blondesdds! père et fils	Strasbourg	67000	France	88.60.15.31	88.60.15.32
BONAP	Bon app'	Marseille	13008	France	91.24.45.40	91.24.45.41
DUMON	Du monde entier	Nantes	44000	France	40.67.88.88	40.67.89.89
FOLIG	Folies gourmandes	Lille	59000	France	20.16.10.16	20.16.10.17
FRANR	France restauration	Nantes	44000	France	40.32.21.21	40.32.21.20
LACOR	La corne d'abondance	Versailles	78000	France	30.59.84.10	30.59.85.11
LAMAI	La maison d'Asie	Toulouse	31000	France	61.77.61.10	61.77.61.11
PARIS	Paris spécialités	Paris	75012	France	(1) 42.34.22.66	(1) 42.34.22.77
SPECD	Spécialités du monde	Paris	75016	France	(1) 47.55.60.10	(1) 47.55.60.20
VICTE	Victuailles en stock	Lyon	69004	France	78.32.54.86	78.32.54.87
VINET	Vins et alcools Chevalier	Reims	51100	France	26.47.15.10	26.47.15.11

Notez que les colonnes PHONE et FAX affichent des valeurs ayant deux formats différents. Vous pouvez désormais décider du format à utiliser dans l'entrepôt de données cible et définir une transformation Validation en conséquence.

Informations associées

[Propriétés de l'affichage des données \[page 730\]](#)

[Définition d'une règle de validation en fonction du profil de colonne \[page 406\]](#)

[Onglet Données \[page 735\]](#)

26.2.2 Onglet Profil

Deux affichages sont disponibles sur l'onglet *Profil* :

- Sans le Profileur de données, l'onglet *Profil* affiche les attributs de colonnes suivants : valeurs distinctes, valeurs nulles, valeur minimale et valeur maximale.
- Si vous avez configuré et que vous utilisez le Profileur de données, l'onglet *Profil* affiche les mêmes attributs de colonnes ci-dessus, ainsi que de nombreuses autres statistiques calculées, telles que la valeur moyenne, la longueur minimale de la chaîne et la longueur maximale de la chaîne, le nombre distinct, le pourcentage distinct, la médiane, la longueur médiane de la chaîne, le nombre du modèle et le pourcentage du modèle.

Informations associées

[Onglet Profil \[page 736\]](#)

[Affichage des attributs de colonne générés par le profileur de données \[page 397\]](#)

26.2.3 Onglet Profil de relation ou Profil de colonne

Le troisième onglet affiché dépend de si vous avez configuré et que vous utilisez le Profileur de données ou non.

- Si vous n'utilisez pas le Profileur de données, l'onglet [Profil de colonne](#) permet de calculer les informations statistiques pour une seule colonne.
- Si vous utilisez le Profileur de données, l'onglet [Relation](#) affiche les incohérences de données entre les deux colonnes à partir desquelles il est possible de déterminer l'intégrité des données entre les deux sources.

Informations associées

[Onglet Profil de colonne \[page 737\]](#)

[Affichage des données de profil de la relation générées par le profileur de données \[page 400\]](#)

26.3 Utilisation de la transformation Validation

Les fonctionnalités Profileur de données et Afficher les données peuvent identifier les anomalies dans les données entrantes. Il est alors possible d'utiliser une transformation Validation pour définir les règles qui séparent les bonnes données des mauvaises. Il est possible d'écrire les mauvaises données dans une table ou un fichier pour une révision ultérieure.

Pour en savoir plus sur la transformation Validation, notamment sur la manière d'implémenter les fonctions réutilisables de validation, voir le *Guide de référence SAP Data Services*.

26.3.1 Analyse du profil de colonne

Il est possible d'obtenir des informations sur le profil de colonne en soumettant des tâches du profileur de colonne.

Par exemple, si vous souhaitez analyser les données dans la table Clients dans l'exemple de base de données Northwinds sur Microsoft SQL Server.

Informations associées

[Soumission des tâches du profileur de colonnes \[page 390\]](#)

26.3.1.1 Analyse des attributs de profilage de colonne

1. Dans la bibliothèque d'objets, cliquez sur la table profilée Clients et sélectionnez [Afficher les données](#).
2. Sélectionnez l'onglet [Profil](#) dans la fenêtre [Afficher les données](#). L'onglet Profil affiche les attributs de profil de colonne indiqués sur l'illustration suivante.

The screenshot shows the 'View Data... - CUSTOMERS(Northwind_DS.DBO)' window. The 'Profil' tab is active, displaying a table of column statistics and a detailed view of the 'PHONE' column's profiling data.

Column	Zeros	Zeros %	Blanks	Blanks %	Patterns
CUSTOMERID	n/a	n/a	0	0%	1
COMPANYNAME	n/a	n/a	0	0%	82
CONTACTNAME	n/a	n/a	0	0%	42
CONTACTTITLE	n/a	n/a	0	0%	12
ADDRESS	n/a	n/a	0	0%	82
CITY	n/a	n/a	0	0%	18
REGION	n/a	n/a	0	0%	7
POSTALCODE	n/a	n/a	0	0%	9
COUNTRY	n/a	n/a	0	0%	8
PHONE	n/a	n/a	0	0%	20
FAX	n/a	n/a	0	0%	20

Profiling data (PHONE.Patterns)

Value	Records	% of total
(999) 999-9999	23	25.27%
(9) 999-9999	13	14.29%
(99) 999-9999	10	10.99%
99.99.99.99	9	9.89%
9999-999999	8	8.79%
(99) 999 99 99	4	4.40%
999-99999999	4	4.40%
(9) 99.99.99.99	2	2.20%
(99) 999 9999	2	2.20%
9999-999999	2	2.20%
99 99 99 99	2	2.20%
9999-9999	2	2.20%

Profiling data (Value,(9) 99.99.99.99)

CONTACTTITLE	ADDRESS	CITY	REGION	POSTALCODE	COUNTRY	PHONE	FAX
Owner	265, boulevard Charonne	Paris	<Null>	75012	France	(1) 42.34.22.66	(1) 42.34.22.77
Marketing Manager	25, rue Lauriston	Paris	<Null>	75016	France	(1) 47.55.60.10	(1) 47.55.60.20

Record: 1 of 2
Last updated: 9/8/05 6:20:00 AM

L'attribut Modèles pour la colonne PHONE indique la valeur 20, ce qui signifie qu'il existe 20 modèles différents.

3. Cliquez sur la valeur 20 sur la colonne de l'attribut [Modèles](#). Le volet [Données de profilage](#) affiche les modèles pour la colonne PHONE individuellement et le pourcentage de lignes pour chaque modèle.
4. Supposons que vos règles d'entreprise imposent que tous les numéros de téléphone en France doivent être au format 99.99.99.99. Cependant, les données de profilage indiquent que deux enregistrements sont au format (9) 99.99.99.99. Pour afficher les colonnes pour ces deux enregistrements dans le volet du bas, cliquez soit sur (9) 99.99.99.99 sous Valeur, soit sur 2 sous Enregistrement. Vous pouvez voir que certains numéros de téléphone en France ont le préfixe (1).

Il est possible d'utiliser une transformation Validation pour identifier les lignes contenant le préfixe non souhaité. Vous pouvez alors corriger les données pour vous conformer à vos règles d'entreprise, puis les recharger.

La section suivante décrit comment configurer la transformation Validation pour identifier les lignes erronée.

Informations associées

[Définition d'une règle de validation en fonction du profil de colonne \[page 406\]](#)

26.3.2 Définition d'une règle de validation en fonction du profil de colonne

Cette section utilise les résultats du Profileur de données et définit la transformation Validation selon l'exemple de règles d'entreprise. A partir de l'exemple précédent du préfixe téléphonique (1) pour les numéros de téléphone en France, la procédure suivante décrit la manière de définir un flux de données et une règle de validation qui identifient ce modèle. Il est alors possible de réviser les données ayant échoué, d'apporter des modifications et de recharger les données.

26.3.2.1 Définition de la règle de validation identifiant un modèle

Cette procédure décrit la façon de définir un flux de données et une règle de validation qui identifie des lignes contenant le préfixe (1) décrit à la section précédente.

1. Créez un flux de données avec la table de Clients comme source, ajoutez une transformation Validation et une cible et reliez les objets.
2. Ouvrez la transformation Validation en cliquant sur son nom.
3. Dans l'éditeur de transformation, cliquez sur [Ajouter](#).
La boîte de dialogue Editeur de règles s'affiche.
4. Saisissez un [Nom](#) et de manière facultative une [Description](#) pour la règle.
5. Vérifiez que la case [Actif](#) est cochée.
6. Pour [Action en cas d'échec](#), sélectionnez [Envoi vers échec](#).
7. Sélectionnez la case d'option [Validation de colonne](#).
 - a. Sélectionnez la [colonne](#) CUSTOMERS.PHONE dans la liste déroulante.
 - b. Pour [Condition](#), dans la liste déroulante, sélectionnez [Modèle de correspondance](#).
 - c. Pour la valeur, saisissez l'expression "[99.99.99.99](#)".
8. Cliquez sur [OK](#).
La règle s'affiche dans la liste de règles.

Après l'exécution du job, les lignes mal formatées s'affichent dans la sortie Fail. Vous pouvez à présent vérifier les données ayant échouées, effectuer des corrections le cas échéant et recharger les données.

Informations associées

[Analyse du profil de colonne \[page 404\]](#)

26.3.3 Editeur de règles

Vous utilisez l'Editeur de règles pour définir ou modifier une règle de validation. Lancez l'Editeur de règles à partir de l'onglet Règles de validation de la transformation Validation, en cliquant sur [Ajouter](#) ou en double-cliquant sur une règle existante pour la modifier.

L'Editeur de règles vous permet de configurer une règle soit en fonction d'une fonction de validation existante, soit en fonction d'une règle définie dans la transformation Validation basée sur une seule entrée de colonne.

Les règles définies avec la fonction de validation contiennent les caractéristiques suivantes :

- Réutilisable
- Accepte plusieurs paramètres d'entrée
- Préférable pour des règles plus complexes
- Peut être créée dans SAP Information Steward, importée et utilisée dans Data Services.

Les règles définies avec une validation de colonne contiennent les caractéristiques suivantes :

- Non réutilisable ; la définition de règle fait partie de la transformation (non partagée en dehors de la transformation)
- Uniquement liée à une colonne d'entrée
- Préférable pour les règles simples

Notez également qu'il existe deux types de fonctions de validation, catégorisées dans la bibliothèque d'objets :

- Importées à partir d'Information Steward : ces fonctions ont été créées dans Information Steward et ne peuvent pas être modifiées dans Data Services
- Créées localement : créées et modifiables dans Data Services

26.3.3.1 Définition d'une règle de validation en fonction d'une colonne

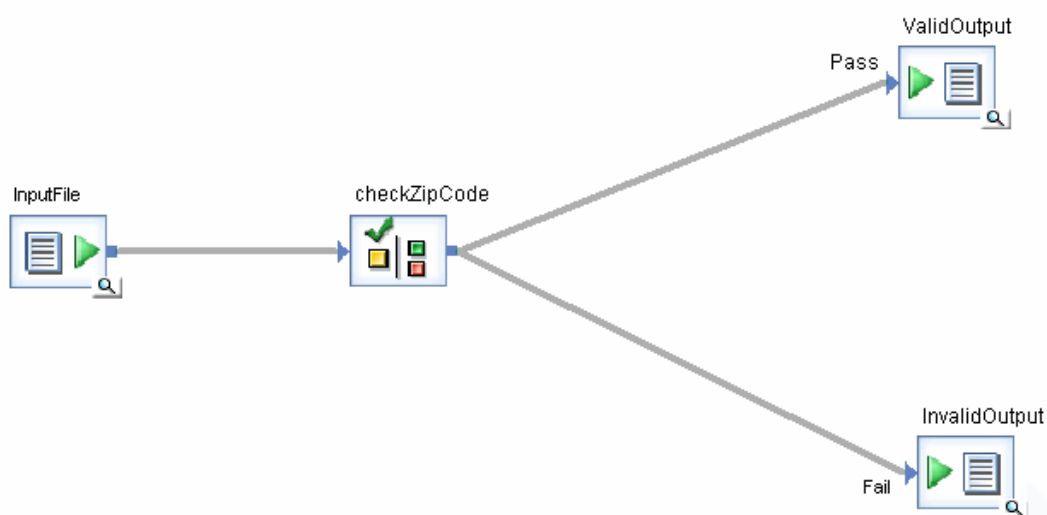
1. Dans un flux de données où la transformation Validation est connectée à un schéma d'entrée, cliquez sur le nom de la transformation pour ouvrir son éditeur.
2. Dans l'onglet Règles de validation, cliquez sur [Ajouter](#).
L'éditeur de règles s'affiche avec la case [Activée](#) cochée par défaut.
3. Saisissez un nom pour la règle et ajoutez éventuellement une description.
4. Sélectionnez une [Action en cas d'échec](#) :
 - Envoi vers échec
 - Envoi vers réussite
 - Envoi vers échec/réussite

- Sélectionnez *Validation de colonne*.
- Sélectionnez une *colonne* dans la liste déroulante.
- Définissez une condition. Toutes les conditions doivent être des expressions booléennes. Sinon, sélectionnez *Condition personnalisée* pour avoir accès à l'éditeur rapide (points de suspension) ou à l'assistant Fonction (bouton *Fonctions*).

❖ Exemple

Définition d'une règle de validation pour un modèle de code postal à cinq chiffres.

L'exemple suivant définit une transformation Validation qui vérifie que les données du code postal correspondent à un nombre à cinq chiffres.



- Ouvrez l'éditeur de la transformation Validation.
- Cliquez sur *Ajouter*.
- Nommez la règle **checkzip**.
- Pour *Action en cas d'échec*, sélectionnez *Envoi vers échec/réussite*.
- Sélectionnez *Validation de colonne*.
- Sélectionnez la colonne *Customers.Zipcode*.
- Sélectionnez la condition *LIKE*.
- Spécifiez comme modèle **99999**, qui correspond à un nombre quelconque à cinq chiffres.
- Cliquez sur *OK*.
- Dans l'onglet *Règles de validation*, pour l'option *Si une règle échoue et affiche Envoi vers réussite*, remplacez par **:**, double-cliquez sur une cellule dans *Colonne* et sélectionnez *Customers.Zipcode*.

11. Double-cliquez dans la cellule sous *Expression* et tapez '*Invalide*'.

Si le code postal ne correspondent pas au modèle, Data Services remplace la valeur par la chaîne *Invalide*.

L'illustration suivante montre les données d'entrée.

Q "C:\work\validation_data\demo\orderinfo2.txt"(OrderInfo)

	OrderNum	OrderDate	Name	Address1	Address2	City	State	Zipcode	Phone
▶	5001	2004.05.12	First National bank	1 First Street	NULL	San Jose	CA	95134	408-933-6821
	5002	NULL	Second National bank	123 First Street	NULL	San Jose	CA	95134	408-933-3456
	3500	2004.05.01	Second National bank	2 First Street	NULL	San Jose	CA	95134	408-933-6821
	5003	2004.05.12	First National bank	1 First Street	NULL	San Jose	CA	NULL	408-933-6821
	1	2004.06.12	First National bank	123 First Street	NULL	San Jose	CA	95134	408-933-6821
	5005	NULL	Third National bank	234 Commercial Street	NULL	San Francisco	CA	95100	650-933-3456
	5006	2004.05.16	First National bank	1 First Street	NULL	San Jose	CA	95134	408-933-6821
	5007	2004.05.16	First National bank	1 First Street	NULL	San Jose	CA	951ab	408-933-6821
	5008	2004.06.18	Third National bank	234 Commercial Street	NULL	San Francisco	CA	95100	650-933-3456
	5009	NULL	Third National bank	234 Commercial Street	NULL	San Francisco	CA	95100	650-933-3456
	5007	2004.05.16	First National bank	1 First Street	NULL	San Jose	CA	1234567	408-933-6821

L'illustration suivante montre les données envoyées à la table Pass/ValidOutput :

Q "C:\work\validation_data\demo\orderinfo2_out.txt"(OrderInfo)

	OrderNum	OrderDate	Name	Address1	Address2	City	State	Zipcode	Phone
▶	5001	2004.05.12	First National bank	1 First Street	NULL	San Jose	CA	95134	408-933-6821
	5002	NULL	Second National bank	123 First Street	NULL	San Jose	CA	95134	408-933-3456
	3500	2004.05.01	Second National bank	2 First Street	NULL	San Jose	CA	95134	408-933-6821
	5003	2004.05.12	First National bank	1 First Street	NULL	San Jose	CA	Invalid	408-933-6821
	1	2004.06.12	First National bank	123 First Street	NULL	San Jose	CA	95134	408-933-6821
	5005	NULL	Third National bank	234 Commercial Street	NULL	San Francisco	CA	95100	650-933-3456
	5006	2004.05.16	First National bank	1 First Street	NULL	San Jose	CA	95134	408-933-6821
	5007	2004.05.16	First National bank	1 First Street	NULL	San Jose	CA	Invalid	408-933-6821
	5008	2004.06.18	Third National bank	234 Commercial Street	NULL	San Francisco	CA	95100	650-933-3456
	5009	NULL	Third National bank	234 Commercial Street	NULL	San Francisco	CA	95100	650-933-3456
	5007	2004.05.16	First National bank	1 First Street	NULL	San Jose	CA	Invalid	408-933-6821

L'illustration suivante montre les données envoyées à la table Fail/InvalidOutput :

Q "C:\work\validation_data\demo\orderinfo2_failed.txt"(FailedOrderInfo)

	Ord...	OrderDate	Name	Address1	Addr...	City	State	Zipcode	Phone	D	DI_ERRORCOLUMNS
▶	5003	2004.05.12	First National bank	1 First Street	NULL	San Jose	CA	NULL	408-933-6821	B	checkZipCode failed rule(s): Zipcode
	5007	2004.05.16	First National bank	1 First Street	NULL	San Jose	CA	951ab	408-933-6821	B	checkZipCode failed rule(s): Zipcode
	5007	2004.05.16	First National bank	1 First Street	NULL	San Jose	CA	1234567	408-933-6821	B	checkZipCode failed rule(s): Zipcode

26.3.3.2 Définition d'une règle de validation à l'aide d'une fonction de validation personnalisée

Cette section explique comment ajouter une nouvelle fonction de validation personnalisée à la bibliothèque d'objets, comment la définir, puis comment l'utiliser pour définir une règle de validation.

1. [Définition d'une fonction de validation personnalisée dans la bibliothèque d'objets \[page 410\]](#)
2. [Définition d'une règle de validation à l'aide d'une fonction de validation \[page 411\]](#)

26.3.3.2.1 Définition d'une fonction de validation personnalisée dans la bibliothèque d'objets

Cette procédure décrit l'ajout et la définition d'une fonction de validation personnalisée dans la bibliothèque d'objets.

Dans cet exemple, la fonction détermine si la colonne ZIP contient des codes postaux à 5 chiffres en contrôlant chaque caractère pour vérifier si c'est un chiffre. Pour vérifier si le caractère est un chiffre, la fonction contrôle si sa valeur ASCII est comprise entre 48 et 57 (correspondant respectivement à un chiffre de 0 à 9).

1. Dans l'onglet *Fonction personnalisée* de la bibliothèque d'objets, cliquez avec le bouton droit sur *Fonctions de validation* et sélectionnez *Nouveau*.
La boîte de dialogue *Fonction personnalisée* s'ouvre avec l'option *Fonction de validation* sélectionnée.
2. Saisissez le nom **ZipIsValid** de la nouvelle fonction et cliquez sur *Suivant*.
L'éditeur rapide s'affiche.
3. Pour définir les paramètres et les variables de cette fonction, sélectionnez l'onglet *Variables* dans la fenêtre de l'éditeur rapide.
4. Cliquez avec le bouton droit sur le nœud *Paramètres* et sélectionnez *Insérer*.
5. Saisissez un nom de paramètre tel que **\$ZipToValidate**.
6. Sélectionnez le type de données approprié, par exemple *varchar* avec une longueur de **5**.
7. Comme *Type de paramètre*, sélectionnez *Entrée*.

i Remarque

Les fonctions de validation peuvent comporter plusieurs paramètres d'entrée. Les paramètres de sortie supplémentaires ne sont pas utilisés.

8. Cliquez sur *OK*.
Le paramètre **\$ZipToValidate** s'affiche sous le nœud *Paramètres*.
9. Ajoutez cinq variables locales nommées **\$Z1** à **\$Z5**, qui représentent les cinq caractères de la colonne. Cliquez avec le bouton droit sur *Local*, sélectionnez *Insérer*, saisissez le nom, sélectionnez le type de données *int* et cliquez sur *OK*.
Les cinq variables s'affichent sous le nœud *Variables locales*.
10. Dans le panneau de texte de l'éditeur rapide, entrez le script de validation suivant :

```
$Z1 = ascii(substr($ZipToValidate, 1, 1));  
$Z2 = ascii(substr($ZipToValidate, 2, 1));
```



```

$Z3 = ascii(substr($ZipToValidate, 3, 1));
$Z4 = ascii(substr($ZipToValidate, 4, 1));
$Z5 = ascii(substr($ZipToValidate, 5, 1));
if ( $Z1 < 48 OR $Z1 > 57 )
    return 0;
if ( $Z2 < 48 OR $Z2 > 57 )
    return 0;
if ( $Z3 < 48 OR $Z3 > 57 )
    return 0;
if ( $Z4 < 48 OR $Z4 > 57 )
    return 0;
if ( $Z5 < 48 OR $Z5 > 57 )
    return 0;
return 1;

```

i Remarque

La transformation Validation prend uniquement en charge les fonctions personnalisées qui renvoient un type de données integer. Si une valeur renvoyée est différente de zéro, Data Services la traite comme VRAIE.

11. Cliquez sur **OK**.

La nouvelle fonction s'affiche dans la bibliothèque d'objets sous le nœud *Fonction de validation*.

Poursuivez avec la procédure suivante pour utiliser la fonction de validation dans une règle de validation.

Présentation des tâches : [Définition d'une règle de validation à l'aide d'une fonction de validation personnalisée \[page 410\]](#)

Prochaine tâche : [Définition d'une règle de validation à l'aide d'une fonction de validation \[page 411\]](#)

26.3.3.2 Définition d'une règle de validation à l'aide d'une fonction de validation

Cette procédure décrit l'utilisation d'une fonction de validation personnalisée pour définir une règle de validation.

1. Ajoutez un flux de données comportant une source et une transformation Validation et connectez-les.
2. Cliquez sur le nom de la transformation pour ouvrir l'éditeur de la transformation Validation.
3. Dans l'onglet *Règles de validation*, cliquez sur *Ajouter*.
L'éditeur de règles s'affiche avec la case *Activée* cochée par défaut.
4. Nommez la règle dans l'éditeur de règles.
5. Sélectionnez *Action en cas d'échec : Envoi vers échec*.
6. L'option *Fonction de validation* étant sélectionnée, sélectionnez la fonction ZipIsValid dans la liste déroulante.
La table Bindings est remplie avec les paramètres requis pour la fonction.
7. Définissez l'argument du paramètre. Pour le paramètre \$ZipToValidate, double-cliquez sur la cellule sous *Argument* et sélectionnez la colonne *ZIP*.
8. Cliquez sur **OK** pour enregistrer la règle.
La règle s'affiche dans la liste Règles de l'onglet *Règles de validation*.

A l'exécution du job, la règle appliquera la fonction de validation réutilisable à la colonne ZIP et contrôlera chaque caractère pour vérifier si c'est un chiffre.

Présentation des tâches : [Définition d'une règle de validation à l'aide d'une fonction de validation personnalisée \[page 410\]](#)

Tâche précédente : [Définition d'une fonction de validation personnalisée dans la bibliothèque d'objets \[page 410\]](#)

26.4 Utilisation de l'Audit

L'audit offre la possibilité de s'assurer qu'un flux de données charge les données correctes dans l'entrepôt. Utilisez l'audit pour effectuer les tâches suivantes :

- Définir les points de vérification pour collecter les statistiques d'exécution sur les données qui sortent des objets. L'audit enregistre ces statistiques dans le référentiel.
- Définir des règles avec ces statistiques d'audit pour s'assurer qu'aux points suivants d'un flux de données, les données correspondent à celles attendues :
 - Extraites des sources
 - Traitées par des transformations
 - Chargées dans les cibles
- Générer une notification d'exécution qui inclut la règle d'audit qui a échoué et les valeurs des statistiques d'audit au moment de l'échec.
- Afficher les statistiques d'audit après l'exécution du job pour aider à identifier les objets dans le flux de données qui peuvent avoir produit les données incorrectes.

i Remarque

Si vous ajoutez un point de vérification avant une opération qui est normalement poussée vers le bas vers le serveur de base de données, la performance peut se dégrader car ces opérations ne peuvent avoir lieu après un point de vérification.

26.4.1 Audit d'objets dans un flux de données

Il est possible de collecter les statistiques d'audit sur les données qui sortent d'un objet, tel qu'une source, une transformation ou une cible. Si une transformation a plusieurs sorties distinctes ou différentes (telles que Validation ou Case), il est possible de réaliser un audit pour chaque sortie indépendamment.

Pour utiliser l'audit, définissez les objets suivant dans la fenêtre [Audit](#) :

Nom d'objet	Description															
Point de vérification	L'objet dans un flux de données dans lequel vous collectez les statistiques d'audit. Il est possible de réaliser un audit pour une source, une transformation ou une cible. Vous identifiez l'objet à vérifier lors de la définition d'une fonction d'audit sur celui-ci.															
Fonction d'audit	Les statistiques d'audit collectées par le logiciel pour une table, un schème de sortie ou une colonne. La table suivante montre les fonctions d'audit qu'il est possible de définir. <table><tr><th>Objet de données</th><th>Fonction d'audit</th><th>Description</th></tr><tr><td>Table ou schéma de sortie</td><td>Nombre</td><td>Cette fonction collecte deux statistiques : - Un nombre correct pour les lignes qui ont été traitées avec succès. - Un nombre d'erreur pour les lignes qui ont généré certains types d'erreurs si vous avez activé la gestion des erreurs. </td></tr><tr><td>Colonne</td><td>Somme</td><td>Somme des valeurs numériques dans la colonne. Les types de données applicables incluent les nombres décimaux, doubles, entiers et réels. Cette fonction comprend uniquement les lignes correctes.</td></tr><tr><td>Colonne</td><td>Moyenne</td><td>Moyenne des valeurs numériques dans la colonne. Les types de données applicables incluent les nombres décimaux, doubles, entiers et réels. Cette fonction comprend uniquement les lignes correctes.</td></tr><tr><td>Colonne</td><td>Total de contrôle</td><td>Total de contrôle des valeurs dans la colonne.</td></tr></table>	Objet de données	Fonction d'audit	Description	Table ou schéma de sortie	Nombre	Cette fonction collecte deux statistiques : - Un nombre correct pour les lignes qui ont été traitées avec succès. - Un nombre d'erreur pour les lignes qui ont généré certains types d'erreurs si vous avez activé la gestion des erreurs.	Colonne	Somme	Somme des valeurs numériques dans la colonne. Les types de données applicables incluent les nombres décimaux, doubles, entiers et réels. Cette fonction comprend uniquement les lignes correctes.	Colonne	Moyenne	Moyenne des valeurs numériques dans la colonne. Les types de données applicables incluent les nombres décimaux, doubles, entiers et réels. Cette fonction comprend uniquement les lignes correctes.	Colonne	Total de contrôle	Total de contrôle des valeurs dans la colonne.
Objet de données	Fonction d'audit	Description														
Table ou schéma de sortie	Nombre	Cette fonction collecte deux statistiques : - Un nombre correct pour les lignes qui ont été traitées avec succès. - Un nombre d'erreur pour les lignes qui ont généré certains types d'erreurs si vous avez activé la gestion des erreurs.														
Colonne	Somme	Somme des valeurs numériques dans la colonne. Les types de données applicables incluent les nombres décimaux, doubles, entiers et réels. Cette fonction comprend uniquement les lignes correctes.														
Colonne	Moyenne	Moyenne des valeurs numériques dans la colonne. Les types de données applicables incluent les nombres décimaux, doubles, entiers et réels. Cette fonction comprend uniquement les lignes correctes.														
Colonne	Total de contrôle	Total de contrôle des valeurs dans la colonne.														
Étiquette d'audit	Le seul nom dans le flux de données généré par le logiciel pour les statistiques d'audit collectées pour chaque fonction d'audit définie. Ces étiquettes permettent de définir des règles d'audit pour le flux de données.															
Règle d'audit	Une expression booléenne dans laquelle vous utilisez les étiquettes d'audit pour vérifier le job. Si vous définissez plusieurs règles dans un flux de données, toutes les règles doivent réussir, sinon l'audit échoue.															
Actions en cas d'échec de l'audit	Une ou plus des trois façons de générer des notifications d'un échec de la règle (ou des règles) d'audit : adresse électronique, script personnalisé, lever une exception.															

26.4.1.1 Fonction d'audit

Cette section décrit les types de données pour les fonctions d'audit et pour les statistiques du nombre d'erreurs.

Types de données

La table suivant montre le type de données par défaut pour chaque fonction d'audit et les types de données autorisés. Il est possible de modifier le type de données dans la fenêtre *Propriétés* pour chaque fonction d'audit de Designer.

Fonctions d'audit	Type de données par défaut	Types de données autorisés
Nombre	Entier	Entier
Somme	Type de colonne audité	Entier, Décimal, Double, Réel
Moyenne	Type de colonne auditée	Entier, Décimal, Double, Réel
Total de contrôle	Varchar(128)	Varchar(128)

Statistique du nombre d'erreurs

Lorsque vous activez une fonction d'audit Nombre, le logiciel collecte deux types de statistiques :

- Nombre de lignes *Correctes* pour les lignes traitées sans erreur.
- Nombre de lignes d'*Erreur* pour les lignes que le job n'a pas pu traiter mais a ignorées pour pouvoir poursuivre le traitement. Les lignes d'erreur peuvent se produire lorsque vous spécifiez l'option *Utiliser le fichier de dépassement de capacité* dans l'éditeur source ou l'éditeur cible.

26.4.1.2 Etiquette d'audit

Le logiciel génère un nom unique pour chaque fonction d'audit définie à un point de vérification. Il est possible de modifier les noms d'étiquettes. Vous pouvez souhaiter modifier un nom d'étiquette pour créer un nom significatif plus court ou pour supprimer les tirets, qui sont autorisés dans les noms de colonnes mais pas dans les noms d'étiquettes.

Génération de noms d'étiquette

Si le point de vérification se trouve sur une table ou sur un schéma de sortie, le logiciel génère les deux étiquettes suivantes pour la fonction d'audit Nombre :

```
$Count<_NomObjet>  
$CountError<_NomObjet>
```

Si le point de vérification se trouve sur une colonne, le logiciel génère une étiquette d'audit sous le format suivant :

```
$ <FonctionAudit_NomObjet>
```

Si le point de vérification se trouve dans un flux de données incorporé, les étiquettes sont sous les formats suivants :

```
$Count<_NomObjet_NomFDIntégré>  
$CountError<_NomObjet_NomFDIntégré>  
$<FonctionAudit_NomObjet_NomFDIntégré>
```

Modification de noms d'étiquette

Il est possible de modifier le nom de l'étiquette d'audit lors de la création de la fonction d'audit, et avant la création d'une règle d'audit qui utilise l'étiquette.

Si vous modifiez le nom d'étiquette après l'avoir utilisé dans une règle d'audit, la règle d'audit n'utilise pas automatiquement le nouveau nom. Vous devez redéfinir la règle avec le nouveau nom.

26.4.1.3 Règle d'audit

Une règle d'audit est une expression booléenne qui comporte un côté gauche, un opérateur booléen et un côté droit.

- Le côté gauche peut être une seule étiquette d'audit, plusieurs étiquettes d'audit qui forment une expression avec un ou plusieurs opérateur(s) mathématique(s), ou une fonction avec des étiquettes d'audit comme paramètres.
- Le côté droit peut être une seule étiquette d'audit, plusieurs étiquettes d'audit qui forment une expression avec un ou plusieurs opérateur(s) mathématique(s), une fonction avec des étiquettes d'audit comme paramètres, ou une constante.

Les expressions booléennes suivantes sont des exemples de règles d'audit :

```
$Count_CUSTOMER = $Count_CUSTDW  
$Sum_ORDER_US + $Sum_ORDER_EUROPE = $Sum_ORDER_DW  
round($Avg_ORDER_TOTAL) >= 10000
```

26.4.1.4 Notification d'audit

Vous pouvez choisir n'importe quelle combinaison parmi les actions suivantes pour la notification d'un échec d'audit. Si vous sélectionnez les trois actions, le logiciel les exécute dans cet ordre :

Action	Description
Liste d'adresses électroniques	Le logiciel envoie une notification à propos de la règle d'audit qui a échoué aux adresses électroniques listées dans cette option. Utilisez une virgule pour séparer la liste des adresses électroniques. Il est possible d'indiquer une variable pour la liste d'adresses électroniques. Cette option doit utiliser la fonction smtp_to pour envoyer le courrier électronique. Par conséquent, vous devez définir le serveur et l'expéditeur pour le protocole SMTP dans le Gestionnaire de serveurs.
Script	Le logiciel exécute le script personnalisé que vous avez créé dans cette option.
Lever une exception	Le job échoue si une règle d'audit échoue et le journal des erreurs indique quelle règle d'audit a échoué. Le job s'arrête dès que la première règle d'audit échoue. Cette action est celle par défaut. Il est possible d'utiliser cette exception d'audit dans un bloc Try/Catch. Il est possible de continuer l'exécution du job dans un bloc Try/Catch. Si vous annulez cette action et qu'une règle d'audit échoue, le job s'achève avec succès et l'audit n'écrit pas de messages dans le journal des jobs. Vous pouvez visualiser la règle qui a échoué dans le rapport Détails de l'audit de l'outil Rapports de métadonnées.

Informations associées

[Affichage des résultats d'audit \[page 423\]](#)

26.4.2 Accès à la fenêtre Audit

Accédez à la fenêtre [Audit](#) depuis l'un des emplacements suivants de Designer :

- Depuis l'onglet Flux de données de la bibliothèque d'objets, faites un clic droit sur le nom du flux de données et sélectionnez l'option [Audit](#).
- Dans l'espace de travail, faites un clic droit sur l'icône du flux de données et sélectionnez l'option [Audit](#).
- Lorsqu'un flux de données est ouvert dans l'espace de travail, cliquez sur l'icône [Audit](#) dans la barre d'outils.

Lorsque vous accédez pour la première fois à la fenêtre [Audit](#), l'onglet Étiquette affiche les sources et les cibles dans le flux de données. Si le flux de données contient plusieurs transformations Query consécutives, la fenêtre [Audit](#) affiche la première requête.

Cliquez sur les icônes en haut à gauche de l'onglet Etiquette pour modifier l'affichage.

Icône	Astuce	Description
	Tout réduire	Réduit les extensions des objets source, cible, et de transformation.

Icône	Astuce	Description
	Afficher tous les objets	Affiche tous les objets du flux de données.
	Afficher la source, la cible et la requête de premier niveau	Affichage par défaut qui montre les objets source, cible et de requête de premier niveau dans le flux de données. Si le flux de données contient plusieurs transformations Query consécutives, seule la requête de premier niveau s'affiche.
	Afficher les objets étiquetés	Affiche les objets qui ont des étiquettes d'audit définies.

26.4.3 Définition de points de vérification, de règles et des actions en cas d'échec.

1. Accédez à la fenêtre [Audit](#).
2. Définissez des points de vérification. Sur l'onglet Etiquette, faites un clic droit sur l'objet que vous souhaitez vérifier et sélectionnez une fonction d'audit ou des propriétés.

Lorsque vous définissez un point de vérification, le logiciel génère les éléments suivants :

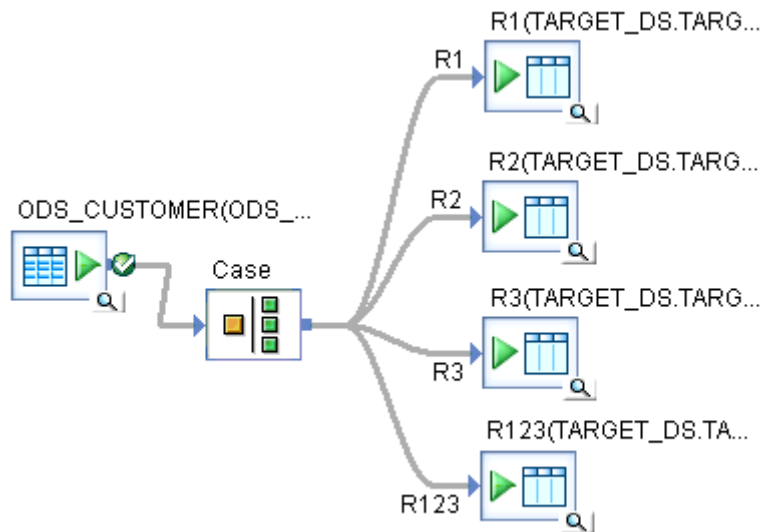
- Une icône d'audit sur l'objet dans le flux de données dans l'espace de travail
- Une étiquette d'audit utilisée pour définir les règles d'audit.

En plus de choisir une fonction d'audit, la fenêtre Propriétés permet de modifier l'étiquette d'audit et le type de données de la fonction d'audit.

Par exemple, le flux de données Case_DF dispose des objets ci-dessous, et vous souhaitez vérifier que toutes les lignes de source sont traitées par la transformation Case.

- Table source ODS_CUSTOMER
- Quatre tables cible :
 - R1 contient les lignes où ODS_CUSTOMER.REGION_ID = 1
 - R2 contient les lignes où ODS_CUSTOMER.REGION_ID = 2
 - R3 contient les lignes où ODS_CUSTOMER.REGION_ID = 3
 - R123 contient les lignes où ODS_CUSTOMER.REGION_ID IN (1, 2 ou 3)
- a. Faites un clic droit sur la table source ODS_CUSTOMER et sélectionnez [Nombre](#).

Le logiciel crée les étiquettes d'audit \$Count_ODS_CUSTOMER et \$CountError_ODS_CUSTOMER, et une icône d'audit apparaît sur l'objet source dans l'espace de travail.



- b. De la même façon, faites un clic droit sur chaque table cible et sélectionnez *Nombre*. La fenêtre Audit affiche les étiquettes d'audit suivantes :

Table cible	Fonction d'audit	Etiquette d'audit
ODS_CUSTOMER	Nombre	\$Count_ODS_CUSTOMER
R1	Nombre	\$Count_ R1
R2	Nombre	\$Count_ R2
R3	Nombre	\$Count_ R3
R123	Nombre	\$Count_ R123

- c. Si vous souhaitez supprimer une étiquette d'audit, faites un clic droit sur l'étiquette, et la fonction d'audit précédemment définie s'affiche avec une coche devant. Cliquez sur la fonction pour supprimer la coche et supprimez l'étiquette d'audit associée.

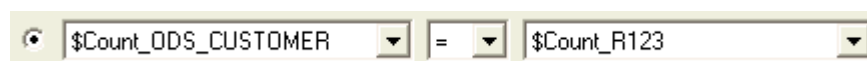
Lorsque vous faites un clic droit sur l'étiquette, vous pouvez également sélectionner Propriétés puis, dans la liste déroulante *Fonction d'audit*, la valeur (Aucun audit).

3. Définissez des règles d'audit. Dans l'onglet Règle de la fenêtre *Audit*, cliquez sur *Ajouter* ; cela active l'éditeur d'expressions de la section Règles d'audit.

Si vous souhaitez comparer les statistiques d'audit pour un objet par rapport à un autre, utilisez l'éditeur d'expression, qui comporte trois zones de texte avec des listes déroulantes :

- Sélectionnez l'étiquette du premier point de vérification dans la première liste déroulante.
- Sélectionnez un opérateur booléen depuis la deuxième liste déroulante. Les options dans l'éditeur fournissent des opérateurs booléens courants. Si vous avez besoin d'un opérateur booléen qui ne se trouve pas dans cette liste, utilisez la zone d'expression personnalisée avec ses fonctions et éditeurs rapides pour saisir l'opérateur.
- Sélectionnez l'étiquette pour le second point de vérification depuis la troisième liste déroulante. Si vous souhaitez comparer la première valeur d'audit à une constant plutôt qu'à une seconde valeur d'audit, utilisez la zone d'expression personnalisée.

Par exemple, pour vérifier que le nombre de lignes de la table source est égal à celui des lignes dans la table cible, sélectionnez les étiquettes d'audit et l'opération booléenne dans l'éditeur d'expression comme suit :



Si vous souhaitez comparer les statistiques d'audit pour un objet ou plus par rapport aux statistiques pour plusieurs autres objets ou constantes, sélectionnez la zone d'expression personnalisée.

- Cliquez sur le bouton d'ellipse pour ouvrir la fenêtre d'éditeur rapide en taille réelle.
- Cliquez sur l'onglet *Variables* sur la gauche et développez le nœud *Étiquettes*.
- Faites glisser la première étiquette d'audit de l'objet vers le volet de l'éditeur.
- Saisissez un opérateur booléen
- Faites glisser les étiquettes d'audit des autres objets auxquels vous souhaitez comparer les statistiques d'audit du premier objet, et placez les opérateurs mathématiques appropriés entre eux.
- Cliquez sur *OK* pour fermer l'éditeur rapide.
- La règle d'audit s'affiche dans l'éditeur personnalisé. Pour mettre à jour la règle dans la zone supérieure Règle d'audit, cliquez sur le titre "Règle d'audit" ou sur une autre option.
- Cliquez sur *Fermer* dans la fenêtre Audit.

Par exemple, pour vérifier que le nombre de lignes de la table source est égal à la somme des lignes dans les trois premières tables cible, faites glisser les étiquettes d'audit et saisissez l'opération booléenne et des signes plus dans l'éditeur rapide comme suit :

```
Count_ODS_CUSTOMER = $Count_R1 + $Count_R2 + $Count_R3
```

- Définissez la mesure à prendre si l'audit échoue.

Il est possible de sélectionner une ou plusieurs mesures suivantes :

Action	Description
Lever une exception	le job échoue si une règle d'audit échoue, et le journal des erreurs indique quelle règle d'audit a échoué. Cette action est celle par défaut. Si vous désactivez cette option et qu'une règle d'audit échoue, le job s'achève avec succès et l'audit n'écrit pas de message au journal des jobs. Vous pouvez visualiser la règle qui a échoué dans le rapport Détails de l'audit dans l'outil Rapports de métadonnées.
Liste d'adresses électroniques	le logiciel envoie une notification à propos de la règle d'audit qui a échoué aux adresses électroniques listées dans cette option. Utilisez une virgule pour séparer la liste des adresses électroniques. Il est possible d'indiquer une variable pour la liste d'adresses électroniques.
Script	le logiciel exécute le script que vous avez créé dans cette option.

- Exécutez le job.

La fenêtre *Propriétés d'exécution* a l'option *Activer l'audit* cochée par défaut. Décochez cette case si vous ne souhaitez pas collecter les statistiques d'audit pour cette exécution de job spécifique.

- Regardez les résultats d'audit.

Il est possible de voir les règles d'audit passées et ayant échoué dans les rapports de métadonnées. Si vous activez le suivi d'audit dans l'onglet Suivi de la fenêtre *Propriétés d'exécution*, vous pouvez voir tous les résultats d'audit dans le journal de surveillance des jobs.

Informations associées

[Audit d'objets dans un flux de données \[page 412\]](#)

[Affichage des résultats d'audit \[page 423\]](#)

26.4.4 Indications pour sélectionner les points de vérification

Voici des indications pour sélectionner des points de vérification :

- Lorsque vous vérifiez les données de sortie d'un objet, l'optimiseur ne peut pas pousser vers le bas les opérations après le point de vérification. Par conséquent, si la performance d'une requête poussée vers le serveur de base de données est plus importante que de regrouper les statistiques d'audit de la source, définissez le premier point de vérification sur la requête ou plus tard dans le flux de données. Par exemple, supposons que le flux de données dispose d'objets source, cible et de requête, et que la requête a une clause WHERE qui est poussée vers le serveur de base de données qui réduit de manière significative le total de données qui sont renvoyées au logiciel. Définissez le premier point de vérification sur la requête, plutôt que sur la source, pour obtenir des statistiques d'audit sur les résultats de la requête.
- Si une fonction pushdown_sql se trouve après le point de vérification, le logiciel ne peut pas l'exécuter.
- Il est uniquement possible de vérifier un chargement par lots qui utilise la méthode API d'Oracle. Pour les autres méthodes de chargements par lots, le logiciel n'a pas accès au nombre de lignes chargées.
- L'audit est désactivé lorsque vous exécutez un job avec le débogueur.
- Il est impossible de vérifier des schèmes MLD ou des jobs en temps réel.
- Il est impossible d'auditer dans un flux de données ABAP, mais il est possible d'auditer la sortie d'un flux de données ABAP.
- Si vous utilisez la fonction d'audit CHECKSUM dans un job qui exécute normalement en parallèle, le logiciel désactive le DOP pour l'ensemble du flux de données. L'ordre des lignes est important pour le résultat du CHECKSUM, et le DOP traite les lignes dans un ordre différent que dans la source.

26.4.5 Audit du flux de données intégré

Il est possible de définir des étiquettes d'audit et des règles d'audit dans un flux de données intégré. Cette section décrit les considérations suivantes lors de l'audit d'un flux de données intégré :

- Activation d'audit dans un flux de données intégré
- Points de vérification non visibles en dehors du flux de données intégré

26.4.5.1 Activation d'audit dans un flux de données intégré

Si vous souhaitez collecter les statistiques d'audit sur un flux de données intégré lors de l'exécution d'un flux de données parent, vous devez activer l'étiquette d'audit du flux de données intégré.

26.4.5.1.1 Activation d'audit dans un flux de données intégré

1. Ouvrez le flux de données parent dans l'espace de travail de Designer
2. Cliquez sur l'icône Audit dans la barre d'outils pour ouvrir la fenêtre Audit.
3. Sur l'onglet Etiquette, développez les objets pour afficher les fonctions d'audit définies dans le flux de données intégré. Si un flux de données est intégré au début ou à la fin du flux de données parent, une fonction d'audit peut exister sur le port de sortie ou sur le port d'entrée.
4. Faites un clic droit sur le nom de la fonction Audit et sélectionnez [Activer](#). Vous pouvez également sélectionner [Propriétés](#) pour modifier le nom de l'étiquette et l'activer.
5. Il est également possible de définir des règles d'audit avec l'étiquette activée.

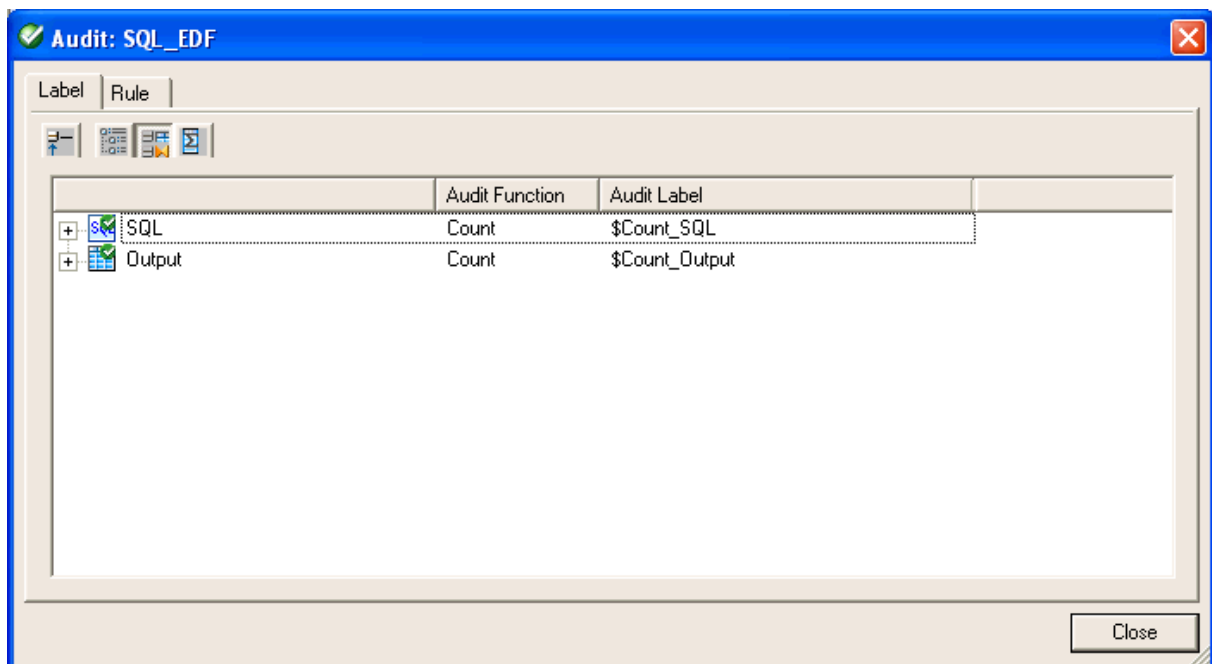
26.4.5.2 Points de vérification non visibles en dehors du flux de données intégré

Lorsque vous intégrez un flux de données au début d'un autre flux de données, les données sont transmises du flux de données intégré au flux de données parent par une seule source. Lorsque vous intégrez un flux de données à la fin d'un autre flux de données, les données sont transmises du flux de données parent au flux de données intégré par une seule cible. Dans les deux cas, certains objets ne sont pas visibles dans le flux de données parent.

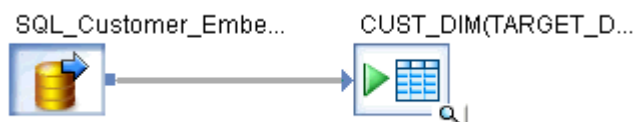
Parce que certains objets ne sont pas visibles dans le flux de données parent, les points de vérification sur ces objets ne sont pas non plus visibles dans le flux de données parent. Par exemple, le flux de données intégré suivant dispose d'une fonction d'audit définie dans la transformation SQL source et d'une fonction d'audit définie dans la table cible.



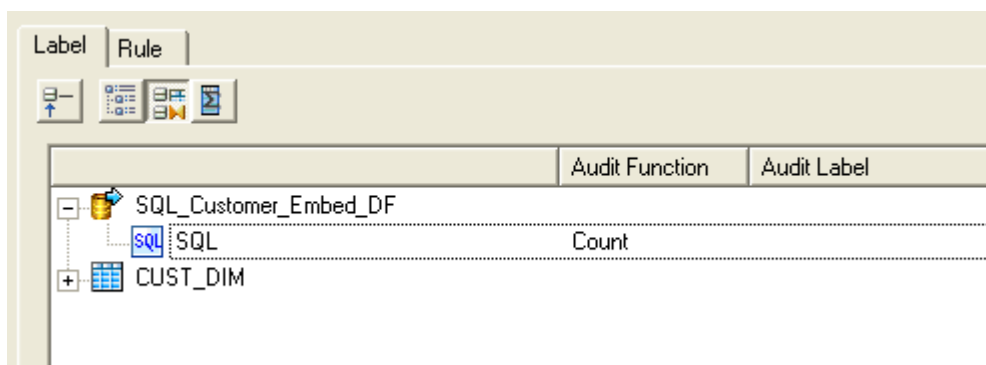
La fenêtre Audit ci-dessous affiche ces deux points de vérification.



Lorsque vous intégrez ce flux de données, la sortie cible devient une source pour le flux de données parent et la transformation SQL n'est plus visible.



Un point de vérification existe toujours pour l'ensemble du flux de données intégré, mais l'étiquette n'est plus applicable. La fenêtre Audit ci-dessous pour le flux de données parent affiche la fonction d'audit définie dans le flux de données intégré, mais n'affiche aucune étiquette d'audit.



Si vous souhaitez vérifier le flux de données intégré, faites un clic droit sur la fonction d'audit dans la fenêtre Audit et sélectionnez [Activer](#).

26.4.6 Résolution des étiquettes d'audit non valides

Une étiquette d'audit peut devenir non valide dans les situations suivantes :

- Si vous supprimez l'étiquette d'audit dans un flux de données intégré activé par un flux de données parent.
- Si vous supprimez ou renommez un objet dans lequel un point de vérification était défini.

26.4.6.1 Résolution des étiquettes d'audit non valides

1. Ouvrez la fenêtre [Audit](#).
2. Développez le nœud Etiquettes non valides pour afficher les étiquettes.
3. Notez les étiquettes que vous souhaitez définir sur n'importe quel nouvel objet dans le flux de données.
4. Après avoir défini une étiquette d'audit correspondante sur un nouvel objet, faites un clic droit sur l'étiquette non valide et sélectionnez [Supprimer](#).
5. Si vous souhaitez supprimer toutes les étiquettes non valides en même temps, faites un clic droit sur le nœud Etiquettes non valides et cliquez sur [Tout supprimer](#).

26.4.7 Affichage des résultats d'audit

Il est possible de voir le statut de l'audit à l'un des endroits suivants :

- Journal de surveillance des jobs
- Si la règle d'audit échoue, les endroits qui affichent l'information d'audit dépendent de l'option [Action en cas d'échec](#) que vous avez sélectionnée :

Action en cas d'échec	Endroits où il est possible d'afficher les informations d'audit
Lever une exception	Journal des erreurs des jobs, rapports de métadonnées
Liste d'adresses électroniques	Message électronique, rapports de métadonnées
Script	Tous les endroits où le script personnalisé envoie les messages d'audit, rapports de métadonnées

Informations associées

[Journal de surveillance des jobs \[page 424\]](#)

[Journal des erreurs des jobs \[page 424\]](#)

[Rapports de métadonnées \[page 424\]](#)

26.4.7.1 Journal de surveillance des jobs

Si vous définissez *Effectuer le suivi des données d'audit* sur *Oui* dans l'onglet Suivi de la fenêtre Propriétés d'exécution, des messages d'audit apparaissent dans le journal de surveillance des jobs. Il est possible de voir des messages pour les règles d'audit qui sont passées et ont échoué.

Les exemples de messages de réussite d'audit suivants apparaissent dans le journal de surveillance des jobs lorsque le *Suivi d'audit* est défini sur *Oui* :

```
Audit Label $Count_R2 = 4. Data flow <Case_DF>.
Audit Label $CountError_R2 = 0. Data flow <Case_DF>.
Audit Label $Count_R3 = 3. Data flow <Case_DF>.
Audit Label $CountError_R3 = 0. Data flow <Case_DF>.
Audit Label $Count_R123 = 12. Data flow <Case_DF>.
Audit Label $CountError_R123 = 0. Data flow <Case_DF>.
Audit Label $Count_R1 = 5. Data flow <Case_DF>.
Audit Label $CountError_R1 = 0. Data flow <Case_DF>.
Audit Label $Count_ODS_CUSTOMER = 12. Data flow <Case_DF>.
Audit Label $CountError_ODS_CUSTOMER = 0. Data flow <Case_DF>.
Audit Rule passed ($Count_ODS_CUSTOMER = (($CountR1 + $CountR2 + $Count_R3)):
LHS=12, RHS=12. Data flow <Case_DF>.
Audit Rule passed ($Count_ODS_CUSTOMER = $CountR123): LHS=12, RHS=12. Data flow
<Case_DF>.
```

26.4.7.2 Journal des erreurs des jobs

Lorsque vous sélectionnez l'option *Lever une exception* et que la règle d'audit échoue, le journal des erreurs des jobs affiche la règle qui a échoué. Les exemples de messages suivants apparaissent dans le journal des erreurs des jobs :

```
Audit rule failed <($Count_ODS_CUSTOMER = $CountR1)> for <Data flow Case_DF>.
```

26.4.7.3 Rapports de métadonnées

Vous pouvez consulter la colonne *Statut de l'audit* dans les rapports Statistiques d'exécution du flux de données de l'outil de rapports de métadonnées. Cette colonne *Statut de l'audit* contient les valeurs suivantes :

Valeur	Description
Non audité	Les statistiques n'ont pas été collectées.
Réussite	Toutes les règles d'audit ont réussi. Cette valeur est un lien avec le rapport Détails de l'audit qui affiche les règles d'audit et les valeurs des étiquettes d'audit.
Informations collectées	Ce statut apparaît lorsque vous définissez des étiquettes d'audit pour collecter des statistiques, mais que vous ne définissez pas de règles d'audit. Cette valeur est un lien avec le rapport Détails de l'audit qui affiche les valeurs des étiquettes d'audit.

Valeur	Description
Échec	La règle d'audit a échoué. Cette valeur est un lien avec le rapport Détails de l'audit qui affiche la règle d'audit qui a échoué et les valeurs des étiquettes d'audit.

27 Qualité des données

27.1 Présentation de la qualité des données

Qualité des données est un terme qui fait référence à l'ensemble de transformations qui fonctionnent ensemble pour améliorer la qualité des données en nettoyant, améliorant, faisant correspondre et consolidant les éléments de données.

La qualité des données est principalement obtenue dans le progiciel à l'aide de quatre transformations :

Transformation	Description
Address Cleanse	Analyse, standardise, corrige et améliore les données d'adresses.
Data Cleanse	Analyse, standardise, corrige et améliore les données client et opérationnelles.
Géocodage	Utilise les coordonnées géographiques, les adresses et les données de points d'intérêt (POI) pour ajouter une adresse, une latitude et une longitude, un recensement et d'autres informations aux enregistrements.
Match	Identifie les enregistrements doublons à plusieurs niveaux lors d'un seul transfert des individus, des membres ou des organisations dans plusieurs tables ou bases de données et les consolide en une source unique.

Informations associées

[A propos du nettoyage de données \[page 426\]](#)

[Address Cleanse \[page 565\]](#)

[Géocodage \[page 457\]](#)

[Stratégies de correspondance \[page 480\]](#)

27.2 Data Cleanse

27.2.1 A propos du nettoyage de données

Le nettoyage de données désigne le processus d'analyse et de standardisation des données.

Les règles d'analyse et autres informations qui définissent le mode d'analyse et de standardisation des données sont stockées dans un package de nettoyage des données. Le Générateur de packages de nettoyage des données de SAP Information Steward fournit une interface utilisateur graphique permettant de créer et d'ajuster les packages de nettoyage des données. Pour respecter vos besoins et standards spécifiques en

matière de nettoyage des données, il est possible de créer un package de nettoyage des données en partant de zéro d'après un exemple de données ou bien d'en adapter un existant ou fourni par SAP.

i Remarque

La fonctionnalité Assistant de nettoyage utilise le package de nettoyage des données fourni par SAP Personne_Société lors du traitement des données. Ne supprimez pas ou ne renommez pas le package de nettoyage des données fourni par SAP. Si vous modifiez le package de nettoyage des données fourni par SAP, les résultats de l'Assistant de nettoyage risquent de s'en trouver altérés.

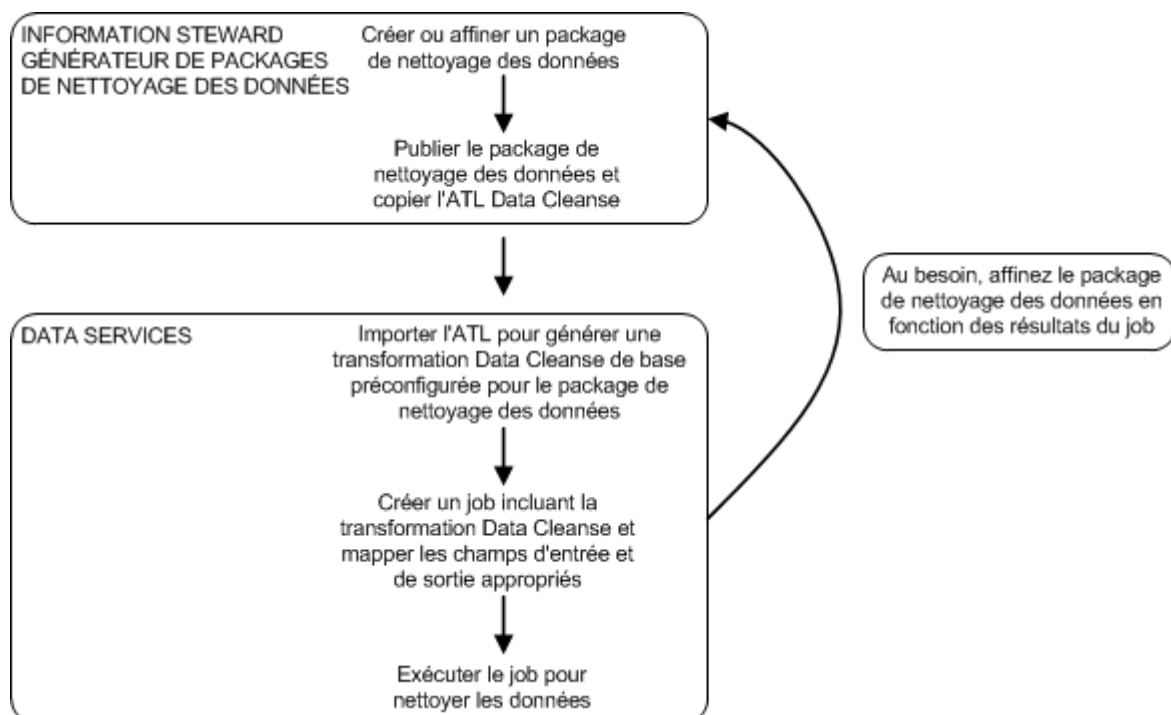
i Remarque

Après la création d'un ou de plusieurs types de contenus personnalisés dans SAP BusinessObjects Data Insight, un package de nettoyage des données nommé CTID_CUSTOM_TYPE s'affiche dans le Générateur de packages de nettoyage des données. N'ouvrez pas ou ne modifiez pas CTID_CUSTOM_TYPE dans le Générateur de packages de nettoyage des données car cela affectera les types de contenus personnalisés.

Un package de nettoyage des données est créé et publié dans le Générateur de packages de nettoyage des données puis référencé par la transformation Data Cleanse dans Data Services à des fins de test et de déploiement en production.

Dans un workflow Data Services, la transformation Data Cleanse identifie et isole des éléments spécifiques de données mixtes, puis analyse et met en forme les données en fonction du package de nettoyage des données référencé ainsi que des options définies directement dans la transformation.

Le diagramme suivant montre comment Data Services et Information Steward fonctionnent ensemble pour vous permettre de développer un package de nettoyage des données correspondant à vos exigences en matière de données, puis de l'appliquer lors du nettoyage de vos données.



27.2.2 Cycle de vie du package de nettoyage des données : développement, déploiement et gestion

Le processus de développement, de déploiement et de gestion d'un package de nettoyage des données est le résultat d'actions et de communications entre l'administrateur de Data Services, le testeur de Data Services et l'administrateur de données du Générateur de packages de nettoyage des données. Les rôles, responsabilités et titres exacts varient selon les organisations, mais comprennent souvent ce qui suit :

Rôle	Responsabilité
Administrateur de données du Générateur de packages de nettoyage des données	Utilise le Générateur de packages de nettoyage des données et possède des connaissances en matière de domaines pour développer et affiner un package de nettoyage des données spécifique à un domaine de données.
Testeur de Data Services	Utilise la transformation Data Services dans un environnement de test Data Services pour nettoyer les données. Envoie des résultats à l'administrateur de données du Générateur de packages de nettoyage des données pour la vérification et l'ajustement.
Administrateur Data Services	Dans un environnement de production Data Services, utilise la transformation Data Cleanse pour nettoyer les données selon les règles et les standards définis dans le package de nettoyage des données sélectionné.

Il existe habituellement trois phases itératives dans un workflow de package de nettoyage des données : création et test (développement), déploiement et gestion.

Phases du workflow du package de nettoyage des données

Phase	Package de nettoyage des données personnalisé	Package de nettoyage des données personnelles et professionnelles
Création et test (développement)	L'administrateur de données crée un package de nettoyage des données personnalisé à l'aide de l'assistant du package de nettoyage des données personnalisé. L'administrateur de données test et ajuste le package de nettoyage des données dans le Générateur de packages de nettoyage des données en fonction des résultats de la standardisation et de l'analyse du nettoyage des données qui s'affichent dans le volet inférieur du mode <i>Conception</i>. Quand l'administrateur de données est satisfait des résultats, il publie le package de nettoyage des données et l'envoie au testeur Data Services pour tester le package de nettoyage des données dans un jeu de données plus important. L'administrateur de données contrôle les résultats non satisfaisants (le cas échéant) du test Data Cleanse et effectue les modifications dans le package de nettoyage des données. Ce processus se poursuit jusqu'à ce que le package de nettoyage des données produise les résultats attendus dans Data Services.	L'administrateur de données crée un package de nettoyage des données personnelles et professionnelles à l'aide d'une copie d'un package de nettoyage des données publié ou en créant un manuellement (non recommandé). L'administrateur de données test et ajuste le package de nettoyage des données dans le Générateur de packages de nettoyage des données en fonction des résultats de la standardisation et de l'analyse du nettoyage des données qui s'affichent dans les exemples de données de l'onglet <i>Valider</i> du Générateur du package de nettoyage des données. Quand l'administrateur de données est satisfait des résultats, il publie le package de nettoyage des données et l'envoie au testeur Data Services pour tester le package de nettoyage des données dans un jeu de données plus important. L'administrateur de données contrôle les résultats non satisfaisants (le cas échéant) du test Data Cleanse et effectue les modifications dans le package de nettoyage des données. Ce processus se poursuit jusqu'à ce que le package de nettoyage des données produise les résultats attendus dans Data Services.

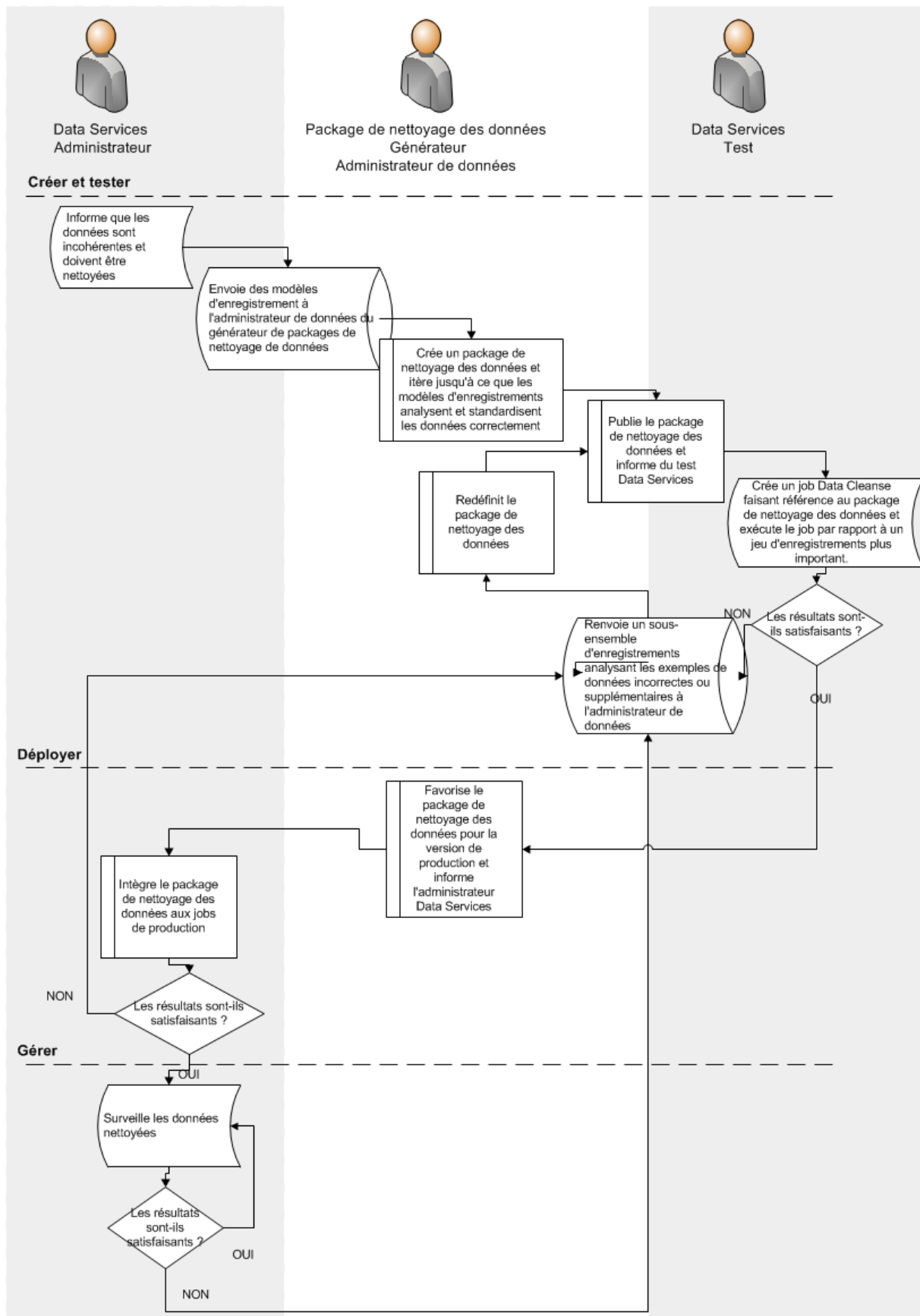
i Remarque

L'onglet *Valider* est disponible uniquement pour des packages de nettoyage des données personnelles et professionnelles à plusieurs domaines.

Pour les packages de nettoyage des données à un domaine, quand l'administrateur de données est satisfait des résultats, il publie le package de nettoyage des données et l'envoie au testeur Data Services pour tester le package de nettoyage des données dans un jeu de données plus important. L'administrateur de données contrôle les ré-

Phase	Package de nettoyage des données personnalisé	Package de nettoyage des données personnelles et professionnelles
		sultats non satisfaisants (le cas échéant) et effectue les modifications dans le package de nettoyage des données. Ce processus se poursuit jusqu'à ce que le package de nettoyage des données produise les résultats attendus dans Data Services.
Déploiement	Lorsque tout le monde est satisfait des résultats du test, l'administrateur de données déploie le package de nettoyage des données en production à l'aide de l'outil de gestion des promotions dans la Central Management Console (CMC). Pour en savoir plus sur la gestion des promotions, voir « Gestion du cycle de vie » dans le guide <i>Information platform services Administrator Guide</i>.	Lorsque tout le monde est satisfait des résultats du test, l'administrateur de données déploie le package de nettoyage des données en production à l'aide de l'outil de gestion des promotions dans la CMC. Pour en savoir plus sur la gestion des promotions, voir « Gestion du cycle de vie » dans le guide <i>Information platform services Administrator Guide</i>.
Maintenance	L'administrateur Information Steward déplace le package de nettoyage des données en phase de maintenance et effectue des mises à jour uniquement lorsque les résultats des jobs planifiés régulièrement ne respectent pas la plage ou lorsque de nouvelles données sont introduites.	L'administrateur Information Steward déplace le package de nettoyage des données en phase de maintenance et effectue des mises à jour uniquement lorsque les résultats des jobs planifiés régulièrement ne respectent pas la plage ou lorsque de nouvelles données sont introduites.

Un workflow classique est illustré par le diagramme ci-dessous :



Pour en savoir plus sur le processus de déplacement d'un package de nettoyage des données du développement vers la production et vers la maintenance, voir le *Guide d'administration SAP BusinessObjects Information Steward*.

Pour en savoir plus sur l'utilisation de l'outil de gestion des promotions de la Central Management Console (CMC) pour déplacer des packages de nettoyage des données (ce qui est désigné sous le terme de promotion d'« objets » ou d'« InfoObjects »), voir le *Guide d'administration de la plateforme de Business Intelligence*.

27.2.3 Configuration de la transformation Data Cleanse

Les conditions préalables pour la configuration de la transformation Data Cleanse incluent :

- L'accès aux packages de nettoyages des données nécessaires.
- L'accès au fichier ATL transféré à partir du générateur de packages de nettoyage des données.
- Les informations sur le mappage du champ d'entrée et de l'attribut (champ de sortie) pour les règles de correspondance des modèles définies par l'utilisateur définies dans l'onglet *Données de référence* de Cleansing Package Builder.

Configurer la transformation Data Cleanse :

1. Importez le fichier ATL transféré à partir du générateur de packages de nettoyage des données.
L'importation du fichier ATL apporte les informations requises et définit automatiquement les options suivantes :
 - Package de nettoyage des données
 - Champ de sortie de filtre
 - Séparateur de mots dans l'entrée
 - Configuration de l'analyseur

i Remarque

Il est possible d'installer et d'utiliser le package de nettoyage des données personnelles et professionnelles fourni par SAP sans modification directement dans Data Services. Pour ce faire, sautez l'étape 1 et définissez manuellement les options requises dans la transformation Data Cleanse.

2. Dans le schéma d'entrée, sélectionnez les champs d'entrée que vous souhaitez mapper et faites les glisser vers les champs appropriés dans l'onglet *Entrée*
 - Les données de nom et de société peuvent être mappées vers des champs discrets ou des champs multilignes.
 - Les données personnalisées doivent être mappées vers des champs multilignes.
 - Les données de téléphone, date, courrier électronique, numéro de sécurité sociale et de modèle défini par l'utilisateur peuvent être mappées vers des champs discrets ou des champs multilignes.
L'analyseur correspondant doit être activé.
3. Dans l'onglet *Options*, sélectionnez les valeurs d'option appropriées pour déterminer la manière dont Data Cleanse doit traiter les données.
Si vous modifiez la valeur par défaut d'une valeur d'option, un triangle vert apparaît à côté du nom de l'option pour indiquer que la valeur a été modifiée.

Le fichier ATL que vous avez importé à l'étape 1 définit certaines options en fonction des informations contenues dans le package de nettoyage des données.

4. Dans l'onglet [Sortie](#), sélectionnez les champs que vous souhaitez sortir de la transformation. Dans le Cleansing Package Builder, les champs de sortie sont appelés attributs.
Assurez-vous d'avoir mappé tous les attributs (champs de sortie) définis dans les modèles définis par l'utilisateur dans les données de référence de Cleansing Package Builder.

Informations associées

[Configurations de transformation \[page 232\]](#)

27.2.4 Classement et définition de priorité des moteurs d'analyse

Pour les entrées multilignes, il est possible de configurer la transformation Data Cleanse pour utiliser uniquement des analyseurs spécifiques et pour indiquer l'ordre dans lequel ils sont exécutés. Le fait de sélectionner soigneusement les analyseurs à utiliser et dans quel ordre peut être bénéfique. Le fait d'éteindre les analyseurs dont vous n'avez pas besoin améliore significativement la vitesse d'analyse et réduit les chances que les données soient mal analysées.

Il est possible de modifier l'ordre des analyseurs pour une entrée multiligne spécifique en modifiant l'option de séquence d'analyseur correspondante dans le groupe des options Configuration_Analyseur de la transformation Data Cleanse. Par exemple, pour modifier l'ordre des analyseurs pour le champ d'entrée Multiligne1, modifiez l'option Sequence_analyseur_Multiligne1.

Pour modifier les analyseurs sélectionnés ou l'ordre des analyseurs, sélectionnez une séquence d'analyseur, cliquez sur [OK](#) sur le message puis utilisez la fenêtre [Options classées](#) pour effectuer les modifications.

Informations associées

[Editeur d'options classées \[page 234\]](#)

27.2.5 Données d'analyse

La transformation Data Cleanse peut identifier et isoler un large éventail de données. Dans la transformation Data Cleanse, vous mappez les champs d'entrée des données aux champs d'entrée appropriés dans la transformation. Les données personnalisées contenant des données opérationnelles ou de produit sont toujours mappées aux champs multilignes. Les données personnelles et professionnelles, les données de téléphone, adresse électronique, date et numéro de sécurité sociale peuvent être mappées aux champs d'entrée ou aux champs d'entrée multilignes.

L'exemple ci-dessous indique la manière dont Data Cleanse analyse les données de produit à partir d'un champ d'entrée multiligne et les affiche dans les champs de sortie discrets. Les données peuvent également être

affichées dans des champs composites, tels que Description standard, qui peuvent être personnalisés dans le Générateur de packages de nettoyage des données pour répondre à vos besoins.

Données d'entrée	Données analysées	
Gant ultra grip profit 2.3 grand noir cuir synthétique élastique avec Velcro Mechanix Wear	Catégorie du produit	Gant
	Taille	Grand
	Article	Cuir synthétique
	Marque déposée	Pro-Fit 2.3 Series
	Style de poignet	Velcro élastique
	Type de la paume	Ultra-Grip
	Couleur	Noir
	Fournisseur	Mechanix Wear
	Description standard	Gant - cuir synthétique, noir, taille : grand, Style de poignet : Velcro élastique, Ultra-Grip, Mechanix Wear.

L'exemple ci-dessous indique la manière dont Data Cleanse analyse les données de nom et de société, et les affiche dans les champs de sortie discrets. Les données peuvent également être affichées dans des champs composites qui peuvent être personnalisés dans le Générateur de packages de nettoyage des données pour répondre à vos besoins.

Données d'entrée	Données analysées		Données composites
Mme Mary Ann Smith Jones, Responsable des comptes, expert-comptable agréé. Jones Inc.	Titre de civilité	Mme	
	Prénom 1	Mary	Prénom_Complet
	Prénom 2	Anne	Mary Anne
	Nom de famille	Smith	Nom_Famille_Complet
	Nom de famille 2	Jones	Smith Jones
	Titre honorifique	Expert-comptable agréé	
	Fonction	Responsable des comptes	
	Société	Jones, Inc.	

Données d'entrée	Données analysées	
James Witt Jr 421-55-2424 jwitt@rdrindustries.com 507-555-3423 20 août, 2003	Prénom 1	James
	Nom de famille 1	Witt
	Titre de maturité	Jr.
	Sécurité sociale	421-55-2424
	Adresse électronique	jwitt@rdrindustries.com
	Téléphone	507.555.3423
	Date	20 août 2003

La transformation Data Cleanse analyse jusqu'à 6 noms par enregistrement, 2 par champ d'entrée. Pour les six noms trouvés, elle analyse les composants tels que le titre de civilité, les prénoms, le nom de famille et le suffixe du nom. Elle envoie ensuite les données vers les champs individuels. La transformation Data Cleanse analyse également jusqu'à 6 titres par enregistrement.

La transformation Data Cleanse analyse jusqu'à 6 noms de société par enregistrement, 2 par champ d'entrée.

27.2.5.1 Analyse des numéros de téléphone

Data Cleanse analyse le plan de numérotation nord-américain (NANP) et les numéros de téléphone internationaux.

Les systèmes de numérotation téléphonique sont différents à travers le monde. Lorsque Data Cleanse analyse un numéro de téléphone, il édite les composants individuels du numéro dans les champs de sortie appropriés.

L'analyse la plus efficace est obtenue lorsqu'un numéro de téléphone a un code pays valide qui apparaît avant le numéro. Si le code pays est absent ou s'il ne figure pas au début de la chaîne, Data Cleanse utilise d'autres ressources pour tenter d'analyser le numéro de téléphone.

Data Cleanse analyse les numéros de téléphone en effectuant d'abord une recherche internationale. Il utilise les codes pays ISO2 lorsqu'ils sont disponibles ainsi que les modèles définis dans le package de nettoyage des données pour chaque pays en vue d'identifier le code correspondant. S'il détecte un pays figurant dans le NANP, il arrête automatiquement son analyse internationale et tente d'analyser le numéro de téléphone comme un numéro nord-américain, en le comparant aux modèles généralement utilisés tels que (234) 567-8901, 234-567-8901 et 2345678901.

Les pays figurant actuellement dans le NANP sont les suivants :

- AS - Samoa américaine
- AI - Anguilla
- AG - Antigua-et-Barbuda
- BS- Bahamas
- BB - Barbade
- BM - Bermudes
- VG - Îles Vierges britanniques
- CA - Canada
- KY - Îles Caïmans
- DM - Dominique
- DO - République dominicaine
- GD - Grenade
- GU - Guam
- JM - Jamaïque
- MS - Montserrat
- MP - Îles Mariannes du Nord
- PR - Porto Rico
- KN - Saint-Kitts-et-Nevis
- LC - Sainte-Lucie
- VC - Saint-Vincent-et-les-Grenadines

- SX - Saint-Martin
- TT - Trinité-et-Tobago
- TC - Îles Turks et Caicos
- US - États-Unis
- VI - Îles Vierges américaines

Data Cleanse offre la possibilité de mettre en forme les numéros nord-américains à la sortie (choix du séparateur, par exemple). Toutefois, Data Cleanse édite les numéros internationaux tels qu'ils ont été entrés, sans mise en forme. De plus, Data Cleanse n'effectue pas une comparaison croisée avec les adresses pour voir si les indicatifs du pays et de la ville du numéro de téléphone correspondent à l'adresse.

Informations associées

[A propos du mappage un par un \[page 444\]](#)

27.2.5.1.1 Analyse des numéros de téléphone par Data Cleanse

Configuration de l'analyse des numéros de téléphone internationaux dans la transformation Data Cleanse

Vous pouvez configurer l'analyse des numéros de téléphone internationaux dans la transformation Data Cleanse à trois niveaux : Enregistrement, Fichier du job et Global. Data Cleanse tente d'analyser les données de téléphone dans l'ordre, comme le montre la table ci-dessous. Si la transformation ne peut pas analyser les données de téléphone via l'un des trois niveaux, elle tente de les analyser comme des données nord-américaines. Si Data Cleanse détecte un code pays qui figure dans le plan de numérotation nord-américain (NANP) à n'importe quel niveau, il ignore les niveaux restants et tente automatiquement d'analyser le numéro comme un numéro nord-américain.

i Remarque

Le pays doit être indiqué dans le package de nettoyage des données pour que Data Cleanse puisse lancer son analyse à l'aide des trois niveaux indiqués dans la table ci-dessous.

Niveau	Processus
Enregistrement	Facultatif : utilisez le champ d'entrée dynamique Option_Pays. Configurez votre flux de données pour inclure Global Address Cleanse avant la transformation Data Cleanse. Vérifiez que la transformation Global Address Cleanse génère le champ ISO_Country_Code_2Char. Mappez ensuite ISO_Country_Code_2Char sur le champ d'entrée dynamique Option_Pays dans votre mappage d'entrée. Avantage : comme Global Address Cleanse détermine le code pays indiqué dans le champ ISO_Country_Code_2Char en fonction des données d'adresse de l'enregistrement, il y a de fortes chances que l'adresse et le numéro de téléphone de l'enregistrement proviennent du même pays. Data Cleanse tente d'analyser les données de téléphone à l'aide du code pays. Si une correspondance est trouvée, Data Cleanse analyse les données de téléphone en fonction du pays concerné, puis passe à l'analyse de l'enregistrement suivant. Si la transformation ne trouve pas de correspondance, l'analyse passe au niveau suivant (fichier du job). Si le code pays figure dans le NANP, Data Cleanse interrompt l'analyse international, ignore les niveaux Fichier du job et Global et tente d'analyser les données de téléphone comme des données nord-américaines. Pour en savoir plus sur les options de transformation dynamique, voir le <i>Guide de référence</i>.

Niveau	Processus
Fichier du job	Facultatif : utilisez l'option <i>Séquence de code pays ISO2</i>, accessible sous <i>Options téléphoniques</i> dans l'onglet <i>Options</i> de Data Cleanse. Définissez une séquence de pays via l'option <i>Séquence de code pays ISO2</i> pour indiquer l'ordre dans lequel Data Cleanse doit rechercher les données de téléphone. Data Cleanse tente de faire correspondre les données de téléphone au premier pays dans la séquence. Si aucune correspondance n'est trouvée, Data Cleanse passe au pays suivant dans la séquence. Global est l'option par défaut pour la séquence. Vous devez conserver Global comme dernière entrée dans la séquence ou supprimer cette option. Avantage : si vous savez que vos données proviennent surtout de certains pays, vous pouvez définir votre séquence avec ces pays et supprimer Global. Dans ce cas, Data Cleanse tente de faire correspondre les données de téléphone à ces pays uniquement et dans l'ordre que vous avez indiqué. Si le code pays figure dans le NANP, Data Cleanse interrompt l'analyse internationale, ignore le niveau Global et tente d'analyser les données de téléphone comme des données nord-américaines. Si vous supprimez Global de la séquence et si Data Cleanse n'a pas trouvé de correspondance avec les pays listés dans la séquence, la transformation tente d'analyser les données de téléphone de l'enregistrement comme des données nord-américaines. Pour ignorer ce niveau d'analyse, ne définissez pas l'option <i>Séquence de code pays ISO2</i>. Data Cleanse tente alors d'analyser les données de téléphone comme des données nord-américaines.

Niveau	Processus
Global	Facultatif. Global est l'option par défaut de <i>Séquence de code pays ISO2</i>. Vous pouvez exclure Global de la séquence, l'inclure après une séquence de codes pays ou inclure uniquement Global dans la séquence. Avantage : si vous ne savez pas précisément quels sont les pays représentés dans vos données et/ou si la transformation Global Address Cleanse n'a pas généré le code pays, la transformation recherche tous les pays listés dans le package de nettoyage des données (sauf les pays figurant dans le NANP) dans le cadre de la recherche globale. i Remarque Le niveau Global n'inclut pas les codes pays dans les critères de recherche. Si le code pays figure dans le NANP, Data Cleanse interrompt l'analyse globale et tente d'analyser les données de téléphone comme des données nord-américaines. Si la transformation ne peut pas analyser les données de téléphone en fonction de l'option Global ou si vous avez supprimé Global de l'option <i>Séquence de code pays ISO2</i>, la transformation tente d'analyser les données de téléphone comme des données nord-américaines.

La transformation indique les données de téléphone non analysées dans un champ de sortie supplémentaire si vous l'avez inclus dans la configuration des champs de sortie.

Informations associées

[Configuration de la transformation Data Cleanse \[page 432\]](#)

[Editeur d'options classées \[page 234\]](#)

27.2.5.2 Analyse des dates

Data Cleansereconnait les dates dans de nombreux formats et les décompose en composants.

Data Cleansepeut analyser jusqu'à six dates de l'enregistrement défini. C'est-à-dire que Data Cleanseidentifie jusqu'à six dates à l'entrée, les décompose dans des composants, et les rend disponibles en sortie, soit dans le format d'origine, soit dans le format standard sélectionné par l'utilisateur.

Informations associées

[A propos du mappage un par un \[page 444\]](#)

27.2.5.3 Analyse des numéros de sécurité sociale des Etats-Unis

Data Cleanse analyse les numéros de sécurité sociale des Etats-Unis qui sont seuls ou sur une ligne d'entrée entourés d'un autre texte.

Champs utilisés

Data Cleanse édite chaque composant d'un numéro de sécurité sociale analysé, c'est-à-dire l'ensemble du numéro, la zone, le groupe et la série.

Comment Data Cleanse analyse les numéros de sécurité sociale.

Data Cleanse analyse les numéros de sécurité sociale selon les étapes suivantes :

1. Il identifie un numéro de sécurité sociale potentiel en regardant les modèles suivants :

Modèle	Chiffres par groupe	Délimité par
nnnnnnnnn	9 chiffres consécutifs	non applicable
nnn nn nnnn	3, 2 et 4 (pour zone, groupe et série)	espaces
nnn-nn-nnnn	3, 2 et 4 (pour zone, groupe et série)	Tous les séparateurs pris en charge

2. Il effectue un contrôle de validité sur les cinq premiers chiffres uniquement. Les résultats possibles de ce contrôle de validité sont :

Résultat	Description
Réussite	Data Cleanse a analysé les données et le numéro de sécurité sociale est édité dans un champ de sortie de numéro de sécurité sociale.
Echec	Data Cleanse n'a pas analysé les données parce qu'il ne s'agit pas d'un numéro de sécurité sociale valide tel que défini par le gouvernement américain. Les données sont sorties comme des données supplémentaires, non analysées.

Contrôle de la validité

Lors d'un contrôle de validité, Data Cleanse ne vérifie pas qu'un numéro de sécurité sociale à 9 chiffres particulier soit émis, ou qu'il s'agisse du bon numéro pour la personne nommée. Il valide à la place les 5 premiers chiffres (zone et groupe). Data Cleanse ne valide pas les 4 derniers chiffres (série), sauf pour confirmer que ce sont des chiffres.

Données des Administrations de sécurité sociale

Data Cleanse valide les cinq premiers chiffres en fonction d'une table de l'Administration de la Sécurité Sociale (http://www.socialsecurity.gov/employer/ssns/HGJune2411_final.txt ). Les règles et les données qui guident ce contrôle sont disponibles à l'adresse <http://www.ssa.gov/history/ssn/geocard.html> . Les informations sur le numéro de sécurité sociale référencées par Data Cleanse sont incluses dans le package de nettoyage des données.

i Remarque

L'administration de la sécurité sociale ne met plus la table à jour. La dernière mise à jour date de juillet 2011. Par conséquent, la validation du numéro de sécurité sociale effectuée dans Data Cleanse se base sur les données de juillet 2011. Pour plus d'informations sur le processus d'affectation de numéro de sécurité sociale de l'administration de la sécurité sociale, voir <http://www.ssa.gov/employer/randomizationfaqs.html> .

Sorties de numéros de sécurité sociale valides

Data Cleanse sort uniquement les numéros de sécurité sociale qui ont réussi la validation. Si un numéro de sécurité sociale apparent échoue la validation, Data Cleanse ne transmet pas le numéro comme un numéro de sécurité sociale analysé, mais comme un numéro non valide.

27.2.5.4 Analyse des adresses électroniques

Lorsque Data Cleanse analyse des données d'entrée qu'il détermine être une adresse électronique, il place les composants de ces données dans des champs spécifiques pour la sortie. Ci-dessous se trouve un exemple d'une adresse électronique simple :

joex@sap.com

En identifiant les différents composants des données (nom d'utilisateur, hôte, etc.) par leur relation entre eux, Data Cleanse peut affecter les données à des attributs spécifiques (champs de sortie).

Champs de sortie utilisés par Data Cleanse

Data Cleanse édite les composants individuels d'une adresse électronique analysée, c'est-à-dire le nom d'utilisateur du courrier électronique, le nom de domaine complet, le domaine principal, le deuxième domaine, le troisième domaine, le quatrième domaine, le cinquième domaine et le nom de l'hôte.

Ce que fait Data Cleanse

Data Cleanse peut exécuter les actions suivantes :

- Analyser une adresse électronique située dans un champ discret ou combinée avec d'autres données dans un champ multiligne.
- Répartir le nom du domaine dans ses sous-éléments.
- Vérifier que l'adresse électronique se trouve dans le bon format.
- Indiquer que l'adresse inclut un fournisseur d'accès à internet (FAI) ou un nom de domaine de courrier électronique référencé dans le type de courrier électronique des données de référence dans Data Cleanse. Cet indicateur est affiché dans le champ de sortie Email_is_ISP.

Ce que Data Cleanse ne vérifie pas

Plusieurs aspects d'une adresse électronique ne sont pas vérifiés par Data Cleanse. Data Cleanse ne vérifie pas :

- si le nom du domaine (la partie à droite du signe @) est enregistré.
- si un serveur de courrier électronique est activé à cette adresse.
- si le nom d'utilisateur (la partie à gauche du signe @) est enregistré sur ce serveur de courrier électronique (le cas échéant).
- si le nom personnel dans l'enregistrement peut être atteint à cette adresse électronique.

Composants d'adresse électronique

Le champ de sortie dans lequel Data Cleanse place les données dépend de la position des données dans l'enregistrement. Data Cleanse suit le système de nom de domaine (DNS) pour déterminer le bon champ de sortie.

Par exemple, si `expat@london.home.office.city.co.uk` correspondait à des données d'entrée, Data Cleanse éditerait les éléments dans les champs suivants :

Champs de sortie	Valeur de sortie
Adresse électronique	expat@london.home.office.city.co.uk
Utilisateur_Email	expat

Champs de sortie	Valeur de sortie
Domaine_Messagerie_Complet	london.home.office.city.co.uk
Domaine_Messagerie_Principal	uk
Domaine_Messagerie_Deuxième	co
Domaine_Messagerie_Troisième	city
Domaine_Messagerie_Quatrième	office
Domaine_Messagerie_Cinquième	home
Domaine_Messagerie_Hôte	london

Informations associées

[A propos du mappage un par un \[page 444\]](#)

27.2.5.5 Analyse des modèles définis par l'utilisateur

Data Cleanse peut analyser les modèles trouvés dans une grande variété de données telles que :

- numéros de compte
- numéros de référence
- bons de commande
- numéros de facture
- NIV (numéro d'identification du véhicule)
- numéros du permis de conduire

Autrement dit, Data Cleanse peut analyser toutes les séquences alphanumériques pour lesquelles vous pouvez définir un modèle.

L'analyseur UDPM (correspondance des modèles définis par l'utilisateur) recherche le modèle dans chaque champ complet.

Les modèles sont définis à l'aide d'expressions régulières dans l'onglet [Données de référence](#) du générateur de packages de nettoyage des données. Faites un contrôle avec le propriétaire du package de nettoyage des données pour déterminer tout mappage requis pour les champs d'entrée et les champs de sortie (attributs).

27.2.5.6 Analyse des noms et numéros de rue

Data Cleanse n'identifie et n'analyse pas les composants individuels d'adresses. Pour analyser les données qui contiennent des informations d'adresse, traitez-les avec les transformations Global Address Cleanse ou U.S. Regulatory Address Cleanse avant Data Cleanse. Si les données d'adresse sont traitées par la transformation Data Cleanse, elles sont généralement éditées dans les champs [Supplément](#).

Informations associées

[Fonctionnement d'Address Cleanse \[page 566\]](#)

27.2.5.7 A propos de l'analyse des données de société

Data Cleanse peut analyser des données de société.

Data Cleanse accepte ces noms de société seuls dans un champ ou avec d'autres données.

Une exception sur la manière dont Data Cleanse recombine les parties de mot contiguës est faite pour les mots qui se terminent par un S, tels que Applebee's ou Macy's. Une chaîne d'entrée Macy's est divisée en trois jetons : MACY, ', s. Puisque le dernier jeton est un S, Data Cleanse combine d'abord les jetons et recherche le terme comprenant l'apostrophe (MACY'S). Si le terme n'est pas trouvé, Data Cleanse recherche le terme sans l'apostrophe (MACYS). Si cette opération échoue, Data Cleanse laisse automatiquement les jetons ensemble (MACY'S) et ajoute la classification FIRM_MISCELLANEOUS au terme. Puisque les mots terminant par S restent automatiquement ensemble, il n'est pas nécessaire d'ajouter au dictionnaire tous les noms de société comportant la marque du possessif.

27.2.5.8 A propos de l'analyse des données de nom et de titre

Data Cleanse peut analyser les données de nom et de titre.

Le nom d'une personne peut comprendre les parties suivantes : titre de civilité, prénoms, noms de famille, suffixes du nom, etc.

Data Cleanse peut accepter jusqu'à deux noms et titres comme composants discrets. Data Cleanse accepte également les données de nom et de titre avec d'autres données ou seules dans un champ. La ligne du nom ou le champ multiligne peut contenir un ou deux noms par champ.

27.2.5.9 A propos du mappage un par un

Le [mappage un par un](#) est une option de la transformation Data Cleanse qui contrôle la manière dont plusieurs analyseurs éditent les données.

L'option [Mappage un par un](#) est disponible pour ces analyseurs :

- Date
- Courrier électronique
- Téléphone

Si l'option est définie sur [Oui](#), la transformation Data Cleanse édite les données analysées de certains champs d'entrée discrets vers leurs champs de sortie correspondants. Les champs de sortie sont « réservés » pour les analyses réalisées à partir de certains champs d'entrée discrets. Cette option affiche plus clairement le mappage du champ d'entrée qui contenait les données d'origine d'après les données analysées dans le champ

de sortie. Par exemple, si les données d'entrée de Téléphone1 à Téléphone5 étaient vides et que le champ Téléphone6 contenait des données, à la sortie, les champs Téléphone1 à Téléphone5 sont toujours vides et Téléphone6 contient les données analysées.

Lorsque l'option *Mappage un par un* est définie sur *Oui*, tous les analyseurs qui utilisent cette option sont définis sur *Oui*. Par exemple, il est impossible d'activer l'option *Mappage un par un* uniquement pour l'analyseur de téléphone.

Si l'option *Mappage un par un* est définie sur *Non*, la transformation Data Cleanse analyse et édite les données dans l'ordre où elles sont entrées dans l'analyseur. Les données ne sont pas nécessairement envoyées vers le même champ que celui auquel elles étaient mappées à la sortie. Les données sont éditées dans le premier champ disponible de la catégorie.

i Remarque

Les exemples de cette section illustrent les champs Date. Les mêmes exemples s'appliquent aussi aux champs Téléphone et Adresse électronique.

❖ Exemple

Champ	Données d'entrée	Données de sortie lorsque l'option est <i>Non</i>	Données de sortie lorsque l'option est <i>Oui</i>
Date1	<vide>	01/01/1968	<vide>
Date2	01/01/1968	02/02/1968	01/01/1968
Date3	02/02/1968	03/03/1968	02/02/1968
Date4	<vide>	04/04/1968	<vide>
Date5	03/03/1968	<vide>	03/03/1968
Date6	04/04/1968	<vide>	04/04/1968

Champs multilignes

Les champs discrets Date, Adresse électronique et Téléphone sont analysés avant les champs multilignes afin que tous les champs non réservés puissent contenir des données des champs multilignes.

❖ Exemple

Champ	Données d'entrée	Données de sortie lorsque l'option est <i>Non</i>	Données de sortie lorsque l'option est <i>Oui</i>
Date1	<vide>	01/01/1968	<vide>
Date2	01/01/1968	02/02/1968	01/01/1968
Date3	<vide>	03/03/1968	<vide>
Date4	02/02/1968	04/04/1968	02/02/1968

Champ	Données d'entrée	Données de sortie lorsque l'option est <i>Non</i>	Données de sortie lorsque l'option est <i>Oui</i>
Date5	<vide>	<vide>	03/03/1968 (non réservé, l'entrée multiligne peut être ajoutée ici)
Date6	<vide>	<vide>	04/04/1968 (non réservé, l'entrée multiligne peut être ajoutée ici)
Multiligne1	03/03/1968 04/04/1968	<vide>	<vide>

Champs Supplément

Quand l'option *Mappage un par un* est définie sur *Oui* et que le champ d'entrée contient plusieurs jeux de données, seul le premier jeu de données est placé dans le champ de sortie correspondant. Tous les autres jeux de données sont placés dans le champ Supplément.

Exemple

Champ	Données d'entrée	Données de sortie lorsque l'option est <i>Non</i>	Données de sortie lorsque l'option est <i>Oui</i>
Date1	01/01/1968 02/02/1968 03/03/1968 04/04/1968 05/05/1968 06/06/1968 07/07/1968 08/08/1968	01/01/1968	01/01/1968
Date2	<vide>	02/02/1968	<vide>
Date3	<vide>	03/03/1968	<vide>
Date4	<vide>	04/04/1968	<vide>
Date5	<vide>	05/05/1968	<vide>
Date6	<vide>	06/06/1968	<vide>
Supplément	<vide>	07/07/1968 08/08/1968	02/02/1968 03/03/1968 04/04/1968 05/05/1968 06/06/1968 07/07/1968 08/08/1968

Informations associées

[Analyse des dates \[page 439\]](#)

[Analyse des adresses électroniques \[page 441\]](#)

27.2.6 A propos de la normalisation des données

Les formes standard des variantes individuelles sont définies dans un package de nettoyage des données à l'aide du générateur de packages de nettoyage des données. De plus, la transformation Data Cleanse peut normaliser les données pour que leur format soit plus cohérent. Les caractéristiques de données pouvant être normalisées par la transformation incluent la case, la ponctuation et les abréviations.

27.2.7 Affectation de descriptions de sexes et de titres de civilité.

Chaque variante dans un package de nettoyage des données a un sexe qui lui ait associée.

Par défaut, le sexe est non affecté. Il est possible d'affecter un sexe à la variante dans le mode Avancé du Générateur de packages de nettoyage des données. Les descriptions de sexe sont : masculin - fort, féminin - fort, masculin - faible, féminin - faible et ambigu.

Les variantes dans le package de nettoyage des données personnelles et professionnelles fourni par SAP ont été affectées d'un sexe.

Il est possible d'utiliser la transformation Data Cleanse pour sortir le sexe associé à la variante dans le champ de sortie GENDER.

Le champ de sortie Titre de civilité inclut toujours les titres de civilité qui font partie des données d'entrée de nom. De plus, lorsque l'option *Affecter des titres de civilités* est définie sur *Oui*, Data Cleanse renseigne le champ de sortie PRENAME lorsqu'un sexe de type "Masculin - probabilité élevée" ou "Féminin - probabilité élevée" est affecté à la variante.

Noms doubles

Lorsque les noms doubles sont analysés, Data Cleanse propose quatre descriptions de sexe supplémentaires : multi-nom féminin, multi-nom masculin, multi-nom mixte et multi-nom ambigu. Ces sexes sont générés avec Data Cleanse en fonction des sexes affectés aux deux noms. Le tableau suivant affiche la manière dont les sexes multi-noms sont affectés :

Nom double	Sexe du premier nom	Sexe du deuxième nom	Sexe affecté pour nom double
Bob et Sue Jones	masculin fort	féminin fort	multi-nom mixte
Bob et Tom Jones	masculin fort	masculin fort	Multi-nom masculins
Sue et Sara Jones	féminin fort	féminin fort	Multi-nom féminins
Bob et Pat Jones	masculin fort	Ambigu	Multi-nom ambigu

Prénoms chinois et japonais

Lorsqu'un prénom a été analysé comme le résultat des règles qui combinent intelligemment les caractères du prénom plutôt que d'inclure le nom donné comme une variante dans le package de nettoyage des données, Data Cleanse génère le sexe en combinant le sexe des caractères individuels qui constituent le prénom à l'aide du tableau suivant.

	<i>Féminin fort</i>	<i>Masculin fort</i>	<i>Féminin faible</i>	<i>Masculin faible</i>	<i>Ambigu</i>
<i>Féminin fort</i>	féminin fort	Ambigu	féminin fort	Ambigu	féminin fort
<i>Masculin fort</i>	Ambigu	masculin fort	Ambigu	masculin fort	masculin fort
<i>Féminin faible</i>	féminin fort	Ambigu	Féminin faible	Ambigu	Féminin faible
<i>Masculin faible</i>	Ambigu	masculin fort	Ambigu	Masculin faible	Masculin faible
<i>Ambigu</i>	féminin fort	masculin fort	Féminin faible	Masculin faible	Ambigu

27.2.8 Préparation d'enregistrements pour correspondance

Si vous planifiez un flux de données qui inclut une correspondance, il est recommandé d'utiliser d'abord Data Cleanse pour normaliser les données afin d'améliorer la précision des correspondances. La transformation Data Cleanse doit être en amont de la transformation Match.

La transformation Data Cleanse peut générer des standards de correspondance ou d'autres noms pour de nombreux champs de nom et de société, ainsi que pour tous les champs de sortie personnalisée. Par exemple, Data Cleanse peut affirmer que Patrick et Patricia sont des correspondances potentielles pour le nom Pat. Les standards de correspondance peuvent aider à surmonter deux types de problèmes de correspondance : orthographe différent (Catherine et Katherine) et surnoms (Pat et Patrick).

Cet exemple indique la manière dont Data Cleanse peut préparer des enregistrements pour la correspondance.

Source de données 1

Enregistrement d'entrée	Enregistrement nettoyé	
Intl Marketing, Inc.	Prénom 1	Pat
Pat Smith, directeur de comptabilité	Standards de correspondance	Patrick, Patricia
	Prénom 2	
	Nom de famille	Smith
	Fonction	Directeur de comptabilité
	Société	Intl. Mktg, Inc.

Source de données 2

Enregistrement d'entrée	Enregistrement nettoyé	
Smith, Patricia R.	Prénom 1	Patricia
International Marketing, Incorp.	Standards de correspondance	

Enregistrement d'entrée	Enregistrement nettoyé	
	Prénom 2	R
	Nom de famille	Smith
	Fonction	
	Société	Intl. Mktg, Inc.

Lorsqu'un package de nettoyage des données n'inclut pas un autre nom, le champ de sortie du standard de correspondance pour ce terme est vide. Dans le cas d'une sortie multi-mots telle qu'un nom de société, lorsqu'aucune des variantes dans le nom de société n'a un autre nom, alors la sortie du standard de correspondance est vide. Cependant, si au moins une variante a un autre nom associé, le standard de correspondance est généré à l'aide de l'autre nom de la variante lorsque disponible, et des variantes pour les mots qui n'ont pas un autre nom.

27.2.9 Données propres à une région

27.2.9.1 À propos des domaines

Un domaine décrit un type spécifique de données ou de contenu.

Les domaines étendent les fonctionnalités de Data Cleanse afin de nettoyer correctement les données selon les standards culturels d'une région. Dans un package de nettoyage des données personnelles et professionnelles fourni par SAP, chacun des paramètres régionaux constitue un domaine. Le tableau suivant illustre la manière dont l'analyse des noms peut varier selon la culture :

Domaine	Nom	Sortie analysée			
		Prénom1	Prénom2	Nom_Famille1	Nom_Famille2
Espagnol	Juan C. Sánchez	Juan	C.	Sánchez	
Portugais	João A. Lopes	João		A.	Lopes
Français	Jean-Christophe Rousseau	Jean-Christophe		Rousseau	
Allemand	Hans Joachim Müller	Hans	Joachim	Müller	
Anglais (Etats-Unis et Canada)	James Andrew Smith	James	Andrew	Smith	

Chaque variante est automatiquement affectée au domaine global et peut être affectée à un ou plusieurs autres domaines. Une variante peut avoir une signification différente dans chaque domaine. Autrement dit, les propriétés associées à une variante donnée, comme le format et la classification standard, peuvent être spécifiques à un domaine. Par exemple, la variante AG a des significations différentes en allemand et en anglais des États-Unis. En allemand, AG est une abréviation de « Aktiengesellschaft » (type de société), nettoyée sous la forme "AG", alors qu'en anglais, AG est une abréviation d'Agriculture, nettoyée sous la forme "Ag". Vous pouvez contrôler la manière dont Data Cleanse nettoie vos données en indiquant le ou les domaines que Data Cleanse doit appliquer et dans quel ordre.

i Remarque

Plusieurs domaines sont pris en charge uniquement par les packages de nettoyage des données personnelles et professionnelles de la version 4.1 ou supérieure. Les variantes des packages de nettoyage des données personnalisés et de nettoyage des données personnelles et professionnelles créés avant Information Steward 4.1 sont affectés uniquement au domaine global.

Domaine global

Le domaine global est un domaine de contenu particulier qui contient toutes les variantes et leurs propriétés associées. Si une variante n'est associée à aucune information spécifique au domaine, le domaine global est utilisé comme domaine par défaut.

Lorsque vous ajoutez une nouvelle variante, elle est d'abord ajoutée au domaine global. Si nécessaire, vous pouvez ensuite l'ajouter à d'autres domaines afin d'affecter toute information propre au domaine. L'ajout d'une variante à un domaine autre que le domaine global est nécessaire uniquement si la variante comporte des propriétés comme le sexe, la classification, le format standard, etc., qui diffèrent entre le domaine global et les autres domaines.

Si vous supprimez une variante du domaine global, elle est également supprimée de tous les autres domaines auxquels elle est associée.

Contrôle du nettoyage propre au domaine

L'administrateur ou le testeur de Data Services peut définir la [séquence de domaines de contenu](#) dans la transformation Data Cleanse afin de contrôler la manière dont Data Cleanse analyse vos données spécifiques au domaine, telles que le nom, le sexe, etc. Dans l'exemple ci-dessous, regardez comment le sexe serait appliqué au nom Jean d'après les informations suivantes :

Nom	Domaine	Sexe
Jean	Global	AMBIGU
Jean	Français	MASCULIN_PROBABILITÉ_ÉLEVÉE
Jean	Anglais (États-Unis et Canada)	FÉMININ_PROBABILITÉ_FAIBLE

Si vous ne voulez favoriser aucune propriété propre au domaine, sélectionnez uniquement [GLOBAL](#). Sexe ambigu sera affecté au nom Jean car ni les informations spécifiques au domaine français, ni celles propres au domaine anglais ne sont prises en compte.

Lorsque les données proviennent d'une région à domaine unique, indiquez un domaine suivi de Global. Par exemple, lorsque vous indiquez EN_US suivi de GLOBAL ([EN_US|GLOBAL](#)), le sexe féminin (probabilité faible) sera affecté au nom Jean.

Lorsque les données proviennent d'une région à plusieurs domaines, sélectionnez la séquence préférée des domaines en terminant par Global. Par exemple, le Benelux (Belgique, Pays-Bas, Luxembourg) comprend les

domaines néerlandais, français et allemand. Selon vos préférences de nettoyage, vous pouvez classer les domaines selon la séquence voulue. Par exemple, pour favoriser le domaine néerlandais, indiquez **NL|FR|DE|GLOBAL**. Quand il existe dans les données une variante comportant différentes propriétés dans les domaines de contenu sélectionnés, le processus de nettoyage utilise d'abord les propriétés du néerlandais s'il en trouve. S'il n'existe aucune propriété spécifique au néerlandais, Data Cleanse utilise les propriétés du français, s'il n'existe aucune propriété spécifique ni au néerlandais ni au français, il utilise les propriétés de l'allemand. Si aucun des trois domaines n'a de propriétés spécifiques, Data Cleanse utilise celles indiquées dans le domaine global.

La Suisse est un autre exemple de région à plusieurs domaines. La Suisse comprend les domaines allemand, français et italien. Selon vos préférences de nettoyage, vous pouvez classer les domaines selon la séquence voulue. Par exemple, pour favoriser le domaine allemand, sélectionnez **DE|FR|IT|GLOBAL**. Quand il existe dans les données une variante comportant différentes propriétés dans les domaines de contenu sélectionnés, le processus de nettoyage utilise d'abord les propriétés de l'allemand s'il en trouve. S'il n'existe aucune propriété spécifique à l'allemand, Data Cleanse utilise les propriétés du français, s'il n'existe aucune propriété spécifique ni à l'allemand ni au français, il utilise les propriétés de l'italien. Si aucune des trois n'a de signification particulière, Data Cleanse utilise celle qui existe dans le domaine global.

Spécification de la séquence de domaines de contenu

Vous pouvez définir la séquence de domaines de contenu à trois niveaux différents.

Niveau	Option
Par défaut	Option de la transformation <i>Séquence de domaines de contenu</i> Data Cleanse. Cette option est obligatoire. En cas d'erreurs ou si les options au niveau intermédiaire ou au niveau supérieur ne sont pas spécifiées, cette option est utilisée.
Niveau intermédiaire	Données dans les options d'entrée Option_Pays, Option_Langue, et/ou Option_Région. En cas d'erreurs ou si les options au niveau supérieur ne sont pas spécifiées, cette option est utilisée. Ces champs d'entrée sont renseignés et mappés à partir des champs de sortie Code_pays_ISO_2car, Langue et Région1 de la transformation Global Address Cleanse, respectivement. L'utilisation de ces champs d'entrée est facultative.
Niveau supérieur	Données dans l'option d'entrée Option_séquence_domaines_contenu. Les données de cette option sont affectées via une transformation Query précédente ou par le biais d'une autre étape de pré-traitement. L'utilisation de ce champ d'entrée est facultative.

❖ Exemple

Au cours du traitement, Data Cleanse affecte la séquence de domaines de contenu en fonction de l'option de niveau supérieur, si l'option est définie et valide.

Niveau	Option	Sortie	Notes
Niveau supérieur	Champ d'entrée : Option_séquence_domaines_contenu=FR Global	La séquence de domaines de contenu est le français et toutes les données qui n'ont pas été identifiées comme relevant du français sont définies par défaut dans le domaine Global.	Cette option écrase les paramètres intermédiaires et par défaut, à moins que cette option ne soit pas valide (par exemple, mal orthographiée) ou non définie.
Niveau intermédiaire	Champs d'entrée : Option_Pays=MX Option_Langue=<non défini> Option_Région=<non défini>	La séquence de domaines de contenu est l'espagnol (ES_MX, de l'Espagne au Mexique) et toutes les données qui n'ont pas été identifiées comme relevant de l'espagnol sont définies par défaut dans le domaine Global.	Cette option écrase les paramètres par défaut, à moins que cette option ne contienne des données non valides ou non définies. Dans la plupart des cas, la séquence de domaines de contenu peut être identifiée en fonction du champ d'entrée Option_Pays uniquement. Toutefois, certains pays peuvent utiliser Option_Langue ou Option_Région pour les aider à déterminer la séquence de domaines de contenu.
Par défaut	Option de transformation Data Cleanse : Séquence de domaines de contenu=EN_US Global	La séquence de domaines de contenu est l'anglais (du Royaume-Uni aux États-Unis et au Canada) et toutes les données qui n'ont pas été identifiées comme relevant de l'anglais sont définies par défaut dans le domaine Global.	Cette option est obligatoire et est utilisée en cas d'erreurs détectées dans les options de niveau intermédiaire et de niveau supérieur.
Lors de l'utilisation du niveau intermédiaire, vous découvrirez que certains pays peuvent utiliser des données d'entrée supplémentaires pour déterminer la séquence de domaines de contenu.			
Champ d'entrée Option_Pays	Champs d'entrée supplémentaires	Séquence de domaines de contenu affectée	Notes
CA (Canada)	Option_Région=QC	FR Global	La séquence de domaines de contenu est toujours affectée à FR Global si Option_Région est défini sur Québec. Les données acceptables incluent Québec ou QC, quelle que soit la casse. Toute autre région valide (telle que BC pour la Colombie-britannique) renvoie une séquence de domaines de contenu EN_US Global.
	Option_Région=BC*	EN_US Global	

Champ d'entrée Option_Pays	Champs d'entrée supplémentaires	Séquence de domaines de contenu affectée	Notes
BE (Belgique)	Option_Langue=FR	FR Global	La séquence de domaines de contenu est toujours affectée à FR Global si Option_Langue=FR. Toute autre langue spécifiée (telle que NL pour le néerlandais ou EN_GB pour l'anglais britannique) ou si cette option n'est pas renseignée, renvoie une séquence de domaines de contenu NL Global.
	Option_Langue=NL*	NL Global	
CH (Suisse)	Option_Langue=FR	FR Global	La séquence de domaines de contenu est toujours affectée à FR Global ou à IT Global si Option_Langue=FR ou IT, respectivement. Toute autre langue spécifiée (telle que RU pour le russe), ou si cette option est vide, renvoie la séquence de domaines de contenu DE Global.
	Option_Langue=IT	IT Global	
	Option_Langue=RU*	DE Global	

i Remarque

*La séquence de domaines de contenu est affectée indépendamment du format de sortie.

Informations associées

[À propos du format de sortie \[page 453\]](#)

27.2.9.2 À propos du format de sortie

D'après le domaine indiqué dans le format de sortie, Data Cleanse utilise certains champs de sortie et met en forme les données de ces champs selon les standards régionaux.

Spécifiez le domaine dans la zone *Format de sortie* de la transformation Data Cleanse.

Selon les standards régionaux, dans certains domaines, un prénom composé est combiné et édité dans le champ `Prénom1`, alors que dans d'autres domaines, le premier prénom est édité dans le champ `Prénom1` et le deuxième prénom dans le champ `Prénom2`.

De même, dans certains domaines, un nom de famille composé est combiné et édité dans le champ `Nom de famille1`, alors que dans d'autres domaines, le premier nom de famille est édité dans le champ `Nom de famille1` et le deuxième dans le champ `Nom de famille2`.

Dans certains domaines, le champ de sortie composite `Personne` comprend le prénom suivi du nom de famille alors que dans d'autres domaines, le champ de sortie composite `Personne` comprend le nom de famille suivi du prénom.

La transformation Data Cleanse requiert l'indication d'un format de sortie, même si les données sont vraiment globales.

Lorsque vous avez des données d'une région à domaine unique, indiquez le domaine. Par exemple, sélectionnez **DE** pour l'Allemagne et **ZH** pour la Chine.

Si vous avez des données d'une région à plusieurs domaines, vous devez sélectionner le domaine préféré. Vos données peuvent être mises en forme différemment selon le domaine sélectionné.

Par exemple, les données des Philippines peuvent être éditées selon le format anglais ou espagnol. Comme l'illustre le tableau suivant, le nom **Juan Carlos Sanchez Cruz** sera édité dans des champs différents selon le format de sortie sélectionné.

Champ de sortie	Format de sortie	
	Anglais	Espagnol
Prénom1	Juan	Juan
Prénom2	Carlos	Carlos
Nom de famille1	Sánchez Cruz	Sánchez
Nom de famille2		Cruz
Personne	Juan Carlos Sánchez Cruz	Juan Carlos Sánchez Cruz

Pour les données du Benelux, vous pouvez choisir d'éditer vos données selon le format hollandais, français ou allemand. Comme l'illustre le tableau suivant, le nom **H. D. BUDJHAWAN** sera édité dans des champs différents selon le format de sortie sélectionné.

Champ de sortie	Format de sortie		
	Néerlandais	Français	Allemand
Prénom1	H.D.	H. D.	H.
Prénom2			D.
Nom de famille1	Budjawan	Budjawan	Budjawan
Nom de famille2			
Personne	H.D. Budjawan	H. D. Budjawan	H. D. Budjawan

Vous pouvez modifier le format de sortie existant ou ajouter un format de sortie propre au domaine en modifiant les règles appropriées dans votre package de nettoyage des données.

Spécification du format de sortie

Vous pouvez définir le format de sortie en trois niveaux différents.

Niveau	Option
Par défaut	Option de la transformation Data Cleanse <i>Format de sortie</i> . Cette option est obligatoire. En cas d'erreurs ou si les options au niveau intermédiaire ou au niveau supérieur ne sont pas spécifiées, cette option est utilisée.
Niveau intermédiaire	Données dans les options d'entrée Option_Pays, Option_Langue, et/ou Option_Région. En cas d'erreurs ou si les options au niveau supérieur ne sont pas spécifiées, cette option est utilisée. Ces champs d'entrée sont renseignés et mappés à partir des champs de sortie Code_pays_ISO_2car, Langue et Région1 de la transformation Global Address Cleanse, respectivement. L'utilisation de ces champs d'entrée est facultative.
Niveau supérieur	Données dans l'option d'entrée Option_format_sortie. Les données de cette option sont affectées via une transformation Query précédente ou par le biais d'une autre étape de pré-traitement. L'utilisation de ce champ d'entrée est facultative.

❖ Exemple

Au cours du traitement, Data Cleanse affecte le format de sortie en fonction de l'option de niveau supérieur, si l'option est définie et valide.

Niveau	Option	Sortie	Notes
Niveau supérieur	Champ d'entrée : Option_format_sortie=FR	Le format de sortie est le français.	Cette option écrase les paramètres intermédiaires et par défaut, à moins que cette option ne soit pas valide (par exemple, mal orthographiée) ou non définie.
Niveau intermédiaire	Champs d'entrée : Option_Pays=MX Option_Langue=<non défini> Option_Région=<non défini>	Le format de sortie est l'espagnol (de l'Espagne au Mexique)	Cette option écrase les paramètres par défaut, à moins que cette option ne contienne des données non valides ou non définies. Dans la plupart des cas, le format de sortie peut être identifié en fonction du champ d'entrée Option_Pays uniquement. Toutefois, certains pays peuvent utiliser Option_Langue ou Option_Région pour aider à déterminer le format de sortie.
Par défaut	Option de transformation Data Cleanse : Format_sortie=EN_US	Le format de sortie est l'anglais (du Royaume-Uni aux États-Unis).	Cette option est obligatoire et est utilisée en cas d'erreurs détectées dans les options de niveau intermédiaire et de niveau supérieur.

Lors de l'utilisation du niveau intermédiaire, vous découvrirez que certains pays peuvent utiliser des données d'entrée supplémentaires pour déterminer le format de sortie.

Champ d'entrée Option_Pays	Champs d'entrée supplémentaires	Format de sortie affecté	Notes
CA (Canada)	Option_Région=QC	FR	Le format de sortie est toujours affecté à FR si Option_Région est défini sur Québec. Les données acceptables incluent Québec ou QC, quelle que soit la casse. Toute autre région valide (telle que BC pour la Colombie-britannique) renvoie un format de sortie EN_US.
	Option_Région=BC*	EN_US	

Champ d'entrée Op- tion_Pays	Champs d'entrée supplémentaires	Format de sortie affecté	Notes
BE (Belgique)	Option_Langue=FR	FR	Le format de sortie est toujours affecté à FR si Option_Langue=FR. Toute autre langue spécifiée (telle que NL pour le néerlandais ou EN_GB pour l'anglais britannique), ou si cette option est vide, renvoie un format de sortie NL.
	Option_Langue=NL*	NL	
CH (Suisse)	Option_Langue=FR	FR	Le format de sortie est toujours affecté à FR ou IT si Option_Langue=FR ou IT, respectivement. Toute autre langue spécifiée (telle que RU pour le russe), ou si cette option est vide, renvoie un format de sortie DE.
	Option_Langue=IT	IT	
	Option_Langue=RU*	DE	

i Remarque

*Le format de sortie est affecté indépendamment de la séquence du domaine de contenu.

27.2.9.3 Personnalisation des titres de civilité par pays

Lorsque le nom d'entrée n'inclut pas de titre de civilité, Data Cleanse génère les titres de civilité anglais Mr. et Ms. Pour modifier ces termes, ajoutez une transformation Query après la transformation Data Cleanse et utilisez la fonction search_replace pour remplacer les termes avec les titres de civilité appropriés pour la région.

27.2.9.4 Numéros personnels d'identification

Data Cleanse peut identifier les numéros de sécurité sociale des Etats-Unis et les séparer dans des composants discrets. Si vos données incluent des numéros personnels d'identification autres que des numéros de sécurité sociale des Etats-Unis, il est possible de créer des règles de modèle défini par l'utilisateur pour identifier les numéros. Les règles de modèle défini par l'utilisateur font partie des packages de nettoyage des données et sont définies dans l'onglet [Modifier les données de référence](#) du Générateur de package de nettoyages des données.

Les règles de modèle défini par l'utilisateur sont analysées dans Data Cleanse avec l'analyseur [UDPM](#). Les numéros de sécurité sociale des Etats-Unis sont analysés dans Data Cleanse avec l'analyseur [SSN](#).

27.2.9.5 Largeur du texte dans les champs de sortie

De nombreux caractères japonais sont représentés à la fois dans des formes pleine chasse et demi-chasse. Les caractères latins peuvent être codés soit dans une forme proportionnelle soit dans une forme pleine chasse. Dans les deux cas, la forme pleine chasse requiert plus d'espace que la forme demi-chasse ou la forme proportionnelle.

Pour normaliser les données, il est possible d'utiliser l'option [Style de largeur de caractère](#) pour définir la largeur de caractère pour tous les champs de sortie sur pleine chasse ou demi-chasse. La valeur de largeur normale reflète la largeur de caractère normalisée en fonction du type de script. Ainsi, certains champs de sortie contiennent des caractères de demi-chasse et d'autres champs contiennent des caractères de pleine chasse. Par exemple, tous les caractères latins de pleine chasse sont normalisés en fonction de leurs formes demi-chasse, et tous les caractères katakana de demi-chasse sont normalisés en fonction de leurs formes pleine chasse. NORMAL_WIDTH n'exige aucun traitement spécial, et est donc le paramètre le plus efficace.

i Remarque

Puisque la largeur de sortie est basée sur la largeur normalisée pour le type de caractère, les données de sortie peuvent être plus larges que les données d'entrée. Il est possible que vous ayez à augmenter la largeur de la colonne dans la table cible.

Pour les tables de modèle, le fait de sélectionner la case [Utiliser le type NVARCHAR pour les colonnes VARCHAR dans les bases de données prises en charge](#) remplace le type de colonne VARCHAR par le type NVARCHAR et prend en compte une plus grande taille de données.

27.3 Géocodage

Introduction aux fonctionnalités de la transformation Geocoder.

La transformation Geocoder utilise des coordonnées géographiques exprimées sous forme de données de latitude, longitude, adresses et point de référence (POI). La transformation Geocoder permet d'ajouter des adresses, des données de latitude et de longitude, des données de recensement ainsi que d'autres informations à vos données.

En fonction des champs d'entrée mappés, la transformation Geocoder offre trois modes de géocodage :

- Géocodage d'adresse
- Géocodage inversé
- Recherche textuelle de POI

Généralement, vous utilisez la transformation Geocoder avec la transformation Global Address Cleanse ou USA Regulatory Address Cleanse.

i Remarque

Nous supprimons la fonctionnalité GeoCensus de la transformation USA Regulatory Address Cleanse pour inclure les fonctionnalités dans la transformation Geocoder. Assurez-vous de mettre à niveau les flux de données existants utilisant les fonctionnalités GeoCensus dans la transformation USA Regulatory Address Cleanse vers la transformation Geocoder. Vous trouverez des instructions sur la mise à niveau de GeoCensus vers la transformation Geocoder dans le *Guide de mise à niveau*.

Informations associées

[GeoCensus \(USA Regulatory Address Cleanse\) \[page 652\]](#)

27.3.1 Géocodage d'adresse

En mode de géocodage des adresses, la transformation Geocoder affecte des données géographiques. En fonction du degré d'intégralité des données d'adresse d'entrée, la transformation Geocoder peut renvoyer plusieurs niveaux de données de latitude et de longitude. Le fait d'ajouter au début les informations de latitude et de longitude à vos données peut aider l'entreprise à cibler certaines tailles de population et d'autres données géographiques régionales.

Préparer des enregistrements pour le géocodage

La transformation Geocoder fonctionne mieux lorsqu'elle a standardisé et corrigé les données d'adresse, donc pour obtenir les informations les plus précises, placez une transformation Address Cleanse avant la transformation Geocoder dans le workflow.

27.3.1.1 Niveaux de latitude et longitude

Niveau numéro principal

Si vos données d'entrée possèdent une adresse complète (y compris le numéro principal), la transformation Geocoder renvoie les coordonnées de latitude et de longitude à l'emplacement exact en fonction du type de répertoire (basé sur l'échelle ou la parcelle de terrain) auquel vous vous inscrivez.

En règle générale, la transformation Geocoder utilise des répertoires de géocodage pour calculer les valeurs de latitude et de longitude d'une adresse par interpolation du point de départ et du point d'arrivée d'un segment de ligne correspondant à une plage d'adresses. Les valeurs de latitude et de longitude peuvent présenter un léger décalage par rapport à l'emplacement exact de la maison. On nomme cela le niveau d'affectation PRI (primary range interpolated).

i Remarque

Si vous souhaitez obtenir une affectation au niveau de l'adresse, le champ d'entrée Numéro_principal doit être mappé et ne peut pas être vide.

La transformation Geocoder prend également en charge les répertoires de parcelle de terrain, qui contiennent les valeurs de latitude et de longitude les plus précises et exactes, en fonction des données de pays disponibles. Les données relatives aux parcelles de terrain sont stockées sous forme de points. Elles ne vous mènent pas à proximité de la maison, mais juste devant la porte. On nomme cela le niveau d'affectation PRE (primary range exact).

Niveau centroïde Code postal

Si une adresse possède un code postal, vous recevez des coordonnées dans la zone de code postal appropriée. La transformation Geocoder possède les niveaux d'affectation Code postal complet (PF), Code postal partiel2 (P2P) et Code postal1 (P1), selon les données de pays disponibles.

Niveau centroïde Localité

Si une adresse possède une localité, vous recevez des coordonnées dans la zone de localité appropriée. La transformation Geocoder possède les niveaux d'affectation Localité1 (L1), Localité2 (L2), Localité3 (L3) et Localité4 (L4), selon les données de pays disponibles.

27.3.1.2 Mappage du champ Géocodage d'adresse

Les tableaux suivants spécifient les champs d'entrée et de sortie requis, facultatifs ou ceux qui ne peuvent pas être mappés en mode Géocodage d'adresse. Plus vous pouvez fournir des données d'entrée, meilleurs sont les résultats que vous obtenez.

Mappage du champ d'entrée

Catégorie du champ d'entrée	Mode Géocodage d'adresse
Adresse	Au moins un champ est requis
Adresse POI	Facultatif
Latitude/Longitude	n/a
Enregistrements max	n/a
Filtre de recherche	n/a

Mappage du champ de sortie

Catégorie du champ de sortie	Mode Géocodage d'adresse
Adresse	n/a
Adresse POI	n/a
Niveau d'affectation	Facultatif
Recensement	Facultatif

Catégorie du champ de sortie	Mode Géocodage d'adresse
Distance	n/a
Code d'info	Facultatif
Latitude/Longitude	Facultatif
Population	Facultatif
Résultats	n/a
Côté de la rue	Facultatif
Code de statut	Facultatif

27.3.1.3 Scénario de géocodage d'adresse

Scénario : utilisez une adresse ou une adresse et un point d'intérêt pour affecter les informations de latitude et de longitude.

Nombre de résultats de sortie : un seul enregistrement

Les sections suivantes décrivent les champs d'entrée requis et facultatif, et les champs de sortie disponibles pour obtenir des résultats pour ce scénario. Nous fournissons également un exemple avec un exemple de données.

Champs d'entrée obligatoires

Pour les champs d'entrée requis, le champ Pays doit être mappé. Plus vous pouvez fournir des données d'entrée, meilleurs sont les résultats que vous obtenez.

Catégorie	Nom du champ d'entrée
Adresse	Pays (requis)
	Localité1-4
	Code_postal1-2
	Nom_principal1-4
	Numéro_principal
	Suffixe_principal1
	Préfixe_principal1
	Type_principal1-4
	Région1-2

Champs d'entrée facultatifs

Catégorie	Nom du champ d'entrée
Adresse POI	Nom_POI
	Type_POI

Champs de sortie disponibles

Tous les champs de sortie sont facultatifs

Catégorie	Nom du champ de sortie
Niveau d'affectation	Niveau_affectation
	Niveau_affectation_localité
	Niveau_affectation_code_postal
Recensement	Ilot_secteur_recensement
	Groupe_ilot_secteur_recensement
	Code_comté_gouv
	Code_localité1_gouv
	Code_région1_gouv
	Metro_Stat_Area_Code
	Code_division_secondaire
	Code_zone_statistique
Code d'info	Code_info
Latitude/Longitude	Latitude
	Latitude_localité
	Latitude_code_postal
	Latitude_numéro_principal
	Longitude
	Longitude_localité
	Longitude_code_postal
Population	Longitude_numéro_principal
	Localité_classe_population1
Côté de la rue	Côté_adresse_principale
Code de statut	Code_Statut

Exemple

Entrée : vous mappez des champs d'entrée contenant les données suivantes :

Nom du champ d'entrée	Valeur d'entrée
Pays	Etats-Unis
Code postal1	54601
Code postal2	4023
Nom_principal1	Devant
Numéro_principal	332
Type_principal1	Rue

Sortie : les champs de sortie mappés contiennent les résultats suivants :

Nom du champ de sortie	Valeur de sortie
Niveau_affectation	PRE
Latitude	43.811616
Longitude	-91.256695

Informations associées

[Compréhension de la sortie \[page 475\]](#)

27.3.2 Géocodage inversé

En mode Géocodage inversé, la transformation Geocoder identifie l'adresse la plus proche ou le point de référence basé sur l'emplacement de référence d'entrée. Basé sur le mappage d'entrée, le géocodage inversé peut s'opérer en quatre modes :

Mode	Description
<i>Attribuer à une entrée d'adresse une sortie d'enregistrement simple.</i>	Utilisez l'adresse et les champs de POI d'adresse facultative pour déterminer un point de référence unique pour ne générer que l'enregistrement unique le plus proche qui correspond au filtre de recherche. Si les champs de filtre de recherche ne sont pas mappés, la transformation Geocoder désigne par défaut le point de référence ou l'adresse la plus proche.

Mode	Description
<i>Attribuer à une entrée d'adresse une sortie d'enregistrement multiple.</i>	Utilisez l'adresse et les champs de POI d'adresse facultative pour déterminer un point de référence unique pour générer les données du point de référence vers les champs de sortie et les enregistrements les plus proches qui correspondent aux champs de filtre de recherche dans le champ Liste_résultats. Le nombre d'enregistrements est déterminé par le champ d'entrée Option_enregistrements_max s'il est mappé et renseigné, ou par l'option Enregistrements max.
<i>Attribuer à une entrée de latitude/longitude une sortie d'enregistrement simple.</i>	Utilisez les champs de latitude et longitude comme point de référence pour ne générer que l'enregistrement unique le plus proche qui correspond aux champs de filtre de recherche facultatifs. Si les champs de filtre de recherche ne sont pas mappés, la transformation Geocoder désigne par défaut le point de référence ou l'adresse la plus proche.
<i>Attribuer à une entrée de latitude/longitude une sortie d'enregistrement multiple.</i>	Utilisez les champs de latitude et longitude comme point de référence pour générer les enregistrements les plus proches qui correspondent aux champs de filtre de recherche facultatifs dans le champ Liste_Résultats. Le nombre d'enregistrements est déterminé par le champ d'entrée Option_enregistrements_max s'il est mappé et renseigné, ou par l'option Enregistrements max. Si les champs de filtre de recherche ne sont pas mappés, la transformation Geocoder désigne par défaut le point de référence ou l'adresse la plus proche.

i Remarque

- Le mappage des champs d'entrée Option_Radius et Option_Distance_Unit permet de définir la distance par rapport au point de référence spécifié et d'identifier une zone dans laquelle des enregistrements de correspondance se trouvent.
- Pour rechercher un ou plusieurs emplacements de type point d'intérêt et/ou adresse, configurez les champs d'entrée Nom_filtre_recherche ou Type_filtre_recherche. Ce paramétrage limite les correspondances de sortie aux critères de recherche.
- Pour renvoyer uniquement une adresse de non-point d'intérêt, saisissez ADDR dans le champ d'entrée Type_filtre_recherche.
- Pour renvoyer un point d'intérêt uniquement, saisissez son nom ou son type.
- Si vous ne définissez aucun filtre de recherche, la transformation renvoie à la fois les adresses et les points d'intérêt.

27.3.2.1 Mappage du champ Géocodage inversé

Les tableaux suivants spécifient les champs d'entrée et de sortie requis, facultatifs ou ceux qui ne peuvent pas être mappés en mode Géocodage inversé. Plus vous pouvez fournir des données d'entrée, meilleurs sont les résultats que vous obtenez.

Mappage du champ d'entrée

Catégorie du champ d'entrée	Entrée d'adresse	Entrée d'adresse	Entrée de latitude/longitude	Entrée de latitude/longitude
	Sortie simple	Sortie multiple	Sortie simple	Sortie multiple
Adresse	Au moins un champ est requis	Au moins un champ est requis	n/a	n/a
Adresse POI	Facultatif	Facultatif	n/a	n/a
Latitude/Longitude	n/a	n/a	Au moins un champ est requis	Au moins un champ est requis
Enregistrements max	n/a	Au moins un champ est requis	n/a	Au moins un champ est requis
Filtre de recherche	Au moins un champ est requis	Au moins un champ est requis	Facultatif	Facultatif

Mappage du champ de sortie

Catégorie du champ de sortie	Entrée d'adresse	Entrée d'adresse	Entrée de latitude/longitude	Entrée de latitude/longitude
	Sortie simple	Sortie multiple	Sortie simple	Sortie multiple
Adresse	Facultatif	n/a	Facultatif	n/a
Adresse POI	Facultatif	n/a	Facultatif	n/a
Niveau d'affectation	Facultatif	Facultatif	Facultatif	n/a
Recensement	Facultatif	n/a	Facultatif	n/a
Distance	Facultatif	n/a	Facultatif	n/a
Code d'info	Facultatif	Facultatif	Facultatif	Facultatif
Latitude/Longitude	Facultatif	Facultatif	Facultatif	n/a
Population	Facultatif	n/a	Facultatif	n/a
Résultats	n/a	Facultatif	n/a	Facultatif
Côté de la rue	Facultatif	Facultatif	Facultatif	n/a

27.3.2.2 Géocodage inversé, scénario 1

Utilisez la latitude et de la longitude pour rechercher un ou plusieurs points d'intérêt ou adresses.

❖ Exemple

L'exemple suivant illustre un scénario qui fait appel à la latitude, à la longitude et à un filtre de recherche pour générer le point d'intérêt le plus proche de la latitude et de la longitude d'entrée.

Entrée : vous mappez des champs d'entrée contenant les données suivantes :

Nom du champ d'entrée	Valeur d'entrée
Latitude	43.811616
Longitude	-91.256695
Nom_filtre_recherche	BUSINESS OBJECTS

Sortie : les champs de sortie mappés contiennent les résultats suivants :

Nom du champ de sortie	Valeur de sortie
Ligne_adresse	332 FRONT ST
Niveau_affectation	PRE
Code_pays	Etats-Unis
Distance	1.3452
Localité1	LA CROSSE
Code postal	54601-4023
Code postal1	54601
Code postal2	4023
Nom_principal1	FRONT
Numéro_principal	332
Type_principal1	ST
Nom_POI	BUSINESS OBJECTS
Type_POI	5800
Région1	WI

Informations associées

[Champs d'entrée obligatoires \[page 466\]](#)

[Champs de sortie disponibles \[page 466\]](#)

[Champs d'entrée facultatifs \[page 466\]](#)

27.3.2.2.1 Champs d'entrée obligatoires

Scénario 1

Pour un résultat sur un seul enregistrement, les champs d'entrée Latitude et Longitude doivent tous deux être mappés. Pour des résultats sur plusieurs enregistrements, les champs d'entrée Latitude, Longitude et Option_Enregistrements_Max doivent tous être mappés.

Catégorie	Résultats sur un seul enregistrement	Résultats sur plusieurs enregistrements
	Nom du champ d'entrée	Nom du champ d'entrée
Latitude/Longitude	Latitude	Latitude
	Longitude	Longitude
Enregistrements max	Sans objet	Option_enregistrements_max

27.3.2.2.2 Champs d'entrée facultatifs

Scénario 1

Catégorie	Résultats sur un seul enregistrement	Résultats sur plusieurs enregistrements
	Nom du champ d'entrée	Nom du champ d'entrée
Filtre de recherche	Option Rayon	Option_rayon
	Nom_filtre_recherche	Nom_filtre_recherche
	Type_filtre_recherche	Type_filtre_recherche

27.3.2.2.3 Champs de sortie disponibles

Scénario 1

Tous les champs de sortie sont facultatifs

Catégorie	Résultats sur un seul enregistrement	Résultats sur plusieurs enregistrements
	Nom du champ de sortie	Nom du champ de sortie
Adresse	Ligne_Adresse	Sans objet
	Code_pays	
	Localité1-4	
	Code postal	
	Code_postal1-2	
	Nom_principal1-4	
	Numéro_Principal	
	Suffixe_principal1	
	Préfixe_principal1	
	Plage_principale_élevée	
	Plage_principale_faible	
	Type_principal1-4	
	Région1-2	
Adresse POI	Nom_POI	Sans objet
	Type_POI	
Niveau d'affectation	Niveau_affectation	Sans objet
	Niveau_affectation_localité	
	Niveau_affectation_code_postal	
Recensement	Illet_secteur_recensement	Sans objet
	Groupe_illet_secteur_recensement	
	Code_comté_gouv	
	Code_localité1_gouv	
	Code_région1_gouv	
	Code_zone_stat_metro	
	Code_division_secondaire	
	Code_zone_stat	
Distance	Distance	Sans objet
Code d'info	Code_info	Code_info

Catégorie	Résultats sur un seul enregistrement	Résultats sur plusieurs enregistrements
	Nom du champ de sortie	Nom du champ de sortie
Latitude/Longitude	Latitude	Sans objet
	Latitude_localité	
	Latitude_code_postal	
	Latitude_numéro_principal	
	Longitude	
	Longitude_localité	
	Longitude_code_postal	
	Longitude_numéro_principal	
Population	Localité_classe_population1	Sans objet
Résultats	Sans objet	Liste_résultats
		Nombre_liste_résultats
Côté de la rue	Côté_adresse_principale	Sans objet

27.3.2.3 Géocodage inversé, scénario 2

Utilisez un point d'intérêt ou une adresse afin de rechercher un ou plusieurs points d'intérêt ou adresses les plus proches.

Par ailleurs, la transformation Geocoder génère les informations de latitude et longitude pour le point de référence d'entrée et les résultats de sortie correspondants.

❖ Exemple

L'exemple suivant illustre un scénario qui fait appel à une adresse et à un filtre de recherche pour générer le point d'intérêt le plus proche de l'adresse d'entrée. La requête se trouve dans la banque (POI de type 6 000) la plus proche de l'emplacement d'entrée. La transformation génère également les informations de latitude et longitude pour le résultat de sortie.

Entrée : vous mappez des champs d'entrée contenant les données suivantes :

Nom du champ d'entrée	Valeur d'entrée
Pays	Etats-Unis
Code postal1	55601
Code postal2	4023
Nom_principal1	Devant
Numéro_Principal	332

Nom du champ d'entrée	Valeur d'entrée
Type_principal1	Rue
Nom_POI	BUSINESS OBJECTS
Type_filtre_recherche	6000

Sortie : les champs de sortie mappés contiennent les résultats suivants :

Nom du champ de sortie	Valeur de sortie
Ligne_Adresse	201 MAIN ST
Niveau_affectation	PRE
Code_pays	Etats-Unis
Distance	0,4180
Localité1	LA CROSSE
Nom_POI	US BANK
Type_POI	6000
Code postal1	54601
Nom_principal1	MAIN
Numéro_Principal	201
Type_principal1	ST
Région1	WI

Informations associées

[Champs d'entrée obligatoires \[page 469\]](#)

[Champs d'entrée facultatifs \[page 470\]](#)

[Champs de sortie disponibles \[page 470\]](#)

27.3.2.3.1 Champs d'entrée obligatoires

Scénario 2

Pour les champs d'entrée obligatoires, au moins un champ d'entrée dans chaque catégorie doit être mappé.

Catégorie	Résultats sur un seul enregistrement	Résultats sur plusieurs enregistrements
	Nom du champ d'entrée	Nom du champ d'entrée
Adresse	Ligne_Adresse	Latitude
	Pays	Longitude
	Localité1-4	
	Code postal	
	Code_postal1-2	
	Nom_principal1-4	
	Numéro_Principal	
	Suffixe_principal1	
	Préfixe_principal1	
	Plage_principale_élevée	
	Plage_principale_faible	
	Type_principal1-4	
	Région1-2	
Enregistrements max	Sans objet	Option_enregistrements_max
Filtre de recherche	Option_rayon	Option_rayon
	Nom_filtre_recherche	Nom_filtre_recherche
	Type_filtre_recherche	Type_filtre_recherche

27.3.2.3.2 Champs d'entrée facultatifs

Scénario 2

Catégorie	Résultats sur un seul enregistrement	Résultats sur plusieurs enregistrements
	Nom du champ d'entrée	Nom du champ d'entrée
Adresse POI	Nom_POI	Nom_POI
	Type_POI	Type_POI

27.3.2.3.3 Champs de sortie disponibles

Scénario 2

Tous les champs de sortie sont facultatifs

Pour un résultat sur un seul enregistrement, les champs de sortie sont les résultats de la recherche spatiale.

Pour des résultats sur plusieurs enregistrements, les champs de sortie dans les catégories Niveau d'affectation et Latitude/Longitude sont les résultats de l'affectation d'adresse de référence. Les champs de sortie dans la catégorie Résultats sont les résultats du géocodage inversé.

Catégorie	Résultats sur un seul enregistrement	Résultats sur plusieurs enregistrements
	Nom du champ de sortie	Nom du champ de sortie
Adresse	Ligne_Adresse	Sans objet
	Code_pays	
	Localité1-4	
	Code postal	
	Code_postal1-2	
	Nom_principal1-4	
	Numéro_Principal	
	Suffixe_principal1	
	Préfixe_principal1	
	Plage_principale_élevée	
	Plage_principale_faible	
	Type_principal1-4	
	Région1-2	
Adresse POI	Nom_POI	Sans objet
	Type_POI	
Niveau d'affectation	Niveau_affectation	Niveau_affectation
	Niveau_affectation_localité	Niveau_affectation_localité
	Niveau_affectation_code_postal	Niveau_affectation_code_postal
Recensement	Ilot_secteur_recensement	Sans objet
	Groupe_ilot_secteur_recensement	
	Code_comté_gouv	
	Code_localité1_gouv	
	Code_région1_gouv	
	Code_zone_stat_metro	
	Code_division_secondaire	
	Code_zone_stat	
Distance	Distance	Sans objet

Catégorie	Résultats sur un seul enregistrement	Résultats sur plusieurs enregistrements
	Nom du champ de sortie	Nom du champ de sortie
Code d'info	Code_info	Code_info
Latitude/Longitude	Latitude	Latitude
	Latitude_localité	Latitude_localité
	Latitude_code_postal	Latitude_code_postal
	Latitude_numéro_principal	Latitude_numéro_principal
	Longitude	Longitude
	Longitude_localité	Longitude_localité
	Longitude_code_postal	Longitude_code_postal
	Longitude_numéro_principal	Longitude_numéro_principal
Population	Localité_classe_population1	Sans objet
Résultats	Sans objet	Liste_résultats
		Nombre_liste_résultats
Côté de la rue	Côté_adresse_principale	Côté_adresse_principale

27.3.3 Recherche textuelle de POI

En mode Recherche textuelle de POI, la transformation Geocoder utilise des champs d'adresse et un nom de POI ou des champs de type sous forme de critères de recherche pour une mise en correspondance avec des points de référence. Les résultats sont affichés dans le champ de sortie XML Liste_Résultats.

Le nombre d'enregistrements est déterminé par le champ d'entrée Option_enregistrements_max s'il est mappé et renseigné, ou par l'option Enregistrements max.

27.3.3.1 Mappage du champ de recherche textuelle de POI

Les tableaux suivants spécifient les champs d'entrée et de sortie requis, facultatifs ou ceux qui ne peuvent pas être mappés en mode Recherche textuelle de POI. Plus vous pouvez fournir des données d'entrée, meilleurs sont les résultats que vous obtenez.

Mappage du champ d'entrée

En mode Recherche textuelle de POI, au moins un champ d'entrée de la catégorie de POI d'adresse doit être mappé, ce qui représente la différence majeure entre les modes Recherche textuelle de POI et Géocodage d'adresse.

Catégorie du champ d'entrée	Mode Géocodage d'adresse
Adresse	Au moins un champ est requis
Adresse POI	Au moins un champ est requis
Latitude/Longitude	n/a
Enregistrements max	Au moins un champ est requis
Filtre de recherche	n/a

Mappage du champ de sortie

Catégorie du champ de sortie	Mode Géocodage d'adresse
Adresse	n/a
Adresse POI	n/a
Niveau d'affectation	n/a
Recensement	n/a
Distance	n/a
Code d'info	Facultatif
Latitude/Longitude	n/a
Population	n/a
Résultats	Facultatif
Côté de la rue	n/a

27.3.3.2 Scénario de recherche textuelle de POI

Scénario : utilisez des informations d'adresse et de point d'intérêt pour identifier une liste des correspondances de points d'intérêt potentielles.

Nombre de résultats de sortie : plusieurs enregistrements. Le nombre d'enregistrements est déterminé par le champ d'entrée Option_Enregistrements_Max (si rempli), ou l'option Enregistrements max.

Les sections suivantes décrivent les champs d'entrée requis, et les champs de sortie disponibles pour obtenir des résultats pour ce scénario. Nous fournissons également un exemple avec un exemple de données.

Champs d'entrée obligatoires

Pour les champs d'entrée obligatoires, un champ d'entrée au moins doit être mappé dans chaque catégorie. Le champ Pays doit être mappé. Plus vous pouvez fournir des données d'entrée, meilleurs sont les résultats que vous obtenez.

Catégorie	Nom du champ d'entrée
Adresse	Pays (requis)
	Localité1-4
	Code_postal1-2
	Nom_principal1-4
	Numéro_principal
	Suffixe_principal1
	Préfixe_principal1
	Type_principal1-4
Adresse POI	Région1-2
	Nom_POI
	Type_POI
Enregistrements max	Option_enregistrements_max

Champs d'entrée facultatifs

Non applicable

Champs de sortie disponibles

Tous les champs de sortie sont facultatifs

Catégorie	Nom du champ de sortie
Code d'info	Code_info
Résultats	Liste_résultats
	Nombre_liste_résultats

Exemple

L'exemple suivant illustre un scénario utilisant la recherche textuelle de POI pour identifier des correspondances de point de référence potentielles (tous les enregistrements "Business Objects" dans Front Street).

Entrée : vous mappez des champs d'entrée contenant les données suivantes :

Nom du champ d'entrée	Valeur d'entrée
Pays	Etats-Unis
Code postal1	54601
Code postal2	4023
Nom_principal1	Devant
Type_principal1	Rue
Nom_POI	BUSINESS OBJECTS
Enregistrements_max_facultatifs	10

Sortie : les champs de sortie mappés contiennent les résultats suivants avec un enregistrement :

Nom du champ de sortie	Valeur de sortie
Liste_résultats	Sortie comme XML, exemple ci-dessous
Nombre_liste_résultats	1

Liste_Résultats XML : le résultat XML pour cet exemple a un seul enregistrement.

```
<RESULT_LIST>
  <RECORD>
    <ADDRESS_LINE>332 FRONT ST</ADDRESS_LINE>
    <ASSIGNMENT_LEVEL>PRE</ASSIGNMENT_LEVEL>
    <COUNTRY_CODE>US</COUNTRY_CODE>
    <DISTANCE>0.3340</DISTANCE>
    <LATITUDE>43.811616</LATITUDE>
    <LOCALITY1>LA CROSSE</LOCALITY1>
    <LONGITUDE>-91.256695</LONGITUDE>
    <POI_NAME>BUSINESS OBJECTS</POI_NAME>
    <POI_TYPE>5800</POI_TYPE>
    <POSTCODE>56001-4023</POSTCODE>
    <POSTCODE1>56001</POSTCODE1>
    <POSTCODE2>4023</POSTCODE2>
    <PRIMARY_NAME1>FRONT</PRIMARY_NAME1>
    <PRIMARY_NUMBER>332</PRIMARY_NUMBER>
    <PRIMARY_TYPE1>ST</PRIMARY_TYPE1>
    <RANKING>1</RANKING>
    <REGION1>WI</REGION1>
  </RECORD>
</RESULT_LIST>
```

27.3.4 Compréhension de la sortie

Informations relatives aux données de sortie Geocoder dont la latitude et la longitude, le niveau d'affectation et la sortie multi-enregistrements.

Latitude et longitude

Les données de latitude et de longitude sont représentées par des degrés décimaux tels que 12.12345, par exemple. Les décimales indiquent le nombre de degrés depuis l'équateur. La latitude représente des

emplacements au nord ou au sud de l'équateur. La longitude représente des emplacements à l'est ou à l'ouest de l'équateur.

- Latitude : de 0 à 90 degrés au nord ou au sud de l'équateur.
 - Les nombres négatifs indiquent le sud de l'équateur.
 - Les nombres positifs indiquent le nord de l'équateur.
- Longitude : de 0 à 180 degrés à l'est ou à l'ouest du méridien de Greenwich à Londres, Angleterre.
 - Les nombres négatifs indiquent un emplacement compris dans les 180 degrés à l'ouest de Greenwich.
 - Les nombres positifs indiquent un emplacement compris dans les 180 degrés à l'est de Greenwich.

Niveau d'affectation

Pour comprendre la précision de l'affectation, il vous suffit de vous reporter au champ de sortie Niveau_Affectation.

Champ de sortie	Description
PRE	Premier niveau d'affectation. L'affectation la plus précise disponible pour l'emplacement exact.
PRI	Deuxième niveau d'affectation. L'affectation la plus précise disponible pour la plage d'adresses principale ou le numéro de rue.
P1 L1	Niveau d'affectation le plus général. Affectation générale basée sur le code postal, P1, ou la localité, L1. P1 ou L1 est généré en fonction de la valeur choisie dans l'option Meilleur niveau d'affectation.

Résultats multiples

Lorsque la sortie consiste en plusieurs enregistrements, les données du champ de sortie Liste_Résultats sont basées sur les données disponibles. Le format de sortie est une chaîne XML.

Données disponibles

Catégorie	Nom du champ de sortie
Adresse	Ligne_Adresse
	Code_pays
	Localité1-4
	Code postal
	Code_postal1-2
	Nom_principal1-4
	Numéro_Principal
	Suffixe_principal1
	Préfixe_principal1
	Plage_principale_élevée
	Plage_principale_faible
	Type_principal1-4
	Région1-2
Adresse POI	Nom_POI
	Type_POI
Niveau d'affectation	Niveau_Affectation
Distance	Distance
Latitude/Longitude	Latitude
	Longitude
Classement	Classement

Informations associées

[Standardisation des données de sortie Geocoder \[page 477\]](#)

27.3.5 Standardisation des données de sortie Geocoder

Les données de géocode de sortie de Geocoder doivent être standardisées à l'aide des transformations Global Address Cleanse ou USA Regulatory Address Cleanse.

La sortie de données d'adresse de la transformation Geocoder doit être standardisée car les données fournisseur utilisées par Geocoder pour affecter des données de géocode ne sont pas standardisées. Avant que les données d'adresse de sortie géographique ne soient utiles pour poursuivre le traitement, standardisez

les données de sortie en suivant la transformation Geocoder avec la transformation de nettoyage d'adresse applicable.

Lorsque vous disposez de résultats de sortie d'enregistrement multiple, la sortie est une chaîne XML qui peut être imbriquée. Supprimez l'imbrication des données de sortie XML avant de les envoyer à une transformation en aval Global Address Cleanse ou USA Regulatory Address Cleanse.

27.3.6 Utilisation d'autres transformations

En règle générale, la transformation Geocoder est combinée à la transformation Global Address Cleanse ou USA Regulatory Address Cleanse

Transformation Global Address Cleanse

Pour que la transformation Geocoder puisse fournir les informations les plus précises possible, assurez-vous que les champs de sortie de la transformation Global Address Cleanse sont bien mappés aux champs d'entrée de la transformation Geocoder comme suit :

Global Address Cleanse (Sortie)	Geocoder (champ d'entrée)
Code_ISO_pays_2Car	Pays
Localité1	Localité1
Code postal 1	Code postal1
Code postal2	Code postal2
Nom_principal1	Nom_principal1
Numéro_principal	Numéro_principal
Suffixe_principal1	Suffixe_principal1
Préfixe_principal1	Préfixe_principal1
Type_principal1	Type_principal1
Symbole_région1	Région1

Il est également recommandé de définir les options de standardisation de la transformation Global Address Cleanse comme suit :

Global Address Cleanse (option de standardisation)	Valeur
Affecter localité	Convert
Style de direction	Short
Style de type principal	Short

→ Conseil

Adresse postale/Adresse administrative : La transformation Global Address Cleanse utilise les adresses postales, tandis que les fournisseurs de la transformation Geocoder, NAVTEQ et TOMBOM, utilisent des

adresses administratives. Malgré l'ajustement postal réalisé dans l'adresse postale, des écarts subsistent. Par exemple, pour une adresse canadienne, la localité associée à la transformation Global Address Cleanse peut être TORONTO, mais le répertoire Geocoder stocke NORTH YORK. Dans ce cas, lorsque le champ Localité1 est mappé au niveau de l'entrée, une erreur de type Adresse introuvable survient. Pour contourner ce problème, il est possible de supprimer le mappage du champ d'entrée Localité1 pour obtenir une meilleure affectation de géocodage.

27.4 Match

27.4.1 Correspondance et consolidation à l'aide de Data Services et Information Steward

Lorsqu'une vérification manuelle est nécessaire, utilisez SAP Data Services pour implémenter automatiquement le processus global de correspondance et de consolidation. Ce processus organise les groupes de correspondance pour la vérification manuelle dans SAP Information Steward.

Pour implémenter ce processus, tenez compte des facteurs suivants :

- Organisez-vous uniquement les groupes de correspondance suspects (groupes qui ne répondent pas au score ni au seuil de correspondance que vous définissez) pour une vérification manuelle ? Ou est-ce que vous organisez tous les résultats de correspondance, y compris les enregistrements uniques pour la vérification de correspondance, puis laissez les administrateurs de données déplacer manuellement les enregistrements suspects et les enregistrements uniques vers le groupe auquel ils appartiennent ? Notez que le groupe cible dans ce cas doit être un groupe suspect ou un groupe « de fiabilité élevée ».
- Voulez-vous autoriser la création des meilleurs enregistrements dans Information Steward ? Ou voulez-vous que la vérification manuelle de Information Steward modifie uniquement les groupes de correspondance suspects si nécessaire et que Data Services traite la création des meilleurs enregistrements une fois la vérification manuelle terminée ?
- Si vous souhaitez activer la création des meilleurs enregistrements dans Information Steward, à quel moment procédez-vous à la création des meilleurs enregistrements pour les groupes de correspondance de fiabilité élevée ?
 - La création manuelle de meilleurs enregistrements pour tous les groupes de correspondance dans un vaste ensemble de données peut ne pas être pratique.
 - Si vous avez déjà consolidé les groupes de correspondance de fiabilité élevée dans Data Services et que vous avez déjà modifié le groupe de correspondance d'origine pour inclure les valeurs des meilleurs enregistrements, vous ne pouvez alors pas organiser ces groupes de correspondance modifiés pour la vérification manuelle dans Information Steward.

Tenez compte des approches suivantes pour la vérification manuelle :

- Engagez des administrateurs de données et des experts de domaines dans la vérification manuelle des groupes de correspondance suspects uniquement. Vous pouvez modifier les groupes de correspondance suspects si nécessaire et créer le meilleur enregistrement pour ces groupes de correspondance.
 1. Le job Data Services effectue des correspondances automatiques, identifie les groupes de correspondance et renseigne uniquement les groupes de correspondance suspects dans une base de données intermédiaire. Dans ce scénario, les « groupes de correspondance de fiabilité élevée » doivent être séparés des « groupes de correspondance suspects ». Le meilleur enregistrement peut être créé

pour les « groupes de correspondance de fiabilité élevée » et ces groupes de correspondance peuvent être consolidés dans Data Services.

Ce processus peut être effectué en routant les « groupes de correspondance de fiabilité élevée » vers une autre transformation Match configurée pour créer les meilleurs enregistrements en dehors des groupes d'enregistrements de correspondance déjà existants.

2. Les administrateurs de données et les experts de domaines vérifient les « groupes de correspondance suspects » dans Information Steward et effectuent les corrections si nécessaire (par exemple, ne pas mettre en correspondance un enregistrement ou déplacer un enregistrement). Ils peuvent créer manuellement le meilleur enregistrement pour ces groupes de correspondance.
- Engagez des administrateurs de données et des experts de domaines dans la vérification et la correction manuelles de tous les groupes de correspondance, mais conservez la création des meilleurs enregistrements et la logique de consolidation dans Data Services.
 1. Le job Data Services effectue les correspondances automatiques et renseigne tous les résultats de correspondance dans la base de données intermédiaire. Une table de résultats de correspondance dans la base de données intermédiaire contient des groupes de correspondance suspects, des groupes de correspondance « de fiabilité élevée » et des enregistrements uniques.
 2. Les administrateurs de données et les experts de domaines vérifient les groupes de correspondance suspects dans Information Steward et effectuent les corrections si nécessaire. Ils peuvent déplacer les enregistrements entre les groupes de correspondance suspects et les groupes de correspondance de « fiabilité élevée » ou inclure un enregistrement unique dans le groupe de correspondance approprié.
 3. Une fois la vérification des correspondances effectuée, tous les groupes de correspondance sont traités dans un autre job Data Services qui crée le meilleur enregistrement et consolide tous les groupes de correspondance.Ce processus peut être effectué par une transformation Match configurée pour créer les meilleurs enregistrements en dehors des groupes d'enregistrements de correspondance déjà existants.

i Remarque

Vous trouverez des exemples de création de meilleurs enregistrements à l'aide de Information Steward ou de Data Services en téléchargeant les modèles Data Quality Management Match. Vous trouverez de plus amples informations sur les modèles à l'adresse <https://wiki.scn.sap.com/wiki/display/EIM/SAP+Data+Services+Blueprints>.

27.4.2 Stratégies de correspondance

Voici quelques exemples de stratégies qui vous aideront à déterminer comment aborder la configuration de votre flux de données pour la correspondance.

SAP Strategy Management	Description
Correspondance simple	Utilisez cette stratégie si vos règles de gestion en matière de correspondance reposent sur un critère de correspondance simple pour identifier des relations dans des données consommateur, métier ou produit.
Membres de la famille	Utilisez cette stratégie si vos règles de gestion en matière de correspondance reposent sur plusieurs niveaux de relations consommateur tels que correspondances d'adresses, correspondances familiales et correspondances individuelles.

SAP Strategy Management	Description
Membres de l'organisation	Utilisez cette stratégie si vos règles de gestion en matière de correspondance reposent sur plusieurs niveaux de relations professionnelles tels que correspondances d'entreprises, correspondances de filiales et correspondances de contacts.
Correspondance de consommateur multinational	Utilisez cette stratégie de correspondance si vos données incluent plusieurs pays et que vos règles de gestion en matière de correspondance diffèrent selon les pays.
Identifier une personne de différentes manières	Utilisez cette stratégie si vos règles de gestion en matière de correspondance reposent sur plusieurs critères de correspondance pour identifier des relations et que vous souhaitez déterminer le point de chevauchement de toutes ces définitions.

Avant de sélectionner une stratégie de correspondance, répondez aux questions suivantes :

- Que contiennent vos données ? (données client, données internationales, etc.)
- Quels champs souhaitez-vous comparer ? (nom de famille, société, etc.)
- Quels sont les avantages et les inconvénients des données dans ces champs ?

→ Conseil

Vous obtiendrez de meilleurs résultats en nettoyant vos données avant la mise en correspondance. Le profilage des données peut aussi vous aider à répondre à cette question.

- Quel résultat final souhaitez-vous obtenir à la fin du job de correspondance ? (un enregistrement par famille, par société, etc.)

27.4.3 Composants de la correspondance

Les composants de base de la correspondance sont les suivants :

- Sets de correspondance
- Niveaux de correspondance
- Critères de correspondance

Sets de correspondance

Un set de correspondance est représenté par une transformation Match dans votre espace de travail. Chaque set de correspondance peut être doté de ses propres groupes de décomposition, critères de correspondance et hiérarchisation.

Un set de correspondance a trois objectifs :

- Autoriser uniquement la comparaison de certaines données dans un ensemble de critères de correspondance (par exemple, exclure les NSS vides, les adresses internationales, etc.).
- Autoriser l'empilement des scénarios de correspondance liés pour créer un set de correspondance multiniveaux.
- Autoriser la prise en compte de plusieurs sets de correspondance pour les associer dans un set de correspondance Associate.

Niveaux de correspondance

Un niveau de correspondance indique le type de correspondance qui sera effectué, par exemple sur l'individu, la famille, la résidence, la société, etc. Un niveau de correspondance ne fait pas référence à un critère spécifique, mais à la catégorie globale de correspondance.

Vous pouvez utiliser autant de niveaux de correspondance que nécessaire. Toutefois, l'assistant Matchvous limite à trois niveaux pendant la configuration (vous pouvez en ajouter par la suite). Vous pouvez définir les niveaux de correspondance d'un set de correspondance avec des limites de plus en plus strictes. La correspondance multiniveaux alimente uniquement les enregistrements qui correspondent de niveau de correspondance en niveau de correspondance (par exemple, résident, famille, individu) à des fins de comparaison.

Composant de correspondance	Description
Nom de famille	L'objectif du type de correspondance Famille est de déterminer si deux personnes doivent être considérées comme membres de la même famille, comme indiqué dans leurs données d'enregistrement. La transformation Matchcompare les données de nom de famille et d'adresse. Une correspondance signifie que les deux enregistrements représentent des membres d'une même famille. Le résultat de la correspondance est un enregistrement par famille.
Individu	L'objectif du type de correspondance Individu est de déterminer si deux enregistrements représentent la même personne, comme indiqué dans les données d'enregistrement. La transformation Matchcompare les données de prénom, de nom de famille et d'adresse. Une correspondance signifie que les deux enregistrements représentent la même personne. Le résultat de la correspondance est un enregistrement par individu.
Résident	L'objectif du type de correspondance Résident est de déterminer si deux enregistrements doivent être considérés comme membres de la même résidence, comme indiqué dans les données d'enregistrement. La transformation Matchcompare les données d'adresse. Une correspondance signifie que les deux enregistrements représentent des membres d'une même résidence. Comparez ce type de correspondance avec le type de correspondance Famille, qui compare aussi les données de nom de famille. Le résultat de la correspondance est un enregistrement par résidence.
Société	L'objectif du type de correspondance Société est de déterminer si deux enregistrements représentent la même société. Ce type de correspondance implique de comparer les données de société et d'adresse. Une correspondance signifie que les deux enregistrements représentent la même société. Le résultat de la correspondance est un enregistrement par société.
Société - Individu	L'objectif du type de correspondance Société - Individu est de déterminer si deux enregistrements représentent la même personne dans la même société, comme indiqué dans les données d'enregistrement. Ce type de correspondance implique de comparer les données de prénom, nom de famille, nom de société et adresse. Une correspondance signifie que les deux enregistrements représentent la même personne dans la même société. Le résultat de la correspondance est un enregistrement par individu et par société.

Critères de correspondance

Les critères de correspondance font référence au champ que vous voulez faire correspondre. Vous pouvez utiliser des options de critères pour spécifier des règles de gestion pour la correspondance sur chacun de ces champs. Ces règles vous permettent de définir à quel niveau d'exactitude les données doivent répondre pour être considérées comme des correspondances.

Vous pouvez, par exemple, définir un niveau de similarité d'au moins 85 % pour les prénoms, mais autoriser aussi l'initiale d'un prénom à correspondre à un prénom complet et autoriser un prénom à correspondre à un deuxième prénom.

- Les critères de correspondance au niveau de la famille peuvent inclure le nom de famille et l'adresse ou le nom de famille et le numéro de téléphone.
- Les critères de correspondance au niveau de l'individu peuvent inclure le nom complet et l'adresse, le nom complet et le NSS ou le nom complet et l'adresse électronique.
- Les critères de correspondance au niveau de la société peuvent inclure le nom de la société et l'adresse, le nom de la société et le code NACE ou le nom de la société et le numéro DUNS (Data Universal Numbering System).

27.4.4 Match Wizard

27.4.4.1 Assistant de correspondance

L'Assistant Match peut rapidement configurer les flux de données de correspondance sans que vous ayez à créer manuellement chaque transformation individuelle nécessaire pour terminer la tâche.

Opérations prises en charge par l'Assistant de correspondance

L'Assistant Match :

- Génère toutes les transformations requises pour exécuter la stratégie de correspondance choisie.
- Applique des valeurs par défaut à vos critères de correspondance selon la stratégie choisie.
- Place les transformations générées dans l'espace de travail, connectées à la plateforme choisie en amont.
- Détecte les champs appropriés en amont et les mappe automatiquement.

Opérations non prises en charge par l'Assistant de correspondance

L'Assistant Match vous fournit une configuration de correspondance de base, qui requiert parfois une personnalisation en fonction de vos règles de gestion.

L'Assistant Match :

- Ne modifie pas les données traitées. Pour corriger des entrées non standard ou des données manquantes, placez l'une des transformations de nettoyage des adresses et une transformation Data Cleanse en amont du processus de correspondance.
- Ne connecte pas les transformations Match générées à une transformation en aval, comme un chargeur. C'est à vous de connecter ces transformations.
- Ne permet pas de définir des valeurs basées sur des règles ou des valeurs de scoring pondéré à des fins de correspondance. L'Assistant Match intègre des "meilleures pratiques" qui définissent ces valeurs à votre place. Vous pouvez modifier les valeurs d'options en fonction de vos règles de gestion.

Informations associées

[Méthode de combinaison \[page 526\]](#)

27.4.4.2 Avant de commencer

Préparer un flux de données pour l'Assistant Correspondance

Pour optimiser la performance de l'Assistant Match, veillez à inclure les éléments suivants avant de le lancer :

- Incluez un lecteur dans votre flux de données. Imaginons que vous vouliez faire correspondre un champ d'entrée spécifique que nos transformations de nettoyage des données ne prennent pas en charge.
- Incluez l'une des transformations de nettoyage des adresses et la transformation Data Cleanse. L'Assistant Match est optimisé si les données que vous faites correspondre ont déjà été nettoyées et analysées dans des champs discrets en amont dans le flux de données.
- Pour faire correspondre des champs d'adresses, veillez à les faire passer par la transformation Data Cleanse. Sinon, ils ne seront pas disponibles pour la transformation Match (et l'Assistant Match). Cette règle s'applique aussi si la transformation Data Cleanse est placée avant une transformation de nettoyage des adresses.

27.4.4.3 Utilisation de l'assistant de correspondance

27.4.4.3.1 Sélection de la stratégie de correspondance

L'Assistant Match commence par vous inviter à sélectionner une stratégie de correspondance en fonction de vos exigences de règles de gestion. Le chemin à suivre dans l'Assistant Match dépend de la stratégie sélectionnée. Utilisez ces descriptions pour décider de la meilleure stratégie à adopter :

- **Correspondance simple.** Utilisez cette stratégie si vos règles de gestion en matière de correspondance reposent sur un critère de correspondance simple pour identifier des relations dans des données consommateur, métier ou produit.

- **Membres de la famille.** Utilisez cette stratégie si vos règles de gestion en matière de correspondance reposent sur plusieurs niveaux de relations consommateur tels que correspondances d'adresses, correspondances familiales et correspondances individuelles.
- **Membres de l'organisation.** Utilisez cette stratégie si vos règles de gestion en matière de correspondance reposent sur plusieurs niveaux de relations professionnelles tels que correspondances d'entreprises, correspondances de filiales et correspondances de contacts.
- **Correspondance de consommateur multinational.** Utilisez cette stratégie de correspondance si vos données incluent plusieurs pays et que vos règles de gestion en matière de correspondance diffèrent selon les pays.

i Remarque

La stratégie Correspondance de consommateur multinational configure un flux de données pour des données Latin1. Pour utiliser la correspondance Unicode, vous devez modifier votre flux de données après sa création.

- **Identifier une personne de différentes manières.** Utilisez cette stratégie si vos règles de gestion en matière de correspondance reposent sur plusieurs critères de correspondance pour identifier des relations et que vous voulez trouver le chevauchement entre toutes ces définitions.

Statistiques source

Si vous voulez générer des statistiques source pour vos rapports, assurez-vous qu'un champ contenant la valeur source physique existe dans toutes les sources de données.

Pour générer des statistiques source pour vos rapports de correspondance, cochez la case [Générer des statistiques pour vos sources](#), puis sélectionnez un champ contenant la valeur source physique.

Informations associées

[Correspondance Unicode \[page 558\]](#)

27.4.4.3.2 Identification des critères de correspondance

Les critères représentent les données que vous voulez utiliser pour déterminer les correspondances. Dans cette fenêtre, vous devez définir ces critères pour chaque set de correspondance que vous utilisez.

Les sets de correspondance comparent les données pour rechercher des enregistrements similaires, indépendamment dans chaque groupe de décomposition désigné (ultérieurement dans l'Assistant de correspondance). Les enregistrements d'un groupe de décomposition ne sont pas comparés à ceux d'un autre groupe de décomposition.

Pour rechercher les données correspondant à tous les champs, utilisez un set de correspondance unique avec plusieurs champs. Pour rechercher les données correspondant uniquement à une combinaison spécifique de champs, utilisez plusieurs sets de correspondance avec deux champs.

En travaillant sur des données d'étudiant ou de pigeon voyageur, un individu peut utiliser le même nom, mais avoir plusieurs adresses valides.

Sélectionnez une combinaison de champs qui indique au mieux les informations qui se chevauchent, telles que le nom de famille et le NSS.

Données1	Données2	Données3	Données4
R. Carson	1239 Whistle Lane	Columbus, Ohio	555-23-4333
Robert T. Carson	52 Sunbird Suites	Tampa, Florida	555-23-4333

1. Entrez le nombre de manières requises pour identifier un individu. Vous obtenez alors le nombre de sets de correspondance correspondant (transformations) dans le flux de données.
2. Le nom du set de correspondance par défaut apparaît dans le champ *Nom*. Sélectionnez un set de correspondance dans la liste *Sets de correspondance* et saisissez un nom plus pertinent, si nécessaire.
3. Pour chaque set de correspondance, sélectionnez le critère à faire correspondre.
Vous affecterez ensuite des champs de transformations en amont à ces critères.
4. Sélectionnez l'option à utiliser pour la comparaison dans la colonne Comparer à l'aide de. Les options varient selon le critère choisi. Les options de comparaison sont les suivantes :
 - *Similarité de champs*
 - *Similarité de mots*
 - *Différence numérique*
 - *Différence numérique en pourcentage*
 - *Proximité géographique*
5. Facultatif : Si vous choisissez de faire une correspondance sur Personnalisé, saisissez un nom pour le critère personnalisé dans la colonne Nom personnalisé.
6. Facultatif : Si vous choisissez d'effectuer une correspondance sur Personnalisé, indiquez à quel point les données doivent correspondre pour que le critère en deux enregistrements soit considéré comme une correspondance. Les valeurs qui en résultent déterminent le degré de similarité attendu lors du processus de comparaison pour ce critère uniquement. Après avoir sélectionné une stratégie, vous pouvez modifier les valeurs des options de règles de comparaison en fonction de vos besoins de correspondance spécifiques. Sélectionnez l'un des éléments suivants dans la liste de la colonne Exactitude personnalisée :
 - *Correspondance exacte* : Les données de ce critère doivent être parfaitement identiques. Aucune variation n'est autorisée dans les données.
 - *Correspondance élevée* : Les données de ce critère doivent comporter un niveau élevé de similarité. Un faible niveau de variation est autorisé dans les données.
 - *Correspondance moyenne* : Les données de ce critère doivent comporter un niveau moyen de similarité. Un niveau moyen de variation est autorisé dans les données.
 - *Correspondance faible* : Les données de ce critère doivent comporter un niveau faible de similarité. Un niveau important de variation est autorisé dans les données.

27.4.4.3.3 Définition des niveaux de correspondance

Les niveaux de correspondance permettent de définir les processus de correspondance à différents niveaux liés de manière logique. Les niveaux de correspondance font référence à la catégorie globale de

correspondance et non aux règles spécifiques de correspondance. Par exemple, une correspondance pour le niveau Domicile correspondra uniquement à des éléments d'adresse, une correspondance pour le niveau Famille correspondra uniquement au nom de famille et une correspondance pour le niveau Individu correspondra uniquement au prénom.

La correspondance multiniveaux peut contenir jusqu'à trois niveaux dans un set de correspondance unique défini de façon incrémentielle. La correspondance multiniveaux alimente uniquement les enregistrements qui correspondent de niveau de correspondance en niveau de correspondance (à savoir résident, famille, individu) à des fins de comparaison.

Pour définir des niveaux de correspondance :

1. Cliquez sur la correspondance de niveau supérieur et saisissez un nom pour le niveau (ou conservez le nom par défaut). Les critères par défaut sont déjà sélectionnés. Si vous ne voulez pas les utiliser, décochez la case.
La sélection des critères par défaut est un bon point de départ pour sélectionner des critères. Vous pouvez ajouter des critères pour chaque niveau afin d'obtenir des correspondances plus précises.
2. Sélectionnez des critères complémentaires pour ce niveau.
3. Si vous voulez utiliser des critères autres que ceux proposés, cliquez sur [Personnalisé](#), puis sélectionnez les critères souhaités.
4. Poursuivez jusqu'à ce que tous les niveaux requis soient renseignés.

27.4.4.3.4 Sélection de pays

Sélectionnez les pays dont les standards postaux peuvent être requis pour pouvoir comparer les données entrantes. Le panneau de gauche affiche une liste de tous les pays disponibles. Le panneau de droite affiche les pays que vous avez déjà sélectionnés.

1. Sélectionnez le nom du pays dans la liste Tous les pays.
2. Cliquez sur [Ajouter](#) pour le déplacer dans la liste Pays sélectionnés.
3. Répétez les étapes 1 et 2 pour chaque pays à inclure.

Vous pouvez aussi sélectionner plusieurs pays à la fois et les ajouter tous en cliquant sur le bouton [Ajouter](#).

Les pays sélectionnés sont gardés en mémoire pour la prochaine session de l'Assistant Match.

27.4.4.3.5 Regroupement des pays dans des axes

Créez des axes pour regrouper les pays par combinaisons logiques selon vos règles de gestion (par exemple, Asie, Europe, Amérique du Sud). Chaque axe crée jusqu'à six sets de correspondance (transformations Match).

1. Sélectionnez le nombre d'axes à créer. La liste des [axes](#) reflète le nombre d'axes sélectionné et leur affecte un numéro d'axe.
2. Pour créer chaque axe, sélectionnez un titre d'axe tel que Axe1.
3. Sélectionnez les pays à inclure dans cet axe.
4. Cliquez sur [Ajouter](#) pour déplacer les pays sélectionnés vers l'axe sélectionné.

Utilisez la liste COUNTRY UNKNOWN (__) pour les données dont le pays d'origine n'a pas été identifié.

Utilisez la liste COUNTRY OTHER (--) pour les données dont le pays d'origine a été identifié, mais n'existe pas dans la liste des pays sélectionnés.

5. Dans Moteurs de comparaison, sélectionnez l'un des moteurs suivants pour chaque axe :

i Remarque

Toutes les transformations Match générées pour l'axe utiliseront le moteur de comparaison sélectionné.

- [LATIN1](#) (par défaut)
- [CHINOIS](#)
- [JAPONAIS](#)
- [COREEN](#)
- [TAIWANAIS](#)
- [AUTRE_NON_LATIN1](#)

Le bouton [Suivant](#) est activé uniquement lorsque tous les axes ont une entrée et tous les pays sont affectés à un axe.

27.4.4.3.6 Sélection de champs de critères

Sélectionnez et désélectionnez des champs de critères pour chaque set de correspondance et niveau de correspondance créés dans votre flux de données. Ces sélections déterminent quels sont les champs comparés pour chaque enregistrement. Certains critères peuvent être sélectionnés par défaut en fonction des données en entrée.

S'il n'existe qu'un seul champ du type de contenu approprié, il n'est pas possible de le modifier pour ces critères dans l'Assistant Match.

Pour activer le bouton [Suivant](#), vous devez sélectionner au moins un champ standard de non-correspondance.

1. Pour chaque champ de critères à inclure, sélectionnez un champ disponible dans la liste déroulante, qui contient des champs de sources en amont. Les champs disponibles sont limités aux types de contenu appropriés pour ces critères. Si aucun champ du type approprié n'est disponible, tous les champs en amont s'affichent dans le menu.
2. Facultatif : désélectionnez les champs de critères que vous ne voulez pas inclure.

27.4.4.3.7 Création de clés de décomposition

Utilisez des clés de décomposition pour créer des groupes gérables de données à comparer. Le set de correspondance compare les données des enregistrements au sein de chaque groupe de décomposition uniquement et non à travers les différents groupes. Le fait d'effectuer des sélections correctes permet de réduire le temps de traitement en évitant les comparaisons de données profondément divergentes.

Les clés de décomposition sont particulièrement importantes lorsque vous traitez de gros volumes de données, car la taille des groupes de décomposition peut affecter le temps de traitement. Même si vos données ne sont pas volumineuses, les groupes de décomposition accélèrent le temps de traitement.

Créez des clés de décomposition qui regroupent des données similaires susceptibles de contenir des correspondances. N'oubliez pas que les enregistrements d'un groupe de décomposition ne seront pas comparés aux enregistrements d'un autre groupe de décomposition.

Par exemple, lorsque vous effectuez une correspondance pour rechercher des adresses en double, basez la clé de décomposition sur le code postal, la ville ou l'Etat pour créer les groupes les plus susceptibles de correspondre. Lorsque vous effectuez une correspondance pour rechercher des individus en double, basez la clé de décomposition sur le code postal et une partie du nom, qui sont les plus susceptibles de correspondre.

Création de clés de décomposition

1. Dans la colonne "Combien de champs ?", sélectionnez le nombre de champs à inclure dans la clé de décomposition.
2. Pour chaque clé de décomposition, sélectionnez les éléments suivants :
 - le ou les champs dans la clé de décomposition
 - le point de départ de chaque champ
 - le nombre d'emplacements à lire dans chaque champ
3. Après avoir défini les clés de décomposition, procédez comme suit :
 - Cliquez sur [Terminer](#). La transformation Match se termine.
 - Si vous effectuez une correspondance multinationale, cliquez sur [Suivant](#) pour accéder à la page Critères de correspondance.

27.4.4.4 Après la configuration

Bien que l'Assistant Match fasse une grande partie du travail, voici quelques étapes à suivre pour exécuter correctement votre job de correspondance. Vous pouvez aussi effectuer d'autres opérations pour affiner votre processus de correspondance.

Se connecter aux transformations en aval

Lorsque l'Assistant Match a terminé, il place les transformations générées dans l'espace de travail, connectées à la transformation en amont que vous avez sélectionnée pour lancer l'Assistant Match. Pour pouvoir exécuter votre job, vous devez connecter chaque port de la dernière transformation à une plateforme en aval. Pour ce faire, cliquez sur un port et faites-le glisser pour le connecter à l'objet souhaité.

Afficher et modifier la nouvelle transformation Match

Pour voir ce que contient la transformation générée par l'Assistant Match, cliquez avec le bouton droit sur la transformation et sélectionnez [Editeur de correspondance](#).

Afficher et modifier les transformations Associate

Pour voir ce que contient la transformation Associate générée par l'Assistant Match, cliquez avec le bouton droit sur la transformation et sélectionnez [Editeur associé](#).

Correspondance multinationale

Pour la stratégie de correspondance des consommateurs multinationaux, l'assistant crée une transformation Match pour chaque axe créé.

⚠ Attention

Si vous supprimez des axes de l'espace de travail une fois qu'ils ont été générés par l'Assistant, vous devez ouvrir la transformation Case et supprimer les règles inutiles.

Informations associées

[Correspondance Unicode \[page 558\]](#)

27.4.5 Transformations pour flux de données de correspondance

Les transformations Match et Associate sont les principales transformations utilisées dans la configuration d'une correspondance dans un flux de données. Ces transformations exécutent les fonctions de correspondance de base.

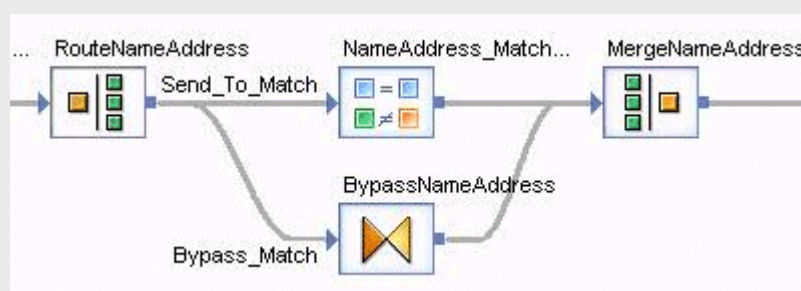
D'autres transformations peuvent être utilisées à des fins spécifiques pour optimiser la correspondance.

Transformation	Utilisation
Case	Route les données vers une transformation Match spécifique (set de correspondance). Cette transformation est couramment utilisée pour envoyer des données américaines et internationales vers différentes transformations. Vous pouvez aussi l'utiliser pour router des enregistrements vides vers une transformation Match.
Merge	Exécute les fonctions suivantes : • Regroupe les données des transformations Match pour la correspondance Association. • Regroupe les enregistrements correspondants et les enregistrements vides après fractionnement par une transformation Case.

Transformation	Utilisation
Query	Crée des champs, exécute des fonctions pour préparer les données pour la correspondance, classe les données, etc.

❖ Exemple

Lorsque vous devez ignorer des enregistrements d'un processus de correspondance spécifique (généralement dans les flux de données associatives et toutes les fois que vous voulez disposer d'enregistrements avec des données vides pour ignorer un processus de correspondance), vous devez utiliser les transformations Case, Query et Merge.



- La transformation Case utilise deux routages : l'un envoie tous les enregistrements qui répondent aux critères vers la transformation Match et l'autre envoie tous les autres enregistrements vers le routage de correspondance ignorée.
- La transformation Query ajoute les champs générés par la transformation Match et vos données de sortie (le schéma de sortie de la transformation Match et celui de la transformation Query doivent être identiques pour que les routes puissent être fusionnées). Le contenu des nouveaux champs de la transformation Query peut être renseigné avec une chaîne vide.
- La transformation Merge fusionne les deux routages en un seul.

27.4.5.1 Suppression de correspondance dans la transformation Match

Imaginons que vous vouliez disposer d'une transformation qui utilise certaines des fonctionnalités d'une transformation Match dans votre flux de données, mais qui n'inclut pas les fonctionnalités de correspondance. Par exemple, dans le cas où vous souhaitez effectuer une hiérarchisation ou une sélection de candidats dans un flux de données ou une partie d'un flux de données sans effectuer de correspondance.

1. Cliquez avec le bouton droit de la souris sur la transformation Match dans la bibliothèque d'objets, puis sélectionnez Nouveau.
2. Dans le champ *Nom du format*, saisissez un nom pertinent pour votre transformation. Il est utile d'indiquer le type de fonction que cette transformation doit exécuter.
3. Cliquez sur *OK*.
4. Glissez-déposez votre nouvelle transformation ainsi configurée dans l'espace de travail et connectez-la à votre flux de données.Match

5. Cliquez avec le bouton droit sur votre nouvelle transformation, puis sélectionnez [Editeur de correspondance](#).
6. Décochez l'option [Exécuter la mise en correspondance](#) dans l'angle supérieur gauche de l'éditeur Match.

Vous pouvez à présent ajouter une opération disponible de votre choix à cette transformation.

27.4.6 Utilisation de l'éditeur de correspondance et de l'éditeur associé

Editeurs

Les éditeurs des transformations Match et Associate permettent de configurer les schémas d'entrée et de sortie. Vous pouvez y accéder en cliquant deux fois sur l'icône de la transformation appropriée dans votre espace de travail.

Les éditeurs Match et Associate permettent de configurer les options de votre transformation. Vous pouvez y accéder en cliquant avec le bouton droit sur la transformation appropriée et en sélectionnant [Editeur de correspondance](#) (ou [Editeur associé](#)).

Ordre de la configuration

→ Conseil

L'ordre dans lequel vous configurez votre transformation Match est important !

Pour commencer, il est recommandé de mapper les champs d'entrée. Si vous ne le faites pas et que vous ajoutez une opération dans l'éditeur Match, il est possible que vous ne puissiez pas afficher un champ à utiliser pour cette opération.

Vous devez ensuite configurer vos options dans l'éditeur Match avant de mapper vos champs de sortie. Le fait d'ajouter des opérations à la transformation Match (telles que ID unique et Grouper statistiques) peut vous fournir des champs utiles générés par la transformation Match à utiliser ultérieurement dans le flux de données ou à ajouter à votre base de données.

→ N'oubliez pas

Vérifiez que vous avez effectué les opérations suivante :

1. Mapper les champs d'entrée.
2. Configurer les options de la transformation.
3. Mapper les champs de sortie.

27.4.7 Sources physiques et logiques

Le suivi des sources de données d'entrée et d'autres sources est indispensable pour générer des rapports de correspondance informatifs.

Le suivi des sources de données d'entrée et d'autres sources (qu'elles soient basées sur une source d'entrée ou sur un élément de données) dans les lignes qui sont lues, à travers l'ensemble du flux de données est indispensable pour générer des rapports de correspondance informatifs. Selon les sources que vous suivez, vous devez créer les champs appropriés dans votre flux de données pour vous assurer que le logiciel génère les statistiques requises (si elles ne figurent pas déjà dans votre base de données). Il existe deux types de sources d'entrée, comme le décrit le tableau ci-dessous.

Source d'entrée	Description
Physique	Nom de fichier ou valeur attribué(e) à la source des données d'entrée.
Logique	Groupe d'enregistrements répartis sur plusieurs sources d'entrée ou ensemble d'enregistrements issus d'une source d'entrée unique.

27.4.7.1 Source d'entrée logique

Effectuer un suivi des diverses sources de données d'entrée pour les rapports les statistiques et les tables de comparaison.

Pour compter le statistiques sources dans la transformation Match (pour le rapport Résumé des statistiques source de correspondance, par exemple), vous devez créer un champ à l'aide d'une transformation Query ou User-Defined (si vos sources de données d'entrée n'en contiennent pas déjà un).

Ce champ effectue un suivi des différentes sources dans un lecteur à des fins de reporting et est utilisé dans l'opération Grouper statistiques de la transformation Match pour générer les statistiques source. Il est également utilisé dans les tables de comparaison pour spécifier les sources à comparer.

27.4.7.2 Source d'entrée physique

Vous effectuez le suivi de votre source de données d'entrée en attribuant à la source physique une valeur dans un champ.

Vous utilisez ensuite ce champ dans les transformations dans lesquelles les statistiques du rapport sont générées.

Pour attribuer cette valeur, ajoutez une transformation Query après la source et ajoutez une colonne avec une constante contenant le nom à affecter à la source.

i Remarque

Si votre source est un fichier plat, vous pouvez utiliser l'option *Inclure la colonne Nom du fichier* pour générer automatiquement une colonne contenant le nom du fichier.

27.4.7.3 Utilisation des sources

Une source est un regroupement d'enregistrements effectué sur la base de caractéristiques de données identifiables. Une source peut correspondre à tous les enregistrements d'un fichier d'entrée ou à tous les enregistrements contenant une valeur donnée dans un champ donné.

Les sources sont abstraites et arbitraires : il n'y a pas de démarcation physique entre elles. L'appartenance source peut recouper des fichiers d'entrée ou des enregistrements de base de données et établir une distinction entre des enregistrements au sein d'un fichier ou d'une base de données selon la manière dont vous définissez la source.

Si vous voulez traiter tous vos enregistrements d'entrée comme des enregistrements éligibles standard de priorité équivalente, il n'est pas nécessaire d'inclure des sources dans votre job.

En règle générale, un utilisateur de correspondances cherche à obtenir une caractéristique ou un ensemble de caractéristiques suffisamment pertinent(e) pour sélectionner le meilleur enregistrement correspondant ou déterminer les enregistrements à inclure ou exclure d'une liste de publipostage, par exemple. Les sources vous permettent d'ajouter ces caractéristiques à un enregistrement du fait de l'appartenance de l'enregistrement dans la source en question.

Avant d'aborder en détails la configuration et l'utilisation des sources, voici certaines des nombreuses raisons pour lesquelles vous pouvez être amené à inclure des sources dans votre job :

- Donner la priorité à un ensemble d'enregistrements par rapport à d'autres. Vous pouvez, par exemple, vouloir donner la priorité aux enregistrements de votre base de données personnelle ou d'une source Suppression au détriment des enregistrements d'un fichier de mise à jour.
- Identifier un ensemble d'enregistrements correspondant à des sources Suppression comme le fichier DMA.
- Configurer un ensemble d'enregistrements qui ne doivent pas être pris en compte dans le statut multi-source. Par exemple, certains expéditeurs utilisent une source fictive d'acheteurs potentiels qui génèrent des rapports lorsqu'ils reçoivent le courrier pour leur permettre d'évaluer la réception de ce courrier. Il s'agit d'enregistrements d'un type spécial.
- Diminuer le temps de traitement en annulant la comparaison au sein d'un ensemble d'enregistrements pour lesquels vous savez qu'aucune correspondance n'existe. Dans ce cas, vous devez savoir qu'il n'existe pas d'enregistrements correspondants dans la source, mais qu'il peut exister des correspondances parmi les sources. Pour diminuer le temps de traitement, vous pouvez configurer des sources et annuler la comparaison au sein de chaque source.
- Obtenir des statistiques de rapport distinctes pour un ensemble d'enregistrements au sein d'une source ou obtenir des statistiques de rapport pour des groupes de sources.
- Protéger une source contre l'écrasement des données par une opération de meilleur enregistrement ou d'ID unique. Vous pouvez protéger les données sur la base de l'appartenance dans une source.

27.4.7.4 Types de sources

Vous pouvez identifier chaque source comme appartenant à l'un des types suivants : Normal, Suppression ou Spécial. Le logiciel peut traiter vos enregistrements différemment selon le type de source.

Source	Description
Normal	Une source Normal est un groupe d'enregistrements considérés comme éligibles.
Suppression	Une source Suppression contient des enregistrements qui excluent souvent l'utilisation d'un enregistrement. Par exemple, si vous utilisez Match pour affiner une source de courrier, une source Suppression peut vous aider à supprimer des enregistrements du courrier. Exemples : • Fichier de préférence de courrier DMA • Sources carcérales de l'American Correctional Association • Sources de personnes n'ayant pas donné leur accord ou non-répondantes • Sources de suppression de cartes de crédit ou chèques sans provision
Spécial	Une source Special est traitée comme une source Normal à une exception près. La source Special n'est pas prise en compte pour déterminer si un groupe de correspondance est de source unique ou multi-source. Une source Spécial peut contribuer à des enregistrements, mais elle n'est pas prise en compte dans la détermination d'un statut multi-source. Par exemple, certaines sociétés utilisent une source de noms de témoins. Il s'agit de noms de personnes signalant la réception de courriers publicitaires, afin que l'expéditeur puisse mesurer la diffusion du courrier. Le contenu de la source fictive n'est pas pris en compte dans le statut multi-source.

L'intérêt d'identifier le type de source est de définir cette identité pour chacun des enregistrements membres de la source. Le type de source joue un rôle important dans le contrôle de la priorité (ordre) des enregistrements dans un groupe de décomposition, dans la manière dont le logiciel traite les enregistrements correspondants (les membres des groupes de correspondance) et dans la manière dont le logiciel génère la sortie (à savoir, s'il inclut ou exclut un enregistrement dans/de sa sortie).

27.4.7.4.1 Définition manuelle de sources d'entrée

Après avoir mappé un champ d'entrée contenant les valeurs source, vous pouvez créer vos sources dans l'éditeur de correspondance.

1. Dans l'éditeur de correspondance, sélectionnez *Options de transformations* dans le volet de l'explorateur situé à gauche, cliquez sur le bouton *Ajouter*, puis sélectionnez *Sources d'entrée*. La nouvelle opération Sources d'entrée apparaît sous Options de transformation dans le volet de l'explorateur. Sélectionnez-la pour afficher les options Source d'entrée.
2. Dans la liste déroulante *Champ de valeur*, sélectionnez le champ contenant la valeur de la source d'entrée.
3. Dans la table *Définir sources*, créez un nom de source, saisissez une valeur source existant dans le champ Valeur pour cette source, puis sélectionnez un type de source.
4. Sélectionnez une valeur dans l'option *Nom de source par défaut*. Ce nom sera utilisé pour tout enregistrement dont la valeur du champ source est vide.

Pensez à cliquer sur le bouton *Appliquer* pour enregistrer vos modifications avant de passer à une autre opération dans l'éditeur de correspondance.

27.4.7.4.2 Définition automatique de sources d'entrée

Pour éviter de définir manuellement vos sources d'entrée, vous pouvez sélectionner l'option *Générer automatiquement des sources* dans l'opération Sources d'entrée.

1. Dans l'éditeur de correspondance, sélectionnez *Options de transformations* dans le volet de l'explorateur situé à gauche, cliquez sur le bouton *Ajouter*, puis sélectionnez *Sources d'entrée*.
La nouvelle opération Sources d'entrée apparaît sous Options de transformation dans le volet de l'explorateur. Sélectionnez-la pour afficher les options Source d'entrée.
2. Dans la liste déroulante *Champ de valeur*, sélectionnez le champ contenant la valeur de la source d'entrée.
3. Sélectionnez une valeur dans l'option *Nom de source par défaut*. Ce nom sera utilisé pour tout enregistrement dont la valeur du champ source est vide.
4. Sélectionnez l'option *Générer automatiquement des sources*.
5. Sélectionnez une valeur dans l'option *Type par défaut*.
Le type par défaut sera affecté à toute source dont le type n'est pas encore défini dans le champ Type.
6. Sélectionnez un champ dans la liste déroulante dans l'option *Champ type*.

La génération automatique de sources crée une source pour chaque valeur unique dans le champ Valeur. Tout enregistrement dont le champ Valeur n'est pas défini se voit attribuer le nom de source par défaut.

27.4.7.5 Groupes source

La fonctionnalité de groupes source permet de gérer les sources à un niveau supérieur. Imaginons, par exemple, que vous ayez loué plusieurs fichiers auprès de deux gestionnaires. Vous définissez cinq sources à utiliser dans le classement des enregistrements. De plus, vous voulez que les statistiques de votre job soient ventilées par gestionnaire et par fichier. Pour ce faire, vous pouvez définir des groupes de sources pour chaque gestionnaire.

Les groupes source affectent principalement les rapports. Toutefois, vous pouvez aussi utiliser des groupes source pour sélectionner des enregistrements de sources multiples en fonction du nombre de groupes source dans lesquels un nom apparaît.

N'oubliez pas que vous ne pouvez pas utiliser les groupes source comme vous utilisez les sources. Par exemple, vous ne pouvez pas donner à un groupe source la priorité sur un autre.

27.4.7.5.1 Création de groupes source

Des sources d'entrée doivent être définies dans une opération Source d'entrée pour pouvoir ajouter cette opération ou définir vos groupes source.

1. Sélectionnez une transformation Match dans votre flux de données, puis sélectionnez *Outils* > *Editeur de correspondance*.
2. Dans l'éditeur de correspondance, sélectionnez *Options de transformations* dans le volet de l'explorateur situé à gauche, cliquez sur le bouton *Ajouter*, puis sélectionnez *Groupes source*.
La nouvelle opération Groupes source apparaît sous l'opération Sources d'entrée dans le volet de l'explorateur. Sélectionnez-la pour afficher les options Groupe source.

3. Confirmez que les sources d'entrée requises se trouvent dans la colonne Sources située à droite.
4. Cliquez deux fois sur la première ligne dans la colonne Groupes source située à gauche, puis saisissez un nom pour votre premier groupe source et appuyez sur Entrée.
5. Sélectionnez une source dans la colonne Sources, puis cliquez sur le bouton [Ajouter](#).
6. Sélectionnez une valeur pour l'option [Action non définie](#).
Cette option spécifie l'action à exécuter si une source d'entrée n'apparaît pas dans un groupe source.
7. Si vous avez sélectionné Par défaut comme action non définie à l'étape précédente, vous devez sélectionner une valeur dans l'option [Groupe source par défaut](#).
Cette option est renseignée avec les groupes source que vous avez déjà définis. Si une source d'entrée n'est pas affectée à un groupe source, elle est affectée à ce groupe source par défaut.
8. Si vous le souhaitez, sélectionnez un champ dans la liste déroulante de l'option [Champ du groupe source](#) contenant la valeur de vos groupes source.

27.4.8 Match preparation

27.4.8.1 Préparation des données pour la correspondance

Utiliser la table comme type de liste de contrôle pour préparer vos données par correspondance.

Liste de contrôle pour la préparation de données pour la correspondance

✓ Terminé	Tâche	Description
	Correction et normalisation des données	La précision des correspondances dépend de la qualité des données utilisées dans la transformation Match. Pour établir une correspondance batch, il est toujours recommandé d'intégrer l'une des transformations de nettoyage d'adresse et une transformation Data Cleanse dans le flux de données avant de lancer la mise en correspondance.
	Filtrage des enregistrements vides	Vous devez filtrer les enregistrements vides avant la mise en correspondance. Vous bénéficierez ainsi de meilleures performances. Utilisez une transformation Case pour diriger les enregistrements vers un autre chemin ou une transformation Query pour filtrer ou bloquer les enregistrements.
	Suppression des mots parasites	Vous pouvez exécuter une recherche et remplacer les mots non significatifs pour le processus de correspondance. Dans le cas d'une correspondance basée sur les données société, des mots tels que Inc., Corp., et Ltd. peuvent être supprimés. Pour cela, vous pouvez utiliser la fonction de recherche et remplacement dans la transformation Query.

✓ Terminé	Tâche	Description
	Supprimer la ponctuation	Pour optimiser votre processus de correspondance, mappez les champs de sortie de transformation Data Cleanse suivants dans votre transformation Match. Ces champs génèrent des données standardisées converties en majuscules, dont la ponctuation a été supprimée, etc. Pour en savoir plus, voir la section Champs de sortie Data Cleanse. • Nom_famille_correspondance • Société_correspondance • Prénom1_correspondance • Prénom2_correspondance • Titre_maturité_correspondance • Personne_correspondance • Téléphone_correspondance • Titre_civilité_correspondance
	Groupes de décomposition	Les groupes de décomposition organisent les enregistrements dans des collections représentant des correspondances potentielles Match, ce qui réduit le nombre de comparaisons à effectuer par la transformation. Pour de meilleures performances, insérez une opération Groupe de décomposition dans la transformation Match.
	Autres noms	Pour obtenir plus facilement une correspondance, vous pouvez inclure des variantes de nom ou de données société dans le processus de correspondance. La variante de « Bill » correspond, par exemple, à « William ». Pour effectuer des comparaisons, il est possible d'utiliser les données d'origine et une ou plusieurs variantes. Selon le type de données, vous pouvez ajouter un nombre compris entre un et cinq de variantes ou autres noms à l'endroit de votre choix. Si, par exemple, les prénoms sont comparés mais ne correspondent pas, les variantes sont comparées. Si les variantes correspondent, les deux résultats ont encore une chance de concorder car les prénoms d'origine n'ont pas été considérés comme une correspondance.
	Autres noms personnalisés	Vous pouvez établir la correspondance d'après des champs de sortie Data Cleanse personnalisés et les alias associés. Mappez les champs de sortie personnalisés de Data Cleanse et les champs personnalisés s'afficheront dans l'onglet Champs de critères de l'éditeur de correspondance.

Informations associées

[Configuration des critères d'autre nom \[page 522\]](#)

27.4.8.1.1 Champs à inclure pour la correspondance

Pour bénéficier des nombreuses fonctionnalités de la transformation Match, vous devez mapper plusieurs champs d'entrée autres que ceux que vous voulez utiliser comme critères de correspondance.

❖ Exemple

Voici certains des champs que vous pouvez inclure : Les noms des champs sont sans importance tant que vous vous souvenez des données qu'ils contiennent.

Contenu du champ	Contient...
Source logique	Valeur qui spécifie de quelle source logique un enregistrement est issu. Ce champ est utilisé dans l'opération Grouper statistiques, dans les tables de comparaison et dans la transformation Associate.
Source physique	Valeur qui spécifie de quelle source physique un enregistrement est issu. (Par exemple, un objet source ou un groupe d'enregistrements sélectionnés par le candidat.) Ce champ est utilisé dans les options de la transformation Match, dans l'opération Sélection du candidat et dans la transformation Associate.
Clés de décomposition	Champ qui contient la valeur de clé de décomposition pour la création de groupes de décomposition. Le fait d'inclure un champ qui contient déjà la valeur de clé de décomposition peut améliorer la performance dans la création des groupes de décomposition, car, dans ce cas, la transformation Match n'a pas à analyser plusieurs champs pour créer la clé de décomposition.
Champs de critères	Champs qui contiennent les données à mettre en correspondance.
Indicateurs de nombre	Valeur Oui ou Non pour indiquer si une source logique doit être comptée dans une opération Grouper statistiques.
Priorité d'enregistrement	Valeur utilisée pour indiquer qu'un enregistrement a la priorité sur un autre lors du classement des enregistrements. Ce champ est utilisé dans les opérations Hiérarchisation du groupe.
Appliquer une pénalité pour champ vide	Valeur Oui ou Non pour indiquer si Match doit appliquer une pénalité pour champ vide à un enregistrement. Ce champ est utilisé dans les opérations Hiérarchisation du groupe.
Valeur ID unique de départ	Valeur ID de départ incrémentée de 1 chaque fois qu'un ID unique est affecté. Ce champ est utilisé dans l'opération ID unique.

Il ne s'agit pas d'une liste complète. Selon les fonctionnalités que vous voulez utiliser, vous pouvez inclure d'autres champs qui seront utilisés dans la transformation Match.

27.4.8.2 Contrôle des comparaisons d'enregistrements

Le contrôle du nombre de comparaisons d'enregistrements dans le processus de correspondance est important pour deux raisons :

- **Vitesse** : En contrôlant le nombre réel de comparaisons, vous pouvez réduire le temps de traitement.
- **Qualité de correspondance** : En regroupant uniquement les enregistrements susceptibles de correspondre, vous êtes assuré d'obtenir de meilleurs résultats dans votre processus de correspondance.

Le contrôle du nombre de comparaisons s'effectue principalement dans la section Formation du groupe de l'éditeur de correspondance avec les opérations suivantes :

- Groupe de décomposition : décomposez vos enregistrements en petits groupes d'enregistrements plus susceptibles de correspondre.
- Sélection du candidat : sélectionnez uniquement les candidats de correspondance dans une table de base de données. Cette opération est principalement utilisée pour les jobs en temps réel.

Vous pouvez aussi utiliser des tables de comparaison pour inclure ou exclure des enregistrements pour comparaison par une source logique.

Informations associées

[Groupes de décomposition \[page 500\]](#)

[Sélection de candidats \[page 503\]](#)

[Tables de comparaison \[page 508\]](#)

27.4.8.2.1 Groupes de décomposition

Lorsque vous créez des groupes de décomposition, vous placez les enregistrements dans des groupes susceptibles de correspondre. Par exemple, un scénario courant consiste à créer des groupes de décomposition basés sur un code postal. Cela permet de s'assurer que les enregistrements de codes postaux différents ne seront jamais comparés, car les chances de trouver un enregistrement correspondant avec un code postal différent sont très minces.

Clés de décomposition

Vous formez des groupes de décomposition en créant une clé de décomposition, à savoir un champ composé de parties d'autres champs ou un champ unique utilisé pour regrouper les enregistrements en fonction de données similaires.

Voici un exemple de clé de décomposition type créée en combinant les cinq chiffres du champ Postcode1 et les trois premiers caractères du champ Address_Primary_Name.

Champ (position de départ : longueur)

Données du champ

Clé de décomposition générée

Postcode1 (1:5)	10101	10101Mai
Address_Primary_Name (1:3)	Main	

Dans cet exemple, tous les enregistrements correspondant à la clé de décomposition générée sont placés dans le même groupe de décomposition et comparés.

Tri des enregistrements dans le groupe de décomposition

Les enregistrements sont triés dans le champ de la clé de décomposition.

Vous pouvez ajouter une opération Hiérarchisation du groupe après l'opération Groupes de décomposition pour indiquer les enregistrements pilotes.

→ N'oubliez pas

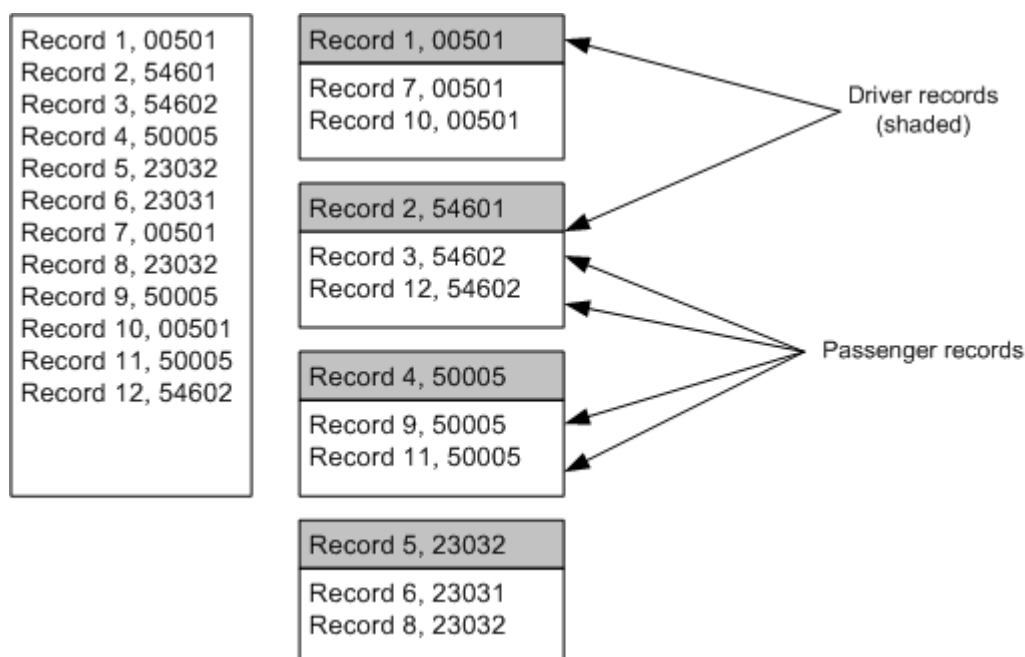
L'ordre est important ! Si vous créez des groupes de décomposition à l'aide d'enregistrements issus d'une source de type Suppression, assurez-vous que les enregistrements de suppression sont les pilotes dans le groupe de décomposition.

Structure des groupes de décomposition

Les groupes de décomposition sont constitués d'enregistrements pilotes et d'enregistrements passagers. L'enregistrement pilote est le premier enregistrement dans le groupe de décomposition et tous les autres enregistrements sont des enregistrements passagers.

L'enregistrement pilote est celui qui pilote le processus de comparaison dans la correspondance. Le pilote est d'abord comparé à tous les passagers.

Cet exemple est basé sur une clé de décomposition qui utilise les trois premiers chiffres du code postal.



Clés de décomposition phonétiques

Vous pouvez aussi utiliser les fonctions Soundex et Double Metaphone pour créer des champs contenant des codes phonétiques, qui peuvent ensuite être utilisés pour former des groupes de décomposition pour la correspondance.

Informations associées

[Critères de correspondance phonétiques \[page 560\]](#)

27.4.8.2.1.1 Création de groupes de décomposition

Il est recommandé de standardiser les données avant de créer les clés de décomposition. Les données peuvent être traitées différemment, par exemple avoir une casse incohérente.

1. Ajoutez une opération Groupes de décomposition au groupe d'options *Formation du groupe*.
2. Ajoutez une ligne dans la *table de clés de décomposition* en cliquant sur le bouton Ajouter.
3. Sélectionnez un champ dans la colonne *Champ* que vous voulez utiliser comme clé de décomposition. Le code postal est une clé de décomposition courante.
4. Sélectionnez la position de début et la longueur (nombre de caractères) à utiliser dans votre clé de décomposition.

Vous pouvez utiliser des nombres entiers négatifs pour indiquer que vous voulez débiter à la fin de la longueur réelle de la chaîne et non de la longueur spécifiée du champ. Par exemple, Champ(-3,3) prend les 3 derniers caractères de la chaîne, qu'elle ait une longueur de 10 ou de 5.

5. Ajoutez des lignes et des champs, si nécessaire.
6. Classez vos lignes en sélectionnant une ligne, puis en cliquant sur les boutons [Déplacer vers le haut](#) et [Déplacer vers le bas](#).
Le classement des lignes permet de s'assurer que les champs sont utilisés dans le bon ordre dans la clé de classification.

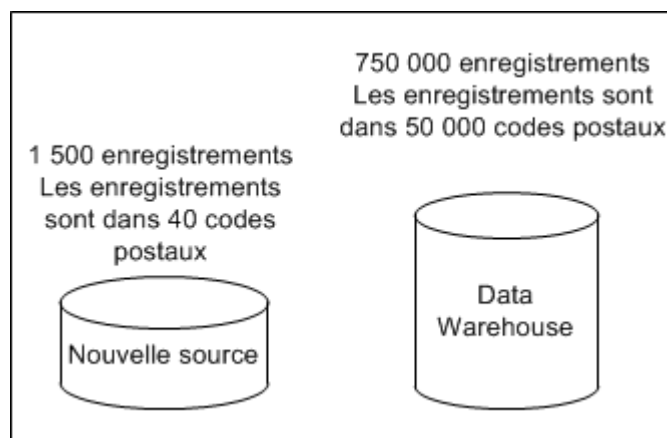
Votre clé de classification est créée.

27.4.8.2.2 Sélection de candidats

Pour accélérer le traitement dans un job de correspondance, utilisez l'opération Sélection du candidat (groupe d'options Formation du groupe) dans la transformation Match pour ajouter des rapports d'une base de données relationnelle à une collection de données existante avant la mise en correspondance. Lors de leur ajout, les enregistrements ne sont pas regroupés de façon logique. Ils sont simplement ajoutés à la fin de la collection de données, enregistrement par enregistrement, jusqu'à ce que la collection atteigne la taille définie.

Supposons que vous disposiez d'une nouvelle source d'enregistrements que vous voulez comparer à votre entrepôt de données dans un job batch. Vous pouvez sélectionner les enregistrements correspondant aux clés de décomposition de la nouvelle source à partir de cet entrepôt. Cette méthode contribue à limiter le nombre de comparaisons à effectuer par la transformation Match.

En voici une illustration simplifiée : supposons que votre job compare une nouvelle base de données source - un petit fichier régional - à une importante base de données nationale comprenant 15 enregistrements de 43 000 codes postaux. Supposons ensuite que vous vouliez former des groupes de décomposition en fonction du code postal seulement.



Notes	Régional	National	Total
Sans la sélection du candidat, la transformation Match lit tous les enregistrements des deux bases de données.	1 500	750 000	751 500

Notes	Régional	National	Total
Avec la sélection du candidat, seuls les enregistrements qui seraient inclus dans un groupe de décomposition sont lus.	1 500	Environ 600 (40 x 15)	2 100

27.4.8.2.2.1 Banques de données et sélection de candidats

Pour utiliser la sélection de candidats, vous devez vous connecter à une banque de données valide. Vous pouvez vous connecter à une banque de données basée sur SQL ou utilisant un cache persistant. Les deux présentent des avantages selon que votre source secondaire est statique (rarement mise à jour) ou dynamique (souvent mise à jour).

Banque de données utilisant un cache persistant

Similaire à toute autre banque de données permettant de charger votre ensemble de candidats. Si la source secondaire que vous utilisez pour votre sélection de candidats est statique (rarement modifiée), vous pouvez envisager de créer un cache persistant à utiliser comme table secondaire plutôt que d'utiliser votre source secondaire directement. Vous pouvez améliorer la performance.

Vous pouvez aussi l'améliorer en utilisant un fichier plat (format de recherche plus simple que SGBDR) pour votre cache persistant. Si la source secondaire n'est pas SGBDR, par exemple un fichier plat, vous ne pouvez pas l'utiliser comme "banque de données". Dans ce cas, vous pouvez créer un cache persistant à partir de cette source de fichier plat, puis l'utiliser pour la sélection de candidats.

i Remarque

Un cache persistant utilisé pour la sélection de candidats doit être créé par un flux de données sur deux octets. Pour ce faire, vous devez modifier l'option Paramètres régionaux dans le Sélecteur de paramètres régionaux Data Services (définissez la page de code sur utf-8). Exécutez le job pour générer le cache persistant. Vous pouvez ensuite redéfinir la page de code sur sa définition initiale, si vous le souhaitez.

Taille du cache

L'amélioration des performances à l'aide du cache persistant dépend aussi de la taille des données de la source secondaire. Plus la taille des données chargées dans le cache persistant augmente, plus les performances risquent de diminuer. Notez aussi que si la table de la source secondaire initiale est correctement indexée et que sa vitesse est optimisée, la création d'un cache persistant (ou même d'un cache pré-chargé) ne présente pas forcément d'intérêt.

Informations associées

[Banques de données de cache persistant \[page 112\]](#)

27.4.8.2.2.2 SQL généré automatiquement ou personnalisé

Dans certains cas, la transformation Match peut générer le SQL. Sinon, vous devez le créer. Cela dépend des options que vous avez définies et de la configuration de votre table secondaire (la table dans laquelle vous sélectionnez des candidats de correspondance).

Servez-vous de cette table pour déterminer si vous pouvez utiliser un SQL généré automatiquement ou si vous devez créer votre propre SQL.

i Remarque

Dans les scénarios suivants, les « données d'entrée » font référence à des champs de clé de décomposition issus d'une transformation en amont de la transformation Match (une transformation Query par exemple) ou à des champs de clé de décomposition issus de l'opération Groupe de décomposition au sein de la transformation Match.

Scénario	Générer automatiquement ou personnaliser ?
Vous disposez d'un seul champ de clé de décomposition dans vos données d'entrée et du même champ dans votre table secondaire.	Générer automatiquement
Vous disposez de plusieurs champs de clé de décomposition dans vos données d'entrée et des mêmes champs dans votre table secondaire.	Générer automatiquement
Vous disposez de plusieurs champs de clé de décomposition dans vos données d'entrée et d'un seul champ de clé de décomposition dans votre table secondaire.	Générer automatiquement
Vous disposez d'un seul champ de clé de décomposition dans vos données d'entrée et de plusieurs champs de clé de décomposition dans votre table secondaire.	Personnaliser
Vous disposez de plusieurs champs de clé de décomposition dans vos données d'entrée, mais d'un format ou d'un nombre de champs différents dans votre table secondaire.	Personnaliser
Vous voulez effectuer votre sélection dans plusieurs sources d'entrée.	Personnaliser

27.4.8.2.2.3 Clés de décomposition et sélection de candidats

Il est recommandé de créer une colonne de clé de décomposition dans la table secondaire (celle qui contient les enregistrements à comparer aux données d'entrée dans le flux de données) correspondant à la clé de décomposition utilisée pour créer les groupes de décomposition dans la transformation Match. Cela simplifie la configuration de l'opération Sélection de candidats. En outre, chaque colonne doit être indexée.

Il est également recommandé de créer et renseigner la base de données utilisée avec un champ de clé de décomposition unique plutôt que d'extraire des sous-chaînes de champs de base de données pour créer votre clé de décomposition. Cela peut améliorer la performance de la sélection de candidats.

i Remarque

Les enregistrements extraits par la sélection de candidats sont ajoutés à la fin d'un groupe de décomposition existant (le cas échéant). Par conséquent, si vous ne reclassez pas les enregistrements à l'aide d'une opération Hiérarchisation du groupe après l'opération Sélection de candidats, les enregistrements de la source d'origine seront toujours les enregistrements pilotes dans les groupes de décomposition. Si vous utilisez la sélection de candidats sur une source Suppression, vous devez reclasser les enregistrements pour que les enregistrements de la source Suppression soient les enregistrements pilotes.

27.4.8.2.2.4 Configuration de la sélection de candidats

Si vous utilisez la Sélection du candidat pour un job en temps réel, veuillez à décocher l'option *Fractionner les enregistrements en groupes de décomposition* dans l'opération Groupe de décomposition de la transformation Match.

Pour accélérer le temps de traitement dans un job de correspondance en temps réel, utilisez l'opération Sélection du candidat (groupe d'options Formation du groupe) dans la transformation Match pour ajouter des enregistrements d'une base de données relationnelle à une collection de données existante avant la correspondance. Lors de leur ajout, les enregistrements ne sont pas regroupés de façon logique. Ils sont simplement ajoutés à la fin de la collection de données, enregistrement par enregistrement, jusqu'à ce que la collection atteigne la taille définie.

1. Dans l'opération Sélection du candidat, sélectionnez une banque de données valide dans la liste déroulante *Banque de données*.
2. Dans la liste déroulante *Type de cache*, sélectionnez l'une des valeurs suivantes :

Option	Description
Aucun_Cache	Capture les données à un moment précis. Les données ne sont pas modifiées tant que le job n'a pas été relancé.
Cache_de_préchargement	Utilisez cette option pour les données statiques.

3. Selon la manière dont vos données d'entrée et votre table secondaire sont structurées, procédez comme suit :
 - Sélectionnez *Générer automatiquement le SQL*. Sélectionnez ensuite l'option *Utiliser une colonne de décomposition à partir de la base de données* (si vous en avez une), puis sélectionnez une colonne dans la liste déroulante *Champ de clé de décomposition*.

i Remarque

Si vous sélectionnez l'option *Générer automatiquement le SQL*, il est recommandé de disposer d'une colonne de clé de décomposition dans la table secondaire, puis de sélectionner l'option *Utiliser une colonne de décomposition à partir de la base de données*. Dans le cas contraire, le SQL créé pourrait être incorrect.

- Sélectionnez *Créer SQL personnalisé*, puis cliquez sur le bouton *Lancer l'éditeur SQL* ou saisissez votre SQL dans la zone d'édition SQL.
- 4. Pour effectuer un suivi de vos enregistrements depuis la source d'entrée, sélectionnez *Utiliser une valeur source constante*.
- 5. Saisissez une valeur qui représente votre source dans l'option *Valeur source physique*, puis sélectionnez un champ contenant cette valeur dans la liste déroulante *Champ source physique*.
- 6. Dans la table *Mappage des colonnes*, ajoutez autant de lignes que nécessaire. Chaque ligne représente un champ qui sera ajouté à la collection.
 - a. Sélectionnez un champ dans la colonne *Nom mappé*.
 - b. Sélectionnez une colonne dans votre table secondaire (ou dans une requête personnalisée) dans l'option *Nom de la colonne*, qui contienne le même type de données que celles spécifiées dans la colonne *Nom mappé*.

Si vous avez déjà défini vos clés de décomposition dans le groupe d'options Groupe de décomposition, les champs utilisés pour créer la clé de décomposition sont publiés ici, avec la colonne Groupe de décomposition définie sur OUI.

27.4.8.2.2.5 Ecriture d'un SQL personnalisé

Utilisation de caractères génériques

Pour éviter les instructions SQL compliquées, vous devez utiliser des caractères génériques (qui sont remplacés par des données d'entrée réelles) dans votre clause WHERE.

Imaginons, par exemple, que la base de données client contienne un champ appelé MatchKey et que l'enregistrement soumis au processus de nettoyage soit doté d'un champ généré appelé MATCH_KEY. Le champ possède un caractère générique [MATCHKEY]. Les enregistrements sélectionnés dans la base de données client et ajoutés à la collection de données existante sont ceux qui contiennent la même valeur dans le champ MatchKey et dans le champ MATCH_KEY de la transaction. Imaginons ici que la valeur réelle soit un numéro de téléphone à 10 chiffres.

Voici un exemple de SQL avec un numéro de téléphone réel à la place du caractère générique [MATCHKEY].

```
SELECT ContactGivenName1, ContactGivenName2, ContactFamilyName, Address1,
Address2, City, Region, Postcode, Country, AddrStreet, AddrStreetNumber,
AddrUnitNumber
FROM TblCustomer
WHERE MatchKey = '123-555-9876';
```

⚠ Attention

Vous devez vous assurer que la performance de l'instruction SQL est optimisée et générera des résultats valides. L'opération Sélection de candidats ne le fait pas à votre place.

Remplacez le caractère générique par les valeurs réelles.

Après avoir testé le SQL avec les valeurs réelles, vous devez remplacer ces valeurs par des caractères génériques ([MATCHKEY], par exemple).

Voici à quoi votre SQL devrait ressembler à présent :

```
SELECT ContactGivenName1, ContactGivenName2, ContactFamilyName, Address1,
Address2, City, Region, Postcode, Country, AddrStreet, AddrStreetNumber,
AddrUnitNumber
FROM TblCustomer
WHERE MatchKey = [MATCHKEY];
```

i Remarque

Les caractères génériques ne peuvent pas être utilisés pour des valeurs de liste, par exemple dans une clause IN :

```
WHERE status IN ([status])
```

Si [status] est une liste de valeurs, l'instruction SQL échoue.

27.4.8.2.3 Tables de comparaison

Les tables de comparaison sont des ensembles de règles qui définissent les enregistrements à comparer, un moyen supplémentaire de créer des groupes de décomposition. Vous utilisez les valeurs de votre source logique pour déterminer les enregistrements à comparer.

En utilisant des tables de comparaison, vous pouvez comparer des enregistrements au sein de sources spécifiques ou comparer des enregistrements à travers différentes sources ou une combinaison des deux.

27.4.8.2.3.1 Configuration d'une table de comparaison

Veillez à inclure un champ contenant une valeur de source logique avant d'ajouter une opération Table de comparaison à la transformation Match (dans le groupe d'options de niveau Match).

Voici un exemple illustrant la configuration d'une table de comparaison. Imaginons que vous disposiez de deux ID (A et B) et que vous vouliez uniquement effectuer des comparaisons entre les différentes sources et non pas au sein des sources.

1. Si la section Correspondance ne contient pas de table de comparaison, cliquez avec le bouton droit de la souris sur **Correspondance** > <Nom du niveau> , puis sélectionnez **Ajouter** > **Comparer** .
2. Définissez l'option **Action par défaut** sur **Aucune_Correspondance**, puis saisissez **Aucune** dans l'option **Valeur source logique par défaut**.
Ainsi, la transformation Match ne comparera pas tout, mais suivra les règles de comparaison définies par les entrées de la table.

i Remarque

Choisissez avec soin les noms des sources logiques. L'entrée « Aucune » dans l'option *Valeur source logique par défaut* ne fonctionnera pas si vous disposez d'un ID source appelé « Aucune ».

3. Dans la table *Actions de comparaison*, ajoutez une ligne, puis définissez la *Valeur pilote* sur *A* et la *Valeur passager* sur *B*.
4. Définissez *Action* sur *Comparer*.

i Remarque

Tenez compte de toutes les valeurs de sources logiques. Les exemples de valeurs saisies ci-dessus supposent que A sera toujours l'ID pilote. Si vous pensez disposer d'un enregistrement pilote avec une valeur autre que A, configurez une entrée de table pour tenir compte de cette valeur et de la valeur ID passager. N'oubliez pas que l'enregistrement pilote est le premier enregistrement lu dans une collection.

Si vous ne renseignez pas les options *Valeur pilote* ou *Valeur passager* dans la table de comparaison, cela signifie que vous voulez comparer toutes les sources. Ainsi, dans le cas d'une valeur pilote A et d'un enregistrement passager vide avec une action Comparer, un enregistrement A sera comparé à tous les autres enregistrements passager.

Il arrive que les données dans les collections soient classées différemment de ce que peut attendre votre table de comparaison ou qu'elles ne soient pas classées du tout. Dans ce cas, le processus de correspondance peut manquer des doublons.

Dans cet exemple, la manière dont vous configurez la ligne de votre table Actions de comparaison signifie que vous vous attendez à ce que l'enregistrement pilote porte la valeur pilote A ; or, s'il porte la valeur B et que l'enregistrement passager porte la valeur A, ils ne seront pas comparés.

Pour prendre en compte les cas où un enregistrement pilote porte la valeur B et l'enregistrement passager la valeur A, par exemple, ajoutez une autre ligne dans la table, qui exécute l'opération inverse. Ainsi, tout enregistrement portant la valeur A ou B sera comparé, indépendamment du pilote ou du passager.

i Remarque

En règle générale, si vous utilisez une source Suppression, vous devez effectuer une comparaison dans les autres sources. Cela permet de s'assurer que toutes les correspondances de ces sources sont supprimées lorsqu'elles dupliquent un enregistrement dans la source Suppression, indépendamment de l'enregistrement pilote.

27.4.8.3 Classement et hiérarchisation des enregistrements

Vous disposez peut-être de sources de données, telles que votre propre entrepôt de données, plus fiables que les enregistrements issus d'autres sources (source louée, par exemple). Vous préférez peut-être aussi les enregistrements récents aux enregistrements plus anciens ou les enregistrements complets à ceux contenant des champs vides. Quelles que soient vos préférences, vous pouvez les exprimer dans le processus de correspondance en utilisant des priorités.

Vous pouvez aussi souhaiter que vos enregistrements au sein d'une opération spécifique (par exemple, correspondance ou meilleur enregistrement) soient exécutés dans un ordre donné. Par exemple, vous pouvez

souhaiter que vos groupes de correspondance soient classés de telle sorte que le premier enregistrement entré soit l'enregistrement le plus récent du groupe. Dans ce cas, vous devez classer vos enregistrements en fonction d'un champ de date.

Il existe deux manières de classer les enregistrements, avant ou après le processus de comparaison :

- Tri des enregistrements en groupes de décomposition groupes de correspondance à l'aide d'une valeur dans un champ.
- Utilisation de scores de pénalité. Ces scores peuvent être définis par champ, par enregistrement ou en fonction de l'appartenance dans une source d'entrée.

Editeur de correspondance

Vous pouvez définir vos priorités et classer vos enregistrements dans l'opération Hiérarchisation du groupe, disponible dans l'opération Formation du groupe et les opérations de traitement post-comparaison de chaque niveau de correspondance dans l'éditeur de correspondance.

Types de priorités

Différents types de priorités sont à prendre en compte :

Priorité	Brève description
Priorité d'enregistrement	Privilégie les enregistrements d'une source d'entrée par rapport à une autre.
Pénalité pour champ vide	Affecte une priorité faible dans les enregistrements contenant un champ spécifique vide.

Classement pré-correspondance

Lorsque vous créez des groupes de décomposition, vous pouvez configurer votre opération ► [Formation du groupe](#) ► [Hiérarchisation du groupe](#) ► pour effectuer un classement (ou un tri) sur un champ en plus du classement sur la clé de décomposition. Cela permet de garantir que l'enregistrement ayant la priorité la plus élevée est le premier enregistrement (pilote) dans le groupe de décomposition.

Il est également important que les sources d'entrée de type Suppression soient les enregistrements pilotes dans un groupe de décomposition.

Classement post-correspondance

Une fois que la transformation Match a créé tous les groupes de correspondance et si le classement est important, vous pouvez utiliser une opération Hiérarchisation du groupe avant une opération Statistiques de groupe, Meilleur enregistrement ou ID unique pour vous assurer que l'enregistrement principal soit le premier dans le groupe de correspondance.

→ Conseil

Si vous n'utilisez pas de pénalité pour champ vide, le classement n'est peut-être pas aussi important pour vous et vous pouvez dans ce cas décider de ne pas utiliser d'opération Hiérarchisation du groupe avant vos opérations post-correspondance. Toutefois, l'opération Meilleur enregistrement peut s'avérer plus performante si vous hiérarchisez les enregistrements et définissez l'option *Publier une seule fois par destination* sur Oui.

Pénalité pour champ vide

Sur deux enregistrements, vous préférerez certainement conserver celui qui contient les données les plus complètes. Vous pouvez utiliser la pénalité pour champ vide pour pénaliser les enregistrements contenant des champs vides.

L'utilisation d'une pénalité pour champ vide est appropriée si vous estimez qu'un champ vide ne doit pas exclure un enregistrement d'une correspondance et que vous souhaitez conserver l'enregistrement le plus complet. Imaginons, par exemple, que vous vouliez accepter un enregistrement comme correspondance même si les champs Titre_de_civilité, Prénom1, Prénom2, Suffixe_principal et/ou Numéro secondaire sont vides. Bien que vous acceptiez ces enregistrements dans vos groupes de correspondance, vous pouvez leur affecter une priorité faible pour chaque champ vide.

27.4.8.3.1 Classement des enregistrements par tri sur un champ

Assurez-vous d'avoir mappé les champs d'entrée dans la transformation Match sur laquelle vous souhaitez effectuer le classement ; sinon, ils n'apparaîtront pas dans la liste déroulante des champs.

Utilisez cette méthode de classement des enregistrements si l'intégralité des données n'est pas un critère important pour vous.

1. Saisissez un nom de hiérarchisation, puis sélectionnez l'onglet Ordre de priorité.
2. Dans la table *Champs prioritaires*, sélectionnez un champ dans la liste déroulante dans la colonne *Champs d'entrée*.
3. Dans la colonne *Ordre des champs*, sélectionnez *Croissant* ou *Décroissant* pour spécifier le type d'ordre. Par exemple, si vous comparez une source Normal à une source Suppression et que vous utilisez un champ ID source pour classer vos enregistrements, vous devez vous assurer que les enregistrements de la source Suppression apparaissent en premier dans le groupe de décomposition.
4. Répétez l'étape 2 pour chaque ligne ajoutée.

5. Classez vos lignes dans la table *Champs prioritaires* à l'aide des boutons *Déplacer vers le haut* et *Déplacer vers le bas*.

La première ligne correspond au classement principal et les autres aux classements secondaires.

27.4.8.3.2 Système de scoring de pénalité

La pénalité pour champ vide est un système de scoring de pénalité. Pour chaque champ vide, vous pouvez attribuer une pénalité en indiquant un nombre entier non négatif.

Vous pouvez attribuer la même pénalité pour chaque champ vide ou une pénalité plus élevée pour les champs que vous considérez plus importants. Par exemple, si vous préparez un courrier à l'intention d'étudiants qui vivent dans des appartements ou des dortoirs, vous pouvez attribuer une pénalité plus élevée pour un champ vide Prénom1 ou un numéro d'appartement.

Champ	Pénalité pour champ vide
Titre de civilité	5
Prénom1	20
Prénom2	5
Suffixe principal	5
Numéro secondaire	20

Dans ce cas, les enregistrements ci-après seraient classés dans l'ordre indiqué (en partant du principe qu'ils proviennent de la même source et que la priorité d'enregistrement n'est donc pas un facteur). Même si le premier enregistrement contient des champs vides (Titre_de_civilité, Prénom2 et Suffixe d'adresse), il doit être généré comme enregistrement principal, car il contient les données considérées comme les plus importantes : Prénom1 et Numéro secondaire.

Titre de civilité (5)	Prénom1 (20)	Prénom2 (5)	Nom de famille	Intervalle principal	Nom principal	Suffixe principal (5)	Numéro secondaire (20)	Pénalité pour champ vide
	Maria		Ramirez	100	Main		6	$5 + 5 + 5 = 15$
Ms.	Maria	A	Ramirez	100	Main	St		20
Ms.			Ramirez	100	Main	St	6	$20 + 5 = 25$

27.4.8.3.3 Interaction de la pénalité pour champ vide avec la priorité d'enregistrement

Les scores de pénalité pour champ vide et priorité d'enregistrement sont regroupés et considérés comme un seul score.

Supposons, par exemple, que vous souhaitiez attribuer une priorité élevée à certains enregistrements de votre base de données personnelle, mais aussi une priorité faible aux enregistrements contenant des champs vides. L'appartenance source est-elle plus importante, même si certains champs sont vides ? Ou est-il plus important de disposer d'un enregistrement aussi complet que possible, même s'il ne provient pas de la base de données personnelle ?

La plupart des utilisateurs mettent la priorité sur leurs enregistrements personnels et ne souhaitent pas que les champs vides suppriment cette priorité. Pour ce faire, définissez une pénalité élevée pour l'appartenance dans une source louée et des priorités faibles pour les champs vides :

Source	Priorité d'enregistrement (points de pénalité)	Champ	Pénalité pour champ vide
Source personnelle	100	Prénom1	20
Source louée A	200	Prénom2	5
Source louée B	300	Suffixe principal	5
Source louée C	400	Numéro secondaire	20

Avec ce système de scoring, un enregistrement de source personnelle a toujours priorité sur un enregistrement de source louée, même si l'enregistrement personnel contient des champs vides. Supposons, par exemple, que les enregistrements ci-dessous figurent dans le même groupe de correspondance.

Même si l'enregistrement personnel contient cinq champs vides, il reçoit uniquement 155 points de pénalité ($100 + 5 + 20 + 5 + 5 + 20$), alors que l'enregistrement de la source A reçoit 200 points de pénalité. Par conséquent, l'enregistrement personnel a la pénalité la plus faible et la priorité la plus élevée.

Source	Prénom1	Prénom2	Famille	Inter- valle principal	Nom principal	No se- con- daire	Code_po- stal	Priorité d'enre- gistre- ment	Pénalité pour champ vide	Total
Personnel			Smith	100	Bren		55343	100	55	155
Source A	Rita	A	Smith	100	Bren	12A	55343	200	0	200
Source B	Rita		Smith	100	Bren	12	55343	300	10	310

Vous pouvez manipuler les scores pour définir la priorité telle que vous la voulez. Dans l'exemple ci-dessus, supposons que vous préféreriez un enregistrement loué contenant des données de prénom à un enregistrement personnel sans données de prénom. Vous pouvez définir le score de pénalité pour champ vide Prénom sur 500 de sorte qu'un champ vide Prénom ait plus de poids qu'une appartenance source quelle qu'elle soit.

27.4.8.3.4 Définition de priorités et pénalités à l'aide des valeurs de champ

Assurez-vous de mapper les champs d'entrée comportant des valeurs de priorité ou de pénalité pour champ vide.

Cette tâche indique à Match les champs qui comportent des valeurs de priorité d'enregistrement et de pénalité pour champ vide pour vos enregistrements et s'il faut les appliquer par enregistrement.

1. Ajoutez une opération Hiérarchisation du groupe à la section Formation du groupe ou Traitement post-correspondance dans l'éditeur de correspondance.
2. Saisissez un nom de hiérarchisation (si nécessaire) et sélectionnez l'onglet *Intégralité de l'enregistrement*.
3. Sélectionnez l'option *Classer les enregistrements selon l'intégralité des données*.
4. Sélectionnez l'option *Définir les champs de pénalité et de priorité*.
 - *Définir les pénalités de champ uniquement* : Cette option vous permet de sélectionner une priorité d'enregistrement par défaut et des pénalités pour champ vide par champ pour générer votre score de priorité.
 - *Définir les pénalités et priorités en fonction de la source d'entrée* : Cette option vous permet de définir une priorité et une pénalité pour champ vide en fonction de l'appartenance dans une source d'entrée.
5. Sélectionnez un champ contenant la valeur de priorité d'enregistrement dans l'option *Champ de priorité d'enregistrement*.
6. Dans l'option *Champ Appliquer une pénalité pour champ vide*, sélectionnez un champ contenant l'indicateur O ou N pour appliquer ou non une pénalité pour champ vide à un enregistrement.
7. Dans l'option *Priorité d'enregistrement par défaut*, entrez une priorité d'enregistrement par défaut à utiliser si un champ de priorité d'enregistrement est vide ou si vous ne spécifiez pas de champ de priorité d'enregistrement.
8. Sélectionnez une valeur *Pénalité pour champ vide appliquée par défaut* (Oui ou Non). Cette sélection détermine si la transformation Match appliquera une pénalité pour champ vide à un enregistrement si vous n'avez pas sélectionné l'option Champ Appliquer une pénalité pour champ vide ou si le champ est vide pour un enregistrement spécifique.
9. Dans la table *Score de pénalité pour champ vide*, sélectionnez un champ dans la colonne *Champ d'entrée* auquel vous voulez affecter des valeurs de pénalité pour champ vide.
10. Dans la colonne *Pénalité pour champ vide*, saisissez une valeur de pénalité pour champ vide à affecter à un enregistrement quelconque contenant un espace vide dans le champ que vous avez sélectionné dans la colonne *Champ d'entrée*.

27.4.8.3.5 Définition de valeurs de pénalité par champ

Cette tâche vous permet de définir votre score de priorité par défaut pour chaque enregistrement et pénalité pour champ vide par champ pour générer votre score de pénalité.

1. Ajoutez une opération Hiérarchisation du groupe à la section Formation du groupe ou Traitement post-correspondance dans l'éditeur de correspondance.
2. Saisissez un nom de hiérarchisation (si nécessaire) et sélectionnez l'onglet *Intégralité de l'enregistrement*.
3. Sélectionnez l'option *Classer les enregistrements selon l'intégralité des données*.

4. Sélectionnez l'option *Définir les pénalités de champ uniquement*.
5. Dans l'option *Priorité d'enregistrement par défaut*, entrez une priorité d'enregistrement par défaut qui sera utilisée dans le score de pénalité pour chaque enregistrement.
6. Sélectionnez une valeur *Pénalité pour champ vide appliquée par défaut* (Oui ou Non). Cette sélection détermine si la transformation Match appliquera une pénalité pour champ vide à un enregistrement si vous n'avez pas sélectionné l'option Champ Appliquer une pénalité pour champ vide ou si le champ est vide pour un enregistrement spécifique.
7. Dans la table *Score de pénalité pour champ vide*, sélectionnez un champ dans la colonne *Champ d'entrée* auquel vous voulez affecter des valeurs de pénalité pour champ vide.
8. Dans la colonne *Pénalité pour champ vide*, saisissez une valeur de pénalité pour champ vide à affecter à un enregistrement quelconque contenant un espace vide dans le champ que vous avez sélectionné dans la colonne *Champ d'entrée*.

27.4.8.4 Hiérarchisation des enregistrements sur la base de l'appartenance source

Indépendamment de la manière dont vous voulez hiérarchiser vos sources (en triant un groupe de décomposition ou en utilisant des scores de pénalité), vous devez vous assurer que les enregistrements de la source de type Suppression sont les enregistrements pilotes dans le groupe de décomposition et le processus de comparaison.

Imaginons, par exemple, que vous soyez une institution de bienfaisance et que vous envoyiez un courrier de sollicitation à vos donateurs actuels ainsi qu'à des noms de deux sources louées. Si un nom apparaît à la fois dans votre source personnelle et votre source louée, vous préférez utiliser celui de votre source personnelle.

Pour l'une des deux sources louées, la source B, supposons aussi que vous puissiez négocier une remise pour les enregistrements non utilisés. Vous voulez utiliser le moins d'enregistrements possibles de la source B afin d'obtenir une remise importante. Par conséquent, vous voulez que les enregistrements de la source B aient la préférence ou priorité la plus faible parmi les trois sources.

Source	Priorité
Source personnelle	Elevée
Source louée A	Moyenne
Source louée B	Faible

Sources de type Suppression et intégralité de l'enregistrement

Si vous voulez utiliser des scores de pénalité, vos sources de type Suppression doivent avoir un score de priorité faible. Ainsi, les enregistrements standard correspondant à un enregistrement de suppression seront probablement des correspondances subordonnées dans un groupe de correspondance et, par conséquent, seront également supprimés. Dans chaque groupe de correspondance, tout enregistrement ayant une priorité plus faible qu'un enregistrement de source Suppression est considéré comme une correspondance de suppression.

Imaginons, par exemple, que vous exécutiez vos fichiers par rapport au fichier de préférence de courrier DMA (liste de personnes ne souhaitant pas recevoir de publipostage). Vous identifiez la source DMA comme source Suppression et vous lui attribuez une priorité de zéro.

Source	Priorité
Source Suppression DMA	0
Source personnelle	100
Source louée A	200
Source louée B	300

Supposons que Match trouve quatre enregistrements correspondants parmi les enregistrements d'entrée.

Enregistrement correspondant (champs Nom unique-ment)				Source	Priorité
	Maria		Ramirez	Personnel	100
Ms.			Ramirez	Source B	300
Mme	Maria	A	Ramirez	Source A	200
Mme	Maria	A	Ramirez	DMA	0

Le groupe de correspondance suivant serait défini. En fonction de leur priorité, Match classerait les enregistrements comme indiqué. Par conséquent, l'enregistrement du fichier de suppression (la source DMA) serait l'enregistrement principal et les autres seraient des correspondances de suppression subordonnées et donc également supprimées.

Source	Priorité
DMA	0 (enregistrement principal)
Personnel	100
Source A	200
Source B	300

27.4.8.4.1 Définition de pénalités en fonction de l'appartenance source

Cette tâche vous permet d'attribuer des scores de priorité et des pénalités pour champ vide à une source d'entrée et d'appliquer ces scores à tout enregistrement appartenant à cette source. Avant d'effectuer cette tâche, assurez-vous que vos sources d'entrée sont définies.

1. Ajoutez une opération Hiérarchisation du groupe à la section Formation du groupe ou Traitement post-correspondance dans l'éditeur de correspondance.
2. Saisissez un nom de hiérarchisation (si nécessaire) et sélectionnez l'onglet *Intégralité de l'enregistrement*.
3. Sélectionnez l'option *Classer les enregistrements selon l'intégralité des données*.
4. Sélectionnez l'option *Définir les pénalités et priorités en fonction de la source d'entrée*.

5. Dans la table *Attributs source*, sélectionnez une source dans la liste déroulante.
6. Saisissez une valeur dans la colonne *Priorité* pour affecter une priorité d'enregistrement à cette source.
N'oubliez pas que plus le score est faible, plus la priorité est élevée. Imaginons, par exemple, que vous vouliez affecter un score très faible (tel que 0) à une source Suppression.
7. Dans la colonne *Appliquer une pénalité pour champ vide*, sélectionnez une valeur Oui ou Non pour déterminer si une pénalité pour champ vide doit être utilisée sur les enregistrements de cette source.
8. Dans l'option *Priorité d'enregistrement par défaut*, entrez une priorité d'enregistrement par défaut qui sera utilisée dans le score de pénalité pour chaque enregistrement qui n'est pas membre d'une source.
9. Sélectionnez une valeur *Pénalité pour champ vide appliquée par défaut* (Oui ou Non). Cette valeur détermine si des pénalités pour champ vide doivent être appliquées à un enregistrement qui n'est pas membre d'une source.
10. Dans la table *Score de pénalité pour champ vide*, sélectionnez un champ dans la colonne *Champ d'entrée* auquel vous voulez affecter des valeurs de pénalité pour champ vide.
11. Dans la colonne *Pénalité pour champ vide*, saisissez une valeur de pénalité pour champ vide à affecter à un enregistrement quelconque contenant un espace vide dans le champ que vous avez sélectionné dans la colonne *Champ d'entrée*.

27.4.8.5 Récupération des données

La récupération des données copie temporairement les données d'un enregistrement passager dans un enregistrement pilote après avoir comparé les deux enregistrements. Les données copiées sont des données trouvées dans l'enregistrement passager, mais manquantes ou incomplètes dans l'enregistrement pilote. La récupération des données empêche la correspondance vide ou la correspondance d'initiales de faire correspondre des enregistrements que vous ne voulez pas faire correspondre.

Prenons, par exemple, le groupe de correspondance suivant. Si vous n'avez pas activé la récupération des données, les enregistrements de la première table appartiennent tous au même groupe de correspondance, car l'enregistrement pilote, qui contient un champ Nom vide, correspond aux deux autres enregistrements.

Enregistrement	Nom	Adresse	Code_postal
1 (pilote)		123 Main St.	54601
2	John Smith	123 Main St.	54601
3	Jack Hill	123 Main St.	54601

Si vous avez activé la récupération des données, le logiciel copie temporairement John Smith du second enregistrement vers l'enregistrement pilote. Résultat : L'enregistrement 1 correspond à l'enregistrement 2, mais l'enregistrement 1 ne correspond pas à l'enregistrement 3 (car John Smith ne correspond pas à Jack Hill).

Enregistre- ment	Nom	Adresse	Code_postal
1 (pilote)	John Smith (copié depuis l'enregistrement ci-dessous)	123 Main St.	54601
2	John Smith	123 Main St.	54601

Enregistre- ment	Nom	Adresse	Code_postal
3	Jack Hill	123 Main St.	54601

L'exemple suivant illustre l'utilisation de cette fonction pour une source Suppression. Imaginons que la source Suppression soit une liste d'adresses de personnes n'ayant pas donné leur accord. Dans ce cas, vous définissez la source Suppression de sorte à ce qu'elle ait la priorité la plus élevée et vous n'activez pas la récupération des données. Le logiciel supprime alors tous les enregistrements correspondant aux enregistrements de la source Suppression.

Par exemple, dans le cas d'un enregistrement de type suppression 123 Main St correspondant à 123 Main St n°2 et 123 Main St Apt C, les deux derniers enregistrements sont supprimés.

27.4.8.5.1 Récupération des données et initiales

Lorsque le champ Nom d'un enregistrement pilote contient une initiale à la place d'un nom complet, le logiciel peut emprunter temporairement le nom complet s'il le trouve dans le champ correspondant d'un enregistrement correspondant. C'est une forme de récupération des données.

Imaginons, par exemple, que les trois enregistrements suivants représentent des enregistrements potentiellement correspondants (le logiciel les a regroupé, par exemple, en tant que membres d'un groupe de décomposition en fonction des données d'adresse et de code postal).

i Remarque

La récupération des initiales a lieu uniquement avec les champs Prénom et Nom de famille.

Enregistrement	Prénom	Nom de famille	Adresse	Notes
357	J	L	123 Main	Driver
391	Juanita	Lopez	123 Main	
839	Joanne	London	123 Main	Enregistrement du classement le plus bas

La première comparaison se fait entre l'enregistrement pilote (357) et l'enregistrement suivant le plus élevé (391). Ces deux enregistrements forment une correspondance. Juanita et Lopez sont temporairement copiés dans les champs Nom de l'enregistrement 357.

La comparaison suivante se fait entre l'enregistrement 357 et l'enregistrement du classement suivant le plus bas (839). Avec la récupération des données, les données de nom de l'enregistrement pilote sont à présent Juanita Lopez ("empruntées" de la première comparaison). Par conséquent, l'enregistrement 839 sera probablement considéré comme ne correspondant pas à l'enregistrement 357.

En conservant plus d'informations pour l'enregistrement pilote, la récupération des données permet d'améliorer la qualité des résultats de correspondance.

Initiales et enregistrements de type suppression

Cependant, si l'enregistrement pilote est un enregistrement de type suppression, il peut être préférable de désactiver la récupération des données pour avoir le plus de chances possibles d'identifier tous les enregistrements correspondant aux données de suppression initialisées. Par exemple, pour supprimer les noms comportant les initiales JL, comme dans le cas ci-dessus, vous rechercherez toutes les correspondances avec JL, indépendamment de l'ordre dans lequel les enregistrements apparaissent dans le groupe de décomposition.

Si vous avez désactivé la récupération des données pour les enregistrements de la source Suppression, vous obtenez le résultat suivant au cours de ces deux comparaisons :

Enregistrement	Prénom	Nom de famille	Adresse	Notes
357	J	L	123 Main	Driver
391	Juanita	Lopez	123 Main	
839	Joanne	London	123 Main	Enregistrement du classement le plus bas

La première comparaison se fait entre l'enregistrement pilote (357) et l'enregistrement suivant le plus élevé (391). Ces deux enregistrements forment une correspondance, car dans l'enregistrement pilote, JL et Juanita Lopez forment une correspondance.

La comparaison suivante se fait entre l'enregistrement pilote (357) et l'enregistrement du classement suivant le plus bas (839). Ici, ces deux enregistrements forment également une correspondance, car dans l'enregistrement pilote, JL correspond à Joanne London.

Comme les deux enregistrements 391 et 839 correspondent à l'enregistrement pilote de type suppression, ils sont tous deux considérés comme correspondances de suppression et, par conséquent, aucun ne figure dans votre sortie.

27.4.8.5.2 Contrôle de la récupération des données à l'aide d'un champ

Vous pouvez utiliser un champ pour contrôler l'activation de la récupération des données. Si la valeur du champ affiche O pour un enregistrement, cela signifie que la récupération des données est activée. Assurez-vous d'avoir mappé au préalable le champ dans la transformation Match que vous voulez utiliser.

1. Ouvrez l'Matchéditeur pour une Matchtransformation.
2. Dans la fenêtre Options de transformation, cliquez dans l'onglet [Récupération des données](#).
3. Sélectionnez l'option Activer la récupération des données, puis sélectionnez une valeur par défaut pour ces enregistrements.
La valeur par défaut sera utilisée si le champ que vous sélectionnez n'est pas renseigné pour un enregistrement donné.
4. Sélectionnez l'option [Spécifier la récupération des données par champ](#), puis sélectionnez un champ dans le menu déroulant.

27.4.8.5.3 Contrôle de la récupération des données par source

Vous pouvez utiliser l'appartenance dans une source d'entrée pour contrôler si la récupération des données est activée ou désactivée pour un enregistrement spécifique. Assurez-vous au préalable de créer vos sources d'entrée.

1. Ouvrez l'Matchéditeur pour une Matchtransformation.
2. Dans la fenêtre Options de transformation, cliquez dans l'onglet *Récupération des données*.
3. Sélectionnez l'option *Activer la récupération des données*, puis sélectionnez une valeur par défaut pour ces enregistrements.
La valeur par défaut sera utilisée si la source d'entrée d'un enregistrement n'est pas spécifiée dans les étapes suivantes.
4. Sélectionnez l'option *Spécifier la récupération des données par source*.
5. Dans la table, sélectionnez une valeur Source, puis Exécuter la récupération des données pour chaque source que vous voulez utiliser.

27.4.9 Match criteria

27.4.9.1 Vue d'ensemble des critères de correspondance

Utilisez des critères de correspondance à chaque niveau de correspondance pour déterminer les scores de seuil pour la correspondance et définir le traitement des différents types de données, telles que données numériques, données de nom, etc. (vos règles de gestion).

Pour ce faire, accédez au groupe d'options Critère dans l'éditeur de correspondance.

Critère de correspondance

Dans la transformation Match, les critères de correspondance représentent les champs à comparer. Imaginons, par exemple, que vous vouliez faire correspondre les 10 premiers caractères d'un prénom et les 15 premiers caractères d'un nom de famille. Vous devez dans ce cas créer deux critères pour spécifier ces exigences.

Les critères permettent à la transformation Match d'identifier le type de données figurant dans le champ d'entrée et par conséquent, les types d'opérations à exécuter sur ces données.

Critères prédéfinis et personnalisés

Il existe deux types de critères :

- Les critères prédéfinis sont disponibles pour les champs généralement utilisés pour la correspondance, tels que les données de nom, d'adresse et autres données. En affectant un critère à un champ, la

transformation Match peut identifier le type de données figurant dans le champ et exécuter des opérations internes pour optimiser les données pour la correspondance sans altérer les données d'entrée réelles.

- Des champs de sortie personnalisés Data Cleanse (définis par l'utilisateur, données non tierces) sont disponibles comme critères prédéfinis. Mappez les champs de sortie personnalisés de Data Cleanse. Les champs personnalisés apparaissent dans l'onglet Champs de critères de l'éditeur de correspondance. Tout autre type de données (tel que numéros de référence ou autres données propriétaires) pour lequel il n'existe pas de critère prédéfini doit être désigné comme critère personnalisé. Certaines fonctions peuvent être exécutées sur des clés personnalisées, telles que la correspondance d'abréviations, de sous-chaînes ou la correspondance numérique, mais la transformation Match ne peut pas effectuer toutes les comparaisons de champs croisés, par exemple certaines fonctions de correspondance de noms.

Options de pré-comparaison de critères de correspondance

La standardisation des données doit se faire en grande partie dans le nettoyage des adresses et dans les transformations Data Cleanse. Toutefois, la transformation Match peut effectuer un pré-traitement par critère (à des fins de correspondance uniquement, sans affecter les données) pour fournir des correspondances plus exactes. Les options permettant de contrôler cette standardisation figurent dans les onglets Options et Comparaisons de champs multiples de l'éditeur Match. Elles incluent les fonctions suivantes :

- Convertir les caractères diacritiques
- Convertir le texte en chiffres
- Convertir en majuscules
- Supprimer la ponctuation
- Paramètres régionaux

Pour plus d'informations sur ces options, voir la section relative à la transformation Match du *Guide de référence*.

27.4.9.1.1 Ajout et classement d'un critère de correspondance

Vous pouvez ajouter autant de critères que vous le voulez à chaque niveau de correspondance dans votre transformation Match.

1. Sélectionnez le niveau de correspondance approprié ou le groupe d'options Critères de correspondance dans l'explorateur d'options de l'éditeur de correspondance, puis cliquez avec le bouton droit de la souris.
2. Sélectionnez *Critère*.
3. Saisissez un nom pour votre critère dans la case *Nom du critère*.
Vous pouvez conserver le nom par défaut du critère prédéfini, mais vous devez saisir un nom de critère pertinent si vous avez choisi un critère personnalisé.
4. Dans l'onglet *Champs de critères*, dans la liste *Critère disponible*, sélectionnez le critère qui représente le mieux les données à faire correspondre. Si aucun critère ne convient, choisissez le critère personnalisé.
5. Dans la table *Mappage des champs de critères*, sélectionnez le nom mappé d'un champ d'entrée contenant les données que vous voulez faire correspondre pour ce critère.
6. Cliquez dans l'onglet *Options*.

7. Configurez les *Options de pré-comparaison* et les *Règles de comparaison*.
Assurez-vous de définir le score de Match et le score de non-correspondance, tous deux requis.
8. Pour activer la comparaison de champs multiples (champs croisés), cliquez dans l'onglet *Comparaisons de champs multiples* et sélectionnez l'option *Comparer plusieurs champs*.
 - a. Sélectionnez le type de comparaison de champs multiples à effectuer :
 - *Tous les champs sélectionnés dans d'autres enregistrements* : Pour comparer chaque champ à tous les champs sélectionnés dans la table dans tous les enregistrements.
 - *Le même champ dans d'autres enregistrements* : Pour comparer chaque champ uniquement au même champ dans tous les enregistrements.
 - b. Dans la table *Champs supplémentaires à comparer*, sélectionnez des champs d'entrée contenant les données à inclure dans la comparaison de champs multiples pour ce critère.

→ Conseil

Vous pouvez utiliser des noms de champs de critères de correspondance personnalisés pour la comparaison de champs multiples en les saisissant dans la colonne *Nom personnalisé*.

i Remarque

Si vous activez la comparaison de champs multiples, les champs de standards de correspondance appropriés sont supprimés de la table *Mappage des champs de critères* dans l'onglet *Champs de critères*. Si vous voulez les inclure dans le processus de correspondance, ajoutez-les à la table *Champs supplémentaires à comparer*.

9. Configurez les options de *pré-comparaison* pour la comparaison de champs multiples.
10. Pour classer vos critères dans l'explorateur d'options de l'éditeur Match (ou de la table Match), sélectionnez un critère, puis cliquez sur les boutons *Déplacer vers le haut* ou *Déplacer vers le bas*.

27.4.9.1.2 Configuration des critères d'autre nom

Assurez-vous que vous avez créé des champs d'autre nom et y avez mappé des données depuis la transformation Data Cleanse.

1. Dans la transformation Match, ajoutez un critère de correspondance Person1_Given_Name1 en le sélectionnant dans la catégorie Person.
L'éditeur ajoute automatiquement ces critères et les critères d'autre nom à la table *Mappage des champs de critères*.
2. Sélectionnez un champ dans la colonne *Nom mappé du champ d'entrée* qui représente le mieux possible les données à comparer.

Pour les données Given_Name2 (deuxième prénom), suivez la procédure décrite ci-dessus en utilisant le critère suivant :

- Person1_Given_Name2
- Person1_Given_Name2_Match_Std1
- Person1_Given_Name2_Match_Std2
- Person1_Given_Name2_Match_Std3

Vous pouvez également utiliser les autres noms pour les champs Prenom, Maturity Postname, Honorary Postname, Firm et Firm Location.

27.4.9.2 Méthodes de correspondance

Il existe plusieurs manières de configurer et de classer les critères pour obtenir les résultats de correspondance souhaités. Chacune présente des avantages et des inconvénients. Par conséquent, examinez-les soigneusement au préalable.

Méthode de correspondance	Description
Basée sur des règles	Permet de contrôler les critères qui déterminent une correspondance. Cette méthode est simple à configurer.
Scoring pondéré	Permet d'attribuer de l'importance ou une pondération à un critère. Toutefois, le scoring pondéré évalue chaque règle avant de déterminer une correspondance, ce qui peut augmenter le temps de traitement.
Méthode de combinaison	Présente les mêmes avantages et inconvénients relatifs que les deux autres méthodes.

27.4.9.2.1 Score de similarité

Le score de similarité représente le pourcentage de similarité de vos données.

Ce score est calculé en interne par l'application lorsque les enregistrements sont comparés. L'application considère les enregistrements comme des correspondances selon ce que vous avez défini pour Score de correspondance et Aucun score de correspondance dans le groupe d'options Critère (et selon d'autres facteurs, abordés ultérieurement).

❖ Exemple

Cet exemple illustre comment les scores de similarité sont déterminés. Veuillez noter ce qui suit :

- La table de comparaison ci-dessous est indiquée à titre d'exemple. Le processus de correspondance ne fonctionne pas ainsi dans la méthode de scoring pondéré, par exemple.
- Seule la première comparaison est considérée comme une correspondance, car le score de similarité est égale à ou supérieur au score de correspondance. La dernière comparaison est considérée comme une non-correspondance, car le score de similarité est inférieur au score de non-correspondance.
- Lorsqu'un critère unique ne peut pas déterminer une correspondance, comme dans le cas de la seconde comparaison dans la table ci-dessous, le processus passe au critère suivant, si possible.

Comparaison	No match	Match	Score de similarité	Correspondance ?
Smith > Smith	72	95	100%	Oui
Smith > Smitt	72	95	80%	Dépend d'autres critères
Smith > Smythe	72	95	72%	Non

Comparaison	No match	Match	Score de similarité	Correspondance ?
Smith > Jones	72	95	20%	Non

27.4.9.2.2 Méthode basée sur des règles

Avec la correspondance basée sur des règles, vous vous fiez uniquement à vos scores de correspondance et non-correspondance pour déterminer des correspondances dans un critère.

❖ Exemple

Cet exemple vous montre comment configurer cette méthode dans la transformation Match.

Critère	Enregistrement A	Enregistrement B	Aucune corres-pondance	Corres-pondance	Score de si-milarité
Prénom1	Mary	Mary	82	101	100
Nom de famille	Smith	Smitt	74	101	80
Adresse élec-tronique	msmith@sap.com	mary.smith@sap.com	79	80	91

En saisissant la valeur 101 dans le score de correspondance pour chaque critère à l'exception du dernier, les critères Prénom1 et Nom de famille ne déterminent jamais une correspondance, bien qu'ils puissent déterminer une non-correspondance.

En définissant les options *Score de correspondance* et *Aucun score de correspondance* pour le critère Adresse électronique sans intervalle, toute comparaison qui atteint le dernier critère doit être une correspondance ou une non-correspondance.

Un score de correspondance de 101 permet de s'assurer que le critère n'entraîne pas la correspondance des enregistrements, car deux champs peuvent être similaires à plus de 100 %.

→ N'oubliez pas

L'ordre est important ! Pour des raisons de performance, vous devez disposer du critère le plus susceptible de générer en premier la décision de correspondance ou non-correspondance dans l'ordre des critères. Cela permet de réduire le nombre de comparaisons de critères.

27.4.9.2.3 Méthode du scoring pondéré

Dans une méthode de mise en correspondance basée sur des règles, l'application donne la même importance (ou la même pondération) à tous les critères. En d'autres termes, si un critère n'atteint pas le score de correspondance spécifié, l'application détermine que les enregistrements ne correspondent pas.

Lorsque vous utilisez la méthode de scoring pondéré, vous vous basez sur le score de contribution total pour déterminer les correspondances au lieu d'utiliser des scores de correspondance et non-correspondance.

Valeurs de contribution

Les valeurs de contribution correspondent à votre manière d'affecter une pondération à des critères individuels. Plus la valeur est élevée, plus le critère a une pondération élevée pour déterminer les correspondances. En règle générale, les critères susceptibles d'avoir une pondération plus élevée que les autres sont les numéros de compte, les numéros de sécurité sociale, les numéros client, le code postal¹ et les adresses.

i Remarque

Toutes les valeurs de contribution pour tous les critères concernés doivent être égales à 100. Il n'est pas nécessaire d'avoir une valeur de contribution pour tous les critères.

Vous pouvez définir la valeur de contribution d'un critère dans l'option Contribution aux scores pondérés du groupe d'options Critère.

Score de contribution et score de contribution total

La transformation Matchgène le score de contribution pour chaque critère en multipliant la valeur de contribution affectée par le score de similarité (pourcentage). Ces scores de contribution individuels sont ensuite ajoutés au score de contribution total.

Score de correspondance pondéré

Dans la méthode de scoring pondéré, les correspondances sont déterminées uniquement en comparant le score de contribution total au score de correspondance pondéré. Si le score de contribution total est égal ou supérieur au score de correspondance pondéré, les enregistrements sont considérés comme une correspondance. Si le score de contribution total est inférieur au score de correspondance pondéré, les enregistrements sont considérés comme une non-correspondance.

Vous pouvez définir le score de correspondance pondéré dans l'option [Score de correspondance pondéré](#) du groupe d'options Niveau.

❖ Exemple

La table suivante illustre la configuration du scoring pondéré. Notez les différents types de scores qui ont été abordés dans ce chapitre. Notez aussi ce qui suit :

- Lors de la configuration du scoring pondéré, l'option [Aucun score de correspondance](#) doit être définie sur -1 et l'option [Score de correspondance](#) doit être définie sur 101. Ces valeurs permettent de s'assurer que ni une correspondance ni une non-correspondance ne peuvent être trouvées à l'aide de ces scores.

- Une valeur de contribution a été affectée au critère Adresse électronique lui attribuant l'importance la plus élevée.

Critère	Enregistrement A	Enregistrement B	Aucune correspondance	Correspondance	Score de similarité	Valeur de contribution	Score de contribution (similarité x valeur de contribution)
Prénom	Mary	Mary	-1	101	100	25	25
Nom	Smith	Smitt	-1	101	80	25	20
Adresse électronique	ms@sap.com	msmith@sap.com	-1	101	84	50	42
							Score de contribution total : 87

Si le score de correspondance pondéré est égal à 87, toute comparaison dont le score de contribution total est égal ou supérieur à 87 est considérée comme une correspondance. Dans cet exemple, la comparaison est une correspondance, car le score de contribution total est égal à 87.

27.4.9.2.4 Méthode de combinaison

Cette méthode associe la correspondance basée sur une règle et par scoring pondéré.

Critère	Enregistrement A	Enregistrement B	Aucune correspondance	Correspondance	Score de similarité	Valeur de contribution	Score de contribution (similarité réelle X valeur de contribution)
Prénom	Mary	Mary	59	101	100	25	25
Nom	Smith	Hope	59	101	22	N/A (aucune correspondance)	N/A
E-mail	ms@sap.com	msmith@sap.com	49	101	N/A	N/A	N/A
							Score de contribution total
							N/A

27.4.9.3 Règles de gestion relatives à la correspondance

Dans le cadre du processus de correspondance, il est essentiel de déterminer comment traiter les diverses formes de données et les différences qui existent entre elles. Si, par exemple, chaque champ d'un enregistrement correspond aux champs d'un autre enregistrement, mais qu'un champ est renseigné dans un enregistrement et non renseigné dans l'autre, voulez-vous que ces enregistrements soient considérés comme des correspondances ? Anticiper votre façon d'agir dans ces situations fait partie de la définition des règles de

gestion. Les règles de gestion se définissent pour la plupart dans les critères de correspondance et certaines options basées sur le nom se définissent dans le groupe d'options Niveau de correspondance.

27.4.9.3.1 Correspondance sur des chaînes, des abréviations et des initiales

Une partie de la définition de vos règles de gestion consiste à définir des paramètres pour la correspondance de chaînes, d'abréviations et d'initiales.

Pour en savoir plus sur la définition de scores pour les trois options suivantes, voir les sections Critères de correspondance du *Guide de référence*.

Option	Description
<i>Score d'ajustement des initiales</i>	Permet de mettre en correspondance des initiales et des mots entiers. Par exemple, « Prestataires internationaux de soins » peut être mis en correspondance avec « PIS ».
<i>Score d'ajustement des abréviations</i>	Permet de mettre en correspondance des mots entiers et des abréviations. Par exemple, « Prestataires internationaux de soins » peut être mis en correspondance avec « Prestataires internat. de soins ».
<i>Score d'ajustement des sous-chaînes</i>	Permet de mettre en correspondance des chaînes développées et des chaînes abrégées. Par exemple, la chaîne « Peinture et décapage Mayfield » peut être mise en correspondance avec « Peinture Mayfield ».

27.4.9.3.2 Correspondance d'abréviation étendue

La correspondance d'abréviation étendue offre des fonctionnalités permettant de gérer des situations non couvertes par les options *Score d'ajustement des initiales*, *Score d'ajustement des sous-chaînes* et *Score d'ajustement des abréviations*. Vous pouvez, par exemple, vous trouver dans l'un des cas suivants :

- Imaginons que vos données contiennent des localités comme La Crosse et New York. Vous disposez aussi de ces mêmes localités indiquées sous la forme LaCrosse et NewYork (sans espaces). Dans le cadre d'une correspondance standard, vous ne pourriez pas désigner ces localités (La Crosse/LaCrosse et New York/NewYork) comme des correspondances à 100 % en raison des espaces. (Il s'agirait normalement d'une correspondance à 94 et 93 %.)
- Imaginons que vos données contiennent Metropolitan Life et MetLife (une abréviation et une combinaison de Metropolitan Life). L'option *Score d'ajustement des abréviations* ne peut pas détecter la combinaison des deux mots.

Si vous êtes dans l'un de ces cas, vous devez utiliser l'option *Score d'ajustement des abréviations ext.*

Fonctionnement du score d'ajustement

Le score que vous définissez dans l'option *Score d'ajustement des abréviations ext* ajuste votre score de similarité pour prendre en compte ces types d'abréviations et de combinaisons dans vos données.

Le score d'ajustement ajoute une pénalité pour la partie non correspondante des mots. Plus le nombre est élevé, plus la pénalité est faible. Un score de 100 signifie aucune pénalité et un score de 0 une pénalité maximum.

♣ Exemple

Chaîne 1	Chaîne 2	Score de sim. lorsque le score d'ajust. est de 0	Score de sim. lorsque le score d'ajust. est de 50	Score de sim. lorsque le score d'ajust. est de 100	Notes
MetLife	Metropolitan Life	58	79	100	
MetLife	Met Life	93	96	100	
MetLife	MetropolitanLife	60	60	60	Ce score est dû à la comparaison des chaînes. Le scoring d'abréviation étendue n'a pas été nécessaire ou utilisé, car les deux chaînes comparées comptent chacune un mot.

27.4.9.3.3 Correspondance de noms

Lorsque vous créez vos règles de gestion, vous devez notamment définir comment les noms doivent être traités dans le processus de correspondance. La transformation Matchvous fournit plusieurs moyens de vous assurer que les variations de noms ou noms multiples, par exemple, sont prises en compte.

i Remarque

Contrairement à d'autres règles de gestion, ces options sont configurées dans le groupe d'options Niveau de correspondance, car elles affectent tous les critères de correspondance appropriés basés sur le nom.

Deux noms, deux personnes

L'option *Nombre de noms devant correspondre* vous permet de contrôler comment la correspondance est effectuée sur les clés de correspondance comportant plusieurs noms (par exemple, comparaison de "John et Mary Smith" et de "Dave et Mary Smith"). Déterminez si un seul nom doit correspondre pour que les enregistrements soient identifiés comme correspondance ou si la transformation Matchdoit ignorer les personnes autres que le premier nom analysé.

Avec cette méthode, vous pouvez demander à ce qu'une personne ou les deux correspondent pour que l'enregistrement corresponde.

Deux noms, une personne

L'option *Comparer Prénom1 avec Prénom2* vous permet aussi de comparer les données Prénom1 (prénom) d'un enregistrement avec les données Prénom2 (deuxième prénom) du deuxième enregistrement. Avec cette option, la transformation Match peut correctement identifier les enregistrements correspondants, comme les deux partiellement indiqués ci-dessous : En règle générale, ces paires d'enregistrement représentent des enfants portant le nom de leurs parents, mais connus sous leur deuxième prénom.

Enregistrement n°	Prénom	Deuxième prénom	Nom de famille	Adresse
170	Leo	Thomas	Smith	225 Pushbutton Dr
198	Tom		Smith	225 Pushbutton Dr

Noms de famille avec trait d'union

L'option *Correspondance sur le nom de famille avec trait d'union* vous permet de contrôler comment la correspondance est effectuée si un champ Family_Name (nom de famille) contient un nom de famille avec trait d'union (par exemple, comparaison de "Smith-Jones" et "Jones"). Déterminez si pour les deux critères, les deux noms doivent correspondre ou juste un nom pour que les enregistrements soient désignés comme correspondance.

27.4.9.3.3.1 Correspondance des noms de famille composés

Le score des sous-chaînes approximatives aide à paramétrer la comparaison des noms de famille composés. Le score des sous-chaînes approximatives est affecté aux mots qui n'ont pas de correspondance dans une chaîne comparée. Cette option assouplit certaines des exigences de l'option Score d'ajustement des sous-chaînes comme suit :

- Il n'est pas nécessaire que les premiers mots correspondent exactement.
- Les mots qui ne correspondent pas peuvent utiliser des ajustements d'initiales et d'abréviations (par exemple, Rodriguez et RDZ).
- Les mots correspondants doivent être dans le même ordre, mais des mots non correspondants peuvent être placés avant ou après les mots correspondants.
- Le score des sous-chaînes approximatives est affecté aux mots et aux espaces sans correspondance dans la chaîne comparée.

L'option Score d'ajustement des sous-chaînes approximatives améliore le score pour certaines correspondances trouvées à l'aide du score d'ajustement des sous-chaînes.

❖ Exemple

Dans une comparaison de CRUZ RODRIGUEZ et GARCIA CRUZ DE RDZ, les scores de similarité sont les suivants :

- Sans définir aucun ajustement, le score de similarité est de 48.

- Si vous définissez le score d'ajustement des sous-chaînes sur 80 et le score d'abréviations sur 80, le score de similarité est de 66.
- Si vous définissez le score d'ajustement des sous-chaînes approximatives sur 80 et le score d'abréviations sur 80, le score de similarité est de 91.

27.4.9.3.4 Correspondance des données numériques

Utilisez l'option *Les mots numériques correspondent parfaitement* pour déterminer si les données comportant une combinaison de chiffres et de lettres doivent correspondre parfaitement. Vous pouvez également spécifier *comment* ces données doivent correspondre. Cette option s'applique le plus souvent aux données d'adresse et aux données personnalisées, comme les numéros de référence.

Le processus de correspondance numérique est le suivant :

1. La chaîne est d'abord décomposée en mots. La décomposition de mots est effectuée sur toutes les ponctuations et les espaces, puis les mots se voient affecter un attribut numérique. Un mot numérique est un mot qui contient au moins un chiffre compris entre 0 et 9. Par exemple, 4L est considéré comme un mot numérique, contrairement à QuatreL.
2. La correspondance numérique est effectuée en fonction des options qui ont été définies (comme indiqué ci-après).

Valeurs d'options et fonctionnement

Valeur d'option	Description
Any_Position	Avec cette valeur, les mots numériques doivent correspondre parfaitement. Toutefois, l'emplacement du mot <i>n'est pas</i> important. Exemple : • Comparaison de numéros et noms de rue : "4932 Main St # 101" et "# 101 4932 Main St" sont considérés comme une correspondance. • Comparaison de numéros et noms de rue : "4932 Main St # 101" et "# 102 4932 Main St" <i>ne sont pas</i> considérés comme une correspondance. • Description de pièces : "ACCU 1.4L 29BAR" et "ACCU 29BAR 1.4L" sont considérés comme une correspondance.
Same_Position	Cette valeur indique que les mots numériques doivent correspondre parfaitement. Toutefois, elle diffère de l'option Any_Position, car l'emplacement du mot est important. Par exemple, 608-782-5000 correspond à 608-782-5000, mais ne correspond pas à 782-608-5000.

Valeur d'option	Description
Any_Position_Consider_Punctuation	Cette valeur effectue la décomposition de mots sur toutes les ponctuations et les espaces, <i>sauf</i> sur le séparateur de décimales (point ou virgule) de sorte à ce que les nombres décimaux ne soient pas fragmentés. Par exemple, la chaîne 123.456 est considérée comme un mot numérique unique, contrairement à deux mots numériques. L'emplacement du mot numérique n'est pas important. Toutefois, les séparateurs de décimales ont un impact sur le processus de correspondance. Exemple : - Description de pièces : "ACCU 29BAR 1.4L" et "ACCU 1.4L 29BAR" sont considérés comme une correspondance. - Description de pièces : "ACCU 1.4L 29BAR" et "ACCU 29BAR 1.4L" <i>ne sont pas</i> considérés comme une correspondance en raison du séparateur de décimales entre le 1 et le 4 dans les deux cas. - Données financières : "25,435" et "25.435" ne sont pas considérés comme une correspondance.
Any_Position_Ignore_Punctuation	Cette valeur est similaire à la valeur Any_Position_Consider_Punctuation, à la différence près que les séparateurs de décimales <i>n'ont pas</i> d'impact sur le processus de correspondance. Exemple : - Description de pièces : "ACCU 29BAR 1.4L" et "ACCU 1.4L 29BAR" sont considérés comme une correspondance. - Description de pièces : "ACCU 1.4L 29BAR" et "ACCU 29BAR 1.4L" sont également considérés comme une correspondance malgré la présence d'un séparateur de décimales entre le 1 et le 4. - Description de pièces : "ACCU 29BAR 1.4L" et "ACCU 1.5L 29BAR" ne sont pas considérés comme une correspondance.

27.4.9.3.5 Correspondance de champs vides

Vous pouvez spécifier dans vos règles de gestion comment la transformation Match traite les comparaisons de champs lorsqu'un ou deux des champs comparés sont vides.

Par exemple, le champ Prénom est vide dans le second enregistrement ci-dessous. Voulez-vous que la transformation Match considère ces enregistrements comme des correspondances ? Que faire si le champ Prénom est vide dans les deux enregistrements ?

Enregistrement n°1	Enregistrement n°2
John Doe	____ Doe
204 Main St	204 Main St
La Crosse WI	La Crosse WI
54601	54601

La transformation Match contient des options qui permettent de contrôler la manière dont les champs sont comparés. Il s'agit des options suivantes :

- Opération pour deux champs vides
- Score pour deux champs vides
- Opération pour un champ vide
- Score pour un champ vide

Opérations pour champs vides

Les options "Opération" sont dotées des valeurs possibles suivantes :

Option	Description
<i>Eval</i>	Si vous sélectionnez Eval, la transformation Matchévalue la comparaison à l'aide du score saisi dans l'option <i>Score pour un champ vide</i> ou <i>Score pour deux champs vides</i> .
<i>Ignorer</i>	Si vous sélectionnez Ignorer, le score de cette règle de champ ne contribue pas au score pondéré global pour la comparaison des enregistrements. En d'autres termes, les deux enregistrements ci-dessus pourraient être considérés comme des doublons malgré le champ vide.

Scores pour champs vides

Les options "Score" contrôlent comme la transformation Matchévalue les comparaisons de champs lorsque le champ est vide dans un enregistrement ou dans les deux. Vous pouvez saisir des valeurs comprises entre 0 et 100.

Pour savoir quel score saisir, déterminez si vous voulez que la transformation Match considère un champ vide comme étant similaire à 0 % à un champ renseigné ou un autre champ vide comme étant similaire à 100 % ou une valeur intermédiaire.

Votre réponse dépend certainement du champ que vous comparez. Il peut être approprié d'attribuer un score élevé à un champ vide si vous effectuez la correspondance sur un prénom, un deuxième prénom ou un nom de société, par exemple.

❖ Exemple

Voici quelques exemples qui vous permettront de mieux comprendre comment les définitions de ces options de correspondances vides peuvent affecter le scoring global des enregistrements.

Opération pour un champ vide pour le champ Prénom1 défini sur Ignorer

Notez que lorsque vous définissez les options de champs vides sur Ignorer, la transformation Match redistribue la contribution attribuée à ce champ aux autres critères et recalcule les contributions des autres champs.

Champs comparés	Enregistrement A	Enregistrement B	% de similitude	Contribution	Score (par champ)
Code postal	54601	54601	100	20 (ou 22)	22

Champs comparés	Enregistrement A	Enregistrement B	% de similitude	Contribution	Score (par champ)
Adresse	100 Water St	100 Water St	100	40 (ou 44)	44
Nom_de_famille	Hamilton	Hammilton	94	30 (ou 33)	31
Prénom1	Mary		—	10 (ou 0)	—
					Score pondéré : 97

Opération pour un champ vide pour le champ Prénom1 défini sur Eval ; score pour un champ vide défini sur 0

Champs comparés	Enregistrement A	Enregistrement B	% de similitude	Contribution	Score (par champ)
Code postal	54601	54601	100	20	20
Adresse	100 Water St	100 Water St	100	40	40
Nom_de_famille	Hamilton	Hammilton	94	30	28
Prénom1	Mary		0	10	0
					Score pondéré : 88

Opération pour un champ vide pour le champ Prénom1 défini sur Eval ; score pour un champ vide défini sur 100

Champs comparés	Enregistrement A	Enregistrement B	% de similitude	Contribution	Score (par champ)
Code postal	54601	54601	100	20	20
Adresse	100 Water St	100 Water St	100	40	40
Nom_de_famille	Hamilton	Hammilton	94	30	28
Prénom1	Mary		100	10	10
					Score pondéré : 98

27.4.9.3.6 Comparaison de champs multiples (champs croisés)

Dans la plupart des cas, vous utilisez un champ unique pour la comparaison. Par exemple, le Champ1 du premier enregistrement est comparé au Champ1 du deuxième enregistrement.

Toutefois, il peut être utile dans certains cas de comparer des champs multiples. Supposons, par exemple, que vous souhaitiez faire correspondre des numéros de téléphone du champ Téléphone à des numéros issus de champs utilisés pour Télécopieur, Portable et Domicile. Vous pouvez le faire grâce à la comparaison de champs multiples.

Lorsque vous activez la comparaison de champs multiples dans l'onglet Comparaison de champs multiples d'un critère de correspondance dans l'éditeur de correspondance, vous pouvez faire correspondre des champs sélectionnés à tous les champs sélectionnés dans chaque enregistrement ou uniquement au même champ dans chaque enregistrement.

i Remarque

Par défaut, Match effectue une comparaison de champs multiples sur les champs où des standards de correspondance sont utilisés. Par exemple, `Personne1_Prénom1` est automatiquement comparé à `Personne1_Prénom_Std_de_correspondance1-6`. La comparaison de champs multiples n'a pas à être explicitement activée et aucune configuration supplémentaire n'est requise pour effectuer ce type de comparaison sur les champs contenant des standards de correspondance.

27.4.9.3.6.1 Comparaison de champs sélectionnés avec tous les champs sélectionnés dans d'autres enregistrements

Lorsque vous comparez chaque champ sélectionné à tous les champs sélectionnés dans d'autres enregistrements, tous les champs définis dans ce critère de correspondance sont comparés.

→ N'oubliez pas

Les champs « sélectionnés » incluent le champ de critères et tous les autres champs définis dans la table [Champs supplémentaires à comparer](#).

- Si une ou plusieurs comparaisons de champs répondent aux définitions du score Match, les deux lignes comparées sont considérées comme des correspondances.
- Si une ou plusieurs comparaisons de champs dépassent le score Aucune correspondance, la règle est considérée comme approuvée et tous les autres critères définis/scoring pondéré sont évalués pour déterminer si les deux lignes sont considérées comme des correspondances.

♣ Exemple

Exemple de comparaison de champs sélectionnés avec tous les champs sélectionnés dans d'autres enregistrements

Vos données d'entrée contiennent deux champs Société.

ID de ligne	Société1	Société2
1	Firstlogic	Postalsoft
2	SAP BusinessObjects	Firstlogic

Avec le score Match défini sur 100 et le score Aucune correspondance défini sur 99, ces deux enregistrements sont considérés comme des correspondances. Voici un résumé du processus de comparaison et des résultats.

- Tout d'abord, la ligne 1 Firm1 (Firstlogic) est comparée à la ligne 2 Firm1 (SAP). Normalement, la comparaison de ces lignes devrait échouer, mais avec la comparaison de champs multiples activée, la décision Aucune correspondance n'est pas prise à ce stade.

- Ensuite, la ligne 1 Société2 est comparée à la ligne 2 Société2 et ainsi de suite jusqu'à ce que toutes les autres comparaisons soient effectuées entre tous les champs sur toutes les lignes. Comme la ligne 1 Société1 (Firstlogic) et la ligne 2 Société2 (Firstlogic) sont identiques à 100 %, les deux enregistrements sont considérés comme des correspondances.

27.4.9.3.6.2 Comparaison de champs sélectionnés avec des champs identiques dans d'autres enregistrements

Lorsque vous comparez chaque champ sélectionné au même champ dans d'autres enregistrements, chaque champ défini dans l'onglet Comparaison de champs multiples d'un critère de correspondance est comparé uniquement au même champ dans d'autres enregistrements. Cela permet de configurer, au sein du critère, une condition OR pour répondre au critère. Chaque champ est utilisé pour déterminer une correspondance : Si Champ_1, Champ_2 ou Champ_3 répond aux critères de correspondance, considérez les enregistrements comme une correspondance. Le score Aucune correspondance d'un champ peut éventuellement répondre aux critères lorsque vous utilisez la comparaison de champs multiples.

→ N'oubliez pas

Les champs « sélectionnés » incluent le champ de critères et tous les autres champs définis dans la table [Champs supplémentaires à comparer](#).

❖ Exemple

Exemple de comparaison de champs sélectionnés avec le même champ dans d'autres enregistrements

Vos données d'entrée contiennent un champ Téléphone, Télécopieur et Téléphone portable. Si des données de ces champs d'entrée sont identiques entre les lignes, les enregistrements sont considérés comme des correspondances.

ID de ligne	Téléphone	Télécopieur	Portable
1	608-555-1234	608-555-0000	608-555-4321
2	608-555-4321	608-555-0000	608-555-1111

Avec un score Match de 100 et un score Aucune correspondance de 99, les numéros de téléphone et de portable ne répondraient pas aux critères de correspondance s'ils étaient définis individuellement. Toutefois, comme les trois champs sont définis dans un seul critère et que les enregistrements sélectionnés sont comparés à des enregistrements identiques, ces enregistrements sont considérés comme une correspondance du fait que le numéro de télécopie est identique à 100 %.

i Remarque

Dans l'exemple ci-dessus, le téléphone portable de la ligne 1 et le téléphone de la ligne 2 ne seraient pas considérés comme correspondance avec l'option [le même champ dans d'autres enregistrements](#) sélectionnée, car dans ce cas, la comparaison se fait uniquement dans le même champ. Si cette comparaison croisée est requise, sélectionnez à la place l'option [tous les champs sélectionnés dans d'autres enregistrements](#).

27.4.9.3.7 Correspondance de proximité

La correspondance de proximité vous permet de faire correspondre des enregistrements en fonction de leur proximité au lieu de comparer la représentation des données sous forme de chaîne. Vous pouvez effectuer une correspondance sur la proximité géographique, numérique et la proximité de date.

Informations associées

[Correspondance sur la proximité géographique \[page 536\]](#)

[Correspondance sur la proximité numérique ou la proximité de date \[page 537\]](#)

27.4.9.3.7.1 Correspondance sur la proximité géographique

L'option Proximité géographique recherche des enregistrements en double basés sur la proximité géographique à l'aide d'informations de latitude et de longitude. Il ne s'agit pas de distance parcourue par la route, mais de distance géographique. Cette option utilise les coordonnées WGS 84 (GPS).

L'option Proximité géographique peut exécuter les fonctions suivantes :

- Rechercher des objets dans une plage radiale. Cette fonction peut être utile à une société qui souhaite envoyer un courrier à ses clients situés à une certaine distance de son site.
- Rechercher l'emplacement le plus proche. Cette fonction peut aider un consommateur à trouver le point de vente le plus proche de son domicile.

27.4.9.3.7.2 Configuration de la correspondance de proximité géographique - Champs de critères

Pour sélectionner des champs pour la correspondance de proximité géographique, procédez comme suit :

1. Accédez à l'éditeur de correspondance et ajoutez un nouveau critère.
2. Dans Critère disponible, développez *Géographique*.
3. Sélectionnez *LATITUDE_LONGITUDE*.
Les deux champs de critères peuvent alors être mappés.
4. Mappez les champs Latitude et Longitude appropriés. Vous devez mapper les deux champs.

27.4.9.3.7.3 Configuration de la correspondance de proximité géographique - Options critères

Pour pouvoir utiliser cette option, vous devez mapper les champs Latitude et Longitude.

Pour effectuer une correspondance de proximité géographique, procédez comme suit :

1. Dans Comparaison des données à l'aide de, sélectionnez *Proximité géographique*. Cela permet de filtrer les options sous Règles de comparaison pour afficher uniquement les options pertinentes.
2. Définissez l'option Unité de distance sur l'une des valeurs suivantes :
 - *Miles*
 - *Feet*
 - *Kilomètres*
 - *Mètres*
3. Entrez la *Distance max.* à prendre en compte pour la plage.
4. Définissez le *score de distance max.*

i Remarque

Une distance égale à la distance max. se voit attribuer un score de distance max. Toute distance inférieure à la distance max. se voit attribuer un score proportionnel compris entre le score de distance max. et 100. Par exemple, une proximité de 10 miles se voit attribuer un score supérieur à une proximité de 40 miles.

i Remarque

Si les données de la distance max. sont susceptibles de varier d'une ligne à l'autre, vous devez entrer dynamiquement les données à l'aide du champ `Option_Field_Algorithm_Geo_Proximity_<logical_name>_Max_Distance`.

27.4.9.3.7.4 Correspondance sur la proximité numérique ou la proximité de date

Les options de proximité numérique de la transformation Match recherchent les doublons en fonction de la proximité numérique des données. Vous pouvez rechercher des doublons en fonction de valeurs numériques et de valeurs de date. Les options suivantes sont disponibles dans l'onglet Options de l'éditeur de critères de correspondance pour la correspondance numérique et de date :

Différence numérique

Recherche les doublons en fonction de la différence numérique pour les valeurs numériques ou les valeurs de date. Par exemple, vous pouvez utiliser cette option pour rechercher des doublons en fonction de valeurs de date dans une plage spécifique (plus ou moins 35 jours, par exemple), indépendamment de la similarité des caractères.

Différence numérique en pourcentage

Recherche les doublons en fonction du pourcentage de différence numérique pour les valeurs numériques. Voici deux exemples dans lesquels cette option peut être utile :

- Domaine des données financières : vous pouvez effectuer une recherche dans les données financières pour trouver tous les paiements hypothécaires mensuels correspondant à 5 % d'une valeur donnée.
- Domaine des données produit : vous pouvez effectuer une recherche dans les données produit pour trouver toutes les tiges d'acier dont la tolérance correspond à 10 % d'une valeur donnée.

27.4.10 Post-match processing

27.4.10.1 Meilleur enregistrement

L'une des opérations clés des tâches de consolidation de données consiste à récupérer les données des enregistrements concordants (autrement dit, les membres des groupes de correspondance) et à transférer ces données dans un "meilleur enregistrement" ou dans tous les enregistrements concordants.

Pour exécuter ces fonctions, vous pouvez ajouter une opération de post-correspondance Meilleur enregistrement.

Opérations dans les groupes de correspondance

Les fonctions que vous exécutez à l'aide de l'opération Meilleur enregistrement impliquent la manipulation ou le transfert des données figurant dans les enregistrements principaux et subordonnés des groupes de correspondance. Les groupes de correspondance regroupent les enregistrements que la transformation Match a identifiés comme concordants en fonction du ou des critères que vous avez créés.

L'enregistrement principal est le premier enregistrement du groupe de correspondance. Vous pouvez déterminer l'enregistrement dont il s'agit en utilisant une opération Hiérarchisation du groupe avant l'opération Meilleur enregistrement.

Les enregistrements subordonnés sont tous les enregistrements restants d'un groupe de correspondance.

Le groupe de correspondance suivant illustre cette utilisation des enregistrements principaux et subordonnés :

Enregistrement	Nom	Téléphone	Date	Classement dans le groupe
1	John Smith		11 avr 2001	Principal
2	John Smyth	788-8700	12 oct 1999	Subordonné
3	John E. Smith	788-1234	22 fév 1997	Subordonné

Enregistrement	Nom	Téléphone	Date	Classement dans le groupe
4	J. Smith	788-3271		Subordonné

Dans la mesure où il s'agit d'un groupe de correspondance, tous les enregistrements sont considérés comme concordants. Comme vous pouvez le constater, chaque enregistrement présente de légères différences. Certains enregistrements comportent des champs vides, d'autres présentent une date plus récente et tous ont des numéros de téléphone différents.

Dans ce groupe de correspondance, vous pouvez effectuer une opération courante consistant à transférer les données mises à jour à tous les enregistrements du groupe de correspondance. Vous pouvez choisir de transférer les données à l'enregistrement principal, à tous les membres subordonnés du groupe de correspondance ou à tous les membres du groupe de correspondance. Le numéro de téléphone le plus récent en constitue un bon exemple.

Autre exemple : la récupération des données utiles dans les enregistrements concordants avant de les ignorer. Lorsque vous exécutez un fichier de permis de conduire par rapport à votre fichier interne, par exemple, vous pouvez récupérer les données de sexe et de date de naissance à ajouter à votre enregistrement interne.

Enregistrements avec la priorité la plus élevée en premier

Les opérations que vous configurez dans le groupe d'options Meilleur enregistrement doivent toujours commencer par le membre avec la priorité la plus élevée du groupe de correspondance (le principal) et se terminer par le dernier subordonné. Ainsi, les données sont récupérées de l'enregistrement avec la priorité la plus élevée à l'enregistrement avec la priorité la plus basse.

Vous devez donc veiller à hiérarchiser les enregistrements correctement en ajoutant une opération de post-correspondance Hiérarchisation du groupe avant l'opération Meilleur enregistrement.

27.4.10.1.1 Stratégies du meilleur enregistrement

Les stratégies fournies permettent de configurer facilement et rapidement des opérations courantes de meilleur enregistrement. Si aucune de ces stratégies ne répond à vos besoins, vous pouvez créer une stratégie du meilleur enregistrement personnalisée en utilisant votre propre code Python.

Les stratégies du meilleur enregistrement servent de critères pour exécuter des actions dans d'autres champs. Si les critères ne sont pas remplis, aucune action n'est exécutée.

❖ Exemple

Dans l'exemple de mise à jour d'un champ Téléphone avec les données les plus récentes, la stratégie de date peut être utilisée avec la priorité la plus récente pour mettre à jour l'enregistrement principal avec le dernier numéro de téléphone dans le groupe de correspondance. Cette partie (mise à jour de l'enregistrement principal avec le dernier numéro de téléphone) correspond à l'action. Vous pouvez aussi mettre à jour tous les enregistrements dans le groupe de correspondance (l'enregistrement principal et tous les enregistrements subordonnés) ou uniquement les enregistrements subordonnés.

! Restriction

La stratégie de date n'analyse pas la date, car elle ignore comment les données sont formatées. Assurez-vous que les données sont prédéfinies au format AAAAMMJJ afin que les comparaisons de chaînes fonctionnent correctement. Pour ce faire, vous pouvez aussi définir une stratégie personnalisée en utilisant le code Python pour analyser la date et utiliser une comparaison de dates.

27.4.10.1.1.1 Stratégies personnalisées du meilleur enregistrement et Python

Dans les stratégies du meilleur enregistrement prédéfinies, la transformation Match génère automatiquement le code Python qu'elle utilise pour le traitement. Ce code contient des variables requises pour la gestion du traitement.

Variables courantes

Les variables courantes contenues dans le code Python généré sont les suivantes :

Variable	Description
SRC	Désigne le champ source.
DST	Désigne le champ de destination.
RET	Désigne la valeur renvoyée, qui indique si la stratégie a réussi (T) ou échoué (F).

Variables NEWDST et NEWGRP

Utilisez les variables NEWDST et NEWGRP pour que la publication des données dans votre action du meilleur enregistrement soit indépendante des champs de stratégie. Si vous n'incluez pas ces variables, les données du champ de stratégie doivent également être mises à jour.

Variable	Description
NEWDST	Indicateur de nouvelle destination Cette variable de chaîne porte la valeur "T" lorsque l'enregistrement de destination est nouveau ou qu'il a été modifié depuis la dernière évaluation de la stratégie et la valeur "F" lorsqu'il n'a pas été modifié depuis la dernière évaluation. La variable NEWDST est utile uniquement si vous publiez vers plusieurs destinations telles que ALL ou SUBS dans l'option <i>Destination de publication</i>.

Variable	Description
NEWGRP	Indicateur de nouveau groupe Cette variable de chaîne porte la valeur "T" lorsque le groupe de correspondance a été modifié depuis la dernière évaluation de la stratégie et la valeur "F" lorsqu'il n'a pas été modifié depuis la dernière évaluation.

Exemple de NEWDST

Le code Python suivant a été généré à partir d'une stratégie NON_BLANK avec les options définies comme suit :

Définition	des options
<i>Stratégie du meilleur enregistrement</i>	NON_BLANK
<i>Priorité de la stratégie</i>	L'option de priorité n'est pas disponible pour la stratégie NON_BLANK.
<i>Champ de stratégie</i>	NORTH_AMERICAN_PHONE1_NORTH_AMERICAN_PHONE_STANDARDIZED.
<i>Destination de publication</i>	TOUS
<i>Publier une seule fois par destination</i>	OUI

Voici à quoi ressemble le code Python.

```
# Setup local temp variable to store updated compare condition
dct = locals()
# Store source and destination values to temporary variables
# Reset the temporary variable when the destination changes
if (dct.has_key('BEST_RECORD_TEMP') and NEWDST.GetBuffer() == u'F'):
    DESTINATION = dct['BEST_RECORD_TEMP']
else:
    DESTINATION =
DST.GetField(u'NORTH_AMERICAN_PHONE1_NORTH_AMERICAN_PHONE_STANDARDIZED')
SOURCE = SRC.GetField(u'NORTH_AMERICAN_PHONE1_NORTH_AMERICAN_PHONE_STANDARDIZED')
if len(SOURCE.strip()) > 0 and len(DESTINATION.strip()) == 0:
    RET.SetBuffer(u'T')
    dct['BEST_RECORD_TEMP'] = SOURCE
else:
    RET.SetBuffer(u'F')
    dct['BEST_RECORD_TEMP'] = DESTINATION
# Delete temporary variables
del SOURCE
del DESTINATION
```

❖ Exemple

NEWDST et NEWGRP

Supposons que vous disposiez de deux groupes de correspondance, chacun contenant trois enregistrements.

Enregistrements	de groupe de correspondance
Groupe de correspondance 1	Enregistrement A
	Enregistrement B
	Enregistrement C
Groupe de correspondance 2	Enregistrement D
	Enregistrement E
	Enregistrement F

Chaque nouvelle destination ou nouveau groupe de correspondance est marqué(e) d'un "T".

NEWGRP (T ou F)	NEWDST (T ou F)	Comparaison
T (nouveau groupe de correspondance)	T (nouvelle destination "A")	Enregistrement A > Enregistrement B
F	F	A > C
F	T (nouvelle destination "B")	B > A
F	F	B > C
F	T (nouvelle destination "C")	C > A
F	F	C > B
T (nouveau groupe de correspondance)	T (nouvelle destination "D")	D > E
F	F	D > F
F	T (nouvelle destination "E")	E > D
F	F	E > F
F	T (nouvelle destination "F")	F > D
F	F	F > E

27.4.10.1.1.2 Création d'une stratégie prédéfinie du meilleur enregistrement

Assurez-vous d'ajouter une opération post-comparaison Meilleur enregistrement au niveau de correspondance approprié dans l'éditeur de correspondance. De plus, pensez à mapper les champs d'entrée pertinents pour qu'ils soient disponibles pour cette opération.

Cette procédure vous permet de générer rapidement les critères requis pour votre action du meilleur enregistrement. Les stratégies disponibles représentent des cas d'utilisation courants.

1. Saisissez un nom pour cette opération de meilleur enregistrement.
2. Sélectionnez une stratégie dans l'option [Stratégie du meilleur enregistrement](#).

3. Sélectionnez une priorité dans l'option *Priorité de la stratégie*.
La sélection des valeurs dépend de la stratégie que vous avez sélectionnée à l'étape précédente.
4. Sélectionnez un champ dans le menu déroulant *Champ de stratégie*.
Le champ que vous sélectionnez ici sert de critère pour déterminer si une action du meilleur enregistrement est exécutée.

♣ Exemple

Le champ de stratégie sélectionné doit contenir des données correspondant à la stratégie que vous créez. Par exemple, si vous utilisez une stratégie de date la plus récente, assurez-vous que le champ sélectionné contient des données de date.

27.4.10.1.3 Création d'une stratégie personnalisée du meilleur enregistrement

1. Ajoutez une opération de meilleur enregistrement à votre transformation Match.
2. Saisissez un nom pour cette opération.
3. Dans l'option *Stratégie du meilleur enregistrement*, sélectionnez Personnalisé.
4. Sélectionnez un champ dans la liste déroulante *Champ de stratégie*.
5. Cliquez sur le bouton *Afficher/Modifier Python* pour créer votre code Python personnalisé et représenter votre stratégie personnalisée.
La fenêtre Editeur Python apparaît.

27.4.10.1.2 Actions du meilleur enregistrement

Les actions du meilleur enregistrement sont les fonctions que vous exécutez sur les données si le critère d'une stratégie est rempli.

♣ Exemple

Supposons que vous souhaitiez mettre à jour les numéros de téléphone de l'enregistrement principal. Cela implique qu'un enregistrement subordonné dans le groupe de correspondance comporte une date plus récente et, par conséquent, qu'il existe certainement un nouveau numéro de téléphone pour la personne concernée.

Dans ce cas, vous devez paramétrer une action qui commande à la transformation Match de mettre à jour le champ des numéros de téléphone dans l'enregistrement principal (action) si une date plus récente existe dans le champ de date (stratégie).

27.4.10.1.2.1 Sources et destinations

Il est important de connaître les différences entre sources et destinations dans une action de meilleur enregistrement.

La source est le champ d'où vous extrayez des données et la destination est l'endroit où vous publiez les données. Une source ou destination peut être un enregistrement principal ou subordonné dans un groupe de correspondance.

❖ Exemple

Dans l'exemple du numéro de téléphone, l'enregistrement subordonné porte la date la plus récente ; ainsi, les données du champ Téléphone (la source) sont utilisées et publiées dans l'enregistrement principal (la destination).

27.4.10.1.2.2 Publication une seule fois ou plusieurs fois par destination

Dans les options de meilleur enregistrement, vous pouvez choisir de publier vers une destination une seule fois ou plusieurs fois par action en définissant l'option *Publier une seule fois par destination*.

Vous pouvez stipuler que votre action de meilleur enregistrement s'arrête après sa première publication de données dans l'enregistrement de destination ou qu'elle poursuive avec les autres enregistrements de groupe de correspondance. Votre décision dépend de la nature des données que vous publiez et des enregistrements dans lesquels vous publiez. Les deux exemples ci-après illustrent les deux cas de figure.

Si vous publiez une seule fois dans chaque enregistrement de destination, une fois les données publiées pour un enregistrement spécifique, la transformation Match passe à l'exécution de l'action de meilleur enregistrement suivante (si plusieurs sont définies) ou à l'enregistrement suivant.

Si vous ne limitez pas l'action de cette manière, toutes les actions sont exécutées à chaque fois que la stratégie renvoie Vrai.

Indépendamment de cette définition, la transformation Match fonctionne toujours par ordre de priorité dans les membres du groupe de correspondance. En publiant dans l'enregistrement n°1 dans le graphique ci-dessous sans limiter la publication à une seule fois, on obtient le résultat suivant :

Groupe de correspondance	Action
Enregistrement n°1 (principal)	
Enregistrement n°2 (subordonné)	Tout d'abord, l'action est tentée en utilisant comme source l'enregistrement ayant la priorité la plus élevée (enregistrement n°2) parmi les autres enregistrements de groupe de correspondance.
Enregistrement n°3 (subordonné)	Ensuite, l'action est tentée en utilisant comme source l'enregistrement suivant ayant la priorité la plus élevée (enregistrement n°3).
Enregistrement n°4 (subordonné)	Enfin, l'action est tentée en utilisant comme source l'enregistrement ayant la priorité la plus faible (enregistrement n°4).

Dans cette dernière tentative, l'enregistrement n°4 était la dernière source utilisée pour l'action et peut donc être une source de données pour l'enregistrement de sortie. Cependant, si vous définissez votre action de meilleur enregistrement de sorte à publier une seule fois par enregistrement de destination, vous obtenez le résultat suivant :

Groupe de correspondance	Action
Enregistrement n°1 (principal)	
Enregistrement n°2 (subordonné)	Tout d'abord, l'action est tentée en utilisant comme source l'enregistrement ayant la priorité la plus élevée (enregistrement n°2) parmi les autres enregistrements de groupe de correspondance. Si cette tentative réussit, la transformation Match considère que cette action de meilleur enregistrement est terminée et passe à l'action suivante (le cas échéant) ou à l'enregistrement de sortie suivant. Si cette tentative échoue, la transformation Match passe au membre du groupe de correspondance ayant la priorité suivante la plus élevée et tente l'opération de publication.
Enregistrement n°3 (subordonné)	
Enregistrement n°4 (subordonné)	

Dans ce cas, l'enregistrement n°2 était la dernière source utilisée pour l'action de meilleur enregistrement ; par conséquent, il est la source des données publiées dans l'enregistrement de sortie.

27.4.10.1.2.3 Création d'une action du meilleur enregistrement

L'action du meilleur enregistrement correspond à la publication des données d'un enregistrement source vers un enregistrement de destination, basée sur les critères de votre stratégie du meilleur enregistrement.

1. Créez une stratégie prédéfinie ou personnalisée.
2. Sélectionnez le ou les enregistrements à publier dans l'option *Destination de publication*.
3. Définissez si vous voulez publier une seule fois ou plusieurs fois vers un enregistrement de destination dans l'option *Publier une seule fois par destination*.
4. Dans la table *Champs d'action du meilleur enregistrement*, sélectionnez votre champ source et votre champ de destination.
Lorsque vous sélectionnez un champ source, la colonne *Champ de destination* est automatiquement renseignée avec les données du même champ. Vous devez modifier le champ de destination s'il ne s'agit pas du champ dans lequel vous voulez publier vos données.
5. Pour créer une action du meilleur enregistrement personnalisée, sélectionnez Oui dans la colonne *Personnalisé*.
Vous pouvez à présent accéder à l'éditeur Python pour créer un code Python personnalisé pour votre action personnalisée.

27.4.10.1.3 Protection de destination

Les opérations Meilleur enregistrement et ID unique de la transformation Match vous permettent de modifier des enregistrements existants dans vos données. Vous pouvez être amené à devoir protéger des données d'enregistrements spécifiques ou des données d'enregistrements de sources d'entrée spécifiques contre tout risque d'écrasement.

L'onglet Protection de destination des opérations de la transformation Match vous permet de protéger vos données contre d'éventuelles modifications.

27.4.10.1.3.1 Protection des enregistrements de destination via des champs

1. Dans l'onglet Protection de destination, sélectionnez *Activer la protection de la destination*.
2. Sélectionnez une valeur dans la liste déroulante de l'option *Protection de destination par défaut*.
Cette valeur détermine si une destination est protégée dans le cas où le champ de protection de destination ne comporte pas de valeur valide.
3. Sélectionnez l'option *Spécifier la protection de destination par champ*, puis sélectionnez un champ dans la liste déroulante *Champ de protection de destination* (ou le *Champ ID unique protégé*).
Le champ sélectionné doit comporter une valeur O ou N pour spécifier l'action.

Tout enregistrement portant la valeur O dans le champ de protection de destination est protégé contre toute modification.

27.4.10.1.3.2 Protection des enregistrements de destination basée sur l'appartenance dans une source d'entrée

Pour pouvoir exécuter cette tâche, vous devez au préalable ajouter une opération Source d'entrée et définir des sources d'entrée.

1. Dans l'onglet Protection de destination, sélectionnez *Activer la protection de la destination*.
2. Sélectionnez une valeur dans la liste déroulante de l'option *Protection de destination par défaut*.
Cette valeur détermine si une destination (source d'entrée) est protégée dans le cas où vous ne définissez pas spécifiquement la source dans la table.
3. Sélectionnez l'option *Spécifier la protection destinataire par source*.
4. Sélectionnez une source d'entrée dans la première ligne de la colonne *Nom de source*, puis sélectionnez une valeur dans la colonne *Destination protégée* (ou *ID unique protégé*).
Répétez cette opération pour chaque source d'entrée pour laquelle vous voulez définir une protection.
N'oubliez pas que si vous ne spécifiez pas de protection pour chaque source, la valeur par défaut est utilisée.

27.4.10.2 ID unique

Un ID unique correspond à un champ dans vos données qui contient une valeur unique associée à un enregistrement ou groupe d'enregistrements. Vous pouvez, par exemple, utiliser un ID unique dans la base de données interne de votre société pour recevoir les mises à jour à un intervalle prédéfini (chaque semaine, chaque mois ou chaque trimestre). L'ID unique s'applique à un enregistrement de données de la même façon qu'un numéro d'identification national peut s'appliquer à une personne, comme dans le cas du numéro de sécurité sociale aux Etats-Unis et du numéro d'assurance national au Royaume-Uni. Il crée et suit les relations entre les données d'une exécution à l'autre. L'opération ID unique permet de définir votre propre ID de départ pour la génération de nouvelles clés ou de paramétrer une affectation dynamique sur la base des données existantes. L'opération de traitement post-correspondance ID unique permet également de commencer à l'endroit où l'ID unique le plus élevé de l'exécution précédente s'est terminé.

Rôle des ID uniques dans les groupes de correspondance

L'ID unique n'affecte pas nécessairement un ID à chaque enregistrement. Vous pouvez affecter le même ID à chaque enregistrement d'un groupe de correspondance (groupes d'enregistrements identifiés comme des correspondances).

Si vous affectez des ID directement à un groupe de décomposition, utilisez l'option *Champ Numéro de groupe* pour indiquer les enregistrements à regrouper. Par ailleurs, veillez à ce que les enregistrements soient triés par numéro de groupe afin que les enregistrements portant le même numéro de groupe apparaissent conjointement.

Si vous affectez des ID à des enregistrements faisant partie d'un groupe de correspondance obtenu à partir du processus de correspondance, vous n'avez pas à utiliser le *champ Numéro de groupe*.

i Remarque

Si vous affectez des ID directement à un groupe de décomposition et que le *champ Numéro de groupe* n'est pas défini, le processus de correspondance considère la collection de données complète comme un seul groupe de correspondance.

27.4.10.2.1 Options de traitement des ID uniques

L'opération de traitement post-comparaison ID unique combine les informations de mise à jour de la source et les informations de la base de données principale pour former une seule source d'informations de groupe de correspondance. L'opération peut ensuite affecter, combiner, fractionner et supprimer des ID uniques si nécessaire. Pour ce faire, utilisez l'option *Traitement de l'opération en cours*.

Opération	Description
Affecter	Affecter un nouvel ID à des enregistrements uniques dépourvus d'ID ou à tous les membres d'un groupe dépourvus d'ID. De plus, l'opération Affecter copie un ID existant si un membre d'un groupe de correspondance est déjà doté d'un ID. Une valeur est affectée à chaque enregistrement. - Les enregistrements d'un groupe de correspondance dans lequel un enregistrement avait un ID unique d'entrée partagent cette valeur avec d'autres enregistrements du groupe de correspondance dépourvu de valeur d'entrée. La première valeur trouvée est partagée. L'ordre est important : si un champ prioritaire peut être séquencé dans l'ordre croissant, placez une opération post-comparaison Hiérarchisation avant l'opération ID unique. - Les enregistrements d'un groupe de correspondance dans lequel deux enregistrement ou plus avaient différentes valeurs d'entrée ID unique conservent chacun leur valeur d'entrée. - Si tous les enregistrements d'un groupe de correspondance ne sont pas dotés d'une valeur d'entrée ID unique, le prochain ID disponible sera affecté à chaque enregistrement du groupe. Si le champ d'entrée GROUP_NUMBER est utilisé, les enregistrements dotés du même numéro de groupe doivent apparaître de manière consécutive dans la collection de données. i Remarque Utilisez le champ d'entrée GROUP_NUMBER uniquement pour traiter un groupe de décomposition susceptible de contenir des groupes de correspondance plus petits. Si le champ GROUP_NUMBER n'est pas spécifié, l'opération ID unique suppose que la collection entière forme un seul groupe.
AffecterCombiner	Exécute à la fois une opération Affecter et Combiner. Une valeur est affectée à chaque enregistrement. - Les enregistrements qui étaient dépourvus d'une valeur ID unique d'entrée et qui ne correspondent pas à d'autres enregistrements contenant une valeur ID unique d'entrée sont affectés au prochain ID disponible. Il s'agit d'enregistrements d'"ajout" qui pourraient être des enregistrements uniques ou des correspondances, mais pas avec un autre enregistrement préalablement affecté à une valeur ID unique. - Les enregistrements d'un groupe de correspondance dans lequel un ou plusieurs enregistrements avaient un ID unique d'entrée avec la même valeur ou des valeurs différentes partagent la première valeur trouvée avec tous les autres enregistrements du groupe de correspondance. L'ordre est important : si un champ prioritaire peut être séquencé dans l'ordre croissant, placez une opération post-comparaison Hiérarchisation avant l'opération ID unique. Si le champ d'entrée GROUP_NUMBER est utilisé, les enregistrements dotés du même numéro de groupe doivent apparaître de manière consécutive dans la collection de données. i Remarque Utilisez le champ d'entrée GROUP_NUMBER uniquement pour traiter un groupe de décomposition susceptible de contenir des groupes de correspondance plus petits. Si le champ GROUP_NUMBER n'est pas spécifié, l'opération ID unique suppose que la collection entière forme un seul groupe.

Opération	Description
Combiner	Garantit que les enregistrements d'un même groupe de correspondance sont dotés du même ID unique. Par exemple, cette opération peut être utilisée pour affecter à tous les membres d'un foyer le même ID unique. Plus précisément, si un foyer compte deux personnes qui partagent le même ID unique et qu'une troisième personne les rejoint avec un ID unique différent, l'opération Combiner peut être utilisée pour affecter le même ID aux trois personnes du foyer. Le premier enregistrement d'un groupe de correspondance portant un ID unique est celui doté de la priorité la plus élevée. Les autres enregistrements du groupe se voient affecter l'ID de cet enregistrement (en supposant que l'enregistrement en question ne soit pas protégé). L'opération Combiner n'affecte pas d'ID unique à un enregistrement qui n'a pas déjà un ID unique. Elle combine simplement l'ID unique des enregistrements d'un groupe de correspondance déjà dotés d'un ID unique. Si le champ d'entrée GROUP_NUMBER est utilisé, les enregistrements dotés du même numéro de groupe doivent apparaître de manière consécutive dans la collection de données. i Remarque Utilisez le champ d'entrée GROUP_NUMBER uniquement pour traiter un groupe de décomposition susceptible de contenir des groupes de correspondance plus petits. Si le champ GROUP_NUMBER n'est pas spécifié, l'opération ID unique suppose que la collection entière forme un seul groupe.
Supprimer	Supprime les ID uniques des enregistrements qui n'en ont plus besoin, à condition qu'ils ne soient pas protégés contre la suppression. Si vous utilisez un fichier et que vous recyclez des ID, ces ID sont ajoutés au fichier. Lors d'une suppression, les enregistrements dotés du même ID unique doivent être regroupés. Lorsque Match détecte qu'un groupe d'enregistrements doté du même ID unique est sur le point d'être supprimé : • Si des enregistrements sont protégés, tous les enregistrements du groupe sont considérés comme étant protégés. • Si le recyclage est activé, l'ID unique est recyclé une fois seulement, même si un groupe d'enregistrements portaient le même ID.

Opération	Description
Fractionner	Modifie les enregistrements uniques d'un groupe fractionné de sorte que les enregistrements n'appartenant pas au même groupe de correspondance porte un ID différent. L'enregistrement avec la priorité la plus élevée au sein du groupe conserve son ID unique. Les autres se voient affecter de nouveaux ID uniques. Pour cette opération, vous devez regrouper vos enregistrements par ID unique plutôt que par numéro de groupe de correspondance. Exemple : • Les enregistrements d'un groupe de correspondance dans lequel deux enregistrements ou plus portaient des valeurs d'entrée ID unique différentes ou des valeurs vides conservent chacun leur valeur d'entrée, renseignée ou vide selon l'enregistrement en question. • Les enregistrements qui étaient dépourvus d'une valeur ID unique d'entrée et qui ne correspondaient pas à d'autres enregistrements contenant une valeur ID unique d'entrée se voient affecter un ID unique vide de sortie. • Pour les enregistrements qui étaient pourvus de la même valeur ID unique d'entrée et qui ne sont plus considérés comme des correspondances, le premier enregistrement porte la valeur d'entrée. Les autres enregistrements se voient affecter de nouvelles valeurs ID unique.

27.4.10.2.2 Protection des ID uniques

La sortie de l'ID unique dépend du fait que le champ d'entrée dans l'enregistrement concerné est doté ou non d'une valeur qui indique que l'ID est protégé.

- Si le champ ID unique protégé n'est pas mappé comme champ d'entrée, Match suppose qu'aucun des enregistrements n'est protégé.
- Deux valeurs sont autorisées dans ce champ : O et N. Tout autre valeur est convertie en O.
La valeur N signifie que l'ID unique n'est pas protégé et que l'ID publié à la sortie peut être différent de l'ID d'entrée.
La valeur O signifie que l'ID unique est protégé et que l'ID publié à la sortie sera identique à l'ID d'entrée.
- Si le champ ID unique protégé est mappé comme champ d'entrée, une valeur autre que N signifie que les données d'entrée de l'enregistrement seront conservées dans le champ ID unique de sortie.

Ces règles de champs protégés s'appliquent à toutes les opérations de traitement ID unique.

27.4.10.2.3 Limitations des ID uniques

Le fait que certaines options dans l'opération ID unique soient basées sur la lecture d'un fichier ou la référence à une valeur de champ peut avoir des conséquences lorsque vous exécutez un environnement de serveur multiple ou serveur en temps réel et que vous partagez un fichier ID unique.

- Si vous lisez un fichier ou écrivez dans un fichier, le fichier ID unique doit figurer dans un système de fichiers partagés.
- Les ID recyclés sont utilisés dans l'ordre premier entré, premier sorti. Lorsque la transformation Match recycle un ID, elle ne vérifie pas s'il est déjà présent dans le fichier. Vous devez vous assurer qu'une valeur ID unique donnée n'est pas recyclée plus d'une fois.

27.4.10.2.4 Affectation d'ID uniques à l'aide d'un champ

1. Dans le groupe d'options ID unique, sélectionnez l'option *Valeur à partir du fichier*.
2. Définissez le nom du fichier et son chemin d'accès dans l'option *Fichier*.

Le fichier doit être au format XML et respecter la structure suivante :

```
<UniqueIdSession>  
  <CurrentUniqueId>477</CurrentUniqueId>  
</UniqueIdSession>
```

i Remarque

La valeur 477 est un exemple de valeur de départ. Toutefois, la valeur doit être au moins égale à 1.

27.4.10.2.5 Affectation d'ID uniques à l'aide d'une constante

De même qu'avec un fichier, vous pouvez affecter un ID unique de départ en définissant cette valeur.

1. Sélectionnez l'option *Valeur constante*.
2. Définissez l'option *Valeur de départ* sur la valeur d'ID souhaitée.

27.4.10.2.6 Affectation d'ID uniques à l'aide d'un champ

L'option de champ permet d'envoyer l'ID unique de départ via un champ dans votre source de données ou à partir d'une transformation User-Defined, par exemple.

L'ID unique de départ est transmis à la transformation Match avant que le premier nouvel ID unique ne soit demandé. Si aucun ID unique n'est reçu, le numéro de départ par défaut est le 1.

⚠ Attention

Attention lorsque vous utilisez l'option de champ ! Le champ que vous utilisez doit contenir la valeur d'ID unique par laquelle vous voulez que débute la numérotation séquentielle. Cela signifie que chaque enregistrement traité doit contenir ce champ et que chaque enregistrement doit avoir la même valeur dans ce champ.

Par exemple, supposons que vous utilisez la valeur 100 000. Lors du traitement, le premier enregistrement ou groupe de correspondance aura un ID de 100 001. Le second enregistrement ou groupe de correspondance aura un ID de 100 002, etc.

La valeur du premier enregistrement transmise à la transformation Match contient la valeur de début de l'incréméntation.

Il est impossible de savoir à l'avance quel enregistrement sera transmis à la transformation Match en premier (en raison du tri, par exemple) ; par conséquent, il est impossible de savoir par quelle valeur débutera l'incréméntation.

1. Sélectionnez l'option *Champ*.
2. Dans l'option *Champ ID unique de départ*, sélectionnez le champ contenant la valeur de l'ID unique de départ.

27.4.10.2.7 Affectation d'ID uniques à l'aide d'un GUID

Vous pouvez utiliser un identificateur global unique (GUID) comme ID unique.

Sélectionnez l'option *GUID*.

Le GUID est également appelé identificateur unique universel (UUID). La variante de l'UUID utilisée pour l'ID unique est une chaîne temporelle de 36 caractères au format : TimeLow-TimeMid-TimeHighAndVersion-ClockSeqAndReservedClockSeqLow-Node

Pour en savoir plus sur l'UUID, voir le document RFC (Request for Comments) sur <http://www.ietf.org/rfc/rfc4122.txt> .

27.4.10.2.8 Recyclage des ID uniques

Si des ID uniques sont supprimés avec l'option de traitement Supprimer, vous pouvez les réécrire dans un fichier pour les utiliser ultérieurement.

1. Dans le groupe d'options ID unique, définissez l'option *Traitement de l'opération en cours* sur *Supprimer*.
2. Sélectionnez l'option *Valeur à partir du fichier*.
3. Définissez le nom du fichier et son chemin d'accès dans l'option *Fichier*.
4. Définissez l'option *Recycler les ID uniques* sur *Oui*. Ce fichier est le même que celui qui peut être utilisé pour affecter un numéro ID de départ.

27.4.10.2.8.1 Utilisation de vos propres ID uniques recyclés

Si vous disposez d'ID que vous voulez recycler et utiliser dans un flux de données, vous pouvez les saisir dans un fichier à utiliser pour le recyclage des ID et la publication d'une valeur de départ pour les ID. Saisissez ces ID dans une balise XML <R></R>. Par exemple :

```
<UniqueIdSession>
  <CurrentUniqueId>477</CurrentUniqueId>
  <R>214</R>
  <R>378</R>
</UniqueIdSession>
```

27.4.10.2.9 Protection de destination

Les opérations Meilleur enregistrement et ID unique de la transformation Match vous permettent de modifier des enregistrements existants dans vos données. Vous pouvez être amené à devoir protéger des données

d'enregistrements spécifiques ou des données d'enregistrements de sources d'entrée spécifiques contre tout risque d'écrasement.

L'onglet Protection de destination des opérations de la transformation Match vous permet de protéger vos données contre d'éventuelles modifications.

27.4.10.2.9.1 Protection des enregistrements de destination via des champs

1. Dans l'onglet Protection de destination, sélectionnez *Activer la protection de la destination*.
2. Sélectionnez une valeur dans la liste déroulante de l'option *Protection de destination par défaut*.
Cette valeur détermine si une destination est protégée dans le cas où le champ de protection de destination ne comporte pas de valeur valide.
3. Sélectionnez l'option *Spécifier la protection de destination par champ*, puis sélectionnez un champ dans la liste déroulante *Champ de protection de destination* (ou le *Champ ID unique protégé*).
Le champ sélectionné doit comporter une valeur O ou N pour spécifier l'action.

Tout enregistrement portant la valeur O dans le champ de protection de destination est protégé contre toute modification.

27.4.10.2.9.2 Protection des enregistrements de destination basée sur l'appartenance dans une source d'entrée

Pour pouvoir exécuter cette tâche, vous devez au préalable ajouter une opération Source d'entrée et définir des sources d'entrée.

1. Dans l'onglet Protection de destination, sélectionnez *Activer la protection de la destination*.
2. Sélectionnez une valeur dans la liste déroulante de l'option *Protection de destination par défaut*.
Cette valeur détermine si une destination (source d'entrée) est protégée dans le cas où vous ne définissez pas spécifiquement la source dans la table.
3. Sélectionnez l'option *Spécifier la protection destinataire par source*.
4. Sélectionnez une source d'entrée dans la première ligne de la colonne *Nom de source*, puis sélectionnez une valeur dans la colonne *Destination protégée* (ou *ID unique protégé*).
Répétez cette opération pour chaque source d'entrée pour laquelle vous voulez définir une protection.
N'oubliez pas que si vous ne spécifiez pas de protection pour chaque source, la valeur par défaut est utilisée.

27.4.10.3 Grouper statistiques

L'opération post-comparaison Grouper statistiques doit être ajoutée à tout niveau de correspondance et toute opération post-comparaison pour lesquels vous avez besoin de statistiques sur vos groupes de correspondance ou vos sources d'entrée.

Cette opération peut aussi compter les statistiques de sources d'entrée logiques que vous avez déjà identifiées avec des valeurs dans un champ (prédéfini) ou de sources logiques que vous spécifiez dans l'opération Sources d'entrée.

Cette opération vous permet aussi d'exclure certaines sources logiques en fonction de vos critères.

i Remarque

Si vous choisissez de compter les statistiques de sources d'entrée dans l'opération Grouper statistiques, Match comptera aussi les statistiques de base sur vos groupes de correspondance.

Grouper champs de statistiques

Lorsque vous incluez une opération Grouper statistiques dans votre transformation Match, les champs suivants sont générés par défaut :

- GROUP_COUNT
- GROUP_ORDER
- GROUP_RANK
- GROUP_TYPE

De plus, si vous choisissez de générer des statistiques source, les champs suivants sont également générés et disponibles pour la sortie :

- SOURCE_COUNT
- SOURCE_ID
- SOURCE_ID_COUNT
- SOURCE_TYPE_ID

27.4.10.3.1 Génération de statistiques de base uniquement

Cette tâche permet de générer des statistiques sur vos groupes de correspondance, telles que le nombre d'enregistrements dans chaque groupe, les enregistrements principaux ou subordonnés, etc.

1. Ajoutez une opération Grouper statistiques à chaque niveau de correspondance souhaité en sélectionnant *Traitement post-correspondance* dans un niveau de correspondance, en cliquant sur le bouton *Ajouter*, puis en sélectionnant *Grouper statistiques*.
2. Sélectionnez *Générer uniquement les statistiques de base*.
3. Cliquez sur le bouton *Appliquer* pour sauvegarder vos modifications.

27.4.10.3.2 Génération de statistiques pour toutes les sources d'entrée

Avant de démarrer cette tâche, assurez-vous d'avoir défini vos sources d'entrée dans l'opération Sources d'entrée.

Utilisez cette procédure pour générer des statistiques pour toutes vos sources au sein du job.

1. Ajoutez une opération Statistiques de groupe au niveau de correspondance approprié.
2. Sélectionnez l'option [Générer des statistiques source à partir des sources d'entrée](#).

Des statistiques sont alors générées pour toutes les sources d'entrée que vous avez définies dans l'opération Sources d'entrée.

27.4.10.3.3 Comptage des statistiques pour les sources d'entrée générées par des valeurs dans un champ

Pour cette tâche, il n'est pas nécessaire de définir des sources d'entrée avec l'opération Sources d'entrée. Vous pouvez spécifier des sources d'entrée pour Match en utilisant des valeurs dans un champ.

Cette tâche vous permet de générer des statistiques pour toutes les sources d'entrée identifiées par des valeurs dans un champ ou de générer des statistiques pour un sous-ensemble de sources d'entrée.

1. Ajoutez une opération Statistiques de groupe au niveau de correspondance approprié.
2. Sélectionnez l'option [Générer des statistiques source à partir des valeurs source](#).
3. Sélectionnez un champ dans la liste déroulante [Champ de source logique](#) contenant les valeurs pour vos sources logiques.
4. Entrez une valeur dans le champ [Valeur source logique par défaut](#). Cette valeur est utilisée si le champ de source logique est vide.
5. Sélectionnez l'une des options suivantes :

Option	Description
Compter toutes les sources	Permet de compter toutes les sources. Si vous sélectionnez cette option, cliquez sur le bouton Appliquer pour sauvegarder vos modifications. Cette tâche est terminée.
Sélectionner les sources à compter	Permet de définir un sous-ensemble de sources d'entrée à compter. Si vous sélectionnez cette option, vous pouvez passer à l'étape 6 de la tâche.

6. Sélectionnez la valeur appropriée dans l'option [Indicateur de nombre par défaut](#). Sélectionnez Oui pour compter toute source non spécifiée dans la [table Définir manuellement des indicateurs de nombre de sources logiques](#). Si vous ne spécifiez pas de source dans la table, vous comptez toutes les sources.
7. Sélectionnez [Générer automatiquement les sources](#) pour compter les sources en fonction d'une valeur dans un champ spécifié dans l'option [Champ de l'indicateur de nombre prédéfini](#). Si vous ne spécifiez pas de source dans la [table Définir manuellement des indicateurs de nombre de sources logiques](#), vous indiquez à la transformation Match de compter toutes les sources en fonction de la valeur (Oui ou Non) de ce champ.
8. Dans la [table Définir manuellement des indicateurs de nombre de sources logiques](#), ajoutez autant de lignes que nécessaire afin d'inclure toutes les sources à compter.

i Remarque

C'est le premier critère sur lequel se base la transformation Match pour déterminer s'il faut compter des sources.

- Ajoutez une valeur source et un indicateur de nombre à chaque ligne pour indiquer à la transformation Match les sources à compter.

→ Conseil

Si vous disposez d'un grand nombre de sources, mais que vous voulez en compter deux uniquement, vous pouvez accélérer le temps de configuration en définissant l'option *Indicateur de nombre par défaut* sur Non et en configurant la *table Définir manuellement des indicateurs de nombre de sources logiques* pour compter ces deux sources. En utilisant la même méthode, vous pouvez configurer Statistiques de groupe de sorte à compter toutes les sources et non pas deux sources uniquement.

27.4.10.4 Sélection de l'indicateur de sortie

En ajoutant une opération Sélection de l'indicateur de sortie à chaque niveau de correspondance souhaité (Traitement post-correspondance), vous pouvez marquer des types d'enregistrement spécifiques pour évaluation ou routage en aval dans votre flux de données.

L'ajout de cette opération génère le champ de sortie Select_Record à inclure dans votre schéma de sortie. Ce champ de sortie est renseigné par un O ou un N selon le type d'enregistrement sélectionné dans l'opération.

Vos résultats apparaîtront dans le rapport Faire correspondre la sélection de sortie à la source d'entrée. Ce rapport vous permet de déterminer de quelle source ou groupe source proviennent les enregistrements et combien de chacun des types d'enregistrements ont été générés par source ou groupe source.

Type d'enregistrement	Description
<i>Unique</i>	Enregistrements n'appartenant à aucun groupe de correspondance. Aucun enregistrement correspondant n'a été trouvé. Peuvent être issus de sources contenant une source Normal ou Special.
<i>Enregistrements principaux de source unique</i>	Membre le plus haut placé d'un groupe de correspondance dont les membres sont tous issus de la même source. Peut être issu d'une source Normal ou Special.
<i>Enregistrements subordonnés de source unique</i>	Enregistrement issu d'une source Normal ou Special et membre subordonné d'un groupe de correspondance.
<i>Enregistrements principaux de sources multiples</i>	Membre le plus haut placé d'un groupe de correspondance dont les membres sont issus de deux sources ou plus. Peut être issu d'une source Normal ou Special.
<i>Enregistrements subordonnés de sources multiples</i>	Enregistrement subordonné d'un groupe de correspondance issu d'une source Normal ou Special dont les membres sont issus de deux sources ou plus.
<i>Correspondances Suppression</i>	Membre subordonné d'un groupe de correspondance incluant un enregistrement de priorité supérieure issu d'une source de type Suppression. Peut être issu d'une source Normal ou Special.

Type d'enregistrement	Description
<i>Enregistrements uniques Suppression</i>	Enregistrement issu d'une source Suppression pour lequel aucun enregistrement correspondant n'a été trouvé.
<i>Enregistrements principaux Suppression</i>	Enregistrement issu d'une source Suppression et membre le plus haut placé d'un groupe de correspondance.
<i>Enregistrements subordonnés Suppression</i>	Enregistrement provenant d'une source de type Suppression et membre subordonné d'un groupe de correspondance.

27.4.10.4.1 Marquage des types d'enregistrements sources pour la sortie possible

1. Dans l'éditeur de correspondance, ajoutez une opération Sélection de l'indicateur de sortie pour chaque niveau de correspondance souhaité.
2. Sélectionnez les types d'enregistrements pour lesquels vous voulez renseigner le champ Select_Record par O.

Le champ de sortie Select_Record peut être généré à partir de la transformation Match pour être utilisé en aval dans le flux de données. Cela peut s'avérer très utile si vous voulez ensuite séparer les correspondances de suppression ou les enregistrements principaux Suppression de vos données (à l'aide d'une transformation Case, par exemple).

27.4.11 Correspondance par association

La correspondance par association combine les résultats de correspondance de deux sets de correspondance (transformations) ou plus pour rechercher des correspondances qui n'ont pas été trouvées dans un set de correspondance unique.

Vous pouvez configurer la correspondance par association dans la transformation Associate. Cette transformation se comporte comme un autre set de correspondance dans votre flux de données, dont vous pouvez dériver des statistiques.

Ce set de correspondance a deux objectifs : Tout d'abord, il fournit un accès à toutes les données générées, à tous les niveaux de correspondance de tous les sets de correspondance. Ensuite, il fournit les résultats chevauchés de plusieurs critères, comme le nom et l'adresse avec le nom et le NSS sous forme d'ID unique. C'est ce que l'on appelle généralement la correspondance par association.

Numéros de groupe

La transformation Associate accepte un champ Numéro de groupe, généré par les transformations Match pour chaque résultat de correspondance qui doit être combiné. La transformation peut ensuite générer un nouveau numéro de groupe associé.

La transformation Associate peut s'exécuter sur tous les enregistrements d'entrée ou sur une collection de données à la fois. La collection de données est requise pour le support en temps réel.

❖ Exemple

Exemple d'association

Imaginons que vous travailliez dans un collège technique et que vous souhaitiez envoyer des informations à tous les élèves avant la rentrée. Vous savez que beaucoup d'entre eux ont une adresse locale temporaire et une adresse permanente.

Dans cet exemple, vous pouvez effectuer une correspondance sur le nom, l'adresse et le code postal dans un set de correspondance et sur le nom et le numéro de sécurité sociale (NSS), disponible pour chaque étudiant, dans un autre set de correspondance.

La transformation Associate combine ensuite les deux sets de correspondance pour générer des groupes de correspondance associés. Cela vous permet d'identifier les personnes susceptibles d'avoir plusieurs adresses et donc d'optimiser vos efforts en termes de marketing et de courrier.

27.4.12 Correspondance Unicode

La correspondance Unicode vous permet de mettre en correspondance des données Unicode. Vous pouvez traiter toutes les données Unicode non Latin1, avec un traitement spécifique pour les données chinoises, japonais, coréennes et taïwanaises (ou CJKT).

Correspondance des données chinoises, japonais, coréennes et taïwanaises

Indépendamment de la langue propre à chaque pays, le processus de correspondance des données CJKT est identique. Par exemple, la transformation Match :

- Considère les caractères demi-largeur et pleine largeur de la même manière.
- Considère les chiffres de script natif et les chiffres arabes de la même manière. Elle peut interpréter des chiffres écrits dans un script natif. Cette fonction peut être contrôlée dans l'option Convertir le texte en chiffres du groupe d'options Critère.
- Inclut des variations pour les caractères des noms de personnes et de sociétés dans les données référentielles.
- Considère les termes d'entreprise tels que Société de la même manière que leur variante (Sté) lors du processus de comparaison. Pour trouver les abréviations, la transformation utilise des variations de script natif de l'alphabet anglais lors de la mise en correspondance des noms de société.
- Ignore les marqueurs facultatifs couramment utilisés pour la ville, le département, etc. dans la comparaison des données d'adresse.
- Traite intelligemment les variations dans un marqueur de bâtiment.

Fonctionnalités de correspondance spécifiques au japonais

Avec les données japonaises, la transformation Match :

- Considère les marqueurs d'îlots (bâtiments) tels que chome et banchi de la même manière que ceux utilisés avec des données avec trait d'union.
- Considère les mots avec ou sans Okurigana de la même manière dans les données d'adresse.
- Considère les variations des marqueurs *no*, *ga*, etc. de la même manière.
- Considère les variations de trait d'union ou ligne en pointillé de la même manière.

Limitations de la correspondance Unicode

La fonction de correspondance Unicode ne prend pas en charge les opérations suivantes :

- Effectuer des conversions de données en chinois simplifié et traditionnel.
- Faire correspondre des scripts non phonétiques tels que le kanji, le chinois simplifié, etc.

Router des enregistrements en fonction de l'ID pays avant la mise en correspondance.

Avant d'envoyer les données Unicode dans le processus de correspondance, vous devez regrouper les données par pays pour séparer les transformations Match. Pour ce faire, vous devez utiliser une transformation Case qui permet de router les données propres au pays en fonction de l'ID pays.

→ Conseil

L'Assistant de correspondance peut exécuter cette tâche si vous utilisez la stratégie multinationale.

Mise en correspondance multi-scripts

La mise en correspondance multi-scripts vous permet de traiter des données susceptibles de contenir plusieurs scripts en convertissant les scripts en Latin1. Imaginons, par exemple, qu'un enregistrement contienne du Latin1 et l'autre du katakana ou l'un du latin et l'autre du cyrillique. Sélectionnez *Oui* pour activer la mise en correspondance multi-scripts. Si vous préférez traiter les données sans les convertir en Latin1, ne modifiez pas l'option Mise en correspondance multi-scripts définie sur *Non*. Voici deux exemples de noms pour lesquels une correspondance a été trouvée à l'aide de la mise en correspondance multi-scripts :

Nom	Peut être mis en correspondance avec...
Viktor Ivanov	Виктор Иванов
Takeda Noburu	スッセフレ

Paramètres régionaux

Cette option permet de définir les paramètres régionaux pour le champ Critère. La définition des paramètres régionaux est recommandée si vous avez l'intention d'utiliser la fonction Convertir le texte en chiffres pour spécifier les paramètres régionaux des données en vue de la conversion à des fins de correspondance. Voici quatre exemples de conversion de texte en chiffres :

Langue	Texte	Chiffres
Français	quatre mille cinq cents soixante-sept	4567
Allemand	dreitausendzwei	3002
Italien	cento	100
Espagnol	ciento veintisiete	127

Pour plus d'informations sur ces options de correspondance, voir la section relative à la transformation Match du *Guide de référence*.

27.4.12.1 Configuration de la correspondance Unicode

1. Utilisez une transformation Case pour router vos données vers une transformation Match qui traite ce type de données.
2. Ouvrez la configuration de la transformation [Address.Japan_MatchBatch](#) Match et enregistrez-la sous un nouveau nom.
3. Définissez l'option [Moteur de correspondance](#) dans les options de la transformation Match sur une valeur qui reflète le type de données traitées.
4. Configurez vos critères et autres opérations. Pour plus d'informations sur les options Critères Match, voir la section relative à la transformation Match du *Guide de référence*.

❖ Exemple

- Dans la mesure du possible, utilisez des critères de composants analysés pour les données d'adresse, de société et de nom tels que Primary_Name ou Person1_Family_Name1.
- Si vous avez analysé des données d'adresse, de société et de nom qui n'ont pas de critère correspondant, utilisez les critères Address_Data1-5, Firm_Data1-3 et Name_Data1-3.
- Pour toutes les autres données sans critère correspondant, utilisez le critère Personnalisé.

27.4.13 Critères de correspondance phonétiques

Dans certains cas, l'utilisation de données phonétiques comme critère peut produire un plus grand nombre de correspondances qu'une comparaison par rapport à d'autres critères comme le nom ou des données société.

Les fonctions Double Metaphone ou Soundex permettent de renseigner un champ et de l'utiliser pour créer des groupes de décomposition ou comme critère dans une correspondance.

Fonction	Description
Double metaphone	Encode la chaîne d'entrée à l'aide de l'algorithme Double Metaphone et renvoie une chaîne.
Soundex	Encode la chaîne d'entrée à l'aide de l'algorithme Soundex et renvoie une chaîne. Utilisez cette fonction pour pousser la fonction vers le niveau de la base de données.

La correspondance par rapport aux données du champ Nom donne des résultats différents de la correspondance basée sur des données phonétiques.

❖ Exemple

Nom	Score de comparaison
Smith	Similaire à 72%
Smythe	

Nom	Clé phonétique (primaire)	Score de comparaison
Smith	SMO	Similaire à 100 %
Smythe	SMO	

27.4.13.1 Options de critères de correspondance phonétiques

Si vous voulez établir la correspondance d'après des données phonétiques, configurez les options des critères comme suit.

Option	Valeur
<i>Comparer les données à l'aide de</i>	Similarité de champs compare l'intégralité des données de champ sous forme de chaîne unique.
<i>Rechercher des lettres transposées</i>	Définir sur n° de logiciel gère les caractères transposés de la même manière qu'elle gère les caractères ne présentant pas de correspondance.
<i>Score d'ajustement des initiales</i>	Définir sur 0 (zéro) (par défaut) pour désactiver le score des initiales.
<i>Score d'ajustement des sous-chaînes</i>	Définir sur 0 (zéro) pour désactiver le score des sous-chaînes (par défaut).
<i>Score d'ajustement des abréviations</i>	Définir sur 0 (zéro) pour désactiver l'option de score des abréviations.

27.4.13.2 Scores de correspondance phonétique

Si vous établissez la correspondance d'après les critères phonétiques seulement, définissez les options de scores de correspondance comme suit.

Option	Valeur
Score de correspondance	Définir sur 100 le score de similarité minimal requis pour que les enregistrements soient considérés comme une correspondance.
Aucun score de correspondance	Définir sur 99 le score de similarité maximal requis pour que les enregistrements soient considérés comme une non-correspondance.

Si vous établissez la correspondance d'après plusieurs critères, dont un critère phonétique, placez le critère phonétique en premier dans l'ordre des critères et définissez vos options de scores de correspondance comme suit :

Option	Valeur
Score de correspondance	Définir sur 101 pour garantir que ce critère ne suffit pas pour le logiciel, à considérer deux enregistrements comme une correspondance et que vous souhaitez envisager d'autres critères dans le cadre du processus de comparaison.
Aucun score de correspondance	Définir sur 99 le score de similarité maximal requis pour que les enregistrements soient considérés comme une non-correspondance.

27.4.13.3 Champs vides de correspondance phonétique

Lorsque vous utilisez des groupes de décomposition, les enregistrements qui ne comportent aucune valeur ne se trouvent pas dans le même groupe que ceux qui comportent une valeur (sauf si vous configurez la correspondance par rapport aux champs vides).

❖ Exemple

Prenons les deux enregistrements d'entrée suivants :

Mr Johnson	100 Main St	La Crosse	WI	54601
Scott Johnson	100 Main St	La Crosse	WI	54601

Une fois ces enregistrements traités par la transformation Data Cleanse, le premier enregistrement comporte un champ de prénom vide et, de ce fait, un champ phonétique vide. Dans ces conditions, vous ne pouvez obtenir aucune correspondance en créant des groupes de décomposition. Si vous ne créez pas de groupes de décomposition, vous ne pouvez obtenir un résultat qu'en utilisant la correspondance des vides.

27.4.13.4 Longueur des données de correspondance phonétique

La longueur que vous affectez à une sortie de fonction phonétique revêt une certaine importance, comme le montre cet exemple.

❖ Exemple

Prénom (nom)	Sortie
S (Johnson)	S
Scott (Johnson)	SKT

Supposez que ces deux enregistrements représentent la même personne. Dans cet exemple, si vous effectuez la décomposition d'après plusieurs caractères, ces enregistrements se trouveront dans des groupes de décomposition différents et ne seront donc pas comparés.

27.4.14 Configuration des rapports de correspondance

Générer des rapports de correspondance visant à faciliter l'analyse des résultats.

Vous disposez de nombreux rapports de correspondance visant à faciliter l'analyse des résultats du processus de correspondance :

- Rapport Contribution de correspondance
- Résumé des critères de correspondance
- Rapport Exemple de correspondance en double
- Rapport Faire correspondre la sélection de sortie à la source d'entrée
- Rapport Fréquence multi-source de correspondance
- Rapport Résumé des statistiques source de correspondance

27.4.14.1 Ajout de Statistiques de groupe dans une transformation Match

Pour générer le rapport Résumé des statistiques source de correspondance, incluez une opération Statistiques de groupe dans vos transformations Match et Associate.

Si vous voulez effectuer un suivi de vos statistiques source d'entrée, vous pouvez ajouter une opération Sources d'entrée dans la transformation Match afin de définir vos sources et choisir de générer des statistiques pour les sources d'entrée dans une opération Statistiques de groupe.

i Remarque

Pour générer des statistiques source d'entrée dans l'opération Statistiques de groupe, vous pouvez également définir des sources d'entrée à l'aide de valeurs de champ. Vous ne devez pas nécessairement intégrer une opération Sources d'entrée dans la transformation Match.

27.4.14.2 Générer des statistiques de rapport

Pour générer les données que vous voulez obtenir dans les rapports de correspondance(autres que le rapport Résumé des statistiques source de correspondance), vous devez définir l'option *Générer les données de rapport* sur Oui dans les transformations Match et Associate.

En activant la génération de données de rapport, vous pouvez obtenir des informations sur les groupes de décomposition, les critères ayant contribué à la création d'une correspondance, etc.

i Remarque

Néanmoins, l'activation de l'option de rapport peut avoir une incidence sur les performances du traitement. Une fois que vous avez testé le flux de données de façon approfondie, il est préférable d'activer la génération de données.

27.4.14.3 Utilisation de noms pour les sets de correspondance, les niveaux et les opérations

Utiliser des noms uniques dans les transformations Match et Associate pour les sets de correspondance, les niveaux et chacune de vos opérations préalables ou postérieures à la correspondance.

Utilisez des noms uniques dans les transformations Match et Associate pour les sets de correspondance, les niveaux et chacune de vos opérations préalables ou postérieures à la correspondance. L'utilisation de noms uniques permet d'obtenir des données plus précises dans vos rapports (Hiérarchisation du groupe et Statistiques de groupe, par exemple) et d'identifier les éléments qui produisent les données que vous recherchez.

27.4.14.4 Insertion des champs de sortie appropriés

Inclure ces trois champs de sortie dans la transformation Match pour publier des données dans le rapport
Exemple de correspondance en double

Champ	Description
Type_correspondance	Indique quel type de correspondance a intégré l'enregistrement à un groupe de correspondance : • vide • D = pilote • R = règle • W = pondéré
Numéro_groupe	Spécifie les enregistrements appartenant au même groupe de correspondance, qui partagent le même numéro de groupe.
Score_correspondance	Fournit les informations suivantes : • Score de similarité des critères ou modèles • Score total pondéré • Espaces vides

27.5 Address Cleanse

Informations sur la préparation des données, la configuration du logiciel et description des éléments de sortie pour le nettoyage des adresses.

Address Cleanse fournit une forme corrigée, complète et standardisée de vos données d'adresse d'origine. Avec la transformation USA Regulatory Address Cleanse et pour certains pays la transformation Global Address Cleanse, Address Cleanse peut également corriger ou ajouter des codes postaux. Avec la transformation DSF2 Walk Sequencer, vous pouvez ajouter à vos données des informations de tournée.

Informations associées

[Préparation des données d'entrée \[page 569\]](#)

[Configuration des fichiers de référence \[page 574\]](#)

[Pays pris en charge \(Global Address Cleanse\) \[page 591\]](#)

27.5.1 Fonctionnement d'Address Cleanse

Les transformations USA Regulatory Address Cleanse et Global Address Cleanse nettoient vos données de la manière suivante :

Processus	Description
Vérification de la cohérence des codes	Elles vérifient que la localité, la région et les codes postaux concordent les uns avec les autres. Par exemple, Address Cleanse peut ajouter (ou vérifier) un code postal en utilisant les informations de localité et de région dans les données d'origine (selon le pays).
Standardisation de l'apparence des lignes d'adresse	Utilise les options de standardisation des adresses définies dans la transformation pour obtenir des adresses qui incluent ou excluent la ponctuation, utilise des textes en majuscules ou minuscules, ou abrège ou écrit en toute lettres les types d'adresses, par exemple.
Identification des adresses non livrables	Pour les enregistrements des États-Unis uniquement, identifie une adresse non livrable, telle qu'une parcelle vide ou un bâtiment condamné.
Affectation de codes de diagnostic	Affecte des codes de diagnostic qui indique le motif des adresses non affectées, ou de la méthode de correction des adresses par exemple.

Informations associées

[Définition des options de standardisation \[page 576\]](#)

27.5.1.1 Transformations Address Cleanse

Tableau qui répertorie les transformations Address Cleanse et leur objectif.

Transformation	Description
DSF2 Walk Sequencer	Ajoute des informations de séquence de distribution aux données lors de l'exécution du séquençage de tournée DSF2 qu'il est possible d'utiliser avec un logiciel de tri préalable pour remplir les conditions pour des remises de séquence de tour. i Remarque Le logiciel ne place pas les données dans l'ordre de la tournée.

Transformation	Description
Global Address Cleanse et moteurs	Nettoie vos données d'adresses de tous les pays pris en charge (hors certification des États-Unis). Vous devez configurer la transformation Global Address Cleanse conjointement à un ou plusieurs moteurs Global Address Cleanse (Canada, Global Address ou USA). Avec cette transformation, vous pouvez créer le Software Evaluation and Recognition Program (SERP) de la poste canadienne (Statement of Address Accuracy Report), le rapport AMAS (Australia Post's Address Matching Processing Summary) de la poste australienne et le rapport SOA (Statement of Accuracy) de Nouvelle-Zélande.
USA Regulatory Address Cleanse	Identifie, analyse, valide et corrige les données d'adresses des États-Unis (dans la page de code Latin 1) conformément au CASS (Coding Accuracy Support System) américain. Peut créer le Formulaire 3553 USPS et éditer plusieurs codes utiles dans vos enregistrements. Vous pouvez également exécuter l'application en mode non-certification et générer des listes de propositions. Certaines options comprennent : DPV, DSF2 (augmentation), eLOT, EWS, GeoCensus, LACS-Link, NCOALink, RDI, SuiteLink, listes de propositions (pas pour la certification) et Z4Change.
Global Suggestion Lists	Effectue des propositions pour les correspondances d'adresses possibles dans les données des États-Unis, du Canada, et des adresses globales. Cette transformation est habituellement utilisée pour le traitement en temps réel et ne standardise pas les adresses. Utilisez une transformation Country ID avant cette transformation dans le flux de données. Si vous désirez également standardiser vos données d'adresses, utilisez la transformation Global Address Cleanse après la transformation Global Suggestion Lists dans le flux de données.
Country ID	Identifie le pays de destination pour l'enregistrement et édite un code ISO. Utilisez cette transformation avant la transformation Global Suggestion Lists dans votre flux de données. (Il n'est pas nécessaire de placer la transformation Country ID avant les transformations Global Address Cleanse ou USA Regulatory Address Cleanse.)

27.5.1.2 Classes des données et champs d'entrée et de sortie

Utilisez les champs d'entrée et de sortie multilignes ou discrets, ou une combinaison des deux. Désignez également le type d'information de sortie à l'aide des classes de champs.

Les transformations Address Cleanse acceptent et éditent les formats de ligne d'adresse discret, multiligne et hybride.

Formats de la ligne d'adresse	Description
Multiligne	Conserver les données d'adresse de sortie dans la même disposition que celle des données d'entrée. Le logiciel applique une abréviation intelligente lorsque les données dépassent les longueurs de champ maximales, et utilise des majuscules et standardise les données de sortie en fonction des options de standardisation définies dans la transformation.

Formats de la ligne d'adresse	Description
Discret	Divise les données d'adresse d'entrée en des éléments d'adresse plus petits pour la sortie. Édite également des champs supplémentaires créés par le logiciel, par exemple les codes de statut ou d'erreur. Le style de certains composants est contrôlé par les options de standardisation, et le logiciel n'applique pas d'abréviation intelligente pour que les composants s'adaptent aux champs de sortie.

Lorsque vous définissez la transformation USA Regulatory Address Cleanse ou Global Address Cleanse transform, vous pouvez inclure des champs de sortie qui contiennent des informations spécifiques en sélectionnant une classe d'adresse Champ et une classe Champ générées.

Classe d'adresse Champ généré	Classe Champ généré
Distribution	<i>Analysé</i> : contient l'entrée analysée avec application d'une certaine standardisation. Les champs soumis à la standardisation sont la localité, la région et le code postal. <i>Meilleur</i> : contient les données analysées si l'adresse n'est pas affectée ou les données corrigées d'une adresse affectée. <i>Corrigé</i> : contient les données affectées après les recherches de répertoire et reste vide si l'adresse n'est pas affectée.
Double	<i>Analysé</i> , <i>Meilleur</i> et <i>Corrigé</i> : contient les détails d'adresse DOUBLE disponibles à l'entrée.
Officiel	<i>Analysé</i> : contient l'entrée analysée avec application d'une certaine standardisation. <i>Meilleur</i> : contient les informations des répertoires définis par le service postal lorsque l'adresse est affectée. Contient l'entrée analysée lorsqu'une adresse n'est pas affectée. <i>Corrigé</i> : contient les informations des répertoires définis par le service postal lorsque l'adresse est affectée et reste vide si l'adresse n'est pas affectée.

27.5.2 Rapports Address cleanse

Le logiciel génère des rapports postaux pour USA et Global Address Cleanse.

De nombreux rapports générés par le logiciel incluent des rapports qui sont requis par les autorités postales mondiales ou le service postal américain.

La transformation USA Regulatory Address Cleanse crée l'USPS Form 3553 (requis pour CASS) et le rapport résumé de NCOALink. La transformation Global Address Cleanse crée des rapports sur vos données, y compris le SERP (Statement of Address Accuracy Report) canadien, le rapport AMAS de la poste australienne et le rapport SOA Nouvelle-Zélande.

La plupart des rapports postaux et de régulation nécessitent que le logiciel soit certifié périodiquement par l'autorité postale correspondante.

27.5.3 Préparation des données d'entrée

Avant de commencer le nettoyage d'adresses, vous devez décider du type de format de ligne d'adresse en entrée. Les transformations USA Regulatory Address Cleanse et Global Address Cleanse acceptent les données d'entrée de la même manière.

⚠ Attention

La transformation USA Regulatory Address Cleanse n'accepte pas les données Unicode. Si un enregistrement d'entrée comporte des caractères extérieurs à la page de code Latin1 (valeur de caractère supérieure à 255), la transformation USA Regulatory Address Cleanse ne traite pas les données. A la place, l'enregistrement d'entrée est envoyé au champ de sortie standardisé correspondant sans aucun traitement. Aucun autre champ de sortie (composant par exemple) n'est renseigné pour cet enregistrement. Si votre base de données Unicode comporte des adresses des Etats-Unis valides en jeu de caractères Latin1, la transformation USA Regulatory Address Cleanse les traite comme d'habitude.

Formats de ligne d'adresse acceptés

Les tableaux suivants répertorient les formats de ligne d'adresse : multiligne, hybride et discret.

i Remarque

Pour tous les formats multiligne et hybride répertoriés, vous n'êtes pas obligés d'utiliser tous les champs multilignes d'un format sélectionné (par exemple, Multiligne1-12). Cependant, vous devez commencer par Multiligne1 et continuer dans l'ordre. Vous ne pouvez pas sauter de numéros, par exemple de Multiligne1 à Multiligne3.

Formats multiligne et multiligne hybride

Exemple 1	Exemple 2	Exemple 3	Exemple 4	Exemple 5
Multiligne1	Multiligne1	Multiligne1	Multiligne1	Multiligne1
Multiligne2	Multiligne2	Multiligne2	Multiligne2	Multiligne2
Multiligne3	Multiligne3	Multiligne3	Multiligne3	Multiligne3
Multiligne4	Multiligne4	Multiligne4	Multiligne4	Multiligne4
Multiligne5	Multiligne5	Localité3	Multiligne5	Multiligne5
Multiligne6	Multiligne6	Localité2	Localité2	Multiligne6
Multiligne7	Multiligne7	Localité1	Localité1	Localité1
Multiligne8	Dernière ligne adresse	Région1	Région1	Région1
Pays (facultatif)	Pays (facultatif)	Code postal (global) ou Code postal1 (ré- gion Etats-Unis)	Code postal (global) ou Code postal1 (ré- gion Etats-Unis)	Code postal (global) ou Code postal1 (ré- gion Etats-Unis)
		Pays (facultatif)	Pays (facultatif)	Pays (facultatif)

Formats de lignes discrets

Exemple 1	Exemple 2	Exemple 3	Exemple 4
Ligne_Adresse	Ligne_Adresse	Ligne_Adresse	Ligne_Adresse
Dernière ligne adresse	Localité3 (global)	Localité2	Localité1
Pays (facultatif)	Localité2	Localité1	Région1
	Localité1	Région1	Code postal (global) ou Code postal1 (région Etats-Unis)
	Région1	Code postal (global) ou Code postal1 (région Etats-Unis)	Pays (facultatif)
	Code postal (global) ou Code postal1 (région Etats-Unis)	Pays (facultatif)	
	Pays (facultatif)		

27.5.4 Détermination des transformations à utiliser

Vous pouvez choisir entre plusieurs transformations Address Cleanse selon ce que vous désirez faire avec vos données. Il existe des transformations pour le nettoyage des données d'adresses globales et/ou des Etats-Unis, le nettoyage basé sur les réglementations USPS, utilisant des règles de gestion pour nettoyer les données et le nettoyage transactionnel des données d'adresses globales.

Informations associées

[Nettoyage des données d'adresses globales \[page 570\]](#)

[Nettoyage des données des Etats-Unis uniquement \[page 571\]](#)

[Nettoyage des données des Etats-Unis et des données globales \[page 572\]](#)

[Nettoyage des données d'adresses à l'aide de plusieurs règles de gestion \[page 572\]](#)

[Nettoyage transactionnel des données d'adresse \[page 572\]](#)

27.5.4.1 Nettoyage des données d'adresses globales

Pour nettoyer les données d'adresses de l'un des pays pris en charge par le progiciel, (y compris le Canada pour la certification SERP (Software Evaluation and Recognition Program) et l'Australie pour la certification AMAS (Address Matching Approval System)), utilisez la transformation Global Address Cleanse dans votre projet avec un ou plusieurs des moteurs suivants :

- Canada
- Adresse globale
- États-Unis

→ Conseil

Le nettoyage des données des États-Unis avec la transformation USA Regulatory Address Cleanse est habituellement plus rapide qu'avec la transformation Global Address Cleanse et le moteur États-Unis. Ce scénario est également vrai même si vous finissez par avoir besoin des deux transformations.

Vous pouvez également utiliser la transformation Global Address Cleanse avec les moteurs Canada, États-Unis et Adresse globale dans un flux de données en temps réel afin de créer des listes de propositions pour ces pays.

Commencer par un exemple de configuration de transformation

Le logiciel comprend plusieurs configurations du modèle de la transformation Global Address Cleanse (qui comprend au moins un moteur), vous pouvez les copier afin de les utiliser pour un projet.

Informations associées

[Pays pris en charge \(Global Address Cleanse\) \[page 591\]](#)

[Nettoyage des données des États-Unis et des données globales \[page 572\]](#)

27.5.4.2 Nettoyage des données des États-Unis uniquement

Pour nettoyer les données des États-Unis, utilisez la transformation USA Regulatory Address Cleanse afin d'obtenir de meilleurs résultats. Avec cette transformation, et lorsque DPV, LACSLink et SuiteLink sont activées, vous pouvez produire un publipostage certifié CASS et un USPS Form 3553. Si vous ne prévoyez pas de traiter des listes certifiées CASS, vous devriez quand même utiliser la transformation USA Regulatory Address Cleanse pour traiter vos données des États-Unis. L'utilisation de la transformation USA Regulatory Address Cleanse sur les données des États-Unis est plus efficace que l'utilisation de la transformation Global Address Cleanse.

Avec la transformation USA Regulatory Address Cleanse, vous pouvez ajouter des informations supplémentaires à vos données, par exemple DSF2, EWS, eLOT, NCOALink et RDI. Vous pouvez aussi traiter des enregistrements individuellement en utilisant des listes de propositions.

Commencer par un exemple de configuration de transformation

Le logiciel comprend plusieurs configurations du modèle de la transformation USA Regulatory Address Cleanse pouvant vous aider à configurer vos projets.

Informations associées

[Introduction aux listes de propositions \[page 658\]](#)

27.5.4.3 Nettoyage des données des Etats-Unis et des données globales

Que devez-vous faire lorsque vous avez dans une base de données des adresses des Etats-Unis devant être certifiées et des adresses d'autres pays ? Dans ce cas, utilisez les deux transformations Global Address Cleanse et USA Regulatory Address Cleanse dans votre flux de données.

→ Conseil

Même si vous ne traitez pas les données américaines pour la certification USPS, vous constaterez que leur nettoyage avec la transformation USA Regulatory Address Cleanse est plus rapide qu'avec la transformation Global Address Cleanse et le moteur Etats-Unis.

27.5.4.4 Nettoyage des données d'adresses à l'aide de plusieurs règles de gestion

Lorsque deux adresses ont une utilisation différente (par exemple, adresse de facturation et adresse d'expédition), utilisez deux transformations Address Cleanse identiques dans un flux de données.

Un ou deux moteurs ?

Lorsque vous utilisez deux transformations Global Address Cleanse pour les données du même pays, elles peuvent partager un moteur. Vous n'avez pas besoin d'avoir deux moteurs du même genre. Que vous utilisiez un ou deux moteurs, cela n'affecte pas le temps de traitement global du flux de données.

Cependant, dans ce cas, vous pouvez éventuellement utiliser deux moteurs distincts (même si les données sont du même pays). Selon vos règles de gestion, vous pouvez avoir à définir les paramètres du moteur différemment pour une adresse de facturation et une adresse d'expédition. Par exemple, dans le groupe Options de standardisation, l'option Langue de pays de sortie peut convertir les données utilisées dans chaque enregistrement en langue officielle du pays ou conserver la langue utilisée dans chaque enregistrement. Par exemple, vous pouvez convertir les données de l'adresse d'expédition mais conserver celles de l'adresse de facturation.

27.5.4.5 Nettoyage transactionnel des données d'adresse

La transformation Global Suggestion Lists, davantage utilisée dans les projets transactionnels, est une manière de remplir les adresses avec un minimum de données. Elle peut aussi faire des propositions pour les

correspondances possibles. Par exemple, les Iles Marshall et les Etats fédérés de Micronésie ont été récemment supprimés du répertoire des adresses des Etats-Unis. Par conséquent, si vous utilisiez auparavant le moteur Etats-Unis, vous devrez maintenant utiliser le moteur Adresse globale. La transformation Global Suggestion Lists permet d'indiquer que ces pays ne figurent plus dans le répertoire des adresses des Etats-Unis.

Ce système de saisie d'adresses facile est idéal dans les environnements de centres d'appels ou dans tout environnement transactionnel où un nettoyage est nécessaire au point d'entrée. C'est également un outil de recherche avantageux lorsque vous devez gérer les mauvaises adresses d'un précédent traitement par lots.

Placez la transformation Global Suggestion Lists après la transformation Country ID et avant une transformation Global Address Cleanse qui utilise un moteur Adresse globale, Canada et/ou Etats-Unis.

Intégration de fonctionnalité

La fonctionnalité Global Suggestion Lists est conçue pour être intégrée à vos applications personnalisées via le service Web. Si vous êtes programmeur à la recherche d'informations sur la manière d'intégrer cette fonctionnalité, voir "Intégrer Global Suggestion Lists" dans le *Guide d'intégration*

Commencer par un exemple de configuration de transformation

Data Quality comprend un modèle de transformation Global Suggestion Lists qui peut vous aider lors de la configuration d'un projet.

Informations associées

[Introduction aux listes de propositions \[page 658\]](#)

27.5.5 Identification du pays de destination

La transformation Global Address Cleanse inclut le traitement Country ID. Par conséquent, vous n'avez pas besoin de placer une transformation Country ID avant la transformation Global Address Cleanse dans votre flux de données.

Dans le groupe d'options Options Country ID de la transformation Global Address Cleanse, vous pouvez définir le pays de destination ou définir si vous désirez exécuter le traitement Country ID.

Pays constant

Si l'ensemble des données provient d'un seul pays, l'Australie par exemple, vous n'avez pas besoin d'exécuter le traitement Country ID ou d'entrer un champ de pays discret. Vous pouvez indiquer le pays à la transformation Global Address Cleanse et elle supposera que tous les enregistrements viennent de ce pays, ce qui peut économiser du temps de traitement.

Affectation par défaut

Vous pouvez exécuter le traitement Country ID si vous utilisez deux moteurs ou plus et que les adresses d'entrée contiennent des données de pays (comme le code ISO à deux caractères ou un nom de pays) ou si vous utilisez un seul moteur et que la source d'entrée contient plusieurs adresses ne pouvant pas être traitées par ce moteur. Les adresses ne pouvant pas être traitées ne sont pas envoyées au moteur. La transformation utilisera par défaut le pays indiqué dans ce groupe d'options.

Informations associées

[Définition d'une constante Pays \[page 593\]](#)

[Définition d'un pays par défaut \[page 592\]](#)

27.5.6 Configuration des fichiers de référence

La transformation USA Regulatory Address Cleanse, la transformation Global Address Cleanse et les moteurs s'appuient sur plusieurs répertoires (fichiers de référence) pour nettoyer les données.

Lorsque vous utilisez des répertoires pour standardiser des adresses et affecter des codes postaux, vous devez utiliser les répertoires les plus à jour. Vous devez par conséquent mettre à jour périodiquement vos fichiers avec de nouveaux répertoires. SAP vous indique les dates actuelles des répertoires et vous informe lorsque vous devez télécharger des répertoires plus récents.

Informations associées

[Au-delà du nettoyage d'adresse de base \[page 599\]](#)

[Pays pris en charge \(Global Address Cleanse\) \[page 591\]](#)

[Traitement des adresses japonaises \[page 577\]](#)

[Traitement des adresses chinoises \[page 587\]](#)

[Certification Nouvelle-Zélande \[page 594\]](#)

27.5.6.1 Répertoires

Pour corriger les adresses et affecter les codes, les transformations Address Cleanse se basent sur les répertoires d'adresses.

Ces répertoires contiennent des types d'informations spécifiques concernant les traitements que vous voulez effectuer sur vos adresses. Par exemple, le service postal américain dispose de répertoires contenant les informations postales de toutes les adresses situées aux États-Unis et leurs territoires. Le logiciel utilise également de nombreux autres répertoires émis par d'autres autorités postales telles que celles du Canada, de la Nouvelle-Zélande et du Japon.

i Remarque

Les transformations Address cleanse prennent en charge de nombreux pays. Toutefois, lorsque vous utilisez la transformation Global Address Cleanse, il est possible que certains pays ne soient pas pris en charge par les répertoires officiels.

En plus des répertoires d'adresses de base, il existe de nombreux répertoires spécialisés. Par exemple, la transformation USA Regulatory Address Cleanse utilise les répertoires spécialisés suivants :

- DPV®
- DSF2®
- Early Warning System (EWS)
- eLOT®
- GeoCensus
- LACSLink®
- NCOALink®
- RDI™
- SuiteLink™
- Z4Change

Ces répertoires spécialisés aident à étendre le nettoyage d'adresses américaines au-delà de l'analyse de base et de la standardisation.

27.5.6.2 Emplacement des fichiers de répertoire

Définissez les emplacements des fichiers de répertoire dans le groupe d'options *Fichiers de référence* de la transformation.

Votre administrateur système doit déjà avoir installé les fichiers nécessaires aux emplacements appropriés selon les besoins de votre société.

⚠ Attention

Les répertoires incompatibles ou obsolètes peuvent rendre le logiciel inutilisable. L'administrateur système doit installer des mises à jour hebdomadaires, mensuelles ou bimensuelles de répertoire pour la transformation USA Regulatory Address Cleanse ; des mises à jour mensuelles de répertoire pour les moteurs de l'Australie et du Canada ; des mises à jour trimestrielles de répertoire pour le moteur d'Adresse globale afin de garantir qu'ils soient compatibles avec le progiciel actuel.

27.5.6.3 Valeurs de substitution des fichiers de répertoire

Utilisez les valeurs de substitution pour définir une valeur constante pour vos emplacements de répertoire.

Si vous commencez avec un modèle de transformation, les options *Fichiers de référence* sont déjà remplies par défaut avec une variable de substitution. Par exemple, la variable `$$RefFilesAddressCleanse` dirige vers le dossier des données de référence du répertoire par défaut du logiciel.

i Remarque

Vous pouvez changer cet emplacement en modifiant le fichier de substitution associé au flux de données. Cette modification est appliquée pour chaque flux de données qui utilise ce fichier de substitution.

Informations associées

[Présentation des paramètres de substitution \[page 328\]](#)

27.5.6.4 Affichage de dates d'expiration de répertoire dans le journal de suivi

Étapes à suivre pour définir le journal de suivi avec des dates d'expiration de répertoire.

Pour inclure des informations relatives à l'expiration d'un répertoire dans le journal de suivi, exécutez les étapes suivantes.

1. Cliquez avec le bouton droit sur l'icône du job applicable dans Designer et sélectionnez *Exécuter*.
2. Dans la fenêtre *Propriétés d'exécution*, ouvrez l'onglet *Options d'exécution* (il doit être déjà ouvert par défaut).
3. Sélectionnez *Imprimer tous les messages de suivi*.

Informations associées

[Utilisation des journaux \[page 372\]](#)

27.5.7 Définition des options de standardisation

La standardisation modifie la manière dont sont présentées les données après une affectation. Le type de modification dépend des options définies dans la transformation. Ces options comprennent entre autres la casse, la ponctuation, la séquence et les abréviations. Cela aide à garantir l'intégrité de vos bases de données, facilite la remise du courrier et donne à vos communications avec les clients une apparence plus professionnelle.

Par exemple, l'adresse suivante a été standardisée au niveau des majuscules, de la ponctuation et de la phrase postale (itinéraire vers la route rurale).

Entrée	Sortie
Multiligne1 = route 1 box 44a	Ligne_adresse = RR 1 BOX 44A
Multiligne2 = stodard wisc	Localité1 = STODDARD
	Région1 = WI
	Code postal1 = 54658

Transformation Global Address Cleanse

Dans la transformation Global Address Cleanse, vous définissez les options de standardisation dans le groupe d'options Options de standardisation.

Vous pouvez standardiser les adresses pour tous les pays et/ou pour des pays individuels (selon vos données). Par exemple, vous pouvez avoir un ensemble d'options de standardisation françaises qui standardise uniquement les adresses en France et un autre ensemble d'options de standardisation globales qui standardise toutes les autres adresses.

Transformation USA Regulatory Address Cleanse

Si vous utilisez la transformation USA Regulatory Address Cleanse, vous définissez les options de standardisation options dans l'onglet [Options](#) de la section Options de standardisation.

27.5.8 Traitement des adresses japonaises

Le moteur d'Adresse globale de la transformation Global Address Cleanse analyse les adresses japonaises. Le premier objectif de cette transformation et de ce moteur est d'analyser et standardiser les adresses japonaises pour les applications de comparaison et nettoyage des données.

i Remarque

Le moteur Japon prend uniquement en charge les données kanji et katakana. Le moteur ne prend pas en charge les données en caractères latins.

Une partie importante de la fonctionnalité d'analyse des adresses repose sur la base de données d'adresses japonaises. Le progiciel renferme des données du Ministère de la gestion publique, des affaires intérieures, des postes et télécommunications (MPT) et d'autres sources de données. La base de données d'adresses avancée comprend une base de données gouvernementale régulièrement mise à jour, qui inclut les codes postaux régionaux mappés aux localités.

Informations associées

[Format d'adresse japonaise standard \[page 578\]](#)

[Formats spéciaux de la région Hokkaido \[page 584\]](#)

[Modèle d'adresse japonaise \[page 585\]](#)

27.5.8.1 Format d'adresse japonaise standard

Tableau comportant les composants d'adresse standard.

Une adresse japonaise habituelle comprend les composants suivants.

Composant d'adresse	Japonais	Anglais	Champ(s) de sortie
Code postal	〒 654-0153	654-0153	Code_postal_complet
Préfecture	兵庫県	Hyogo-ken	Région1_complète
Ville	神戸市	Kobe-shi	Localité1_complète
Quartier	須磨区	Suma-ku	Localité2_complète
District	南落合	Minami Ochiai	Localité3_complète
Numéro d'îlot	1 丁目	1 chome	Nom_principal_complet1
Numéro de sous-îlot	25 番地	25 banchi	Nom_principal_complet2
Numéro de rue	2 号	2 go	Numéro_principal_complet

Une adresse peut également comporter un nom de bâtiment, un numéro d'étage et un numéro de chambre.

Informations associées

[Composants d'adresse japonaise \[page 579\]](#)

27.5.8.1.1 Composants d'adresse japonaise

Descriptions de chaque composant d'adresse dans une adresse japonaise.

Composants d'adresse japonaise	Description
Code postal	Les codes postaux japonais sont au format <i>nnn-nnnn</i>. Les trois premiers chiffres représentent la zone. Les quatre derniers chiffres représentent une localisation dans la zone. Les localisations possibles sont district, sous-district, îlot, sous-îlot, bâtiment, étage et société. Les codes postaux doivent être écrits en chiffres arabes. Le symbole du bureau de poste 〒 est facultatif. Avant 1998, le code postal comprenait 3 ou 5 chiffres. Certaines anciennes bases de données peuvent encore refléter l'ancien système.
Préfecture	Les préfectures sont les régions. Le Japon comprend quarante-sept préfectures. Vous pouvez omettre la préfecture pour certaines villes bien connues.
Ville	Les noms des villes japonaises comportent le suffixe 市 (-shi). Dans certaines parties des régions de Tokyo et Osaka, les gens omettent le nom de la ville. Dans certains villages insulaires, ils utilisent le nom de l'île avec un suffixe 島 (-shima) à la place du nom de la ville. Dans certaines zones rurales, ils utilisent le nom du comté avec le suffixe 郡 (-gun) à la place du nom de la ville.
Quartier	Une ville est divisée en quartiers. Le nom du quartier comporte le suffixe 区 (-ku). Le composant quartier est omis pour les petites villes, les villages insulaires et les zones rurales qui n'ont pas de quartiers.
District	Un quartier est divisé en districts. Lorsqu'il n'existe pas de quartiers, la petite ville, le village insulaire ou la zone rurale est divisée en districts. Le nom du district peut comporter le suffixe 町 (-cho/-machi), mais il est souvent omis. 町 a deux prononciations possibles, mais une seule est correcte pour un district particulier. Dans les très petits villages, les gens utilisent le nom du village avec le suffixe 村 (-mura) à la place du district. Lorsqu'un village ou un district est sur une île ayant le même nom, le nom de l'île est souvent omis.

Composants d'adresse japonaise	Description
Sous-district	Principalement dans les zones rurales, un district peut être divisé en sous-districts, marqués par le préfixe 字(aza-). De plus, un sous-district peut être divisé en sous-districts marqués par le préfixe 小字(koaza-), signifiant petit aza. koaza peut être abrégé en aza. Un sous-district peut aussi être marqué du préfixe 大字(oaza-), qui signifie grand aza. Oaza peut également être abrégé en aza. Voilà les combinaisons possibles : • oaza • aza • oaza et aza • aza et koaza • oaza et koaza i Remarque Les caractères 大字(oaza-), 字(aza-) et 小字(koaza-) sont souvent omis.
Parcelle de sous-districts	Un sous-district aza peut être divisé en parcelles de sous-district numérotées qui sont identifiées par le suffixe 部(-bu), signifiant <i>pièce</i>. Le caractère 部 est souvent omis. Les parcelles peuvent être numérotées de plusieurs manières : • Chiffres arabes (1, 2, 3, 4, etc.) 石川県七尾市松百町 8 部 3 番地 1 号 • Lettres katakana dans l'ordre iroha (イ, 口, ハ, ニ, etc.) 石川県小松市里川町ナ部 23 番地 • Chiffres kanji, ce qui est très rare (甲, 乙, 丙, 丁, etc.) 愛媛県北条市上難波甲部 311 番地
Sous-division	Un district ou sous-district rural (oaza/aza/koaza) est parfois divisé en sous-divisions, marquées par le suffixe 地割(-chiwari), qui signifie division de terrain. Le préfixe facultatif est 第(dai-). Les exemples d'adresses suivants illustrent les sous-divisions : 岩手県久慈市旭町 10 地割 1 番地 岩手県久慈市旭町第 10 地割 1 番地

Composants d'adresse japonaise	Description
Numéro d'îlot	Un district est divisé en îlots. Le numéro d'îlot comprend le suffixe 丁目 (-chome). Les districts comprennent en général de 1 à 5 îlots, mais ils peuvent en contenir davantage. Le numéro d'îlot peut être écrit avec un chiffre kanji. Les adresses japonaises ne comprennent pas de nom de rue. 東京都渋谷区道玄坂 2 丁目 2 5 番地 1 2 号 東京都渋谷区道玄坂二丁目 2 5 番地 1 2 号
Numéro de sous-îlot	Un îlot est divisé en sous-îlots. Le nom de sous-îlot inclut le suffixe 番地 (-banchi), qui signifie "terrain numéroté". Le suffixe 番地 (-banchi) peut être abrégé en 番 (-ban).
Numéro de rue	Chaque maison possède un numéro unique. Le numéro de rue comprend le suffixe 号 (-go), qui signifie numéro.
Variations des numéros d'îlot, de sous-îlot et de maison	Les données des numéros d'îlot, de sous-îlot et de maison peuvent varier.
Tirets	Les marqueurs de suffixe 丁目(chome), 番地 (banchi) et 号 (go) peuvent être remplacés par des tirets. 東京都文京区湯島 2 丁目 18 番地 12 号 東京都文京区湯島 2-18-12 Parfois, les numéros d'îlot, de sous-îlot et de maison sont combinés ou omis. 東京都文京区湯島 2 丁目 18 番 12 号 東京都文京区湯島 2 丁目 18 番地 12 東京都文京区湯島 2 丁目 18-12
Aucun numéro d'îlot	Parfois, le numéro d'îlot est omis. Par exemple, ce quartier de Tokyo a des districts numérotés et aucun numéro d'îlot n'est inclus. 二番町 signifie district numéro 2. 東京都千代田区 二番町 9 番地 6 号
Noms de bâtiment	Les noms des appartements ou des bâtiments sont souvent inclus après le numéro de maison. Lorsqu'un nom de bâtiment comprend le nom du district, ce dernier est souvent omis. Lorsqu'un bâtiment est bien connu, les numéros d'îlot, de sous-îlot et de maison sont souvent omis. Lorsqu'un nom de bâtiment est long, il peut être abrégé ou écrit à l'aide de son acronyme en lettres anglaises. Pour une description des attributs, voir Suffixes de noms de bâtiment communs [page 582].

Composants d'adresse japonaise	Description
Numéros de bâtiment	Les numéros de chambre, d'appartement, etc. suivent le nom du bâtiment. Les numéros de bâtiment peuvent comprendre le suffixe 号室 (-goshitsu). Les numéros d'étage au-dessus du niveau du sol peuvent comprendre le suffixe 階 (-kai) ou la lettre F. Les numéros d'étage en dessous du niveau du sol peuvent comprendre le suffixe 地下<n>階 (chika <n> kai) ou les lettres B <n>F (où <n> représente le numéro d'étage). Une résidence peut comprendre plusieurs bâtiments nommés Bâtiment A, Bâtiment B, etc., marqués par le suffixe 棟 (-tou). Pour voir des exemples d'adresse avec des numéros de bâtiment, voir Exemples de numéros de bâtiment japonais [page 583].

Informations associées

[Format d'adresse japonaise standard \[page 578\]](#)

27.5.8.1.2 Suffixes de noms de bâtiment communs

Tableau qui contient les suffixes de noms de bâtiment japonais communs.

Dans les adresses japonaises, les noms des appartements ou des bâtiments sont souvent inclus après le numéro de maison. Lorsqu'un nom de bâtiment comprend le nom du district, ce dernier est souvent omis. Lorsqu'un bâtiment est bien connu, les numéros d'îlot, de sous-îlot et de maison sont souvent omis. Lorsqu'un nom de bâtiment est long, il peut être abrégé ou écrit à l'aide de son acronyme en lettres anglaises.

Suffixe	Écriture latine	Traduction
ビルディング	birudingu	bâtiment
ビルヂング	birudingu	bâtiment
ビル	biru	bâtiment
センター	senta-	centre
プラザ	puraza	place
パーク	pa-ku	parc
タワー	tawa-	tour
会館	kaikan	entrée

Suffixe	Ecriture latine	Traduction
棟	tou	bâtiment (unité)
庁舎	chousha	bâtiment de bureau gouvernemental
マンション	manshon	condominium
団地	danchi	résidence
アパート	apa-to	appartement
荘	sou	villa
住宅	juutaku	logement
社宅	shataku	logement de fonction
官舎	kansha	domicile officiel

27.5.8.1.3 Exemples de numéros de bâtiment japonais

Tableau qui contient des exemples d'adresses japonaises comprenant des numéros de bâtiment.

Zone du bâtiment	Exemple
Premier étage au-dessus du sol.	東京都千代田区二番町 9 番地 6 号 バウエプタ 3 F
Deuxième étage en sous-sol.	東京都渋谷区道玄坂 2-25-12 シティバンク地下 2 階
Bâtiment A Chambre 301	兵庫県神戸市須磨区南落合 1-25-10 須磨パークヒルズ A 棟 301 号室
Bâtiment A Chambre 301	兵庫県神戸市須磨区南落合 1-25-10 須磨パークヒルズ A-301

Informations associées

[Composants d'adresse japonaise \[page 579\]](#)

27.5.8.2 Formats spéciaux de la région Hokkaido

La région d'Hokkaido a deux formats d'adresse spéciaux.

Format régional d'Hokkaido

Format	Description
Super-îlot	Un format spécial de super-îlot existe uniquement dans la préfecture d'Hokkaido. Un super-îlot, marqué du suffixe 条 (-joh), est d'un niveau plus important que l'îlot. Le numéro de super-îlot ou d'îlot peut contenir une direction 北 (Nord), 南 (Sud), 東 (Est) ou 西 (Ouest). L'exemple d'adresse suivant illustre un super-îlot 4 Joh. 北海道札幌市西区二十四軒 4 条 4 丁目 1 3 番地 7 号
Sous-districts numérotés	Un autre format régional d'Hokkaido est le sous-district numéroté. Un nom de sous-district peut être marqué du suffixe 線 (-sen) signifiant numéro au lieu du suffixe 字 (-aza). Lorsqu'un sous-district comporte un suffixe 線, l'îlot peut avoir le suffixe 号 (-go) et le numéro de maison ne comporte pas de suffixe. L'adresse ci-dessous contient le sous-district 4 sen, puis un îlot numéroté 5 go. 北海道旭川市西神楽 4 線 5 号 3 番地 1 1

27.5.8.3 Autres formats spéciaux d'adresse japonaise

Tableau qui contient les autres formats spéciaux d'adresse pour les adresses japonaises.

Format	Description
Orthographe acceptée	Les noms de villes, districts, etc., peuvent avoir plusieurs orthographes acceptés parce que plusieurs manières d'écrire certains sons en Japonais sont acceptées.
Numérotation acceptée	Lorsque l'îlot, le sous-îlot, la maison ou le district comporte un numéro, celui-ci peut être écrit en chiffres arabes ou kanji. Par exemple, 二番町 signifie numéro de district 2 et dans l'exemple suivant, il correspond à Niban-cho. 東京都千代田区二番町九番地六号

Format	Description
Adresses de boîte postale	Les adresses de boîte postale contiennent le code postal, Localité1, la préfecture, le nom du bureau de poste, le marqueur de boîte et le numéro de boîte. i Remarque La transformation Global Address Cleanse reconnaît les adresses de boîte postale se trouvant uniquement dans la base de données LOPC (Large Organization Postal Code). L'adresse peut avoir l'un des formats suivants : • Préfecture, Localité1, nom de bureau de poste, marqueur de boîte (私書箱) et numéro de boîte postale. • Code postal, préfecture, Localité1, nom de bureau de poste, marqueur de boîte (私書箱) et numéro de boîte postale. L'exemple d'adresse suivant illustre une adresse de boîte postale : Bureau de poste d'Osaka marqueur de boîte n°1 大阪府大阪市大阪支店私書箱 1 号
Format LOPC (Large Organization Postal Code)	Le service postal peut affecter un code postal unique à une organisation importante, par exemple le service client d'une entreprise majeure. Une organisation peut avoir jusqu'à deux codes postaux uniques selon le volume de courrier qu'elle reçoit. L'adresse peut avoir l'un des formats suivants : • Adresse, nom de société • Code postal, adresse, nom de société L'exemple suivant est une adresse au format LOPC. 100-8798 東京都千代田区霞が関 1 丁目 3 - 2 日本郵政 株式会社

27.5.8.4 Modèle d'adresse japonaise

Cette adresse a été traitée par la transformation Global Address Cleanse et le moteur Global Address.

Entrée

0018521 北海道札幌市北区北十条西 1 丁目 12 番地 3 号創生ビル 1 階 101 号室札幌私書箱センター

Champs de ligne d'adresse

Nom_principal1	1
Type_principal1	丁目
Nom_principal2	12
Type_principal2	番地
Numéro_principal	3
Description_numéro_principal	号
Nom_immeuble1	創生ビル
Numéro_étage	1
Description_étage	階
Numéro_unité	101
Description_unité	号室
Adresse_principale	1 丁目 12 番地 3 号
Adresse_secondaire	創生ビル 1 階 101 号室
Adresse_principale_secondaire	1 丁目 12 番地 3 号 創生ビル 1 階 101 号室

Champs de la dernière ligne

Pays	日本
Code_ISO_pays_3Chiffres	392
Code_ISO_pays_2Car	JP
Code postal1	001
Code postal2	8521
CodePostal_Complet	001-8521
Région1	北海
Description_région1	道
Nom_localité1	札幌
Description_localité1	市

Champs de la dernière ligne

Nom_localité2	北
Description_localité2	区
Nom_localité3	北十条西
Dernière ligne	001-8521 北海道 札幌市 北区 北十条西

Société

Société	札幌私書箱センター

Champs non analysés

Code_Statut	S0000
Type_Affectation	F
Type_Adresse	S

27.5.9 Traitement des adresses chinoises

Le moteur d'adresses globales de la transformation Global Address Cleanse analyse les adresses chinoises. Le premier objectif de cette transformation et de ce moteur est d'analyser et de normaliser les adresses pour la correspondances des données et les applications de nettoyage.

27.5.9.1 Format d'adresse chinoise

Les adresses chinoises commencent par le code postal, suivi par la région administrative la plus grande (par exemple, la province) et continuent jusqu'à la plus petite unité (par exemple, le numéro de la chambre et le destinataire).

Lorsque les personnes envoient des courriers entre deux préfectures chinoises, elles incluent souvent la région administrative la plus grande dans l'adresse. Les adresses contiennent des informations détaillées sur l'endroit où le courrier sera livré. Les bâtiments le long des rues sont numérotés de façon séquentielle, parfois avec des nombres pairs d'un côté, et des nombres impairs de l'autre. Quelquefois, les nombres pairs et impairs se trouvent du même côté de la rue.

Le tableau suivant décrit les formats d'adresse chinoise.

Format	Description
Code Postal	Un numéro à 6 chiffres qui identifie le point de livraison cible de l'adresse. Il comporte souvent le préfixe 邮编.
Pays	Le nom complet de la Chine est 中华人民共和国 (République populaire de Chine, ou RPC). Pour le courrier livré en Chine, les adresses domestiques omettent souvent le nom de pays de l'adresse cible.
Province	Équivalent d'un État aux États-Unis. La Chine a 34 divisions de niveau provincial, dont : • Des provinces (省 shěng) • Des régions autonomes (自治区 zìzhìqū) • Des municipalités (直辖市 zhíxíashì) • Des régions administratives spéciales (特别行政区 tèbié xíngzhèngqū)
Préfecture	Les divisions de niveau préfectoral sont le deuxième niveau de la structure administrative, et incluent : • Les préfectures (地区 dìqū) • Les préfectures autonomes (自治州 zìzhìzhōu) • Les villes-préfectures (地级市 dìjīshì) • Les ligues (盟 méng)
District	Une sous-division de Préfecture qui comprend : • Les districts (县 xiàn) • Les districts autonomes (自治县 zìzhìxiàn) • Les villes-districts (县级市 xiànjíshì) • Les districts urbains (市辖区 shìxiáqū) • Les bannières (旗 qí) • Les bannières autonomes (自治旗 zìzhìqí) • Les zones forestières (林区 línqū) • Les districts spéciaux (特区 tèqū)
Canton	La division de niveau cantonal qui inclut : • Les cantons (乡 xiāng) • Les cantons ethniques (民族乡 mínzúxiāng) • Les bourgs ((镇 zhèn) • Les sous-districts (街道办事处 jiēdào bànshìchù) • Les offices publics de district (区公所 qūgōngsuǒ) • Les sumu (苏木 sūmù) • Les sumu ethniques (民族苏木 mínzúsūmù)
Village	Les villages incluent : • Les communautés résidentielles (社区居民委员会 jūmín-wēiyuánhùi) • Les quartiers ou communautés (社区) • Les villages (村民委员会 cūnmínwēiyuánhùi) ou Les groupes de village (村民小组 cūnmínxiǎozǔ) • Les villages administratifs (行政村 xíngzhèngcūn)

Format	Description
Informations de rue	Spécifient le point de livraison où le destinataire peut être trouvé. En Chine, les informations sur la rue ont souvent la forme Nom de rue > Numéro de rue. Par exemple, 上海市浦东新区晨晖路 1001 号 - Nom de rue : le nom de rue est généralement suivi par l'un de ces suffixes 路, 大道, 街, 大街,, etc. - Numéro de rue : Le numéro de rue est suivi par le suffixe 号. Le numéro de rue est un numéro unique dans une rue.
Communauté résidentielle	Peut être utilisé pour la livraison de courrier, particulièrement pour certaines communautés résidentielles connues dans les grandes villes. Le nom de rue et le numéro de rue peuvent être omis. La communauté résidentielle n'a pas un standard d'appellation et ne doit pas forcément être suivie d'un marqueur typique. Cependant, elle est souvent suivie par les suffixes typiques tels que 新村, 小区,, etc.
Nom de bâtiment	Le bâtiment est souvent suivi par le marqueur de bâtiment, tel que 大厦, 大楼, bien qu'il ne soit pas strictement exigé (par exemple 中华大厦). Le nom du bâtiment dans les communautés résidentielles est souvent représenté par un nombre avec un suffixe 号, 幢. Exemple : 上海市浦东新区晨晖路 100 弄 10 号 101 室.
Adresse métropolitaine commune	Inclut le nom du district, qui est commun pour les zones métropolitaines dans les grandes villes.
Adresse rurale	Inclut le nom du village, qui est commun pour les adresses rurales.

Informations associées

[Composants d'adresses métropolitaines communes \[page 589\]](#)

[Composants d'adresse rurale chinoise \[page 590\]](#)

27.5.9.1.1 Composants d'adresses métropolitaines communes

Tableau des composants d'adresses métropolitaines communes.

Composant d'adresse	Chinois	Anglais	Champ de sortie
Code postal	510030	510030	Code_postal_complet
Pays	中国	Chine	Pays

Composant d'adresse	Chinois	Anglais	Champ de sortie
Province	广东省	Province du Guang-dong	Région1_complet
Nom de la ville	广州市	Guangzhou	Localité1_complet
Nom du district	越秀区	Yuexiu	Localité2_complet
Nom de la rue	西湖路	Xihu Road	Nom_principal_complet1
Numéro de rue	99 号	No. 99	Numéro_principal_complet

27.5.9.1.2 Composants d'adresse rurale chinoise

Tableau des composants d'adresse rurale chinoise

Composant d'adresse	Chinois	Anglais	Champ de sortie
Code postal	5111316	5111316	Code_postal_complet
Pays	中国	Chine	Pays
Province	广东省	Province du Guang-dong	Région1_complet
Nom de la ville	广州市	Guangzhou	Localité1_complet
Ville du niveau du district	增城市	Zengcheng	Localité2_complet
Nom de la ville	荔城镇	Licheng Town	Localité3_complet
Nom du village	联益村	Lianyi	Localité4_complet
Nom de la rue	光大路	Guangda Road	Nom_principal_complet1
Numéro de rue	99 号	No. 99	Numéro_principal_complet

27.5.9.2 Exemple d'adresse chinoise

Cette adresse a été traitée par la transformation Global Address Cleanse et le moteur Global Address.

Entrée

510830 广东省广州市花都区赤坭镇广源路 1 号星辰大厦 8 层 809 室

Champs de la ligne d'adresse

Nom_principal1	广源
Type_principal1	路
Numéro_Principal	1
Description_numéro_principal	号
Bâtiment_nom1	星辰大厦
Numéro_étage	8
Description_étage	层
Numéro_unité	809
Description_unité	室
Adresse_principale	广源路 1 号
Adresse_Secondaire	星辰大厦 8 层 809 室
Adresse_secondaire_principale	广源路 1 号星辰大厦 8 层 809 室

Champs de la dernière ligne d'adresse

Pays	中国
Code_postal_complet	510168
Région1	广东
Description_région1	省
Nom_localité1	广州
Description_localité1	市
Nom_localité2	花都
Description_localité2	区
Nom_Localité3	赤坭
Description_localité3	镇
Dernière ligne adresse	510830 广东省广州市花都区赤坭镇

Champs non analysés

Code_Statut	S0000
Type_affectation	S
Type_adresse	S

27.5.10 Pays pris en charge (Global Address Cleanse)

Plusieurs pays sont pris en charge par la transformation Global Address Cleanse. Le niveau de correction varie selon le pays et le moteur utilisé. La couverture complète de toutes les adresses d'un pays n'est pas garantie.

Pour le moteur d'Adresse globale, la prise en charge des pays dépend des ensembles de répertoires postaux achetés.

Pour le Japon, le niveau d'affectation dépend des données fournies par le Ministère de la gestion publique, des affaires intérieures, des postes et télécommunications (MPT)

Lors du traitement Country ID, la transformation peut identifier plusieurs pays. Toutefois, les moteurs de la transformation Global Address Cleanse peuvent ne pas fournir la correction d'adresses pour l'ensemble de ces pays.

Informations associées

[Modification du nom par défaut de pays de sortie \[page 592\]](#)

27.5.10.1 Modification du nom par défaut de pays de sortie

Lorsque vous utilisez le moteur des Etats-Unis pour traiter des adresses des Samoa Américaines, de Guam, des îles Mariannes du Nord, de Palau, de Porto Rico et des Iles vierges américaines, la région de sortie est AS, GU, MP, PW, PR ou VI, respectivement. Les Etats-Unis (EU) constituent cependant le pays de sortie.

Si vous ne souhaitez pas les Etats-Unis comme pays de sortie lors du traitement des adresses avec le moteur des Etats-Unis, définissez l'option *Utiliser le nom de pays postal* sur *Non*.

Ces étapes indiquent comment activer l'option Utiliser le nom de pays postal, dans la transformation Global Address Cleanse.

1. Ouvrez la transformation *Global Address Cleanse*.
2. Dans l'onglet *Options*, développez ► *Options de standardisation* ► *Pays* ► *Options* ►.
3. Pour l'option *Utiliser le nom de pays postal*, sélectionnez *Non*.

Informations associées

[Pays pris en charge \(Global Address Cleanse\) \[page 591\]](#)

27.5.10.2 Définition d'un pays par défaut

i Remarque

Exécutez un traitement d'ID de pays uniquement dans les cas suivants :

- Vous utilisez deux ou plusieurs moteurs et vos adresses d'entrée contiennent des données de pays (par exemple, le code ISO à deux caractères ou une désignation du pays).

- Vous utilisez un moteur mais vos données d'entrée contiennent des adresses de plusieurs pays.

1. Ouvrez la transformation *Global Address Cleanse*.
2. Dans l'onglet *Options*, développez *Options d'ID de pays* puis l'option de *Mode d'ID de pays*, sélectionnez *Affecter*.
Cette valeur indique à la transformation d'utiliser l'*ID de pays* pour affecter le pays. Si l'*ID de pays* ne peut pas affecter le pays, il utilisera la valeur figurant dans *Désignation du pays*.
3. Pour l'option *Désignation du pays*, sélectionnez le pays que vous souhaitez utiliser comme pays par défaut.
La transformation utilisera ce pays uniquement si l'*ID de pays* ne peut pas affecter un pays. Si vous ne souhaitez pas utiliser de pays par défaut, sélectionnez *Aucun*.
4. Pour l'option *Code de script*, sélectionnez le type de code de script qui représente vos données.
L'option *LATN* indique un code de script pour la plupart des types de données. Cependant, si vous traitez les données relatives au Japon, sélectionnez *KANA*.

Informations associées

[Identification du pays de destination \[page 573\]](#)

[Définition d'une constante Pays \[page 593\]](#)

27.5.10.3 Définition d'une constante Pays

1. Ouvrez la transformation *Global Address Cleanse*.
2. Dans l'onglet *Options*, développez *Options d'ID pays* puis pour l'option *Mode d'ID de pays*, sélectionnez *Constante*.
Cette valeur indique à la transformation de prendre les informations relatives au pays à partir des options *Désignation du pays* et *Code de script* (au lieu d'exécuter le traitement ID de pays).
3. Pour l'option *Désignation du pays*, sélectionnez le pays qui représente toutes vos données d'entrée.
4. Pour l'option *Code de script*, sélectionnez le type de code de script qui représente vos données.
L'option *LATN* indique un code de script pour la plupart des types de données. Cependant, si vous traitez les données relatives au Japon, sélectionnez *KANA*.

Informations associées

[Identification du pays de destination \[page 573\]](#)

[Définition d'un pays par défaut \[page 592\]](#)

27.5.11 Certification Nouvelle-Zélande

La certification Nouvelle-Zélande permet de traiter les adresses de Nouvelle-Zélande et de les qualifier pour les remises sur les publipostages de la poste néo-zélandaise.

27.5.11.1 Activer la certification Nouvelle-Zélande

Vous devez acheter les données du répertoire de Nouvelle-Zélande et demander un numéro de client à la poste néo-zélandaise pour pouvoir utiliser l'option de certification Nouvelle-Zélande.

Pour traiter les adresses de Nouvelle-Zélande qui qualifient pour les remises sur les publipostages :

1. Dans la transformation Global Address Cleanse, activez ► [Rapport et analyse](#) ► [Générer les données du rapport](#) ►.
2. Dans la transformation Global Address Cleanse, définissez ► [Options de pays](#) ► [Désactiver la certification](#) ► sur [Non](#).

i Remarque

Le progiciel ne produit pas le rapport SOA (Statement of Accuracy) de Nouvelle-Zélande quand cette option est définie sur [Oui](#).

3. Dans la transformation Global Address, renseignez toutes les options applicables du sous-groupe ► [Adresse globale](#) ► [Options de rapport](#) ► [Nouvelle-Zélande](#) ►.
4. Dans la transformation Global Address Cleanse, définissez ► [Moteurs](#) ► [Adresse globale](#) ► sur [Oui](#).

Après avoir exécuté le job et produit le rapport SOA (Statement of Accuracy) de Nouvelle-Zélande, vous devez renommer ce dernier ainsi que le journal de production SOA de Nouvelle-Zélande avant d'envoyer votre publipostage.

Informations associées

[Rapport Statement of Accuracy de Nouvelle-Zélande \[page 594\]](#)

27.5.11.2 Rapport Statement of Accuracy de Nouvelle-Zélande

Le rapport SOA (Statement of Accuracy) de Nouvelle-Zélande comprend des informations statistiques sur le nettoyage d'adresses de Nouvelle-Zélande.

27.5.11.3 Journal de production SOA de Nouvelle-Zélande

Le journal de production est un fichier texte ASCII délimité par des barres verticales (avec un enregistrement d'en-tête).

Le journal de production est identique au rapport SOA (Statement of Accuracy) de Nouvelle-Zélande sauf qu'il se trouve dans un fichier texte ASCII délimité par des barres verticales (avec un enregistrement d'en-tête). Le logiciel crée le journal de production SOA en extrayant des données de la table Sendrightaddraccuracy dans le référentiel. Le logiciel ajoute un nouvel enregistrement à la table Sendrightaddraccuracy chaque fois qu'un fichier est traité avec l'option DÉACTIVER_CERTIFICATION définie sur *Non*. Si l'option DÉACTIVER_CERTIFICATION est définie sur *Oui*, le logiciel ne produit pas le rapport SOA et aucune entrée n'est ajoutée à la table Sendrightaddraccuracy.

Les expéditeurs de publipostage **doivent** conserver le fichier journal de production pendant au moins 2 ans.

L'emplacement par défaut du journal de production SOA est `<REP_INSTALL>\Data Services
\DataQuality\certifications\CertificationLogs.`

27.5.11.4 Exigences concernant le publipostage

Le rapport et le journal de production SOA ne sont requis que si vous envoyez les données traitées pour un publipostage et désirez bénéficier de remises d'affranchissement.

- Envoyez le journal de production SOA au moins une fois par mois.
- Envoyez un rapport SOA pour chaque fichier traité pour les remises sur publipostages.

Informations associées

[Certification Nouvelle-Zélande \[page 594\]](#)

[Rapport Statement of Accuracy de Nouvelle-Zélande \[page 594\]](#)

[Journal de production SOA de Nouvelle-Zélande \[page 595\]](#)

27.5.11.5 Modification du modèle de certification de Nouvelle-Zélande

Suivez les étapes indiquées ci-dessous pour modifier le modèle de certification de Nouvelle-Zélande.

1. Importez le fichier `nz_sendright_certification.atl` situé dans le dossier `DataQuality\certifications` dans lequel vous avez installé le logiciel.

L'emplacement par défaut est `<REP_INSTALL>\Data Services\DataQuality\certifications.`

L'importation ajoute les objets suivants au référentiel.

- Le projet `DataQualityCertifications`

- Le job Job_DqBatchNewZealand_SOAProductionLog
 - Le flux de données DF_DqBatchNewZealand_SOAProductionLog
 - La banque de données DataQualityCertifications
 - Le format de fichier DqNewZealandSOAProductionLog
2. Modifiez l'objet de la banque de données DataQualityCertifications en suivant les étapes indiquées dans [Modification de la banque de données \[page 596\]](#).
 3. Conservez l'emplacement par défaut pour le journal de production SOA ou utilisez un autre emplacement en modifiant la configuration du paramètre de substitution :
 - a. Dans Designer, sélectionnez ► **Outils** ► *Configurations des paramètres de substitution* ►.
 - b. Remplacez le chemin d'accès dans Configuration1 pour le paramètre de substitution \$
\$CertificationLogPath par l'emplacement de votre choix.

L'emplacement par défaut du journal de production SOA est <REP_INSTALL>\Data Services
\DataQuality\certifications\CertificationLogs.
 4. Exécutez le job Job_DqBatchNewZealand_SOAProductionLog.
Le job crée un journal de production SOA nommé SOAPerc_SOAExpDate_SOAIId.txt à l'emplacement par défaut ou à celui indiqué dans la configuration du paramètre de substitution.
 5. Renommez le fichier SOAPerc_SOAExpDate_SOAIId.txt en utilisant les données du dernier enregistrement dans le fichier journal et le format de nommage de fichier décrit dans la section [Rapport Statement of Accuracy de Nouvelle-Zélande \[page 594\]](#).

Informations associées

[Rapport Statement of Accuracy de Nouvelle-Zélande \[page 594\]](#)

27.5.11.6 Modification de la banque de données

Après avoir téléchargé le fichier modèle .zip vers le dossier approprié, l'avoir décompressé et avoir importé le fichier .atl dans le logiciel, vous devez modifier la banque de données DataQualityCertifications.

Pour modifier la banque de données :

1. Sélectionnez l'onglet *Banques de données* de la bibliothèque d'objets locale, faites un clic droit sur DataQualityCertifications et sélectionnez *Modifier*.
2. Cliquez sur *Avancé* pour développer la fenêtre Modifier la banque de données DataQualityCertifications.

i Remarque

Passez l'étape 3 si vous avez Microsoft SQL Server 2000 or 2005 comme type de base de données de la banque de données.

3. Cliquez sur *Modifier*.
4. Trouvez la colonne pour votre type de base de données, modifiez la *Configuration par défaut* par **Oui**, et cliquez sur *OK*.

i Remarque

Si vous utilisez une version d'Oracle autre que Oracle 9i; effectuez les sous-étapes suivantes :

- a. Dans la barre d'outils, cliquez sur [Créer une configuration](#).
 - b. Saisissez vos informations, dont la version de la base de données Oracle que vous utilisez, puis cliquez sur [OK](#).
 - c. Cliquez sur [Fermer](#) sur la fenêtre Nouvelles valeurs ajoutées - Objets modifiés.
 - d. Dans la nouvelle colonne qui apparaît à droite des colonnes précédentes, sélectionnez [Oui](#) pour la [Configuration par défaut](#).
 - e. Saisissez les informations pour les options [Nom de la connexion à la base de données](#), [Nom de l'utilisateur](#) et [Mot de passe](#).
 - f. Dans [DBO](#), saisissez le nom du schéma.
 - g. Dans [Code Page](#), sélectionnez [cp1252](#) puis cliquez sur [OK](#).
5. Sur la fenêtre Modifier la banque de données DataQualityCertifications, saisissez vos informations de connexion au référentiel à la place des valeurs CHANGE_THIS. (Il est possible que vous deviez modifier trois ou quatre options, en fonction du type de référentiel).
6. Développez le groupe [Alias](#) et saisissez votre nom de propriétaire à la place de la valeur CHANGE_THIS. Si vous utilisez Microsoft SQL Server, définissez cette valeur sur [DBO](#).
7. Cliquez sur [OK](#).
- Si la fenêtre se ferme sans aucun message d'erreur, alors la base de données est correctement connectée.

27.5.12 Option de la liste de propositions de Global Address Cleanse

L'option de traitement de la liste de propositions de la transformation Global Address Cleanse est utilisée dans les projets opérationnels pour compléter et remplir les adresses qui ont des données minimales. Les listes de propositions peuvent fournir des propositions pour les correspondances possibles si aucune correspondance exacte n'est trouvée.

Cette option est avantageuse dans les situations où vous souhaitez extraire les adresses pas entièrement affectées par un processus automatisé et effectuer dans le système une recherche de liste des correspondances possibles. En fonction de l'adresse d'entrée donnée, la transformation Global Address Cleanse effectue une recherche tolérante aux erreurs dans le répertoire d'adresses et renvoie une liste des correspondances possibles. A partir de la liste de propositions renvoyée, vous pouvez sélectionner la proposition correcte et mettre à jour la base de données en conséquence.

i Remarque

Aucune certification avec les listes de propositions : Si vous utilisez le moteur Canada ou le moteur d'adresse globale pour l'Australie et la Nouvelle-Zélande, il est impossible de certifier votre publipostage pour la certification SERP, AMAS ou de Nouvelle-Zélande.

Commencer par un modèle de transformation

Si vous souhaitez utiliser la fonctionnalité des listes de propositions, le mieux est de commencer avec les exemples de transformations qui sont configurés à cette fin. Le modèle de transformation GlobalSuggestions_AddressCleanse est configuré pour nettoyer les données d'adresses de type Latin-1 dans tous les pays pris en charge à l'aide de la fonctionnalité de liste de propositions.

Informations associées

[Extraction de chaînes XML de qualité des données à l'aide de la fonction extract_from_xml \[page 295\]](#)

27.5.13 Transformation Global Suggestion List

La transformation Global Suggestion List vous permet d'effectuer des requêtes sur les adresses avec un minimum de données (utilisation de caractères génériques) et peut fournir une liste de propositions pour les correspondances possibles.

C'est un outil précieux dans un environnement de centre d'appels où les opérateurs doivent effectuer un minimum de saisie (nombre de frappes de touches) pour rechercher l'adresse de livraison de l'appelant. Par exemple, si l'opérateur est en ligne avec un appelant du Royaume-Uni, l'application propose le code postal et la plage d'adresses. Global Suggestion List est utilisée pour rechercher l'adresse avec une saisie rapide.

La transformation Global Suggestion List requiert l'entrée du code pays ISO à deux caractères. Par conséquent, vous pouvez éventuellement souhaiter placer une transformation comme Country ID qui sort le champ Code_Pays_ISO_2Car avant la transformation Global Suggestion Lists. La transformation Global Suggestion List est disponible pour les moteurs Canada, Adresse globale et Etats-Unis.

i Remarque

Aucune certification avec les listes de propositions : Si vous utilisez les moteurs Canada, Adresse globale ou Etats-Unis pour l'Australie et la Nouvelle-Zélande, vous ne pouvez pas certifier votre publipostage pour la certification SERP, CASS, AMAS ou Nouvelle-Zélande.

i Remarque

Cette option ne prend pas en charge le traitement des données d'adresses japonaises ou chinoises.

Commencer par un modèle de transformation

Pour utiliser la transformation Global Suggestion List, il vaut mieux commencer par l'un des modèles de transformation configuré spécifiquement. Les modèles de transformation suivants sont disponibles.

Modèle de transformation	Description
GlobalSuggestions	Modèle de transformation configuré pour générer une liste de propositions pour les données d'adresses Latin-1 dans tous les pays pris en charge.
UKSuggestions	Modèle de transformation configuré pour générer une liste de propositions pour les données d'adresses partielles au Royaume-Uni.

27.6 Au-delà du nettoyage d'adresse de base

La transformation USA Regulatory Address Cleanse propose des fonctionnalités de nettoyage d'adresse supplémentaires pour les adresses américaines. Ces fonctionnalités étendent le nettoyage d'adresse au-delà de l'analyse et le retraitement de base. Pour en savoir plus sur la transformation USA Regulatory Address Cleanse et ses options, voir le *Guide de référence*.

27.6.1 USPS DPV®

Delivery Point Validation® est un produit USPS conçu pour aider les utilisateurs à valider l'exactitude de leurs informations d'adresse. DPV compare les informations de code postal2 aux répertoires DPV afin d'identifier les adresses connues et les problèmes potentiels au niveau de l'adresse susceptibles de la rendre non livrable.

DPV est disponible pour les données des États-Unis, uniquement dans la transformation USA Regulatory Address Cleanse.

i Remarque

Le traitement DPV est requis pour la certification CASS. Si vous n'effectuez pas le traitement pour la certification CASS, vous pouvez choisir d'exécuter vos jobs en mode non certifié et toujours activer DPV

⚠ Attention

Si vous choisissez de désactiver le traitement DPV, le progiciel ne générera pas la documentation requise par CASS et votre publipostage ne sera pas éligible pour les remises sur les affranchissements.

Informations associées

[Activer DPV \[page 604\]](#)

[Mode non certifié \[page 605\]](#)

27.6.1.1 Avantages de DPV

DPV peut être bénéfique dans les domaines suivants :

- Publipostage : DPV aide à écarter les plis non distribuables (UAA) et aide à réduire les coûts du publipostage.
- Qualité d'informations : DPV augmente le niveau de précision des données en permettant de vérifier la maison, la suite ou l'appartement individuel d'une adresse, au lieu de vérifier uniquement l'immeuble.
- Taux d'affectation amélioré : DPV peut augmenter le taux d'affectation à l'aide d'un bris d'égalité DPV pour résoudre une égalité lorsque les autres méthodes de bris d'égalité ne sont pas concluantes.
- Prévention de la fraude par vente par correspondance : DPV peut éliminer les transports de marchandises aux personnes qui réalisent des commandes frauduleuses en vérifiant les adresses de livraison valides et les agences de boîte postale (CMRA).

27.6.1.2 Sécurité DPV

L'USPS a instauré des processus qui surveillent l'utilisation de DPV. Chaque société qui achète la fonctionnalité DPV doit signer un accord juridique qui affirme qu'elle n'essayera pas de faire mauvais usage du produit DPV. Si un utilisateur abuse du produit DPV, l'USPS a le droit d'interdire à l'utilisateur d'utiliser DPV à l'avenir.

27.6.1.2.1 fausses adresses DPV positives

L'USPS a inclus des fausses adresses positives dans les répertoires DPV comme une sécurité supplémentaire pour prévenir des abus de DPV. Selon votre type d'utilisateur et vos codes de clé de licence, le logiciel adopte un comportement différent lorsqu'il rencontre une adresse avec faux positif. Le tableau suivant décrit les comportements pour chaque type d'utilisateur :

Type d'utilisateur	Comportement du logiciel	A consulter
Utilisateurs finaux	Le traitement DPV s'interrompt.	Obtention d'un code de déblocage DPV auprès du Support SAP.
Utilisateurs finaux avec un contrat pour le Stop processing alternative	Le traitement DPV se poursuit.	Envoi des journaux de faux positifs à USPS
Fournisseurs de services	Le traitement DPV se poursuit.	Envoi des journaux de faux positifs à USPS

Informations associées

[Stop Processing Alternative \[page 601\]](#)

[Obtention d'un code de déblocage DPV auprès de SAP \[page 602\]](#)

[Envoi de journaux de faux positifs DPV à l'USPS \[page 603\]](#)

27.6.1.2.2 Stop Processing Alternative

Les utilisateurs finaux peuvent établir un contrat Stop Processing Alternative avec l'USPS et SAP.

Le fait d'établir un contrat Stop processing permet d'ignorer les verrouillages de répertoires futurs. Stop Processing Alternative n'est pas une option dans le logiciel, il s'agit d'un code de clé que vous obtenez auprès du Support SAP.

Vous devez d'abord obtenir les autorisations nécessaires auprès de l'USPS, puis fournir les preuves d'autorisation au Support SAP. Le support fournit ensuite un code de clé qui désactive la fonction de verrouillage de répertoire dans le logiciel.

→ N'oubliez pas

Lorsque vous avez obtenu le code de clé Stop Processing Alternative de la part du Support SAP, saisissez-le dans le Gestionnaire de licences SAP. Une fois le code de clé Stop Processing Alternative placé, le logiciel prend les mesures suivantes lorsqu'un faux positif est rencontré :

- Marque l'enregistrement comme un faux positif.
- Génère un fichier journal qui contient la fausse adresse positive.
- Note les chemins d'accès aux fichiers journaux dans le journal des erreurs.
- Génère un rapport US Regulatory Locking Report qui contient le chemin d'accès au fichier journal.
- Continue de traiter les jobs.

Bien que le job continue le traitement, vous devez envoyer le fichier journal de faux positifs à l'USPS pour les informer qu'une fausse adresse positive a été détectée. L'USPS doit libérer la liste avant que vous puissiez l'utiliser pour le traitement.

Informations associées

[Envoi de journaux de faux positifs DPV à l'USPS \[page 603\]](#)

27.6.1.2.3 journaux DPV de faux positifs

Le logiciel génère un fichier journal de faux positif à chaque fois qu'il rencontre un enregistrement de faux positif, sans tenir compte de la configuration du job. Le logiciel crée un fichier journal distinct pour chaque liste de diffusion qui contient un faux positif. Si plusieurs faux positifs existent dans une liste de diffusion, le logiciel les écrit tous dans le même fichier journal.

27.6.1.2.4 Verrouillage DPV pour utilisateurs finaux

Ce comportement de verrouillage s'applique aux utilisateurs finaux ou aux utilisateurs qui sont des clients DSF2 et qui ont désactivé DSF2 dans le job

Lorsque le logiciel trouve une fausse adresse positive, le traitement DPV est interrompu pour le reste du flux de données. Le logiciel prend également les mesures suivantes :

- Marque l'enregistrement comme adresse avec faux positif.
- Émet un message dans le journal des erreurs affirmant qu'une fausse adresse DPV positive a été rencontrée.
- Écrit l'adresse avec faux positif et le code de blocage dans le journal des erreurs.
- Continue de traiter les flux de données sans le traitement DPV.
- Génère un code de blocage.
- Génère un journal de faux positifs.
- Génère un rapport US Regulatory Locking Report qui contient la fausse adresse positive et le code de verrouillage. (La génération de rapport doit être activée dans la transformation USA Regulatory Address Cleanse.)

Pour restaurer la fonctionnalité DPV, l'utilisateur doit obtenir un code de déblocage DPV auprès du Support SAP.

Informations associées

[Obtention d'un code de déblocage DPV auprès de SAP \[page 602\]](#)

27.6.1.2.5 Obtention d'un code de déblocage DPV auprès de SAP

Ces étapes peuvent être appliquées pour des utilisateurs finaux ne disposant pas d'accord Stop Processing Alternative avec l'USPS. Lorsque vous recevez un message de traitement indiquant que des adresses avec faux positifs DPV figurent sur votre liste d'adresses, utilisez l'utilitaire de déblocage SAP USPS pour obtenir un code de déblocage.

1. Naviguez vers <https://support.sap.com/keys-systems-installations/keys/usps-unlock-codes.html> pour ouvrir la page Codes de déblocage USPS sur le SAP Support Portal.
2. Cliquez sur [Retrieve USPS Unlock Code \(Récupération d'un code de déblocage USPS\)](#).
3. Cliquez sur [Search \(Rechercher\)](#) et sélectionnez un système applicable Data Services dans la liste.
4. Saisissez le code de blocage trouvé dans le fichier `dpvx.txt` (l'emplacement est indiqué dans l'option [DPV Path \(Chemin d'accès DPV\)](#) dans le groupe de fichiers de référence).
5. Sélectionnez DPV comme type de blocage.
6. Sélectionnez BOJ-EIM-DS comme composant.
7. Saisissez l'adresse de blocage listée dans le fichier `dpvx.txt`.
8. Joignez le fichier `dpv1####.log` (l'emplacement est indiqué dans l'option [USPS Log Path \(Chemin d'accès du journal USPS\)](#) dans le groupe de fichiers de référence).
9. Cliquez sur [Submit \(Soumettre\)](#).
Le code de déblocage s'affiche.

10. Copiez le code de déblocage et collez-le dans le fichier `dpvw.txt`, en remplaçant tous les contenus du fichier par le code de déblocage (l'emplacement est indiqué dans l'option *DPV path (Chemin d'accès de DPV)* du groupe de fichiers de référence).
11. Supprimez l'enregistrement ayant causé le blocage de la base de données et supprimez le fichier `dpv1####.log` avant de traiter à nouveau la liste.

→ Conseil

N'oubliez pas que vous pouvez utiliser le code de déblocage une seule fois. Si le logiciel détecte un autre faux positif (même s'il s'agit du même enregistrement), vous devrez récupérer un nouveau code de déblocage LACSLink.

Si un code de déblocage n'a pas pu être généré, un message est de nouveau créé, il est traité par un ingénieur d'étude de l'Assurance client technique (pendant les heures d'ouverture normales).

i Remarque

Si vous êtes un utilisateur final disposant d'un accord Stop Processing Alternative, suivez les étapes pour envoyer le journal de faux positifs au USPS.

27.6.1.2.6 Envoi de journaux de faux positifs DPV à l'USPS

Les prestataires de service doivent suivre ces étapes une fois qu'ils ont reçu un message de traitement indiquant que de fausses adresses positives DPV figurent dans leur liste d'adresses. Les utilisateurs finaux disposant d'un accord Stop Processing Alternative doivent suivre ces étapes une fois qu'ils ont reçu un traitement du message indiquant que de fausses adresses positives DPV figurent dans leur liste d'adresses.

1. Envoyez un e-mail à USPS NCSC à l'adresse Dsf2stop@usps.gov et incluez les informations suivantes :
 - Saisissez « DPV False Positive » sur la ligne Objet.
 - Joignez le fichier `dpv1####.log` ou les fichiers précédemment générés par le logiciel (l'emplacement est indiqué dans l'option de répertoire *Chemin d'accès du journal USPS* dans l'option de répertoire du groupe de fichiers de référence)L'USPS NCSC utilise les informations pour déterminer si la liste peut être renvoyée à l'expéditeur.
2. Une fois que l'USPS NCSC a validé la liste contenant l'enregistrement bloqué ou de faux positifs :
 - Supprimez le fichier journal ou les fichiers correspondants.
 - Supprimez l'enregistrement qui a causé le blocage de la liste et retraitez le fichier.

Si vous êtes un utilisateur final qui ne dispose pas d'un accord Stop Processing Alternative, suivez les étapes permettant de récupérer le code de déblocage DPV auprès du Support SAP.

Informations associées

[Obtention d'un code de déblocage DPV auprès de SAP \[page 602\]](#)

27.6.1.3 Répertoires mensuels DPV

Les répertoires DPV sont livrés mensuellement avec les répertoires USPS conformément aux indications USPS.

Les répertoires expirent en 105 jours. La date sur les répertoires DPV doit être identique à celle du répertoire Adresse.

Ne renommez aucun fichier. DPV ne fonctionnera pas si les noms des fichiers sont modifiés. Vous trouverez ci-dessous une liste de répertoires DPV.

- `dpva.dir`
- `dpvb.dir`
- `dpvc.dir`
- `dpvd.dir`
- `dpv_vacant.dir`
- `dpv_no_stats.dir`

27.6.1.4 Informations requises pour la configuration des jobs

Lorsque vous configurez le traitement DPV, les options suivantes dans le groupe d'informations de licence USPS sont requises :

- *Nom de la société du client*
- *Adresse de la société du client*
- *Localité de la société du client*
- *Région de la société du client*
- *Code postal de la société du client1*
- *Code postal de la société du client2*

27.6.1.5 Activer DPV

i Remarque

DPV est requis pour CASS.

Outre les informations requises relatives à la société client que vous saisissez dans le groupe USPS License Information, activez les options suivantes pour exécuter le traitement DPV :

1. Ouvrez la transformation USA Regulatory Address Cleanse.
2. Ouvrez l'onglet *Options*, développez le groupe Options d'affectation et sélectionnez Oui pour l'option *Activer DPV*.
3. Dans le groupe de fichiers de référence, saisissez le chemin de vos répertoires DPV dans l'option *Chemin d'accès de DPV*.

i Remarque

DPV peut être exécuté uniquement lorsque l'emplacement pour tous les répertoires DPV a été saisi et aucun des fichiers de répertoire DPV n'a été renommé.

4. Définissez un référentiel pour le fichier journal DPV dans l'option *Chemin d'accès du journal USPS*. Utilisez la variable de substitution \$\$CertificationLogPath si vous l'avez configurée.
5. Dans le groupe Rapport et analyse, sélectionnez Oui pour l'option *Générer données de rapport*.

27.6.1.6 Champs de sortie DPV

Plusieurs champs de sortie sont disponibles pour signaler les résultats du traitement DPV :

- DPV_CMRA
- Note_Bas_Page_DPV
- DPV_No_Stat
- Statut_DPV
- DPV_Vacant

Pour accéder à une description complète de ces champs de sortie, consultez le *Guide de référence* ou les informations de l'Aide Data Services qui apparaissent lorsque vous ouvrez l'onglet Sortie de la transformation USA Regulatory Address Cleanse.

27.6.1.7 Mode non certifié

Il est possible de configurer les jobs en ayant DPV désactivé si vous n'êtes pas un client CASS mais que vous souhaitez ajouter un code postal2 aux adresses. L'option de non CASS, *Affecter code postal2 non validé DPV*, permet au logiciel d'affecter un code postal2 lorsqu'une adresse ne confirme pas un DPV.

⚠ Attention

Si vous choisissez de désactiver le traitement DPV, le logiciel ne génère pas la documentation requise pour CASS et le publipostage n'est pas éligible pour les remises postales.

27.6.1.7.1 Activer le mode non certifié

Pour exécuter un job en mode non certifié, suivez les étapes de configuration suivantes

1. Dans le groupe Opérations d'affectation, activez l'option *Activer DPV* sur Non.
2. Dans le groupe d'options Non certifié, activez l'option *Désactiver la certification* sur Oui.
3. Dans le groupe d'options Non certifié, activez l'option *Affecter code postal2 non validé DPV* sur Oui.

⚠ Attention

Le logiciel efface toutes les informations relatives au code postal2 de vos données si vous désactivez le traitement DPV et l'option *Affecter code postal2 non validé DPV*. Ceci comprend les informations relatives au code postal2 fournies dans votre fichier d'entrée.

27.6.1.8 Performances de DPV

En raison du temps supplémentaire requis pour exécuter le traitement DPV, vous pouvez constater une modification du temps de traitement. Le temps de traitement peut varier avec la fonctionnalité DPV selon le système d'exploitation, la configuration système et d'autres variables pouvant être uniques à votre environnement opérationnel.

Vous pouvez réduire le temps requis pour le traitement DPV en chargeant les répertoires DPV dans la mémoire système avant le traitement.

27.6.1.8.1 Utilisation de la mémoire

Vous devrez peut-être installer de la mémoire supplémentaire sur votre système d'exploitation pour le traitement DPV. Un minimum de 768 Mo est recommandé pour le traitement avec DPV activé.

Pour déterminer la quantité de mémoire requise pour une exécution avec DPV activé, contrôlez la taille des répertoires DPV (récemment environ 600 Mo¹) et ajoutez-la à la quantité de mémoire requise pour exécuter le logiciel.

La taille des répertoires DPV varie en fonction de la quantité de nouvelles données dans chaque version de répertoire.

Assurez-vous que votre ordinateur dispose de suffisamment de mémoire disponible avant d'effectuer le traitement DPV.

Pour connaître l'espace disque requis pour la mise en cache des répertoires, voir le document *Supported Platforms (Plateformes prises en charge)* sur SAP Support Portal. Vous trouverez des informations sur les liens dans le tableau des ressources d'informations SAP (voir le lien ci-dessous).

27.6.1.8.2 Mise en cache des répertoires DPV

Pour mieux gérer l'utilisation de la mémoire lorsque vous avez activé le traitement DPV, choisissez de cacher les répertoires DPV.

Pour configurer votre job pour la mise en mémoire cache DPV, suivez ces étapes :

1. Dans le groupe Transform Performance, activez l'option *Mettre en mémoire cache répertoires DPV* sur oui.

¹ ; la taille des répertoires est susceptible de changer à chaque fois que de nouveaux répertoires DPV sont installés

2. Dans le même groupe, activez l'option *Action de mémoire cache insuffisante* sur l'une des valeurs suivantes :

Option	Description
Erreur	Le logiciel génère une erreur et interrompt la transformation.
Continuer	Le logiciel tente de poursuivre l'initialisation sans mise en mémoire cache.

27.6.1.8.3 Exécution de plusieurs jobs avec DPV

Lors de l'exécution de plusieurs jobs DPV et du chargement de répertoires dans la mémoire, il est recommandé d'ajouter une pause de 10 secondes entre les jobs pour laisser du temps à la mémoire pour se libérer. Pour plus d'informations pour effectuer correctement ce paramétrage, consultez le manuel de votre système d'exploitation.

Si vous n'ajoutez pas une pause de 10 secondes entre les jobs, il est possible que votre système n'ait pas assez de temps pour libérer la mémoire utilisée pour la mise en cache des répertoires à partir du premier job. Le prochain job en attente de traitement peut créer une erreur ou accéder aux répertoires à partir du disque s'il n'y a pas assez de mémoire pour cacher les répertoires. Cela peut aboutir à une dégradation de la performance.

27.6.1.9 Informations DPV dans le document US Addressing Report

Le document US Addressing Report est généré automatiquement lorsque le reporting est activé dans le job. Les sections suivantes du rapport US Addressing Report contiennent les informations DPV :

- Codes de retour DPV
- Résumé Delivery Point Validation (DPV)

Pour plus d'informations sur le rapport US Addressing Report, ou sur d'autres rapports Data Quality, consultez le guide *Management Console Guide*.

27.6.1.10 Indicateurs Aucune statistique DPV

L'USPS utilise les indicateurs Aucune statistique pour marquer les adresses qui rentrent dans la catégorie Aucune statistique. Le logiciel utilise la table No Stats lorsque DPV ou DSF2 sont allumés dans le job. L'USPS place les adresses sans statistique dans trois catégories :

- Les adresses qui n'ont pas encore de livraison établie.
- Les adresses qui reçoivent le courrier dans le cadre d'un dépôt.
- Les adresses qui ont été vacantes pendant une certaine durée.

27.6.1.10.1 Table No Stats

Vous devez installer la table No Stats `dpv_no_stats.dir` avant que le logiciel effectue le traitement DPV ou DSF2. La table No Stats est fournie par SAP BusinessObjects avec l'installation du répertoire DPV.

Le logiciel contrôle automatiquement la table No Stats dans le dossier du répertoire que vous avez indiqué dans la configuration du job. Le logiciel effectue le traitement DPV ou DSF2 en fonction du statut d'installation du répertoire.

<code>dpv_no_stats.dir</code>	Résultats
Installé	Le logiciel sort automatiquement les indicateurs Aucune statistique lorsque vous incluez le champ de sortie DPV_NoStats dans le job.
Non installé	Le logiciel ignore automatiquement le traitement Aucune statistique et n'émet pas un message d'erreur. Le logiciel effectue le traitement DPV mais ne renseigne pas le champ de sortie DPV_NoStats.

27.6.1.10.2 Champ de sortie Aucune statistique

Utilisez le champ de sortie DPV_NoStats pour enregistrer les informations de l'indicateur Aucune statistique dans un fichier de sortie.

Aucune statistique signifie que l'adresse est une propriété vacante, qu'elle reçoit du courrier dans le cadre d'un dépôt, ou qu'elle n'a pas encore de livraison établie.

Informations associées

[Champs de sortie DPV \[page 605\]](#)

27.6.1.11 Indicateurs Vacants DPV

Le logiciel fournit des informations sur les vacants dans les champs de sortie et des rapports à l'aide des nombres de vacants DPV. La table de recherche des vacants USPS DPVest fournie par SAP BusinessObjects avec l'installation du répertoire DPV.

L'USPS utilise des indicateurs vacants DPV pour marquer les adresses qui rentrent dans la catégorie vacante. Le logiciel utilise les indicateurs vacants DPV lorsque vous avez activé DPV ou DSF2 dans le job.

→ Conseil

L'USPS définit comme vacant tous les points de livraison qui étaient actifs dans le passé mais qui sont actuellement non occupés (généralement au-delà de 90 jours) et qui ne reçoivent actuellement aucune

livraison de courrier. L'adresse pourrait recevoir une livraison à nouveau dans le futur. "Vacant" ne s'applique pas aux adresses saisonnières.

27.6.1.11.1 Champ de sortie de l'attribut d'adresse DPV

Les indicateurs vacants pour les adresses affectées sont disponibles dans le champ de sortie DPV_Vacant.

i Remarque

Le rapport US Addressing Report contient des nombres de Vacant DPV dans la section Summary DPV.

Informations associées

[Champs de sortie DPV \[page 605\]](#)

27.6.2 LACSLink

LACSLink est un produit USPS disponible pour les enregistrements des États-Unis, uniquement avec la transformation USA Regulatory Address Cleanse. Le traitement LACSLink est requis pour la certification CASS.

LACSLink met à jour les adresses lorsque l'adresse physique ne change pas, mais que l'adresse postale a été modifiée. Par exemple, lorsque la municipalité change les adresses de routes rurales en adresses de nom de rue. Les conversions de routes rurales facilitent la localisation des adresses rurales pour les personnels de la police, des pompiers, des ambulances et de la poste. LACSLink convertit également les adresses lorsque les rues sont renommées ou les boîtes postales renumérotées.

La technologie LACSLink garantit que les données restent privées et sécurisées tout en vous permettant un accès facile à celles-ci. LACSLink est une partie intégrée du traitement des adresses, ce n'est pas une étape supplémentaire. Pour obtenir les nouvelles adresses, vous devez déjà avoir les anciennes données d'adresses.

Informations associées

[Fonctionnement de LACSLink \[page 610\]](#)

[Contrôle de l'utilisation de mémoire pour le traitement LACSLink \[page 617\]](#)

[Désactivation de LACSLink \[page 615\]](#)

[Sécurité LACSLink \[page 610\]](#)

27.6.2.1 Fonctionnement de LACSLink

LACSLink fournit une nouvelle adresse si elle existe. LACSLink suit ces étapes lors du traitement d'une adresse :

1. La transformation USA Regulatory Address Cleanse standardise l'adresse d'entrée.
2. La transformation recherche une adresse correspondante dans les données LACSLink.
3. Si une correspondance est trouvée, la transformation édite l'adresse convertie en LACSLink et les autres informations LACSLink.

Informations associées

[Contrôle de l'utilisation de mémoire pour le traitement LACSLink \[page 617\]](#)

[LACSLink \[page 609\]](#)

27.6.2.2 Sécurité LACSLink

USPSa institué des processus qui surveillent l'utilisation de LACSLink. Chaque société qui achète la fonctionnalité LACSLink doit signer un accord juridique qui affirme qu'elle n'essayera pas de faire mauvais usage du produit LACSLink. Si un utilisateur abuse du produit LACSLink, l'USPSa le droit d'interdire à l'utilisateur d'utiliser LACSLink à l'avenir.

27.6.2.2.1 Adresses avec faux positif LACSLink

USPS a inclus des adresses avec faux positif dans les répertoires LACSLink, mettant ainsi en place une sécurité supplémentaire pour éviter le détournement de LACSLink. Selon votre type d'utilisateur et vos codes de clé de licence, le progiciel adopte un comportement différent lorsqu'il rencontre une adresse avec faux positif. Le tableau suivant décrit les comportements pour chaque type d'utilisateur :

Type d'utilisateur	Comportement du progiciel	A consulter
Utilisateurs finaux	Le traitement LACSLink s'interrompt.	Obtention du code de déblocage LACSLink auprès du Support SAP
Utilisateurs finaux disposant d'un accord Stop Processing Alternative	Le traitement LACSLink se poursuit.	Envoi des journaux de faux positifs à USPS
Fournisseurs de services	Le traitement LACSLink se poursuit.	Envoi des journaux de faux positifs à USPS

Informations associées

[Stop Processing Alternative \[page 601\]](#)

[Obtention d'un code de déblocage LACSLink auprès de SAP \[page 611\]](#)

[Envoi de journaux de faux positifs LACSLink au USPS \[page 612\]](#)

27.6.2.2.2 Journaux de faux positifs LACSLink

Le progiciel génère un fichier journal de faux positif chaque fois qu'il rencontre un enregistrement avec faux positif, quelle que soit la configuration de votre job. Un fichier journal distinct est créé pour chaque liste de diffusion contenant un faux positif. Si plusieurs faux positifs existent dans une liste de diffusion, le logiciel les écrit tous dans le même fichier journal.

27.6.2.2.3 Blocage de LACSLink pour les utilisateurs finaux

Un tel blocage est pertinent pour les utilisateurs finaux ou les utilisateurs qui disposent d'une licence DSF2 lorsque DSF2 est désactivé dans le job.

Lorsque le progiciel identifie une adresse avec faux positif, le traitement LACSLink est abandonné pour le reste du traitement du job. Le progiciel effectue les opérations suivantes :

- Il marque l'enregistrement comme adresse avec faux positif.
- Dans le journal des erreurs, il génère un message indiquant qu'une adresse avec faux positif LACSLink a été identifiée.
- Il indique l'adresse avec faux positif et le code de blocage dans le journal des erreurs.
- Continue de traiter les flux de données sans le traitement LACSLink.
- Il génère un code de blocage.
- Il génère un journal des erreurs de faux positifs.
- Il génère un rapport US Regulatory Locking Report contenant l'adresse avec faux positif et le code de blocage (la création du rapport doit être activée dans la transformation USA Regulatory Address Cleanse).

Pour restaurer la fonctionnalité LACSLink, l'utilisateur doit obtenir un code de déblocage LACSLink auprès du Support SAP.

27.6.2.2.4 Obtention d'un code de déblocage LACSLink auprès de SAP

Ces étapes peuvent être appliquées pour des utilisateurs finaux ne disposant pas d'accord Stop Processing Alternative avec l'USPS. Lorsque vous recevez un message de traitement indiquant que des adresses avec faux positifs LACSLink figurent sur votre liste d'adresses, utilisez l'utilitaire de déblocage SAP USPS pour obtenir un code de déblocage.

1. Naviguez vers <https://support.sap.com/keys-systems-installations/keys/usps-unlock-codes.html> pour ouvrir la page Codes de déblocage USPS sur le SAP Support Portal.
2. Cliquez sur *Retrieve USPS Unlock Code (Récupération d'un code de déblocage USPS)*.
3. Cliquez sur *Search (Rechercher)* et sélectionnez un système applicable Data Services dans la liste.
4. Saisissez le code de blocage trouvé dans le fichier `lacsx.txt` (l'emplacement est indiqué dans l'option *LACSLink Path (Chemin d'accès LACSLink)* dans le groupe de fichiers de référence).
5. Sélectionnez LACSLink comme type de blocage.
6. Sélectionnez BOJ-EIM-DS comme composant.
7. Saisissez l'adresse de blocage listée dans le fichier `lacsx.txt`.
8. Joignez le fichier `lacs1####.log` (l'emplacement est indiqué dans l'option *USPS Log Path (Chemin d'accès du journal USPS)* dans le groupe de fichiers de référence).
9. Cliquez sur *Submit (Soumettre)*.
Le code de déblocage s'affiche.
10. Copiez le code de déblocage et collez-le dans le fichier `lacsx.txt`, en remplaçant tous les contenus du fichier par le code de déblocage (l'emplacement est indiqué dans l'option *LACSLink path (Chemin d'accès LACSLink)* dans le groupe de fichiers de référence).
11. Supprimez l'enregistrement ayant causé le blocage de la base de données et le fichier `lacs1####.log` avant de traiter à nouveau la liste.

→ Conseil

N'oubliez pas que vous pouvez utiliser le code de déblocage une seule fois. Si le logiciel détecte un autre faux positif (même s'il s'agit du même enregistrement), vous devrez récupérer un nouveau code de déblocage LACSLink.

Si un code de déblocage n'a pas pu être généré, un message est de nouveau créé, il est traité par un ingénieur d'étude de l'Assurance client technique (pendant les heures d'ouverture normales).

i Remarque

Si vous êtes un utilisateur final disposant d'un accord Stop Processing Alternative, suivez les étapes pour envoyer le journal de faux positifs au USPS.

27.6.2.2.5 Envoi de journaux de faux positifs LACSLink au USPS

Les prestataires de service doivent suivre ces étapes une fois qu'ils ont reçu un message de traitement indiquant que de fausses adresses positives LACSLink figurent dans leur liste d'adresses. Les utilisateurs finaux disposant d'un accord Stop Processing Alternative doivent suivre ces étapes une fois qu'ils ont reçu un traitement du message indiquant que de fausses adresses positives LACSLink figurent dans leur liste d'adresses.

1. Envoyez un e-mail au USPS à l'adresse Dsf2stop@usps.gov. Incluez les éléments suivants :
 - Saisissez « LACSLinkFaux positif » sur la ligne Objet
 - Joignez le fichier `lacs1####.log` ou les fichiers générés par le logiciel (l'emplacement est indiqué dans l'option *Fichiers journal UPS* dans le groupe de fichiers de référence).

Le USPS NCSC utilise les informations pour déterminer si la liste peut être renvoyée à l'expéditeur ou pas.

2. Une fois que USPS NCSC a validé la liste contenant l'enregistrement bloqué ou de faux positif :

- Supprimez le fichier journal ou les fichiers correspondants.
- Supprimez l'enregistrement qui a causé le blocage de la liste et retraitez le fichier.

Si vous êtes un utilisateur final qui ne dispose pas d'un accord Stop Processing Alternative, suivez les étapes permettant de récupérer le code de déblocage LACSLink auprès du Support SAP.

Informations associées

[Obtention d'un code de déblocage LACSLink auprès de SAP \[page 611\]](#)

27.6.2.3 Conditions relatives au traitement des adresses

La transformation ne traite pas toutes les adresses à l'aide de LACSLink lorsque ce service est activé. Voici les conditions en fonction desquelles vos données sont transmises au traitement LACSLink :

- L'adresse est disponible dans le répertoire d'adresses et elle est marquée en tant qu'enregistrement convertible LACS dans le répertoire d'adresses.
- L'adresse est disponible dans le répertoire d'adresses et bien qu'une affectation par défaut à un contrat d'itinéraire rural ou d'autoroute ait été effectuée, l'enregistrement n'a pas été marqué comme convertible LACS.
- L'adresse n'est pas disponible dans le répertoire d'adresses mais l'enregistrement contient suffisamment d'informations pour être envoyé à LACSLink.

Le tableau suivant affiche, par exemple, une adresse qui a été trouvée dans le répertoire d'adresses en tant qu'adresse convertible LACS.

Adresse d'origine	Après conversion LACSLink
RR2 BOX 204	463 SHOWERS RD
DU BOIS PA 15801	DU BOIS PA 15801-66675

27.6.2.4 Modèle de configuration de transformation

Le traitement LACSLink est activé par défaut dans le modèle de configuration de transformation car il est nécessaire dans le cadre de la certification CASS. Le modèle de configuration de transformation se nomme USARegulatory_AddressCleanse ; il est disponible sous le groupe USA_Regulatory_Address_Cleanse dans la bibliothèque d'objets.

27.6.2.5 Fichiers de répertoire LACSLink

SAP fournit les fichiers de répertoire LACSLink avec la mise à jour du répertoire national des Etats-Unis. Les fichiers du répertoire LACSLink nécessitent environ 600 Mo d'espace supplémentaire sur le disque dur. Les répertoires LACSLink comprennent les éléments suivants :

- `lacsw.txt`
- `lacsx.txt`
- `lacsy.ll`
- `lacsx.ll`

⚠ Attention

Les répertoires LACSLink doivent résider sur le disque dur, dans le même répertoire que les fichiers de prise en charge LACSLink. Ne renommez aucun fichier. LACSLink ne s'exécutera pas si les noms de fichier ont été modifiés.

27.6.2.6 Activation de LACSLink

Le service LACSLink est activé par défaut dans la transformation USA Regulatory Address Cleanse. Si vous devez réactiver l'option, procédez comme suit :

1. Ouvrez la transformation USA Regulatory Address Cleanse et sélectionnez l'onglet *Options*.
2. Développez le groupe Options de traitement.
3. Sélectionnez Oui pour l'option *Enable LACSLink*.
4. Saisissez le chemin d'accès LACSLink pour l'option *LACSLink Path* dans le groupe Fichiers de référence. Vous pouvez utiliser la variable de substitution `$$RefFilesAddressCleanse` si elle est configurée.
5. Renseignez les champs requis dans le groupe Informations de licence USPS.

27.6.2.6.1 Informations à indiquer dans la configuration du job

Tous les utilisateurs qui se servent de LACSLink doivent indiquer les informations requises dans le groupe Informations de licence USPS. Les options suivantes doivent être traitées :

- Nom de la société du client
- Adresse de la société du client
- Localité de la société du client
- Région de la société du client
- Code postal de la société du client1
- Code postal de la société du client2
- Téléphone de la société du client

27.6.2.6.2 Désactivation de LACSLink

Le service LACSLink est activé par défaut dans la configuration de la transformation USA Regulatory Address Cleanse car il intervient dans le traitement CASS. Vous devez donc désactiver la certification CASS afin de désactiver LACSLink.

1. Dans la configuration de la transformation USA Regulatory Address Cleanse, ouvrez l'onglet [Options](#).
2. Ouvrez le groupe Options non certifiées.
3. Sélectionnez Oui pour l'option [Désactiver la certification](#).
4. Ouvrez le groupe Opération d'affectation.
5. Sélectionnez Non pour l'option [Enable LACSLink](#).

Informations associées

[LACSLink \[page 609\]](#)

27.6.2.6.3 Motifs d'erreurs

Si certaines informations du groupe Informations de licence USPS sont absentes de votre job et que DPV et/ou LACSLink est/sont activé(s) dans ce job, vous obtenez des messages d'erreur dans les cas suivants :

Motif de l'erreur	Description
Options requises manquantes	Lorsque la configuration de votre job ne comprend pas les paramètres obligatoires du groupe Informations de licence USPS et que DPV et/ou LACSLink est/sont activé(s) dans ce job, le progiciel génère une erreur de vérification.
Répertoire de fichiers journaux non accessible en écriture	Si le chemin d'accès que vous avez indiqué pour l'option Chemin d'accès au journal USPS dans le groupe Fichiers de référence n'est pas accessible en écriture, le progiciel génère une erreur.

27.6.2.7 Champs de sortie LACSLink

Plusieurs champs de sortie sont disponibles pour signaler les résultats du traitement LACSLink.

Vous devez activer LACSLink et ajouter ces champs de sortie dans la configuration de votre job pour que le progiciel enregistre des informations dans ces champs.

Nom du champ	Longueur	Description
LACSLINK_QUERY	50	Retourne l'adresse avant conversion, uniquement complétée lorsque LACSLink est activé et qu'une recherche LACSLink a été tentée. Cette adresse se présente au format USPS standard (conformément à l'USPS Publication 28). Cependant, lorsqu'une adresse possède à la fois un indicateur d'unité et une unité secondaire, l'indicateur est remplacé par le caractère « # ». Non renseigné : aucune recherche LACSLink n'a été entreprise.
LACSLINK_RETURN_CODE	2	Renvoie le statut de correspondance pour le traitement LACSLink : A = correspondance de l'enregistrement LACSLink. Une adresse convertie est fournie dans les champs de données d'adresse. 00 = aucune correspondance et aucune adresse convertie. 09 = LACSLink a rapproché une adresse d'entrée d'une ancienne adresse correspondant à une adresse "high-rise default". Aucune nouvelle adresse n'est fournie. 14 = un enregistrement LACSLink a été identifié mais les données n'ont pas pu être converties en adresse livrable. 92 = une correspondance avec l'enregistrement LACSLink a été établie après suppression du numéro secondaire de l'adresse d'entrée. Non renseigné = aucune recherche LACSLink n'a été entreprise.

Nom du champ	Longueur	Description
LACSLINKINDICATOR	1	Retourne le statut de conversion des adresses traitées par LACSLink. Y = adresse convertie par LACSLink (LACSLink_Return_Code a la valeur A). N = adresse recherchée via LACSLink mais non convertie. F = adresse correspondant à un faux positif. S = la conversion LACSLink a été effectuée mais les informations secondaires ont dû être supprimées. vide : aucune recherche LACSLink n'a été entreprise.

27.6.2.8 Contrôle de l'utilisation de mémoire pour le traitement LACSLink

Les performances de la transformation s'améliorent considérablement si vous mettez en cache les répertoires LACSLink. Pour connaître l'espace disque requis pour la mise en cache des répertoires, voir le document [Matrice de disponibilité des produits SAP Data Services pour Data Services<version> dans la PAM](#).

Si vous ne disposez pas de la mémoire système adéquate pour charger les répertoires LACSLink et que le paramètre *Action de mémoire cache insuffisante* a la valeur Erreur, un message d'erreur de vérification est affiché lors de l'exécution et la transformation s'interrompt. Si vous sélectionnez l'option Continuer, la transformation tente de poursuivre le traitement LACSLink sans mise en cache.

Ouvrez l'onglet *Options* de la configuration de la transformation USA Regulatory Address Cleanse dans votre flux de données. Pour charger les répertoires USA Regulatory Address Cleanse dans la mémoire système, procédez comme suit :

1. Ouvrez le groupe d'options Performance de la transformation.
2. Sélectionnez Oui pour l'option *Mise en cache des répertoires LACSLink*.

Informations associées

[LACSLink \[page 609\]](#)

27.6.2.9 Informations LACSLink dans le document US Addressing Report

Le document US Addressing Report est généré automatiquement lorsque le reporting est activé dans le job. Le tableau suivant énumère les sections LACSLink du document US Addressing Report :

Section	Informations
Locatable Address Conversion (LACS-Link) Summary	Enregistre le nombre et le pourcentage pour les informations suivantes : • Adresses converties LACSLink • Adresses non converties LACSLink
Codes de retour LACSLink	Enregistre le nombre et le pourcentage pour les informations suivantes : • Converted • Secondary dropped • No match • Can't convert • High-rise default

27.6.2.10 USPS Form 3553

Le formulaire USPS Form 3553 indique le comptage LACSLink. Le champ LACS/LACSLink affiche le nombre d'enregistrements qui possèdent un indicateur LACSLinkY ou S si le traitement LACSLink est activé. Si le traitement LACSLink n'est pas activé, ce champ affiche la valeur du comptage des codes LACS.

27.6.3 SuiteLink™

SuiteLink est une option de la transformation USA Regulatory Address Cleanse.

SuiteLink utilise un répertoire USPS contenant plusieurs fichiers d'informations sur des adresses indexées, par exemple les numéros secondaires et les désignations d'unité pour des emplacements identifiés comme bâtiments par défaut.

SuiteLink vous permet de créer des adresses actualisées et complètes en ajoutant des numéros de suite aux adresses commerciales de bâtiment. Avec les informations secondaires relatives à l'adresse ajoutée à vos adresses, la majorité de vos pièces sont triées par séquence de livraison et livrées rapidement avec précision.

SuiteLink est requis pour CASS

SuiteLink est requis pour un traitement en mode CASS (et la valeur Non est attribuée à l'option [Désactiver la certification](#)). Si vous avez désactivé SuiteLink dans votre configuration de job mais que vous êtes en mode CASS, un message d'erreur est émis et le traitement est interrompu.

27.6.3.1 Avantages de SuiteLink

Pour les entreprises qui dépendent d'un site Web, d'e-mail ou de commandes au point de vente, SuiteLink est un remarquable outil permettant de faire des économies. Il en est de même pour les entreprises dont les clients habitent dans des bâtiments avec activités multiples qui apprécieront d'avoir leurs documents marketing, leurs extraits de compte et leurs commandes livrés à proximité.

Les informations relatives au numéro secondaire de vos adresses permettent un séquençage de la livraison plus performant et à moindre coût ainsi que des remises sur le port.

i Remarque

SuiteLink permet de préparer des listes de mailing conformes à CASS.

27.6.3.2 Comment fonctionne SuiteLink

Le logiciel utilise les données des répertoires SuiteLink pour ajouter des numéros de suite aux adresses applicables. Le logiciel fait correspondre le nom d'une société, l'adresse d'un bâtiment et le code postal2 certifié CASS de votre base de données aux données du SuiteLink. Lorsqu'il y a une correspondance, le logiciel crée une adresse professionnelle complète contenant le numéro de suite.

❖ Exemple

Affectez un numéro de suite

Cet exemple présente un enregistrement traité via SuiteLink et l'enregistrement de sortie avec le numéro de suite affecté.

L'enregistrement d'entrée contient les éléments suivants :

- Nom de la société (dans le champ d'entrée SOCIETE)
- Adresse de bâtiment connue
- Code postal2 certifié CASS

Le répertoire SuiteLink contient les éléments suivants :

- les numéros de l'adresse secondaire
- les désignations de l'unité

L'enregistrement de sortie contient les éléments suivants :

- le numéro de suite correct

l'enregistrement d'entrée	l'enregistrement de sortie
Telera	TELERA
910 E Hamilton Ave FI2	910 E HAMILTON AVE STE 200
Campbell CA 95008 0610	CAMPBELL CA 95008 0625

27.6.3.3 Répertoire SuiteLink

Le répertoire SuiteLink est distribué mensuellement.

Vous devez utiliser des répertoires SuiteLink avec un répertoire `zip4us.dir` pour le même mois. (Saisissez le chemin `zip4us.dir` dans l'option [Répertoire d'adresses](#) du groupe de fichiers de référence dans la transformation USA Regulatory Address Cleanse.) Par exemple, le répertoire SuiteLink Décembre 2011 peut être utilisé avec uniquement le répertoire `zip4us.dir` Décembre 2011.

Vous ne pouvez pas utiliser un répertoire SuiteLink datant de plus de 60 jours à partir de sa date de validation. Le logiciel vous prévient 15 jours avant l'expiration du répertoire. Comme dans le cas de tous les répertoires, le logiciel ne traite pas vos enregistrements avec un répertoire SuiteLink qui a expiré.

27.6.3.4 Activer SuiteLink

SuiteLink est activé par défaut dans toute configuration d'exemple de transformation définie conformément à CASS (et l'option Désactiver la certification est activée sur Non). Par exemple, SuiteLink est activé si vous utilisez la configuration d'exemple de transformation USA Regulatory_AddressCleanse.

1. Ouvrez la transformation USA Regulatory Address Cleanse dans votre flux de données.
2. Ouvrez l'onglet [Options](#).
3. Développez le groupe Opération d'affectation et activez l'option [Enable SuiteLink](#) sur Oui.
4. Dans le groupe de fichiers de référence, saisissez le chemin d'accès au répertoire SuiteLink dans l'option [SuiteLink Path](#). Vous pouvez utiliser la variable de substitution `$$RefFilesAddressCleanse` si vous l'avez configurée avec l'emplacement de répertoire contenant vos répertoires SuiteLink.
5. Facultatif : dans le groupe d'option Transform Performance, activez l'option [Cache SuiteLink Directories](#) sur Oui de telle sorte que les répertoires SuiteLink soient mis en mémoire cache dans la mémoire.

i Remarque

Vérifiez que vous avez une mémoire RAM suffisante pour pouvoir mettre en mémoire cache les répertoires SuiteLink avant d'activer cette option.

27.6.3.5 Amélioration de la vitesse de traitement de SuiteLink

Vous pouvez augmenter la vitesse de traitement de SuiteLink en chargeant les répertoires SuiteLink dans la mémoire. Pour activer cette option, accédez au groupe Performance de transformation et activez les [Cache SuiteLink Directories](#) sur Oui.

27.6.3.6 SuiteLink Codes de retour dans l'US Addressing Report

Les informations relatives au code de retour SuiteLink sont disponibles dans la section Codes retour SuiteLink du US Addressing Report.

L'US Addressing Report indique le nombre d'enregistrements et le pourcentage pour les codes retour suivants :

A = Des données secondaires existent et une affectation est effectuée

00 = Une recherche a été tentée mais aucune affectation

27.6.4 USPS DSF2®

DSF2 est un produit avec une licence USPS que vous pouvez utiliser pour valider les adresses, ajouter des informations de séquence de livraison et ajouter des attributs d'adresse DSF2 aux adresses.

Les deux fonctionnalités DSF2 suivantes sont prises en charge dans Data Services :

- DSF2 Augment dans la transformation USA Regulatory Address Cleanse
- Walk Sequence de DSF2 dans la transformation DSF2 Walk Sequencer

i Remarque

Les données USPS DSF2 sont disponibles uniquement pour les licenciés USPS certifiés par DSF2.

Informations associées

[Séquençement de tournée de DSF2 \[page 626\]](#)

27.6.4.1 Validation des adresses

DSF2 permet de réduire la quantité de plis non distribuables (PND) et garde des prix de publipostage bas. DSF2 utilise DPV® pour valider les adresses et identifier les adresses non précises ou incomplètes.

Informations associées

[Vue d'ensemble DPV \[page 599\]](#)

27.6.4.2 Ajout d'attributs d'adresse

DSF2 ajoute des attributs d'adresse (informations sur les adresses) à vos données. Utilisez les informations des attributs pour créer des publipostages plus ciblés.

27.6.4.3 Ajout d'informations de séquence de livraison

DSF2 ajoute des informations de séquence de livraison aux données. Vous pouvez les utiliser pour remplir les conditions des remises de tournée. Ces informations sont parfois appelées "walk sequencing" (séquençement de tournée) ou "pseudo sequencing" (pseudo séquençement).

Informations associées

[Séquençement de tournée \[page 626\]](#)

[Pseudo séquençement \[page 626\]](#)

27.6.4.4 Avantages de DSF2

Tous ceux qui souhaitent cibler leur courrier vers des types spécifiques d'adresses, et ceux qui souhaitent gagner des remises postales supplémentaires apprécieront ce que DSF2 peut effectuer.

Les données d'attributs d'adresses DSF2 dotent les expéditeurs de la connaissance sur les adresses supérieure et inférieure à ce qui est nécessaire pour mettre précisément en forme les adresses. Les données d'attributs d'adresse permettent aux expéditeurs de produire plus de publipostage ciblé.

Par exemple, si vous prévoyez d'envoyer un bon pour votre entreprise de services d'entretien de pelouse, vous ne souhaitez pas l'envoyer aux personnes habitant dans un appartement (elles n'ont sûrement pas de pelouse). Vous souhaitez que votre bon soit envoyé à des adresses résidentielles ne se situant pas dans des immeubles d'habitation.

Avec les informations DSF2, il est possible de faire une tournée de vos publipostages pour obtenir les meilleures remises postales possibles grâce à l'utilisation de la transformation DSF2 Walk Sequencer.

27.6.4.5 Devenir un licencié DSF2

Avant de pouvoir effectuer le traitement DSF2 dans le logiciel, vous devez exécuter les procédures de certification USPS DSF2 et obtenir la licence d'USPS.

Une partie de la certification consiste à traiter les jobs test dans Data Services pour prouver que le logiciel est conforme au contrat de licence. Lorsque vous êtes prêt à effectuer ces tests, contactez l'assistance des utilisateurs professionnels SAP BusinessObjects pour obtenir l'accès aux fonctionnalités DSF2 dans Data Services.

Pour en savoir plus, voir la rubrique « Certification DSF2 » dans le *Guide de référence* (section « Certifications USPS »).

27.6.4.6 Répertoires DSF2

Le traitement DSF2 nécessite les données suivantes :

Données	Notes
Répertoires DPV	Le logiciel utilise les répertoires DPV pour vérifier les adresses et pour identifier les adresses inexactes. SAP fournit les répertoires DPV avec la livraison du répertoire national des Etats-Unis. i Remarque Les répertoires DPV sont inclus avec les tables DSF2. N'utilisez pas les répertoires DPV inclus dans les tables DSF2. Utilisez les répertoires DPV de SAP avec la livraison du répertoire national des Etats-Unis.
Répertoires eLOT	Le logiciel utilise les répertoires eLOT pour affecter des numéros de tournée. SAP fournit les répertoires eLOT avec la livraison du répertoire national des Etats-Unis. i Remarque Les répertoires eLOT sont inclus avec les tables DSF2. N'utilisez pas les répertoires eLOT inclus avec les tables DSF2. Utilisez les répertoires eLOT de SAP avec la livraison du répertoire national des Etats-Unis.
Tables DSF2	Le logiciel utilise les tables DSF2 pour affecter des attributs d'adresse. i Remarque Les tables DSF2 sont fournies par USPS et non pas par SAP. De plus, les tables DSF2 incluent les répertoires DPV et eLOT. N'utilisez pas les répertoires DPV et eLOT inclus avec les tables DSF2. Utilisez les répertoires DPV et eLOT de SAP avec la livraison du répertoire national des Etats-Unis.
Fichier de statistiques de livraisons	Le logiciel utilise le fichier de statistiques de livraison pour fournir des nombres d'adresses commerciales et résidentielles par ZIP Code (Code postal1), par itinéraire du transporteur (code de tri). SAP fournit le fichier de statistiques de livraison avec la livraison du répertoire national des Etats-Unis.

Vous devez indiquer l'emplacement de ces fichiers de répertoires dans la transformation USA Regulatory Address Cleanse, excepté pour le fichier de statistiques de livraison. Définissez l'emplacement du fichier de statistiques de livraison (dsf.dir) dans la transformation DSF2 Walk Sequencer. Pour remplir les conditions DSF2, vous devez également installer tous les mois des répertoires mis à jour.

27.6.4.7 Traitement DSF2 Augment

Configurez le traitement DSF2 Augment dans la transformation USA Regulatory Address Cleanse.

Le traitement DSF2 exige des informations DPV ; par conséquent, activez le DPV dans la configuration du job.

Si vous prévoyez d'utiliser les informations de sortie du traitement Augment de DSF2 pour le traitement de tournée, vous devez également activer eLOT.

i Remarque

DSF2 Augment est disponible uniquement en mode batch. Il est impossible d'ajouter des informations Augment à vos données en temps réel.

27.6.4.7.1 Expiration des répertoires DSF2 Augment

Les répertoires DSF2 sont distribués tous les mois. Vous devez utiliser les répertoires DSF2 avec les répertoires nationaux des Etats-Unis qui sont étiquetés pour le même mois. Par exemple, les répertoires DSF2 Mai 2011 peuvent uniquement être utilisés avec les répertoires nationaux Mai 2011.

Les données DSF2 Augment expirent en 60 jours au lieu de l'expiration en 105 jours pour les répertoires nationaux des Etats-Unis. Du fait que tous les répertoires doivent avoir la même date de base (MM/AAAA), les utilisateurs DSF2 pour lesquels l'option Activer DSF2 Augment est définie sur Oui doivent mettre à jour tous les répertoires nationaux des Etats-Unis et les autres répertoires qu'ils utilisent (LACSLink ou DPV, par exemple) en même temps que les répertoires DSF2 Augment. Le logiciel rappelle aux utilisateurs de mettre à jour les répertoires par un message d'avertissement qui apparaît 15 jours avant que le répertoire expire.

→ N'oubliez pas

Comme avec tous les répertoires, le logiciel ne traite pas les enregistrements avec des répertoires DSF2 expirés.

27.6.4.7.2 Identification du licencié DSF2

Lorsque vous effectuez un traitement DSF2, vous devez fournir les informations suivantes : la société avec la licence DSF2 et le client pour lequel la société traite le job.

Vous devez compléter les options suivantes dans le groupe Informations de licence USPS pour le traitement DSF2 :

- ID du licencié DSF2
- Nom du licencié
- List Owner NAICS Code
- ID de liste
- Nom de la société du client
- Adresse de la société du client

- Localité de la société du client
- Région de la société du client
- Code postal de la société du client1
- Code postal de la société du client2
- Date de réception de la liste
- Date de retour de la liste

i Remarque

Si vous effectuez des traitements DSF2 et NCOALink dans la même instance de la transformation USA Regulatory Address Cleanse, alors les informations que vous saisissez dans le groupe Informations de la licence USPS doivent s'appliquer aux traitements DSF2 et NCOALink. Si, par exemple, l'*ID de liste* est différent pour le DSF2 et le NCOALink, vous devez inclure deux transformations USA Regulatory Address Cleanse : l'une pour NCOALink et l'autre pour DSF2.

27.6.4.7.3 Activer DSF2 Augment

Avant de pouvoir traiter avec DSF2, vous devez d'abord devenir un licencié certifié.

En plus des informations requises sur la société du client que vous avez saisies dans le groupe Informations de licence USPS, définissez les options suivantes pour effectuer un traitement DSF2 Augment :

1. Dans la transformation USA Regulatory Address Cleanse, ouvrez l'onglet *Options*.
2. Développez le groupe Rapport et analyses et définissez l'option *Générer les données de rapport* sur Oui.
3. Développez le groupe Fichiers de référence et saisissez le chemin pour les options *Chemin d'accès de Augment de DSF2*, *Chemin d'accès de DPV*, et *Répertoire eLOT*, ou utilisez la variable de substitution \$RefFilesAddressCleanse si vous l'avez configurée.
4. Dans le groupe Fichiers de référence également, saisissez un chemin pour l'option *Chemin d'accès du journal USPS*, ou utilisez la variable de substitution \$\$CertificationLogPath si vous l'avez configurée.
5. Facultatif. Développez le groupe Performance de la transformation et définissez les options *Cache DPV Directories* et *Cache DSF2 Augment Directories* sur Oui.
6. Développez le groupe Opérations d'affectation et définissez les options *Enable DSF2 Augment*, *Enable DPV*, et *Enable eLOT* sur Oui.
7. Incluez les champs de sortie des attributs d'adresse DSF2 dans la configuration du fichier de sortie

27.6.4.7.4 Champs de sortie DSF2

Lorsque vous effectuez un traitement DSF2 Augment dans le logiciel, les attributs d'adresse sont disponibles dans les champs de sortie suivants pour chaque adresse qui a été affectée. Assurez-vous d'inclure les champs qui contiennent les informations nécessaires pour la configuration du fichier de sortie :

- Indicateur_Gestion_DSf2
- Type_Livraison_DSf2
- Nombre_Dépôts_DSf2

- Indicateur_Dépôts_DSF2
- Ind_Enseignement_DSF2
- Indicateur_Conversion_LCAS_DSF2
- Type_Enregistrement_DSF2
- Indicateur_Saisonnier_DSF2
- Indicateur_Renvoi_DSF2

i Remarque

Une sortie vide dans l'un de ces champs signifie que l'adresse n'a pas été recherchée dans les répertoires DSF2.

27.6.4.7.5 Amélioration de la vitesse de traitement DSF2

Il est possible de cacher les données DSF2 pour améliorer la vitesse du traitement DSF2.

Pour cacher les données DSF2, définissez l'option *Mise en cache des répertoires DFS2 Augment* dans le groupe Performance de la transformation sur Oui. Le logiciel cache uniquement les répertoires nécessaires pour l'ajout d'attributs d'adresse.

27.6.4.8 Walk sequencing de DSF2

Lorsque vous exécutez le Walk sequencing de DSF2 dans le progiciel, celui-ci ajoute à vos données des informations de tournée de distribution que vous pouvez utiliser avec le logiciel de tri préalable pour bénéficier de remises sur les tournées.

→ N'oubliez pas

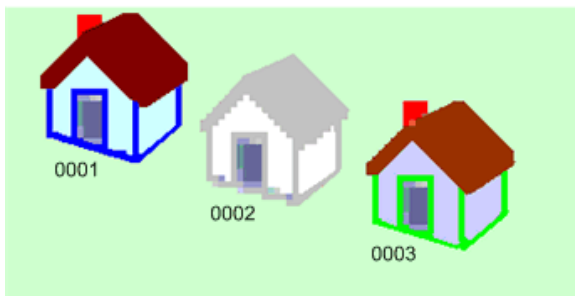
Le logiciel ne place pas les données dans un ordre de tournée.

Incluez la transformation DSF2 Walk Sequencer pour permettre le séquençement de tournée.

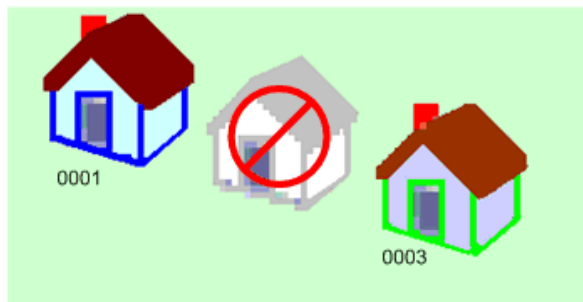
27.6.4.8.1 Pseudo séquençement

Le séquençement de tournée DSF2 est souvent appelé pseudo séquençement car il imite le séquençement de tournée d'USPS. Là où les numéros de tournée d'USPS couvrent toutes les adresses, le traitement DSF2 fournit des numéros de pseudo séquençement uniquement pour les adresses dans ce fichier particulier.

Walk sequencing includes every U.S. address, whether or not they're included in your list or mailing.



Pseudo sequencing includes only the addresses in your list. In the illustration below, assume the middle house is not included in your list.



Le progiciel utilise les données DSF2 pour affecter des numéros de tournée pour toutes les adresses qui sont des points de livraison confirmés par DPV (Statut_DPV = Y). Les autres adresses présentes dans le fichier de sortie qui ne sont pas des points de livraisons valides confirmés par DPV (Statut_DPV = S, N ou D) reçoivent le numéro de séquence "0000". Toutes les autres adresses ont un numéro de séquence vide.

i Remarque

Lorsque vous réalisez une tournée pour le courrier avec le logiciel, souvenez-vous des points suivants :

- (Batch uniquement) Le séquençement de tournée DSF2 est disponible uniquement en mode batch. Il est impossible d'affecter des numéros de séquence en temps réel.
- Retraiter si vous avez effectué des modifications de fichier. Si les données sont modifiées de quelque façon que ce soit, vous devez réaffecter des numéros de séquence. Les numéros de séquence sont uniquement valides pour le fichier de données au moment où vous le traitez.

27.6.4.9 Création de clé de décomposition

Les clés de décomposition créent des groupes gérables de données. Elles sont créées lorsqu'il y a deux champs ou plus à comparer.

La transformation DSF2 Walk Sequencer forme automatiquement des groupes de décomposition avant d'ajouter des informations de tournée aux données. Le logiciel crée des groupes de décomposition en fonction des champs Code postal1 et Itinéraire_par_code_de_tri.

Définissez les options pour que le logiciel configure comme vous le souhaitez les champs dans le groupe Configuration de collection de données. Le fait de conserver les paramètres par défaut optimise le flux de données et permet au logiciel de rendre les clés de décomposition cohérentes dans toutes les données.

Option	Valeur par défaut
Remplacer les valeurs nulles par un espace	Oui
Remplir la chaîne à droite par un espace	Oui

27.6.4.10 Activer Walk sequencing de DSF2

Pour activer la tournée DSF2, incluez la transformation DSF2 Walk Sequencer dans le flux de données.

Le fichier d'entrée pour la transformation DSF2 Walk Sequencer doit avoir été pré-traité avec le logiciel certifié par CASS (tel que la transformation USA Regulatory Address Cleanse). Pour obtenir une remise postale supplémentaire, incluez les informations du champs de sortie Indicateur_Commercial_DSf2 à partir du logiciel certifié par CASS.

Lorsque vous configurez le traitement de tournée DSF2, les options suivantes dans le groupe Informations de licence USPS sont requises :

- Nom du licencié
- ID du licencié DSF2
- ID de liste

En plus des champs Information de licence USPS requis, effectuez les paramétrages suivants dans la transformation DSF2 Walk Sequencer :

1. Facultatif. Sélectionnez Oui ou Non dans l'option *Exécuter en tant que processus distinct* dans le groupe Commun. Sélectionnez Non si vous regroupez des statistiques DSF2. Sélectionnez Oui pour enregistrer le temps de traitement (si vous n'avez pas besoin des statistiques DSF2).
2. Saisissez le chemin du fichier et le nom du fichier (dsf.dir) dans le répertoire Statistiques de livraison dans l'option *Répertoire Delstats* dans le groupe Fichiers de référence. Vous pouvez utiliser le paramètre de substitution \$\$RefFilesAddressCleanse si vous l'avez configuré.
3. Saisissez l'emplacement du site de traitement dans l'option *Emplacement du site* du groupe Options de traitement. Cette option s'applique uniquement si vous disposez de plus d'un emplacement de site pour le traitement DSF2.
4. Effectuez les paramétrages suivants dans le groupe Configuration de la collection de données :
 - Sélectionnez Oui ou Non dans l'option *Remplacer les valeurs nulles par un espace* selon vos préférences.
 - Sélectionnez Oui ou Non dans l'option *Remplir la chaîne à droite par un espace* selon vos préférences.
 - Sélectionnez Oui ou Non dans l'option *Données pré-triées* (facultatif). Nous recommandons de conserver le paramètre par défaut Non pour que Data Services trie les données en fonction des champs de la clé de décomposition (au lieu d'utiliser un autre programme logiciel).

27.6.4.11 Champs d'entrée de tournée DSF2

Ci-dessous se trouve une liste des champs d'entrée de la tournée DSF2.

i Remarque

Ces champs doivent avoir été sortis du traitement du logiciel certifié par CASS avant de pouvoir être utilisés comme entrées pour la transformation DSF2 Walk Sequencer :

- Code postal1
- Code postal 2
- Itinéraire_par_code_tri

- LOT
- Ordre_LOT
- Point_livraison
- Statut_DPV
- Indicateur_gestion_DSF2 (facultatif)

Le logiciel utilise les informations de ces champs pour déterminer la façon dont les enregistrements doivent être organisés (classés en fonction de la tournée) s'ils ont été utilisés dans une liste de diffusion. Le logiciel ne modifie pas physiquement l'ordre de la base de données. Le logiciel affecte des numéros de tournée à chaque enregistrement en fonction des informations qu'il a regroupées à partir de ces champs d'entrée.

i Remarque

Tous les champs sont requis excepté le champ Indicateur_Gestion_DSF2.

Le champ facultatif Indicateur_Gestion_DSF2 aide le logiciel à déterminer si l'enregistrement remplit les conditions pour les remises de saturation. Les remises de saturation sont déterminées par le pourcentage d'adresses résidentielles dans chaque itinéraire de transporteur. Consultez le *USPS Domestic Mail Manual* pour plus de détails sur tous les aspects du publipostage commercial et des remises de tri.

27.6.4.12 Champs de sortie de tournée DSF2

Le logiciel sort les informations sur les numéros de tournée vers les champs suivants :

- Remise_Livr_Active
- Remise_sat_résidentielle
- Remise_itinéraire_par_code_tri
- Remise_tournée
- Numéro_tournée

27.6.4.13 Reporting DSF2

Il existe des rapports et des fichiers journaux générés par le logiciel pour Augment et Walk sequencing de DSF2.

Trouvez les informations complètes sur ces rapports et fichiers journaux dans le guide *Management Console Guide*.

Rapport Delivery Sequence Invoice

USPS exige que vous soumettiez le rapport Delivery Sequence Invoice si vous demandez des remises de tournée DSF2 pour ce job.

US Addressing Report

- Le rapport US Addressing Report est généré par la transformation USA Regulatory Address Cleanse.
- Les sections Second Generation Delivery Sequence File Summary et Address Delivery Types du US Addressing Report affichent les nombres et pourcentages des adresses dans le fichier qui correspondent aux différentes catégories DSF2 (si NCOALink est activé). Les informations sont listées pour les pré-traitement et post-traitement NCOALink.

Fichier journal de statistiques DSF2 Augment

USPS exige que les licenciés DSF2 enregistrent les informations sur leurs traitements dans le fichier journal DSF2. USPS détermine les contenus du fichier journal DSF2 et exige que vous les lui soumettiez une fois par mois.

Les fichiers journaux sont à la disposition des utilisateurs avec les autorisations de l'administrateur et de l'opérateur.

27.6.4.13.1 Fichier journal de statistiques DSF2 Augment

Si vous créez une liste de publipostage certifiée DSF2, envoyez le fichier journal de statistiques DSF2 Augment à USPS.

La transformation génère et stocke le fichier journal de statistiques DSF2 Augment dans votre référentiel. Exportez le fichier journal à l'aide de la console de gestion Data Services.

Le logiciel utilise la convention d'appellation suivante pour le fichier journal : [ID_licencié_DSf2] [mm] [aa].dat

USPS détermine les contenus du fichier journal et requiert que vous le lui envoyiez le troisième jour ouvrable de chaque mois par e-mail. Pour en savoir plus sur ce processus, lisez le document *Conditions requises en matière de performances des licenciés DSF2*, disponible sur le site Web PostalPro d'USPS à l'adresse https://postalpro.usps.com/DSF2_LPR .

Rétention de fichier journal et suppression automatique

Le logiciel supprime périodiquement les fichiers journaux. Le paramètre par défaut est 30 jours. Si une période de 30 jours n'est pas suffisante, nous vous recommandons d'étendre le paramétrage par défaut en une période plus longue, par exemple 50 jours. En outre, USPS nécessite que vous conserviez les données USPS, afin de pouvoir les examiner sur plusieurs années. Par conséquent, nous vous recommandons d'exporter les fichiers journaux depuis votre référentiel vers un répertoire local, créé spécialement pour les données USPS. Le mieux est d'exporter ces fichiers tous les mois. En outre, cela a l'avantage de réduire l'espace occupé dans votre référentiel, ce qui permet un traitement plus efficace.

Vous pouvez prolonger la période de rétention des fichiers journaux dans la CMC (Central Management Console). Accédez aux paramètres de l'application Data Services. L'option est [période de rétention de l'historique](#).

Assurez-vous de vous familiariser avec les règles de rétention des informations USPS en effectuant un contrôle avec votre contrat de licence DSF2.

27.6.5 Présentation de NCOALink®

Le standard Move Update de l'USPS permet aux utilisateurs et à l'USPS de réduire le nombre d'enregistrements renvoyés en raison d'une adresse obsolète. NCOALink s'inscrit dans cette optique. La mise à jour de déplacements représente le processus de contrôle des adresses dans la base de données NCOA (National Change of Address) permettant de garantir la mise à jour des données avec les adresses actuelles.

Lors du traitement des données avec NCOALink, vous mettez à jour les enregistrements de particuliers ou de sociétés ayant déménagé et rempli un formulaire COA (Change of Address) avec l'USPS. Les autres programmes inclus dans Move Update et pris en charge dans la transformation USA Regulatory Address Cleanse comprennent notamment ANKLink® et SuiteLink®.

L'USPS exige la conformité de vos listes avec les standards Move Update afin de qualifier des tarifs postaux préférentiels pour les courriers prioritaires pré-triés. Vous pouvez répondre à cette exigence grâce au processus NCOALink.

i Remarque

Mover ID est le nom sous lequel SAP Data Services est certifié pour NCOALink.

Informations associées

[SuiteLink™ \[page 618\]](#)

27.6.5.1 Importance de la mise à jour des déplacements

L'USPS exige la mise à jour des déplacements pour tous les courriers prioritaires pré-triés. Pour aider les expéditeurs à répondre à cette exigence, l'USPS propose certaines options, notamment NCOALink.

La rétention d'informations d'adresse précises de vos contacts implique l'utilisation d'une méthode de l'USPS afin de recevoir les nouvelles adresses de vos contacts. La mise à jour des déplacements s'avère non seulement une pratique commerciale avisée, mais elle est également requise pour tous les expéditeurs prioritaires souhaitant des taux de prétriage ou d'automatisation. L'USPS développant ses exigences en matière de mise à jour des déplacements et renforçant les réglementations existantes, la mise à jour des déplacements deviendra de plus en plus importante.

27.6.5.2 Fonctionnement de NCOALink

Lorsque le traitement des adresses avec NCOALink est activé, le progiciel suit les étapes suivantes :

1. La transformation USA Regulatory Address Cleanse standardise les adresses d'entrée. NCOALink nécessite des données d'adresses standardisées et analysées en tant qu'entrées.
2. Le progiciel recherche dans la base de données NCOALink les enregistrements correspondant à vos enregistrements standardisés et analysés.
3. Si une correspondance est trouvée, le progiciel reçoit les informations de déplacement, notamment la nouvelle adresse, le cas échéant.
4. Le progiciel recherche les enregistrements de déplacement renvoyés de la base de données NCOALink afin d'affecter des codes postaux et autres codes.
5. Selon la sélection de classe de champ, le fichier de sortie contient :
 - L'adresse d'entrée originale. Il s'agit des valeurs complètes et correctes présentes dans les répertoires, standardisées selon toute définition configurée dans le groupe Options de standardisation dans l'onglet Options. (CORRECT)
 - Il s'agit des composants d'adresse mis à jour avec les données d'adresses contenant les mises à jour de déplacement. (DEMENAGEMENT MIS A JOUR)

i Remarque

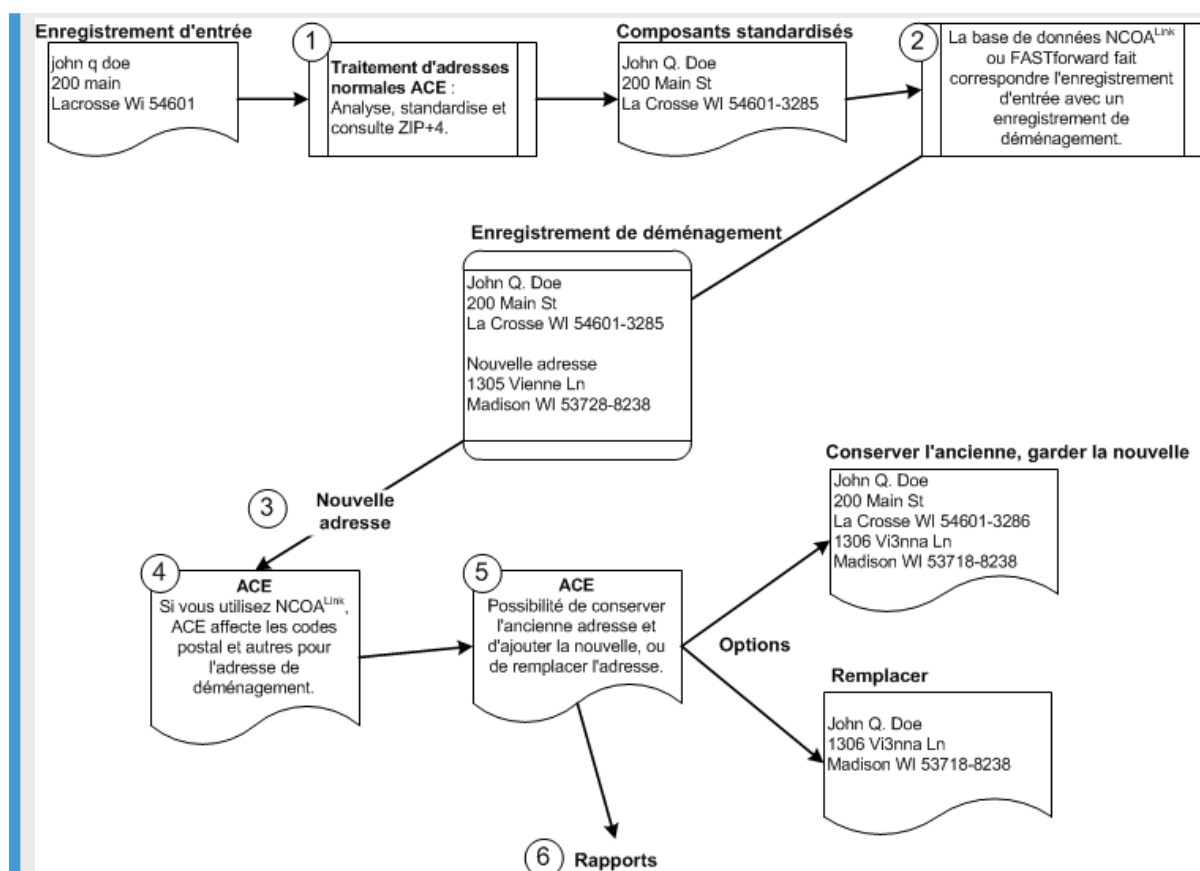
La transformation recherche les informations d'adresses contenant des déplacements mis à jour dans les Répertoires d'adresses nationaux américains. Lorsque l'adresse de déplacement mis à jour est introuvable dans les répertoires d'adresses nationaux américains, le progiciel remplit les champs Déplacement mis à jour avec les informations trouvées dans les répertoires Move Update uniquement. Les champs Déplacement mis à jour remplis suite à la standardisation des répertoires d'adresses nationaux américains ne seront pas mis à jour.

- Il s'agit des données d'adresses avec déplacement mis à jour si elles existent et correspondent aux répertoires d'adresses nationaux américains. Le champ peut aussi contenir les données d'adresses originales si aucun déplacement n'existe ou s'il ne correspond pas aux répertoires d'adresses nationaux américains. (MEILLEUR)

Selon la définition de l'option *Appliquer déplacement aux champs standardisés* dans le groupe NCOALink, les composants standardisés peuvent contenir des adresses originales ou avec les déplacements mis à jour.

6. Le progiciel génère les rapports et fichiers journaux requis pour la conformité USPS.

❖ Exemple



1. NCOALink nécessite des données d'adresses standardisés et analysées en tant qu'entrées. Ainsi, avant d'effectuer le traitement de NCOALink, le progiciel exécute son traitement normal des données d'adresses.
2. Le progiciel recherche dans la base de données NCOALink un enregistrement correspondant à votre enregistrement standardisé et analysé.
3. Le progiciel reçoit les informations de déplacement, notamment la nouvelle adresse, le cas échéant.
4. Le progiciel recherche l'enregistrement de déplacement renvoyé de la base de données NCOALink afin d'affecter des codes postaux et autres codes.
5. Selon votre choix, le progiciel peut conserver l'ancienne adresse et ajouter la nouvelle ou bien remplacer l'ancienne adresse par la nouvelle.
6. Le progiciel génère les rapports et fichiers journaux requis pour la conformité USPS.

27.6.5.3 Niveaux des fournisseurs NCOALink

Les utilisateurs de NCOALink appartiennent à l'une des trois catégories de fournisseurs suivantes. Spécifiez le fournisseur de services dans le groupe d'options Informations de licence USPS sous [Provider Level](#).

i Remarque

Seuls les niveaux de fournisseurs pris en charge dans vos codes clés enregistrés s'affichent dans la liste de sélection.

Niveau du fournisseur	Description
Full Service Provider (FSP)	Fournit le traitement de NCOALink à des tiers.
Limited Service Provider (LSP)	Fournit le traitement de NCOALink à des tiers et en interne.
End User Mailer (EUM)	Fournit le traitement de NCOALink à des listes en interne uniquement.

27.6.5.4 Gestionnaires et administrateurs de listes NCOALink

Un gestionnaire ou un administrateur de listes peuvent être propriétaires des listes de l'utilisateur de NCOALink traitées. Lorsqu'un gestionnaire ou un administrateur de listes est impliqué, ajoutez les informations de contact dans le groupe NCOALink sous ► [Liste de détails du contact](#) ► [Détails du contact](#) ►.

Gestionnaire

Transmet l'activité à un fournisseur de services NCOALink.

Administrateur de listes

L'administrateur de listes stocke et gère les listes. Les administrateurs de listes et les gestionnaires présentent deux différences majeures :

- Les administrateurs de listes n'envoient pas de fichiers des mises à jour de déplacements au propriétaire de la liste.
- Les administrateurs de listes peuvent disposer d'une licence NCOALink.

Si un administrateur de listes, un gestionnaire, ou les deux, sont impliqués dans un job, vous devez remplir une liste de détails du contact distincte pour chacun. Vous pouvez dupliquer un groupe d'options en cliquant avec le bouton droit de la souris sur le nom du groupe et en sélectionnant [Dupliquer l'option](#).

27.6.5.5 ANKLink (Address not known, adresse inconnue)

Chaque année, les courriers UAA (Undeliverable-as-addressed, adresse destinataire incorrecte) engendrent de nombreux frais pour l'USPS et l'industrie du traitement du courrier. Le progiciel fournit l'option NCOALink en tant que solution complémentaire pour les courriers UAA. NCOALink vous donne également accès aux données USPS's ANKLink.

27.6.5.5.1 Données ANKLink

ANKLink est un sous-ensemble de NCOALink. Vous pouvez demander des données ANKLink au NCSC (National Customer Support Center, centre national de support client) de l'USPS en composant le 1-800-589-5766 ou par e-mail à l'adresse ncoalink@usps.gov. ANKLink SAP ne fournit pas les données.

Le logiciel détecte toute utilisation de données ANKLink. Ainsi, vous ne devez pas spécifier l'utilisation d'ANKLink lors de la configuration d'un job.

27.6.5.5.2 Prise en charge d'ANKLink pour les niveaux de fournisseurs NCOALink

Le logiciel prend en charge trois niveaux de fournisseurs NCOALink définis par l'USPS. Les options du logiciel varient selon le niveau du fournisseur et sont activées suivant le package logiciel acquis. Le tableau suivant indique les niveaux de fournisseurs et la prise en charge :

Niveau du fournisseur	Fournit des services à des tiers	Données de COA (en mois)	Données reçues de l'USPS	Prise en charge d'ANKLink
Full Service Provider (FSP)	Oui Les services tiers doivent représenter au moins 51 % de l'ensemble du traitement.	48	Hebdomadaires	Non (aucun avantage)
Limited Service Provider (LSP)	Oui Les LSP peuvent fournir des services à des tiers et utiliser le produit en interne.	18	Hebdomadaires	Oui
End User Mailer (EUM)	Non	18	Mensuelles	Oui

→ Conseil

Si vous êtes un EUM NCOALink, vous pouvez demander un contrat Stop Processing Alternative de l'USPS. Une fois approuvé par l'USPS, vous pouvez acquérir la fonctionnalité Stop processing alternative du logiciel qui permet la continuité du traitement DPV et LACSLink après la détection d'un enregistrement d'adresse faux positif.

Informations associées

[Stop Processing Alternative \[page 601\]](#)

27.6.5.6 Performances du progiciel

Lors de nos tests, le progiciel s'exécutait plus lentement lorsque NCOALink était activé. La vitesse de traitement dépend de l'ordinateur exécutant le progiciel et du pourcentage d'enregistrements d'entrées affectés par un déplacement (un nombre plus élevé de déplacements équivaut à un ralentissement des performances).

Informations associées

[Optimisation des performances de traitement de NCOALink \[page 641\]](#)

27.6.5.7 Démarrage de NCOALink

Avant de démarrer le traitement de NCOALink, il convient d'effectuer les tâches suivantes :

- Terminer le processus de certification d'USPS pour devenir un utilisateur final ou un fournisseur de services NCOALink. Pour en savoir plus sur la certification, voir la section Certification NCOALink via le lien ci-dessous.
- Comprendre les stratégies de sortie disponibles et les options d'optimisation des performances.
- Configurer le job.

27.6.5.8 Solutions proposées par le USPS et SAP

NCOALink et les conditions requises pour les licences associées ont créé une nouvelle dimension dans la relation entre les expéditeurs (vous), l'USPS et les fournisseurs. Il est important de clarifier les solutions proposées par chacun.

27.6.5.8.1 Mise à jour des déplacements : une décision commerciale qui vous appartient

NCOALink permet de remplacer l'ancienne adresse d'une personne par sa nouvelle adresse. En tant que fournisseur de services, vous devez décider si vous acceptez les mises à jour de déplacements liées aux déménagements de familles ou d'une seule personne uniquement. L'USPS recommande de ne prendre cette décision qu'après une mûre réflexion sur vos relations client. Étudiez les exemples suivants :

- Si vous envoyez des chèques, des relevés de compte ou toute autre correspondance pour laquelle vous avez une responsabilité fiduciaire, la mise à jour des déplacements représente une garantie sérieuse. L'USPS préconise de vérifier chaque déplacement par l'envoi d'une carte postale double ou toute autre correspondance à réponse facile avant de modifier tout enregistrement financier sur la nouvelle adresse.

- Si vos relations commerciales concernent l'un des époux uniquement, la mise à jour des déplacements doit être gérée avec prudence, en respectant tout divorce ou séparation. Encore une fois, il peut être censé consacrer davantage de temps et d'argent à la confirmation de chaque déplacement avant de mettre à jour un enregistrement de manière permanente.

27.6.5.8.2 Exigences de sécurité de NCOALink

En raison de la sensibilité et de la confidentialité des données de changement d'adresse, l'USPS impose de strictes procédures de sécurité aux fournisseurs de logiciels utilisant et fournissant le traitement de NCOALink.

L'une des responsabilités incombant au fournisseur du logiciel consiste à contrôler que chaque entrée de liste pour la transformation USA Regulatory Address Cleanse contient au moins 100 enregistrements uniques. Ainsi, la transformation USA Regulatory Address Cleanse contrôle la présence de 100 enregistrements uniques dans le fichier de sortie. Ces contrôles augmentent le délai de vérification, mais sont requis par l'USPS et doivent être effectués.

Si le logiciel ne détecte pas 100 enregistrements uniques dans vos données, il signale une erreur et interrompt le traitement.

Le processus de contrôle de la présence de 100 enregistrements uniques représente une étape de pré-traitement. Ainsi, si le logiciel ne détecte pas ces 100 enregistrements uniques, le fichier d'entrée ne comprendra aucune sortie de statistique et aucun traitement ne sera effectué.

Informations associées

[Démarrage de NCOALink \[page 636\]](#)

27.6.5.8.2.1 Méthode de contrôle des 100 enregistrements uniques par le progiciel

Lorsque NCOALink est activé dans le job, le progiciel contrôle la présence de 100 enregistrements uniques avant d'effectuer tout traitement sur les données. Le progiciel contrôle l'ensemble de la base de données pour détecter les 100 enregistrements uniques. Si ces derniers sont détectés, le job est traité normalement. Toutefois, si le progiciel ne détecte pas ces enregistrements, il envoie une erreur indiquant que les données d'entrées ne contiennent pas 100 enregistrements uniques ou que les enregistrements ne sont pas suffisants pour déterminer une unicité.

Pour la recherche des 100 enregistrements uniques, un enregistrement représente tous les champs d'entrées mappés concaténés dans le même ordre que celui de leur mappage dans la transformation. Chaque enregistrement doit être identique à un autre enregistrement pour pouvoir être considéré comme semblable (non unique).

❖ Exemple

Comparaison d'enregistrements

L'exemple ci-dessous illustre la manière dont le progiciel concatène les champs dans chaque enregistrement et détermine les enregistrements non uniques. La première et la dernière ligne de cet exemple ne sont pas uniques.

332	FRONT	STREET	NORTH	LACROSSE	WI	54601
332	FRONT	STREET	SOUTH	LACROSSE	WI	54601
331	FRONT	STREET	SOUTH	LACROSSE	WI	54601
332	FRONT	STREET	NORTH	LACROSSE	WI	54601

27.6.5.8.2 Recherche d'enregistrements uniques dans plusieurs threads

Parfois, la liste d'entrée comprend 100 enregistrements uniques mais l'utilisateur reçoit toujours un message d'erreur indiquant que la liste n'inclut pas ces 100 enregistrements uniques. Cela peut se produire lorsque la liste contient un faible volume de données. Pour contourner ce problème, les utilisateurs peuvent ajuster la définition du degré de parallélisme (DOP) dans le job.

Faible volume de données et degré de parallélisme > 1

Lorsqu'un job NCOALink est configuré avec un degré de parallélisme supérieur à 1, chaque thread vérifie les enregistrements uniques présents dans la première collection traitée, puis partage avec les autres threads les connaissances acquises sur les enregistrements uniques détectés. Le premier thread ayant terminé le traitement de sa collection compte les enregistrements uniques trouvés par tous les threads jusqu'à ce point et décide si le contrôle des 100 enregistrements minimums est satisfaisant. Ce thread n'est pas nécessairement le thread 1. Par exemple, votre liste comporte 3 050 enregistrements et le degré de parallélisme est défini sur 4. Si le nombre d'enregistrements par collection est de 1 000, chaque thread aura une collection de 1 000 enregistrements, à l'exception du dernier thread qui n'en comportera que 50. Le thread traitant les 50 enregistrements terminera probablement sa collection en premier et prendra la décision d'échec ou de réussite avant de détecter les 100 enregistrements uniques. Vous pouvez terminer l'exécution de ce job en diminuant le degré de parallélisme défini. Dans cet exemple, vous pouvez le diminuer pour le définir sur 3.

27.6.5.9 A propos du fichier quotidien de suppression NCOALink

Si vous êtes un fournisseur de services, vous devez télécharger le fichier de suppression quotidien et l'installer dans le même dossier que les répertoires NCOALink avant d'exécuter quotidiennement le traitement de NCOALink.

Le fichier quotidien de suppression contient des enregistrements en attente de suppression à partir des données NCOALink. Par exemple, si Martine Dupond remplit un formulaire de changement d'adresse auprès de l'USPS puis ne déménage pas, son enregistrement se trouvera dans le fichier quotidien de suppression. Le changement d'adresse étant stocké dans les répertoires NCOALink mis à jour toutes les semaines ou tous les mois, le fichier quotidien de suppression est nécessaire pendant la période intermédiaire, jusqu'à la nouvelle mise à jour des répertoires NCOALink.

Téléchargez le fichier quotidien de suppression à partir du site Web de gestion du produit électronique de l'USPS à l'adresse <https://epf.usps.gov/> . Avant de pouvoir accéder au fichier, vous devez remplir et soumettre le formulaire Electronic Product Fulfillment Web Access Request Form (Form 5116) à l'USPS. Ce formulaire est disponible à partir du même site Web.

i Remarque

Si vous êtes un utilisateur final, seul le fichier quotidien de suppression est nécessaire pour traiter les fichiers Etape I ou Etape II. Il n'est pas requis pour le traitement normal de NCOALink.

Voici quelques points importants à connaître à propos du fichier quotidien de suppression :

- La vérification du progiciel échouera si NCOALink est activé, si un test d'étape est effectué et si le fichier quotidien de suppression n'est pas installé.
- La transformation USA Regulatory Address Cleanse prend uniquement en charge la version ASCII du fichier quotidien de suppression.
- Ne renommez pas le fichier quotidien de suppression. Il doit être nommé `dailydel.dat`.
- Le progiciel produira un avertissement de vérification si le fichier quotidien de suppression date de plus de trois jours.

27.6.5.9.1 Installation du fichier quotidien de suppression NCOALink

Pour télécharger et installer le fichier quotidien de suppression NCOALink, suivez les étapes suivantes :

1. Accédez au site USPS Electronic Product Fulfillment à l'adresse <https://epf.usps.gov/> .
2. Avant le premier téléchargement du fichier quotidien de suppression, vous devez remplir et envoyer par télécopie le formulaire PS Form 5116 (Electronic Product Fulfillment Web Access Request Form) au Service de licence de l'USPS (USPS Licensing Department).
Lors du remplissage du formulaire, assurez-vous de sélectionner l'une des options NCOALink ou NCOALink with ANKLink selon vos besoins. Vous pouvez alors accéder au fichier quotidien de suppression.
3. Connectez-vous au site USPS Electronic Product Fulfillment et téléchargez le fichier quotidien de suppression NCOALink [TEXTE] vers un emplacement à partir duquel le fichier `.tar` peut être extrait. Si les bloqueurs de fenêtres publicitaires sont activés sur le navigateur de l'ordinateur, vous devez les remplacer.
4. Extrayez le fichier `dailyDeletes_txt.tar`.
5. Copiez le fichier `dailydel.dat` vers l'emplacement de stockage de vos répertoires NCOALink.
6. Répétez les étapes 3 à 5 quotidiennement avant d'effectuer le traitement de NCOALink.

27.6.5.10 Stratégies de fichiers de sortie

Vous pouvez configurer le fichier de sortie pour répondre à vos besoins. Selon la Sélection de classe de champ définie, les composants de votre fichier de sortie contiennent les informations de type Correct, Déplacement mis à jour ou Meilleur :

- **CORRECT** : produit l'adresse d'entrée originale. Il s'agit des valeurs complètes et correctes présentes dans les répertoires, standardisées selon toute définition configurée dans le groupe Options de standardisation dans l'onglet Options. (CORRECT)
- **DEPLACEMENT MIS A JOUR** : produit les composants d'adresse mis à jour avec les données d'adresses contenant les mises à jour de déplacement.

i Remarque

La transformation recherche les informations d'adresses contenant des déplacements mis à jour dans les Répertoires d'adresses nationaux américains. Lorsque l'adresse de déplacement mis à jour est introuvable dans les répertoires d'adresses nationaux américains, le progiciel remplit les champs Déplacement mis à jour avec les informations trouvées dans les répertoires Move Update uniquement. Les champs Déplacement mis à jour remplis suite à la standardisation des répertoires d'adresses nationaux américains ne seront pas mis à jour.

- **MEILLEUR** : produit les données d'adresses avec déplacement mis à jour si elles existent et correspondent aux répertoires d'adresses nationaux américains. Le champ peut aussi contenir les données d'adresses originales si aucun déplacement n'existe ou s'il ne correspond pas aux répertoires d'adresses nationaux américains.

Selon la définition de l'option *Appliquer déplacement aux champs standardisés* dans le groupe d'options NCOA, les composants standardisés peuvent contenir des adresses originales ou avec les déplacements mis à jour.

Par défaut, l'option de sortie *Appliquer déplacement aux champs standardisés* est définie sur Oui et le progiciel met à jour les champs standardisés pour inclure les détails relatifs à l'adresse mise à jour disponible grâce à NCOALink.

Pour conserver les anciennes adresses dans les composants standardisés et ajouter les nouvelles adresses au fichier de sortie, vous devez définir l'option *Appliquer déplacement aux champs standardisés* sur Non. Vous pouvez ensuite utiliser les champs de sortie tels que NCOALINK_RETURN_CODE pour déterminer la survenance d'un déplacement. L'une des manières de configurer le fichier de sortie consiste à répliquer le format de fichier d'entrée, puis d'ajouter des champs supplémentaires pour les données de déplacement. Dans les enregistrements de sortie non affectés par un déplacement, la plupart des champs ajoutés seront vides. Il est également possible de créer un second fichier de sortie spécifique pour les enregistrements de déplacement. Deux méthodes sont applicables :

- Produire chaque enregistrement une seule fois, placer les enregistrements de déplacement dans le second fichier de sortie puis l'ensemble des autres enregistrements dans le fichier de sortie principal.
- Produire les enregistrements de sortie deux fois : la première fois dans le fichier de sortie principal et la seconde fois dans le second fichier de sortie.

Ces deux méthodes nécessitent l'utilisation d'un filtre de sortie déterminant si un enregistrement est un déplacement.

27.6.5.11 Optimisation des performances de traitement de NCOALink

Plusieurs facteurs affectent les performances lors du traitement de données NCOALink. Le facteur le plus critique est en général le volume d'accès au disque qui se produit. La plus efficace manière de réduire l'accès au disque consiste souvent à disposer d'une mémoire disponible suffisante pour les données de cache. D'autres facteurs critiques, tels que la vitesse du disque dur, le temps d'accès et le taux élevé de transfert, peuvent affecter les performances. Lorsque le délai d'accès au disque est réduit, les performances de l'UC deviennent significatives.

Informations associées

[Recherche d'enregistrements uniques dans plusieurs threads \[page 638\]](#)

27.6.5.11.1 Systèmes d'exploitation et processeurs

Les calculs impliqués dans la majeure partie du traitement du progiciel et de NCOALink sont parfaitement adaptés aux microprocesseurs intégrés dans la majorité des ordinateurs, tels que ceux fabriqués par Intel et AMD. Les processeurs RISC tels que ceux intégrés dans la plupart des systèmes UNIX sont en général considérablement plus lents pour ce type de calculs. En effet, un PC commun peut souvent exécuter un job unique via le progiciel et NCOALink environ deux fois plus rapidement qu'un système UNIX commun. Si vous recherchez une méthode rentable pour traiter des jobs uniques, un serveur Windows ou un poste de travail rapide peuvent produire d'excellents résultats. La majorité des systèmes UNIX intègrent plusieurs processeurs et peuvent au mieux traiter plusieurs jobs simultanément.

Vous devez pouvoir augmenter le degré de parallélisme (DOP) dans les propriétés du flux de données pour optimiser l'utilisation du cœur ou du processeur de votre système. L'augmentation du degré de parallélisme dépend de la complexité du flux de données.

27.6.5.11.2 Mémoire

Le traitement de NCOALink exploite un grand volume de données, exprimé en gigaoctets. Le volume exact dépend du niveau de votre fournisseur de services, du format de données et de la version spécifique des données de l'USPS.

En général, si les performances s'avèrent critiques et notamment si vous êtes un fournisseur de services complets NCOALink exécutant fréquemment de nombreux jobs volumineux avec des millions d'enregistrements, vous devez obtenir autant de mémoire que possible. Vous souhaitez mettre l'ensemble du jeu de données NCOALink en cache. Vous devriez pouvoir mettre l'ensemble du jeu de données NCOALink en cache en utilisant 20 Go de mémoire RAM, avec une mémoire restante suffisante pour le système d'exploitation.

27.6.5.11.3 Stockage de données

Dans la mesure du possible, le disque dur utilisé pour les données NCOALink doit être entièrement dédié à ce processus, au moins lorsque le job est en cours d'exécution. D'autres processus tentant d'utiliser ce même lecteur de disque physique peuvent largement réduire les performances de NCOALink.

Pour obtenir des taux de transfert encore plus élevés, vous pouvez opter pour l'utilisation d'un système RAID (redundant array of independent disks).

Lorsque le logiciel accède directement aux données NCOALink, et non par l'intermédiaire d'un cache, la fonctionnalité la plus significative du disque dur est son temps moyen d'accès.

27.6.5.11.4 Format de données

Le logiciel prend en charge les versions de fichiers plats et de hachage des données NCOALink. Si vous disposez d'une mémoire cache suffisante pour l'ensemble du jeu de données du fichier de hachage, ce format peut fournir les meilleures performances. Les données du fichier plat étant considérablement moins importantes, une part plus importante d'entre elles peut ainsi être mise en cache dans un volume donné de mémoire RAM. Toutefois, accéder aux données du fichier plat implique des recherches binaires nettement plus longues que l'accès direct utilisé avec le format de fichier de hachage.

27.6.5.11.5 Utilisation de la mémoire

La quantité de mémoire optimale dépend de nombreux facteurs différents. L'option "Automatique" s'avère en général utile pour décider de la quantité de mémoire à utiliser, mais un ajustement manuel de cette quantité peut être nécessaire dans de nombreux cas.

27.6.5.11.6 Astuces sur les performances

Plusieurs facteurs peuvent augmenter ou diminuer la vitesse de traitement de NCOALink. Il est possible d'en contrôler certains, mais d'autres peuvent être liés à votre activité. Prenez en considération les facteurs suivants :

- Taille de la mémoire cache : l'utilisation d'un volume insuffisant de mémoire pour la mise en cache NCOALink peut engendrer un accès au fichier aléatoire non nécessaire et des recherches de disque dur fastidieuses. L'utilisation d'un trop grand volume de mémoire peut engendrer la lecture de fichiers volumineux à partir du disque dans la mémoire cache même si une fraction infime de données seulement sera utilisée. Le volume de mémoire cache le plus adapté à votre environnement peut nécessiter certains tests afin de déterminer la meilleure option pour votre configuration et la taille de job classique.
- Emplacement du répertoire : il est préférable que les répertoires NCOALink se situent sur un disque SSD local ou sur un lecteur RAM virtuel. L'utilisation d'un disque SSD local ou d'un lecteur RAM virtuel élimine toutes les E/S pour NCOALink lors du traitement du job. Si les répertoires se trouvent sur un disque dur, il est préférable d'utiliser un disque dur local défragmenté. Le disque dur doit être uniquement utilisé pour les données NCOALink lors de l'exécution du job.

- Taux de correspondance : plus vous possédez d'enregistrements avec des déplacements transférables, plus le traitement sera lent. L'extraction et le décodage de nouvelles adresses prennent du temps, ainsi la mise à jour régulière d'une liste de diffusion optimisera la vitesse de traitement de cette liste.
- Format d'entrée : idéalement, vous devez fournir à la transformation USA Regulatory Address Cleanse des champs discrets pour le prénom, le deuxième prénom et le nom du destinataire, ainsi que le titre de civilité et la désignation. Si votre entrée ne présente qu'une seule ligne de nom, la transformation devra tout d'abord prendre le temps de l'analyser avant de contrôler les données NCOALink.
- Taille du fichier : les fichiers les plus volumineux se traitent relativement plus rapidement que les fichiers plus petits. Il existe une surcharge lors du traitement de tout job, mais si un job inclut des millions d'enregistrements, quelques secondes de surcharge deviennent insignifiantes.

27.6.5.12 Activation du traitement NCOALink

Vous devez avoir accès aux fichiers suivants :

- Répertoires NCOALink
- Version actuelle du fichier quotidien de suppression USPS
- Fichiers de données DPV
- Fichiers de données LACSLink

Si vous utilisez une copie de l'exemple de configuration de transformation, USARegulatoryNCOALink_AddressCleanse, NCOALink, DPV et LACSLink sont déjà activés.

1. Ouvrez la transformation *USA Regulatory Address Cleanse*, puis l'onglet *Options*.
2. Définissez les valeurs des options selon votre configuration.

Pour en savoir plus sur les champs de transformation USA Regulatory Address Cleanse, voir le *Guide de référence*. La table ci-dessous indique les champs obligatoires pour certains niveaux de fournisseurs.

Groupe d'options	Nom ou sous-groupe de l'option	Utilisateur final sans contrat pour le Stop Processing Alternative	Utilisateur final avec un contrat pour le Stop Processing Alternative	Fournisseur de services complets ou restreints (Full/Limited service provider)
Informations de licence USPS	<i>Nom du licencié</i>	oui	oui	oui
	<i>Code NAICS du propriétaire de la liste</i>	oui	oui	oui
	<i>ID de liste</i>	non	non	oui
	<i>Nom de la société du client</i>	non	oui	oui
	<i>Adresse de la société du client</i>	non	oui	oui
	<i>Localité de la société du client</i>	non	oui	oui
	<i>Région de la société du client</i>	non	oui	oui
	<i>Code postal de la société du client1</i>	non	oui	oui

Groupe d'options	Nom ou sous-groupe de l'option	Utilisateur final sans contrat pour le Stop Processing Alternative	Utilisateur final avec un contrat pour le Stop Processing Alternative	Fournisseur de services complets ou restreints (Full/Limited service provider)
	<i>Code postal de la société du client²</i>	non	oui	oui
	<i>Téléphone de la société du client</i>	non	non	non
	<i>Fréquence de traitement de liste</i>	oui	oui	oui
	<i>Date de réception de la liste</i>	non	non	oui
	<i>Date de retour de la liste</i>	non	non	oui
	<i>Niveau du fournisseur</i>	oui	oui	oui
NCOALink	<i>Sous-groupe Détails PAF</i>	non	non	Toutes les options sont obligatoires, à l'exception des options Nom de la société mère du client et Autre nom de la société du client.
	<i>Sous-groupe Options du fournisseur de services</i>	non	non	Toutes les options sont obligatoires, à l'exception des options Nom de la société de l'acheteur et Code postal du courrier.

→ Conseil

Si vous êtes un fournisseur de services et si vous devez fournir des détails de contact à plusieurs gestionnaires, développez le groupe NCOALink, cliquez avec le bouton droit de la souris sur [Détails du contact](#), puis cliquez sur [Dupliquer l'option](#). Un groupe supplémentaire de champs de détails de contact sera ajouté en-dessous du groupe original.

Informations associées

[A propos du fichier quotidien de suppression NCOALink \[page 638\]](#)

[Stratégies de fichiers de sortie \[page 640\]](#)

[Stop Processing Alternative \[page 601\]](#)

27.6.5.13 Fichiers journaux NCOALink

Le logiciel génère automatiquement les fichiers journaux requis par USPS et les nomme selon les exigences USPS. Il génère ces fichiers journaux dans le référentiel d'où vous pouvez les exporter en utilisant la Console de gestion.

Il crée un seul fichier journal par ID de licence. Au début de chaque mois, le logiciel lance de nouveaux fichiers journaux. Chaque fichier journal est ensuite ajouté avec les informations sur chaque job NCOALink traité lors de ce mois pour cet ID de licence spécifique.

Le logiciel produit les fichiers journaux suivants associés aux déplacements :

- Journal CSL (Customer Service log, journal de service client)
- Journal PAF d'informations clients (Processing Acknowledgement Form, formulaire de confirmation de traitement)
- Journal BALA (Broker/Agent/List Administrator, Gestionnaire/Agent/Administrateur de liste)

Les fichiers journaux PAF d'informations clients et BALA ne sont pas requis pour les utilisateurs finaux. Le tableau suivant indique les fichiers journaux requis pour les fournisseurs de services restreints ou complets :

Fichier journal	Description
CSL	Ce fichier journal contient un enregistrement par liste traitée. Chaque enregistrement détaille les résultats du traitement de changement d'adresse.
Journal PAF d'informations clients	Ce fichier journal contient les informations fournies pour le PAF. Le fichier journal répertorie chaque entrée PAF unique. Si une liste est traitée avec les mêmes informations PAF, celles-ci ne s'affichent qu'une seule fois dans le fichier journal. Lorsque les informations de contact pour l'administrateur de listes ont changé, les informations pour l'administrateur de listes ainsi que pour le gestionnaire correspondant sont écrites dans le fichier journal PAF.
BALA	Ce fichier journal contient toutes les informations de contact saisies pour le gestionnaire ou l'administrateur de listes. Le fichier journal répertorie une fois seulement les informations pour chaque gestionnaire ou administrateur de listes. USPS exige des fournisseurs de services le fichier journal Gestionnaire/Agent/Administrateur de listes, y compris dans les jobs n'impliquant pas de gestionnaire ou d'administrateur de listes. Le logiciel génère ce fichier journal pour chaque job si vous êtes un fournisseur de services certifié.

Rétention de fichier journal et suppression automatique

Vous devez soumettre les fichiers journaux NCOALink à l'USPS mensuellement. Le logiciel supprime les fichiers journaux sur une base périodique (30 jours par défaut), qui peut être contrôlée par les paramètres d'application de Data Services dans la Central Management Console. Pour éviter de perdre les informations du journal mensuellement, définissez la *Période de rétention de l'historique* à plus de 31 jours (nous recommandons un paramètre de 50 jours).

En plus d'envoyer tous les mois les fichiers journaux à l'USPS, vos données doivent être disponibles durant quelques années après le traitement du job pour que l'USPS les examine. (Assurez-vous d'être informé des règles d'USPS actuelles pour la rétention des données en contrôlant le contrat de licence USPS). Pour garantir que vous conservez tous les rapports et les journaux requis avant la suppression des données dans le référentiel, il est recommandé d'exporter les rapports et les journaux requis du référentiel vers un dossier local sur une base mensuelle. Cela permet également d'éviter que le contenu du référentiel ne devienne trop volumineux et ne génère l'« expiration » du processus d'exportation en raison du volume des statistiques conservées.

27.6.5.13.1 Log file retention and deletion

The USPS requires NCOALink users to retain the log files for several years, but the software deletes the log files after a short period of time.

As an NCOALink provider, the USPS requires you to submit NCOALink log files every month. However, the software deletes log files on a periodic basis after 30 days by default. You can change the period of time by setting the *History Retention Period* option in the Data Services Application Settings group in the Central Management Console (CMC). To avoid losing monthly log information, set the option to more than 31 days; we recommend a setting of 50 days.

In addition to sending monthly log files to the USPS, you are required to have the data available for the USPS to examine for several years after you process the job. Make sure you are aware of current USPS rules for data retention by checking your USPS licensing agreement. To ensure that you retain all required reports and logs before the data is deleted from the repository, we recommend that you export the required reports and logs to a local folder on a monthly basis.

Exporting the log files to a local folder also prevents the repository contents from becoming so large that the export process “times out” because the volume of statistics retained.

27.6.5.13.2 Noms des fichiers journaux

Le logiciel respecte le modèle d'attribution de noms de fichiers d'USPS pour les fichiers journaux suivants :

- Journal de services clients
- Journal PAF d'informations clients
- Journaux de l'administrateur de listes, de l'agent et du gestionnaire

Le tableau ci-dessous décrit le modèle d'attribution de noms des fichiers journaux NCOALink. Par exemple, P1234C10.DAT est un fichier journal PAF généré en décembre 2010 pour un licencié avec l'ID 1234.

Caractère 1		Caractères 2 à 5	Caractère 6		Caractères 7 à 8	Extension
Type de journal		ID de plate-forme	Mois		Année	
B	Journal du gestionnaire	Contient exactement quatre caractères	1	janvier	Deux caractères, par exemple 10 pour l'année 2010	.DAT
C	Journal de services clients		2	février		
P	Journal PAF		3	mars		
			4	avril		
			5	mai		
			6	juin		
			7	juillet		
			8	août		
			9	septembre		
			A	octobre		
			B	novembre		
			C	décembre		

Par exemple, P1234C10.DAT est un fichier journal PAF généré en décembre 2010 pour un licencié avec l'ID 1234.

27.6.6 USPS eLOT®

eLOT est disponible pour les enregistrements des États-Unis, uniquement dans la transformation USA Regulatory Address Cleanse.

eLOT donne une étape d'avance à Line of Travel. Le LOT d'origine a affiné l'itinéraire de distribution du facteur au niveau de l'immeuble (niveau Code postal2) en distinguant si une adresse est du côté pair ou impair d'une rue ou d'une route.

eLOT affine l'itinéraire de la tournée de distribution du facteur au niveau de la maison (point de distribution). Cela permet de trier vos publipostages jusqu'à un niveau plus précis.

Informations associées

[Activer eLOT \[page 648\]](#)

[Configuration des fichiers de référence \[page 574\]](#)

27.6.6.1 Activer eLOT

1. Ouvrez la transformation USA Regulatory Address Cleanse.
2. Ouvrez l'onglet *Options*, développez le groupe Options d'affectation et sélectionnez Oui pour l'option *Activer eLOT*.
3. Dans le groupe *Fichiers de référence*, définissez le chemin de votre répertoire eLOT.
Vous pouvez utiliser la variable de substitution \$\$RefFilesAddressCleanse pour cette option si vous l'avez configurée.

27.6.7 Early Warning System (EWS)

EWS aide à réduire le courrier mal adressé lorsque des points de livraison valides sont créés entre les mises à jour de répertoires nationaux. EWS est disponible pour les données des Etats-Unis, uniquement dans la transformation USA Regulatory Address Cleanse.

27.6.7.1 Présentation de EWS

La fonctionnalité EWS est la solution au problème du courrier mal dirigé causé par des points de livraison valides qui apparaissent entre les mises à jour du répertoire national. Par exemple, supposons que 300 Main Street est une adresse valide et que 300 Main Avenue n'existe pas. Un courrier adressé à 300 Main Avenue est affecté à 300 Main Street sur l'hypothèse que l'expéditeur s'est trompé sur le suffixe.

Supposons maintenant qu'une maison est terminée au 300 Main Avenue. Le nouveau propriétaire s'inscrit aux utilitaires et au courrier, mais cela peut prendre plusieurs mois avant que le point de livraison soit listé dans le répertoire national. Tout le courrier destiné à la nouvelle maison au 300 Main Avenue sera mal envoyé au 300 Main Street jusqu'à ce que le point de livraison soit ajouté au répertoire national.

La fonctionnalité EWS résout ce problème en utilisant un répertoire supplémentaire qui informe les utilisateurs CASS de l'existence du 300 Main Avenue bien longtemps avant qu'il apparaisse dans le répertoire national. Lorsque vous utilisez le traitement EWS, l'adresse anciennement mal dirigée reçoit par défaut une affectation à 5 chiffres.

27.6.7.2 Commencer par un exemple de configuration de transformation

Si vous souhaitez utiliser la transformation USA Regulatory Address Cleanse avec l'option EWS activée, le mieux est de commencer par l'exemple de configuration de transformation pour le traitement EWS appelé : USARegulatoryEWS_AddressCleanse.

27.6.7.3 Répertoire EWS

Le répertoire EWS contient quatre mois de données continues. Chaque semaine, l'USPS ajoute de nouvelles données et fait glisser une semaine d'anciennes données. L'USPS publie ensuite les dernières données EWS. Chaque vendredi, SAP convertit les données dans notre format `EWaammjj.zip` et les publie sur le SAP Software Download Center destiné aux utilisateurs professionnels à l'adresse <https://support.sap.com/swdc> .

1. Sélectionnez *Address Directories and Reference Data* (Répertoires d'adresses et Données de référence).
2. Cliquez sur *Address Directories* (Répertoires d'adresses).
3. Sélectionnez *Early Warning System Directories* pour accéder au téléchargement.

27.6.7.4 Activer EWS

EWS est déjà activé lorsque vous utilisez l'exemple de transformation EWS du logiciel `USARegulatoryEWS_AddressCleanse`. Ces étapes indiquent comment activer manuellement EWS.

1. Ouvrez la transformation `USA Regulatory Address Cleanse`.
2. Ouvrez l'onglet *Options* et développez le groupe `Assignment Options`.
3. Sélectionnez `Activer` pour l'option *Activer EWS*.
4. Développez le groupe de fichiers de référence et saisissez un chemin pour l'option *Répertoire EWS* ou utilisez la variable de substitution `$$RefFilesAddressCleanse` si vous l'avez configurée.

Informations associées

[Early Warning System \(EWS\) \[page 648\]](#)

27.6.8 USPS RDI®

L'option RDI (Residential Delivery Indicator) est disponible dans la transformation `USA Regulatory Address Cleanse`. RDI détermine si une adresse donnée correspond ou non à un domicile.

Les expéditeurs de colis peuvent trouver des informations RDI très précieuses parce que certains services de distribution facturent des taux supérieurs pour distribuer aux adresses résidentielles. D'autre part, USPS n'ajoute pas de suppléments pour les distributions résidentielles. Lorsque vous pouvez reconnaître une adresse comme étant un domicile, vous augmentez la motivation à expédier le colis avec USPS plutôt qu'avec un concurrent qui applique un supplément résidentiel.

Selon USPS, 91 % des adresses des Etats-Unis sont résidentielles. USPS incite à encourager l'utilisation de RDI par les expéditeurs de colis.

Vous pouvez utiliser RDI si vous traitez les données pour la certification CASS ou si vous effectuez un traitement en mode non certifié. En outre, RDI ne nécessite pas l'utilisation du traitement DPV.

27.6.8.1 Commencez par un exemple de transformation

Pour utiliser la fonctionnalité RDI avec la transformation USA Regulatory Address Cleanse, il est préférable de commencer par l'exemple de configuration de transformation USARegulatoryRDI_AddressCleanse.

Les exemples de transformations sont situés dans l'onglet Transformations de la bibliothèque d'objets. Cet exemple est situé sous transformations USA_Regulatory_Address_Cleanse.

27.6.8.2 Fonctionnement de RDI

Une fois que vous avez installé les répertoires RDI et que vous avez activé le traitement RDI, le logiciel détermine si l'adresse représentée par un code postal à 11 chiffres (code postal1, code postal2 et le DPBC) est une adresse résidentielle. (Le logiciel peut parfois procéder de la même façon avec un code postal2).

Le logiciel indique si une adresse correspond ou non à un domicile dans le composant de sortie RDI_INDICATOR.

L'utilisation de la fonctionnalité RDI implique quelques étapes seulement :

1. Installez les répertoires RDI.
2. Indiquez l'emplacement des répertoires.
3. Activez le traitement RDI dans le logiciel.
4. Exécutez le job.

Informations associées

[Activer RDI \[page 651\]](#)

27.6.8.3 Fichiers de répertoire RDI

SAP fournit les fichiers de répertoire RDI avec la mise à jour du répertoire national des Etats-Unis.

RDI nécessite les répertoires suivants :

Fichier	Description
rts.hs11	Pour des recherches de code postal à 11 chiffres (code postal2 plus DPBC). Ce fichier est utilisé lorsqu'une adresse contient un code postal à 11 chiffres. La détermination est basée sur le Delivery Point.

Fichier	Description
<code>rts.hs9</code>	Pour des recherches de code postal à 9 chiffres (code postal2). Ce fichier est un fichier ZIP+4. Ceci est possible uniquement lorsque les adresses de ce fichier ZIP+4 sont destinées à tous les domiciles ou à aucun domicile.

27.6.8.3.1 Spécification du chemin d'accès aux répertoires RDI

Dans le groupe de fichiers de référence, indiquez l'emplacement de vos répertoires RDI dans l'option [RDI Path](#). Si le traitement RDI est désactivé, le logiciel ignore le paramétrage [RDI Path](#).

27.6.8.4 Activer RDI

Si vous utilisez une copie de l'exemple de transformation `USARegulatoryRDI_AddressCleanse` dans votre flux de données, RDI est déjà activé. Cependant si vous commencez par une transformation `USA Regulatory Address Cleanse`, vérifiez que vous activez RDI et définissez l'emplacement pour les répertoires RDI suivants : `rts.hs11` et `rts.hs9`.

1. Ouvrez la transformation [USA Regulatory Address Cleanse](#).
2. Dans l'onglet [Options](#), développez le groupe de fichiers de référence et saisissez l'emplacement des répertoires RDI dans l'option [Chemin d'accès RDI](#) ou utilisez la variable de substitution `$RefFilesAddressCleanse` si vous l'avez configurée.
3. Développez le groupe d'options d'affectation et sélectionnez [Oui](#) pour l'option [Activer RDI](#).

27.6.8.5 Champ de sortie RDI

Pour RDI, le logiciel utilise toujours un composant sortant contenant toujours un seul caractère. Le composant `RDI` est renseigné uniquement lorsque la valeur `Oui` est attribuée à l'option [Activer RDI](#) dans le groupe d'options d'affectation.

Champ Job/Vues	Longueur	Description
<code>RDI_INDICATOR</code>	1	Ce champ contient la valeur <code>RDI</code> composée d'une des valeurs suivantes : <code>O</code> = L'adresse est une adresse de domicile.

Champ Job/Vues	Longueur	Description
		N = L'adresse n'est pas une adresse de domicile.

27.6.8.6 RDI dans rapports

Quelques rapports du logiciel disposent d'informations supplémentaires en raison de la fonctionnalité RDI.

27.6.8.6.1 CASS Instruction, USPS Form 3553

USPS Form 3553 contient une entrée pour le nombre de résidences. (L'enregistrement d'en-tête CASS contient également ces informations).

27.6.8.6.2 Fichiers de statistiques

Les fichiers de statistiques contiennent des comptages et des pourcentages.RDI

27.6.9 GeoCensus (USA Regulatory Address Cleanse)

L'option GeoCensus de la transformation USA Regulatory Address Cleanse fournit un codage géographique et de recensement permettant une analyse avancée des ventes et du marketing. Elle est disponible uniquement pour les enregistrements des Etats-Unis.

i Remarque

La fonctionnalité GeoCensus de la transformation USA Regulatory Address Cleanse sera obsolète dans une prochaine version. Il est recommandé de mettre à niveau tout flux de données utilisant actuellement la fonctionnalité GeoCensus pour utiliser la transformation Geocoder. Pour obtenir des instructions concernant la mise à niveau de GeoCensus vers la transformation Geocoder, voir le *Guide de mise à niveau*.

Informations associées

[Comment fonctionne GeoCensus \[page 653\]](#)

[Répertoires GeoCensus \[page 654\]](#)

[Activer le codage GeoCensus \[page 656\]](#)

27.6.9.1 Comment fonctionne GeoCensus

Avec l'utilisation de GeoCensus, la transformation USA Regulatory Address Cleanse peut ajouter la latitude, la longitude et les codes de recensement tels que le secteur de recensement ou la zone statistique métropolitaine (MSA) aux enregistrements, en fonctions des codes ZIP+4. MSA est une agrégation des comtés des Etats-Unis dans les zones statistiques métropolitaines affectées par le US Office of Management and Budget. Il est possible d'appliquer les codes GeoCensus lors de la normalisation d'adresses et l'affectation du code postal² pour un traitement simple, « en une seule passe ».

La transformation ne peut pas, par elle-même, ajouter des données démographiques aux enregistrements. La transformation établit les bases en donnant les coordonnées du recensement via des champs de sortie. Pour ajouter des informations démographiques, vous avez besoin d'une base de données démographique d'un autre fournisseur. Lorsque vous en obtenez une, nous vous recommandons d'utiliser le processus de correspondance pour faire correspondre les enregistrements avec la base de données démographique, et de transférer les informations démographiques dans les enregistrements (Vous utiliseriez les informations MSA et de secteur de recensement comme des critères de correspondance, puis la transformation Best Record pour enregistrer les revenus et d'autres informations.)

De la même façon, la transformation ne dessine aucune carte. Cependant, il est possible d'utiliser la latitude et la longitude affectées par la transformation comme une entrée pour les applications de mappage tiers. Ces applications permettent de tracer les emplacements des clients et de filtrer la base de données pour couvrir une zone géographique en particulier.

27.6.9.2 Le logiciel fournit les coordonnées de recensement

Le logiciel ne peut pas, par lui-même, ajouter des données démographiques aux enregistrements. Le logiciel établit seulement les bases en donnant les coordonnées du recensement via des champs de sortie. Pour ajouter des informations démographiques, vous avez besoin d'une base de données démographique d'un autre fournisseur. Lorsque vous en obtenez une, nous vous recommandons d'utiliser la transformation Match pour faire correspondre vos enregistrements avec la base de données démographique, et de transférer les informations démographiques dans vos enregistrements (en termes techniques, vous utilisez les informations MSA et de secteur/îlot de recensement comme des critères de correspondance, puis vous utilisez l'opération de post-comparaison du meilleur enregistrement dans la transformation Match pour transférer les revenus et d'autres informations.

De la même façon, le logiciel ne dessine aucune carte. Cependant, il est possible d'utiliser la latitude et la longitude affectées par le logiciel comme une entrée pour le logiciel de mappage tiers. Ces programmes permettent de tracer les emplacements des clients et de filtrer la base de données pour couvrir une zone géographique en particulier.

Informations associées

[Meilleur enregistrement \[page 538\]](#)

27.6.9.3 Obtenir le meilleur des données GeoCensus

Il est possible de combiner le GeoCensus avec la fonctionnalité du logiciel de mappage pour afficher les informations géoaméliorées. GeoCensus aide une entreprise à construire ses stratégies de vente et de marketing. Ci-dessous se trouvent quelques manières d'utiliser les données GeoCensus, avec ou sans produit de mappage.

Type d'informations	Comment GeoCensus peut aider
Etudes de marché	Il est possible d'utiliser les applications de mappage pour analyser la pénétration de marché, par exemple. Les sociétés qui s'efforcent d'obtenir une meilleure compréhension de leurs marchés utilisent des études de marché. De cette façon, elles peuvent voir les données de vente, de marketing et démographiques sur des cartes, tableaux et graphiques. Le résultat est un programme marketing mieux ciblé. Vous comprenez où se trouvent vos clients et la pénétration que vous avez acquise dans les marchés choisis.
Modélisation prédictive et marketing ciblé	Il est possible de cibler avec plus de précision les clients pour les campagnes de réponse directes à l'aide de sélections géographiques. La modélisation prédictive ou d'autres techniques d'analyses permettent d'identifier les caractéristiques du client idéal. Cette méthode incorpore les informations démographiques utilisées pour enrichir votre base de données client. A partir de cette analyse, il est possible d'identifier les meilleurs clients potentiels pour les programmes de publipostage ou de télémarketing.
Planification des médias	Pour mieux appuyer vos décisions publicitaires, il est possible que vous souhaitiez utiliser la planification des médias. Le fait d'associer un affichage visuel des marchés clés avec une vue des médias peut aider une société à utiliser de façon plus stratégique le budget publicitaire.
Gestion du territoire	Les données GeoCensus offrent une image du marché plus précise pour la société. Elles peuvent aider à distribuer les territoires et les quotas de vente plus équitablement.
Ventes directes	L'utilisation des données GeoCensus avec les outils d'étude de marché et le logiciel de mappage permet de suivre les clients potentiels regroupés à partir des activités marketing.

27.6.9.4 Répertoires GeoCensus

Les chemins et les noms de fichier pour les répertoires suivants doivent être définis dans le groupe d'options Fichiers de référence de la transformation USA Regulatory Address Cleanse avant de pouvoir commencer le traitement GeoCensus. Il est possible d'utiliser la variable de substitution \$\$RefFilesDataCleanse.

Nom de répertoire	Description
ageo1-10	Les répertoires GeoCensus de niveau d'adresse sont exigés si vous sélectionnez Adresse pour l'option Mode Geo sous le groupe Opérations d'affectation.
cgeo2.dir	Les répertoires GeoCensus de niveau centroïde sont exigés si vous sélectionnez Adresse pour l'option Mode Geo sous le groupe Opérations d'affectation.

27.6.9.5 Options du mode GeoCensus

Pour activer le GeoCensus dans la transformation, vous devez choisir un mode dans l'option [Mode Geo](#) dans le groupe Opérations d'affectation.

Mode	Description
Adresse	Traite le GeoCensus de niveau d'adresse uniquement.
Les deux	Essaye d'effectuer une affectation de GeoCensus de niveau d'adresse en premier. Si aucune affectation n'est réalisée, il essaye de faire une affectation de GeoCensus de niveau centroid.
Centroid	Traite le GeoCensus de niveau centroid uniquement.
Aucune	Désactive le traitement GeoCensus.

27.6.9.6 Champs de sortie GeoCensus

Vous devez inclure au moins l'un des champs de sortie générés suivants dans la transformation USA Regulatory Address Cleanse si vous prévoyez d'utiliser l'option GeoCensus :

- Code_Comté_AGeo
- AGeo_Latitude
- AGeo_Longitude
- AGeo_CodeMCD
- AGeo_Code_Emplacement
- AGeo_Code_Section
- Code_Stat_AGeo
- CGeo_CodeBSA
- CGeo_Latitude
- CGeo_Longitude
- CGeo_Code_Metro
- CGeo_Code_Section

Trouvez des descriptions de ces champs dans le *Guide de référence*.

27.6.9.7 Exemple de configuration de transformation

Pour traiter avec la fonctionnalité GeoCensus dans la transformation USA Regulatory Address Cleanse, le mieux est de commencer avec l'exemple de configuration de la transformation créé avec GeoCensus. Trouvez l'exemple de configuration, USARegulatoryGeo_AddressCleanse sous USA_Regulatory_Address_Cleanse dans la bibliothèque d'objets.

27.6.9.8 Activer le codage GeoCensus

Si vous utilisez une copie de l'exemple de fichier de transformation `USARegulatoryGeo_AddressCleanse` dans votre flux de données, GeoCensus est déjà activé. Cependant, si vous commencez par une transformation `USA Regulatory Address Cleanse`, vérifiez que vous définissez l'emplacement du répertoire et l'option [Mode Geo](#).

1. Ouvrez la transformation [USA Regulatory Address Cleanse](#).
2. Dans l'onglet [Options](#), développez le groupe [Fichiers de référence](#).
3. Définissez les emplacements pour les répertoires `cgeo.dir` et `ageo1-10.dir` en fonction du Mode Geo que vous sélectionnez.
4. Développez le groupe Opération d'affectation et sélectionnez [Adresse](#), [Centroid](#) ou [Les deux](#) pour l'option [Mode Geo](#).

Si vous sélectionnez [Aucun](#), la transformation n'exécutera pas le traitement GeoCensus.

Informations associées

[GeoCensus \(USA Regulatory Address Cleanse\) \[page 652\]](#)

27.6.10 Z4Change (USA Regulatory Address Cleanse)

L'option `Z4Change` est basée sur un répertoire USPS du même nom. L'option `Z4Change` est disponible uniquement dans la transformation `USA Regulatory Address Cleanse`.

27.6.10.1 Utilisation de Z4Change pour gagner du temps

L'utilisation de l'option `Z4Change` permet un gain de temps considérable en matière de traitement, par rapport à l'exécution de tous les enregistrements via le processus d'affectation classique ZIP+4.

`Z4Change` est plus rentable pour les bases de données importantes et stables, par exemple les bases de données de clients, abonnés réguliers etc. Dans nos tests, en fonction des fichiers où 5 % des enregistrements ont été concernés par une modification ZIP+4, la durée totale de traitement batch était du tiers du temps de traitement normal.

Lorsque vous utilisez la transformation de manière interactive, c'est-à-dire traitement d'une adresse à la fois, l'utilisation de `Z4Change` présente un moindre intérêt.

27.6.10.2 règles USPS

`Z4Change` doit être utilisé uniquement pour mettre à jour une base de données qui a subi précédemment un processus de validation complet. USPS nécessite que la liste de diffusion passe par un processus d'affectation complet tous les trois ans.

27.6.10.3 Répertoire Z4Change

Le répertoire `Z4Changez4change.dir` est mis à jour mensuellement, il est disponible uniquement si vous avez acheté l'option Z4Change pour la transformation USA Regulatory Address Cleanse.

Le répertoire Z4Change contient une liste de tous les codes ZIP et ZIP+4 dans le pays.

27.6.10.4 Initialisation par un exemple de transformation

Pour utiliser la fonctionnalité Z4Change dans la transformation USA Regulatory Address Cleanse, il est préférable de commencer par l'exemple de transformation `USARegulatoryZ4Change_AddressCleanse`.

27.6.10.5 Activer Z4Change

Si vous utilisez une copie du modèle de fichier de configuration de transformation Z4Change (`USARegulatoryZ4Change_AddressCleanse`) dans votre flux de données, Z4Change est déjà activé. Cependant si vous commencez à partir d'une transformation USA Regulatory Address Cleanse, vérifiez que vous définissez l'emplacement du répertoire et l'option *Mode Z4Change*.

1. Ouvrez la transformation *USA Regulatory Address Cleanse*.
2. Dans l'onglet *Options*, développez le groupe de fichiers de référence.
3. Activez l'emplacement du répertoire `z4change.dir` dans l'option *Répertoire Z4Change*.
4. Développez le groupe d'options Z4Change et sélectionnez *Oui* pour l'option *Activer Z4Change*.
5. Dans le groupe d'options Z4Change, saisissez le mois et l'année de la mise à jour la plus récente des enregistrements d'entrée ZIP+4 dans l'option *Dernière date d'affectation ZIP+4*.

27.6.11 Présentation des listes de propositions

Le traitement de liste de propositions est utilisé dans des projets opérationnels avec les transformations USA Regulatory Address Cleanse, Global Address Cleanse et Global Suggestion List

Utilisez des listes de propositions pour compléter et renseigner des adresses contenant un minimum de données. Les listes de propositions peuvent fournir des propositions pour les correspondances possibles si aucune correspondance exacte n'est trouvée. Cette section aborde uniquement les listes de propositions dans la transformation USA Regulatory Address Cleanse.

i Remarque

Le traitement de liste de propositions n'est pas disponible pour le traitement batch. En outre, si vous avez activé des listes de propositions, vous ne pouvez pas bénéficier de remises CASS et le logiciel ne créera pas la documentation CASS requise.

Informations associées

[Transformation Global Suggestion List \[page 598\]](#)

27.6.11.1 Introduction aux listes de propositions

Dans l'idéal, lorsque la transformation USA Regulatory Address Cleanse recherche une adresse dans les fichiers de référence USPS (Ville/ ZIP City File), elle recherche un seul enregistrement correspondant avec la combinaison correspondante localité/région/code postal. Puis lors de la recherche dans le répertoire national ZIP+4 USPS, le logiciel doit trouver un seul enregistrement correspondant à l'adresse.

Liens de séparation

Il est parfois impossible de relier une adresse d'entrée à un enregistrement correspondant dans le répertoire. Parfois, le logiciel peut trouver plusieurs enregistrements de répertoire correspondant à peu près aux données d'entrée.

Lorsque le logiciel ne trouve pas une correspondance parfaite, il constitue une liste des correspondances proches et les présente comme propositions. Lorsque vous choisissez une proposition, le logiciel essaie à nouveau d'affecter l'adresse.

❖ Exemple

Dernière ligne incomplète

La dernière ligne ci-dessous étant incomplète, le logiciel n'a pas pu sélectionner de façon fiable l'une des quatre localités. Mais si vous en sélectionnez une, le logiciel peut poursuivre le processus d'affectation.

Enregistrement d'entrée	Correspondances possibles dans les répertoires Ville/ ZIP City File
Ligne1=1000 vine	La Crosse, WI 54603
Ligne2lacr wi	Lancaster, WI 53813
	La Crosse, WI 54601
	Larson, WI 54947

❖ Exemple

Direction manquante

Le même processus peut se produire avec les lignes d'adresse. Une direction manquante constitue un problème commun. Dans l'exemple ci-dessous, la direction peut être aussi bien le Nord que le Sud. Le logiciel ne dispose d'aucune base pour choisir une direction plutôt qu'une autre.

Enregistrement d'entrée	Correspondances possibles dans le répertoire ZI+4

Ligne1 = 615 losey blvd	600-699 Losey Blvd N

Ligne2 = 54603	600-698 Losey Blvd S

❖ Exemple

Suffixe manquant

Un suffixe manquant pourrait se traduire par un comportement similaire à celui de l'exemple ci-dessus.

Enregistrement d'entrée	Correspondances possibles dans le répertoire ZI+4

Ligne1 = 121 dorn	100-199 Dorn Pl

Ligne2 = 54601	101-199 Dorn St

❖ Exemple

Noms de rue mal orthographiés

Un nom de rue mal orthographié ou incomplet pourrait également se traduire par une proposition d'adresses.

Enregistrement d'entrée	Correspondances possibles dans le répertoire ZI+4

Ligne1 = 4100 marl	4100-4199 Marshall 55421

Ligne2 = minneapolis mn	4100-4199 Maryland 55427

27.6.11.1.1 Informations supplémentaires requises

Lorsque le logiciel crée une liste de propositions, vous devez disposer d'une certaine base pour sélectionner l'une des correspondances possibles. Parfois, vous avez besoin de plus d'informations avant de choisir une proposition.

Exemple

- Les opérateurs qui prennent des informations par téléphone peuvent demander des informations supplémentaires au client pour choisir la liste de propositions.
- Les opérateurs qui saisissent des données à partir d'un bon de réduction un peu taché peuvent choisir une proposition en fonction des informations apparaissant sur la partie non tachée.

27.6.11.1.2 Règle CASS

USPS n'autorise pas SAP Data Services à générer un USPS Form 3553 lorsque les listes de propositions sont utilisées dans l'affectation des adresses. USPS suppose que les utilisateurs peuvent être tentés de faire des estimations, ce qui peut provoquer un mauvais acheminement du courrier, dont la gestion est coûteuse pour USPS.

Donc, quand vous utilisez la fonctionnalité de liste de propositions, vous ne pouvez pas obtenir de USPS Form 3553 couvrant les adresses que vous affectez. Le formulaire est disponible uniquement quand vous exécutez un traitement par lots avec l'option *Désactiver la certification* définie sur Non.

Vous devez effectuer un traitement par lots des adresses des processus en temps réel afin d'être en conformité avec CASS. Le logiciel génère ensuite un USPS Form 3553 qui couvre la totalité de la base de données de publipostage et votre liste peut être éligible aux remises d'affranchissement.

27.6.11.1.3 Intégration de fonctionnalité

La fonctionnalité Listes de propositions est conçue pour être intégrée à vos propres applications personnalisées via le service Web. Pour plus d'informations sur l'intégration de Data Services aux applications Web, reportez-vous au *Guide d'intégration*.

27.6.11.1.4 Exemple de modèle de listes de propositions

Pour utiliser la fonction de listes de propositions, il est préférable de commencer par l'un des modèles de transformations configurés spécifiquement. Le modèle de transformation est appelé `USARegulatorySuggestions_Address_Cleanse`. Il est disponible pour la transformation USA Regulatory Address Cleanse.

27.6.12 Rapport de statistiques de sources de données multiples

Statistiques basées sur des groupes logiques

Pour la transformation USA Regulatory Address Cleanse, une base de données d'entrée peut être une compilation de listes, chaque liste contenant un champ avec un identifiant unique. Cet identifiant unique peut être un nom, un nombre mais il doit se situer dans le même champ dans toutes les listes.

Le logiciel recueille des statistiques pour chaque liste à l'aide du champ d'entrée `Data_Source_ID`. Vous mappez le champ contenant l'identifiant unique dans votre liste, dans le champ d'entrée `Data_Source_ID` du logiciel. Lorsque le logiciel génère des rapports, certains de ces rapports contiendront un résumé pour toute la liste et un résumé distinct par liste en fonction de la valeur mappée dans le champ `Data_Source_ID`.

! Restriction

Afin de se conformer aux restrictions en matière de rapport NCOALink, la transformation USA Regulatory Address Cleanse ne prend pas en charge le traitement de plusieurs listes de publipostage associées aux différents PAF. Par conséquent, pour le traitement de NCOALink, tous les enregistrements du fichier d'entrée sont considérés comme une liste de publipostage unique et sont signalés comme tel dans le fichier Customer Service Log (CSL).

! Restriction

La fonctionnalité Regrouper les statistiques par source de données n'est pas prise en charge lorsque l'option *Activer uniquement l'analyse* ou *Activer uniquement Geo* dans le groupe des options non certifiées est activée sur Oui.

Informations associées

[Regroupement de statistiques par liste \[page 662\]](#)

27.6.12.1 Champ ID_Source_Données

Le logiciel effectue le suivi des statistiques pour chaque liste en fonction du champ d'entrée ID_Source_Données.

❖ Exemple

Dans cet exemple, cinq listes de publipostage sont combinées en une liste comme entrée dans la transformation USA Regulatory Address Cleanse. Chaque liste a un champ commun nommé ID_Liste et un identifiant unique dans le champ ID_Liste : N, S, E, W, C. Le mappage d'entrée est le suivant :

Nom du champ d'entrée de la transformation	Nom de la colonne du schéma d'entrée	Type
ID_SOURCE_DONNÉES	ID_LISTE	varchar(80)

Pour obtenir des statistiques DPV pour chaque ID_Liste, traitez le job puis ouvrez le rapport US Addressing.

La première section de résumé DPV du rapport US Addressing liste le Résumé cumulé, qui indique les totaux pour l'ensemble d'entrées complet. Les sections suivantes du résumé DPV listent les résumés par Data_Source_ID. L'exemple du tableau ci-dessous indique les comptages et les pourcentages de l'ensemble de la base de données (résumé cumulatif) et pour Data_Source_ID « N ».

Statistique	Comptage de résumé cumulé DPV	Résumé DPV pour % ID_Source_Données « N »	%
Adresses validées DPV	1 968	3,94	214 4,28
Adresses DPV non valides	3 032	6,06	286 5,72

Statistique	Comptage de résumé cumulé DPV	%	Résumé DPV pour ID_Source_Données « N »	%
Adresses validées CMRA	3	0,01	0	0,00
Adresses vacantes DPV	109	0,22	10	0,20
Adresses NoStats DPV	162	0,32	17	0,34

Informations associées

[Grouper les rapports de statistiques \[page 664\]](#)

27.6.12.2 Regroupement de statistiques par liste

Avant de configurer la transformation USA Regulatory Address Cleanse, pour regrouper des statistiques par liste, déterminez le champ qui identifie uniquement chaque liste. Par exemple, une liste de diffusion constituée de plusieurs sources doit contenir des listes avec un champ nommé LIST_ID qui identifie de manière unique chaque liste.

1. Ouvrez la transformation USA Regulatory Address Cleansedans le flux de données puis cliquez sur l'onglet [Options](#).
2. Développez le groupe Rapport et Analyse et sélectionnez Oui pour les options [Générer les données du rapport](#)et [Regrouper les statistiques par source de données](#).
3. Cliquez sur l'onglet [Entrée](#)puis sur le champ [Nom de la colonne du schéma d'entrée](#)en regard du champ Data_Source_ID pour identifier une liste de manière unique.
Un menu déroulant s'affiche.
4. Cliquez sur ce menu déroulant et sélectionnez le champ d'entrée à partir de la base de données que vous avez sélectionnée comme champ commun pour identifier une liste de manière unique. Dans le scénario ci-dessus, ce serait le champ LIST_ID.
5. Poursuivez avec les tâches de configuration de job restantes et exécutez votre job.

27.6.12.3 Champ source physique et Résumé cumulé

Certains rapports incluent un rapport par liste en fonction du champ ID_Source_Données (identifié dans le pied de page du rapport par « Champ source physique ») et un résumé de la liste entière (identifié dans le pied de page du rapport par « Résumé cumulé »). Cependant, les rapports Address Standardization, Address Information Code et USA Regulatory Locking n'incluent pas un résumé cumulé. Les enregistrements de ces rapports sont triés par la valeur ID Data Source.

i Remarque

Lorsque vous activez NCOALink, le logiciel indique un résumé par liste uniquement pour les sections suivantes du rapport COALink Processing Summary :

- NCOALink Move Type Summary
- NCOALink Return Code Summary
- ANKLink Return Code Summary

Cas particuliers

Dans certains cas, les termes « Résumé cumulé » et « Champ source physique » ne s'afficheront pas dans les sections de pied de page du rapport.

- Lorsque l'option *Regrouper les statistiques par source de données* est définie sur Non.
- Lorsque l'option *Regrouper les statistiques par source de données* est définie sur Oui et qu'une seule valeur d'ID de source de données figure dans la liste mais qu'elle est nulle.

27.6.12.3.1 Formulaire 3553 USPS et reporting de groupe

Le formulaire 3553 USPS comprend un résumé de la liste entière et un rapport par liste basé sur le champ ID_Source_Données.

❖ Exemple

Cumulative Summary

Le formulaire 3553 USPS désigne le résumé de la liste entière par les mots « Cumulative Summary ». Ils figurent dans le pied de page comme le montre le surlignage dans le modèle de rapport Cumulative Summary ci-dessous. En outre, le Cumulative Summary du formulaire 3553 USPS contient le nombre total de listes du job dans la section B, champ numéro 5, *Nombre de listes* (surligné ci-dessous).

4. List Name or ID No. (If using ID No., number must start with ID#) G		5. Number of Lists 15		6. Total Records Submitted for Processing 500	
C. Output					
Output Rating	1. Total Coded	2. Validation Period		Output Rating	1. Total Coded
a. ZIP + 4/DPV Confirmed	22125	From 6/29/2010	To 12/26/2010	d. 5-Digit Coded	54193
b. Z4Change Processed	0			e. CRRT Coded	46727
c. DirectDPV		From	To	f. eLOT Assigned	22125
D. Mailer					
I certify that the mailing submitted with this form has been coded (as indicated above) using CASS-Certified software meeting all of the requirements listed in the DIMM Section 708.			3. Name and Address of Mailer		
1. Mailer's Signature			2. Date Signed		
E. Qualitative Statistical Summary (QSS)					
For Informational Purposes Only: QSS is solely made available for the list processor's review and analysis. This information is not to be considered by the U.S. Postal Service® personnel in determining rate eligibility under any circumstances.					
High Rise Default 1704	High Rise Exact 9485	RR Default 30	RR Exact 467	LACS LACS 637	EWS 0
				Suite LACS 34	
Privacy Notice: For information regarding our Privacy Policy, visit USPS.COM®.					

PS Form 3553, September 2008

Cumulative Summary

Page 1 of 11

❖ Exemple

Physical Source Field

Le formulaire 3553 USPS désigne le résumé de chaque liste par les mots *Physical Source Fields* suivis de la valeur de l'ID de la source de données. Ils figurent dans le pied de page comme le montre le surlignage dans le modèle ci-dessous. Les données du rapport concernent uniquement cette liste.

4. List Name or ID No. (If using ID No., number must start with ID#) G		5. Number of Lists 1		6. Total Records Submitted for Processing 500		
C. Output						
Output Rating		1. Total Coded		2. Validation Period		
a. ZIP + 4/DPV Confirmed ▶		175		From 7/12/2010 To 1/8/2011		
b. Z4Change Processed ▶		0				
c. DirectDPV ▶				From To		
Output Rating		1. Total Coded		2. Validation Period		
d. 5-Digit Coded ▶		477		From 7/12/2010 To 7/12/2011		
e. CRRT Coded ▶		396		From 7/12/2010 To 10/10/2010		
f. eLOT Assigned ▶		175		From 7/12/2010 To 10/10/2010		
D. Mailer						
I certify that the mailing submitted with this form has been coded (as indicated above) using CASS-Certified software meeting all of the requirements listed in the DMM Section 708.			3. Name and Address of Mailer			
1. Mailer's Signature			2. Date Signed			
			A B C D			
E. Qualitative Statistical Summary (QSS)						
For Informational Purposes Only: QSS is solely made available for the list processor's review and analysis. This information is not to be considered by the U.S. Postal Service® personnel in determining rate eligibility under any circumstances.						
High Rise Default	High Rise Exact	RR Default	RR Exact	LACS Line	EWS	Suite Line
15	71	0	4	4	0	1
Privacy Notice: For information regarding our Privacy Policy, visit USPS.COM®.						
PS Form 3553, September 2008						
Physical Source Field: 1				Page 2 of 11		

27.6.12.3.2 Grouper les rapports de statistiques

Les rapports indiquant deux statistiques cumulées (résumés pour la liste de diffusion entière) et regroupant des statistiques (en fonction du champ source physique) incluent les rapports suivants :

- Résumé de validation d'adresse
- Résumé de type d'adresse
- US Addressing

Les rapports n'incluant aucun résumé cumulé comprennent les éléments suivants :

- Résumé du code d'information de l'adresse
- Normalisation de l'adresse
- US Regulatory Locking

Informations associées

[Champ ID_Source_Données \[page 661\]](#)

27.7 Tables de référentiel pour les statistiques de qualité de données

Il est possible de générer des statistiques de qualité de données pour les transformations Data Cleanse, Geocoder, et Global Address Cleanse.

Lorsque vous activez les définitions des statistiques de qualité de données dans la transformation et exécutez votre job, le logiciel envoie les statistiques aux tables dans le référentiel. En fonction de la table que vous choisissez de générer, les statistiques contiennent des informations résumées et non résumées concernant les processus de nettoyage et d'affectation.

Deux raisons principales dans la génération et l'utilisation de statistiques de qualité de données :

- Afficher les résultats agrégés et les informations résumées concernant la qualité de vos données.
- Explorer des enregistrements spécifiques que le logiciel a identifié comme de mauvaise qualité.

Activez les tables de statistiques de qualité de données dans la transformation applicable en effectuant les paramètres dans le groupe *Paramètres des statistiques Data Quality* sous *Rapport et analyse*.

Informations associées

[A propos du nettoyage de données \[page 426\]](#)

[Géocodage \[page 457\]](#)

[Nettoyage des données d'adresses globales \[page 570\]](#)

27.8 Prise en charge de Data Quality pour les types de données natives

Les transformations Data Quality traitent les types de données entrantes en tant que données de caractères avec quelques exceptions.

Ainsi, si un type de données non-caractères est mappé en tant qu'entrée, le logiciel convertit les données en chaîne de caractères avant de les transmettre par l'intermédiaire des transformations Data Quality.

Certains types de données Data Quality sont reconnus et traités en tant que types de données similaires à des entrées. Par exemple, si un champ de type de données est mappé vers un champ d'entrée de type de données Data Quality, le logiciel présente les avantages suivants :

- Tri : la transformation reconnaît et trie les données entrantes comme le type de données spécifié.
- Efficacité : la quantité de données converties en données caractères est réduite, optimisant le traitement.

27.8.1 Définitions du type de données Qualité des données

Les transformations Data Quality disposent de quatre attributs pour définir le champ :

- Nom
- Type
- Longueur
- Echelle

Ces attributs sont listés dans les onglets Entrée et Sortie de l'éditeur de transformation.

Dans l'onglet Entrée, l'attribut Nom est listé dans la colonne *Nom du champ d'entrée de la transformation*. Les attributs Type, Longueur et Echelle sont listés dans la colonne *Type* au format <type>(<longueur>, <échelle>).

L'onglet Sortie contient également les quatre attributs de champ listés ci-dessus. L'attribut Nom est listé dans la colonne *Nom_Champ*. Les attributs Type, Longueur et Echelle sont listés dans la colonne *Type* au format <type>(<longueur>, <échelle>).

27.9 Data Quality support pour valeurs nulles

Les transformations Data Quality traitent les valeurs nulles comme telles sauf si des données sont définies pour renseigner la zone à la sortie.

Une zone de VALEUR NULLE est transférée au traitement avec le marqueur VALEUR NULLE conservé sauf si des données sont définies pour renseigner la zone à la sortie. Lorsque des données sont définies, la zone est renseignée avec ces dernières au lieu de la VALEUR NULLE. L'avantage de ce processus est que les transformations Data Quality traitent un marqueur VALEUR NULLE comme étant inconnu au lieu de vide.

i Remarque

Si toutes les zones d'un enregistrement contiennent la VALEUR NULLE, la transformation ne traitera pas l'enregistrement et ce dernier ne fera pas partie des statistiques et des rapports.

Informations associées

[Guide de référence : langage de script, syntaxe de langage, valeurs nulles, valeurs nulles et chaînes vides \[page 712\]](#)

28 Query transform

Use the Query transform anywhere in a data flow to perform SQL type operations on source data such as grouping, joining, and ordering.

The Query transform serves many purposes based on the situation. Many of the examples we use throughout the Data Services documentation involve using a Query transform. For example, see how to use the Query transform to flatten a nested schema in the Tutorial located at https://help.sap.com/viewer/p/SAP_DATA_SERVICES.

Perform the following operations with the Query transform:

- Ensure that the output contains only distinct rows and no duplicates
- Construct inner and left outer joins using multiple input schemas, and include conditions and ranking
- Use joins proposed by the transform
- Output specific data based on filters or manual selection
- Map columns from input to output schemas
- Remap schemas when you change the data coming into the Query transform
- Perform transformations and functions on data
- Perform data nesting and unnesting
- Add new columns, nested schemas, and function results to the output schema
- Assign primary keys to output columns

[Entrée de schéma et Sortie de schéma \[page 667\]](#)

Les volets *Entrée de schéma* et *Sortie de schéma* déterminent les champs qui demeurent dans votre flux de données, ainsi que le mappage de champs.

[Schéma d'entrée de transformation Query \[page 669\]](#)

[Schéma de sortie de transformation Query \[page 671\]](#)

[Commandes de schémas de sortie \[page 672\]](#)

Lorsque vous cliquez avec le bouton droit de la souris sur un élément, comme une colonne, dans le volet *Schéma de sortie*, vous pouvez choisir à partir de plusieurs commandes pour contrôler le contenu de sortie.

[Recherche dans un schéma d'entrée ou de sortie \[page 674\]](#)

28.1 Entrée de schéma et Sortie de schéma

Les volets *Entrée de schéma* et *Sortie de schéma* déterminent les champs qui demeurent dans votre flux de données, ainsi que le mappage de champs.

Le logiciel vous aide à renseigner les volets des schémas dans la transformation Query en fonction des objets et des connexions dans votre flux de données. Vous pouvez basculer le focus dans l'*Éditeur de requêtes* pour modifier le schéma actuel. Vous pouvez également compléter automatiquement le volet *Sortie de schéma* avec un schéma de la table cible.

Il existe plusieurs manières de modifier le schéma actuel dans une transformation Query :

- Sélectionnez un schéma dans le volet *Sortie de schéma*.
- Faites un clic droit sur un schéma, une colonne ou une fonction dans le volet *Sortie de schéma* et sélectionnez *En faire l'élément actuel*.
- Double-cliquez sur un schéma dans le volet *Sortie de schéma*.

Pour compléter automatiquement un schéma de sortie vide avec un schéma de la table cible, connectez la transformation Query à une table cible dans votre flux de données. Si le volet *Sortie de schéma* est déjà renseigné, le logiciel ne remplace pas ces mappages.

Il existe plusieurs techniques pour mapper un schéma d'entrée dans un schéma de sortie :

- Glissez-déposez les colonnes ou les schémas imbriqués de la zone du schéma d'entrée vers la zone du schéma de sortie pour des mappages de colonnes simples. Si vous déposez une seule unique en plus d'une colonne existante, vous pouvez la remapper.
 - Sélectionnez *Remapper la colonne* pour mettre à jour uniquement le mappage de la colonne.
 - Sélectionner *Remapper avec le type de données* pour mettre à jour le mappage de la colonne avec un type de données.
 - Sélectionnez *Insérer au-dessus* ou *Insérer en dessous* pour ajouter une colonne comme nouveau mappage.
- Cliquez avec le bouton droit sur le schéma actuel dans le volet *Sortie de schéma* et sélectionnez *Nouvelle colonne de sortie* ou *Nouveau schéma de sortie*. Vous pouvez fournir des mappages de colonnes simples en faisant glisser des colonnes d'entrée vers les nouvelles colonnes de sortie. Pour les mappages complexes, utilisez la zone Options.
- Cliquez avec le bouton droit de la souris sur le schéma actuel dans le volet *Sortie de schéma* pour affecter et inverser des paramètres de clé primaire. Une icône de clé indique les clés primaires.
- Cliquez avec le bouton droit sur le schéma actuel dans le volet *Sortie de schéma* et sélectionnez *Nouvelle fonction*.
 - Seules les fonctions étant dans le référentiel sont disponibles.
 - Ajoutez des fonctions d'adaptateur et des fonctions RFC ou BAPI SAP. Ces types de fonctions renvoient plusieurs colonnes, contrairement aux fonctions utilisées dans les mappages et les clauses Where, qui renvoient des valeurs uniques.
 - Mappez tous les paramètres d'entrée de premier niveau de la fonction aux paramètres d'entrée de la requête.
- Cliquez avec le bouton droit sur le schéma actuel dans le volet *Sortie de schéma* et sélectionnez *Supprimer l'imbrication* pour aplanir les schémas de sortie imbriqués. Utilisez l'option *Supprimer l'imbrication* lorsque la source dispose d'un schéma imbriqué, comme un fichier XML. Vous mappez les colonnes à partir de cette source dans un schéma de table cible plat.
- Cliquez avec le bouton droit sur les éléments dans le volet *Sortie de schéma* et sélectionnez une commande. Les éléments doivent figurer dans le schéma actuel.

Rubrique parent : [Query transform \[page 667\]](#)

Informations associées

[Schéma d'entrée de transformation Query \[page 669\]](#)

[Schéma de sortie de transformation Query \[page 671\]](#)

[Commandes de schémas de sortie \[page 672\]](#)

[Recherche dans un schéma d'entrée ou de sortie \[page 674\]](#)

[Schéma de sortie de transformation Query \[page 671\]](#)

[Commandes de schémas de sortie \[page 672\]](#)

28.2 Schéma d'entrée de transformation Query

La zone Schéma d'entrée affiche l'entrée de tous les schémas pour la transformation Query sous forme d'arborescence hiérarchique. Chaque schéma d'entrée peut comporter aucun ou plusieurs des éléments suivants :

- Colonnes
- Schémas imbriqués

Les icônes précédant les colonnes sont des combinaisons des graphiques suivants :

Icône	Description
	Clé primaire.
	Colonne non utilisée dans le mappage de sortie.
	Colonne utilisée dans les mappages de sortie.

La liste [Entrée](#) située en haut à gauche de l'éditeur de requêtes indique le schéma actuellement sélectionné. Au fur et à mesure de la sélection de schémas ou de colonnes dans la zone Schéma d'entrée, la liste [Entrée](#) affiche le schéma correspondant. À l'inverse, vous pouvez sélectionner un schéma dans la liste [Entrée](#) pour vous déplacer facilement vers un schéma voulu.

Vous pouvez effectuer un clic droit sur des éléments de la zone du schéma d'entrée et sélectionner les commandes de menu suivantes :

Commande	Éléments applicables	Effet
Copier	Colonnes, schémas	Enregistre une copie des éléments sélectionnés dans le presse-papiers en les conservant dans la zone du schéma d'entrée.
Rechercher	N'importe où dans la zone du schéma d'entrée	Localise un élément de sortie correspondant au nom ou à la description saisi(e).
Actualiser	N'importe où dans la zone du schéma d'entrée	Actualise l'affichage de la zone du schéma d'entrée.
Parent	Colonnes	Sélectionne le schéma parent de la colonne sélectionnée.

Commande	Éléments applicables	Effet
Réduire	Colonnes, schémas	Réduit un schéma sélectionné ou le schéma parent d'une colonne sélectionnée (pour faciliter l'affichage/la navigation).
Générer le fichier DTD...		Génère un format DTD qui correspond à la structure du schéma sélectionné (NRDM ou relationnelle). Génère tous les types de données en varchar.
Générer le schéma XML...		Génère un schéma XML qui correspond à la structure du schéma sélectionné (NRDM ou relationnelle). Tous les types de données correspondent à ceux du schéma sélectionné.
Propager la colonne à partir de...	Colonnes	Déplace un schéma de colonne sélectionné d'une source ou transformation en amont vers le schéma d'entrée via des objets intermédiaires. Des mappages simples sont créés dans chaque objet sans modification du type de données ou des données elles-mêmes.
Mapper vers sortie	Colonnes, schémas	Crée un mappage simple de la zone du schéma d'entrée vers la zone du schéma de sortie.
Créer un format de fichier...		Crée un format de fichier à partir d'un schéma de table relationnelle. Tous les types de données correspondent à ceux du schéma de table d'origine.
Propriétés	Colonnes, schémas	Affiche les propriétés de l'élément sélectionné. Vous ne pouvez pas modifier les propriétés.

Rubrique parent : [Query transform \[page 667\]](#)

Informations associées

[Entrée de schéma et Sortie de schéma \[page 667\]](#)

[Schéma de sortie de transformation Query \[page 671\]](#)

[Commandes de schémas de sortie \[page 672\]](#)

[Recherche dans un schéma d'entrée ou de sortie \[page 674\]](#)

28.3 Schéma de sortie de transformation Query

La zone du schéma de sortie affiche la sortie de la transformation Query sous forme d'arborescence hiérarchique. Le schéma de sortie peut comporter un ou plusieurs des éléments suivants :

- Colonnes
- Schémas imbriqués
- Fonctions

Les icônes précédant les colonnes sont des combinaisons des graphiques suivants :

Icône	Description
	Clé primaire.
	Colonne comportant un mappage simple. Un mappage simple est une seule colonne ou une expression sans colonne d'entrée (c'est-à-dire une expression qui ne varie pas selon l'entrée).
	Colonne comportant un mappage complexe. Un mappage complexe est tout mappage autre que simple.
	(Croix rouge superposée sur une icône) Mappage incorrect. Data Services n'effectue pas de validation complète pendant la conception ; il est donc possible que l'éditeur ne marque pas un mappage incorrect. Pour obtenir une validation complète, sélectionnez ► Validation ► Valider ►.

Le volet Sortie de schéma affiche les éléments suivants :

- Le schéma actuel dans la liste Sortie de schéma en haut et dans la zone du schéma de sortie. Le schéma actuel détermine :
 - Les éléments de sortie pouvant être modifiés (ajout, mappage ou suppression).
 - Le périmètre des onglets de SELECT à ORDER BY dans la zone des options.
- Les schémas non actuels sont estompés dans la zone du schéma de sortie.

Rubrique parent : [Query transform \[page 667\]](#)

Informations associées

[Entrée de schéma et Sortie de schéma \[page 667\]](#)

[Schéma d'entrée de transformation Query \[page 669\]](#)

[Commandes de schémas de sortie \[page 672\]](#)

[Recherche dans un schéma d'entrée ou de sortie \[page 674\]](#)

[Entrée de schéma et Sortie de schéma \[page 667\]](#)

28.4 Commandes de schémas de sortie

Lorsque vous cliquez avec le bouton droit de la souris sur un élément, comme une colonne, dans le volet *Schéma de sortie*, vous pouvez choisir à partir de plusieurs commandes pour contrôler le contenu de sortie.

La table suivante contient des descriptions pour les commandes disponibles à la sélection.

Commande	Éléments applicables	Effet
<i>Couper</i>	Tout	Supprime les éléments sélectionnés du volet <i>Sortie de schéma</i> et enregistre une copie des éléments dans le presse-papiers.
<i>Copier</i>	Tout	Enregistre une copie des éléments sélectionnés dans le presse-papiers en les conservant dans le volet <i>Sortie de schéma</i> .
<i>Coller</i>	Tout	Insère les éléments enregistrés dans le presse-papiers à l'emplacement actuel du curseur (doit figurer dans le schéma actuel). Visible uniquement si le presse-papiers contient des éléments. Si le curseur chevauche une colonne existante, vous êtes invité à insérer au-dessus, insérer en dessous, remapper la colonne ou annuler.
i Remarque Le fait de copier un élément d'entrée et de le coller dans le volet <i>Sortie de schéma</i> procure un mappage simple de l'élément d'entrée à l'élément de sortie. Vous pouvez également glisser l'élément d'entrée vers le volet <i>Sortie de schéma</i>.		
<i>Supprimer</i>	Tout	Supprime les éléments sélectionnés du volet <i>Sortie de schéma</i> (sans effectuer de copie).
<i>Rechercher...</i>	Tout	Localise un élément de sortie correspondant au nom ou à la description saisi(e).
<i>En faire l'élément actuel</i>	Tout sauf le schéma actuel	Fait du schéma sélectionné ou du schéma de l'élément sélectionné le schéma actuel.
<i>Nouvelle colonne de sortie...</i>	Schémas	Ajoute au schéma actuel une colonne de sortie ayant le nom et les propriétés saisis.
<i>Nouveau schéma de sortie...</i>	Schémas	Ajoute au schéma actuel un schéma imbriqué ayant le nom saisi.

Commande	Éléments applicables	Effet
<i>Nouvel appel de fonction...</i>	Schémas	Ajoute au schéma actuel un appel de fonction ou de procédure stockée. La fonction ou la procédure doit déjà être importée dans le référentiel. Vous pouvez ajouter des fonctions d'adaptateur et des fonctions RFC ou BAPI SAP. Ces fonctions renvoient plusieurs colonnes, contrairement aux fonctions utilisées dans les mappages et les clauses Where, qui renvoient des valeurs uniques. Dans la fenêtre <i>Définir les paramètres d'entrée</i>, mappez tous les paramètres d'entrée de premier niveau de la fonction aux paramètres d'entrée de la requête.
<i>Modifier l'appel de fonction</i>	Fonctions	Permet de modifier la fonction sélectionnée. Dans la fenêtre <i>Définir les paramètres d'entrée</i>, mappez tous les paramètres d'entrée de premier niveau de la fonction aux paramètres d'entrée de la requête.
<i>Propager la colonne à partir de</i>	Colonnes	Déplace un schéma de colonne sélectionné d'une source ou transformation en amont vers le schéma de sortie via des objets intermédiaires. Des mappages simples sont créés dans chaque objet sans modification du type de données ou des données elles-mêmes.
<i>Supprimer l'imbrication</i>	Schémas imbriqués	Le schéma sélectionné est aplani ou réimbriqué dans le schéma parent de la sortie de la requête. Un schéma non imbriqué n'apparaît pas dans la transformation ou la cible suivante ; seules ses colonnes apparaissent. Les schémas non imbriqués s'affichent dans le volet <i>Sortie de schéma</i> de la transformation Query sous forme d'icônes de table avec une flèche noire pointant vers la gauche.
<i>Imbriquer avec sous-schémas</i>	Schémas imbriqués	Ré-imbrique le schéma sélectionné et tous les sous-schémas dans le schéma parent de la sortie de la requête.
<i>Supprimer l'imbrication avec sous-schémas</i>	Schémas imbriqués	Aplanit le schéma sélectionné et tous les sous-schémas dans le schéma parent de la sortie de la requête. Les schémas et sous-schémas non imbriqués n'apparaissent pas dans la transformation ou la cible suivante ; seules les colonnes apparaissent. Les schémas et sous-schémas non imbriqués s'affichent dans le volet <i>Sortie de schéma</i> de la transformation Query sous forme d'icônes de table avec une flèche noire pointant vers la gauche.
<i>Clé primaire</i>	Colonnes	Active l'attribut de clé primaire de la colonne (une coche s'affiche en regard de la commande) ou le désactive (aucune coche ne s'affiche). Une icône de clé indique qu'une colonne est une clé primaire.
<i>Facultatif</i>	Schémas	Active/désactive un schéma pour le rendre facultatif.

Commande	Éléments applicables	Effet
Générer le fichier DTD	Tout	Génère un format DTD qui correspond à la structure du schéma sélectionné (NRDM ou relationnelle). Génère tous les types de données en varchar.
Générer le schéma XML	Schéma	Génère un schéma XML qui correspond à la structure du schéma sélectionné (NRDM ou relationnelle). Tous les types de données correspondent à ceux du schéma sélectionné.
Créer un format de fichier	Schéma	Crée un format de fichier à partir d'un schéma de table relationnelle. Tous les types de données correspondent à ceux du schéma de table d'origine.
Propriétés	Tout	Affiche les propriétés de l'élément sélectionné.

Rubrique parent : [Query transform \[page 667\]](#)

Informations associées

[Entrée de schéma et Sortie de schéma \[page 667\]](#)

[Schéma d'entrée de transformation Query \[page 669\]](#)

[Schéma de sortie de transformation Query \[page 671\]](#)

[Recherche dans un schéma d'entrée ou de sortie \[page 674\]](#)

28.5 Recherche dans un schéma d'entrée ou de sortie

1. Dans l'onglet Rechercher de l'éditeur de requêtes, saisissez le terme de recherche dans la zone [Rechercher](#) ou sélectionnez-le dans les précédents termes de recherche de la liste déroulante.
2. Dans la liste [Schémas](#), sélectionnez les schémas dans lesquels effectuer la recherche.
3. Dans la liste [Éléments](#), sélectionnez les types de mappage dans lesquels effectuer la recherche.
4. Dans la liste [Where](#), sélectionnez les propriétés dans lesquelles effectuer la recherche.

i Remarque

Vous pouvez rechercher dans une propriété ou dans toutes, mais pas dans deux ou trois propriétés spécifiques à la fois.

5. Cochez la case [Respecter la casse](#) pour limiter la recherche selon l'emploi des majuscules dans les termes entrés.
6. Cliquez sur [Rechercher](#).
Designer recherche dans la requête les mots spécifiés en respectant les contraintes définies.

i Remarque

Designer recherche les colonnes chargées dans la mémoire. Si les colonnes ne sont pas chargées dans la mémoire, développez le schéma pour charger les colonnes dans la mémoire avant de cliquer sur [Rechercher](#) et de rechercher les colonnes.

Toutes les correspondances sont affichées dans la zone située sous les contraintes de la recherche. Lorsque vous cliquez pour sélectionner un nom de table ou de colonne, celle-ci est automatiquement mise en surbrillance et affichée dans la zone du schéma d'entrée ou de sortie correspondant.

Initialement, Designer liste les colonnes correspondantes dans l'ordre où elles apparaissent dans les schémas. Si vous recherchez dans les deux schémas, Designer liste d'abord la première correspondance trouvée dans le schéma d'entrée et en dernier la dernière correspondance trouvée dans le schéma de sortie. Vous pouvez trier la liste des correspondances par propriété. Chaque fois que vous cliquez sur un intitulé de propriété, Designer retreuve les correspondances, en parcourant l'ordre d'origine, l'ordre croissant et l'ordre décroissant.

Les icônes en forme de flèches confirment le type de colonne et de tri. Par exemple, si vous triez les données selon la propriété [Description](#) et par ordre croissant, une flèche vers le haut apparaît en regard de l'intitulé [Description](#). Cliquez à nouveau sur l'intitulé et une flèche vers le bas apparaît pour indiquer que les données sont maintenant triées par ordre alphanumérique décroissant. Cliquez à nouveau et la liste des correspondances reprend son ordre de tri d'origine.

Présentation des tâches : [Query transform \[page 667\]](#)

Informations associées

[Entrée de schéma et Sortie de schéma \[page 667\]](#)

[Schéma d'entrée de transformation Query \[page 669\]](#)

[Schéma de sortie de transformation Query \[page 671\]](#)

[Commandes de schémas de sortie \[page 672\]](#)

29 Fonctions et procédures

Les fonctions et les procédures sont identiques, car elles reprennent toutes deux les valeurs d'entrée. Cependant, elles se distinguent de par l'utilisation qu'elles font des données.

- Les fonctions prennent les valeurs d'entrée et produisent une valeur renvoyée.
- Les procédures prennent les valeurs d'entrée et effectuent un ensemble d'opérations sans renvoyer de valeur spécifique.

Les valeurs de sortie comprennent les éléments suivants :

- Paramètres transmis dans un flux de données
- Valeurs d'une colonne de données
- Variables définies à l'intérieur d'un script

[A propos des fonctions \[page 676\]](#)

Dans Data Services, les fonctions prennent les valeurs d'entrée et produisent une valeur renvoyée.

[À propos des procédures \[page 690\]](#)

SAP Data Services prend en charge les procédures stockées, qui désignent des objets exécutables ou des entités nommées que vous stockez dans une base de données.

29.1 A propos des fonctions

Dans Data Services, les fonctions prennent les valeurs d'entrée et produisent une valeur renvoyée.

Certaines fonctions produisent des valeurs identiques ou similaires à celles des transformations. Toutefois, les fonctions opèrent sur des valeurs uniques et les transformations sur des jeux de données. Utilisez des fonctions seules ou incluez les fonctions dans les expressions. Data Services offre de nombreuses fonctions intégrées. Vous pouvez également créer des fonctions personnalisées en utilisant le langage de script et l'éditeur rapide de Data Services.

Data Services peut exécuter des conversions de types de données implicites sur les valeurs de date, heure, date/heure et d'intervalle.

Pour plus d'informations sur l'utilisation des fonctions, consultez le *Guide de Designer*.

[Fonctions comparées aux transformations \[page 677\]](#)

Certaines fonctions peuvent produire des valeurs identiques ou similaires à celles des transformations. Toutefois, les fonctions et les transformations opèrent dans un périmètre différent.

[Opération d'une fonction \[page 677\]](#)

L'opération de la fonction détermine l'endroit où vous pouvez appeler la fonction.

[Arithmétique dans les fonctions de dates \[page 678\]](#)

Le logiciel exécute des conversions de types de données implicites sur les valeurs de date, heure, date/heure et d'intervalle.

[Insertion de fonctions dans les expressions \[page 678\]](#)

Cette section vous explique où et comment inclure des fonctions dans les expressions.

[Fonctions personnalisées \[page 681\]](#)

[Fonctions intégrées \[page 685\]](#)

Informations associées

[Éditeur rapide et outils fonctionnels \[page 252\]](#)

[Langage de script \[page 706\]](#)

29.1.1 Fonctions comparées aux transformations

Certaines fonctions peuvent produire des valeurs identiques ou similaires à celles des transformations. Toutefois, les fonctions et les transformations opèrent dans un périmètre différent.

- Les fonctions opèrent sur des valeurs uniques, comme les valeurs de colonnes spécifiques dans un jeu de données.
- Les transformations opèrent sur des jeux de données en créant, en mettant à jour et en supprimant des lignes de données.

SAP Data Services ne prend pas en charge les fonctions incluant des tables comme paramètres d'entrée ou de sortie, à l'exception de fonctions importées à partir d'Applications SAP.

29.1.2 Opération d'une fonction

L'opération de la fonction détermine l'endroit où vous pouvez appeler la fonction.

Par exemple, la fonction `Lookup database` opère en tant que fonction itérative. La fonction de recherche peut mettre en cache des informations sur la table et les colonnes si elle opère entre des appels de fonction. Au contraire, les fonctions de conversion, comme `to_char`, opèrent comme des fonctions sans état. Les fonctions de conversion opèrent indépendamment à chaque itération. Une fonction d'agrégat, comme `max`, nécessite un set de valeurs pour opérer. Ni la fonction `lookup` (itérative) ni la fonction `max` (agrégat) ne peuvent être appelées à partir d'un script ou d'une structure conditionnelle où le contexte ne prend pas en charge l'opération de ces fonctions.

Le type de fonction détermine l'endroit où une fonction peut être utilisée :

Type	Description
<i>Agrégat</i>	Génère une seule valeur à partir d'un set de valeurs. Les fonctions d'agrégat, comme max, min et count, utilisent le jeu de données spécifié par l'expression dans l'onglet <i>Group by</i> d'une requête. Peuvent être appelées uniquement à partir d'une transformation Query, et non à partir de fonctions ou scripts personnalisés.
<i>Itératives</i>	Maintient les informations sur l'état d'une invocation à l'autre. La durée de vie des informations sur l'état d'une fonction itérative correspond à la durée de vie de l'exécution de la requête dans laquelle elles sont incluses. La fonction lookup est une fonction itérative. Peuvent être appelées uniquement à partir d'une transformation Query, et non à partir de fonctions ou de scripts.
<i>Sans état</i>	Les informations sur l'état ne sont pas maintenues d'une invocation à l'autre. Les fonctions sans état comme to_char ou month peuvent être utilisées partout où les expressions sont autorisées.

29.1.3 Arithmétique dans les fonctions de dates

Le logiciel exécute des conversions de types de données implicites sur les valeurs de date, heure, date/heure et d'intervalle.

Informations associées

[Calcul de date \[page 855\]](#)

29.1.4 Insertion de fonctions dans les expressions

Cette section vous explique où et comment inclure des fonctions dans les expressions.

Vous pouvez utiliser des fonctions dans les objets suivants :

- Transformations (Query, Case, SQL, Map_Operation)
- Objets de script
- Structures conditionnelles
- Autres fonctions personnalisées

Avant d'utiliser une fonction, vous devez déterminer si l'opération correspondante est pertinente dans l'expression que vous créez.

Exemple :

- La fonction max ne peut pas être utilisée dans un script ni une structure conditionnelle, qui ne comprend aucune valeur à traiter.

- Des paramètres peuvent être générés par un workflow, mais non par un flux de données.

Vous disposez des deux éditeurs suivants pour ajouter une fonction existante à une expression : Ces étapes sont :

- **Editeur rapide**
Intégré à d'autres fenêtres d'éditeurs tels que l'Editeur de script, l'Editeur conditionnel et l'Editeur de requêtes Query, l'Editeur rapide propose une syntaxe à code de couleurs, un menu contextuel, des raccourcis clavier et une liste des variables disponibles, des formats de type de données et des fonctions que vous pouvez utiliser afin de définir une fonction.
- **Assistant Fonction**
L'assistant Fonction permet de définir les paramètres d'une fonction existante. L'assistant Fonction apporte l'aide la plus importante lors de la définition de fonctions complexes.

[Création d'une expression comprenant une fonction existante \[page 679\]](#)

[Modification d'un appel de fonction existant dans une expression \[page 680\]](#)

Informations associées

[Éditeur rapide et outils fonctionnels \[page 252\]](#)

29.1.4.1 Création d'une expression comprenant une fonction existante

1. Accédez au script, à la requête ou à l'éditeur conditionnel où vous incluez l'expression.
2. Activez l'éditeur rapide et commencez à saisir votre expression.
3. Pour inclure la fonction, cliquez sur le bouton *Fonctions*.

Designer ouvre la fenêtre Sélectionner la fonction de l'Assistant Fonction.

4. Sélectionnez une catégorie dans la zone *Catégories de fonctions*.

Une liste des fonctions de cette catégorie s'affiche dans la zone Nom de la fonction. Les fonctions affichées dépendent de l'objet utilisé. Par exemple, les fonctions disponibles pour les flux de données ABAP sont un sous-ensemble de celles disponibles pour les flux de données.

Dans certains cas, il n'est pas logique d'utiliser une fonction, même si elle est disponible. Par exemple, la fonction SQL peut être appelée dans une expression de mappage ou dans une clause WHERE, mais elle aura pour effet d'insérer une instruction SQL dans l'instruction SQL générée pour exécuter un flux de données.

5. Sélectionnez une fonction spécifique dans la liste *Nom de la fonction*.

Une description de la fonction s'affiche sous les zones.

6. Cliquez sur *Suivant*.
7. Saisissez les valeurs requises par la fonction dans les zones de texte.

La page affichée pour chaque fonction est unique. Chaque page est conçue pour vous aider à construire la fonction en cours. L'exemple ci-dessous illustre la mise en forme la plus courante, mais des fonctions plus complexes peuvent utiliser des mises en forme différentes.

Cliquez dans une zone pour afficher une description du paramètre en bas de la fenêtre. Utilisez la flèche vers le bas pour sélectionner des paramètres d'entrée. Utilisez le bouton de l'éditeur rapide (...) pour afficher une zone de saisie plus grande.

8. Cliquez sur [Terminer](#).

La fonction et les paramètres s'affichent dans l'éditeur rapide.

Présentation des tâches : [Insertion de fonctions dans les expressions \[page 678\]](#)

Informations associées

[Modification d'un appel de fonction existant dans une expression \[page 680\]](#)

29.1.4.2 Modification d'un appel de fonction existant dans une expression

Divers éditeurs permettent de modifier les appels de fonction. Vous pouvez, par exemple, modifier les appels de fonction dans un éditeur rapide, dans un éditeur de script, etc. Toutefois, lorsqu'il s'agit de modifier des appels de fonction complexes, l'assistant Fonction permet de gagner du temps étant donné que l'analyse des différents éléments de la fonction entraîne une décomposition en arguments distincts. Cela permet ensuite d'identifier les options qu'il est nécessaire de modifier.

i Remarque

Si la fonction contient des espacements particuliers, des sauts de ligne ou des commentaires, il se peut que l'utilisation de l'assistant Fonction ne soit pas la meilleure solution. En effet, si vous l'utilisez pour modifier un appel de fonction existant, l'assistant Fonction supprime la mise en forme, les espaces et les commentaires. Une fois les modifications effectuées dans l'assistant Fonction, le texte de la fonction apparaît sur une seule ligne avec des espaces minimaux entre les paramètres.

1. Accédez à l'expression qui contient l'appel de fonction que vous souhaitez modifier.
2. Cliquez avec le bouton droit de la souris sur le texte de l'appel de fonction, puis sélectionnez [Modifier la fonction](#).
L'assistant Fonction s'ouvre.
3. Apportez les modifications souhaitées à l'appel de fonction, puis cliquez sur [Terminer](#).
L'assistant Fonction se ferme et vous pouvez voir le résultat de votre appel de fonction.

Présentation des tâches : [Insertion de fonctions dans les expressions \[page 678\]](#)

Informations associées

[Création d'une expression comprenant une fonction existante \[page 679\]](#)

[Opération d'une fonction \[page 677\]](#)

[Insertion de fonctions dans les expressions \[page 678\]](#)

29.1.5 Fonctions personnalisées

Vous pouvez créer vos propres fonctions à l'aide de l'éditeur rapide, en rédigeant des fonctions de script dans le langage de SAP Data Services. Les fonctions personnalisées enregistrées apparaissent dans l'assistant Fonction et l'éditeur rapide sous la catégorie *Personnalisé*. Elles sont également affichées dans la bibliothèque d'objets sous l'onglet *Fonctions personnalisées*. Vous pouvez modifier et supprimer les fonctions personnalisées présentes dans la bibliothèque d'objets.

L'onglet *Fonctions personnalisées* de la bibliothèque d'objets affiche plusieurs catégories de fonctions personnalisées :

- Fonctions personnalisées
- Fonctions de validation :
 - Importé depuis Information Steward : ces fonctions ont été créées dans Information Steward et importées ; elles ne sont pas modifiables dans Data Services.
 - Créé localement : fonctions de validation personnalisées réutilisables, créées dans Data Services.

Lors de la création de vos propres fonctions, prenez en compte les recommandations suivantes :

- Les fonctions peuvent appeler d'autres fonctions.
- Les fonctions ne peuvent pas s'appeler elles-mêmes.
- Les fonctions ne peuvent pas intervenir dans un cycle d'appels récursifs. La fonction A ne peut pas appeler la fonction B qui appelle la fonction A, par exemple.
- Les fonctions renvoient une valeur.
- Les fonctions peuvent avoir des paramètres d'entrée, de sortie ou les deux. Néanmoins, les flux de données ne peuvent pas transmettre de paramètres de type sortie ou entrée/sortie.

Avant de créer une fonction personnalisée, vous devez connaître les valeurs d'entrée et de sortie et les valeurs renvoyées, ainsi que les types de données correspondants. La valeur renvoyée est prédéfinie comme un retour.

[Create a custom function \[page 682\]](#)

Create custom functions for validation or when there is no built-in function for your purposes.

[Création d'une fonction personnalisée \[page 682\]](#)

[Modification d'une fonction existante \[page 684\]](#)

[Répliquer une fonction personnalisée \[page 684\]](#)

[Suppression d'une fonction personnalisée \[page 685\]](#)

Informations associées

[Langage de script \[page 706\]](#)

29.1.5.1 Create a custom function

Create custom functions for validation or when there is no built-in function for your purposes.

A custom function consists of return values, variables, and parameters. Create custom functions in the Smart Editor. Begin to create a custom function in one of two ways:

- From the *Custom Functions* tab in the object library.
- From the *Tools* menu.

Each time that you add a return, variable, or parameter, the Smart Editor uses a default naming convention. For example, create a new variable and the Smart Editor names it "\$NewVariable1." Change the name if you want. Make sure that you include the dollar sign in front of the name that you enter.

When you accept the default name, the ending number increases by one for each subsequent variable. So when you create a second new variable, the default name is "\$NewVariable2," and so on.

When you have finished configuring the various parts of your function, start to build it in the Smart Editor. Use the Validate feature to check your syntax and to ensure that the function is correct.

The variables and parameters that you create for a custom function are local to each function. Therefore, you can't access them in the *Variables and Parameters* dialog box that opens when you select **Tools** > *Variables*. If you edit the custom function, you can view the variables and parameters in the smart editor library, in the *Variables* tab.

Parent topic: [Fonctions personnalisées \[page 681\]](#)

Informations associées

[Création d'une fonction personnalisée \[page 682\]](#)

[Modification d'une fonction existante \[page 684\]](#)

[Répliquer une fonction personnalisée \[page 684\]](#)

[Suppression d'une fonction personnalisée \[page 685\]](#)

29.1.5.2 Création d'une fonction personnalisée

1. Sélectionnez **Outils** > *Fonctions personnalisée*.
2. Dans la liste Fonction personnalisée, cliquez avec le bouton droit et sélectionnez *Nouveau*.

Vous pouvez aussi cliquer avec le bouton droit dans la bibliothèque d'objets et sélectionner Nouveau dans l'onglet Fonctions personnalisées

3. Saisissez le nom de la nouvelle fonction.
4. Saisissez une description de la fonction.
5. Cliquez sur [Suivant](#) pour ouvrir l'éditeur rapide.

L'éditeur rapide permet de définir le type de renvoi, la liste de paramètres et les variables locales à utiliser dans la fonction.

6. Dans l'onglet Variables, cliquez avec le bouton droit sur [Renvoi](#) et sélectionnez [Propriétés...](#)

Par défaut, le type de données du renvoi est défini sur int. Pour le modifier, sélectionnez un autre type de données de renvoi dans la liste Type de données. Cliquez sur OK.

7. Dans l'onglet Variables, cliquez avec le bouton droit sur [Paramètres](#) et sélectionnez [Insérer](#).
8. Définissez les propriétés du paramètre en sélectionnant un type de données et un type de paramètre (Entrée, Sortie ou Entrée/Sortie).

i Remarque

Les flux de données ne peuvent pas transmettre de paramètres de variables de type Sortie ou Entrée/Sortie.

9. Cliquez sur [OK](#).

Répétez les étapes 7 à 9 pour chaque paramètre requis dans votre fonction. Après l'ajout d'un paramètre, le menu contextuel permet de choisir où insérer chacun des nouveaux paramètres.

Utilisez ce menu pour créer, supprimer ou modifier les variables et les paramètres.

10. Pour définir des variables utilisées par la fonction, mais non transmises hors de celle-ci, cliquez avec le bouton droit sur [Local](#) et sélectionnez [Insérer](#).
11. Sélectionnez un type de données dans la fenêtre Propriétés de la variable et cliquez sur [OK](#).

Répétez cette étape pour chaque variable requise dans votre fonction.

12. Complétez le texte de votre fonction.
13. Cliquez sur l'icône [Valider](#) pour valider votre fonction.



Si elle comporte des erreurs de syntaxe, ces erreurs s'affichent dans un volet intégré sous l'éditeur.

14. Pour savoir où se produisent les erreurs dans le texte, double-cliquez sur une erreur.

L'éditeur rapide redessine pour indiquer l'emplacement de l'erreur.

15. Lorsque la fonction est valide, cliquez sur [OK](#) pour l'enregistrer, ainsi que les variables et paramètres qu'elle contient.

Les variables et paramètres d'une fonction personnalisée existante sont locaux pour chaque fonction. Ils ne sont donc pas affichés dans la fenêtre Variables et paramètres (accessible depuis [Outils](#) > [Variables](#)). Les variables et paramètres des fonctions personnalisées peuvent être affichés dans l'onglet Variables de la bibliothèque de l'éditeur rapide, lorsque vous modifiez la fonction personnalisée.

Présentation des tâches : [Fonctions personnalisées \[page 681\]](#)

Informations associées

[Create a custom function \[page 682\]](#)

[Modification d'une fonction existante \[page 684\]](#)

[Répliquer une fonction personnalisée \[page 684\]](#)

[Suppression d'une fonction personnalisée \[page 685\]](#)

[Éditeur rapide et outils fonctionnels \[page 252\]](#)

29.1.5.3 Modification d'une fonction existante

Sélectionnez ► *Outils* ► *Fonctions personnalisées* ►, puis sélectionnez la fonction et *Suivant*.

L'éditeur rapide s'ouvre en affichant la fonction ainsi que les variables et paramètres la concernant dans l'onglet Variables.

Autres possibilités :

- Accédez à l'onglet Fonction personnalisée de la bibliothèque d'objets et double-cliquez sur la fonction à modifier.
- Accédez à l'onglet Fonction personnalisée de la bibliothèque d'objets, cliquez avec le bouton droit sur la fonction et sélectionnez *Modifier*.

Présentation des tâches : [Fonctions personnalisées \[page 681\]](#)

Informations associées

[Create a custom function \[page 682\]](#)

[Création d'une fonction personnalisée \[page 682\]](#)

[Répliquer une fonction personnalisée \[page 684\]](#)

[Suppression d'une fonction personnalisée \[page 685\]](#)

[Fonctions personnalisées \[page 681\]](#)

[Modification d'un appel de fonction existant dans une expression \[page 680\]](#)

29.1.5.4 Répliquer une fonction personnalisée

1. Dans la bibliothèque d'objets, cliquez avec le bouton droit sur une fonction personnalisée et sélectionnez *Répliquer*.

L'éditeur de fonctions personnalisées s'ouvre. Seul le champ Nom de la fonction est activé.

2. Saisissez un nouveau nom pour la fonction.

Le nom doit être valide et différent de celui d'origine.

3. Cliquez sur *Terminer*.

La nouvelle fonction personnalisée apparaît dans la bibliothèque d'objets.

Présentation des tâches : [Fonctions personnalisées \[page 681\]](#)

Informations associées

[Create a custom function \[page 682\]](#)

[Création d'une fonction personnalisée \[page 682\]](#)

[Modification d'une fonction existante \[page 684\]](#)

[Suppression d'une fonction personnalisée \[page 685\]](#)

29.1.5.5 Suppression d'une fonction personnalisée

Accédez à l'onglet Fonction personnalisée de la bibliothèque d'objets, cliquez avec le bouton droit sur la fonction et sélectionnez *Supprimer*.

Si vous supprimez une fonction de la liste de fonctions personnalisées, vous devez supprimer les références à la fonction dans les expressions conditionnelles, celles des scripts, des requêtes et des autres fonctions personnalisées.

Présentation des tâches : [Fonctions personnalisées \[page 681\]](#)

Informations associées

[Create a custom function \[page 682\]](#)

[Création d'une fonction personnalisée \[page 682\]](#)

[Modification d'une fonction existante \[page 684\]](#)

[Répliquer une fonction personnalisée \[page 684\]](#)

29.1.6 Fonctions intégrées

SAP Data Services propose un jeu de fonctions intégrées.

Le progiciel permet d'exécuter des fonctions de base de données et d'application, des fonctions personnalisées et la plupart des fonctions intégrées en parallèle dans les transformations avec lesquelles elles sont utilisées.

Vous pouvez exécuter les différentes fonctions gourmandes en ressources, telles que `lookup_ext` et `count_distinct`, sous forme de flux de données secondaire séparé utilisant des ressources distinctes (de la mémoire et de l'ordinateur).

[Fonctions de base de données et d'application \[page 686\]](#)

[Définition d'une fonction `sap\_openhub\_processchain\_execute` \[page 687\]](#)

[Définition d'une fonction `search\_replace` \[page 687\]](#)

[Définition et activation de la fonction `smtp\_to` \[page 689\]](#)

Informations associées

[Exécution de flux de données distribués \[page 224\]](#)

29.1.6.1 Fonctions de base de données et d'application

Vous pouvez importer les métadonnées de fonctions de base de données et d'application et les utiliser dans les applications Data Services. Vous pouvez aussi importer des fonctions et procédures stockées.

Au moment de l'exécution, Data Services transmet les informations appropriées à la base de données ou à l'application depuis laquelle la fonction a été importée.

Les métadonnées d'une fonction incluent les paramètres d'entrée et de sortie et leurs types de données. S'il existe des restrictions sur les données transmises à la fonction (requérant, par exemple, des valeurs en majuscules ou limitant les données à une plage spécifique), vous devez les appliquer dans l'entrée. Vous pouvez soit tester les données avant l'extraction, soit inclure une logique dans le flux de données qui appelle la fonction.

Rubrique parent : [Fonctions intégrées \[page 685\]](#)

Informations associées

[Définition d'une fonction `sap\_openhub\_processchain\_execute` \[page 687\]](#)

[Définition d'une fonction `search\_replace` \[page 687\]](#)

[Définition et activation de la fonction `smtp\_to` \[page 689\]](#)

[À propos des procédures \[page 690\]](#)

[Guide de Designer : Informations de procédures et de fonctions stockées importées \[page 103\]](#)

29.1.6.2 Définition d'une fonction sap_openhub_processchain_execute

1. Pour accéder à l'Assistant Fonctions pour sap_openhub_processchain_execute à partir de l'Editeur de script, cliquez sur [Fonctions](#) ou ... en haut de la fenêtre.
2. Sélectionnez [sap_openhub_processchain_execute](#) à partir de la liste de fonctions. La fenêtre [Définir les paramètres](#) apparaît.
3. Sélectionnez un nom de banque de données Source SAP BW à partir de la liste déroulante. Vous pouvez aussi sélectionner une variable de substitution à partir de la liste. Si vous saisissez un nom de banque de données, insérez des guillemets simples autour du nom.
4. Sélectionnez le nom d'une table Open Hub à partir de la liste déroulante. Seuls les noms des tables Open Hub importées apparaissent dans la liste.
5. Sélectionnez le nom d'une chaîne de processus à partir de la liste déroulante.
6. Indiquez une variable qui obtiendra l'ID journal BW pour la chaîne de processus après l'exécution de la fonction. Vous devez définir la variable avant de pouvoir la voir.
7. Indiquez une variable qui obtiendra la description du statut après l'exécution de la fonction. Vous devez définir la variable avant de pouvoir la voir.

Présentation des tâches : [Fonctions intégrées \[page 685\]](#)

Informations associées

[Fonctions de base de données et d'application \[page 686\]](#)

[Définition d'une fonction search_replace \[page 687\]](#)

[Définition et activation de la fonction smtp_to \[page 689\]](#)

29.1.6.3 Définition d'une fonction search_replace

1. Cliquez avec le bouton droit de la souris sur le schéma de sortie d'une transformation Query, puis cliquez sur [Nouvel appel de fonction](#).
2. Choisissez la catégorie [Fonctions de chaîne](#), puis cliquez sur [search_replace](#).
3. Cliquez sur [Suivant](#).
L'assistant Fonction search_replace s'ouvre.
4. Sélectionnez le type de recherche à effectuer.

Type de recherche	Description
Chaîne complète	Recherche et remplace les correspondances de l'intégralité du contenu de l'expression d'entrée.

Type de recherche	Description
Mot	Remplace uniquement le mot qui correspond à la valeur de recherche. Les données pour lesquelles il n'y a pas de correspondance ne sont pas modifiées. Un mot est défini comme un jeu de caractères séparé des autres données par des espaces.
Sous-chaîne	Remplace uniquement une sous-chaîne de caractères particulière se trouvant à côté ou entre d'autres caractères de l'expression d'entrée. Les données pour lesquelles il n'y a pas de correspondance ne sont pas modifiées. Dans le cas d'une recherche de sous-chaîne, les valeurs de recherche sont ordonnées de la plus longue à la plus courte, et chaque sous-chaîne n'est remplacée qu'une seule fois.

- Pour ignorer les différences de casse entre les valeurs de recherche spécifiées et les valeurs des expressions d'entrée, décochez la case *Sensible à la casse*.
- Définissez les expressions d'entrée à rechercher. Vous pouvez sélectionner un nom de colonne dans le schéma d'entrée ou vous servir de toute expression utilisant une ou plusieurs colonnes d'entrée.
- Définissez les noms des colonnes de sortie et les longueurs pour les valeurs de remplacement. Vous pouvez renommer la colonne de sortie en sélectionnant la valeur et en appuyant sur la touche **F2**.
- Configurez la table de recherche et de remplacement.
 - Pour une table interne, sélectionnez *Interne*, puis spécifiez des valeurs de recherche et les valeurs de remplacement correspondantes.
Les valeurs destinées à une table interne sont stockées dans le référentiel de métadonnées comme partie intégrante de la définition de la fonction.
 - Pour une table relationnelle externe, sélectionnez *Externe*. Spécifiez un format de fichier ou une banque de données valide comme source, puis sélectionnez les colonnes contenant les valeurs de recherche et de remplacement.
 - Pour un SQL personnalisé, sélectionnez *SQL personnalisé*. Choisissez la banque de données source, puis spécifiez les colonnes contenant les valeurs de recherche et de remplacement. Définissez le texte du SQL personnalisé à exécuter.
- Spécifiez la valeur de remplacement par défaut pour toutes les lignes qui ne correspondent pas aux valeurs de recherche définies. Si la valeur de remplacement est une chaîne fixe, elle doit se trouver entre guillemets simples.

Remarque

La valeur de remplacement par défaut s'applique uniquement dans le cas d'une recherche de chaîne complète et non pour les recherches portant sur un mot ou sur une sous-chaîne.

Si aucune valeur de remplacement n'est définie, les valeurs d'origine sont conservées ; aucun remplacement n'est effectué pour les expressions non trouvées.

- Pour obliger la fonction `search_replace` à s'exécuter dans un processus à part, cochez la case *Exécuter en tant que processus séparé*. La table de recherche et de remplacement est toujours chargée en mémoire, de sorte que pour les tables de grande taille, l'exécution de cette fonction en tant que processus distinct peut se traduire par une amélioration de la performance.

Exemple

Recherche et remplacement de sous-chaîne tenant compte de la casse

Dans cet exemple, la valeur de recherche la plus longue est remplacée en premier. En conséquence, la sous-chaîne AUS ne devient pas AUSA mais est correctement remplacée par Australie.

Table de recherche et de remplacement interne :

Valeur de recherche	Valeur de remplacement
US	Etats-Unis
AUS	Australie

Résultat de la recherche et du remplacement :

Valeur dans l'entrée	Valeur dans la sortie
Californie, US	Californie, Etats-Unis
Melbourne, AUS	Melbourne, Australie

Présentation des tâches : [Fonctions intégrées \[page 685\]](#)

Informations associées

[Fonctions de base de données et d'application \[page 686\]](#)

[Définition d'une fonction sap_openhub_processchain_execute \[page 687\]](#)

[Définition et activation de la fonction smtp_to \[page 689\]](#)

29.1.6.4 Définition et activation de la fonction smtp_to

1. Dans l'éditeur de fonctions, cliquez sur la catégorie de fonctions [Système](#).
2. Cliquez sur le nom de fonction [smtp_to](#).
3. Cliquez sur [Suivant](#).
4. Définissez les paramètres d'entrée comme décrit précédemment dans la table [WHERE](#) de cette description de fonction.
5. Cliquez sur [Terminer](#).
6. Ouvrez le Gestionnaire de serveurs Data Services :
 - Sous Windows, cliquez sur ► [Démarrer](#) ► [Programmes](#) ► [SAP Data Services](#) <x . x> ► [Gestionnaire de serveurs Data Services](#) ►.
 - Sous UNIX, exécutez le Gestionnaire de serveurs en saisissant :

```
$ cd $LINK_DIR/bin/
```

```
$ . ./al_env.sh
```

```
$ ./svrcfg
```

7. La fenêtre de l'utilitaire Gestionnaire de serveurs Data Services s'ouvre.

8. Entrez les paramètres du serveur SMTP :
 - *Nom du serveur* : saisissez le nom ou l'adresse IP du serveur SMTP (par exemple, mail.société.com).
 - *Expéditeur* : saisissez l'adresse électronique qui figurera dans le champ *De* de l'e-mail.
9. Dans la fenêtre Gestionnaire de serveurs, cliquez sur *Appliquer*.

Présentation des tâches : [Fonctions intégrées \[page 685\]](#)

Informations associées

[Fonctions de base de données et d'application \[page 686\]](#)

[Définition d'une fonction sap_openhub_processchain_execute \[page 687\]](#)

[Définition d'une fonction search_replace \[page 687\]](#)

29.2 À propos des procédures

SAP Data Services prend en charge les procédures stockées, qui désignent des objets exécutables ou des entités nommées que vous stockez dans une base de données.

Une procédure stockée consiste en une ou plusieurs instructions SQL précompilées. En appelant une procédure stockée à partir de Data Services, vous pouvez appeler la logique applicative que vous avez déjà codée. Cette possibilité d'utiliser votre logique applicative existante vous permet de développer rapidement et aisément des tâches d'extraction de données et de gestion des données. Vous pouvez également utiliser les procédures stockées aux fins suivantes :

- Gérer les règles de logique applicative et fournir un seul point de contrôle pour vous assurer de l'exactitude et du respect des règles
- Réduire les frais généraux liés au réseau des applications client et serveur. Le logiciel stocke les procédures sur le serveur de base de données et conserve les plans d'exécution compilés des procédures dans le dictionnaire de données.

Data Services prend en charge l'utilisation des procédures stockées pour les bases de données suivantes :

- DB2
- ODBC
- Oracle
- Microsoft SQL Server
- SAP HANA
- SAP SQL Anywhere
- SAP ASE
- SAP Sybase IQ
- Teradata

Appelez des procédures stockées à partir des jobs que vous créez et exécutez dans Data Services.

Les procédures stockées doivent exister dans la base de données avant de pouvoir utiliser la procédure dans Data Services. Créez une procédure stockée dans une base de données à l'aide des outils client fournis avec la

base de données, comme Oracle SQL *Plus. Une fois la procédure stockée créée, les utilisateurs disposant des autorisations d'exécution pour les procédures peuvent l'appeler. Après l'importation d'une procédure stockée dans Data Services, utilisez-la comme une fonction dans les jobs Data Services.

Les procédures stockées incluent des paramètres. Chaque paramètre comprend les éléments suivants :

- Nom
- Type de données
- Mode (par exemple, IN, INOUT, ou OUT)

Une procédure stockée peut utiliser une valeur de paramètre par défaut ou une valeur NULL pour son entrée ; elle peut produire plusieurs valeurs de paramètre de sortie.

Pour obtenir des informations complètes sur l'utilisation des procédures, voir le *Guide de Designer*.

[Exigences \[page 691\]](#)

[Création de procédures stockées dans une base de données \[page 692\]](#)

[Importation de métadonnées pour des procédures stockées \[page 697\]](#)

[Structure d'une procédure stockée \[page 698\]](#)

[Appel de procédures stockées \[page 699\]](#)

[Vérification du statut d'exécution \[page 704\]](#)

29.2.1 Exigences

Pour utiliser des procédures stockées avec Data Services, la configuration doit répondre aux exigences suivantes :

- La version du client doit correspondre à celle du serveur.
- Seules les procédures stockées définies par l'utilisateur peuvent être utilisées. Data Services ne prend pas en charge les procédures stockées fournies par un système de bases de données.
- Les procédures stockées ou les fonctions stockées définies par l'utilisateur doivent répondre aux exigences supplémentaires qui suivent :
 - Le type de retour (type de données de la valeur du résultat) doit être un type de données pris en charge par Data Services, comme `varchar`, `date` ou `int`.

i Remarque

Cette version ne prend pas en charge le type de données long dans les procédures stockées.

- Le nom de la procédure stockée - la combinaison du nom de la banque de données, du propriétaire et de la procédure - doit être unique. Data Services importe uniquement la première procédure ou fonction portant un nom particulier.
Par exemple, si vous utilisez Oracle, que plusieurs procédures ou fonctions portant le même nom se trouvent dans un package et que vous souhaitez toutes les utiliser, vous devez renommer les procédures, en donnant un nom unique à chacune. Une procédure est surchargée lorsque plusieurs versions existent dans un package particulier. La surcharge n'est pas prise en charge dans Data Services.
- Data Services valide le nom d'utilisateur spécifié dans la banque de données qui a été utilisé pour importer la procédure stockée. Cela garantit que le flux de données qui appelle une procédure stockée respecte les restrictions associées à cette procédure, comme :

- Les restrictions d'entrée de données
- Les droits d'exécution
- Pour utiliser des procédures stockées avec Teradata, vous devez d'abord activer l'utilisation des procédures stockées dans les options de pilote ODBC Teradata et invoquer les procédures stockées dans le même mode de session que celui dans lequel elle a été activée.
 Dans Windows, dans l'écran *Options du pilote ODBC Teradata* utilisé pour la configuration du pilote ODBC Teradata, cochez la case *Désactiver la conversion CALL vers EXEC*, définissez le mode de session souhaité dans la liste déroulante *Mode de session*, puis cliquez sur *OK* pour enregistrer la modification. Vous devez invoquer les procédures stockées dans le même mode de session.
 Pour les plateformes UNIX, modifiez le fichier `odbc.ini` pour définir l'option **Désactiver la conversion CALL vers EXEC=YES**. Définissez l'option de mode de session comme vous le souhaitez, puis enregistrez le fichier. Vous devez invoquer les procédures stockées dans le même mode de session.
 Pour plus d'informations sur la configuration du pilote ODBC Teradata, voir la documentation relative à cette configuration.

29.2.2 Création de procédures stockées dans une base de données

Cette section fournit :

- Un exemple de scénario concernant l'utilisation d'une procédure stockée
- Des conseils relatifs à la création de procédures stockées sur chaque base de données
- Exemple de syntaxe pour des bases de données DB2, Oracle, SAP HANA, SQL Server et SAP ASE basées sur l'exemple de scénario

Puisque la syntaxe varie entre les bases de données, les instructions SQL nécessaires pour créer la procédure stockée d'un scénario varient également. Les outils client utilisés pour transmettre les instructions SQL au serveur de base de données varient aussi. Reportez-vous à la documentation de la base de données pour des descriptions détaillées et des exemples concernant la création de procédures stockées.

Dans l'exemple suivant, la base de données source contient une procédure stockée `Get_emp_rec` qui extrait le nom et la date d'embauche d'un employé à partir de la table `Employé` à l'aide d'un numéro d'employé donné. La procédure stockée prend un paramètre d'entrée `Emp_number` et retourne deux valeurs `Emp_name` et `Hire_date` via des paramètres de sortie.

Le schéma de la table de base de données `Employé` est le suivant :

Nom	Null?	Type
empno	Non Null	Integer
ename	Null	Varchar(20)
job	Null	Varchar(9)
mgr	Null	Integer
hiredate	Null	Date

Nom	Null?	Type
sal	Null	Decimal(7,2)
comm	Null	Decimal(7,2)
deptno	Null	Integer

[Création de procédures stockées dans DB2 \[page 693\]](#)

[Création de procédures stockées dans Oracle \[page 694\]](#)

[Création de procédures stockées dans SAP HANA \[page 695\]](#)

SAP Data Services prend en charge les procédures stockées SAP HANA comportant zéro, un ou plusieurs paramètres de sortie.

[Création de procédures stockées dans Microsoft SQL Server ou SAP ASE \[page 696\]](#)

29.2.2.1 Création de procédures stockées dans DB2

Dans DB2, un paramètre peut être IN, OUT ou INOUT. Tous les types de données SQL peuvent être utilisés comme type de données d'un paramètre. Les types de données définis par l'utilisateur ne sont pas pris en charge.

❖ Exemple

La syntaxe DB2 pour la procédure stockée

```
CREATE PROCEDURE GET_NAME_USING_ID (IN NID INTEGER, OUT outVar varchar(20))
language SQL
reads sql data
BEGIN
select first_name into outVar FROM CONTACT where id = NID;
END
```

Reads sql data et language SQL représentent deux des nombreuses options prises en charge par les procédures stockées DB2. Reportez-vous à la documentation DB2 relative aux options des procédures stockées DB2 et à leur signification.

Rubrique parent : [Création de procédures stockées dans une base de données \[page 692\]](#)

Informations associées

[Création de procédures stockées dans Oracle \[page 694\]](#)

[Création de procédures stockées dans SAP HANA \[page 695\]](#)

[Création de procédures stockées dans Microsoft SQL Server ou SAP ASE \[page 696\]](#)

[Création de procédures stockées dans une base de données \[page 692\]](#)

29.2.2.2 Création de procédures stockées dans Oracle

Oracle prend à la fois en charge les procédures stockées et les fonctions stockées. Toute procédure stockée qui retourne une valeur est appelée fonction stockée. Oracle est la seule base de données qui autorise des valeurs renvoyées avec des types de données autre que le type entier.

Dans Oracle, les procédures stockées sont créées à l'aide de l'instruction `CREATE [OR REPLACE] PROCEDURE` et les fonctions stockées sont créées à l'aide de l'instruction `CREATE [OR REPLACE] FUNCTION`. L'option `OR REPLACE` permet de remplacer la définition existante d'une procédure ou d'une fonction portant le même nom.

Oracle prend en charge les paramètres `IN`, `OUT` et `INOUT`. Un paramètre `INOUT` permet de transmettre un paramètre, de le modifier et de retourner la valeur modifiée. Vous pouvez utiliser le mot clé `DEFAULT` ou l'opérateur d'affectation pour donner une valeur par défaut au paramètre `IN`. Lorsqu'un paramètre `IN` a une valeur par défaut, vous pouvez omettre le paramètre de la liste d'arguments lors de l'appel de la procédure. Si vous spécifiez une valeur d'argument dans l'appel, la valeur spécifiée remplace la valeur par défaut. Vous devez spécifier les paramètres `INOUT` et `OUT`.

❖ Exemple

Syntaxe

La syntaxe Oracle pour la procédure stockée :

```
CREATE OR REPLACE PROCEDURE
get_emp_rec (Emp_Number IN NUMBER,
Emp_Name OUT VARCHAR, Emp_Hiredate OUT DATE) AS
BEGIN
SELECT ename, hiredate
INTO Emp_Name, Emp_Hiredate
FROM Employee
WHERE empno = Emp_Number;
END;
/
```

Dans cet exemple, les paramètres sont déclarés en tant que `IN` et `OUT`, mais Oracle prend aussi en charge un type de paramètre `INOUT`, qui permet de transmettre un paramètre, de le modifier et de retourner la valeur modifiée. Les paramètres `INOUT` peuvent prendre les valeurs par défaut. Autrement dit, le paramètre peut être omis de la liste des paramètres réels lorsque vous appelez la procédure. Les paramètres `INOUT` et `OUT` doivent être spécifiés.

Un package Oracle est une collection encapsulée d'objets de programme reliés (par exemple des procédures, fonctions, variables, curseurs et exceptions) stockés ensemble dans une base de données. Data Services permet d'importer des procédures ou des fonctions créées dans des packages et de les utiliser comme des procédures ou des fonctions de niveau supérieur.

Data Services ne prend pas en charge la surcharge des procédures ou des fonctions. La surcharge d'une procédure implique la création de plusieurs procédures portant le même nom dans le même package, chacune reprenant des arguments d'un nombre ou d'un type de données différent. Si plusieurs procédures ou fonctions portant le même nom se trouvent dans le même package (et que vous souhaitez toutes les utiliser), vous devez les renommer, en donnant un nom unique à chacune. Sinon, Data Services importe uniquement la première procédure ou fonction portant ce nom.

Rubrique parent : [Création de procédures stockées dans une base de données \[page 692\]](#)

Informations associées

[Création de procédures stockées dans DB2 \[page 693\]](#)

[Création de procédures stockées dans SAP HANA \[page 695\]](#)

[Création de procédures stockées dans Microsoft SQL Server ou SAP ASE \[page 696\]](#)

[Création de procédures stockées dans une base de données \[page 692\]](#)

29.2.2.3 Création de procédures stockées dans SAP HANA

SAP Data Services prend en charge les procédures stockées SAP HANA comportant zéro, un ou plusieurs paramètres de sortie.

Data Services prend en charge les types de données scalaires pour les paramètres d'entrée et de sortie. Data Services ne prend pas en charge les types de données de table. Si vous essayez d'importer une procédure avec type de données de table, le logiciel envoie une erreur. Data Services ne prend pas en charge les types de données tels que binary, blob, clob, nclob ou varbinary pour les paramètres de procédures SAP HANA.

Les procédures peuvent être appelées depuis un script ou une transformation Query sous forme de nouvel appel de fonction.

❖ Exemple

Syntaxe

Syntaxe SAP HANA de la procédure stockée :

```
CREATE PROCEDURE GET_EMP_REC (IN EMP_NUMBER INTEGER, OUT EMP_NAME
VARCHAR(20), OUT EMP_HIREDATE DATE) AS
BEGIN
    SELECT ENAME, HIREDATE
        INTO EMP_NAME, EMP_HIREDATE
    FROM EMPLOYEE
    WHERE EMPNO = EMP_NUMBER;
END;
```

Limitations

SAP HANA fournit une prise en charge limitée des fonctions définies par l'utilisateur qui peuvent renvoyer une ou plusieurs valeurs scalaires. Ces fonctions définies par l'utilisateur sont habituellement écrites en L. Si vous utilisez des fonctions définies par l'utilisateur, limitez-les à la liste de projections et à la clause GROUP BY d'une requête d'agrégation au-dessus d'un cube OLAP ou d'une table de colonnes. Ces fonctions ne sont pas prises en charge par Data Services.

Les procédures SAP HANA ne peuvent pas être appelées depuis une clause WHERE.

Rubrique parent : [Création de procédures stockées dans une base de données \[page 692\]](#)

Informations associées

[Création de procédures stockées dans DB2 \[page 693\]](#)

[Création de procédures stockées dans Oracle \[page 694\]](#)

[Création de procédures stockées dans Microsoft SQL Server ou SAP ASE \[page 696\]](#)

[Création de procédures stockées dans une base de données \[page 692\]](#)

29.2.2.4 Création de procédures stockées dans Microsoft SQL Server ou SAP ASE

Dans Microsoft SQL Server et SAP ASE, une procédure stockée peut porter un numéro de groupe (entre 1 et 32767), qui est un entier facultatif utilisé pour regrouper les procédures du même nom pour pouvoir les annuler à l'aide d'une seule instruction `DROP PROCEDURE`.

Le groupe numéro 1 est automatiquement affecté à la première procédure que vous créez avec un nom. Bien que vous puissiez exécuter cette procédure sans spécifier le numéro de groupe, vous devez appeler les autres procédures portant le même nom de procédure avec le numéro de groupe. Par exemple, exécutez `orderproc; 2` et ainsi de suite. Data Services importe la procédure avec le groupe numéro 1, c'est-à-dire la première procédure, sauf si vous spécifiez le nom et le numéro de groupe, comme `orderproc; 2`.

Un paramètre peut être un paramètre d'entrée (par défaut) ou un paramètre de sortie. Un paramètre ne peut pas être à la fois d'entrée et de sortie. Dans l'exemple suivant, le premier paramètre `@Emp_Number` est un paramètre d'entrée et les deux autres paramètres, `@Emp_Name` et `@Emp_Hiredate`, sont des paramètres de sortie. Les paramètres de sortie sont spécifiés à l'aide du mot clé `OUTPUT`. Les paramètres d'entrée peuvent prendre les valeurs par défaut (valeurs par défaut de la base de données). Si une valeur par défaut est définie, la procédure peut être exécutée sans spécifier de valeur pour ce paramètre. Vous spécifiez un nom de paramètre qui commence par le caractère "@".

Une valeur renvoyée entière facultative est automatiquement ajoutée à chaque procédure stockée pour stocker son statut de retour. Dans l'exemple suivant, si l'instruction `SELECT` aboutit, la valeur "0" est retournée comme valeur renvoyée. Le nom du paramètre de la valeur renvoyée est `RETURN_VALUE`. Les procédures stockées Microsoft SQL Server et SAP ASE peuvent être utilisées comme des fonctions stockées Oracle dans Data Services, car elles incluent cette valeur renvoyée.

Une fonction définie par l'utilisateur est une nouvelle fonctionnalité de SQL Server 2000. Data Services prend en charge les procédures stockées définies par l'utilisateur, mais pas les fonctions définies par l'utilisateur.

❖ Exemple

La syntaxe Microsoft SQL Server et SAP ASE

```
CREATE PROCEDURE GET_EMP_REC @Emp_Number int, @Emp_Name varchar(20) OUTPUT,  
@Emp_Hiredate datetime OUTPUT AS  
SELECT @Emp_Name = ename, @Emp_Hiredate = hiredate  
FROM Employee  
WHERE empno = @Emp_Number  
RETURN 0
```

Rubrique parent : [Création de procédures stockées dans une base de données \[page 692\]](#)

Informations associées

[Création de procédures stockées dans DB2 \[page 693\]](#)

[Création de procédures stockées dans Oracle \[page 694\]](#)

[Création de procédures stockées dans SAP HANA \[page 695\]](#)

[Création de procédures stockées dans une base de données \[page 692\]](#)

29.2.3 Importation de métadonnées pour des procédures stockées

Importez des procédures stockées dans Data Services à l'aide de la commande Importer par nom. Vous devez connaître le nom de la procédure stockée, de la fonction stockée ou du package que vous souhaitez importer. Utilisez le nom de la procédure stockée, le nom du package uniquement, ou, si vous utilisez une base de données Microsoft SQL Server ou SAP Sybase ASE, utilisez le nom de la procédure stockée et le numéro de groupe, par exemple `orderproc;2`.

La création et la structure des procédures stockées ressemblent beaucoup à celles des fonctions. Les procédures stockées importées apparaissent sous l'onglet Banques de données de la bibliothèque d'objets. Chacune d'entre elles est listée dans la catégorie Fonctions de la banque de données de base de données à partir de laquelle la procédure a été importée. Elle s'affiche avec la syntaxe suivante :

```
<datastore>.<owner>.<proc_name>.
```

Si une procédure stockée a été définie avec un package Oracle, le nom de la procédure s'affiche sous la forme suivante :

```
<datastore>.<owner>.<package>.<proc_name>.
```

Pour afficher la signature de la procédure, cliquez avec le bouton droit sur la procédure stockée importée, puis sélectionnez Ouvrir (ou cliquez simplement deux fois sur la procédure).

- Une procédure stockée a une signature qui se compose de ses paramètres définis comprenant deux paramètres de sortie ajoutés par Data Services lors de l'importation, `AL_SP_RETCODE` et `AL_SP_ERRMSG`.
- Si la base de données que vous utilisez fournit un code de retour pour une procédure stockée, ce code s'affiche dans la signature sous la forme `AL_SP_RETURN`.
- Chaque paramètre possède un nom, un type de données et un mode (`IN`, `INOUT` ou `OUT`). Le mode est indiqué par des icônes affichées avant le nom du paramètre.
- Les fonctions stockées Oracle possèdent une signature qui se compose des paramètres et d'un type de données de retour. Lorsque vous appelez une fonction, la valeur renvoyée est nommée `AL_SP_RETURN`.

i Remarque

Si vous changez la signature d'une procédure stockée, par exemple, vous ajoutez un paramètre, ré-importez les métadonnées pour la procédure à partir de la base de données.

Informations associées

Vérification du statut d'exécution [page 704]

29.2.4 Structure d'une procédure stockée

Exemple d'informations importées pour une procédure stockée :

- Nom et propriétaire
- Type de données retournées si la procédure stockée a une valeur ou un code de retour
- Paramètres

Pour chaque paramètre, Data Services importe les informations suivantes :

- Nom
- Mode (IN, OUT ou INOUT)
- Type de données (longueur, précision et échelle le cas échéant)

i Remarque

Dans une déclaration de procédure stockée Oracle, il est illégal de limiter la longueur des paramètres char et varchar2 ainsi que de limiter les paramètres de nombre avec une précision ou une échelle. Ces limites proviennent des arguments de la procédure. Par conséquent, la longueur d'un paramètre char ou varchar ainsi que la précision et l'échelle d'un paramètre de type de données nombre ne sont pas stockées dans le dictionnaire de données.

Cependant, Data Services n'autorise pas la valeur 0 pour la longueur d'un type de données varchar ou la précision d'un type de données nombre. (Data Services convertit les types de données char dans Oracle en varchar.) Lors de l'importation d'une procédure stockée, Data Services remplit la longueur par défaut d'un paramètre varchar ou la précision et l'échelle d'un paramètre nombre. Pour varchar, Data Services définit la longueur sur 4000. Pour les nombres, Data Services définit la précision et l'échelle sur les valeurs définies dans la section Avancé de la fenêtre Créer une banque de données. Lors de l'exécution, la longueur de varchar et la précision et l'échelle du nombre provient des arguments de la procédure.

- Position au sein de la liste des paramètres
- Peut être de valeur Null (qu'une valeur par défaut existe ou non)

Data Services omet un paramètre si une ou plusieurs des conditions de type de données suivantes sont vérifiées :

- N'est pas un type de données pris en charge par Data Services ou est `long`
- Le type de données est absent du dictionnaire de base de données
- Est un type de données composite (par exemple table, enregistrement ou curseur)

Après l'importation d'une procédure stockée par Data Services, Designer signale des avertissements sur les paramètres omis.

i Remarque

Même si une procédure stockée présente un paramètre omis, la procédure peut être appelée si une valeur par défaut est définie pour le paramètre. Si le paramètre omis n'a pas de valeur par défaut, il n'est pas possible d'appeler la procédure et Data Services renvoie une erreur d'exécution.

Si une procédure stockée présente une valeur renvoyée, son type de résultat (le type de données de la valeur du résultat) doit être un type de données de serveur de base de données pris en charge par Data Services. Sinon, la procédure ne peut pas être importée et utilisée au sein d'applications Data Services.

Pour chaque procédure stockée, Data Services ajoute automatiquement deux paramètres de sortie facultatifs, `AL_SP_RETCODE` et `AL_SP_ERRMSG`. Ces paramètres permettent de vérifier le statut d'exécution de la procédure stockée.

Informations associées

[Vérification du statut d'exécution \[page 704\]](#)

29.2.5 Appel de procédures stockées

Vous pouvez appeler une procédure stockée tout comme vous appelez des fonctions personnalisées dans Data Services.

- Lorsque vous appelez une procédure stockée à l'intérieur d'une fonction personnalisée, conditionnelle ou de script, Data Services évalue l'appel en tant qu'instruction SQL indépendante.
- Lorsque vous appelez une procédure stockée dans un mappage de colonne d'une requête ou dans une clause WHERE, (Oracle uniquement), Data Services peut associer l'instruction SQL de la procédure stockée à l'instruction SQL de la requête.
- Lorsque vous appelez une procédure stockée dans une transformation SQL, vous devez utiliser une syntaxe appropriée. Data Services ne valide pas d'instruction SQL dans une transformation SQL.
- Une procédure stockée peut être utilisée au sein d'une expression. Exemple :

```
ora_ds.acta_user.get_emp_sal (EMPLOYEE.NAME) + 10  
ora_ds.acta_user.get_emp_sal (EMPLOYEE.NAME) < 100
```

[Paramètres d'entrée de procédures stockées \[page 700\]](#)

[Appel de procédures stockées à partir de scripts \[page 701\]](#)

[Appel de procédures stockées à partir de requêtes \[page 701\]](#)

[Sans l'Assistant Fonction \[page 703\]](#)

Informations associées

[À propos des procédures \[page 690\]](#)

[Paramètres d'entrée de procédures stockées \[page 700\]](#)

[Appel de procédures stockées à partir de scripts \[page 701\]](#)

[Appel de procédures stockées à partir de requêtes \[page 701\]](#)

[Sans l'Assistant Fonction \[page 703\]](#)

29.2.5.1 Paramètres d'entrée de procédures stockées

Les procédures stockées peuvent nécessiter des valeurs d'entrée pour certains paramètres sans en nécessiter pour d'autres. Vous devez déterminer les exigences de la procédure stockée et préparer les entrées appropriées. Dans Data Services, les paramètres d'entrée peuvent être des constantes ou des expressions. En outre, Data Services prend en charge ce qui suit :

- Si un paramètre a une valeur par défaut, vous pouvez l'omettre dans l'appel en ne le spécifiant pas dans l'Assistant Fonction. La base de données utilisera la valeur par défaut du paramètre lors de l'évaluation de l'appel de la procédure stockée.
- L'ensemble des paramètres `OUT` ou `INOUT` doit être spécifié à l'exception des deux paramètres ajoutés par Data Services lors de l'importation : `AL_SP_RETCODE` et `AL_SP_ERRMSG`.
- Tous les paramètres `OUT` ou `INOUT` doivent être des variables (sauf dans un schéma de sortie de requête). Avant d'être utilisée, une variable doit être déclarée.
- Si une procédure stockée retourne une valeur (comme une fonction Oracle ou une procédure stockée Microsoft SQL Server ou SAP ASE), vous pouvez utiliser la procédure dans une expression.
- Vous pouvez appeler une procédure stockée en tant qu'étape dans un schéma de sortie d'une requête, d'un script ou d'une fonction personnalisée. Exemple :

```
As an Oracle user you create two Oracle stored procedures: GET_EMP_SALARY and
SET_EMP_SALARY
CREATE OR REPLACE FUNCTION GET_EMP_SALARY(Emp_Name IN VARCHAR)
RETURN NUMBER AS
Return_value NUMBER;
BEGIN
SELECT sal
INTO Return_value
FROM Employee
WHERE ename = Emp_Name;
RETURN Return_value;
END;
CREATE OR REPLACE PROCEDURE SET_EMP_SALARY(Emp_Name IN VARCHAR, Emp_Sal IN
NUMBER) AS
BEGIN
UPDATE Employee
SET sal = Emp_Sal
WHERE ename = Emp_Name;
END;
```

Rubrique parent : [Appel de procédures stockées \[page 699\]](#)

Informations associées

[Appel de procédures stockées à partir de scripts \[page 701\]](#)

[Appel de procédures stockées à partir de requêtes \[page 701\]](#)

[Sans l'Assistant Fonction \[page 703\]](#)

[À propos des procédures \[page 690\]](#)

[Appel de procédures stockées \[page 699\]](#)

29.2.5.2 Appel de procédures stockées à partir de scripts

Après l'importation des procédures stockées dans une banque de données Oracle, vous pouvez les utiliser dans un script. Exemple :

```
$new_sal = ora_ds.bodi_user.get_emp_salary('MARTIN') + 100;  
ora_ds.bodi_user.set_emp_salary('MARTIN', $new_sal);
```

Rubrique parent : [Appel de procédures stockées \[page 699\]](#)

Informations associées

[Paramètres d'entrée de procédures stockées \[page 700\]](#)

[Appel de procédures stockées à partir de requêtes \[page 701\]](#)

[Sans l'Assistant Fonction \[page 703\]](#)

[À propos des procédures \[page 690\]](#)

[Appel de procédures stockées \[page 699\]](#)

29.2.5.3 Appel de procédures stockées à partir de requêtes

Vous pouvez appeler des procédures stockées à partir de transformations Query dans le mappage et les clause WHERE ou dans un schéma de sortie de transformation Query.

Rubrique parent : [Appel de procédures stockées \[page 699\]](#)

Informations associées

[Paramètres d'entrée de procédures stockées \[page 700\]](#)

[Appel de procédures stockées à partir de scripts \[page 701\]](#)

[Sans l'Assistant Fonction \[page 703\]](#)

[À propos des procédures \[page 690\]](#)

[Appel de procédures stockées \[page 699\]](#)

[Importation de métadonnées pour des procédures stockées \[page 697\]](#)

29.2.5.3.1 Mappage et clauses WHERE

Dans une clause WHERE ou dans le mappage de colonne d'une requête, vous pouvez utiliser une procédure stockée dans une expression, soumise aux mêmes contraintes que les autres expressions. Toutefois, lorsque

des procédures stockées sont utilisées dans le mappage de colonne d'une requête ou dans une clause WHERE, SAP Data Services peut combiner l'instruction SQL de l'appel de la procédure stockée à l'instruction SQL de la requête et soumettre une seule requête à la base de données.

Pour que le logiciel combine les instructions SQL, la procédure stockée doit répondre aux conditions suivantes :

- Elle doit provenir d'une base de données Oracle.
- Il doit s'agir d'une fonction stockée.
- Tous ses paramètres doivent être des paramètres IN ; aucun ne peut être OUT ou INOUT.
- Toutes les valeurs des paramètres doivent être spécifiées. Vous ne pouvez pas laisser un paramètre non spécifié pour, par exemple, utiliser la valeur par défaut.
- Toutes les valeurs des paramètres doivent être des constantes ou des expressions contenant uniquement des constantes ou des noms de colonne de table de base de données ; aucune d'entre elles ne peut posséder de fonctions ni d'opérations.
- Les types de données des paramètres de la procédure et le type de retour (le type de données de sa valeur de résultat) doivent être des types de données pris en charge. Le logiciel ne doit pas omettre des paramètres lors de l'importation de la procédure.
- Les opérations DDL (Data Definition Language) comme la création de tables, de vues, etc. et les instructions de contrôle des transactions telles que `COMMIT`/`ROLLBACK` ne peuvent pas être effectuées à l'intérieur de la procédure stockée.

En outre, lors de l'importation de la fonction, dans la fenêtre Importer par nom, vous devez activer la case à cocher Joignable depuis l'expression SQL.

Si les cinq premières conditions sont remplies et que la case *Joignable depuis l'expression SQL* est activée, le logiciel envoie un appel de procédure stockée vers Oracle. Si un appel de procédure stockée contient des instructions DDL, il termine la transaction en cours par `COMMIT` ou `ROLLBACK`, ou bien il émet les commandes `ALTER SESSION` ou `ALTER SYSTEM`. Par conséquent, dans la fenêtre Importer par nom, vous devez désactiver l'option *Joignable depuis l'expression SQL* pour vous assurer que l'appel de fonction ne sera pas renvoyé vers la base de données.

Lorsqu'un appel de procédure stockée ne peut pas être renvoyé vers la base de données, il y est soumis sous la forme d'une instruction SQL séparée.

Informations associées

[À propos des procédures \[page 690\]](#)

[Appel de procédures stockées \[page 699\]](#)

29.2.5.3.2 Schéma de sortie de transformation Query

Lorsqu'une procédure stockée peut fournir au moins une valeur renvoyée ou contient des paramètres `INOUT` ou `OUT` différents de `AL_SP_RETCODE` ou `AL_SP_ERRMSG`, vous pouvez l'appeler en l'incluant sous forme d'objet dans le schéma de sortie d'une requête.

Lorsque vous utilisez une procédure stockée dans un schéma de sortie de requête, vous devez mapper les paramètres `INOUT` ou `OUT` et les valeurs renvoyées à une colonne du schéma de sortie. Dans ce cas, vous ne pouvez pas mapper ces paramètres à des variables.

Informations associées

[À propos des procédures \[page 690\]](#)

[Appel de procédures stockées \[page 699\]](#)

29.2.5.4 Sans l'Assistant Fonction

Vous pouvez entrer un appel d'une procédure stockée existante dans l'onglet *Mappage* ou *Where* de l'éditeur de requête.

Lorsque vous entrez un appel de procédure stockée sans utiliser l'Assistant Fonction :

- Faites correspondre les valeurs des paramètres que vous entrez à la signature de la procédure stockée dans Data Services. Les valeurs doivent être spécifiées dans le même ordre que les paramètres sont définis.
- Utilisez `AL_UNSPECIFIED_PARAM` à la place d'un paramètre absent, comme lorsque vous souhaitez utiliser la valeur par défaut d'un paramètre `IN`.
- Utilisez la convention d'affectation des noms qui convient pour la procédure stockée :
 - `<banque de données>.<propriétaire>.<nom_proc>`
 - `<banque de données>.<propriétaire>.<package>.<nom_proc>`.

Si le nom respecte la casse dans la base de données (et non pas tout en majuscules), saisissez le nom tel qu'il apparaît dans la base de données et utilisez les guillemets doubles (") autour du nom pour conserver la casse.

Les deux paramètres de sortie ajoutés par Data Services lors de l'importation, `AL_SP_RETCODE` et `AL_SP_ERRMSG`, sont facultatifs. Il n'est pas nécessaire de fournir d'arguments pour ces deux paramètres si les valeurs qui y sont stockées ne vous intéressent pas.

Rubrique parent : [Appel de procédures stockées \[page 699\]](#)

Informations associées

[Paramètres d'entrée de procédures stockées \[page 700\]](#)

[Appel de procédures stockées à partir de scripts \[page 701\]](#)

[Appel de procédures stockées à partir de requêtes \[page 701\]](#)

[À propos des procédures \[page 690\]](#)

[Appel de procédures stockées \[page 699\]](#)

[Vérification du statut d'exécution \[page 704\]](#)

29.2.5.4.1 Insertion d'une procédure stockée dans le schéma de sortie d'une requête

1. Dans le diagramme d'un flux de données, cliquez sur le nom d'une requête pour ouvrir l'éditeur de requêtes.
2. Dans l'éditeur de requêtes, cliquez avec le bouton droit sur le nom de la requête dans le schéma de sortie et sélectionnez *Nouvel appel de fonction*.

Vous pouvez aussi cliquer avec le bouton droit sur un objet de procédure stockée existant dans la zone du schéma de sortie de l'éditeur de requêtes et sélectionner *Modifier l'appel de fonction*.

SAP Data Services fournit un Assistant Fonction pour vous aider à inclure les paramètres d'entrée et de sortie ainsi que la valeur de renvoi pour un appel de procédure. Les procédures stockées importées sont référencées pour chaque banque de données à la première page de l'Assistant Fonction.

3. Sélectionnez une catégorie de fonction. Notez que les noms de fonctions changent selon la catégorie sélectionnée.
4. Sous le nom de la fonction, sélectionnez une procédure stockée, puis cliquez sur *Suivant*.
5. Dans la fenêtre *Définir les paramètres d'entrée*, saisissez les paramètres d'entrée de la procédure stockée, puis cliquez sur *Suivant*.
6. Dans la fenêtre *Sélectionner le paramètre de sortie*, sélectionnez les paramètres de sortie de la procédure que vous voulez mapper à la sortie de la requête.

Dans la zone *Tous les paramètres de sortie*, sélectionnez les paramètres à mapper, puis cliquez sur la flèche pour les déplacer dans la zone *Paramètres de sortie sélectionnés*.

Vous pouvez mapper plusieurs paramètres INOUT ou OUT et la valeur de renvoi comme colonnes de sortie d'un seul appel de procédure.

7. Cliquez sur *Terminer*.

Le logiciel ajoute ces paramètres au schéma de sortie de la requête. Dans cet exemple, le schéma de sortie de la transformation Query comporte deux colonnes, EMP_NAME et EMP_HIREDATE.

29.2.6 Vérification du statut d'exécution

Pour vous permettre de surveiller le statut d'exécution d'une procédure stockée, Data Services ajoute deux paramètres à la signature de chaque procédure stockée lors de leur importation :

- AL_SP_RETCODE, varchar(256). Trois valeurs sont possibles :
 - ACTA_SP_OK : l'appel de la procédure stockée a abouti.
 - ACTA_SP_DB_CONN_EXCEPTION : la connexion à la base de données n'a pas pu être créée en raison d'une erreur de connexion, d'un utilisateur, mot de passe ou nom d'hôte non valide.
 - ACTA_SP_CALL_ERROR : la connexion a été effectuée, mais l'appel a échoué dans la base de données. Vous trouverez les détails relatifs à l'origine de l'erreur dans AL_SP_ERRMSG.
- AL_SP_ERRMSG, varchar(1024)

Data Services ajoute ces deux paramètres à la fin de la signature, à la suite des paramètres définis par l'utilisateur.

Avec ces paramètres, il existe deux techniques que vous pouvez utiliser pour gérer les erreurs en cas d'échec de la procédure stockée pendant l'exécution :

- Renvoyez une erreur qui arrête immédiatement le job. Un message d'erreur est enregistré dans le fichier `errorlog.txt`.
Pour utiliser cette méthode, ne mappez pas sous forme de paramètres de sortie les deux paramètres ajoutés par Data Services aux procédures stockées importées.
- Enregistrez les messages d'erreur Oracle dans le paramètre `AL_SP_ERRMSG`. Cette méthode permet à la requête, au script, à la structure conditionnelle ou à la fonction personnalisée que contient l'appel de la procédure stockée de poursuivre son exécution.
Pour utiliser cette méthode, définissez les paramètres pour `AL_SP_RETCODE` et `AL_SP_ERRMSG`. Contrôlez la logique d'application à l'aide du statut d'exécution stocké dans `AL_SP_RETCODE`. Par exemple, si vous stockez la valeur de `AL_SP_ERRMSG` dans une variable, vous pouvez utiliser la fonction d'impression pour imprimer la variable dans le journal des jobs. Avec cette méthode, le message d'erreur n'est pas enregistré dans le fichier `errorlog.txt`.

30 Langage de script

Utilisez le langage de script SAP Data Services pour écrire des scripts, des fonctions personnalisées et des expressions, également appelées fonctions du script utilisateur. Vous pouvez aussi utiliser le langage de script pour écrire des expressions telles que les expressions de mappage de colonnes complexe et les conditions de clause WHERE.

Les fonctions personnalisées sont également appelées fonctions du script utilisateur. Utilisez aussi le langage de script pour écrire des expressions telles que les expressions de mappage de colonnes complexe et les conditions de clause WHERE.

Le logiciel utilise le langage de script en interne pour enregistrer les objets créés dans les tables du référentiel. Exportez les objets enregistrés vers un fichier avec l'extension `.atl`. Importez ensuite le fichier dans un autre référentiel. Vous pouvez exporter et importer des objets pour des tâches comme la migration vers un nouvel environnement ou la sauvegarde d'un référentiel avant une mise à niveau.

[Utilisation de la langue de script \[page 706\]](#)

[Syntaxe du langage \[page 707\]](#)

La syntaxe utilisée dans le langage de script peut être utilisée dans une expression et dans un script. Avec le langage de script, vous pouvez affecter des valeurs à des variables, appeler des fonctions et utiliser les opérateurs mathématiques et de chaînes standard.

[Modèles de scripts \[page 715\]](#)

30.1 Utilisation de la langue de script

1. Ouvrez l'éditeur d'un objet dans l'espace de travail de Designer, par exemple un script, une expression conditionnelle ou une transformation.
2. Dans l'un de ces éditeurs, utilisez l'éditeur rapide pour entrer un script ou une expression en utilisant la syntaxe de langage décrite dans cette section.

Présentation des tâches : [Langage de script \[page 706\]](#)

Informations associées

[Syntaxe du langage \[page 707\]](#)

[Modèles de scripts \[page 715\]](#)

[Syntaxe du langage \[page 707\]](#)

[Modèles de scripts \[page 715\]](#)

30.2 Syntaxe du langage

La syntaxe utilisée dans le langage de script peut être utilisée dans une expression et dans un script. Avec le langage de script, vous pouvez affecter des valeurs à des variables, appeler des fonctions et utiliser les opérateurs mathématiques et de chaînes standard.

[Syntaxe des instructions de scripts \[page 707\]](#)

Les instructions de scripts sont les étapes détaillées écrites dans le flux de calcul.

[Syntaxe des références de colonnes et de tables dans les expressions \[page 708\]](#)

Le langage de script Data Services reconnaît les noms de colonnes et de tables sans syntaxe particulière.

[Chaînes \[page 708\]](#)

Cette section aborde trois situations où est utilisée la syntaxe de chaînes.

[Variables \[page 710\]](#)

Les noms de variables doivent être précédés du signe dollar (\$).

[Interpolation de variables \[page 710\]](#)

[Fonctions et procédures stockées \[page 711\]](#)

[Valeurs nulles \[page 712\]](#)

Indique les valeurs nulles à l'aide du mot-clé NULL.

[Debugage et validation \[page 714\]](#)

Rubrique parent : [Langage de script \[page 706\]](#)

Informations associées

[Utilisation de la langue de script \[page 706\]](#)

[Modèles de scripts \[page 715\]](#)

30.2.1 Syntaxe des instructions de scripts

Les instructions de scripts sont les étapes détaillées écrites dans le flux de calcul.

Les instructions d'un objet de script ou d'une fonction personnalisée doivent se terminer par un point-virgule (;). Les lignes de commentaire doivent commencer par un caractère dièse (#).

Informations associées

[Modèles de scripts \[page 715\]](#)

30.2.2 Syntaxe des références de colonnes et de tables dans les expressions

Le langage de script Data Services reconnaît les noms de colonnes et de tables sans syntaxe particulière.

Les expressions désignent une combinaison de constantes, d'opérateurs, de fonctions et de variables évalués sur une valeur d'un certain type de données. Les expressions peuvent être utilisées dans les instructions de scripts ou ajoutées aux objets des flux de données. Comme les expressions peuvent être utilisées dans les objets des flux de données, elles comportent souvent des noms de colonnes.

Aucune syntaxe particulière n'est requise pour les noms de colonne ou de table. Par exemple, vous pouvez indiquer la colonne `start_date` comme entrée d'une fonction de la façon suivante :

```
to_char(start_date, 'dd.mm.yyyy')
```

La colonne `start_date` doit figurer dans le schéma d'entrée de la requête.

S'il existe plusieurs colonnes du même nom dans le schéma d'entrée d'une requête, indiquez laquelle est incluse dans une expression en qualifiant le nom de la colonne avec le nom de la table. Par exemple, indiquez la colonne `start_date` de la table `status` comme suit :

```
status.start_date
```

Notez que les noms de tables et de colonnes faisant partie de chaînes SQL peuvent requérir une syntaxe spéciale selon le SGBD qui évalue l'expression SQL. Par exemple, sélectionnez toutes les lignes de la colonne `LAST_NAME` de la table `CUSTOMER` comme suit :

```
sql('oracle_ds','SELECT CUSTOMER.LAST_NAME  
FROM CUSTOMER')
```

Comme le nom de la colonne est entièrement en majuscules dans Oracle, aucune syntaxe particulière n'est requise. Cependant, sélectionnez toutes les lignes de la colonne `start_date` dans la table Oracle `Status` comme suit :

```
sql('oracle_ds','SELECT "Status"."start_date"  
FROM "Status"')
```

Les noms de tables et de colonnes sont définis dans Oracle avec une casse spéciale. Pour conserver cette casse, mettez les identificateurs entre doubles guillemets.

30.2.3 Chaînes

Cette section aborde trois situations où est utilisée la syntaxe de chaînes.

- Guillemets
- Caractères d'échappement
- Espaces vides de fin

[Guillemets \[page 709\]](#)

Le type de guillemet utilisé dans les chaînes dépend de l'utilisation d'identifiants ou de constantes.

[Caractères d'échappement \[page 709\]](#)

Si une constante contient un guillemet simple ('), une barre oblique inverse (\) ou un autre caractère spécial, ces caractères doivent faire l'objet d'un échappement.

[Espaces vides de fin \[page 710\]](#)

Data Services ne supprime pas les espaces vides de fin des chaînes utilisées dans les scripts ou les fonctions personnalisées. Pour supprimer les espaces vides de fin, utilisez la fonction `rtrim` ou `rtrim_blank`.

Informations associées

[Guillemets \[page 709\]](#)

[Caractères d'échappement \[page 709\]](#)

[Espaces vides de fin \[page 710\]](#)

[Valeurs nulles et chaînes vides \[page 712\]](#)

30.2.3.1 Guillemets

Le type de guillemet utilisé dans les chaînes dépend de l'utilisation d'identifiants ou de constantes.

Identificateur	Nom d'un objet (par exemple, table, colonne, flux de données ou fonction).
Constante	Valeur fixe utilisée dans les calculs. Il existe deux types de constantes : • Constantes de chaîne (par exemple, 'Hello World' ou '1995.01.23') • Constantes numériques (par exemple, 2.14)

Les identificateurs requièrent des guillemets s'ils contiennent des caractères spéciaux (non alphanumériques). Par exemple, `"calculer des grands nombres"` requiert des guillemets doubles en raison des espaces vides.

Utilisez des guillemets simples pour les constantes de chaîne.

30.2.3.2 Caractères d'échappement

Si une constante contient un guillemet simple ('), une barre oblique inverse (\) ou un autre caractère spécial, ces caractères doivent faire l'objet d'un échappement.

Par exemple, les caractères suivants doivent être précédés d'un caractère d'échappement pour être évalués correctement dans une chaîne. Data Services utilise la barre oblique inverse (\) comme caractère d'échappement.

Caractère	Exemple
single quote (')	'World\'s Books'
backslash (\)	'C:\\temp'

30.2.3.3 Espaces vides de fin

Data Services ne supprime pas les espaces vides de fin des chaînes utilisées dans les scripts ou les fonctions personnalisées. Pour supprimer les espaces vides de fin, utilisez la fonction `rtrim` ou `rtrim_blank`.

30.2.4 Variables

Les noms de variables doivent être précédés du signe dollar (\$).

- Les variables locales utilisées dans un script ou une expression doivent être définies dans le job ou le workflow qui appelle le script à l'aide de la fenêtre [Variables et paramètres](#).
- Les variables globales utilisées dans un script ou une expression doivent être définies au niveau du job à l'aide de la fenêtre [Variables et paramètres](#).
- La valeur renvoyée doit être transmise hors de la fonction en utilisant l'instruction suivante :

```
RETURN (<expression>)
```

où expression définit la valeur devant être renvoyée.

- Les variables locales utilisées dans une fonction personnalisée doivent être définies à l'aide de l'éditeur rapide.
- Les variables et paramètres existants affichés dans l'éditeur rapide sont filtrés par le contexte à partir duquel est ouvert l'éditeur rapide.

Informations associées

[Guide de Designer : Variables et paramètres \[page 308\]](#)

[Création d'une fonction personnalisée \[page 682\]](#)

30.2.5 Interpolation de variables

Vous pouvez incorporer des expressions dans des chaînes de constantes ; Data Services évalue les variables et remplace la valeur dans la chaîne. L'opérateur de concaténation (||) n'est pas requis.

Voici un exemple d'utilisation de l'opérateur de concaténation :

```
$st_date=sql('warehouse_ds','SELECT max(start_timestamp)
FROM dw_process_log
WHERE table_name=\'\'||$Table_Name||\'\'
print('The value of the start date
is: \'||sysdate()+5);
```

L'instruction peut être simplifiée comme suit :

```
$st_date=sql('warehouse_ds','SELECT max(start_timestamp)
FROM dw_process_log
WHERE table_name={$Table_Name}'
print('The value of the start date
is:[sysdate()+5]');
```

Des accolades ({}) et des crochets ([]) entourent les expressions incorporées :

- Les crochets ([]) indiquent que la valeur de l'expression doit être remplacée.
- Les accolades ({}) indiquent que la valeur de l'expression doit être entre guillemets simples.

Les chaînes comprenant des accolades ou des crochets provoquent une erreur de traitement. Vous pouvez contourner ce problème en faisant précéder les accolades ou crochets d'une barre oblique inverse (\).

30.2.6 Fonctions et procédures stockées

Un script, une expression ou une fonction personnalisée peut appeler la plupart des fonctions intégrées ou personnalisées. Une fonction ne peut pas s'appeler elle-même ou appeler une autre fonction qui produirait un appel récursif. Par exemple, la fonction A ne peut pas appeler la fonction A et la fonction A ne peut pas appeler la fonction B si la fonction B appelle la fonction A.

Vous pouvez également utiliser les fonctions et procédures stockées importées d'un SGBD. Utilisez la casse exacte des noms de fonctions et procédures stockées importées.

- Les fonctions intégrées sont répertoriées par catégorie dans l'Assistant Fonction et l'éditeur rapide.
- Vous trouverez la liste des fonctions personnalisées dans l'Assistant Fonction et l'éditeur rapide.
- Les fonctions et procédures stockées importées sont répertoriées dans l'Assistant Fonction et l'éditeur rapide sous la banque de données utilisée pour les importer.

Informations associées

[Création d'une fonction personnalisée \[page 682\]](#)

[Importation de métadonnées pour des procédures stockées \[page 697\]](#)

30.2.7 Valeurs nulles

Indique les valeurs nulles à l'aide du mot-clé `NULL`.

Par exemple, vous pouvez vérifier si une colonne (`COLX`) est nulle ou pas :

```
COLX IS NULL
COLX IS NOT NULL
```

Le logiciel ne vérifie pas les valeurs nulles dans les colonnes de données. Utilisez la fonction `nvl` pour supprimer les valeurs nulles.

30.2.7.1 Valeurs nulles et chaînes vides

Cette section aborde les règles concernant les valeurs `NULLES` et les chaînes vides.

Les deux règles s'appliquant aux chaînes vides sont les suivantes.

- Lorsque vous affectez une chaîne vide à une variable de type `varchar`, la valeur de la variable sera traitée sous la forme d'une chaîne de longueur zéro. Une erreur se produit si vous affectez une chaîne vide à une variable qui n'est pas de type `varchar`. Pour affecter une valeur `null` à une variable d'un type quelconque, utilisez la constante `NULL`.
- En tant que constante (`' '`), le logiciel traite la chaîne vide sous la forme d'une valeur `varchar` de longueur zéro. Utilisez la constante `NULL` pour la valeur nulle.

i Remarque

Oracle ne fait pas de distinction entre une chaîne vide et une valeur `null`. Lorsque vous insérez une chaîne vide ou une valeur `null` dans une colonne `varchar`, Oracle traite la chaîne vide et la valeur `null` comme des valeurs `nulles`. Par conséquent, le logiciel traite la valeur sous forme d'une valeur `NULLE`.

Les trois règles suivantes sont celles des valeurs `NULLES` et les chaînes vides dans les expressions conditionnelles :

- Par rapport à une valeur nulle, les opérateurs de comparaison Egal à (`=`) et Pas égal à (`<>`) sont toujours évalués sur `FALSE`. Ce résultat `FALSE` inclut la comparaison d'une variable ayant une valeur `null` par rapport à une constante `NULL`.
La table suivante affiche les résultats de comparaison pour les affectations de variables `$var1 = NULL` et `$var2 = NULL` :

Condition	Se traduit par	Renvoie
<code>If (NULL = NULL)</code>	<code>NULL</code> est égal à <code>NULL</code>	<code>FALSE</code>
<code>If (NULL != NULL)</code>	<code>NULL</code> n'est pas égal à <code>NULL</code>	<code>FALSE</code>
<code>If (NULL = '')</code>	<code>NULL</code> est égal à une chaîne vide	<code>FALSE</code>
<code>If (NULL != '')</code>	<code>NULL</code> n'est pas égal à une chaîne vide	<code>FALSE</code>
<code>If ('bbb' = NULL)</code>	<code>bbb</code> est égal à <code>NULL</code>	<code>FALSE</code>

Condition	Se traduit par	Renvoie
If ('bbb' != NULL)	bbb n'est pas égal à NULL	FALSE
If ('bbb' = '')	bbb est égal à une chaîne vide	FALSE
If ('bbb' != '')	bbb n'est pas égal à une chaîne vide	TRUE
If (\$var1 = NULL)	NULL est égal à NULL	FALSE
If (\$var != NULL)	NULL n'est pas égal à NULL	FALSE
If (\$var1 = '')	NULL est égal à une chaîne vide	FALSE
If (\$var != '')	NULL n'est pas égal à une chaîne vide	FALSE
If (\$var1 = \$var2)	NULL est égal à NULL	FALSE
If (\$var != \$var2)	NULL n'est pas égal à NULL	FALSE

La table suivante affiche les résultats de comparaison pour les affectations de variable `$var1 = ''` et `$var2 = ''`:

Condition	Se traduit par	Renvoie
If (\$var1 = NULL)	Chaîne vide est égal à NULL	FALSE
If (\$var != NULL)	Chaîne vide n'est pas égal à NULL	FALSE
If (\$var1 = '')	Chaîne vide est égal à une chaîne vide	TRUE
If (\$var != '')	Chaîne vide n'est pas égal à une chaîne vide	FALSE
If (\$var1 = \$var2)	Chaîne vide est égal à Chaîne vide	TRUE
If (\$var != \$var2)	Chaîne vide n'est pas égal à Chaîne vide	FALSE

- Utilisez les opérateurs IS NULL et IS NOT NULL pour tester la présence de valeurs nulles. Par exemple, en supposant qu'une variable est affectée : `$var1 = NULL`;

Condition	Se traduit par	Renvoie
If ('bbb' IS NULL)	bbb est NULL	FALSE
If ('bbb' IS NOT NULL)	bbb n'est pas NULL	TRUE
If ('' IS NULL)	Chaîne vide est NULL	FALSE
If ('' IS NOT NULL)	Chaîne vide n'est pas NULL	TRUE
If (\$var1 IS NULL)	NULL est NULL	TRUE
If (\$var IS NOT NULL)	NULL n'est pas NULL	FALSE

- Lors de la comparaison de deux variables, testez toujours les valeurs nulles. Dans ce scénario, vous ne testez pas une variable ayant une valeur nulle par rapport à une constante NULL (comme dans la première règle). Testez chaque variable et chaque branche en conséquence ou testez dans l'expression conditionnelle, comme illustré dans la deuxième ligne de la table ci-dessous.

Condition	Recommandation
<code>if (\$var1 = \$var2)</code>	Ne comparez pas sans tester explicitement les valeurs nulles. L'utilisation de cette logique n'est pas recommandée car toute comparaison relationnelle avec une valeur nulle renvoie FALSE.
<code>if (((\$var1 IS NULL) AND (\$var2 IS NULL)) OR (\$var1 = \$var2))</code>	Exécutera la branche VRAI si \$var1 et \$var2 ont une valeur nulle ou si aucune des deux n'a de valeur nulle, mais qu'elles sont égales.

30.2.8 Debugage et validation

Si vous utilisez l'éditeur rapide pour créer un script ou une fonction personnalisée, sélectionnez l'icône [Valider](#) dans la barre d'outils (ou cliquez avec le bouton droit et sélectionnez [Valider](#)). Les erreurs sont listées dans la fenêtre de l'éditeur rapide en dessous de la zone de texte. Double-cliquez sur chaque erreur pour savoir où elle s'est produite dans le script.

Autres possibilités :

- Sélectionnez la commande [► Deboguer ► Valider ► Vue actuelle ►](#) pour contrôler la syntaxe des expressions utilisées dans la définition de l'objet en cours. Les erreurs sont affichées dans la fenêtre Sortie.
- Sélectionnez la commande [► Deboguer ► Valider ► Tous les objets de la vue ►](#) pour contrôler la syntaxe des expressions utilisées dans la définition de l'objet en cours et les définitions de tous les objets appelés par l'objet en cours. Les erreurs sont affichées dans la fenêtre Sortie.

Par défaut, Data Services répertorie dix erreurs avant d'interrompre l'analyse. En signalant dix erreurs pour chaque validation, Data Services permet de raccourcir le cycle d'édition-validation-test. Il est possible de réparer plusieurs erreurs lors de chaque itération du cycle.

Pour modifier ce paramètre par défaut, sélectionnez [► Outils ► Options ► Job Server ► Général ►](#) et définissez les options comme suit :

Section :	Entrez "Parser".
Clé :	Entrez "NumErrors".
Valeur :	Entrez un nombre positif.

Pour chaque erreur, Data Services fournit une description du problème.

Data Services fournit également trois options de menu pour les erreurs. Cliquez avec le bouton droit sur une erreur pour utiliser ces options de menu :

Afficher : l'option [Afficher](#) permet d'afficher le texte de l'erreur dans une fenêtre distincte plus petite pour faciliter la lecture.

Accéder à l'erreur : l'option [Accéder à l'erreur](#) place une icône d'erreur près de la ligne qui a provoqué l'erreur. Vous pouvez aussi double-cliquer sur l'erreur pour placer cette icône. Cette option n'est cependant pas disponible pour toutes les erreurs.

Copier : l'option [Copier](#) copie le texte de l'erreur dans le presse-papiers pour pouvoir s'en servir ailleurs.

Lorsque vous exécutez un job, vous pouvez rencontrer dans une expression des erreurs qui sont renvoyées depuis le système d'exploitation ou le SGBD. Lors de l'exécution d'un job, Data Services envoie au SGBD le plus d'opérations possible à évaluer. Le SGBD peut évaluer une partie de l'expression et renvoyer une valeur qui contribue à évaluer le reste de l'expression.

30.3 Modèles de scripts

Les exemples suivants illustrent la mise en forme d'un script pour les fonctions Square et Repeat String.

Rubrique parent : [Langage de script \[page 706\]](#)

Informations associées

[Utilisation de la langue de script \[page 706\]](#)

[Syntaxe du langage \[page 707\]](#)

30.3.1 Fonction Square

Cette fonction accepte un entier et renvoie le carré de la valeur d'entrée. Pour définir ce script, fournissez les valeurs suivantes à Data Services :

Description	Valeur	Arguments
Nom de la fonction	Square	
Valeur renvoyée	Return	int
Valeur d'entrée	input_int	int, input

Le texte du script de la fonction est le suivant :

```
RETURN($input_int * $input_int);
```

30.3.2 Fonction RepeatString

Cette fonction accepte une chaîne d'un caractère et un entier indiquant le nombre de fois où le caractère d'entrée doit être répété. Elle sort la chaîne créée ou une valeur nulle si le caractère d'entrée ou le nombre de répétitions est NULL. Elle utilise une variable pour suivre les composants de la chaîne lors de sa génération.

Pour définir ce script, fournissez les valeurs suivantes à Data Services :

Description	Valeur	Arguments
Nom de la fonction	RepeatString	
Valeur renvoyée	Return	VARCHAR (255),allow NULL
Chaîne d'entrée	InputCharacter	VARCHAR (1), input, allow NULL
Entier d'entrée	RepeatNumber	INT, input, allow NULL
Variable interne	Partial	VARCHAR (255), allow NULL

Le texte de la fonction est le suivant :

```

IF ($InputCharacter = NULL)
    BEGIN
        $Partial = NULL;
        RETURN NULL;
    END
IF ($RepeatNumber = NULL)
    RETURN NULL;
IF ($RepeatNumber > 255)
    $RepeatNumber = 255;
IF ($RepeatNumber < 0)
    $RepeatNumber = 0;
BEGIN
    WHILE ($RepeatNumber != 0)
    BEGIN
        $Partial = ($Partial || $InputCharacter);
        $RepeatNumber = ($RepeatNumber - 1);
    END
    RETURN $Partial;
END

```

31 Fonctionnement de l'extraction

Le processus d'extraction utilise sa connaissance inhérente de la sémantique des mots et du contexte linguistique dans lequel ces mots apparaissent pour rechercher des entités et des faits.

Il crée des modèles spécifiques pour extraire des entités et des faits en fonction de règles système. Vous pouvez ajouter des entrées dans un dictionnaire et écrire des règles personnalisées pour personnaliser une sortie d'extraction.

L'exemple suivant indique comment un contenu non structuré peut être transformé en une information structurée pour être traité et analysé ultérieurement.

❖ Exemple

Texte d'exemple et informations relatives à l'extraction

Exemple de vos données : "Monsieur Jones est très ennuyé par rapport à Green Insurance Corp. L'offre relative à son véhicule sinistré n'est pas assez importante. Il déclare que Soleil lui a proposé 1 250 euros mais sa voiture est estimée partout entre 2 500 et 4 500 euros. M. Dupont souhaiterait que Soleil s'aligne sur d'autres concurrents."

La transformation Entity Extraction traite le texte en configurant le sentiment avec des règles personnalisées. Elle identifie le groupe d'une manière logique pour identifier les différentes parties de l'entité et en identifiant les faits.

Les tableaux suivants indiquent des résultats partiels avec des informations marquées comme entité, types et sous-types.

Entité	Type	Sous-type
M. Dupont	Personne	N/A
Green Insurance	Organisation	Commercial
Green	Divers_propriété	N/A
1 250,00 euros	Devise	

Le tableau suivant indique des résultats partiels avec des informations marquées comme faits, types, sous-faits et types de sous-faits.

Fait	Type	Sous-fait	Type de sous-fait
M. Jones est très ennuyé par rapport à Green Insurance Corp.	Sentiment	Très ennuyé	SentimentNégatifFort

Fait	Type	Sous-fait	Type de sous-fait
Jones souhaiterait que Green s'aligne sur d'autres concurrents.	Demande	N/A	N/A

31.1 Utilisation de la transformation Entity Extraction

Utilisez la transformation Entity Extraction pour extraire des informations à partir d'un contenu au format texte, HTML, XML, ou binaire (tel que PDF) et générer une sortie structurée.

Utilisez la sortie générée de plusieurs façons en fonction de votre workflow. Utilisez-la comme entrée dans une autre transformation, ou écrivez dans plusieurs sources de sortie, par exemple une table de base de données ou un fichier plat. La transformation génère la sortie dans l'encodage `UTF-16`. La liste suivante fournit des exemples pour indiquer le moment où utiliser la transformation Entity Extraction:

- Pour rechercher des informations et de relations spécifiques issues d'une quantité importante de texte liée à un domaine générale. Par exemple, une société est intéressée par l'analyse des commentaires client reçus dans un texte en forme libre suite au lancement d'un nouveau produit.
- Pour relier des informations structurées à partir d'un texte non structuré avec des informations structurées existantes pour effectuer de nouvelles connexions. Par exemple, un service de poursuite pénale essaie d'établir des connexions entre divers crimes et des personnes impliquées à l'aide de sa propre base de données et des informations disponibles dans divers rapports au format texte.
- Pour analyser et établir des rapports sur les problèmes de qualité d'un produit, comme les réparations et retours excessifs pour certains produits, en effectuant des connexions entre diverses formes de données. Par exemple, vous pourriez avoir des informations de produits dans un format de données structurées, comme des produits, des pièces, des clients et des fournisseurs. Mais d'autres données contiennent du texte concernant les problèmes liés aux produits et les réclamations client. Une troisième source de données contient des enregistrements de maintenance, journaux de réparation, escalades produit et journaux de centre de support.

31.2 Entités, types, sous-types et faits

La transformation Entity Extraction extrait des informations de vos données et crée des types et sous-types d'entités. Elle catégorise ensuite les données dans différents types de faits.

Entités

Dans Data Services, les entités sont les noms des personnes, les lieux et les choses que la transformation Entity Extraction extrait de vos données.

❖ Exemple

Entités

Entité	Type
Paris	Localité
M. Joe Smith	Personne

Sous-types

Les entités peuvent avoir des sous-types, lesquels sont des sous-catégories d'une entité. Les sous-types indiquent une autre classification d'une entité. Les sous-types sont des spécifications hiérarchiques d'un type d'entité permettant de distinguer différentes variétés sémantiques du même type d'entité.

❖ Exemple

Types et sous-types

Entité	Type	Sous-type
Boeing 747	Véhicule	Air
Mercedes-Benz SL500	Véhicule	Terre
SAP	Organisation	Commercial

Faits

Les entités par elles-mêmes indiquent uniquement le fait qu'elles soient présentes dans un document. Les faits fournissent des informations sur la manière dont les entités sont liées. Le terme « Fait » couvre plusieurs significations : modèles complexes incluant une ou plusieurs entités, une relation entre une ou plusieurs entités ou un prédicat concernant une entité.

Les faits indiquent un modèle. La transformation convertit le modèle en règles d'extraction qui extraient des informations comme les sentiments, les événements ou les relations. Les faits indiquent dans quel contexte différentes entités sont reliées dans le texte. La transformation Entity Extraction utilise également les règles d'extraction personnalisées que vous créez.

Tout comme les entités, les faits ont des types et des sous-types.

- Les types de faits identifient la catégorie d'un fait ; par exemple, des sentiments et des demandes.
- Les sous-types de faits ou sous-faits sont des éléments d'informations clés intégrés dans un fait.

- Les types de sous-faits sont comme des catégories associées au sous-fait.

❖ Exemple

Exemples de faits et de types de faits

Fait	Type
SAP a acquis Business Objects dans de bonnes conditions.	Fusion et acquisition
Monsieur Joe Smith est préoccupé par ses réservations auprès de la compagnie aérienne.	Sentiment

31.3 Différences entre traitement des données de texte et transformations Data Cleanse

La transformation Entity Extraction propose une fonctionnalité similaire à la transformation Data Cleanse dans certains cas, notamment concernant les fonctions de personnalisation.

La documentation décrit les différences entre les deux transformations et celle qui doit être utilisée conformément à vos objectifs. La transformation Text Data Processing et la transformation Entity Extraction sont logiques pour un contenu non structuré et la transformation Data Cleanse permet de standardiser et de nettoyer des données structurées. Le tableau suivant décrit certaines des principales différences. Dans de nombreux cas, l'utilisation d'une combinaison transformations Entity Extraction/Data Cleanse générera les données les mieux adaptées à vos analyses et rapports Business Intelligence.

Critères	Text Data Processing	Data Cleanse
Type d'entrée	Texte non structuré nécessitant une analyse linguistique pour générer des informations pertinentes.	Données structurées représentées comme champs d'enregistrements.
Taille d'entrée	Supérieure à 5 Ko du texte.	Inférieure à 5 Ko du texte.
Périmètre d'entrée	Normalement domaine général avec variations.	Domaine de données spécifique avec variations limitées.
Tâche correspondante	Découverte de contenu, réduction bruit, correspondance de modèle et relation entre différentes entités.	Recherche dans le dictionnaire, reconnaissance des dépendances.
Utilisation potentielle	Identifie les informations potentiellement significatives à partir d'un contenu non structuré et les extrait sous un format pouvant être archivé dans un référentiel.	Garantit la qualité des données pour la correspondance et l'archivage dans un référentiel tel que Meta Data Management.

Critères	Text Data Processing	Data Cleanse
Sortie	Crée des annotations sur le texte source sous forme d'entités, de types d'entités, de faits et leur décalage, longueur etc. L'entrée n'est pas modifiée.	Crée des champs analysés et standardisés. L'entrée est modifiée si nécessaire.

32 Conception et débogage

Cette section traite des fonctionnalités suivantes de Designer pouvant être utilisées pour concevoir et déboguer les jobs :

- Utilisez la fonctionnalité **Afficher les cas d'emploi** pour déterminer l'impact de la modification d'un objet de métadonnées (par exemple, sur une table). Voyez quels flux de données utilisent le même objet.
- Utilisez la fonctionnalité **Afficher les données** pour visualiser des échantillons de données source, de transformation et cible dans un flux de données après exécution d'un job.
- Utilisez la fonctionnalité **Affichage des données lors de la conception** pour visualiser et analyser l'entrée et la sortie d'un jeu de données en temps réel lors de la conception d'une transformation.
- Utilisez le **Débogueur interactif** pour définir des points de contrôle et des filtres entre les transformations dans un flux de données et afficher les données de job ligne par ligne durant l'exécution d'un job.
- Utilisez l'**Affichage des différences** pour comparer les métadonnées d'objets similaires et leurs propriétés.
- Utilisez la fonctionnalité **Audit du flux de données** pour vérifier que les bonnes données sont traitées par un objet source, de transformation ou cible.

Informations associées

[Utilisation de Afficher les cas d'emploi \[page 722\]](#)

[Utilisation de l'affichage des données \[page 727\]](#)

[Utilisation de l'affichage des données lors de la conception \[page 738\]](#)

[Utilisation du débogueur interactif \[page 741\]](#)

[Comparaison d'objets \[page 755\]](#)

[Utilisation de l'Audit \[page 412\]](#)

32.1 Utilisation de Afficher les cas d'emploi

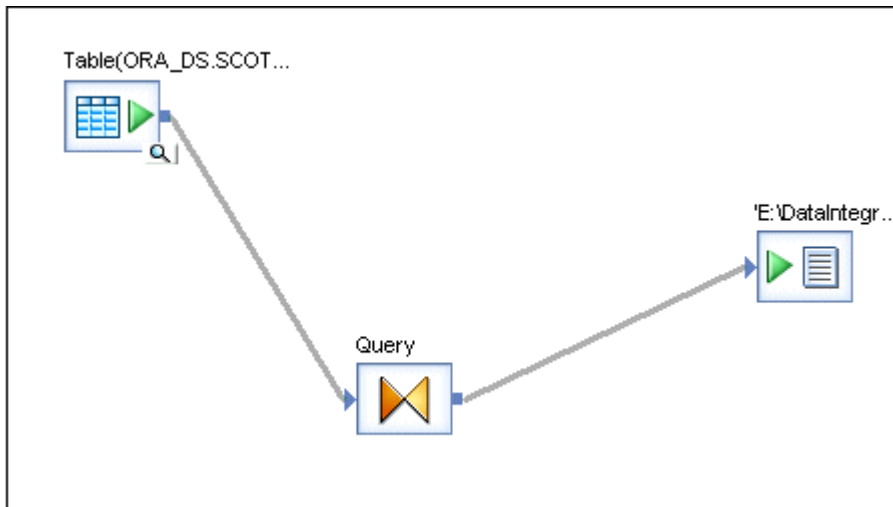
Lorsque vous enregistrez un job, un workflow ou un flux de données, le logiciel enregistre une liste des objets liés dans votre référentiel.

Le logiciel préserve les données de la relation parent-enfant. Vous pouvez consulter les relations à l'aide du menu **Outils** (► **Outils** ► **Afficher utilisation** ►). Il est également possible d'accéder aux informations en cliquant sur l'icône **Afficher utilisation** dans la barre d'outils.

Par exemple, lorsque le flux de données parent suivant est enregistré, le logiciel enregistre également les connexions entre celui-ci et ses trois enfants :

- une table source
- une transformation Query

- un fichier cible



Vous pouvez utiliser ces données de relation parent/enfant pour déterminer par exemple l'impact d'une modification éventuelle sur les autres flux de données qui utilisent la même objet.

Par exemple, lors de la gestion d'un flux de données, vous pouvez avoir besoin de supprimer une définition de table source et de réimporter la table (ou de modifier son schéma). Avant de le faire, recherchez tous les flux de données utilisant la même table et mettez-les à jour si nécessaire.

Designer compte l'utilisation d'un objet comme le nombre de fois qu'il est utilisé dans un but unique.

Par exemple, dans un flux de données intitulé DF1, si la table DEPT est utilisée deux fois comme source et une fois comme cible, le nombre dans la colonne *Utilisation* de la bibliothèque d'objets est 2. Ce cas doit être rare. Par exemple, une table n'est pas souvent jointe à elle-même dans la conception d'un job.

Informations associées

[Limitations \[page 726\]](#)

32.1.1 Accès à Afficher utilisation depuis la bibliothèque d'objets

Le Designer de la bibliothèque d'objets contient une colonne *Utilisation* qui affiche le nombre de relations parent-enfant dans un objet.

Affichage des relations parent-enfant pour les types d'objets suivants :

- Jobs
- Workflows
- Flux de données
- Transformations (Data Quality, Traitement des données de texte et ABAP dans le dossier plateforme)

- Banques de données
- Format

Lorsque la valeur d'utilisation est supérieure à zéro, vous pouvez afficher des informations de relations.

1. Cliquez avec le bouton droit sur le nom d'objet dans la bibliothèque d'objets et sélectionnez [Afficher utilisation](#) dans le menu déroulant.

La boîte de dialogue [Sortie](#) s'ouvre dans la partie inférieure de l'espace de travail.

- L'onglet [Informations](#) contient des lignes pour chaque parent de l'objet sélectionné.
- Les colonnes affichent le type d'objet et son nom.
- La colonne [En tant](#) indique comment l'objet sélectionné est utilisé par le parent.

2. Double-cliquez sur un objet parent pour l'afficher dans l'espace de travail.

L'espace de travail affiche le diagramme d'objet avec l'objet enfant mis en surbrillance.

Une fois qu'un parent est ouvert dans l'espace de travail, vous pouvez double-cliquer à nouveau sur une ligne de la fenêtre [Sortie](#) afin d'afficher les informations d'utilisation pour un autre objet.

Objet	L'espace de travail affiche
Parent	Diagramme de l'objet parent
Enfant	S'il s'agit d'un objet enfant du même parent, l'objet mis en surbrillance passe à l'enfant dans le diagramme.

→ Conseil

C'est une option importante, parce qu'un objet enfant de la boîte de dialogue [Sortie](#) peut ne pas correspondre au nom utilisé dans son parent. Vous pouvez personnaliser les noms d'objets de l'espace de travail pour les sources et les cibles. Le progiciel enregistre le nom utilisé dans chaque parent et le nom utilisé dans la bibliothèque d'objets. L'onglet [Informations](#) de la fenêtre [Sortie](#) affiche le nom utilisé dans la bibliothèque d'objets. Les noms des objets utilisés dans les parents sont visibles uniquement en ouvrant le parent dans l'espace de travail.

3. Sélectionnez un objet sur une autre ligne dans la boîte de dialogue [Sortie](#). L'espace de travail se modifie afin d'afficher le diagramme de l'objet sélectionné.

Informations associées

[À propos de la boîte de dialogue Sortie \[page 725\]](#)

[Limitations \[page 726\]](#)

[Accès à Afficher utilisation depuis la fenêtre Variables et paramètres \[page 314\]](#)

32.1.1.1 À propos de la boîte de dialogue Sortie

Lorsque vous décidez d'apporter des modifications à un objet spécifique, les informations contenues dans la boîte de dialogue *Sortie* d'*Afficher utilisation* peuvent être particulièrement utiles.

Cliquez avec le bouton droit sur un nom d'objet ayant un chiffre supérieur à zéro (0) dans la colonne *Utilisation* de la bibliothèque d'objets du Designer afin d'ouvrir la boîte de dialogue *Sortie*.

L'onglet *Informations* de la boîte de dialogue *Sortie* affiche les informations d'utilisation. Il contient une ligne pour chaque parent de l'objet sélectionné. La boîte de dialogue affiche le type et le nom de l'objet sélectionné dans l'en-tête de la première colonne. La colonne *En tant que* fournit davantage de contexte sur la façon dont l'objet sélectionné est utilisé par le parent, le cas échéant. Par exemple, les fichiers XML et messages, tables, fichiers plats, etc., peuvent avoir une valeur de « Source » ou « Cible » dans la colonne *En tant que*, ce qui signifie que l'objet est utilisé en tant que source ou cible.

Les tables suivantes affichent d'autres valeurs possibles dans la colonne *En tant que*.

Fichiers plats et tables uniquement

En tant que	Description
Lookup()	Table/fichier de recherche utilisé dans une fonction lookup
Lookup_ext()	Table/fichier de recherche utilisé dans une fonction lookup_ext
Lookup_seq()	Table/fichier de recherche utilisé dans une fonction lookup_seq

Tables uniquement

En tant que	Description
Comparaison	Table utilisée dans la transformation Table Comparison
Génération de clé	Table utilisée dans la transformation Key Generation

32.1.2 Accès à Afficher les cas d'emploi depuis l'espace de travail

Depuis le diagramme ouvert d'un objet de l'espace de travail (un flux de données par exemple), vous pouvez voir si un objet parent ou enfant est utilisé.

La fenêtre Sortie répertorie	la façon d'accéder à Afficher utilisation
Jobs utilisant un flux de données spécifique	Ouvrez un flux de données dans l'espace de travail. Sélectionnez ► Afficher ► Utilisation de la barre de menus ou sélectionnez le bouton Afficher utilisation dans la barre d'outils.
Liste d'objets parents utilisant un objet sélectionné dans un flux de données.	Ouvrez un flux de données dans l'espace de travail et sélectionnez l'objet de flux de données applicable. Sélectionnez ► Afficher ► Utilisation de la barre de menus ou sélectionnez le bouton Afficher utilisation dans la barre d'outils.

Informations associées

[Accès à Afficher utilisation depuis la fenêtre Variables et paramètres \[page 314\]](#)

32.1.3 Limitations

Points importants à savoir lors de l'utilisation d'**Afficher utilisation**.

- Le logiciel ne prend pas en charge **Afficher utilisation** dans les référentiels centraux.
- L'onglet **Informations** de la fenêtre **Sortie** affiche uniquement les paires parents/enfants.

❖ Exemple

Par exemple, pour une table, un flux de données est le parent. Si la table est également utilisée par un grand-parent (un workflow, par exemple), le workflow n'est pas répertorié dans l'affichage de la fenêtre **Sortie** pour une table. Pour voir la relation entre un flux de données et un workflow, ouvrez le workflow dans l'espace de travail, puis cliquez avec le bouton droit sur un flux de données et sélectionnez l'option **Afficher utilisation**.

- Le logiciel n'enregistre pas les relations parent/enfant entre les fonctions. Exemple :

❖ Exemple

Si la fonction A appelle la fonction B et que la fonction A n'est dans aucun flux de données ou script, la zone **Utilisation** de la bibliothèque d'objets affichera zéro pour les deux fonctions. Le fait que la fonction B soit utilisée une fois dans la fonction A n'est pas pris en compte.

- ❖ Exemple

Si la fonction A est enregistrée dans un flux de données, la colonne *Utilisation* de la bibliothèque d'objets affiche 1 pour les deux fonctions.

32.2 Utilisation de l'affichage des données

L'affichage des données fournit un moyen d'analyser et capturer un échantillon des données produites par chaque étape d'un job, même si son exécution échoue. Affichez les données source importées, les données modifiées des transformations et les données finales des cibles. Après l'importation d'une source de données, vous pouvez contrôler à tout moment le statut de ces données, avant et après traitement des flux de données.

Utilisez l'affichage des données pour contrôler les données lors des jobs de conception et de test afin de vous assurer que votre conception renvoie les résultats attendus. À l'aide d'un ou plusieurs volets d'affichage des données, vous pouvez afficher et comparer les échantillons de données de différentes étapes. Les informations d'affichage des données sont affichées dans des panneaux intégrés pour faciliter la navigation entre les flux et les données.

Utilisez l'affichage des données pour examiner :

- Les sources et les cibles
L'affichage des données permet de voir les données avant d'exécuter un job. Fort des détails des données, vous pouvez concevoir des jobs de qualité supérieure. Vous pouvez parcourir et analyser les données des tables et fichiers de la bibliothèque d'objets et aussi voir les données de ces mêmes objets dans des jobs existants. Après exécution du job, vous pouvez à nouveau vous reporter aux données source.
- Les transformations
- Les lignes d'un diagramme

i Remarque

- L'affichage des données affiche les données BLOB comme <blob>.
- L'affichage des données n'est pas pris en charge pour les IDocs de SAP. Pour SAP et PeopleSoft, les options de l'onglet Profil de table et Profil de colonne ne sont pas prises en charge pour les hiérarchies.

Informations associées

[Affichage des données passées par les transformations \[page 753\]](#)

[Utilisation du débogueur interactif \[page 741\]](#)

32.2.1 Accessing View Data

32.2.1.1 Affichage des données des sources et cibles

Vous pouvez afficher les données des sources et cibles depuis deux emplacements différents :

1. [Le bouton Afficher les données](#)

Le bouton Afficher les données apparaît sur les objets source et cible lorsque vous les faites glisser dans l'espace de travail. Cliquez sur le bouton Afficher les données (icône loupe) pour ouvrir un volet d'affichage des données pour cet objet source ou cible.

2. [La bibliothèque d'objets](#)

Affichez les données des objets source ou cible potentiels des onglets Banques de données ou Formats.

Ouvrez un volet d'affichage des données depuis la bibliothèque d'objets à l'aide d'une des méthodes suivantes :

- Cliquez avec le bouton droit sur un objet de table et sélectionnez [Afficher les données](#).
- Cliquez avec le bouton droit sur une table et sélectionnez [Ouvrir](#) ou [Propriétés](#).

La fenêtre Métadonnées de table, Editeur de format XML ou Propriétés s'ouvre. Vous pouvez sélectionner l'onglet Afficher les données depuis l'une de ces fenêtres.

Pour afficher les données d'un fichier, celui-ci doit exister physiquement et être disponible depuis le système d'exploitation de votre ordinateur. Pour afficher les données d'une table, elle doit provenir d'une base de données prise en charge.

Informations associées

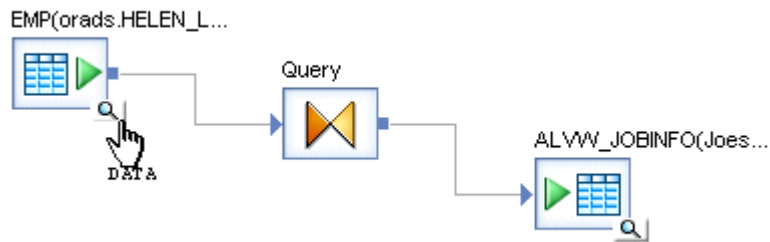
[Affichage des données dans l'espace de travail \[page 728\]](#)

32.2.2 Affichage des données dans l'espace de travail

L'affichage des données est accessible dans l'espace de travail quand les boutons loupe apparaissent sur les objets qualifiés d'un flux de données. Cela signifie :

Pour les sources et cibles, les fichiers doivent exister physiquement et être accessibles depuis Designer, les tables doivent provenir d'une base de données prise en charge.

Pour ouvrir un volet d'affichage des données dans l'espace de travail de Designer, cliquez sur le bouton loupe d'un objet de flux de données.



Un grand volet d'affichage des données s'ouvre sous la zone de l'espace de travail. Cliquez sur le bouton loupe d'un autre objet et un autre volet s'affiche en-dessous de la zone d'espace de travail (notez que la première zone de volet se réduit, afin de permettre l'affichage du second volet).

ALVW_PARENT_CHILD(Joes.DI_REPO)

PARENT_OBJ	PARENT_OBJ_TYPE	PARENT_DE
New_DataFlow	DataFlow	No descriptio
New_DataFlow	DataFlow	No descriptio
RT_TestConnectivity	DataFlow	No descriptio
RT_TestConnectivity	DataFlow	No descriptio
RT_TestConnectivity	DataFlow	No descriptio
Job_TestConnectivity	Job	No descriptio
New_DataFlow	DataFlow	No descriptio
New_DataFlow	DataFlow	No descriptio
Job_TestConnectivity	Job	No descriptio
RT_TestConnectivity	DataFlow	No descriptio
New_DataFlow18	DataFlow	No descriptio

Record: 7 of 28

ALVW_JOBINFO(Joes.DI_REPO)

JOB_NAME	JOB_ID
BUG_18950_JOB	39
di_job_al_mach_info	47
New_Job	61
SimpleJob_1	51

Record: 1 of 4

Vous pouvez ouvrir deux volets d'affichage des données pour un affichage simultané. Lorsque les deux volets sont remplis et que vous cliquez sur un autre bouton d'affichage des données, un petit menu s'affiche, il contient les icônes de placement des fenêtres. La zone noire de chaque icône indique le volet que vous désirez remplacer par un nouveau jeu de données. Cliquez sur une option du menu et les données du dernier objet sélectionné remplacent celles du volet correspondant.

La description ou le chemin correspondant au bouton d'affichage des données sélectionné s'affiche en haut du volet.

- Pour les sources et les cibles, la description est le nom complet de l'objet :
 - `<NomObjet>` (`<BanqueDeDonnées.Propriétaire>`) pour les tables ;
 - `<NomFichier>` (`<NomFormatFichier>`) pour les fichiers.
- Pour les boutons d'affichage des données placés sur une ligne, le chemin comprend le nom de l'objet de gauche, une flèche et le nom de l'objet de droite.
Par exemple, si vous sélectionnez un bouton d'affichage des données placé sur la ligne entre la requête nommée Query et la cible nommée ALVW_JOBINFO(joes.DI_REPO), le chemin sera le suivant :

```
Query -> ALVW_JOBINFO(Joes.DI_REPO)
```

Vous pouvez aussi rechercher le volet d'affichage des données associé à un objet ou une ligne en :

- Faisant passer le curseur sur un bouton d'affichage des données d'un objet ou d'une ligne. Designer met en surbrillance le volet d'affichage des données de l'objet.

- Recherchant les boutons d'affichage des données gris sur les objets et les lignes. Designer affiche les boutons d'affichage des données des objets ouverts avec un arrière-plan gris plutôt que blanc.

Informations associées

[Affichage des données passées par les transformations \[page 753\]](#)

32.2.3 Propriétés de l'affichage des données

Vous pouvez accéder aux propriétés de l'affichage des données à l'aide des boutons de la barre d'outils ou du menu contextuel (par clic droit).

L'affichage des données se fait dans les lignes et les colonnes d'une grille de données. Le nombre de lignes affichées est déterminé par une combinaison de plusieurs conditions :

- Taille d'échantillon : nombre de lignes échantillonnées dans la mémoire. La taille d'échantillon par défaut est de 1 000 lignes pour les objets source et cible importés. La taille d'échantillon maximale est de 5 000 lignes. La taille d'échantillon pour les objets source est définie depuis **► Outils ► Options ► Designer ► Général ► Taille d'échantillon de l'affichage des données ►**.
Lorsqu'il utilise le débogueur interactif, le progiciel emploie l'option Fréquence d'échantillonnage des données au lieu de la taille d'échantillon.
- Filtrage
- Tri

Si le jeu de données d'origine est plus petit ou si vous utilisez des filtres, le nombre de lignes retournées peut être inférieur à la valeur par défaut.

Vous pouvez voir quelles conditions ont été appliquées dans la [barre de navigation](#).

Informations associées

[Filtrage \[page 731\]](#)

[Tri \[page 732\]](#)

[Démarrage et arrêt du débogueur interactif \[page 744\]](#)

32.2.3.1 Filtrage

Vous pouvez mettre l'accent sur différents ensembles de lignes d'un échantillon de données local ou nouveau en plaçant des conditions d'extraction sur les colonnes.

32.2.3.1.1 Afficher et ajouter des filtres

1. Dans la barre d'outils de l'affichage des données, cliquez sur le bouton Filtres ou cliquez avec le bouton droit sur la grille et sélectionnez [Filtres](#). 

La fenêtre Filtres s'ouvre.

2. Création de filtres

La fenêtre Filtres comporte trois colonnes :

- a. Colonne : sélectionnez un nom dans la première colonne. Sélectionnez [{supprimer le filtre}](#) pour supprimer le filtre.
- b. Opérateur : sélectionnez un opérateur dans la deuxième colonne.
- c. Valeur : saisissez dans la troisième colonne une valeur qui utilise un des trois formats de type de données suivants.

Type de données	Format
Entier, double, réel	standard
date	aaaa.mm.jj
time	hh24:mm:ss
datetime	aaaa.mm.jj hh24:mm:ss
varchar	'abc'

3. Dans la zone de liste déroulante [Concaténer tous les filtres à l'aide de](#), sélectionnez un opérateur ([ET](#), [OU](#)) pour que le moteur l'utilise dans la concaténation des filtres.

Chaque ligne de cette fenêtre est considérée comme un filtre.

4. Pour voir de quelle manière le filtre affecte l'ensemble actuel de lignes renvoyées, cliquez sur [Appliquer](#).

5. Pour enregistrer les filtres et fermer la fenêtre Filtres, cliquez sur [OK](#).

Vos filtres sont enregistrés pour l'objet en cours et l'échantillon local est mis à jour pour afficher les données filtrées comme indiqué dans la boîte de dialogue Filtres. Pour utiliser les filtres avec un nouvel échantillon, voir *Utilisation de l'actualisation*.

Informations associées

[Utilisation de l'actualisation \[page 733\]](#)

32.2.3.1.2 Ajouter un filtre à la cellule sélectionnée.

1. Sélectionnez une cellule dans la grille de l'échantillon de données.
2. dans la barre d'outils de l'affichage des données, cliquez sur le bouton Ajouter un filtre ou cliquez sur la cellule avec le bouton droit et sélectionnez [Ajouter un filtre](#). 

L'option Ajouter un filtre ajoute la nouvelle condition de filtrage, <colonne> = <valeur de la cellule>, puis ouvre la fenêtre Filtres pour que vous puissiez afficher ou modifier le nouveau filtre.

3. Quand vous avez terminé, cliquez sur [OK](#). 

Pour supprimer les filtres d'un objet, accédez à la barre d'outils d'affichage des données et cliquez sur le bouton Supprimer les filtres ou cliquez sur la grille avec le bouton droit et sélectionnez [Supprimer les filtres](#). Tous les filtres sont supprimés de l'objet en cours.

32.2.3.2 Tri

Vous pouvez cliquer sur un ou plusieurs en-têtes de colonne de la grille de données pour trier les données. Une flèche apparaît sur l'en-tête pour indiquer l'ordre de tri : croissant (flèche vers le haut) ou décroissant (flèche vers le bas).

Pour modifier l'ordre de tri, cliquez à nouveau sur l'en-tête de colonne. La priorité d'un tri se fait de gauche à droite dans la grille.



Pour supprimer le tri d'un objet, dans la barre d'outils, cliquez sur le bouton Supprimer le tri ou cliquez sur la grille avec le bouton droit et sélectionnez [Supprimer le tri](#).

Informations associées

[Utilisation de l'actualisation \[page 733\]](#)

32.2.3.3 Utilisation de l'actualisation



Pour charger un autre échantillon de données de la base de données à l'aide de nouveaux paramètres de filtre et de tri, utilisez la commande [Actualiser](#). Après avoir modifié le filtrage et le tri, cliquez sur le bouton Actualiser de la barre d'outils ou cliquez sur la grille de données avec le bouton droit et sélectionnez [Actualiser](#).



Pour interrompre une opération d'actualisation, cliquez sur le bouton Arrêter. Pendant que le progiciel actualise les données, toutes les commandes d'affichage des données sont désactivées, sauf le bouton [Arrêter](#).

32.2.3.4 Utilisation de l'option Afficher/masquer les colonnes

Vous pouvez restreindre le nombre de colonnes affichées dans l'affichage des données en utilisant l'option [Afficher/masquer les colonnes](#) depuis :

- La barre d'outils.
- Le menu contextuel (clic droit).
- Le menu contextuel fléché, situé à droite du bouton [Afficher/masquer les colonnes](#) de la barre d'outils. Cette option est disponible uniquement si le nombre total de colonnes de la table est au maximum de dix. Sélectionnez une colonne pour l'afficher.

Vous pouvez également "masquer rapidement" une colonne en cliquant avec le bouton droit sur son en-tête et en sélectionnant [Masquer](#) dans le menu.

32.2.3.4.1 Afficher ou masquer les colonnes



1. Cliquez sur le bouton Afficher/masquer les colonnes de la barre d'outils ou cliquez avec le bouton droit sur la grille de données et sélectionnez [Afficher/masquer les colonnes](#).
La fenêtre Paramètres de colonne s'ouvre.
2. Sélectionnez les colonnes à afficher ou cliquez sur l'un des boutons suivants : [Afficher](#), [Afficher tout](#), [Masquer](#) ou [Masquer tout](#).
3. Cliquez sur [OK](#).

32.2.3.5 Ouverture d'une nouvelle fenêtre.



Pour voir davantage de l'échantillon de données que vous visualisez dans un volet d'affichage des données, ouvrez une fenêtre d'affichage des données de taille normale. Depuis un volet d'affichage des données, cliquez sur le bouton Ouvrir une fenêtre de la barre d'outils pour activer une fenêtre distincte d'affichage des données de taille normale. Autrement, vous pouvez cliquer avec le bouton droit et sélectionner dans le menu la commande *Ouvrir dans une nouvelle fenêtre*.

32.2.4 Options de la barre d'outils d'affichage des données

Les options suivantes sont disponibles sur les volets d'affichage des données

Icône	Option	Description
	Ouvrir dans une nouvelle fenêtre	Ouvre le volet d'affichage des données dans une fenêtre plus grande.
	Enregistrer sous	Enregistre les données du volet d'affichage des données.
	Imprimer	Imprime les données du volet d'affichage des données.
	Copier la cellule	Copie les données d'une cellule du volet d'affichage des données.
	Actualiser les données	Extrait un autre échantillon des données existantes dans le volet d'affichage des données à l'aide de nouveaux paramètres de filtre et de tri.
	Ouvrir la fenêtre Filtres	Ouvre la fenêtre Filtres.
	Ajouter un filtre	Ajoute un filtre à la cellule sélectionnée.
	Supprimer le filtre	Supprime tous les filtres d'un volet d'affichage des données.
	Supprimer le tri	Supprime les paramètres de tri de l'objet sélectionné.
	Afficher/masquer la navigation	Affiche ou masque la barre de navigation qui s'affiche sous la table de données.
	Afficher/masquer les colonnes	

Informations associées

[Ouverture d'une nouvelle fenêtre. \[page 733\]](#)

[Utilisation de l'actualisation \[page 733\]](#)

[Filtrage \[page 731\]](#)

[Tri \[page 732\]](#)

[Ajouter un filtre à la cellule sélectionnée. \[page 732\]](#)

[Utilisation de l'option Afficher/masquer les colonnes \[page 733\]](#)

32.2.5 Onglets d'affichage des données

Le panneau d'affichage des données des objets comporte trois onglets :

- Onglet Données
- Onglet Profil
- Onglet Profil de colonne

Utilisez les options des onglets pour obtenir un profil complet d'un objet source ou cible. L'onglet Données est toujours disponible. Les onglets Profil et Relation sont pris en charge par le profileur de données. Sans le profileur de données, les onglets Profil et Profil de colonne sont pris en charge pour certaines sources ou cibles.

Informations associées

[Affichage des résultats du profileur \[page 396\]](#)

32.2.5.1 Onglet Données

L'onglet Données permet d'utiliser les propriétés d'*affichage des données*. Il indique également les schémas imbriqués tels que ceux utilisés dans les fichiers et messages XML. Lorsqu'une colonne fait référence à des schémas imbriqués, elle est ombrée en jaune et une petite icône de table s'affiche dans l'en-tête de colonne.

Informations associées

[Propriétés de l'affichage des données \[page 730\]](#)

32.2.5.1.1 Affichage d'un schéma imbriqué

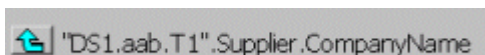
1. Double-cliquez sur une cellule.

La grille de données se met à jour afin d'afficher les données dans la cellule sélectionnée ou la table imbriquée. 

Dans la zone Schéma, la valeur de la cellule sélectionnée est marquée d'une icône spéciale. En outre, les tables et colonnes du chemin sélectionné sont affichées en bleu alors que les références du schéma imbriqué sont affichées en gris.

Dans la zone Données, les données sont affichées pour les colonnes. Les références du schéma imbriqué sont affichées entre chevrons, par exemple `<NomSociété>`.

2. Continuez à utiliser le côté grille de données du panneau pour naviguer. Exemple :
 - Sélectionnez une colonne imbriquée de niveau inférieur et double-cliquez sur une cellule pour mettre à jour la grille de données.
 - Cliquez sur la flèche en haut de la grille de données pour monter dans la hiérarchie.



- Voyez le chemin complet vers la colonne ou la table sélectionnée qui est affiché à droite du bouton Explorer en arrière. Utilisez le chemin et la grille de données pour naviguer dans les schémas imbriqués.

32.2.5.2 Onglet Profil

Si vous utilisez le profileur de données, l'onglet Profil affiche les attributs de profil que vous avez sélectionnés dans l'option [Soumettre la demande de profil de colonne](#).

L'onglet Profil permet de calculer les informations statistiques de tout ensemble de colonnes sélectionné. Cette fonctionnalité facultative n'est pas disponible pour les colonnes avec schémas imbriqués ou pour le type de données LONG.

Informations associées

[Exécution d'une tâche du profileur \[page 390\]](#)

32.2.5.2.1 Utilisation de l'onglet Profil sans le profileur de données

1. Sélectionnez une ou plusieurs colonnes.

Comme les calculs de mise à jour ont un impact sur les performances, sélectionnez uniquement les noms des colonnes dont vous avez besoin pour cette opération de profilage.

Vous pouvez également cliquer avec le bouton droit pour utiliser les options du menu Sélectionner tout ou Désélectionner tout.

2. Cliquez sur [Mettre à jour](#).

3. Les statistiques s'affichent dans la grille Profil.

La grille comporte six colonnes.

Colonne	Description
Colonne	Noms des colonnes de la table actuelle. Sélectionnez les noms dans cette colonne, puis cliquez sur Mettre à jour pour remplir la grille de profil.
Valeurs distinctes	Nombre total de valeurs distinctes dans cette colonne.
Valeurs NULL	Nombre total de valeurs NULL dans cette colonne.
Min.	Valeur minimale de toutes les valeurs de cette colonne.
Max.	Valeur maximale de toutes les valeurs de cette colonne.
Dernière mise à jour	Heure à laquelle cette statistique a été calculée.

Triez les valeurs de cette grille en cliquant sur les en-têtes des colonnes. Notez que les colonnes Min. et Max. ne peuvent pas être triées.

Outre la mise à jour des statistiques, vous pouvez cliquer sur le bouton [Enregistrements](#) de l'onglet Profil pour compter le nombre total d'enregistrements physiques dans l'objet que vous profilez.

Le progiciel enregistre les valeurs calculées auparavant dans le référentiel et les affiche jusqu'à la prochaine mise à jour.

32.2.5.3 Onglet Profil de colonne

L'onglet Profil de colonne permet de calculer des informations statistiques pour une seule colonne. Si vous utilisez le profileur de données, l'onglet Relation s'affiche à la place de Profil de colonne.

i Remarque

Cette fonctionnalité facultative n'est pas disponible pour les colonnes avec schémas imbriqués ou pour le type de données LONG.

Informations associées

[Affichage des données de profil de la relation générées par le profileur de données \[page 400\]](#)

32.2.5.3.1 Calcul des statistiques d'utilisation d'une valeur pour une colonne

1. Saisissez un nombre dans la zone [Premières](#).

Ce nombre est utilisé pour rechercher les valeurs les plus souvent utilisées dans la colonne. La valeur par défaut est 10, c'est-à-dire que le logiciel renvoie les 10 valeurs les plus souvent utilisées.

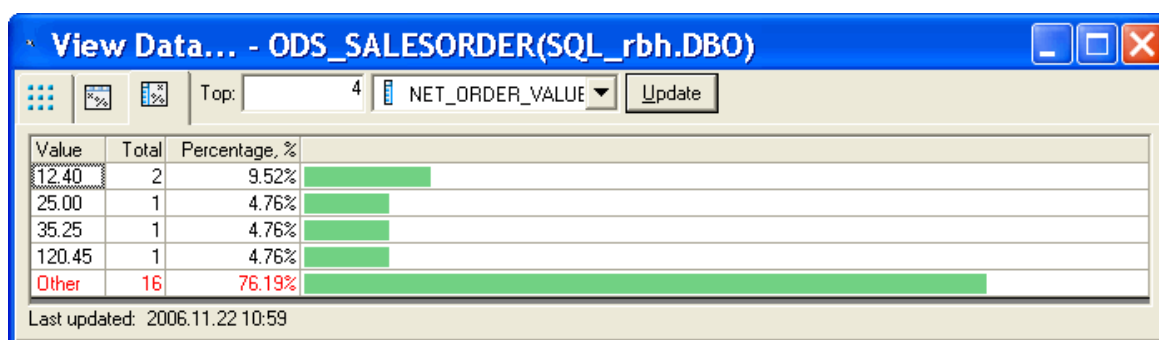
2. Sélectionnez un nom de colonne dans la zone de liste déroulante.
3. Cliquez sur [Mettre à jour](#).

La grille Profil de colonne affiche les statistiques de la colonne indiquée. La grille comporte trois colonnes :

Colonne	Description
Valeur	Une valeur première (plus souvent utilisée) trouvée dans la colonne indiquée ou Autre (valeurs restantes utilisées moins fréquemment).
Total	Nombre total de lignes dans la colonne indiquée qui contient cette valeur.
Pourcentage	Pourcentage de lignes dans la colonne indiquée comportant cette valeur par rapport au nombre total de valeurs de la colonne.

Le logiciel renvoie un nombre de valeurs correspondant au maximum au nombre indiqué dans la zone [Premières](#), plus une valeur supplémentaire appelée [Autre](#).

Donc, si vous saisissez 4 dans la zone Premières, vous obtiendrez jusqu'à cinq valeurs renvoyées (les quatre valeurs les plus utilisées dans la colonne indiquée, plus la catégorie Autre). Les résultats sont enregistrés dans le référentiel et affichés jusqu'à l'exécution d'une nouvelle mise à jour.



Par exemple, les résultats statistiques de la table précédente indiquent que parmi les quatre valeurs les plus souvent utilisées dans la colonne NET_ORDR_VALUE, 9,52 % utilisent la valeur 12.40, 4,76 % utilisent la valeur 25.00, etc. Vous pouvez également voir que les quatre valeurs les plus souvent utilisées (les "quatre premières") le sont dans approximativement 24 % des cas, puisque seulement 76 % sont affichés dans la catégorie Autre. Pour cet exemple, le nombre total de lignes comptées pendant le calcul pour chacune des premières valeurs est de 21.

32.3 Utilisation de l'affichage des données lors de la conception

L'affichage des données lors de la conception permet d'afficher et d'analyser en temps réel l'entrée et la sortie pour un jeu de données lors de la conception d'une transformation. Le flux de données ne doit pas nécessairement être complet ni valide bien qu'il doive utiliser une source accessible et valide contenant les données.

L'affichage des données lors de la conception permet de visualiser les données lorsque vous concevez une transformation pour être sûr d'obtenir les résultats escomptés. L'affichage des données lors de la conception est disponible sous forme de volets d'entrée et de sortie dans l'éditeur de transformation ; vous pouvez ainsi comparer les données avant et après leur traitement via la transformation.

Pour utiliser l'affichage des données lors de la conception pour les transformations Global Address Cleanse, Geocoder et USA Regulatory Address Cleanse, vous devez avoir accès aux données de référence de Data Quality (répertoires).

Informations associées

[Affichage des données lors de la conception \[page 739\]](#)

[Configuration de l'affichage des données lors de la conception \[page 740\]](#)

32.3.1 Affichage des données lors de la conception

1. Pour afficher les données d'entrée et de sortie des transformations dans l'éditeur de transformation, sélectionnez ► [Déboguer](#) ► [Afficher les données au moment de la conception](#) ►.

Les volets de données d'entrée et de sortie s'ouvrent dans l'éditeur de transformation. Chaque volet peut contenir plusieurs vues avec des onglets selon le nombre d'entrées et de sorties de la transformation sélectionnée. Bien que la plupart des transformations comportent une seule sortie, la transformation Validation en comporte trois, par exemple.

2. Pour afficher les données relatives à une entrée ou sortie différente, cliquez sur l'onglet pertinent dans le volet Afficher les données.
3. Pour afficher les données d'entrée et de sortie automatiquement après avoir modifié une transformation, sélectionnez ► [Déboguer](#) ► [Affichage automatique](#) ►.
4. Pour filtrer le nombre de lignes de données affichées dans les volets, sélectionnez ► [Déboguer](#) ► [Filtrer le jeu de données d'entrée](#) ►. La définition du filtre peut contribuer à améliorer les performances lorsque la fonction Design-Time Data Viewer est paramétrée en vue d'une mise à jour automatique au moment de la conception de la transformation.
Par défaut, le filtre affiche les 50 premières lignes. Vous pouvez configurer le nombre de lignes affichées dans la fenêtre [Options](#).
5. Pour fermer les volets Design-Time Data Viewer, cliquez sur x en haut à droite du volet.

Informations associées

[Configuration de l'affichage des données lors de la conception \[page 740\]](#)

32.3.2 Configuration de l'affichage des données lors de la conception

Vous pouvez configurer le nombre de lignes de données affichées dans les volets d'affichage des données lors de la conception ainsi que la durée autorisée pour les mises à jour avant l'expiration du délai.

1. Pour configurer les options d'affichage des données lors de la conception, sélectionnez ► [Déboguer](#) ► [Options](#) .
Une fenêtre s'affiche avec les options disponibles.
2. Modifiez les options voulues.

Option	Description
Nombre de lignes à lire depuis une source filtrée	Indique le nombre de lignes lues depuis la source si le point de menu Filtrer le jeu de données d'entrée est sélectionné. Par défaut, le filtre lit les 50 premières lignes de données depuis la source.
Intervalle de temporisation pour le mode automatique	Indique la durée (en secondes) autorisée pour mettre à jour les données si le point de menu Affichage automatique est sélectionné avant l'expiration du délai et le renvoi d'une erreur.
Intervalle de temporisation pour le mode manuel	Indique la durée (en secondes) autorisée pour mettre à jour les données si le point de menu Affichage automatique est désélectionné avant l'expiration du délai et le renvoi d'une erreur.

3. Pour fermer la fenêtre, cliquez sur [OK](#).

32.3.3 Définition de variables pour les expressions

Vous pouvez définir des variables à utiliser dans les expressions de transformations. La valeur de la variable est utilisée dans les calculs de données lors de la conception. Vous pouvez également définir des valeurs pour les variables globales, uniquement utilisées dans le cadre des calculs lors de la conception.

1. Pour configurer les options d'affichage des données lors de la conception, sélectionnez ► [Déboguer](#) ► [Options](#) .
Une fenêtre s'affiche avec les options disponibles.
2. Dans la section Variables, saisissez le nom, le type de données et la valeur de la variable.
3. Pour importer une variable globale, cliquez sur le bouton [Importer](#). Toutes les variables globales de chaque job du référentiel alimentent la table.
4. Saisissez la valeur de chaque variable globale importée.
Les valeurs de variables définies dans les options d'affichage des données lors de la conception sont seulement utilisées pour les calculs de données lors de la conception.
5. Pour supprimer une variable de la table, sélectionnez-la et cliquez sur le bouton [Supprimer](#).
6. Pour fermer la fenêtre, cliquez sur [OK](#).

32.4 Utilisation du débogueur interactif

Designer comprend un débogueur interactif qui permet d'examiner et modifier les données ligne par ligne (pendant l'exécution d'un job en mode débogage) en plaçant des filtres et des points d'arrêt sur les lignes d'un diagramme de flux de données. Le débogueur interactif fournit des options puissantes pour déboguer un job.

i Remarque

Une mise à niveau du référentiel est nécessaire pour utiliser cette fonctionnalité.

32.4.1 Avant de lancer le débogueur interactif

Comme pour l'exécution d'un job, vous pouvez lancer le débogueur interactif depuis le menu [Débogage](#) lorsqu'un job est actif dans l'espace de travail. Sélectionnez [Démarrer le débogage](#), définissez les propriétés pour l'exécution, puis cliquez sur [OK](#). Le mode de débogage commence. Le mode débogage propose les fenêtres, les menus et les boutons de barre d'outils du débogueur interactif que vous pouvez utiliser pour contrôler le rythme du job et afficher les données en mettant en pause l'exécution du job à l'aide de filtres et de points d'arrêt.

En mode débogage, toutes les autres fonctionnalités de Designer sont placées en lecture seule. Pour quitter le mode débogage et repasser les autres fonctionnalités de Designer en lecture/écriture, cliquez sur le bouton [Arrêter le débogage](#) dans la barre d'outils du débogueur interactif.

Toutes les commandes du débogueur interactif sont listées dans le menu Débogage de [Designer](#). Designer active les commandes appropriées au fur et à mesure de l'avancement d'une session de débogage interactif.

Cependant, avant de lancer une session de débogage, définissez éventuellement les éléments suivants :

- Filtres et points d'arrêt
- Port du débogueur interactif entre Designer et un moteur.

32.4.1.1 Définition des filtres et points d'arrêt

Vous pouvez définir n'importe quelle combinaison de filtres et points d'arrêt dans un flux de données avant de lancer le débogueur interactif. Le débogueur utilise les filtres et fait une pause aux points d'arrêt définis.

Si vous ne configurez pas de filtres ou points d'arrêt prédéfinis :

- Designer optimisera l'exécution du job de débogage. Cela signifie souvent que la première transformation de chaque flux de données d'un job est poussée vers la base de données source. Par conséquent, vous ne pouvez pas afficher les données d'un job entre la source et la première transformation, sauf en plaçant un point d'arrêt prédéfini sur cette ligne.
- Vous pouvez mettre manuellement un job en pause en utilisant une option de débogage nommée [Suspendre le débogage](#) (le job est mis en pause avant qu'il ne rencontre la prochaine transformation).

Informations associées

[Optimiseur de poussée vers le bas \[page 754\]](#)

32.4.1.1.1 Définir un filtre ou un point d'arrêt

1. Dans l'espace de travail, ouvrez le job à déboguer.
2. Ouvrez un de ses flux de données.
3. Cliquez avec le bouton droit sur la ligne que vous voulez examiner et sélectionnez [Définir le filtre/point d'arrêt](#).

Une ligne est une ligne entre deux objets d'un diagramme d'espace de travail.

La fenêtre Point d'arrêt s'ouvre. Sa barre de titre affiche les objets auxquels la ligne est connectée.

4. Définissez et activez un filtre ou un point d'arrêt à l'aide des options de cette fenêtre.

Un filtre de débogage fonctionne comme une simple transformation Query avec une clause WHERE.

Utilisez un filtre pour réduire un jeu de données lors de l'exécution d'un job de débogage. Notez que les expressions complexes ne sont pas prises en charge dans un filtre de débogage.

Placez un filtre de débogage sur une ligne entre une source et une transformation ou entre deux transformations. Si vous placez un filtre et un point d'arrêt sur la même ligne, le progiciel applique d'abord le filtre. Le point d'arrêt peut voir uniquement les lignes filtrées.

Comme pour les filtres, vous pouvez placer un point d'arrêt entre une source et une transformation ou entre deux transformations. Un point d'arrêt est l'emplacement où l'exécution d'un job de débogage est mise en pause et vous rend le contrôle.

Choisissez d'utiliser un point d'arrêt avec ou sans conditions.

- Si vous utilisez un point d'arrêt sans condition, l'exécution du job est suspendue pour la première ligne transmise à un point d'arrêt.
- Si vous utilisez un point d'arrêt avec condition, l'exécution du job est suspendue pour la première ligne transmise au point d'arrêt respectant la condition.

Une condition de point d'arrêt s'applique à la post-image pour les types de ligne UPDATE, NORMAL et INSERT et à la pré-image pour un type de ligne DELETE.

Au lieu de sélectionner un point d'arrêt conditionnel ou non conditionnel, vous pouvez également utiliser l'option Arrêt après n lignes. Dans ce cas, l'exécution est suspendue lorsque le nombre de lignes indiqué passe par le point d'arrêt.

5. Cliquez sur [OK](#).

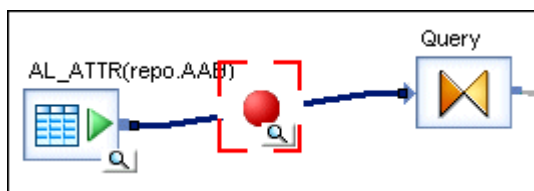
L'icône Point d'arrêt activé s'affiche sur la ligne sélectionnée.

Le progiciel fournit les conditions suivantes de filtre et de point d'arrêt :

Icône	Description
	Point d'arrêt désactivé

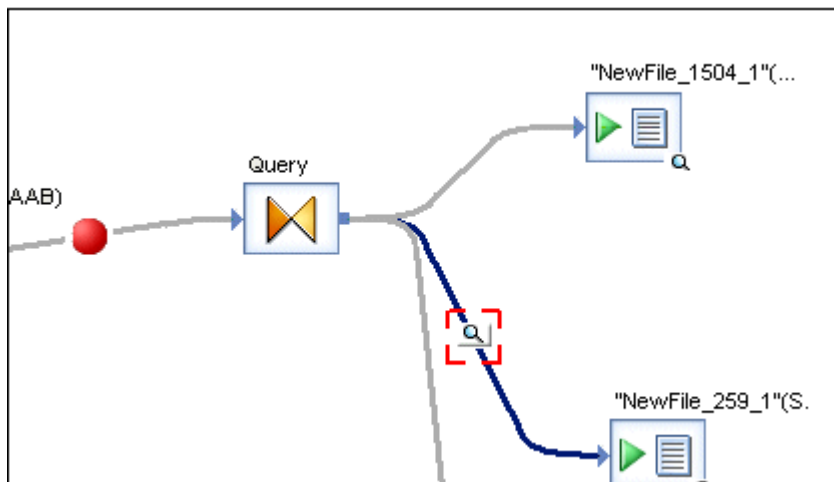
Icône	Description
	Point d'arrêt activé
	Filtre désactivé
	Filtre activé
	Filtre et point d'arrêt désactivés
	Filtre et point d'arrêt activés
	Filtre activé et point d'arrêt désactivé
	Filtre désactivé et point d'arrêt activé

Outre les icônes de filtre et de point d'arrêt qui peuvent apparaître sur une ligne, le débogueur met une ligne en surbrillance lorsqu'il s'y met en pause. Une case rouge de localisation indique également l'emplacement actuel dans le flux de données. Par exemple, lorsque vous lancez le débogueur interactif, le job est suspendu au niveau de votre point d'arrêt. La case de localisation s'affiche au-dessus de l'icône du point d'arrêt, comme l'illustre le diagramme suivant :



Un bouton d'affichage des données s'affiche également au-dessus du point d'arrêt. Vous pouvez utiliser ce bouton pour ouvrir et fermer les volets d'affichage des données.

Au fur et à mesure de l'avancement du débogueur dans les calculs du flux de données, les lignes suivantes sont mises en surbrillance et la case de localisation s'affiche sur votre position actuelle.



Informations associées

[Volets \[page 746\]](#)

32.4.1.2 Modification du port du débogueur interactif

Designer utilise un port vers un moteur pour démarrer et arrêter le débogueur interactif. Le port du débogueur interactif est 5001 par défaut.

32.4.1.2.1 Modification du port du débogueur interactif

1. Sélectionnez **Outils** > **Options** > **Designer** > **Environnement**.
2. Saisissez une valeur dans la zone *Débogueur interactif*.
3. Cliquez sur **OK**.

32.4.2 Démarrage et arrêt du débogueur interactif

Un job doit être actif dans l'espace de travail pour pouvoir démarrer le débogueur interactif. Vous pouvez sélectionner un job dans la bibliothèque d'objets ou dans la zone de projet pour l'activer dans l'espace de travail. Quand un job est actif, Designer active l'option *Démarrer le débogage* dans le menu et la barre d'outils *Débogage*.

32.4.2.1 Démarrage du débogueur interactif

1. Dans la zone de projet, cliquez avec le bouton droit sur un job et sélectionnez *Démarrer le débogage*.

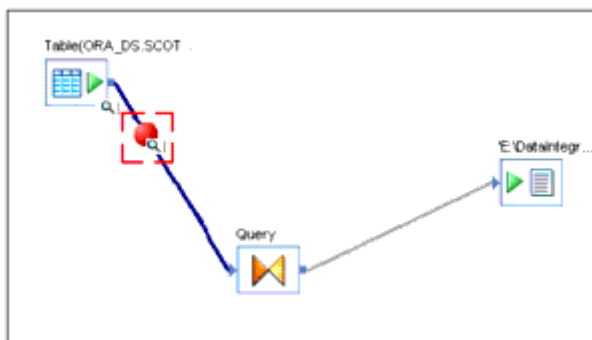
Autrement, vous pouvez cliquer sur un job dans la zone de projet, puis :

- Appuyez sur **Ctrl**+**F8**.
- Dans le menu *Déboguer*, cliquez sur *Démarrer le débogage*.
- Cliquez sur le bouton *Démarrer le débogage* dans la barre d'outils.

Vous trouverez également davantage d'informations sur les options de Suivi et variable globale. La fenêtre Propriétés de débogage s'ouvre. La fenêtre Propriétés de débogage comprend trois paramètres similaires à la fenêtre Propriétés d'exécution (utilisée lorsque vous désirez simplement exécuter un job).

Les options propres à la fenêtre Propriétés de débogage sont :

- *Modèle de débit de données* - nombre de lignes mises en cache pour chaque ligne lorsqu'un job est exécuté à l'aide du débogueur interactif. Par exemple, dans le diagramme de flux de données suivant, si la table source comporte 1 000 lignes et que vous définissez le *Modèle de débit de données* sur 500, Designer affiche jusqu'à 500 des dernières lignes qui passent par une ligne sélectionnée. Le débogueur affiche la dernière ligne traitée lorsqu'il atteint un point d'arrêt.



- *Quitter le débogueur lorsque le job est terminé* - cliquez pour arrêter le débogueur et revenir au mode normal après exécution du job. Option non sélectionnée par défaut.
2. Entrez les propriétés de débogage à utiliser ou employez celles par défaut.
 3. Cliquez sur *OK*.

Le job sélectionné dans la zone de projet commence à s'exécuter en mode débogage. Designer :

- Affiche les fenêtres du débogueur interactif.
- Ajoute *Débogage* du job *<NomJob>* à sa barre de titre.
- Activez le menu *Déboguer* approprié et les options de la barre d'outils. La fenêtre Propriétés de débogage s'ouvre.



- Affiche l'icône de débogage dans la barre de statut.
- Place l'interface utilisateur en lecture seule.

i Remarque

Vous ne pouvez pas effectuer d'opérations qui affectent votre référentiel (comme supprimer des objets d'un flux de données) lors de l'exécution d'un job en mode débogage.

Quand le débogueur rencontre un point d'arrêt, il suspend l'exécution du job. Vous contrôlez alors l'exécution du job. Les fenêtres du débogueur interactif affichent des informations sur l'exécution du job jusqu'à ce point. Elles se mettent également à jour lorsque vous poursuivez manuellement les étapes du job ou permettez au débogueur de continuer l'exécution.

Informations associées

[Guide de référence : objets, description d'objets, job batch, options d'exécution \[page 366\]](#)

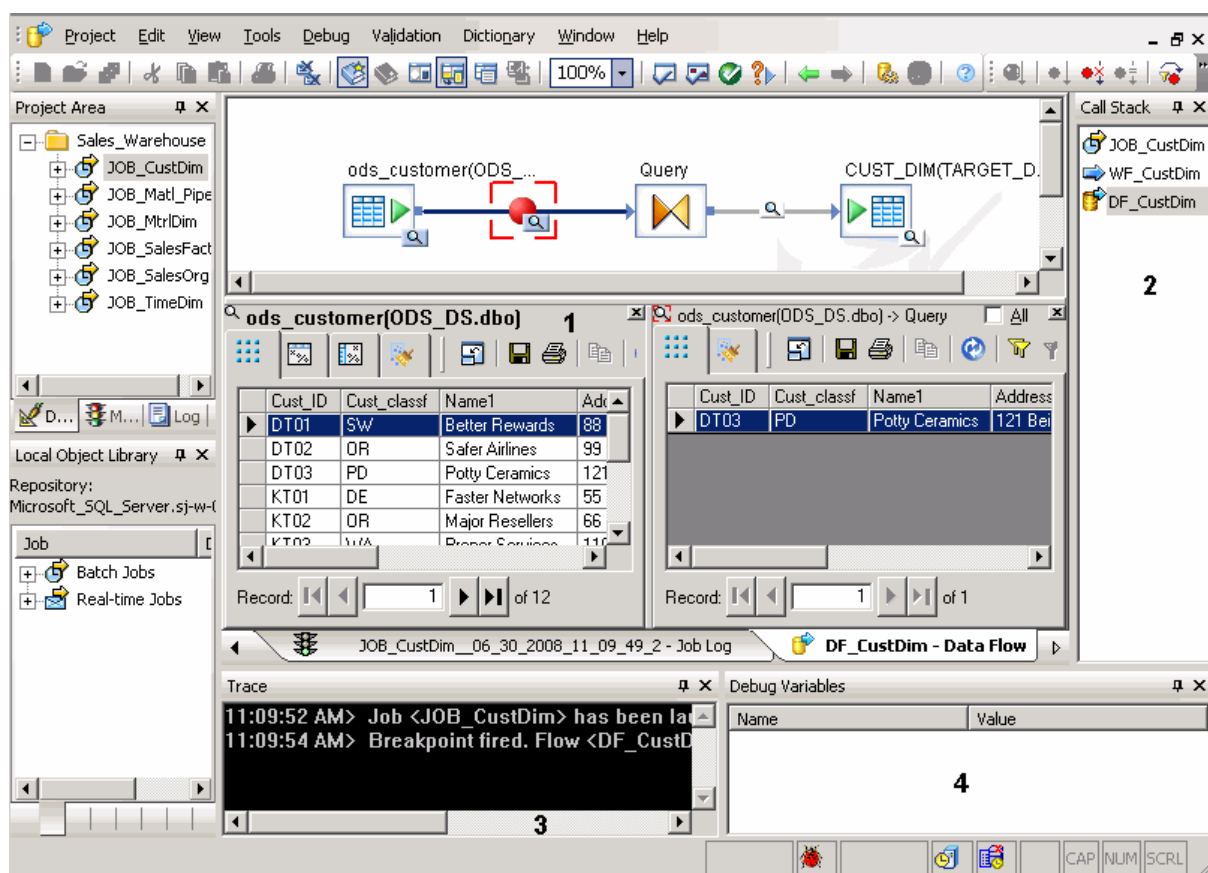
32.4.2.2 Arrêter un job en mode débogage et quitter le débogueur interactif



Cliquez sur le bouton *Arrêter le débogage* dans la barre d'outils, appuyez sur *Maj+F8*, ou bien dans le menu *Déboguer*, cliquez sur *Arrêter le débogage*.

32.4.3 Volets

Lorsque vous démarrez un job dans le débogueur interactif, Designer affiche trois volets supplémentaires ainsi que les volets d'affichage des données sous l'espace de travail. Le diagramme suivant illustre les emplacements par défaut de ces volets.



1. Volets d'affichage des données
2. Volet Pile d'appels
3. Volet Suivi
4. Volet Variables de débogage

Chaque volet est ancré dans la fenêtre de Designer. Pour déplacer un volet du débogueur, double-cliquez sur sa barre de commande pour le libérer, puis cliquez et faites glisser sa barre de titre pour l'ancrer à nouveau.

Designer enregistre la mise en forme créée lorsque vous arrêtez le débogueur interactif. La mise en forme est conservée pour votre prochaine session de Designer.

Vous pouvez redimensionner ou masquer un volet du débogueur à l'aide de ses boutons de commande. Pour afficher ou masquer manuellement un volet du débogueur, utilisez le menu ou la barre d'outils [Débogage](#).

Informations associées

[Options et barre d'outils du menu Débogage \[page 751\]](#)

32.4.3.1 Fenêtre Pile d'appels

La fenêtre Pile d'appels liste les objets rencontrés jusqu'à présent sur le chemin (avant la fin du job, la rencontre d'un point d'arrêt ou la mise en pause).

Par exemple, pour le job JOB_CustDim qui inclut un workflow et un flux de données, il est possible que la fenêtre Pile d'appels s'affiche :

JOB_CustDim

WF_CustDim

DF_CustDim

Vous pouvez double-cliquer sur un objet de la fenêtre Pile d'appels pour l'ouvrir dans l'espace de travail. De même, si vous cliquez sur un objet d'un diagramme, la fenêtre Pile d'appels le met en surbrillance.

32.4.3.2 Fenêtre Suivi

La fenêtre Suivi affiche le statut de sortie du débogueur et les erreurs. Exemple :

```
11:22:06 AM> Job <GO> has been launched in debug mode.  
11:22:07 AM> Breakpoint fired. Flow <aaa>: from <Query> to  
<"NewFile_773" (Simple)>
```

Lorsque le job se termine, cette fenêtre affiche :

Job <NomJob> terminé. Arrêter le débogueur.

Lorsque le job se termine, le débogueur vous donne une dernière possibilité d'examiner les données. Pour quitter le débogueur, sélectionnez le bouton *Arrêter le débogage* dans la barre d'outils, appuyez sur *Shift+F8* ou sélectionnez *Déboguer Arrêter le débogage*.

32.4.3.3 Fenêtre Variables de débogage

La fenêtre Variables de débogage affiche les variables globales utilisées par le job à chaque point d'arrêt.

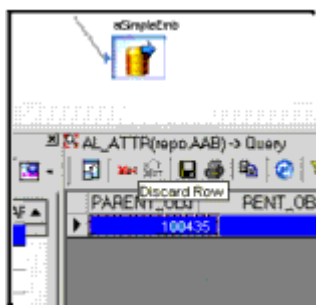
32.4.3.4 Volet d'affichage des données

Le volet d'affichage des données des lignes utilise les mêmes options de barre d'outils et de navigation que celles décrites pour la fonctionnalité Afficher les données.

Les options suivantes du volet d'affichage des données sont propres au débogueur interactif :

- Permet d'afficher les données qui passent par les lignes.
- Affiche (au-dessus de la barre d'outils d'affichage des données) les noms des objets auxquels une ligne se connecte en utilisant le format : <NomTable> (<NomBanqueDonnées>.<NomPropriétaireTable>) -> <NomRequête>.

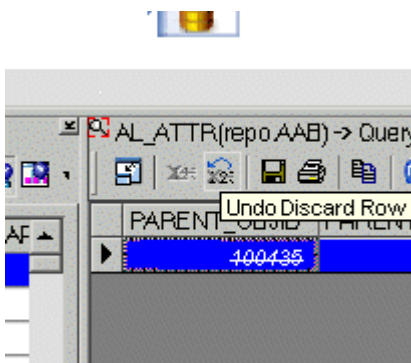
- Affiche par défaut les données une ligne à la fois.
- Fournit la case à cocher *Tout* qui permet de voir plusieurs lignes de données traitées.
- Permet de modifier les données d'une cellule.
Par exemple, vous pouvez éventuellement corriger une erreur temporairement pour poursuivre une exécution du débogueur. Vous pouvez corriger la conception du job ultérieurement pour éliminer l'erreur définitivement.
Pour modifier les données d'une cellule :
 - Décochez la case *Tout* pour qu'une seule ligne soit affichée.
 - Double-cliquez sur une cellule ou cliquez dessus avec le bouton droit et sélectionnez *Modifier la cellule*.
- Utilise une propriété nommée *Fréquence d'échantillonnage des données*.
- Permet de marquer une ligne qui ne doit pas être traitée par la prochaine transformation.
Pour ignorer une ligne dans l'étape suivante du traitement d'un flux de données, sélectionnez-la et cliquez sur *Ignorer la ligne*.



Les données d'une ligne ignorée apparaissent dans le volet d'affichage des données au format "barré" (par exemple, 100345).

Si vous ignorez une ligne accidentellement, vous pouvez annuler immédiatement cette action.

Sélectionnez la ligne ignorée et cliquez sur *Annuler Ignorer la ligne*.



Autrement, cliquez avec le bouton droit de la souris sur la ligne et sélectionnez Ignorer la ligne ou *Annuler Ignorer la ligne* dans le menu contextuel.

Par exemple, si une source d'un flux de données comporte quatre lignes et que vous définissez la fréquence d'échantillonnage des données sur 2 en démarrant le débogueur, il affiche la première ligne traitée à un point d'arrêt prédéfini.

The screenshot shows the Oracle Designer interface. At the top, a diagram illustrates a data flow: a table named 'MATS_DEPT(david.A...)' is connected via a red circle (representing a join or filter) to a 'Query' object, which is then connected to a table named 'Y(david.A)'. Below the diagram, two data tables are displayed side-by-side.

Table 1: MATS_DEPT(david.A)

DEPTNO	DNAME	LOC	ST
10	Accounting	New York	NY
20	Research	Dallas	TX
30	Sales	Chicago	IL
40	Operations	Boston	MA

Record: 1 of 4

Table 2: MATS_DEPT(david.A) -> Query

DEPTNO	DNAME	LOC	ST
10	Accounting	New York	NY

Record: 1 of 1



Si vous utilisez l'option [Ligne suivante](#), la ligne suivante au même point d'arrêt est affichée. Pour voir les deux lignes, cochez la case [Tout](#) dans le coin supérieur droit de ce volet. La ligne affichée en bas de la table est la dernière ligne traitée. A ce niveau, vous avez affiché deux lignes étant passées par une ligne. Si vous cliquez à nouveau sur [Ligne suivante](#), seules les deux dernières lignes traitées s'affichent car vous avez défini la taille d'échantillon sur 2.

Informations associées

[Utilisation de l'affichage des données \[page 727\]](#)

32.4.3.5 Fenêtre Filtres et points d'arrêt



Vous pouvez gérer les filtres et les points d'arrêt du débogueur de manière interactive grâce à la fenêtre Filtres/Points d'arrêt. Cette fenêtre s'ouvre depuis le menu ou la barre d'outils Débogage.

Les lignes comportant des filtres ou des points d'arrêt sont listées tout à gauche de la fenêtre Filtres/Points d'arrêt. Pour les gérer, sélectionnez la ou les lignes à modifier, sélectionnez une commande dans la liste et cliquez sur Exécuter. Vous pouvez également sélectionner une seule ligne sur la gauche et afficher/modifier ses filtres et points d'arrêt du côté droit de cette fenêtre. Lorsque vous avez fini d'utiliser la fenêtre Filtres/Points d'arrêt, cliquez sur OK.

32.4.4 Options et barre d'outils du menu Débogage

Après avoir démarré le débogueur interactif, vous pouvez accéder aux options appropriées depuis le menu et la barre d'outils *Déboguer* de Designer.

Image	Option	Description	Commandes par touches
	Exécuter	Ouvre la fenêtre Propriétés d'exécution, dans laquelle vous pouvez sélectionner les propriétés d'un job puis l'exécuter en dehors du mode débogage. Disponible quand un job est actif dans l'espace de travail.	F8
	Lancer le débogage	Ouvre la fenêtre Propriétés de débogage, dans laquelle vous pouvez sélectionner les propriétés d'un job puis l'exécuter en mode débogage (démarrer le débogueur). Les autres opérations de Designer sont en lecture seule jusqu'à l'arrêt du débogueur. Disponible quand un job est actif dans l'espace de travail.	Ctrl+F8
	Arrêter le débogage	Arrête une exécution en mode débogage et quitte le débogueur. Toutes les opérations de Designer sont rétablies en lecture/écriture.	Shift+F8
	Suspendre le débogage	Permet de mettre manuellement le débogueur en pause. Vous pouvez utiliser cette action à la place d'un point d'arrêt.	Aucune

Image	Option	Description	Commandes par touches
	Étape suivante	Permet de passer manuellement à la ligne suivante d'un flux de données en franchissant une transformation dans l'espace de travail. Utilisez cette option pour voir la première ligne d'un jeu de données après sa transformation. L'espace de travail affiche un carré rouge sur la ligne pour indiquer le chemin utilisé. Si la transformation de l'étape suivante comporte plusieurs sorties, Designer fournit un menu contextuel où vous pouvez sélectionner la branche de calcul à emprunter.	F10
	Ligne suivante	Permet de rester au point d'arrêt actuel et d'afficher la ligne de données suivante du jeu de données.	F11
	Continuer	Permet de rendre le contrôle du job à Designer. - vous utilisiez l'option Suspendre le débogage ; - un autre point d'arrêt soit rencontré ; - le job soit terminé.	Ctrl+F10. Le débogueur continue jusqu'à ce que :
	Afficher les filtres/points d'arrêt	Affiche tous les filtres et points d'arrêt existant dans un job. Si l'option n'est pas sélectionnée, tous les filtres et points d'arrêt sont masqués dans la vue. Cette option est toujours disponible dans Designer.	Aucune

Image	Option	Description	Commandes par touches
	Définir les filtres/points d'arrêt...	Ouvre une boîte de dialogue où vous pouvez définir, supprimer, modifier, activer ou désactiver les filtres et les points d'arrêt. Vous pouvez aussi définir des conditions pour les points d'arrêt. Dans l'espace de travail, vous pouvez cliquer avec le bouton droit sur une ligne et sélectionner la même option dans un menu contextuel. Disponible quand un flux de données est actif dans l'espace de travail.	F9
	Filtres/Points d'arrêt...	Ouvre une boîte de dialogue où vous pouvez gérer plusieurs filtres et points d'arrêt dans un flux de données. Propose aussi la même fonctionnalité que la fenêtre Définir les filtres/points d'arrêt. Cette option est toujours disponible dans Designer.	Alt+F9
	Pile d'appels	Affiche ou masque la fenêtre Pile d'appels.	Aucune
	Variables	Affiche ou masque la fenêtre Déboguer les variables.	
	Suivi	Affiche ou masque la fenêtre Suivi.	Aucune

32.4.5 Affichage des données passées par les transformations

Pour afficher les données passées par les transformations, exécutez le job en mode débogage.

32.4.5.1 Affichage des données passées par les transformations

1. Dans la zone de projet, cliquez avec le bouton droit sur un job et cliquez sur [Démarrer le débogage](#).
La fenêtre Propriétés de débogage s'ouvre.
2. Décochez la case [Quitter l'outil de débogage lorsque le job est terminé](#).

3. Vous pouvez saisir une valeur dans la zone de texte *Fréquence d'échantillonnage des données* ou laisser la valeur 500 par défaut.
4. Cliquez sur *OK*.

32.4.5.2 Affichage des exemples de données en mode débogage

1. En restant en mode débogage à la fin du job, dans la zone de projet, cliquez sur le nom du flux de données à afficher.
2. Cliquez sur le bouton Afficher les données figurant sur une ligne du flux de données.
3. Parcourez les données pour les examiner. Quand c'est fait, cliquez sur le bouton *Arrêter le débogage* de la barre d'outils.

32.4.6 Optimiseur de poussée vers le bas

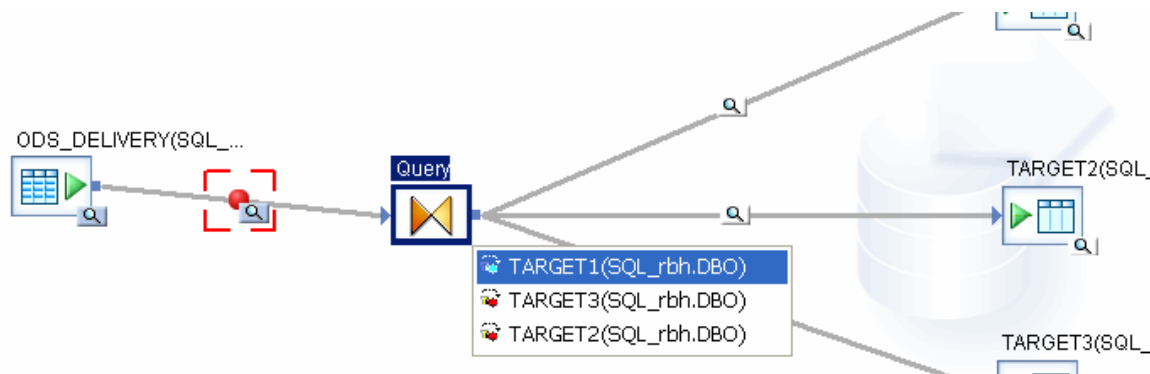
Quand le logiciel exécute un job, il pousse normalement vers le bas autant d'opérations que possible, vers la base de données source, afin de maximiser les performances. Comme le débogueur interactif requiert l'exécution d'un job, les règles suivantes de poussée vers le bas s'appliquent :

- Transformations Query
La première transformation suivant un objet source dans un flux de données est optimisée dans le débogueur interactif et poussée vers le bas, vers la base de données source, si les deux objets répondent aux critères de poussée vers le bas et si vous n'avez pas placé de point d'arrêt sur la ligne avant la première transformation.
- Points d'arrêt
Le logiciel ne pousse vers le bas aucune opération si vous placez un point d'arrêt prédéfini. Les points d'arrêt prédéfinis sont des points d'arrêt définis avant le démarrage du débogueur interactif. Après le démarrage du débogueur interactif, si la première transformation est poussée vers le bas, la ligne est désactivée durant la session de débogage. Vous ne pouvez pas placer de point d'arrêt sur cette ligne ni utiliser le volet d'affichage des données.
- Filtres
Si l'entrée d'un filtre prédéfini est une source de base de données, elle est poussée vers le bas. Les filtres prédéfinis sont des filtres du débogueur interactif définis avant son démarrage.

32.4.7 Restrictions

- Le débogueur interactif peut être utilisé pour examiner les flux de données. Les options de débogage ne sont pas disponibles au niveau du workflow.
- Une mise à niveau du référentiel est nécessaire pour utiliser cette fonctionnalité.
- Le débogueur ne peut pas être utilisé avec les flux de données ABAP.

- Tous les objets d'un flux de données doivent avoir un nom unique. Par exemple, si une transformation comporte plusieurs sorties, vous pouvez choisir le chemin à utiliser. Si certains de ces objets ont le même nom, le résultat de votre sélection est imprévisible.



32.5 Comparaison d'objets

Le progiciel permet de comparer deux objets quelconques et leurs propriétés à l'aide de l'utilitaire Affichage des différences. Vous pouvez comparer :

- deux objets différents
- différentes versions du même objet
- un objet de la bibliothèque d'objets locale à son homologue de la bibliothèque d'objets centrale

Vous pouvez comparer seulement les objets de niveau supérieur ou inclure les dépendants de l'objet dans la comparaison.

Les objets doivent être de même type. Par exemple, vous pouvez comparer un job à un autre job ou une fonction personnalisée à une autre, mais vous ne pouvez pas comparer un job à un flux de données.

32.5.1 Comparer deux objets différents

1. Dans la bibliothèque d'objets locale ou centrale, cliquez avec le bouton droit sur le nom d'un objet.
2. Dans le menu contextuel, mettez en surbrillance *Comparer* et dans le sous-menu, cliquez sur l'une des options suivantes (leur disponibilité dépend de l'objet sélectionné) :

Option	Description
<i>Objet à référentiel central</i>	Compare l'objet sélectionné à son homologue dans la bibliothèque d'objets centrale.
<i>Objet et dépendants à référentiel central</i>	Compare l'objet sélectionné et ses dépendants à son homologue dans la bibliothèque d'objets centrale.
<i>Objet à...</i>	Compare l'objet sélectionné à un autre objet de type similaire.

Option	Description
Objet et dépendants à...	Compare l'objet sélectionné et ses dépendants à un autre objet de type similaire.

Le curseur se change en icône de cible.

3. Lorsque vous déplacez le curseur sur un objet éligible pour la comparaison, le curseur cible change de couleur. Cliquez sur l'objet voulu.

La fenêtre [Affichage des différences](#) s'ouvre dans l'espace de travail.

La fenêtre identifie les éléments modifiés par une combinaison d'icônes, de couleurs et d'ombrage d'arrière-plan. Certaines de ces propriétés sont configurables.

Selon le type d'objet, les volets montrent des éléments tels que les propriétés de l'objet et celles des connexions (liens) entre ses objets enfants.

32.5.2 Comparaison de deux versions du même objet

Si vous travaillez dans un environnement multi-utilisateurs et utilisez une bibliothèque d'objets centrale, vous pouvez comparer deux objets ayant des versions ou des étiquettes différentes.

1. Dans la bibliothèque d'objets centrale, cliquez avec le bouton droit sur le nom d'un objet et dans le menu contextuel, cliquez sur [Afficher l'historique](#).
2. Dans la fenêtre [Historique](#), appuyez sur la touche Ctrl en cliquant sur les deux versions ou étiquettes à comparer.
3. Cliquez sur [Afficher les différences](#) ou [Afficher les différences avec dépendants](#).

La fenêtre [Affichage des différences](#) s'ouvre dans l'espace de travail.

4. Fermez la fenêtre [Historique](#).

Informations associées

[Affichage de l'historique de l'objet \[page 924\]](#)

32.5.3 Présentation de la fenêtre Affichage des différences

Le premier objet sélectionné s'affiche dans le volet gauche de la fenêtre et le deuxième dans le volet droit. L'emplacement de chaque objet figure après son nom.

La fenêtre Affichage des différences comprend les fonctionnalités suivantes :

- barre d'outils
- barre de navigation
- barre d'état

- menu contextuel

Lorsqu'une fenêtre Affichage des différences est active, la fenêtre principale de Designer comporte également un menu nommé Affichage des différences. La section suivante décrit ces fonctionnalités.

Plusieurs fenêtres Affichage des différences peuvent être ouvertes en même temps dans l'espace de travail. Pour actualiser une fenêtre Affichage des différences, appuyez sur la touche F5.

Le développement ou la réduction de tout ensemble de propriétés développe ou réduit également l'ensemble de propriétés correspondant de l'objet comparé.

32.5.3.1 Barre d'outils

La barre d'outils comprend les boutons suivants.

Boutons de navigation

- Première différence (Alt+Orig)
- Différence précédente (Alt+flèche gauche)
- Différence actuelle
- Différence suivante (Alt+flèche droite)
- Dernière différence (Alt+Fin)

Boutons de filtres

- Activer les filtres - cliquez pour ouvrir la boîte de dialogue Filtres.
 - [Masquer les éléments non exécutables](#) - sélectionnez cette option pour supprimer de la vue les éléments qui n'affectent pas l'exécution du job.
 - [Masquer les éléments identiques](#) - sélectionnez cette option pour supprimer de la vue les éléments qui ne comportent pas de différences.
- Désactiver les filtres - supprime tous les filtres appliqués à la comparaison

Afficher les niveaux

Afficher le niveau 1 affiche uniquement les objets sélectionnés pour la comparaison ; Afficher le niveau 2 développe le niveau suivant, etc. Afficher tous les niveaux développe tous les niveaux des deux arbres.

Rechercher (Ctrl+F)

Cliquez sur [Rechercher](#) pour ouvrir une boîte de dialogue de recherche de texte.

Ouvrir dans une nouvelle fenêtre

Cliquez pour ouvrir l’Affichage des différences actif dans une fenêtre distincte. Vous devez fermer cette fenêtre avant de continuer.

32.5.3.2 Barre de navigation

La barre de navigation verticale contient des barres de couleur qui représentent chacune des différences de la comparaison. Les couleurs correspondent à celles de la barre d’état pour chaque différence. Une flèche dans la barre de navigation indique la différence actuellement en surbrillance dans les volets. Vous pouvez cliquer sur la barre de navigation pour sélectionner une différence (la pointe du curseur comportera une étoile sur elle). Les parenthèses violettes dans la barre indiquent la portion de la comparaison actuellement affichée dans les volets.

Informations associées

[Navigation dans les différences \[page 760\]](#)

32.5.3.3 Barre d'état

La barre d’état en bas de la fenêtre comporte une clé qui illustre le schéma de couleur et les icônes qui identifient les différences entre les deux objets.

Icône	Différence	Description
	Supprimé	L’élément n’apparaît pas dans l’objet du volet droit.
	Modifié	Les différences entre les éléments sont mises en évidence par un texte bleu (par défaut)
	Inséré	L’élément a été ajouté à l’objet du volet droit.

Icône	Différence	Description
	Consolidé	Cette icône apparaît près d'un élément si les éléments qu'il renferme comportent des différences. Développez l'élément en cliquant sur son signe plus pour afficher les différences.

Vous pouvez modifier la couleur de ces icônes en cliquant avec le bouton droit dans la fenêtre Affichage des différences et en cliquant sur Configuration.

La barre d'état comprend aussi une référence pour la différence actuellement sélectionnées dans la comparaison (par exemple, la différence actuellement en surbrillance est la 9ème sur un total de 24 différences dans la comparaison).

La barre d'état indique également qu'au moins un filtre est appliqué à cette comparaison.

Informations associées

[Menu contextuel \[page 759\]](#)

32.5.3.4 Menu contextuel

Un clic droit sur le corps de la fenêtre Affichage des différences permet d'afficher un menu contextuel qui contient toutes les commandes de la barre d'outils, plus :

- Afficher — Bascule pour afficher ou masquer la barre d'état, la barre de navigation ou la barre d'outils secondaire (barre d'outils supplémentaire qui s'affiche en haut de la fenêtre, elle peut être utile si l'Affichage des différences est ouvert dans une fenêtre distincte).
- Mise en forme — Utilisez cette commande pour repositionner la barre de navigation.
- Configuration — Cliquez pour modifier les options d'affichage des éléments comportant des différences.

Informations associées

[Modification du schéma de couleurs \[page 759\]](#)

32.5.3.4.1 Modification du schéma de couleurs

La barre d'état située en bas de la fenêtre Affichage des différences montre le schéma de couleurs actuellement utilisé pour identifier les éléments supprimés, modifiés, insérés ou consolidés dans les volets de comparaison. Vous pouvez personnaliser ce schéma de couleurs de la manière suivante.

1. Cliquez avec le bouton droit sur le corps de la fenêtre Affichage des différences pour afficher la barre d'outils contextuelle.

2. Cliquez sur Configuration pour ouvrir la fenêtre Configuration.
3. Cliquez sur un marqueur (Inséré, Supprimé, Modifié ou Consolidé) pour le modifier.
4. Cliquez sur l'échantillon de couleurs pour ouvrir la palette de couleurs
5. Cliquez sur une couleur de base ou créez une couleur personnalisée.
6. Cliquez sur **OK**.
7. Cliquez sur un autre marqueur pour le modifier ou sur **OK** pour fermer la fenêtre Configuration.

32.5.3.4.2 Modification de l'ombrage d'arrière-plan

Les éléments comportant des différences s'affichent avec une couleur d'arrière-plan grise par défaut. Vous pouvez personnaliser cet arrière-plan.

1. Cliquez avec le bouton droit sur le corps de la fenêtre Affichage des différences pour afficher la barre d'outils contextuelle.
2. Cliquez sur Configuration pour ouvrir la fenêtre Configuration.
3. Cliquez sur un marqueur pour le modifier ou cochez la case *Appliquer à tous les marqueurs*.
4. Cliquez sur l'échantillon d'arrière-plans pour ouvrir la palette de couleurs
5. Cliquez sur une couleur de base ou créez une couleur personnalisée.
6. Cliquez sur **OK**.
7. Pour appliquer des couleurs d'arrière-plan différentes aux divers marqueurs, cliquez sur le marqueur pour le modifier et répétez les étapes 4 à 6.
8. Cliquez sur **OK** pour fermer la fenêtre Configuration.

32.5.3.5 Menu Affichage des différences

Lorsqu'une fenêtre Affichage des différences est active dans l'espace de travail, la fenêtre principale de Designer comporte un menu nommé Affichage des différences. Le menu contient les mêmes commandes que la barre d'outils.

32.5.4 Navigation dans les différences

La fenêtre Affichage des différences propose plusieurs options de navigation dans les différences.

Vous pouvez naviguer dans les différences entre les objets à l'aide des boutons de navigation de la barre d'outils. Par exemple, un clic sur le bouton Différence suivante met en surbrillance l'élément suivant qui diffère de l'objet comparé d'une manière quelconque. L'élément est marqué avec l'icône appropriée et seul le texte différent est mis en surbrillance avec la couleur affectée à ce type de différence.

Vous pouvez également utiliser la barre de navigation. Sélectionnez un élément dans un des deux volets comportant une différence. Notez qu'une flèche apparaît près de la barre colorée qui correspond à cet élément. Vous pouvez cliquer sur ces barres pour passer à différents endroits de la comparaison, par exemple pour afficher uniquement les éléments insérés (de couleur verte par défaut). Les parenthèses violettes dans la barre

indiquent la portion de la comparaison actuellement affichée dans les volets. Utilisez la barre de défilement d'un des volets pour ajuster la vue entre parenthèses

Pour les éléments de texte comme les scripts, cliquez sur la loupe pour afficher le texte dans un ensemble de nouveaux volets qui s'affichent sous les volets principaux de l'objet. Utilisez les barres de défilement de ces volets pour les parcourir. Cliquez sur la loupe (ou tout autre élément) pour fermer les volets de texte.

32.6 Calcul des mappages de colonnes

SAP Data Services peut calculer des informations sur les tables et colonnes cible ainsi que les sources utilisées pour les remplir, par exemple pour l'impact et le lignage ou les rapports de documentation automatique.

Le calcul des mappages de colonnes remplit la vue interne ALVW_MAPPING et la table AL_COLMAP_NAMES. La vue ALVW_MAPPING fournit des données actuelles aux applications de reporting sur les métadonnées comme Analyse d'impact et de lignage. Si vous devez générer un rapport sur un flux de données qui traite des données imbriquées (NRDM), effectuez une requête sur la table AL_COLMAP_NAMES à l'aide d'un rapport personnalisé.

Dès qu'un calcul de mappage de colonnes est en cours, Designer affiche une icône de statut en bas à droite de la fenêtre. Vous pouvez double-cliquer sur cette icône pour annuler le processus.

Pour calculer les mappages de colonnes, vous pouvez :

- Activer l'option dans Designer pour calculer automatiquement les mappages de colonnes.
- Exécuter manuellement le processus de mappage de colonnes depuis Designer ou l'application Analyse d'impact et de lignage dans la Console de gestion.

32.6.1 Calculer automatiquement les mappages de colonnes

Pour activer l'option permettant de calculer automatiquement les informations de mappage des colonnes, dans Designer, sélectionnez ► **Outils** ► **Options** ► **Designer** ► **Général** ► **Calculer automatiquement les mappages de colonnes** ►. Cette option est sélectionnée par défaut.

Notez que si l'option de Designer **Calculer automatiquement les mappages de colonnes** est désélectionnée, toutes les modifications ultérieures apportées au flux de données impliquent de recalculer manuellement les mappages de colonnes pour garantir que la vue ALVW_MAPPING et la table AL_COLMAP_NAMES comportent les informations les plus récentes.

32.6.2 Calculer automatiquement les mappages de colonnes

Si l'option de Designer **Calculer automatiquement les mappages de colonnes** est désélectionnée et que vous désirez générer des rapports, vous pouvez calculer manuellement les mappages. Vous pouvez calculer manuellement les mappages de colonnes à tout moment dans Designer ou dans la Console de gestion.

Dans Designer, cliquez avec le bouton droit de la souris sur la bibliothèque d'objets et sélectionnez

► [Répertoire](#) ► [Calculer les mappages de colonnes](#) ►.

Dans la Console de gestion :

1. Sélectionnez [Analyse d'impact et de lignage](#).
2. Ouvrez le panneau de configuration [Paramètres](#).
3. Cliquez sur l'onglet [Actualiser les données d'utilisation](#).
4. Sélectionnez le Job Server associé au répertoire que vous désirez utiliser.
5. Cliquez sur [Calculer les mappages de colonnes](#).

Dans l'onglet [Présentation](#) de l'Analyse d'impact et de lignage, vous pouvez développer l'option [Calcul du mappage de colonnes du flux de données](#) pour afficher une liste des flux de données et le statut de calcul de chacun d'eux. Si le calcul de mappage est terminé, l'indicateur [Statut](#) est coché.

32.7 Contournement de workflows et flux de données spécifiques

Vous pouvez contourner l'exécution d'un workflow ou d'un flux de données au moment de la conception.

L'attribut de contournement peut accélérer le processus de test lors de la conception des jobs en vous permettant d'exécuter certains workflows ou flux de données plutôt que de les exécuter tous. Notez que tous les objets figurant dans un workflow ou flux de données contourné (des scripts, par exemple) sont également contournés.

i Remarque

Vous devez créer des paramètres de substitution à utiliser avec l'option [Contournement](#). Par exemple, dans la fenêtre [Editeur de paramètres de substitution](#), vous pouvez créer \$\$BYPASSEnable avec la valeur Oui et \$\$BYPASSDisable avec la valeur Non (ou toute autre valeur différente de Oui).

Les limitations suivantes s'appliquent :

- Les workflows et flux de données contournés sont ignorés en mode de débogage. Le contournement fonctionne uniquement en mode de conception.
- Les flux de données et flux de données intégrés ABAP ne sont pas pris en charge.
- Les attributs de contournement ne sont pas pris en charge lors de l'exportation et sont alors supprimés.

Il existe plusieurs moyens de vérifier si un objet a été contourné :

- Après la validation, un message s'affiche dans l'onglet [Avertissement](#) de la fenêtre de [sortie](#) (par exemple, Le workflow <nom du workflow> est contourné). Vous pouvez cliquer sur le message pour accéder directement à l'objet contourné.
- Un message s'affiche dans le journal de suivi.
- Une icône (cercle rouge avec une ligne en travers) apparaît dans l'angle inférieur gauche de l'objet de workflow ou de flux de données dans l'espace de travail Designer.

i Remarque

Cette icône reste affichée tant que l'attribut de contournement (activé ou désactivé) est appliqué à l'objet.

Informations associées

[Modification des propriétés d'un flux de données \[page 205\]](#)

[Ordre d'exécution dans les workflows \[page 238\]](#)

[Création d'un paramètre de substitution \[page 333\]](#)

32.7.1 Contournement d'un flux de données simple ou d'un workflow unique

Vous pouvez activer l'attribut de contournement pour un flux de données simple ou un workflow unique.

1. Cliquez avec le bouton droit de la souris sur un flux de données ou un workflow dans l'espace de travail Designer ou la zone de projet et sélectionnez [Propriétés](#).
2. Dans l'onglet [Général](#), sélectionnez une variable de substitution dans la liste déroulante [Contournement](#).
3. Cliquez sur [Appliquer](#).

32.7.2 Contournement de flux de données ou workflows multiples

Vous pouvez activer l'attribut de contournement pour des flux de données ou des workflows multiples.

1. Sélectionnez des flux de données et des workflows dans l'espace de travail Designer.
2. Cliquez avec le bouton droit de la souris et sélectionnez [Contournement](#).
La fenêtre [Définir le contournement](#) s'affiche.
3. Sélectionnez une variable de substitution dans la liste déroulante [Contournement](#).
4. Cliquez sur [OK](#).

32.7.3 Désactivation du contournement

Après avoir créé et testé votre job, vous devez désactiver l'attribut de contournement avant de passer le job en mode de production.

Pour désactiver l'attribut de contournement au niveau du job, cliquez avec le bouton droit de la souris sur un flux de données ou un workflow dans la zone de projet Designer et sélectionnez [Supprimer le contournement](#).

Pour désactiver l'attribut de contournement au niveau de l'appel pour des flux de données ou des workflows multiples, effectuez les opérations suivantes :

1. Sélectionnez des flux de données et des workflows dans l'espace de travail Designer.
2. Cliquez avec le bouton droit de la souris et sélectionnez [Contournement](#).
La fenêtre [Définir le contournement](#) s'affiche.

3. Sélectionnez *{Pas de contournement}* dans la liste déroulante *Contournement*.
4. Cliquez sur *OK*.

33 Mécanismes de récupération

Des mécanismes de récupération sont disponibles dans SAP Data Services pour des jobs batch uniquement.

Informations associées

[Récupération suite à un échec d'exécution de job \[page 765\]](#)

[Récupération automatique de jobs \[page 766\]](#)

[Récupération manuelle de jobs à l'aide de tables de statuts \[page 774\]](#)

[Traitement de données avec problèmes \[page 774\]](#)

33.1 Récupération suite à un échec d'exécution de job

Si un job SAP Data Services n'est pas correctement exécuté, vous devez régler les problèmes qui ont empêché l'exécution correcte du job et exécuter le job à nouveau.

Cependant, au cours de l'échec de l'exécution du job, certains flux de données du job peuvent être exécutés, certaines tables peuvent avoir été chargées, partiellement chargées ou altérées. Vous devez donc créer vos jobs de mouvement de données de telle sorte que vous puissiez restaurer le job, en d'autres termes exécuter à nouveau le job et récupérer toutes les données sans données dupliquées ou manquantes.

Vous pouvez utiliser diverses techniques pour récupérer des données d'échec d'exécutions de job. Cette section présente deux techniques :

- Récupération automatique de jobs Fonctionnalité logicielle permettant d'exécuter des jobs qui ont échoué en mode de récupération.
- Récupération manuelle de jobs à l'aide de tables de statuts — Technique de conception permettant d'exécuter à nouveau des jobs sans tenir compte des résultats partiels d'une exécution précédente.

Vous pouvez être amené à utiliser une combinaison de ces techniques en fonction des relations entre des flux de données dans votre application.

Si vous n'utilisez pas ces techniques, vous serez peut-être amené à rejeter les modifications manuellement à partir de tables cible en cas d'interruptions au cours de l'exécution de job.

Informations associées

[Récupération automatique de jobs \[page 766\]](#)

[Récupération manuelle de jobs à l'aide de tables de statuts \[page 774\]](#)

33.2 Récupération automatique de jobs

La récupération automatique permet au logiciel d'enregistrer le résultat de chaque étape correctement exécutée dans un job. Si un job échoue, vous pouvez décider d'exécuter à nouveau le job en mode de récupération. En mode de récupération, le logiciel récupère les résultats pour des étapes correctement exécutées et exécute à nouveau les étapes non exécutées ou qui ont échoué sous les mêmes conditions que le job d'origine. A des fins de récupération, le logiciel considère les étapes déclenchant des exceptions comme des étapes ayant échoué même si l'étape est interceptée dans un bloc Try/Catch.

33.2.1 Activation de la récupération automatique

Pour utiliser la fonctionnalité de récupération automatique, vous devez l'activer au cours de l'exécution initiale d'un job. Le logiciel sauvegarde les résultats d'étapes correctement exécutées lorsque la fonction de récupération automatique est activée.

33.2.1.1 Exécuter un job à partir de Designer avec la récupération activée

1. Dans la zone de projet, sélectionnez le nom du job.
2. Cliquez avec le bouton droit et sélectionnez *Exécuter*.

Le logiciel vous invite à sauvegarder toutes les modifications.

3. Vérifiez que l'option *Activer la récupération* est sélectionnée dans la fenêtre *Propriétés d'exécution*.

Si cette case n'est pas cochée, le logiciel n'enregistre pas les résultats des étapes au cours du job et ne peut pas récupérer le job en cas d'échec. Dans ce cas, vous devez exécuter les opérations de récupération manuellement.

33.2.1.2 Exécuter un job avec récupération activée depuis l'Administrateur

Lorsque vous planifiez ou exécutez un job à partir de l'Administrateur, cochez la case *Activer la récupération*.

33.2.2 Marquage des unités de récupération

Dans certains cas, les étapes d'un workflow dépendent les unes des autres et doivent être exécutées ensemble. En raison de la dépendance, vous devez désigner le workflow comme "unité de récupération". Lorsqu'un workflow correspond à une unité de récupération, l'ensemble du workflow doit être correctement exécuté. Si le

workflow n'est pas exécuté correctement, le logiciel exécute l'intégralité du workflow au cours de la récupération, même les étapes exécutées correctement auparavant.

Cependant il existe des exceptions au traitement de l'unité de récupération. Par exemple, lorsque vous indiquez qu'un workflow ou un flux de données doit être exécuté une fois, un job n'exécutera jamais à nouveau ce workflow ou ce flux de données une fois qu'il a été correctement exécuté, sauf si ce workflow ou ce flux de données figure dans un workflow d'unité de récupération à nouveau exécuté et qui n'a pas été correctement exécuté ailleurs en dehors de l'unité de récupération.

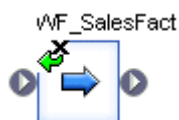
Il est recommandé de ne pas marquer un workflow ou un flux de données comme *Exécuter une fois* lorsque le workflow ou un workflow parent est une unité de récupération.

33.2.2.1 Indiquer un workflow comme unité de récupération

1. Dans la zone de projet, sélectionnez le workflow.
2. Cliquez avec le bouton droit et sélectionnez *Propriétés*.
3. Cochez la case *Restaurer en tant qu'unité* puis cliquez sur *OK*.

Au cours de la récupération, le logiciel considère ce workflow comme une unité. Si l'ensemble du workflow est correctement exécuté, c'est-à-dire sans erreur, au cours d'une exécution précédente, le logiciel récupère les résultats de l'exécution précédente. Si une étape du workflow n'a pas été correctement exécutée, l'ensemble du workflow est ré-exécuté au cours de la récupération.

Dans le diagramme Espace de travail, la croix noire et la flèche verte indiquent qu'un workflow est une unité de récupération.



33.2.3 Exécution en mode de récupération

Si l'exécution d'un job avec récupération automatique activée échoue, vous pouvez ré-exécuter ce job en mode de récupération.

Comme dans le cas de tout échec d'exécution de job, vous devez déterminer et supprimer la cause de la défaillance et ré-exécuter le job en mode de récupération. Si vous devez apporter des modifications au job lui-même pour corriger la défaillance, vous ne devez pas utiliser la récupération automatique. À la place, exécutez le job comme une première exécution.

En mode de récupération, le logiciel exécute les étapes ou les unités de récupération qui n'ont pas été correctement exécutées précédemment, cela comprend les étapes qui ont échoué et les étapes qui ont déclenché une exception mais qui ont été correctement exécutées, par exemple, celles d'un bloc Try/Catch. Comme lors d'une exécution classique de job, le logiciel exécute les étapes en parallèle si elles ne sont pas connectées dans les diagrammes de workflow et en série si elles sont connectées.

Informations associées

[Garantie d'un chemin d'exécution correct \[page 768\]](#)

33.2.3.1 Exécuter un job en mode de récupération à partir de Designer

1. Dans la zone de projet, sélectionnez le nom du job (qui a échoué).
2. Cliquez avec le bouton droit et sélectionnez *Exécuter*.

Le logiciel vous invite à sauvegarder tous les objets ayant des modifications non sauvegardées.

3. Vérifiez que la case *Récupération à partir de l'échec d'exécution* est cochée.

Cette option n'est pas disponible lorsqu'un job n'a pas encore été exécuté, lorsque l'exécution précédente a été correctement effectuée ou lorsque le mode de récupération a été désactivé au cours de l'étape précédente.

Lorsque vous sélectionnez Récupérer à partir du dernier échec d'exécution, le logiciel récupère les résultats de toutes les étapes correctement exécutées précédemment et exécute ou exécute à nouveau toutes les autres étapes.

Si vous désactivez cette option, le logiciel exécute ce job de nouveau en exécutant toutes les étapes.

Lorsque vous planifiez ou exécutez un job (qui a échoué) à partir de l'Administrateur, cochez la case *Récupérer à partir de la dernière exécution qui a échoué*.

33.2.4 Garantie d'un chemin d'exécution correct

Le système de récupération automatique nécessite qu'un job en mode de récupération soit à nouveau exécuté exactement comme précédemment.

Il est important que le job de récupération soit exécuté exactement comme précédemment. Si l'exécution du job a été autorisée à des conditions modifiées, supposez qu'une fonction `<sysdate>` retourne une nouvelle date de contrôle des données extraites, puis que les nouvelles données chargées dans les cibles ne correspondent plus aux données correctement chargées dans la cible au cours de la première exécution du job.

Par exemple, supposez que l'exécution d'un job de mise à jour quotidienne de nuit charge correctement des tables de dimension dans un magasin. Cependant, alors que le job est en cours d'exécution, le journal de base de données dépasse et arrête le chargement à partir de tables de faits. Le jour suivant, l'Administrateur tronque le fichier journal et exécute à nouveau le job en mode de récupération. Le job de récupération ne recharge pas les tables de dimension car elles ont été correctement chargées lors de l'exécution précédente.

Pour garantir que les tables de faits sont chargées avec les données qui correspondent aux données déjà chargées dans les tables de dimension, le job de récupération doit utiliser les mêmes critères d'extraction que ceux utilisés par le job d'origine lors du chargement des tables de dimension. Si le job de récupération a utilisé de nouveaux critères d'extraction, par exemple l'extraction de données de base à la date système actuelle, les données des tables de faits ne correspondront pas aux données précédemment extraites dans les tables de dimension.

En outre, si le job de récupération utilise de nouvelles valeurs, l'exécution de job peut suivre un chemin complètement différent via des étapes conditionnelles ou des blocs Try/Catch.

Si une valeur de variable globale est modifiée lors d'une récupération en tant que workflow d'unité, il n'est pas certain qu'une référence de variable globale en aval comporte la valeur mise à jour lors de l'exécution en mode de récupération. Tel est le cas lors de l'exécution en mode de récupération, si l'échec de l'exécution précédente ne figure pas dans le workflow contenant la modification de variable.

Lorsque la récupération est activée, le logiciel enregistre les résultats à partir des types suivants d'étapes :

- Workflows
- Flux de données batch
- Instructions de script
- Fonctions personnalisées (type sans état uniquement)
- Fonction SQL
- fonction exec
- fonction get_env
- fonction rand
- fonction sysdate
- fonction systime

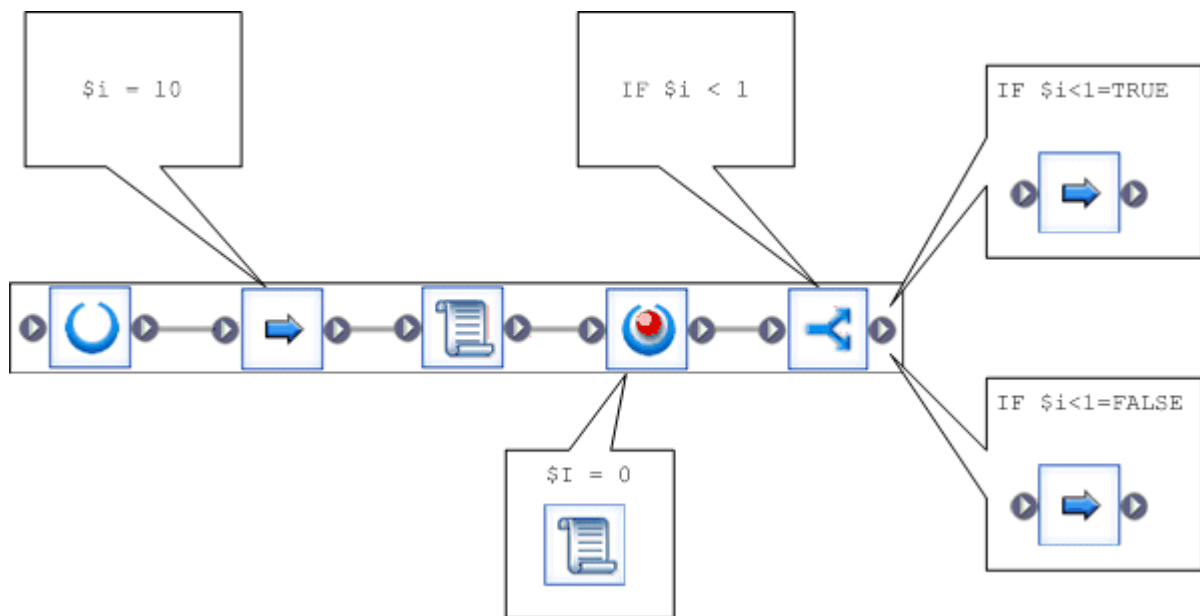
33.2.5 Utilisation de blocs Try/Catch avec récupération automatique

SAP Data Services ne sauvegarde pas le résultat d'un bloc Try/Catch pour le réutiliser au cours de la récupération. Si une exception a été générée dans un bloc Try/Catch, au cours de la récupération, le logiciel exécute l'étape qui a généré l'exception et les étapes suivantes.

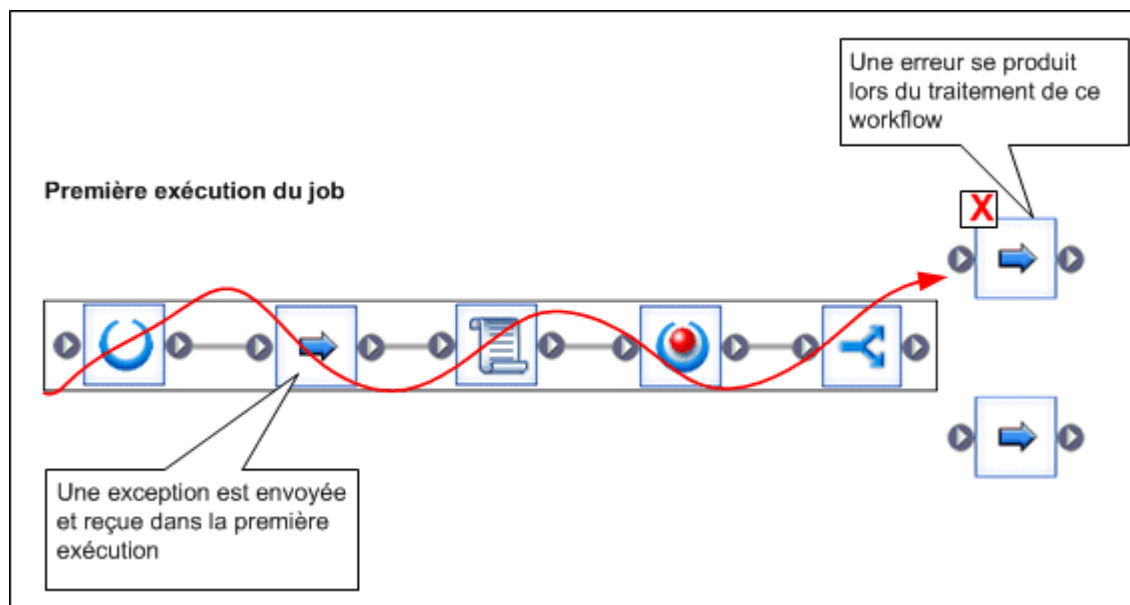
Le chemin d'exécution via le bloc Try/Catch devant être différent dans le job récupéré, l'utilisation d'un ensemble de variables dans le bloc Try/Cath pourrait modifier les résultats au cours de la récupération automatique.

Supposez, par exemple que vous créez un job qui définit une variable `$i`, que vous avez déterminée dans un bloc Try/Catch. Si une exception survient, vous pouvez définir une variable `$i`. Les étapes suivantes sont basées sur la valeur de `$i`.

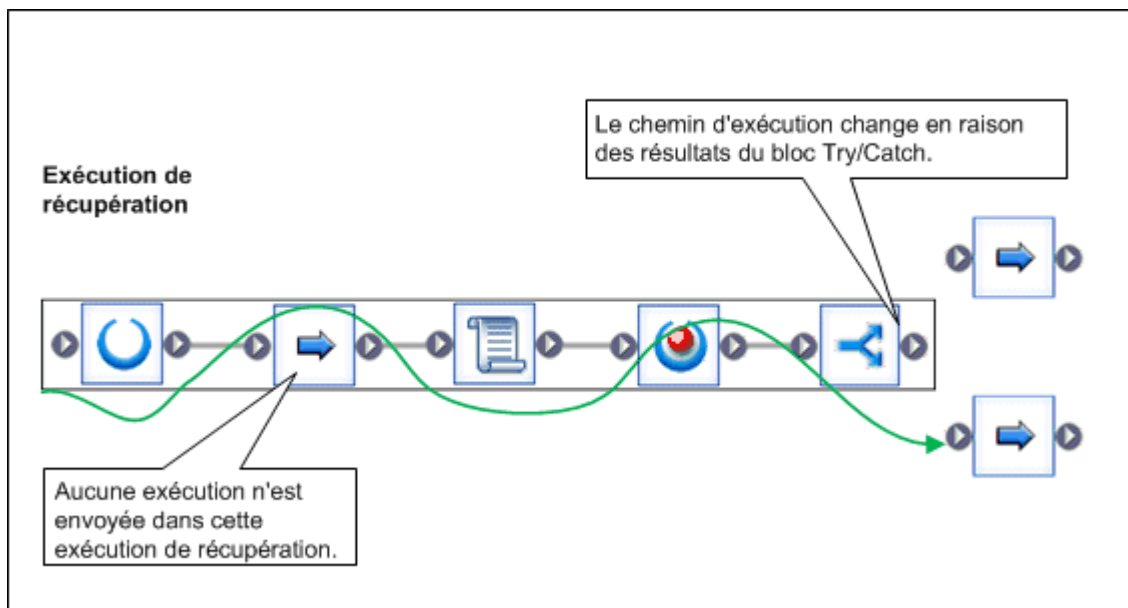
Logique d'exécution de job



Au cours de la première exécution de job, le premier workflow contient une erreur qui génère une exception détectée. Cependant, le job échoue dans le workflow suivant.



Vous corrigez l'erreur et exécutez le job en mode de récupération. Au cours de la récupération, le premier workflow ne génère plus l'exception. Par conséquent, la valeur de la variable `<$i>` est différente et le job sélectionne un workflow suivant différent avec différents résultats.



Pour garantir des résultats corrects, lorsqu'un job contient un bloc Try/Catch, n'utilisez pas d'ensemble de valeurs dans le bloc Try/Catch et dans toutes les étapes suivantes.

33.2.6 Garantie que les données ne sont pas dupliquées dans des cibles

Définissez des workflows permettant une récupération correcte de jobs. Un flux de données peut être partiellement complété lors d'une exécution incomplète. Parmi les lignes requises certaines seulement ont pu alors être insérées dans une table. Vous ne souhaitez pas insérer des lignes dupliquées au cours de la récupération lorsque le flux de données est à nouveau exécuté.

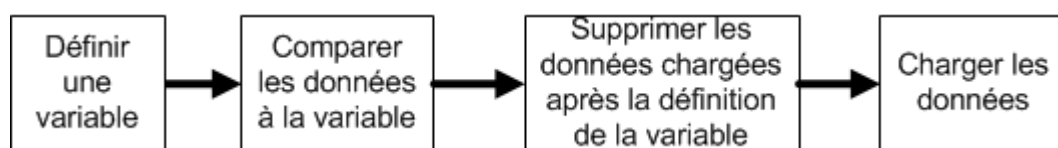
Vous pouvez utiliser plusieurs méthodes pour garantir que vous n'insérez pas de lignes dupliquées :

- Créez le flux de données pour remplacer entièrement la table cible au cours de chaque exécution. Cette technique peut être optimisée lorsque de nombreuses modifications sont apportées à la table cible, par rapport à la taille de la table. Vous pouvez utiliser des techniques d'optimisation, par exemple les options de chargement par lots pour améliorer la performance globale.
- Activez l'option de chargement de correction automatique pour la table cible. Cette option vérifie qu'il existe des lignes dans la table cible avant d'en ajouter de nouvelles. L'utilisation de l'option de chargement par lots peut cependant ralentir les jobs exécutés en mode de non-récupération lorsque ce n'est pas nécessaire. Utilisez cette technique lorsque la table cible est grande et peu de modifications y sont apportées.
- Incluez une commande SQL à exécuter avant le chargement de la table. Les commandes SQL de préchargement peuvent supprimer des mises à jour de base de données partielles apparues au cours de l'exécution incomplète d'une étape dans un job. En règle générale, la commande SQL de préchargement supprime des lignes en fonction d'une variable, c'est-à-dire avant le début de l'étape d'insertion.

33.2.7 Utilisation de SQL de préchargement pour autoriser des flux de données ré-exécutables

Pour utiliser des commandes SQL de préchargement pour supprimer des mises à jour de base de données partielles, les tables doivent contenir un champ qui permet de déterminer quand une ligne a été insérée. Créez une commande SQL de préchargement qui supprime des lignes en fonction de la valeur de ce champ.

Par exemple, supposez qu'une table contienne une colonne qui enregistre l'horodatage d'une insertion de ligne. Vous pouvez créer un script avec une variable qui enregistre l'horodatage actuel avant l'insertion de nouvelles lignes. Dans les options de la table cible, ajoutez une commande SQL de préchargement qui supprime toutes les lignes avec un horodatage supérieur à celui enregistré par la variable.



Au cours de l'exécution initiale, aucune ligne ne correspond aux critères de suppression. Au cours de la récupération, la valeur de la variable n'est pas réinitialisée. (La valeur de la variable est définie dans un script exécuté correctement au cours de l'exécution initiale.) Les lignes insérées au cours du chargement partiel précédent de base de données doivent correspondre à ces critères et la commande SQL de préchargement doit les supprimer.

Pour utiliser correctement des commandes SQL de préchargement, vous devez définir des variables et les transférer correctement aux flux de données.

33.2.7.1 Utilisation de commandes SQL de préchargement pour garantir une récupération correcte

1. Déterminez les valeurs appropriées que vous pouvez utiliser pour suivre des enregistrements insérés dans vos tables.

Par exemple, si chaque ligne dans une table est marquée d'un horodatage d'insertion, vous pouvez utiliser la valeur de la fonction `sysdate()` pour déterminer quand une ligne a été ajoutée à cette table.

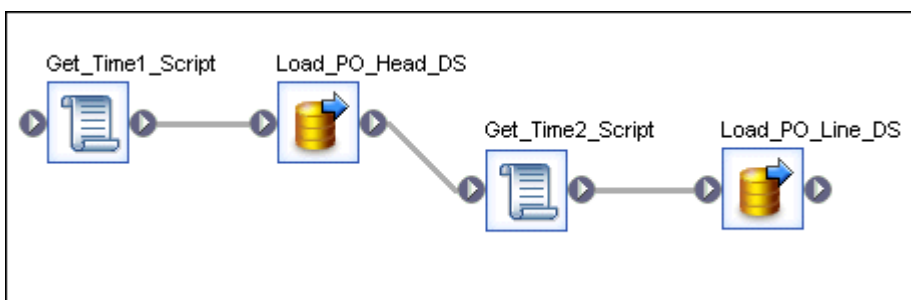
2. Créez des variables pouvant archiver les valeurs de "suivi".

Les variables sont spécifiques au job ou au workflow. Si un workflow est une unité de récupération, créez les variables de "suivi" pour ce workflow au niveau du job, sinon, créez vos variables de suivi au niveau du workflow. Généralement, vous ne souhaitez pas réinitialiser des variables de suivi au cours de la récupération car lorsque les variables sont réinitialisées, la commande SQL de préchargement ne fonctionnera pas correctement.

3. Créez des scripts qui définissent les variables sur les valeurs appropriées.

Les scripts sont des étapes uniques de jobs ou de workflows. Vous devez créer un script distinct qui définit les variables requises avant que chaque flux de données ou workflow ne charge une table. Si un workflow correspond à une unité de récupération, créez les scripts pour ce workflow au niveau du job ; sinon créez les scripts au niveau du workflow.

- Connectez les scripts aux flux de données ou aux workflows correspondants.



- Créez des paramètres pour transférer les informations relatives à la variable du job ou du workflow où vous avez créé la variable aux flux de données, qui utilise la variable de suivi dans la commande SQL de préchargement.
- Insérez les commandes SQL de préchargement appropriées qui suppriment tous les enregistrements insérés au cours des exécutions précédentes qui ont échoué.

Les commandes SQL de préchargement référencent les paramètres contenant la variable de suivi en supprimant les lignes insérées après l'activation de la variable.

Par exemple, supposez que la table `<PO_ITEM>` enregistre l'horodatage vers date dans la colonne `<TIMESTMP>`. Vous avez créé une variable `<$load_time>` qui enregistre la valeur de la fonction `sysdate()` avant le début du chargement et vous avez transféré cette variable au flux de données qui charge la table `<PO_ITEM>` dans un paramètre nommé `<$load_time>`. Puis votre commande SQL de préchargement doit supprimer tous les enregistrements de la table si la valeur dans `<TIMESTMP>` est supérieure à la valeur dans `<$load_time>`.

Supprimez de `<PO_ITEM>` si `<TIMESTMP>` `>` [`<$load_time>`]

Informations associées

[Définir une variable locale \[page 318\]](#)

[Scripts \[page 250\]](#)

33.3 Récupération manuelle de jobs à l'aide de tables de statuts

Vous pouvez créer vos jobs et workflows pour pouvoir les récupérer manuellement en cas d'échec d'une exécution. Un job conçu pour une récupération manuelle doit avoir certaines caractéristiques :

- Vous pouvez exécuter le job de façon répétitive.
- Le job implémente des étapes spécifiques pour récupérer des données lorsqu'une étape n'a pas été correctement exécutée précédemment.

Vous pouvez utiliser une table des statuts d'exécution pour créer des jobs pouvant être exécutés plusieurs fois sans duplication des lignes cibles. La table enregistre un statut d'exécution de job. Une valeur "Echec" indique à SAP Data Services un chemin d'accès de récupération.

Pour implémenter un workflow avec un chemin d'exécution de récupération :

- Définissez un indicateur indiquant lorsque le workflow est en cours d'exécution en mode de récupération.
- Archivez la valeur de l'indicateur dans une table de statuts.
- Vérifiez la valeur de l'indicateur dans la table de statuts avant d'exécuter un workflow pour déterminer le chemin à exécuter dans le workflow.
- Mettez à jour la valeur de l'indicateur lorsque le workflow est correctement exécuté.

Par exemple, vous pourriez créer un workflow qui utilise l'option de chargement de correction automatique lorsqu'une précédente exécution n'est pas correcte. Ce workflow aurait cinq étapes comme l'illustre la figure :

1. Récupérez la valeur de code qui indique si l'exécution précédente est correcte ou si elle a échoué, à partir de la table de statuts. Archivez cette valeur dans une variable, par exemple `<$recovery_needed>`.
2. Dans une structure conditionnelle, évaluez la variable `<$recovery_needed>`.
3. Si une récupération est requise, exécutez le flux de données de récupération `recover_customer`. Ce flux de données charge les données à l'aide de l'option de correction automatique de chargement.
4. Si aucune récupération n'est requise, exécutez le flux de données de non-récupération `load_customer`. Ce flux de données charge les données sans option de correction automatique de chargement.
5. Mettez à jour la valeur de l'indicateur dans la table de statuts pour indiquer que l'exécution est correcte.

33.4 Traitement de données avec problèmes

Les jobs peuvent ne pas donner les résultats que vous attendez en raison de problèmes avec les données. Dans certains cas, le logiciel est incapable d'insérer une ligne. Dans d'autres cas, le logiciel peut insérer des lignes avec des informations manquantes. Vous pouvez concevoir vos flux de données pour anticiper et traiter ces types de problèmes. Par exemple, vous pouvez avoir un flux de données qui écrit des lignes avec des informations manquantes dans un fichier spécifique que vous pouvez vérifier ultérieurement.

Cette section décrit des mécanismes que vous pouvez utiliser pour anticiper et traiter des problèmes de données. Cette section aborde notamment trois techniques :

- Utilisation de fichiers de dépassement

- Filtrage pour valeurs manquantes ou erronées
- Traitement de faits avec dimensions manquantes

33.4.1 Utilisation de fichiers de dépassement

Une ligne qui ne peut pas être insérée constitue un problème de données général. Utilisez le fichier de dépassement pour traiter ce type de problème de données. Lorsque vous indiquez un fichier de dépassement et que SAP Data Services ne peut pas charger une ligne dans une table, le logiciel écrit la ligne sur le fichier de dépassement à la place. Le journal de suivi indique le flux de données où le chargement a échoué et l'emplacement du fichier.

Pour toute table utilisée comme cible, vous pouvez activer l'option d'utilisation d'un fichier de dépassement dans l'onglet *Options*. Lorsque vous indiquez un fichier de dépassement, donnez un nom de chemin complet pour garantir que le logiciel crée un fichier unique lorsque plusieurs fichiers sont créés dans le même job. Par défaut, le nom du fichier de dépassement est le nom de la table cible.

Lorsque vous sélectionnez l'option de fichier de dépassement, vous sélectionnez les données que le logiciel écrit dans le fichier, concernant les lignes pour lesquelles le chargement a échoué : soit les données de la ligne, soit les commandes SQL requises pour charger la ligne. Si vous sélectionnez des données, vous pouvez utiliser le logiciel pour lire les données à partir du fichier de dépassement, les nettoyer et les charger dans la table cible. Si vous sélectionnez des commandes SQL, vous pouvez les utiliser pour charger la cible manuellement lorsqu'elle est accessible.

Plusieurs raisons peuvent être à l'origine de l'échec d'un chargement, par exemple :

- Capacité de mémoire saturée
- Paramètres de colonne de dépassement
- Valeurs clés doubles

Vous pouvez utiliser les informations de dépassement pour identifier des données non valides dans votre source ou les problèmes du mouvement de données. Toute nouvelle exécution écrasera le fichier de dépassement existant.

i Remarque

Vous ne pouvez pas utiliser des fichiers de dépassement lors d'un chargement dans une structure de transfert BW.

33.4.2 Filtrage pour valeurs manquantes ou erronées

Une valeur manquante ou non valide dans les données source constitue un autre problème de données commun. Les requêtes dans les flux de données vous permettent d'identifier des valeurs manquantes ou non valides dans des données sources. Vous pouvez également décider d'inclure ces données dans la cible ou de les ignorer.

Par exemple, supposez que vous extrayiez des données d'une source et que vous sachiez que certains numéros de téléphone et noms de client manquent. Vous pouvez utiliser un flux de données pour extraire des données de la source, charger ces données dans une cible et filtrer les valeurs nulles dans un fichier pour votre contrôle.

Le flux de données comporte les cinq étapes suivantes :

1. Il extrait les données de la source.
2. Il sélectionne le jeu de données à charger dans la source et applique de nouvelles clés. (Cette étape est exécutée à l'aide de la fonction Key_Generation.)
3. Il charge le jeu de données dans la cible, à l'aide de l'option Chargement par lots pour une meilleure performance.
4. Il utilise le jeu de données pour lequel de nouvelles clés ont été générées dans l'étape 2 et sélectionne des lignes avec noms de client et numéros de téléphone manquants.
5. Il écrit les ID client pour les lignes avec données manquantes dans un fichier.

Maintenant, supposez que vous souhaitiez charger des lignes avec des noms de client manquants dans votre cible. Vous pouvez insérer une autre requête dans le flux de données pour garantir que SAP Data Services n'insère pas de lignes incomplètes dans la cible. La nouvelle requête filtre les lignes avec noms de client manquants avant de charger des lignes dans la cible. La requête de données manquantes continue à collecter ces lignes avec les lignes avec numéros de téléphone manquants. Dans cette version de l'exemple, la transformation Key_Generation ajoute des clés pour les nouvelles lignes avant d'insérer le jeu de données filtrées dans la cible.

Maintenant le flux de données comporte six étapes.

1. Il extrait les données de la source.
2. Il sélectionne le jeu de données à charger dans la cible en filtrant les lignes sans nom de client.
3. Il génère des clés pour les lignes avec noms de client.
4. Il charge le jeu de données valide (lignes avec noms de client) dans la cible à l'aide de l'option Chargement par lots pour une meilleure performance.
5. Il utilise une transformation Query distincte pour sélectionner les lignes de la source qui ne comportent ni noms ni numéros de téléphone.
Notez que le logiciel ne charge pas de lignes avec noms de client manquants dans la cible ; cependant le logiciel charge les lignes avec numéros de téléphone manquants.
6. Il écrit les ID client pour les lignes avec données manquantes dans un fichier.

Vous pourriez ajouter plus de requêtes dans le flux de données pour sélectionner des valeurs manquantes ou non valides pour un contrôle ultérieur.

33.4.3 Traitement de faits avec dimensions manquantes

Un autre problème de données se produit lorsque SAP Data Services recherche une table de dimensions et ne peut pas trouver les valeurs requises pour compléter une table de faits.

Vous pouvez traiter ce problème de plusieurs façons :

- Sortir la ligne posant problème de la table de faits.
En règle générale, ce n'est pas une bonne idée car l'analyse effectuée sur les faits ne tiendra pas compte de cette ligne.
- Noter la ligne ayant généré l'erreur mais la charger quand-même dans la table cible.
Vous pouvez marquer la ligne comme étant erronée ou transférer les informations relatives à la ligne dans un fichier d'erreurs comme dans les exemples de [Filtrage pour valeurs manquantes ou erronées \[page 775\]](#).
- Régler le problème via un programme.

Selon les données manquantes, vous pouvez insérer une nouvelle ligne dans la table de dimensions, ajouter des informations d'une source secondaire ou utiliser une autre méthode fournissant des données en dehors du chemin de haute performance normal.

33.5 Echange de métadonnées

SAP Data Services vous permet d'exporter des métadonnées dans un fichier XML. Après avoir créé le fichier, vous devez l'importer manuellement dans un autre outil.

Informations associées

[Echange de métadonnées \[page 777\]](#)

33.5.1 Echange de métadonnées

Vous pouvez échanger des métadonnées entre le logiciel et des outils tiers à l'aide de fichiers XML et de l'option [Echange de métadonnées](#).

Le logiciel prend en charge deux formats d'échange de données intégrés :

- CWM 1.0 XML/XMI 1.1
CWM (le Common Warehouse Metamodel)— est une spécification qui permet l'échange des métadonnées de l'entrepôt de données entre outils, plateformes et référentiels dans des environnements hétérogènes distribués.
- ERwin 4.x XML

Le logiciel peut aussi utiliser :

- MIMB (le *Meta Integration® Model Bridge*)
MIMB est un utilitaire autonome Windows qui convertit les modèles de métadonnées. En utilisant MIMB avec le logiciel, vous pouvez échanger des métadonnées avec tous les formats pris en charge par MIMB. Si MIMB est installé, les formats supplémentaires qu'il prend en charge sont indiqués dans la fenêtre Echange de métadonnées.

33.5.1.1 Importation de fichiers de métadonnées dans le logiciel

Vous pouvez importer des métadonnées ERwin Data Modeler 4.x XML dans une banque de données.

33.5.1.1.1 Importer des métadonnées à l'aide de l'option Echange de métadonnées

1. Dans le menu *Outils*, sélectionnez *Echange de métadonnées*.
2. Dans la fenêtre Echange de métadonnées, sélectionnez *Importer les métadonnées depuis le fichier*.
3. Dans la case *Format de métadonnées*, sélectionnez *ERwin 4.x XML* dans la liste des formats disponibles.
4. Spécifiez le *Nom du fichier source* (saisissez-le directement ou cliquez sur *Parcourir* pour le rechercher).
5. Sélectionnez le *Nom de la banque de données cible* dans la liste des banques de données.
6. Cliquez sur *OK* pour terminer l'importation.

33.5.1.2 Exportation de fichiers de métadonnées du logiciel

Vous pouvez exporter des métadonnées dans un fichier lisible par d'autres outils.

33.5.1.2.1 Exporter des métadonnées à l'aide de l'option Echange de métadonnées

1. Dans le menu *Outils*, sélectionnez *Echange de métadonnées*.
2. Dans la fenêtre Echange de métadonnées, sélectionnez *Exporter des métadonnées Data Services vers un fichier*.
3. Sélectionnez un *format de métadonnées* pour la cible dans la liste des formats disponibles.

Si MIMB est installé et que vous sélectionnez un format pris en charge par MIMB, cochez la case *Visuel* pour ouvrir l'application MIMB à la fin du processus d'exportation.

Si vous ne cochez pas la case *Visuel*, les métadonnées sont exportées sans ouvrir l'application MIMB. L'utilisation de l'application MIMB offre des options de configuration supplémentaires permettant de structurer les métadonnées dans le fichier exporté.

4. Spécifiez le nom du fichier source (saisissez-le directement ou cliquez sur *Parcourir* pour le rechercher).

Lorsque vous recherchez le fichier, vous ouvrez une fenêtre de navigateur type.

Recherchez l'un des formats/types de fichiers suivants :

Format	Type de fichier
DI CWM 1.0 XML/XMI 1.1	XML
DI ERwin 4.x XML	XML
Format MIMB (uniquement si installé)	Tout

Après avoir sélectionné un fichier, cliquez sur *Ouvrir*.

5. Sélectionnez le *Nom de la banque de données source* dans la liste des banques de données.

6. Cliquez sur [OK](#) pour terminer l'exportation.

33.6 Chargement du fichier de données volumineux avec option de récupération

Le chargement du fichier de données volumineux avec option de récupération est une extension du mécanisme de récupération existant dans Data Services. Les versions précédentes de Data Services prenaient en charge le mécanisme de récupération uniquement jusqu'à la couche de flux de données. Par conséquent, si le moteur s'est bloqué pour un flux de données, tout le flux de données a dû être redémarré depuis le début en mode de récupération.

Cependant, s'il existe des données volumineuses dans le flux de données échoué, le redémarrage du processus engendrera la duplication du travail de nettoyage et de redémarrage. La situation de redémarrage empire si le job doit être interrompu périodiquement (par exemple à cause d'une déconnexion du réseau, d'une expiration de la base de données ou d'un redémarrage de la machine).

Pour éviter cette situation, vous pouvez activer la fonctionnalité de récupération du chargement des données volumineuses lors de la lecture d'un fichier plat source volumineux dans les tables de bases de données. Par exemple, lorsque vous chargez un fichier d'entrée volumineux dans une base de données, il se peut que le job dure plusieurs jours. Si le moteur se bloque ou échoue au cours d'un job, ce dernier peut être repris à partir du dernier point de contrôle échoué. Ce point de contrôle inclut le nom de fichier échoué et le décalage du fichier qui a été traité jusqu'à présent. La fonctionnalité de récupération évite le redémarrage du flux de données et garantit la progression du chargement à partir du point d'échec.

33.6.1 Activation de l'option de récupération pour le chargement de données volumineuses

Pour utiliser la fonctionnalité de récupération automatique pour le chargement de données volumineuses, vous devez l'activer lors de l'exécution initiale d'un job. Ainsi, en cas d'échec d'un job, celui-ci peut reprendre à partir du dernier point contrôlé et poursuivre, au lieu de reprendre depuis le début.

Pour exécuter un job de chargement de données volumineuses à partir de Designer avec la récupération activée :

1. Dans la zone de projet, sélectionnez le nom du job.
2. Cliquez avec le bouton droit et sélectionnez [Exécuter](#). Le logiciel vous invite à sauvegarder toutes les modifications.
3. Vérifiez que l'option [Activer la récupération](#) est sélectionnée dans la fenêtre [Propriétés d'exécution](#).

i Remarque

La case [Activer la récupération](#) doit être cochée pour que le logiciel enregistre les étapes au cours du job. Si vous ne cochez pas la case et que le job échoue, vous ne pourrez pas le récupérer. Dans ce cas, vous devrez exécuter les opérations de récupération manuellement et relancer le job.

4. Dans la table cible du flux de données qui contient la source du fichier de données volumineuses, contrôlez l'existence d'une clé primaire. S'il n'existe aucune clé, vous pouvez soit ajouter la colonne de clé primaire

dans la table cible, soit générer la colonne ID de ligne en tant que colonne de clé primaire dans le fichier source, puis la transmettre aux tables cible.

- a. Pour ajouter la clé primaire à la table cible : si le schéma d'origine de la table cible ne comporte pas de clé primaire, activez [Utiliser des clés d'entrée](#) dans les options de la table cible et assurez-vous que l'entrée dans la table cible a une clé primaire. Par exemple, la colonne ID de ligne que vous avez saisie dans le fichier source peut être transmise à la table cible en tant que clé primaire.
- b. Pour générer la colonne ID de ligne pour le fichier source : saisissez un nom de colonne valide dans le champ d'entrée ID DE LIGNE, situé sous ► [Lecteur de fichiers](#) ► [Général](#) ►.

i Remarque

Le système ajoute automatiquement un ID de ligne au schéma en tant que première colonne en haut. Le nom de colonne est le nom que vous avez saisi et la colonne Type de données est double.

Une colonne de clé primaire est requise pour une récupération rapide, car le moteur utilise la correction automatique du chargement pour le dernier job échoué. Si vous ne sélectionnez pas la colonne de clé primaire, la correction automatique du chargement est fortement ralentie. Pour remplacer cette correction lente, vous pouvez ajouter la colonne ID DE LIGNE manuellement si la table cible ne comporte pas de clé primaire.

Data Services active automatiquement la correction automatique du chargement pour le batch de récupération, indépendamment du paramètre cible. De plus, pour HANA, Netezza et Sybase IQ, la correction automatique du chargement du batch de récupération utilise le chargement par lots pour une meilleure performance.

33.6.2 Limitations

Certaines limitations doivent être respectées pour qu'un fichier puisse être chargé en mode de récupération. Le fichier ainsi que le flux de données et le job qu'il contient doit répondre aux critères suivants :

- Le fichier plat est délimité ou de largeur fixe.
- Le fichier a un séparateur de lignes.
- Le job est exécuté avec [Activer la récupération](#).
- Le job n'est pas un job en temps réel.
- Le job n'est pas exécuté dans le débogueur.
- Le flux de données n'est pas dans un workflow continu.
- Le flux de données n'est pas audité.
- Le flux de données n'est pas un flux de données ABAP.
- Le flux de données ne s'exécute pas en mode de flux de données secondaire.
- Le flux de données ne contient aucun chargeur qui ne se connecte pas au lecteur de fichier.
- Le flux de données ne comporte pas les transformations et fonctions suivantes :
 - Row_Generation
 - Date_Generation
 - gen_row_num
 - gen_row_num_by_group
 - Source CDM
 - Cible incluse dans la transaction.
- L'aval de la source du fichier ne comporte pas les informations suivantes :

- Transformation Query qui a un ordre par clause
- Transformation Query qui a un groupe par clause
- Transformation Query qui sélectionne des lignes distinctes
- Transformation Query qui a une fonction d'agrégat
- Transformation Query qui a une jointure dans laquelle le fichier est une source interne

i Remarque

Vous pouvez utiliser le classement de jointure pour contrôler le classement de la source liée à d'autres tables et fichiers dans un flux de données. Le système joint les sources ayant des classements supérieurs avant les sources possédant des classements inférieurs.

- Transformation Hierarchy_Flattening
- Transformation Table_Comparison
- Transformation Reverse_Pivot

i Remarque

La transformation Reverse_Pivot en aval du gros fichier est autorisée uniquement si l'option "Les données d'entrée sont regroupées" est cochée.

- Validation
- Merge

i Remarque

La transformation Merge en aval du gros fichier est autorisée uniquement si toutes les entrées de cette transformation proviennent directement ou indirectement du fichier.

- Transformation Data Quality qui utilise une clé de décomposition, telle que la transformation Match.
- Soit la table cible a une clé primaire, soit l'entrée dans la table cible a une clé primaire et l'option "Utiliser la clé d'entrée" de la cible est définie sur Oui.

Si l'une des conditions ci-dessus n'est pas remplie, le système n'autorise pas la récupération du fichier. Cependant, cela n'a aucun effet sur les autres fichiers source ou flux de données.

Si le flux de données contient plusieurs sources de fichier, le système sélectionne une seule source à exécuter en mode de récupération.

34 Capture de données modifiées

Lorsque vous avez une grande quantité de données à mettre à jour régulièrement et peu de temps d'arrêt système pour la maintenance planifiée d'un entrepôt de données, mettez à jour les données au fil du temps ou effectuez un chargement de delta. L'actualisation complète et la capture des données modifiées (CDM) sont deux méthodes de chargement de delta couramment utilisées.

34.1 Actualisation complète

L'actualisation complète est facile à implémenter et à gérer. Cette méthode garantit qu'aucune donnée ne sera négligée ou perdue en raison d'erreurs techniques ou de programmation. Pour un environnement comportant une quantité de données source gérable, l'actualisation complète est une méthode simple à utiliser pour effectuer un chargement de delta vers un système cible.

34.2 Capture des modifications uniquement

À la fin d'un chargement initial, vous pouvez choisir d'extraire uniquement les données nouvelles ou modifiées et de mettre à jour le système cible.

L'identification et le chargement des seules données modifiées est appelé capture des données modifiées (CDM). Sont comprises uniquement les données incrémentielles ayant été modifiées depuis le dernier cycle d'actualisation. SAP Data Services agit sous forme de mécanisme pour localiser et extraire uniquement les données incrémentielles ayant été modifiées depuis la dernière actualisation.

L'amélioration des performances et la conservation d'un historique sont les deux motifs les plus importants justifiant l'utilisation de la capture des données modifiées.

- Les performances sont améliorées parce qu'il y a moins de données à extraire, transformer et charger, le job dure habituellement moins longtemps.
- Si le système cible doit suivre l'historique des modifications afin que ces données puissent être correctement analysées au fil du temps, la méthode de capture des données modifiées peut fournir un enregistrement de ces modifications.

❖ Exemple

Si un client déménage d'une région de vente dans une autre, la simple mise à jour de la fiche client pour tenir compte de la nouvelle région affecte toute analyse par région dans le temps, parce que les achats effectués par ce client avant son déménagement sont attribués à la nouvelle région.

34.3 CDM basée sur la source et basée sur la cible

La capture des données modifiées peut être basée sur la source ou la cible.

34.3.1 CDM basée sur la source

La capture des données modifiées basée sur la source extrait uniquement les lignes modifiées de la source. Elle est parfois appelée extraction incrémentielle. Cette méthode est privilégiée parce qu'elle améliore les performances en extrayant le dernier nombre de lignes.

SAP Data Services offre un accès aux données modifiées basées sur la source proposées par divers fournisseurs de logiciels. Le tableau suivant répertorie les sources de données prises en charge ainsi que les produits et techniques de capture des données modifiées.

Source de données	Produits ou techniques
Oracle 9i et versions ultérieures	Utilisez les packages CDM d'Oracle pour gérer les tables CDM. Ces packages utilisent un modèle de publication et d'abonnement. Vous pouvez créer une banque de données CDM pour les sources Oracle à l'aide de Designer. Vous pouvez également utiliser Designer pour créer des tables CDM dans Oracle, puis les importer pour les utiliser dans des jobs.
Sources de données d'ordinateur central (Adabas, DB2 UDB pour z/OS, IMS, SQL/MP, VSAM, fichiers plats) avec accès via Attunity Connect	Pour les sources de données d'ordinateur central qui utilisent Attunity pour se connecter au progiciel, vous pouvez utiliser Attunity Streams 4.6.
Bases de données Microsoft SQL Server	Utilisez le serveur de réplication Microsoft SQL pour capturer les données modifiées des bases de données SQL Server.
Autres sources	Utilisez les champs de date et heure pour comparer les exécutions de jobs de capture des données modifiées basés sur la source. Cette technique utilise un horodatage de création et/ou modification sur chaque ligne. Vous pouvez comparer les lignes en utilisant comme référence l'heure de la dernière mise à jour. Cette méthode s'appelle CDM basée sur l'horodatage.
Sources SAP ODP	Définissez les critères de filtre CDM lors de l'importation d'une source ODP.
i Remarque L'interface d'extracteurs SAP a été étendue et remplacée par des sources OPD. Consultez le guide Nouveautés pour les sources Operational Data Provisioning [ODP].	
SAP Sybase Replication Server	Utilisez SAP Sybase Replication Server pour capturer les données modifiées.

Informations associées

[Utilisation de CDC avec des sources Oracle \[page 784\]](#)

[Utilisation de la CDM avec les sources d'un ordinateur central Attunity \[page 798\]](#)

[Utilisation de la CDM avec les bases de données Microsoft SQL Server \[page 821\]](#)

[Utilisation de la CDM avec des sources basées sur l'horodatage \[page 835\]](#)

[Utilisation de la CDM avec SAP Replication Server \[page 805\]](#)

34.3.2 CDM basée sur la cible

La capture des données modifiées basée sur la cible extrait toutes les données de la source, mais charge uniquement les lignes modifiées dans la cible.

La capture des données modifiées basée sur la cible est utile quand vous désirez capturer l'historique, mais que vous n'avez pas la possibilité d'utiliser la capture des données modifiées basée sur la source. Le progiciel propose la comparaison de tables pour prendre en charge cette méthode.

34.4 Utilisation de CDC avec des sources Oracle

Utiliser CDC Oracle pour limiter le nombre de lignes lues régulièrement.

Si votre environnement doit garder à jour de grandes quantités de données, la fonctionnalité de CDM (Capture des données modifiées) d'Oracle est une solution simple pour limiter le nombre de lignes devant être lues régulièrement par le progiciel. Une source qui lit uniquement les opérations les plus récentes (INSERT, UPDATE, DELETE) permet de concevoir des chargements de delta plus légers et plus rapides.

34.4.1 Présentation de la CDM pour les bases de données Oracle

Avec Oracle 9i ou supérieur, SAP Data Services gère l'environnement CDM en accédant aux packages de publication et d'abonnement d'Oracle. Oracle publie les données modifiées de la table d'origine dans sa table CDM.

Designer permet de créer ou importer les tables CDM et de créer des abonnements pour accéder aux données de la table CDM. Des abonnements distincts permettent à chaque utilisateur de garder une trace de la dernière ligne modifiée à laquelle il a accédé.

Vous pouvez également activer des points de contrôle pour les abonnements afin que le progiciel lise uniquement les dernières modifications dans la table CDM.

Oracle utilise les termes suivants pour la Capture des données modifiées :

Terme	Description
Table des modifications (CDM)	Table relationnelle contenant les données modifiées résultant des opérations DML effectuées sur une table source.
Set de modifications	Groupe de tables CDM transactionnellement cohérentes. Par exemple, les tables CommandeClient et PosteVente doivent être dans un set de modifications pour garantir que les modifications d'une commande et de ses postes individuels soient capturées ensemble.
Source de modification	Base de données qui contient un ou plusieurs sets de modifications
Editeur	Personne qui capture et publie les données modifiées. L'éditeur est habituellement un administrateur de base de données qui crée et gère les objets de schéma qui constituent la base de données source et la base de données intermédiaire.
Mode de publication	Indique quand et comment capturer les données modifiées. Pour en savoir plus, voir le tableau suivant sur les modes de publication.
Base de données source	Base de données de production qui contient les données que vous avez extraites pour votre chargement initial. La base de données source contient les tables source.
Base de données intermédiaire	Base de données où sont publiées les données modifiées. Selon le mode de publication, la base de données intermédiaire peut être identique ou différente de la base de données source.
Abonné	Utilisateur pouvant accéder aux données publiées dans les tables CDM.
Abonnement	Contrôle l'accès aux données modifiées d'une ou plusieurs tables source dans un seul set de modifications. Un abonnement contient une ou plusieurs vues d'abonné.
Vue d'abonné	Données modifiées pour lesquelles l'éditeur a accordé le droit d'accès à l'abonné afin de les utiliser.

Oracle prend en charge les modes de publication suivants :

Synchrone

- Les données sont capturées à l'aide de déclencheurs internes dans les tables source afin de stocker les modifications dans les tables CDM.
- Les données capturées sont disponibles en temps réel.
- Les tables CDM doivent se trouver dans la base de données source.

Commentaires :

- Ajoute des données inutiles à la base de données source au moment de la capture.
- Disponible dans Oracle 9i et 10G.

HotLog asynchrone

- Les données sont capturées à l'aide des fichiers redo log ou des journaux d'archives de la base de données source.
- Les données capturées sont disponibles pratiquement en temps réel.
- Un set de modifications contient plusieurs tables CDM et doit se trouver localement dans la base de données.

Commentaires :

- Améliore les performances parce que les données sont capturées hors connexion.
- Disponible uniquement dans Oracle 10G.

AutoLog asynchrone

- Les données sont capturées à l'aide des fichiers redo log gérés par les services de transport de journal qui automatisent le transfert de la base de données source à la base de données intermédiaire.

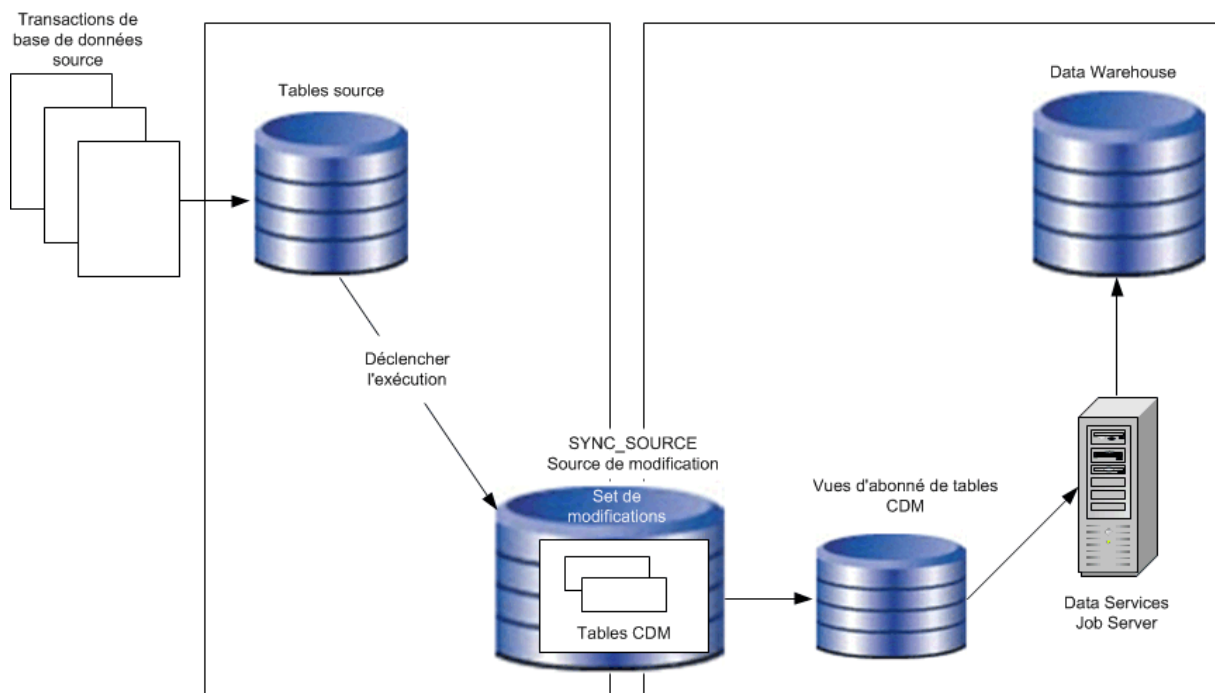
- La disponibilité des données capturées dépend de la fréquence des modifications des fichiers redo log dans la base de données source.
- Un set de modifications contient plusieurs tables CDM et peut être distant ou local par rapport à la base de données.

Commentaires :

- Améliore les performances parce que les données sont capturées hors connexion.
- Disponible uniquement dans Oracle 10G.

34.4.1.1 CDM Oracle en mode synchrone

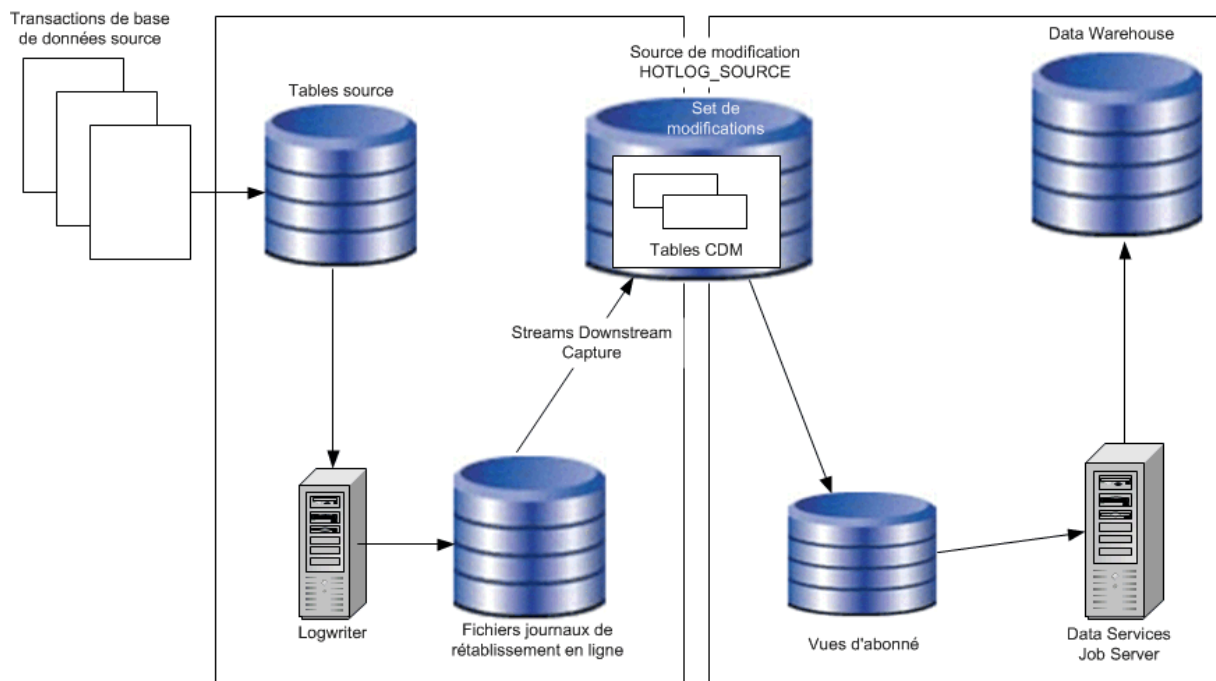
Le diagramme suivant illustre la manière dont les données modifiées sont transférées des tables CDM Oracle vers SAP Data Services en mode synchrone.



Lorsqu'une transaction modifie une table source, des déclencheurs internes capturent les données modifiées et les stockent dans la table CDM correspondante.

34.4.1.2 CDM Oracle en mode HotLog asynchrone

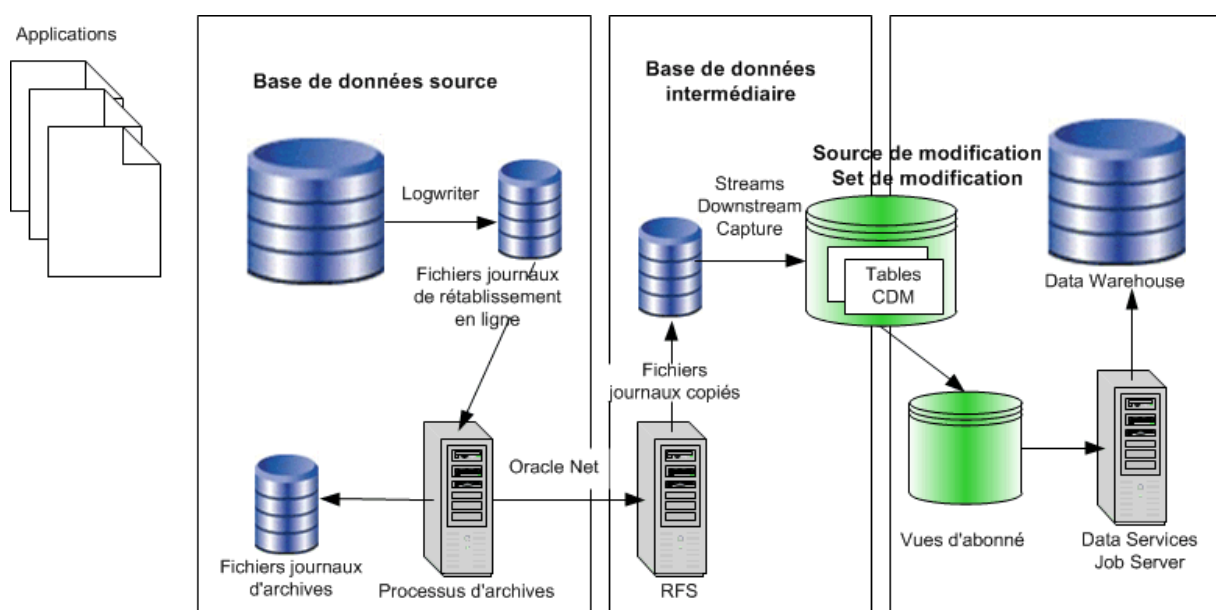
Le diagramme suivant illustre la manière dont les données modifiées sont transférées des tables CDM Oracle vers SAP Data Services en mode HotLog asynchrone.



Lorsqu'une transaction modifie une table source, Logwriter enregistre les modifications dans les fichiers log redo en ligne. Les processus des flux Oracle remplissent automatiquement les tables CDM lors de la validation des transactions.

34.4.1.3 CDM Oracle en mode AutoLog asynchrone

Le diagramme suivant illustre la manière dont les données modifiées sont transférées des tables CDM Oracle vers SAP Data Services en mode AutoLog asynchrone.



Lorsque le journal active la base de données source, Oracle archive le fichier redo log et copie les fichiers redo log en ligne dans la base de données intermédiaire. Les processus des flux Oracle remplissent les tables CDM à partir des fichiers journaux copiés.

i Remarque

Le processus d'archivage Oracle requiert une connectivité ininterrompue via Oracle Net pour envoyer les fichiers redo log au serveur de fichiers distant.

34.4.2 Configuration de la CDM Oracle

Configuration système requise pour que le serveur de base de données source Oracle puisse suivre les modifications.

La table suivante répertorie les configurations système requises sur votre serveur de base de données source Oracle pour le suivi des modifications.

Conditions requises	Notes
Installez les packages CDM d'Oracle.	Ces packages sont installés par défaut. Toutefois, si un package CDM doit être réinstallé, ouvrez le répertoire <code>Admin</code> d'Oracle, puis recherchez et exécutez le script SQL d'Oracle <code>initcdc.sql</code> . • La CDM synchrone est disponible avec Oracle Standard Edition et Enterprise Edition. • La CDM asynchrone est disponible avec Oracle Enterprise Edition uniquement.
Activez Java.	
Définissez les droits du propriétaire de la table source	Pour que les tables CDM puissent être créées, purgées et supprimées si nécessaire.
Activez les droits pour les propriétaires de la banque de données	Les droits comprennent les éléments suivants : • Droits SELECT pour tables CDM • SELECT_CATALOG_ROLE • EXECUTE_CATALOG_ROLE
	G
Activez les déclencheurs système d'Oracle pour la CDM synchrone	

Conditions requises	Notes
Pour la CDM AutoLog asynchrone	• L'administrateur de la base de données source doit créer un dictionnaire de données LogMiner pour permettre aux services de transport de journal d'envoyer ce dictionnaire de données à la base de données intermédiaire. Oracle met à jour automatiquement le dictionnaire de données avec toutes les opérations DDL des tables source qui se produisent durant la CDM pour conserver la cohérence entre les tables intermédiaires et les tables source. • L'administrateur de la base de données source doit également obtenir la valeur du SCN de la version du dictionnaire de données. Si vous souhaitez utiliser Designer pour créer des tables CDM, vous devez indiquer le SCN dans l'assistant. • L'éditeur (habituellement l'administrateur de la base de données source) doit configurer les services de transport de journal pour copier les fichiers redo log du système de la base de données source vers le système de la base de données intermédiaire et enregistrer automatiquement les fichiers redo log.

Informations associées

[Création de tables CDM Oracle dans le logiciel \[page 790\]](#)

34.4.3 Création d'une banque de données CDM pour Oracle

Pour accéder aux tables CDM, créez une banque de données CDM à l'aide de Designer. Une banque de données CDM est une banque de données en lecture seule qui peut uniquement accéder aux tables. Comme pour les autres banques de données, vous pouvez créer, modifier une banque de données CDM et y accéder depuis l'onglet [Banques de données](#) de la bibliothèque d'objets.

1. Créez une banque de données de base de données avec l'option [Type de base de données](#) définie sur [Oracle](#).
2. Cochez la case [CDM](#).
3. Sélectionnez une [version Oracle](#).

Designer permet uniquement de sélectionner les versions d'Oracle qui prennent en charge les packages CDM.

4. Indiquez le nom de votre base de données intermédiaire (la base de données source des modifications où sont publiées les données modifiées) dans [Nom de la connexion](#).
5. Saisissez l'[Utilisateur](#) et le [Mot de passe](#) de la base de données intermédiaire, puis cliquez sur [OK](#).

Vous pouvez utiliser cette banque de données pour parcourir et importer les tables CDM.

34.4.4 Importation de données CDC dans des tables

Créez d'abord une table CDC dans Oracle pour chaque table source que vous souhaitez lire. Vous pouvez ensuite importer cette table CDC à l'aide de SAP Data Services. Créez des tables CDC de l'une des façons suivantes :

- Utilisez un utilitaire Oracle
- Utilisez Designer Data Services

34.4.4.1 Utilisation des tables CDM existantes

1. Importez une table CDM Oracle en cliquant avec le bouton droit sur le nom de la banque de données CDM dans la bibliothèque d'objets, puis en sélectionnant [Ouvrir](#), [Importer par nom](#) ou [Rechercher](#).

Si vous sélectionnez Ouvrir, vous pouvez parcourir la banque de données pour rechercher les tables CDM existantes à l'aide de l'Explorateur de banque de données.

2. Quand vous avez trouvé la table à importer, cliquez dessus avec le bouton droit et sélectionnez [Importer](#).

34.4.4.2 Création de tables CDM Oracle dans le logiciel

Le progiciel donne la possibilité de créer des tables CDM Oracle pour tous les modes de publication :

- CDM synchrone
 - CDM HotLog asynchrone
 - CDM AutoLog asynchrone
1. Dans la bibliothèque d'objets, cliquez avec le bouton droit sur une banque de données CDM et sélectionnez [Ouvrir](#).
 2. Dans l'Explorateur de banque de données, cliquez avec le bouton droit sur l'espace blanc dans la section [Métadonnées externes](#), puis sélectionnez [Nouvelle](#).

L'assistant Nouvelle table CDM s'ouvre. L'assistant permet d'ajouter une table CDM.

i Remarque

Si l'Explorateur de banque de données s'ouvre et qu'il n'existe aucune table CDM dans votre banque de données, cet assistant s'ouvre automatiquement.

3. Sélectionnez le mode de publication dans la première page de l'assistant.

Si votre base de données source est Oracle 9i, vous ne pouvez sélectionner que le mode synchrone. Les modes asynchrones sont désactivés.

Si votre base de données source est Oracle 10G, l'assistant sélectionne par défaut le mode HotLog asynchrone.

Si votre base de données source utilise le mode de publication AutoLog asynchrone, sélectionnez AutoLog asynchrone et fournissez les informations de connexion à la base de données source suivantes :

Champ	Description
Nom de la connexion	Nom de la base de données où se trouve la source des modifications. Utilisez le nom du service de la configuration du service Oracle Net.
Nom d'utilisateur	Nom d'utilisateur de l'administrateur de la base de données source.
Mot de passe	Mot de passe de l'utilisateur de la source des modifications.

4. Cliquez sur [Suivant](#). La deuxième page de l'assistant s'affiche.
5. Indiquez les informations de la table source sur la deuxième page de l'assistant.
 - a. Cliquez sur le bouton [Rechercher](#) pour voir une liste de tables externes non CDM disponibles dans cette banque de données. Pour filtrer une recherche, saisissez les valeurs d'un propriétaire et/ou un nom de table. Vous pouvez utiliser un caractère générique (%) pour effectuer une correspondance de modèle pour les valeurs Nom ou Propriétaire.
 - b. (Facultatif) Sélectionnez [Obtenir les pré-images](#) si vous désirez suivre les pré et post-images dans la nouvelle table CDM.
 - c. Cliquez sur un nom dans la liste des tables renvoyées et cliquez sur [Suivant](#) pour créer une table CDM en utilisant la table sélectionnée comme table source.
6. Indiquez le [Propriétaire de la table CDM](#) pour la nouvelle table CDM.

Par défaut, le nom du propriétaire de la nouvelle table CDM est le nom du propriétaire de la banque de données. Le nom du propriétaire de la table source est également affiché dans la zone de liste déroulante des propriétaires de table CDM. Si le nom que vous désirez utiliser n'est pas dans la liste, saisissez un autre nom de propriétaire.
7. Indiquez le [Nom de la table CDM](#) pour la nouvelle table CDM.

Par défaut, le progiciel génère un nom de table selon la convention suivante : CDM__NomTableSource.
8. Toutes les colonnes sont sélectionnées par défaut. Indiquez les colonnes à inclure ou exclure de la table CDM selon une des méthodes suivantes : supprimez la coche de la case près du nom de chaque colonne à exclure ou bien cliquez sur [Désélectionner tout](#) et placez une coche près du nom de chaque colonne à inclure.
9. Pour le mode de publication synchrone :
 - a. Cliquez sur [Terminer](#). Designer se connecte à l'instance Oracle, crée la table CDM sur le serveur Oracle et importe les métadonnées de la table dans le référentiel. Toutes les tables que le progiciel importe via une banque de données CDM contiennent une colonne qui indique quelle opération effectuer pour chaque ligne. Pour une table CDM Oracle, cette colonne s'appelle Operation\$. Outre cette colonne, Oracle en ajoute d'autres en créant une table CDM. Ces colonnes ont toutes un signe dollar (\$) comme suffixe.
 - b. Cliquez sur [OK](#) dans la boîte de dialogue d'information. Celle-ci confirme qu'Oracle a créé une table CDM, puis l'a importée dans le progiciel.
10. Pour le mode de publication asynchrone (HotLog ou AutoLog), cliquez sur [Suivant](#).
11. Pour le mode de publication HotLog asynchrone, indiquez les informations du set de modifications dans la quatrième page de l'assistant.

- a. Si vous désirez ajouter cette table de modifications à un set de modifications existant afin de conserver la cohérence transactionnelle des modifications avec les tables du set de modifications, sélectionnez un nom dans la liste déroulante de [Modifier le nom du set](#). Sinon, vous pouvez créer un set de modifications en saisissant son nom.
 - b. Sélectionnez [Arrêter la capture sur DDL](#) si une erreur DDL se produit et que vous ne voulez pas capturer les données.
 - c. Sélectionnez [Définir la période de rétention](#) pour activer les zones de texte [Date de début](#) et [Date de fin](#).
 - d. Cliquez sur [Terminer](#).
Designer se connecte à l'instance Oracle, crée la table CDM sur le serveur Oracle et importe les métadonnées de la table dans le référentiel du logiciel. Toutes les tables que le progiciel importe via une banque de données CDM contiennent une colonne qui indique quelle opération effectuer pour chaque ligne. Pour une table CDM Oracle, cette colonne s'appelle Operation\$. Outre cette colonne, Oracle en ajoute d'autres en créant une table CDM. Ces colonnes ont toutes un signe dollar (\$) comme suffixe.
12. Pour le mode de publication AutoLog asynchrone, indiquez les informations du set de modifications et de la source de modification dans la quatrième page de l'assistant.
- a. Si vous désirez ajouter cette table de modifications à un set de modifications existant afin de conserver la cohérence transactionnelle des modifications avec les tables du set de modifications, sélectionnez un nom dans la liste déroulante de [Modifier le nom du set](#). Sinon, vous pouvez créer un set de modifications en saisissant son nom.
 - b. Si vous désirez ajouter cette table de modifications à une source de modification existante, sélectionnez un nom dans la liste déroulante de [Modifier le nom de source](#).
 - c. Pour créer une source de modification, saisissez le nom de la source de modification CDM et le nom de la base de données source. Vous pouvez obtenir ce nom dans la table Global_Name de la base de données source, valeur du SCN de la version du dictionnaire de données.
 - d. Sélectionnez [Arrêter la capture sur DDL](#) si une erreur DDL se produit pendant la capture des données et que vous ne voulez pas la poursuivre.
 - e. Sélectionnez [Définir la période de rétention](#) pour activer les zones de texte [Date de début](#) et [Date de fin](#).
 - f. Cliquez sur [Terminer](#).
Designer se connecte à la base de données intermédiaire Oracle, crée la table CDM dans la source de modification et importe les métadonnées de la table dans le référentiel du logiciel. Toutes les tables que le progiciel importe via une banque de données CDM contiennent une colonne qui indique quelle opération effectuer pour chaque ligne. Pour une table CDM Oracle, cette colonne s'appelle Operation\$. Outre cette colonne, Oracle en ajoute d'autres en créant une table CDM. Ces colonnes ont toutes un signe dollar (\$) comme suffixe.

Informations associées

[Utilisation des pré-images \[page 795\]](#)

34.4.5 Affichage d'une table CDM importée

Pour afficher une table CDM importée :

1. Recherchez votre banque de données CDM dans la bibliothèque d'objets.
2. Développez le dossier [Tables](#).
3. Double-cliquez sur un nom de table ou cliquez avec le bouton droit et sélectionnez [Ouvrir](#).

Lorsque le progiciel importe une table CDM, il ajoute également deux colonnes au schéma de la table : DI_SEQUENCE_NUMBER avec le type de données integer et DI_OPERATION_TYPE avec le type de données varchar(1).

Un schéma de table CDM Oracle importée ressemble au suivant :

	Description
CDC_ORDERS	
DI_SEQUENCE_NUMBER	DI-generated integer for sequencing
DI_OPERATION_TYPE	DI-generated row operation type
OPERATION\$	Colonnes de table CDM Oracle
CSCN\$	
COMMIT_TIMESTAMP\$	
RSID\$	
USERNAME\$	
TIMESTAMP\$	
ORDER_ID	Colonnes source Oracle
ORDER_DATE	
CUSTOMER_ID	
ORDER_TOTAL	
SALES_REP_ID	
PROMOTION_ID	

Cet exemple comporte huit colonnes de contrôle ajoutées à la table d'origine.

- Deux sont générées par le progiciel :
 - DI_SEQUENCE_NUMBER - Contient un entier pour le séquençement.
 - DI_OPERATION_TYPE - Contient le type d'opération de la ligne.
- Six colonnes de contrôle Oracle : OPERATION\$, CSCN\$, COMMIT_TIMESTAMP\$, RSID\$, USERNAME\$, TIMESTAMP\$

i Remarque

Les colonnes de contrôle Oracle varient selon les options ayant été sélectionnées à la création de la table CDM. Toutes les colonnes de contrôle Oracle se terminant par un signe dollar (\$).

Informations associées

[Colonne DI_SEQUENCE_NUMBER \[page 794\]](#)

[Colonne DI_OPERATION_TYPE \[page 794\]](#)

34.4.5.1 Colonne DI_SEQUENCE_NUMBER

La colonne `DI_SEQUENCE_NUMBER` commence par zéro au début de chaque extraction. Ce champ est incrémenté d'un chaque fois que le progiciel lit une ligne, sauf s'il rencontre une paire de pré et post-images pour une opération UPDATE. La pré-image et la post-image reçoivent le même numéro de séquence. Cette colonne de séquençement fournit un moyen de rassembler les paires d'images si elles sont séparées en raison de la conception du flux de données.

Informations associées

[Utilisation des pré-images \[page 795\]](#)

34.4.5.2 Colonne DI_OPERATION_TYPE

Les valeurs possibles de la colonne `DI_OPERATION_TYPE` sont :

- I pour INSERT
- D pour DELETE
- B pour une pré-image d'une opération UPDATE
- U pour une post-image d'une opération UPDATE

Lorsque le progiciel lit une ligne d'Oracle, il contrôle les valeurs de la colonne `Operation$` et les traduit en ses propres valeurs dans la colonne `DI_OPERATION_TYPE`.

Operation\$	DI_OPERATION_TYPE
I	I
D	D
UO, UU	B
UN	U

34.4.6 Configuration d'une table source CDM Oracle

Quand vous faites glisser une table de banque de données CDM dans un flux de données, elle devient automatiquement un objet source.

1. Faites glisser une table de banque de données CDM dans un flux de données.
2. Cliquez sur le nom de cet objet source pour ouvrir son éditeur de table source.
3. Cliquez sur l'onglet Options CDM.
4. Indiquez une valeur pour le [nom d'abonnement CDM](#)
Pour en savoir plus, consultez le *Guide de référence*.

34.4.6.1 Utilisation des points de contrôle

Lorsqu'un job de SAP Data Services s'exécute avec les points de contrôle activés, le progiciel utilise le nom d'abonnement de la table source pour lire le plus récent ensemble de lignes ajoutées. Si vous n'activez pas les points de contrôle, le job lit toutes les lignes de la table, ce qui augmente le temps de traitement.

Pour utiliser les points de contrôle, saisissez un nom dans la zone *Nom d'abonnement CDM* de l'Editeur de table source et sélectionnez l'option *Activer le point de contrôle*.

i Remarque

Pour éviter les problèmes de corruption de données, ne réutilisez pas les flux de données qui utilisent les banques de données CDM parce que chaque fois qu'une table source extrait des données, elle utilise le même nom d'abonnement. C'est-à-dire que des jobs identiques, selon le moment où ils sont exécutés, peuvent obtenir des résultats différents et laisser des points de contrôle à différents endroits de la table. Par exemple, lorsque vous migrez des jobs CDM de test en production, un scénario de meilleure pratique consisterait à modifier le nom d'abonnement pour le job de production afin que le job de test, s'il est à nouveau exécuté, n'affecte pas les résultats du job de production.

34.4.6.2 Utilisation des pré-images

Si vous voulez récupérer les pré-images des lignes UPDATE avant que l'opération de mise à jour soit appliquée à la source, le progiciel peut développer la ligne UPDATE en deux lignes : une pour la pré-image de la mise à jour et une pour la post-image de la mise à jour. La pré-image d'une ligne de mise à jour est l'image de la ligne avant qu'elle ne soit modifiée et la post-image d'une ligne de mise à jour fait référence à l'image de la ligne après application de la modification.

Comportement par défaut : un lecteur CDM récupère uniquement les post-images. En ne récupérant pas les pré-images, moins de lignes passent par le moteur, ce qui permet de diminuer le temps d'exécution du job.

Vous pouvez utiliser les pré-images pour :

- Mettre à jour les clés primaires

i Remarque

Dans la plupart des cas, lorsque les tables source sont mises à jour, leurs clés primaires n'ont pas besoin d'être actualisées.

- Effectuer les calculs des modifications entre les données des colonnes
Par exemple, vous pouvez calculer la différence entre le nouveau et l'ancien salaire d'un employé en examinant la différence entre les valeurs des champs de salaire.

34.4.6.2.1 Capturer les pré-images des lignes de mise à jour

1. Au moment de la création d'une table CDM, vérifiez que la table CDM Oracle est également configurée pour récupérer les pré-images complètes.

Si vous créez une table CDM Oracle à l'aide de Designer, vous pouvez sélectionner *Obtenir les pré-images* dans ce but.

2. Sélectionnez *Obtenir les pré-images pour chaque ligne de mise à jour* dans l'éditeur source de la table CDM.

Si la table CDM sous-jacente n'est pas correctement configurée, l'activation de l'option Obtenir les pré-images pour chaque ligne de mise à jour n'a aucun effet.

Après avoir sélectionné l'option *Obtenir les pré-images pour chaque ligne de mise à jour*, le progiciel traite deux lignes lors de chaque mise à jour. Outre l'impact sur les performances de cette augmentation du volume de données, les paires de pré et post-images peuvent être séparées ou perdues selon la conception de votre flux de données. Cela provoque des problèmes d'intégrité des données.

La transformation Map_CDC_Operation peut résoudre les problèmes, mais des résultats indésirables peuvent toujours se produire en raison d'erreurs de programmation. Lors de l'utilisation de fonctions et transformations qui réordonnent, redirigent, éliminent et multiplient le nombre de lignes dans un flux de données (par exemple, en raison de l'utilisation des clauses GROUP BY ou ORDER BY dans une requête) soyez conscients de l'impact possible sur les cibles.

34.4.7 Création d'un flux de données avec une source CDM Oracle

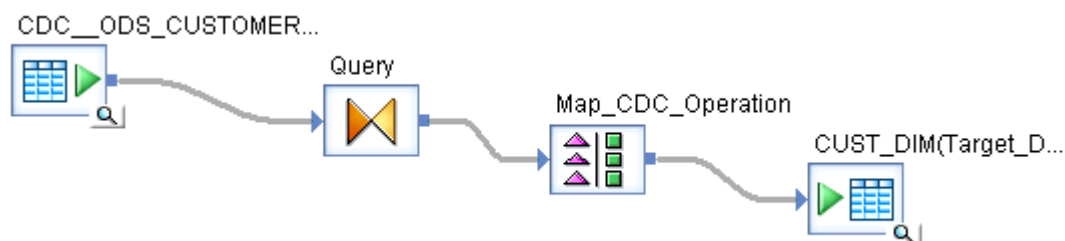
Pour utiliser une source CDM Oracle, vous utilisez une transformation Query pour supprimer les colonnes de contrôle Oracle et la transformation Map_CDC_Operation pour interpréter les colonnes de contrôle et appliquer les actions appropriées.

1. Dans le volet *Bibliothèque d'objets Designer*, faites glisser la table CDM Oracle, ainsi que les transformations Query et Map_CDC_Operation vers la zone de l'espace de travail du flux de données.

i Remarque

Un flux de données ne peut contenir qu'une seule source CDM.

2. Configurez la table CDM.
3. Ajoutez la table cible appropriée et connectez les objets.



4. Dans le volet *Zone de projet*, double-cliquez sur une transformation. L'*Editeur de requêtes* s'ouvre et affiche la transformation.
5. Mappez les colonnes de contrôle *Data Services* et les colonnes de la table source que vous voulez dans la table cible.

La transformation Map_CDC_Operation utilise les valeurs de la colonne de la zone *Colonne d'opération sur ligne* pour exécuter l'opération appropriée sur la ligne source pour la table cible. Pour une table source

CDM Oracle, la colonne `DI_OPERATION_TYPE` est automatiquement sélectionnée comme Colonne de l'opération sur la ligne.

Les opérations peuvent être `INSERT`, `DELETE` ou `UPDATE`. Par exemple, si l'opération est `DELETE`, la ligne correspondante est supprimée de la table cible.

Informations associées

[Configuration d'une table source CDM Oracle \[page 794\]](#)

34.4.8 Maintaining CDC tables and subscriptions

34.4.8.1 Purge des tables CDM

Purgez régulièrement les tables CDM pour qu'elles ne grossissent pas indéfiniment.

i Remarque

Le progiciel ne propose pas cette fonctionnalité. Consultez la documentation Oracle pour savoir comment purger les données qui ne sont plus utilisées par aucun abonné.

34.4.8.2 Suppression des abonnements ou tables CDM Oracle

La fonction de purge d'Oracle ne purge aucune donnée n'ayant pas été lue par tous les abonnements. En conséquence, une bonne pratique consiste à supprimer tous les abonnements qui ne sont plus nécessaires. Vous pouvez supprimer les tables CDM Oracle et leurs abonnements depuis la fenêtre *Explorateur de banque de données* de Designer.

1. Dans le volet *Bibliothèque d'objets locale*, cliquez avec le bouton droit de la souris sur une banque de données CDM, puis sur *Ouvrir*.
2. Dans la fenêtre *Explorateur de banque de données*, cliquez sur *Métadonnées du référentiel*.
3. Cliquez avec le bouton droit sur une table, pointez sur *Maintenance CDM* et suivez une de ces méthodes :
 - Pointez sur *Annuler l'abonnement*.
 - *Annuler l'abonnement* ouvre la liste des abonnements créés dans le logiciel pour la table sélectionnée. Les abonnements Oracle sont associés à ces noms d'abonnement. Sélectionnez chaque nom d'abonnement pour le supprimer d'Oracle et du référentiel.
 - Cliquez sur *Supprimer la table*.
 - Cette option supprime la table CDM dans Oracle et dans le référentiel.

34.4.9 Restrictions

Les restrictions suivantes s'appliquent à l'utilisation de la CDM avec les sources Oracle :

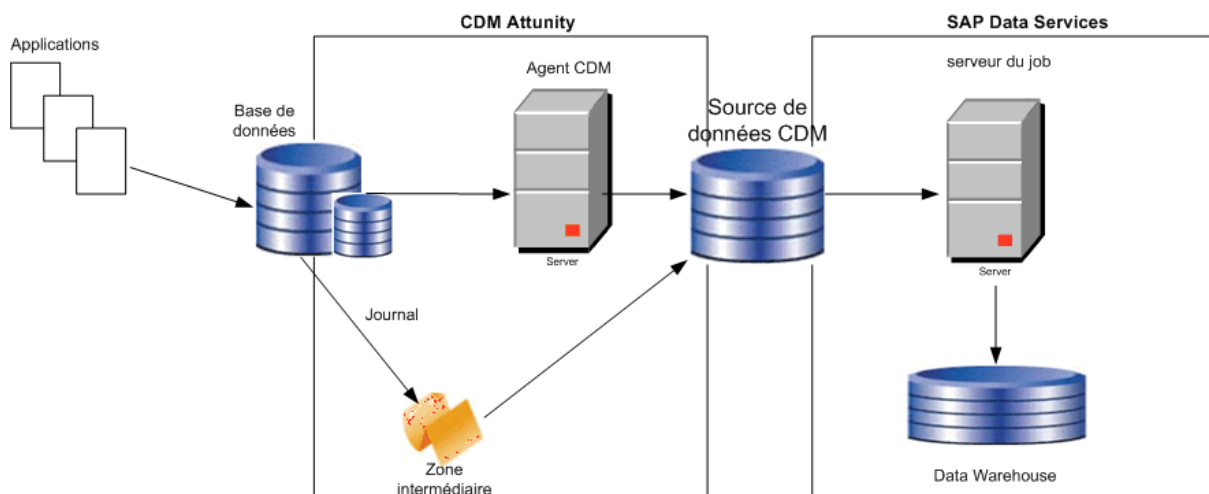
- Il est impossible d'utiliser les transformations et fonctions suivantes pour une table source importée avec une banque de données CDM en raison de l'existence des colonnes générées par SAP Data Services pour les tables CDM. Le progiciel ne peut pas comparer ces colonnes ou rechercher leur contenu.
 - Transformations `Table_Comparison`, `Key_Generation` et `SQL`
 - Toutes les fonctions de base de données comme `lookup`, `lookup_ext`, `key_generation`, `sql` et `total_rows`
- Une seule source CDM peut être créée dans un flux de données.
- La CDM Oracle capture les instructions DML, y compris `INSERT`, `DELETE` et `UPDATE`. Cependant, la CDM Oracle ne prend pas en charge les opérations suivantes parce qu'elles désactivent tous les déclencheurs de base de données :
 - Les instructions `INSERT` en chemin direct
 - L'instruction `multi_table_insert` en mode DML parallèle
- Si vous utilisez les points de contrôle et exécutez votre job en mode de récupération, le job récupéré commencera à examiner le job au début de la table CDM. Les points de contrôle sont ignorés.

34.5 Utilisation de la CDM avec les sources d'un ordinateur central Attunity

Si votre environnement doit garder à jour de grandes quantités de données, la fonctionnalité CDM d'ordinateur central est une solution simple pour limiter le nombre de lignes devant être lues régulièrement.

Si votre environnement doit garder à jour de grandes quantités de données, la fonctionnalité CDM d'ordinateur central est une solution simple pour limiter le nombre de lignes devant être lues régulièrement. Une source qui lit uniquement les opérations les plus récentes (`INSERT`, `UPDATE`, `DELETE`) permet de concevoir des chargements de delta plus légers et plus rapides.

SAP Data Services capture les données modifiées sur les sources de données d'un ordinateur central Attunity et les applique à un système cible. Le diagramme suivant illustre le chemin emprunté par les données, de la CDM Attunity à SAP Data Services.



- L'agent CDM d'Attunity surveille dans le journal de la base de données les modifications de certaines tables. Après la première demande de capture des modifications, l'agent CDM stocke un contexte qu'il utilise comme marqueur pour ne pas capturer à nouveau les modifications antérieures à ce dernier.
 - L'agent CDM envoie les données modifiées à une zone intermédiaire facultative. Les avantages d'une zone intermédiaire sont :
 - Une seule analyse de journal peut extraire les modifications de plusieurs tables. Sans zone intermédiaire, plusieurs analyses de journal, une pour chaque table modifiée, sont nécessaires pour extraire les modifications.
 - Extrait uniquement les modifications validées, ce qui représente un traitement moindre que d'extraire chaque modification. Le traitement est également moindre lors de la récupération d'un job échoué parce que le processus de récupération ne nécessite pas de renvoyer les modifications non validées.
- Toutefois, une zone intermédiaire implique le stockage et le traitement supplémentaires de données inutiles. Consultez la documentation CDM Attunity pour en savoir plus.
- La CDM Attunity Connect envoie les modifications aux sources de données CDM via lesquelles le progiciel peut accéder aux modifications à l'aide de connexions standard ODBC ou JDBC.

34.5.1 Configuration de la CDM Attunity

Si vous utilisez actuellement Attunity comme connexion à SAP Data Services pour extraire les données des sources de l'ordinateur central, créez une source de données CDM Attunity dans Attunity Studio. Les étapes suivantes résument la procédure d'utilisation de l'assistant d'Attunity Studio pour créer une source de données CDM.

- Indiquez votre source de données.
- Selon votre source de données, choisissez une des méthodes suivantes pour capturer les modifications et indiquer l'emplacement du journal :
 - VSAM sous CICS – Par flux de journal CICS
 - DB2 sur les plate-formes OS/390 et z/OS – Par journal DB2
 - DB2 sous OS/400 – Par journal DB400
 - DISAM sous Windows – Par journal

Pour obtenir une liste complète des sources de données prises en charge, consultez le document CDM Attunity Connect.

- Sélectionnez un nom pour votre agent CDM.
- Indiquez si vous voulez capturer les pré-images des opérations de mise à jour. Si vous n'indiquez pas cette option dans Attunity Studio, vous ne capturerez pas les pré-images, même si vous sélectionnez l'option [Obtenir une pré-image pour chaque ligne de mise à jour](#).
- Sélectionnez les tables pour lesquelles il faut surveiller les modifications.

L'assistant d'Attunity Studio génère les composants suivants que vous devez indiquer dans l'Editeur de banque de données lorsque vous définissez une banque de données CDM Attunity :

- Un nom de source de données CDM que vous indiquez dans l'option [Source de données](#). Attunity génère par défaut la source de données CDM sur le même ordinateur que l'agent CDM. Vous avez la possibilité de placer la source de données CDM sur le client (même ordinateur que SAP Data Services). Obtenez le nom de cet ordinateur pour l'indiquer dans l'option [Emplacement de l'hôte](#).
- Un espace de travail pour que l'agent CDM gère la file d'attente des événements de capture des modifications. Vous indiquez le nom de l'espace de travail dans l'option [Espace de travail Attunity](#).

Pour en savoir plus, voir la section Configuration de la CDM dans *Attunity Connect : la solution de capture des données modifiées*.

34.5.2 Configuration du progiciel pour la CDM sur les sources d'ordinateur central

Pour utiliser SAP Data Services afin de lire et de charger les données modifiées depuis des sources d'ordinateur central en utilisant Attunity, appliquez les procédures suivantes dans Designer :

- Créez une banque de données CDM pour Attunity
- Importez les métadonnées pour les tables d'Attunity
- Configurez une source CDM d'ordinateur central
- Créez des jobs en temps réel utilisant des métadonnées

34.5.2.1 Création des banques de données CDM

L'option de banque de données CDM est disponible pour toutes les interfaces d'ordinateur central pour SAP Data Services.

Informations associées

[Qu'est-ce qu'une banque de données \[page 82\]](#)

34.5.2.1.1 Création d'une banque de données CDM pour Attunity

1. Ouvrez l'Editeur de banque de données.
2. Saisissez un nom pour la banque de données.
3. Dans la zone *Type de banque de données*, sélectionnez Base de données.
4. Dans la zone *Type de base de données*, sélectionnez Attunity_Connector.
5. Cochez la case *Activer la CDM* pour activer la fonctionnalité de CDM. Vous pouvez activer la CDM pour les sources de données suivantes : Pour obtenir la liste actuelle des sources de données, consultez le site Web d'Attunity.
 - VSAM sous CICS
 - DB2 UDB pour z/OS
 - DB2 UDB pour OS/400
6. Dans la zone *Source de données*, indiquez le nom de la source de données CDM Attunity. Vous pouvez indiquer plusieurs sources de données pour une banque de données, mais vous ne pouvez pas joindre deux tables CDM. Vous pouvez éventuellement indiquer plusieurs sources de données dans une banque de données Attunity pour faciliter la gestion. Si vous pouvez accéder à l'ensemble des tables CDM via une banque de données Attunity, il est plus facile de créer une banque de données, de saisir les informations de connexion une seule fois et d'importer les tables.

Si vous faites une liste de plusieurs noms de source de données pour une banque de données du connecteur Attunity, assurez-vous de respecter les conditions suivantes :

- N'indiquez pas des sources de données standard avec des sources de données CDM dans la même banque de données. Le progiciel importe les données des sources de données Attunity standard différemment des sources de données CDM.
 - Toutes les sources de données Attunity doivent être accessibles avec le même nom d'utilisateur et mot de passe.
 - Toutes les sources de données Attunity doivent utiliser le même espace de travail. Lorsque vous configurez l'accès aux sources de données dans Attunity Studio, utilisez le même nom d'espace de travail pour chaque source de données.
7. Dans la zone *Emplacement de l'hôte*, indiquez le nom de l'hôte qui héberge le démon de la source de données Attunity.
 8. Dans la zone *Port*, indiquez le numéro de port du démon Attunity. La valeur par défaut est 2551.
 9. Indiquez le nom de l'espace de travail du serveur Attunity utilisé par l'agent CDM pour gérer la file d'attente des événements de capture des modifications pour la source de données CDM.
 10. Renseignez le reste de la boîte de dialogue et cliquez sur *OK*.

Vous pouvez maintenant utiliser la nouvelle connexion de banque de données pour importer les tables de métadonnées dans le référentiel actuel.

Une fois enregistrée, cette banque de données devient une banque de données CDM.

34.5.3 Importation des données CDM de l'ordinateur central.

Utilisez une banque de données CDM pour importer des métadonnées de table CDM.

Dans la bibliothèque d'objets, cliquez avec le bouton droit sur le nom de la banque de données et sélectionnez [Ouvrir](#), [Importer par nom](#) ou [Rechercher](#). Pour la maintenance CDM, seules les tables CDM sélectionnées lors de la configuration d'un CDM Attunity sont visibles lorsque vous parcourez les métadonnées externes. Les fonctions et modèles ne sont pas disponibles parce que la banque de données CDM d'Attunity est en lecture seule.

L'opération d'importation de SAP Data Services ajoute les colonnes suivantes à la table d'origine :

Nom de colonne	Type de données	Source de la colonne
DI_SEQUENCE_NUMBER	integer	Généré par SAP Data Services.
DI_OPERATION_TYPE	varchar(1)	Généré par SAP Data Services
Contexte	varchar(128)	Fourni par Attunity Streams
Timestamp	varchar(26)	Fourni par Attunity Streams
TransactionID	varchar(4)	Fourni par Attunity Streams
Opération	varchar(12)	Fourni par Attunity Streams
tableName	varchar(256)	Fourni par Attunity Streams

Informations associées

[Configuration de la CDM Attunity \[page 799\]](#)

34.5.3.1 Colonne DI_SEQUENCE_NUMBER

La colonne `DI_SEQUENCE_NUMBER` commence par zéro au début de chaque extraction. Ce champ est incrémenté d'un chaque fois que le progiciel lit une ligne, sauf s'il rencontre une paire de pré et post-images.

La pré-image et la post-image reçoivent le même numéro de séquence. Cette colonne de séquençement fournit un moyen de rassembler les paires d'images si elles sont séparées en raison de la conception du flux de données.

Vous pouvez configurer Attunity Streams pour récupérer les pré-images des lignes UPDATE avant que le progiciel n'applique l'opération UPDATE à la cible. Notez que si vous ne configurez pas Attunity Streams pour capturer les pré-images dans la base de données, le progiciel éliminera les lignes.

Lors du déroulement du flux de données, si les pré et post-images sont séparées ou réparties en plusieurs lignes (par exemple, en utilisant des clauses GROUP BY ou ORDER BY dans une requête), vous pouvez perdre l'ordre des lignes.

La transformation Map_CDC_Operation permet de restaurer l'ordre d'origine des paires d'images en utilisant la colonne DI_SEQUENCE_NUMBER comme *Colonne de séquençement*.

Informations associées

[Utilisation des pré-images \[page 795\]](#)

34.5.3.2 Colonne DI_OPERATION_TYPE

SAP Data Services génère des valeurs dans la colonne DI_OPERATION_TYPE. Les valeurs valides pour cette colonne sont :

- I pour INSERT
- D pour DELETE
- B pour une pré-image d'une opération UPDATE
- U pour une post-image d'une opération UPDATE

34.5.4 Configuration d'une source CDM d'ordinateur central

Quand vous faites glisser une table de banque de données CDM dans un flux de données, elle devient automatiquement un objet source.

34.5.4.1 Configuration d'une table CDM d'ordinateur central

1. Faites glisser une table de banque de données CDM dans un flux de données.

La table devient automatiquement un objet source.

2. Cliquez sur le nom de cet objet source pour ouvrir son éditeur de table source.
3. Cliquez sur l'onglet Options CDM.
4. Indiquez une valeur pour le *nom d'abonnement CDM*

Pour en savoir plus, consultez le *Guide de référence*.

Informations associées

[Utilisation des points de contrôle de l'ordinateur central \[page 804\]](#)

[Utilisation des pré-images \[page 795\]](#)

34.5.5 Utilisation des points de contrôle de l'ordinateur central

Les agents CDM d'Attunity lisent les sources de l'ordinateur central et chargent les données modifiées dans une zone intermédiaire ou directement dans la source de données CDM. Les lignes de données modifiées s'ajoutent au précédent chargement dans la source de données CDM.

Lorsque vous activez des points de contrôle, un job CDM de SAP Data Services utilise le nom d'abonnement pour lire le plus récent set de lignes ajoutées et pour marquer la fin de la lecture. Si les points de contrôle ne sont pas activés, le job CDM lit toutes les lignes de la source de données CDM d'Attunity et le temps de traitement augmente.

Pour utiliser les points de contrôle, dans l'Editeur de table source, saisissez le *Nom de l'abonnement CDM* et sélectionnez l'option *Activer le point de contrôle*.

Si vous activez les points de contrôle et exécutez votre job CDM en mode de récupération, le job récupéré commence à examiner la source de données CDM au dernier point de contrôle.

i Remarque

Pour éviter les problèmes de corruption de données, ne réutilisez pas les flux de données qui utilisent les banques de données CDM parce que chaque fois qu'une table source extrait des données, elle utilise le même nom d'abonnement. C'est-à-dire que des jobs identiques, selon le moment où ils sont exécutés, peuvent obtenir des résultats différents et laisser des points de contrôle à différents endroits du fichier. Lorsque vous migrez des jobs CDM de test en production, un scénario de meilleures pratiques consiste à modifier le nom d'abonnement pour le job de production. Si le job de test est de nouveau exécuté, cela n'affecte donc pas les résultats du job de production.

34.5.5.1 Utilisation des pré-images de sources d'ordinateur principal

La pré-image d'une ligne de mise à jour correspond à l'image de la ligne avant la modification de cette dernière.

Suivez les étapes ci-dessous pour capturer des lignes de mise à jour de pré-image :

1. Vérifiez qu'Attunity Streams est configuré pour récupérer des pré-images complètes.
2. Sélectionnez l'option *Obtenir les pré-images pour chaque ligne de mise à jour* dans l'éditeur source de la table CDM.

Le progiciel sous-jacent de capture CDM basée sur les journaux doit être configuré correctement, sinon, l'activation de l'option Obtenir les pré-images pour chaque ligne de mise à jour dans le progiciel n'a aucun effet.

Après avoir coché l'option *Obtenir les pré-images pour chaque ligne de mise à jour*, le progiciel traite deux lignes pour chaque mise à jour. Outre l'impact sur les performances de cette augmentation du volume de données, les paires de pré et post-images pourraient être séparées ou perdues selon la conception de votre flux de données, ce qui provoquerait des problèmes d'intégrité des données.

La transformation Map_CDC_Operation peut résoudre les problèmes, mais des résultats indésirables peuvent toujours se produire en raison d'erreurs de programmation. Lors de l'utilisation de fonctions et transformations

qui réordonnent, redirigent, éliminent et multiplient le nombre de lignes dans un flux de données, soyez conscients de l'impact possible sur les cibles.

Informations associées

[Utilisation des pré-images \[page 795\]](#)

34.5.6 Restrictions

Les restrictions suivantes s'appliquent à cette fonctionnalité :

- Il est impossible d'utiliser les transformations et fonctions suivantes pour une table source importée avec une banque de données CDM en raison de l'existence des colonnes générées par SAP Data Services pour les tables CDM. Le progiciel ne peut pas comparer ces colonnes ou rechercher leur contenu.
 - Transformations Table_Comparison, Key_Generation et SQL
 - Toutes les fonctions de base de données comme `lookup`, `lookup_ext`, `key_generation`, `sql` et `total_rows`
- Une seule source CDM peut être créée dans un flux de données.

34.6 Utilisation de la CDM avec SAP Replication Server

Data Services, combiné avec la fonctionnalité de réplication de SAP Replication Server, offre deux mécanismes qui permettent de capturer, transformer et propager des données volumineuses vers un entrepôt de données en temps réel.

L'utilisation combinée de la CDM et de SAP Replication Server est idéale si votre environnement doit maintenir à jour de gros volumes de données avec un minimum d'impact sur les systèmes opérationnels. Les sources CDM SAP Replication Server dans les flux de données lisent uniquement les opérations les plus récentes (INSERT, UPDATE, DELETE), ce qui permet de concevoir des chargements delta plus légers et plus rapides.

Les deux méthodes d'utilisation de SAP Replication Server sont les suivantes :

Méthode	Avantages
Modélisation SAP PowerAMC	Permet une exécution en mode batch pendant les heures creuses pour appliquer les modifications à la base de données cible.

Méthode	Avantages
Fonctions intégrées et workflow continu	Cette méthode permet d'exploiter les fonctionnalités suivantes : • Configuration automatique pour la capture des données en temps réel. • Récupération automatique en cas d'échec d'un job à l'aide de l'option de correction automatique du lecteur CDM. • Fonctionnalités de restauration spécifiques de la file d'attente delta du serveur de réplication. Cette fonctionnalité peut s'avérer utile si la base de données cible est corrompue et que la base de données doit être activée sur le statut actuel après une restauration à partir de la sauvegarde de la base de données cible. • Transition automatique du chargement initial au chargement delta. • Fonctionnalité de correction automatique du lecteur permettant au chargeur de charger les données dans la base de données cible à l'aide de l'option de chargement par lots.

34.6.1 Aperçu de l'utilisation d'un workflow continu et de fonctions

Vous pouvez utiliser SAP Replication Server pour la capture des données modifiées à l'aide d'un workflow continu et de fonctions permettant de capturer les données en temps réel. Cette méthode est une alternative plus simple à l'utilisation de la modélisation SAP PowerAMC.

Le processus est le suivant :

- Data Services se connecte directement à la base de données source à l'aide de son connecteur natif (la bibliothèque OCI pour Oracle, par exemple) pour :
 - Importer les métadonnées des tables source dans son référentiel
 - Effectuer le chargement initial de la base de données cible
- Data Services communique avec l'agent de réplication à l'aide de `ct_lib` pour :
 - Valider les configurations globales requises
 - Marquer les tables source pour réplication
 - Obtenir le mappage des types de données du serveur de réplication sur la base de données source
- Data Services communique avec le serveur de réplication à l'aide de `ct_lib` pour :
 - Configurer la file d'attente cible
 - Configurer les définitions et abonnements de la réplication pour la base de données source
 - Lire les données CDM de la file d'attente

Le processus d'activation de l'environnement est le suivant :

1. Contrôle de l'installation du logiciel requis.
2. Configuration du serveur et de l'agent de réplication pour fonctionner avec Data Services.
3. Dans Data Services, configuration des connexions à la banque de données et importation des tables à surveiller.
4. Dans Data Services, conception du job CDM à l'aide d'un workflow continu et des fonctions intégrées fournies.

34.6.1.1 Configuration requise

Pour permettre la capture des données modifiées à l'aide des fonctions intégrées de SAP Replication Server et Data Services avec un workflow continu, assurez-vous d'avoir installé les composants suivants :

Composant	Fonction	Référence
SAP Replication Server	Capture les données modifiées de la base de données source et les stocke dans des files d'attente sur le serveur de réplication où elles peuvent ensuite être récupérées par un job Data Services.	Voir la documentation en ligne SAP Sybase à l'adresse suivante : http://infocenter.sybase.com/help/index.jsp 
Agent de réplication	Communique entre la base de données source et le serveur de réplication.	Pour en savoir plus sur la base de données source, voir le <i>Guide de la base de données principale de l'agent de réplication</i> (documentation en ligne SAP Sybase).
Data Services Designer	Sert à concevoir le workflow continu pour capturer les données modifiées.	

34.6.1.2 Configuration du serveur et de l'agent de réplication

Lorsque vous utilisez le serveur de réplication et l'agent de réplication pour CDM, suivez la procédure ci-dessous :

En règle générale, il est recommandé d'utiliser une instance dédiée du serveur de réplication/de l'agent de réplication hétérogène pour capturer les données modifiées, car les jobs Data Services gèrent leurs propres configurations. Pour exécuter une configuration automatique, les jobs suspendent/reprennent l'agent de réplication hétérogène et les connexions de rétention de l'opération.

Serveur de réplication

- Définissez le jeu de caractères du serveur de réplication sur UTF8.
- Data Services doit pouvoir se connecter directement au serveur de réplication principal pour gérer les configurations de la file d'attente des connexions de rétention de l'opération, la définition et l'abonnement de la réplication et pour utiliser les données modifiées de la file d'attente.
- L'utilisateur Data Services qui se connecte au serveur de réplication doit disposer des autorisations suivantes :
 - connecter la cible
 - créer un objet
 - souscrire à un abonnement primaire
- Il est recommandé que le fichier journal du serveur de réplication ne soit pas placé dans le même système de fichiers que le fichier de partition, car de gros volumes de données répliquées dans la partition pourraient interférer avec la taille du fichier journal. Vous pouvez modifier l'emplacement du fichier journal comme suit :

- `rs_init` : modifiez le paramètre `rs.rs_rs_errorlog`. Par exemple : `rs.rs_rs_errorlog : /home/user/sample_repserver/sample_rs.log`
- Si vous utilisez le fichier `RUN` pour lancer le serveur de réplication, modifiez-le pour changer l'emplacement du journal. Par exemple : `-E/home/user/sample_repserver/sample_rs.log`
- Sur le serveur de réplication, vous pouvez configurer une partition qui s'étend automatiquement. Par exemple :

```
create auto partition path partition2 on '/home/.../partition2' with auto
expand size = 1024 max size = 10240
```

La partition logique `partition2` est créée dans un répertoire appelé `partition2`. Si besoin, un fichier de partition de 1 024 Mo est créé sous `partition2` jusqu'à ce que la limite soit atteinte. Pour contrôler une partition automatique sur le serveur de réplication :

```
admin auto_part_path
```

Pour en savoir plus sur l'utilisation de cette fonctionnalité, voir le *Manuel de référence du serveur de réplication*.

- Pour accéder à des instructions générales sur le dimensionnement, consultez le dernier guide relatif au dimensionnement du serveur de réplication.

Agent de réplication

- Data Services requiert l'utilisateur et le mot de passe d'administrateur pour se connecter à l'agent de réplication hétérogène.
- L'utilisateur de l'agent de réplication SAP ASE doit être affecté au rôle `rs_replication`.

i Remarque

Data Services modifie la configuration globale et les configurations au niveau des tables requises pour capturer et traiter les données modifiées.

Dépannage

Si la réplication ne se déroule pas comme prévu, vérifiez les points suivants :

- Vérifiez le journal de l'agent de réplication pour déterminer si l'agent est en cours d'exécution et quelle en est la raison s'il ne l'est pas.
- La partition sur le serveur de réplication est pleine. Dans le journal du serveur de réplication, recherchez les messages tels que :

```
SQM_ADD_SEGMENT ('102:1'): Going to wait for a segment to be freed.
Administrative action is needed
```

Pour résoudre ce problème, exécutez la commande suivante sur le serveur de réplication :

```
alter partition <partition name> expand size = <number of segments>
```

Pour vérifier l'espace disque actuel :

```
admin disk_space
```

- La mémoire est insuffisante sur le serveur de réplication. Dans le journal du serveur de réplication, recherchez les messages tels que :

```
Replication Agent for rao21.orcl1 is sleeping due to memory controls (EXEC threshold '90 percent') being triggered.
```

Pour résoudre ce problème, exécutez la commande suivante sur le serveur de réplication :

```
configure replication server set memory_limit to '<N megabytes>' :
```

Pour vérifier l'utilisation de la mémoire actuelle :

```
admin stats, mem_in_use
```

- Si vous utilisez la base de données système du serveur de réplication intégré (ERSSD) comme source et que vous vous apercevez qu'elle consomme une quantité importante de mémoire physique, vous pouvez limiter la taille de la mémoire cache à l'aide du commutateur `-ch SQL Anywhere`.
 1. Fermez l'instance du serveur de réplication.
 2. Ouvrez le fichier `<rs_instance_name>.cfg` de l'instance du serveur de réplication (SAMPLE_RS.cfg, par exemple) et ajoutez le paramètre `erssd_start_cmd` pour indiquer la commande à utiliser pour lancer la base de données ERSSD (le paramètre `erssd_start_cmd` vous permet de spécifier une commande différente pour lancer ERSSD). L'exemple suivant montre comment modifier la configuration par défaut de la base de données ERSSD SQL Anywhere pour limiter la taille de la mémoire cache à 1 Go pour la base de données ERSSD existante.

```
erssd_start_cmd=$SYBASE/$SYBASE_REP/ASA16/bin/dbspawn -f -q dbsrv16 -s  
none -ti 0 -x  
"tcpip(PORT=11751;DOBROAD=NO;BLISTENER=NO)" -ch lg -o $SYBASE/$SYBASE_REP/  
/samp_repserver/errorlog/SAMPLE_RS_ERSSD.out $SYBASE/$SYBASE_REP/  
samp_repserver/dbfile/SAMPLE_RS_ERSSD.db
```

3. Redémarrez l'instance du serveur de réplication.

Pour en savoir plus sur le paramètre `erssd_start_cmd`, voir le *Guide de configuration du serveur de réplication*.

34.6.1.3 Création de la banque de données CDM

Cette procédure décrit comment créer une banque de données pour utiliser le serveur de réplication en vue de capturer les données modifiées dans les bases de données source (principales) prises en charge.

Vous devez ajouter une connexion de la banque de données source vers la base de données source (principale) et importer les tables source concernées.

1. Dans l'onglet *Banques de données* de la bibliothèque d'objets, ouvrez l'éditeur de banque de données (cliquez avec le bouton droit de la souris et sélectionnez *Nouveau*).
2. Saisissez un nom pour la banque de données.
3. Pour le *Type de banque de données*, sélectionnez *Base de données*.
4. Pour le *Type de base de données*, sélectionnez la base de données source.

5. Dans la zone de liste déroulante en regard du type de base de données, sélectionnez *Serveur de réplication CDM*.
6. Sélectionnez la *Version de la base de données*.
7. Saisissez le *Nom du serveur de la base de données*.
8. Saisissez un *Nom d'utilisateur* et un *Mot de passe* de base de données.
9. Cliquez sur *Avancé* et saisissez les valeurs CDM suivantes :

Option	Description
Nom du serveur de réplication	Le nom du serveur de réplication doit correspondre à celui défini dans le fichier d'interface <code>sql.ini</code> . De plus, le nom doit correspondre au nom du serveur de réplication auquel l'agent de réplication est connecté.
Nom d'utilisateur du serveur de réplication	L'utilisateur du serveur de réplication doit être autorisé à créer un objet, souscrire à un abonnement primaire et connecter la cible.
Mot de passe du serveur de réplication	Le mot de passe pour accéder au système.
Période de rétention	Définissez la valeur correspondant à la propriété <code>save_interval</code> de connexion au serveur de réplication en minutes. Par exemple, si l'intervalle de sauvegarde de votre base de données cible est de 24 heures, définissez la période de rétention sur $24 \times 60 = 1\,440$ pour définir le nombre équivalent de minutes pendant lesquelles conserver les données modifiées au cours des dernières 24 heures.
Nom de l'agent de réplication	Le nom de l'agent de réplication doit correspondre à celui défini dans le fichier d'interface <code>sql.ini</code> .
Nom d'utilisateur de l'agent de réplication	L'utilisateur de l'agent de réplication doit être un utilisateur de type "sa". S'il utilise la base de données source SAP ASE, l'utilisateur de l'agent de réplication doit être affecté au rôle "replication_role".
Mot de passe de l'agent de réplication	Le mot de passe pour accéder au système.

10. Cliquez sur *Appliquer* ou sur *OK* pour enregistrer la banque de données.

34.6.1.4 Création du flux de données

Cette procédure décrit comment créer un flux de données en vue de capturer les données modifiées à l'aide des fonctions intégrées et d'un workflow continu.

- Vous devez ajouter une connexion de la banque de données source vers la base de données source (principale) et importer les tables source concernées.
 - Vous devez ajouter une banque de données CDM et importer la ou les tables correspondantes.
1. Ajoutez un nouveau job.
 2. Ajoutez une variable globale intitulée, par exemple, `$initialload` avec une valeur par défaut de 0.
 3. Ajoutez une expression conditionnelle (intitulée `initial_load`, par exemple) avec les paramètres suivants :
 - a. If `$initialload = 1`
 - b. Ajoutez un script intitulé, par exemple, `initialloadCCDmarker`.

- c. Dans le script, ajoutez un appel à la fonction prédéfinie `begin_initial_load()`.
 - d. Ajoutez un flux de données intitulé, par exemple, `initial_load`.
 - e. Ouvrez le flux de données et ajoutez la table source initiale comme source.
 - f. Complétez le flux de données en ajoutant une cible (les transformations supplémentaires sont facultatives).
 - g. Ouvrez les options du chargeur et sélectionnez *Supprimer les données de la table avant le chargement*.
4. Retournez dans le job et ajoutez un workflow continu.
 5. Ouvrez le workflow continu.
 - a. Ajoutez un script intitulé, par exemple, `MarkBeginCDCLoad`.
 - b. Dans le script, ajoutez un appel à la fonction prédéfinie `begin_delta_load()`.
 6. Ajoutez un flux de données CDM :
 - a. Ajoutez une table CDM de la banque de données CDM correspondant à la table source utilisée dans le flux de données `initial_load`.
 Dans l'éditeur source, vous pouvez sélectionner *Activer la correction automatique*. Cette option peut être utile lorsque le chargeur peut identifier des enregistrements de la table source dans la base de données cible à l'aide de clés dans la table source pour exécuter des opérations de correction automatique générées par le lecteur.
 - a. Ajoutez une transformation `Map_CDC_Operation`.
 - b. Ajoutez une cible et connectez les objets.
 7. Lorsque vous exécutez le job pour la première fois, accédez à l'onglet *Variable globale* et définissez la variable `$initialload` sur 1 pour initialiser (ou réinitialiser) la banque de données cible.

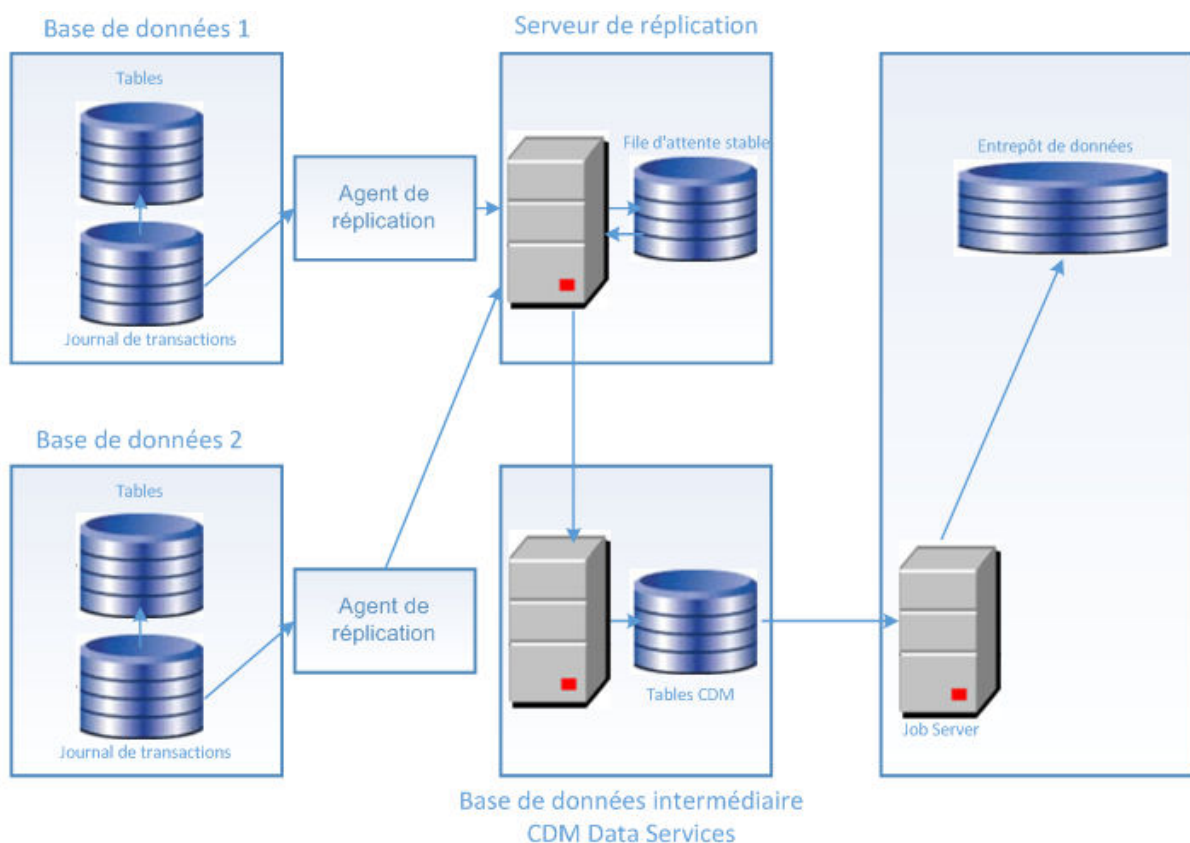
Les données modifiées sont répliquées dans la cible.

Si la base de données cible est corrompue en raison d'une erreur, vous pouvez restaurer les enregistrements modifiés des transactions en souffrance du serveur de réplication. Utilisez la fonction intégrée de Data Services `restore_repserver_cdb_backlogged_transactions`.

34.6.2 Présentation de la méthode de modélisation SAP PowerAMC

Vous pouvez utiliser SAP Replication Server pour la capture des données modifiées à l'aide de la modélisation SAP PowerAMC. Cette méthode constitue une alternative à l'utilisation des fonctions intégrées avec un workflow continu.

À l'aide du modèle de mouvements de données de SAP PowerAMC, Data Services gère l'environnement CDM en créant une base de données intermédiaire CDM ainsi que des définitions et abonnements de réplication SAP Replication Server. Lors du déploiement des définitions et abonnements de réplication ainsi que des définitions de la base de données intermédiaire CDM, SAP Replication Server publie les données modifiées de la table source dans sa table CDM correspondante figurant dans la base de données intermédiaire CDM. La source de la table CDM dans le flux de données Data Services lit les données modifiées les plus récemment à partir du dernier point de contrôle depuis la table CDM et les applique à la cible. Le diagramme suivant illustre le flux des données modifiées du serveur de réplication vers Data Services.



34.6.2.1 Configuration requise

Pour activer la capture des données modifiées à l'aide de SAP Replication Server et de la modélisation SAP PowerAMC, assurez-vous d'avoir installé les composants suivants.

Composant de conception	Fonction	Référence
SAP PowerAMC	Outil de modélisation pour la création des définitions de réplication, des abonnements pour SAP Replication Server et des définitions de la base de données intermédiaire CDM Data Services. Vous avez besoin de cet outil ainsi que de Data Services Designer pour modéliser le système source conformément au flux de données CDM du système cible.	Voir la documentation en ligne SAP Sybase à l'adresse suivante : http://infocenter.sybase.com/help/index.jsp
Designer Data Services	Permet de concevoir le flux de données pour la transformation des données modifiées avant l'application au système cible.	

Composants d'exécution	Fonction	Référence
SAP Replication Server	Capture les données modifiées dans la base de données source et les publie dans la base de données intermédiaire CDM Data Services.	Documentation en ligne SAP Sybase
outil isql	Déploie les définitions de réplication et les abonnements sur le serveur de réplication en activant les tables source des bases de données source pour la réplication et crée les définitions de base de données CDM Data Services dans la base de données intermédiaire.	<i>Guide de l'utilitaire SAP Sybase Adaptive Server Enterprise</i>
SAP ASE	Fractionne les données modifiées CDM Data Services.	Documentation en ligne SAP Sybase
Planification des jobs SAP ASE	Purge les anciennes données modifiées dans la base de données intermédiaire CDM (partie de la gestion de l'environnement CDM).	i Remarque Installez l'agent de planification des jobs dans le cadre de SAP ASE 15.7.1.
Job Server de Data Services	Transforme les données modifiées et les charge dans la cible.	

34.6.2.2 Modélisation du processus de réplication pour la base de données intermédiaire CDM Data Services

PowerAMC automatise le processus de configuration du serveur de réplication SAP Sybase qui inclut le marquage de la table source comme répliquée, la création des définitions et abonnements de réplication et la définition de la base de données intermédiaire CDM Data Services.

i Remarque

La table de la base de données source à répliquer doit comporter une clé primaire.

Les étapes suivantes résument la procédure de modélisation du processus de réplication pour la base de données intermédiaire CDM Data services dans PowerAMC :

1. Supprimez le fichier `SybaseRepServerDataServicesCDC100.xem` de l'emplacement d'installation Data Services `%<DS_COMMON_DIR>%/ext/cdc` dans le dossier de PowerAMC ► [Fichiers de ressources](#) ► [Définitions de modèle étendues](#) ►.

i Remarque

L'extension PowerAMC Data Services fait partie du package Data Services.

2. Créez un modèle de mouvement de données dans PowerAMC en sélectionnant ► [Modèle](#) ► [Nouveau](#) ► [Modèle de mouvement de données](#) ►. Saisissez le *Nom du modèle* et, sous *Extensions*, sélectionnez *Serveur de réplication 15.5* dans la liste déroulante.

3. Liez-y l'*extension SAP Data Services CDC 14.2.0*. Pour ce faire, accédez à ► *Modèle* ► *Extensions* ► *Liste des extensions* ►. Cliquez sur l'icône de lien d'une extension et sélectionnez l'*extension SAP Data Services CDC 14.2.0*.

i Remarque

Cette extension fonctionne uniquement avec la modélisation de mouvements de données de serveur de réplication 15.7.

4. Exécutez l'*Assistant de réplication* de PowerAMC en sélectionnant ► *Outils* ► *Assistant de réplication* ►. Le processus du serveur de réplication, le modèle de données physique source et le modèle de données physique sont créés.

i Remarque

L'Assistant de réplication exécute une rétroconception des bases de données source pour importer les objets de base de données dans le modèle de mouvements de données. Ne choisissez que des tables et les propriétés de table suivantes : Clés primaires, Clés étrangères, Clés alternatives et index pour créer le modèle de base de données physique répliquée.

5. Dans la fenêtre *Sélection de publication*, dans la boîte de dialogue *Assistant de réplication*, sélectionnez *Créer ou sélectionner une publication individuelle pour regrouper les tables répliquées* puis sélectionnez le type de réplication *Définitions de réplication*.

i Remarque

Lorsque l'assistant de réplication se termine, PowerAMC crée les définitions de réplication par défaut (appelées Articles) et le modèle de données physiques pour la base de données distante avec toutes les tables sélectionnées dans le modèle de données physiques source et les colonnes et propriétés de table correspondantes.

6. Dans l'*Assistant de réplication*, dans l'écran *Sélection à distance du modèle de données physiques*, sélectionnez le *SGBD Sybase AS Enterprise 15.7*.
7. Enregistrez le modèle de mouvements de données et le modèle de données physiques.
8. Sélectionnez l'icône de processus de réplication dans le diagramme du modèle de mouvement de données et sélectionnez *Ajouter des définitions CDM Data Services* dans le menu contextuel pour exécuter l'extension Data Services et créer ce qui suit :
 - a. les définitions de réplication CDM
 - b. les tables intermédiaires CDM
 - c. les tables système CDM (DS_CDC_TABLES_MAP, DS_CHANGE_RETENTION et DS_PURGE_TIMESTAMP)
 - d. l'utilisateur SAPDS

i Remarque

La sélection automatique de l'option *Ajouter des définitions CDM Data Services* modifie le propriétaire des tables répliquées, la procédure stockée de purge, la table intermédiaire CDM et les tables système CDM par SAPDS. Cela définit le nom de la table intermédiaire CDM comme étant le même que le celui de la table distante de l'article avec le suffixe "_SAPCDC". Cela crée ou remplace les chaînes de fonction : rs_delete, rs_insert, rs_update et rs_get_textptr pour les colonnes Blob. La table

intermédiaire CDM définit les colonnes supplémentaires s'ajoutant aux colonnes RemoteTable pour stocker les variables système du serveur de réplication :

Chaîne de fonction	Son rôle
DS_SRCDB	stocke la variable système « rs_origin_db »
DS_SRCDB_SRV	stocke la variable système « rs_origin_ds »
DS_SRCDB_ID	stocke la variable système « rs_origin »
DS_SRCDB_COMMIT_TIME	stocke la variable système « rs_origin_commit_time »
DS_SRCDB_ORIGIN_QID	stocke « rs_origin_qid »
DS_SRCDB_OP_TYPE	stocke les types d'opération de modification « I » : insérer, « B » : pré-image de mise à jour, « U » : post-image de mise à jour, « D » : supprimer.
DS_CDCDB_PUBLISH_TIME	stocke la date et l'heure de publication de la base de données CDM lorsque l'enregistrement est inséré par le serveur de réplication.

La commande [Ajouter des définitions CDM Data Services](#) définit également les attributs de la « Table intermédiaire CDM » avec le nom de la table intermédiaire CDM pour chaque article.

De plus, l'extension [SAP Data Services CDC 14.2.0](#) ajoute deux attributs étendus, "Période de rétention" et "Table intermédiaire CDM", à l'objet Article. La valeur par défaut des attributs étendus "Période de rétention" est de 14 jours. Cette valeur peut être modifiée.

34.6.2.3 Configuration de l'environnement

Le processus suivant décrit comment configurer Data Services et le serveur de réplication pour capturer les données modifiées à l'aide de la méthode de modélisation SAP PowerAMC.

1. Créez la base de données intermédiaire CDM Data Services sur le serveur SAP ASE.
2. Configurez le serveur de réplication pour ajouter les bases de données source et la base de données intermédiaire CDM.
3. Configurez l'environnement d'exécution de la base de données intermédiaire CDM Data Services sur le serveur SAP ASE.
4. Déployez les définitions et abonnements de réplication sur le serveur de réplication.
5. Configurez l'environnement d'exécution des bases de données principales.

34.6.2.3.1 Création de la base de données intermédiaire CDM Data Services sur le serveur SAP ASE

La procédure suivante décrit comment créer une base de données intermédiaire CDM Data Services sur le serveur SAP ASE.

1. Connectez-vous en tant qu'administrateur de base de données et créez une base de données correspondant au nom de la base de données intermédiaire CDM Data Services, conformément au modèle de mouvements de données de PowerAMC.
2. Créez un utilisateur de connexion SAPDS. Verrouillez cet utilisateur pour empêcher la connexion.

i Remarque

Cet utilisateur est destiné à la CDM Data Services et ne doit être utilisé à aucune autre fin. Tous les objets de la base de données intermédiaire Data Services sont la propriété de l'utilisateur SAPDS.

3. Sélectionnez le modèle de données physiques de la base de données intermédiaire Data Services dans le modèle de mouvements de données de PowerAMC et cliquez avec le bouton droit de la souris sur le contexte pour sélectionner *Générer une base de données* et générer le script DLL de la base de données CDM.

i Remarque

Cette opération génère automatiquement les DLL de tous les objets de la base de données nécessaires à la base de données intermédiaire CDM Data Services.

4. Vérifiez le script de base de données CDM généré sous forme de fichier.
5. Déployez le script DDL à l'aide de l'outil isql. Saisissez le code :

```
isql-Usa-Ppwd -Spds -Dpdb -i DS_CDC_staging_DDL_script.sql-  
oDS_CDC_staging_DDL_script.out
```

où :

- **<sa>** est le nom de connexion de l'administrateur système sur le serveur de données intermédiaire CDM Data Services.
- **<pwd>** est le mot de passe correspondant au nom de connexion de l'administrateur système.
- **<pds>** est le nom du serveur de données intermédiaire CDM Data Services.
- **<pdb>** est le nom de la base de données intermédiaire CDM Data Services.

34.6.2.3.2 Configuration du serveur de réplication pour ajouter les bases de données source et la base de données intermédiaire CDM

La procédure suivante décrit comment configurer le serveur de réplication pour ajouter une base de données source et une base de données intermédiaire CDM.

1. Ajoutez la base de données source et la base de données intermédiaire CDM au fichier d'interface de l'hôte du serveur de réplication.

2. Exécutez rs_init pour ajouter la base de données source et la base de données intermédiaire CDM.

34.6.2.3.3 Configuration de la base de données intermédiaire CDM Data Services sur le serveur SAP ASE

La procédure suivante décrit comment configurer la base de données intermédiaire CDM Data Services sur le serveur SAP ASE.

1. Sélectionnez le modèle de données physiques de la base de données intermédiaire Data Services dans le modèle de mouvements de données de PowerAMC et cliquez avec le bouton droit de la souris sur le contexte pour sélectionner *Générer un script*.
2. Sélectionnez l'extension *SAP Data Services CDC 14.2.0* dans la boîte de dialogue. Un fichier de script est généré avec le suffixe `_DataServices_CDC_DB_Runtime_Script.sql`. Ce fichier crée l'environnement d'exécution de la base de données intermédiaire CDM Data Services et fournit les fonctionnalités suivantes :
 - a. Vérification de l'ajout de la base de données intermédiaire au serveur de réplication via l'utilitaire `rs_init`.
 - b. Vérification de l'exécution du script de création de la base de données intermédiaire CDM Data Services.
 - c. Création de l'utilisateur de maintenance du serveur de réplication défini dans le modèle de mouvements de données (si cet utilisateur n'a pas encore été créé). Vous pouvez changer de mot de passe après la création à des fins de protection contre l'accès non autorisé.
 - d. Ajout de l'utilisateur de maintenance à la base de données.
 - e. Ajout du nom de la table cible et du nom de la table intermédiaire CDM à la table `DS_CDC_TABLES_MAP` pour chaque définition de réplication.
 - f. Ajout du nom de la table intermédiaire CDM et de la période de rétention à la table `DS_CHANGE_RETENTION` pour chaque définition de réplication.
 - g. Octroi de droits de sélection, d'insertion, de mise à jour et de suppression sur les tables intermédiaires CDM.
 - h. Octroi de droit d'exécution sur la procédure stockée `PURGE_OLD_CHANGE_DATA`.
 - i. Octroi à l'utilisateur de maintenance du serveur de réplication du droit d'exécution du job planifié `PURGE_OLD_CHANGE_DATA` dans la base de données du planificateur des jobs SAP ASE.
 - j. Création d'un job planifié `PURGE_OLD_CHANGE_DATA` dans la base de données du planificateur des jobs SAP ASE. Le nom du job est le suivant : `SAPDS_<nom de la base de données intermédiaire CDM>_PURGE_CHANGE_OLD_DATA`.
3. Déployez le script DDL à l'aide de l'outil isql.

L'utilisateur de la base de données qui se connecte doit être un superutilisateur. Saisissez le code :

```
isql -Usa -Ppwd -Spds -Dpdb -i DS_CDC_staging_Runtime script.sql -o DS_CDC_staging_Runtime_script.out
```

où :

- `<sa>` est le nom de connexion de l'administrateur système sur le serveur de données intermédiaire CDM Data Services.
- `<pwd>` est le mot de passe correspondant au nom de connexion de l'administrateur système.
- `<pds>` est le nom du serveur de données intermédiaire CDM Data Services.

- `<pdb>` est le nom de la base de données intermédiaire CDM Data Services.

34.6.2.3.4 Déploiement des définitions et abonnements de réplication sur le serveur de réplication

La procédure suivante décrit comment déployer des définitions et abonnements de réplication sur le serveur de réplication.

1. Sélectionnez le processus de réplication dans le modèle de mouvements de données de PowerAMC et cliquez avec le bouton droit de la souris sur le contexte pour sélectionner *Générer un script*.
2. Définissez la valeur *Créer une connexion* sur *false* dans la boîte de dialogue et cliquez sur *OK* pour générer les définitions de réplication, les abonnements des définitions des chaînes de fonction CDM et le fichier de script RCL du serveur de réplication.

i Remarque

Il est recommandé de créer une connexion à la base de données à l'aide de l'outil `rs_init`.

3. Déployez le script RCL à l'aide de l'outil `isql` sur le serveur de réplication que vous avez configuré pour la base de données intermédiaire CDM et les bases de données source.

L'utilisateur du serveur de réplication qui se connecte doit être un superutilisateur. Saisissez le code :

```
isql -Usa -Ppwd -Spds -Dpbd -i RCL_script.sql -o RCL_script.out
```

où :

- `<sa>` est le nom de connexion de l'administrateur système sur le serveur de données intermédiaire CDM Data Services.
- `<pwd>` est le mot de passe correspondant au nom de connexion de l'administrateur système.
- `<pds>` est le nom du serveur de données intermédiaire CDM Data Services.
- `<pdb>` est le nom de la base de données intermédiaire CDM Data Services.

34.6.2.3.5 Configuration de l'environnement de la base de données principale

La procédure suivante décrit comment configurer l'environnement de la base de données principale.

1. Sélectionnez la base de données principale dans le modèle de mouvements de données de PowerAMC, cliquez avec le bouton droit de la souris sur le contexte pour sélectionner *Générer un script* et sélectionnez *Serveur de réplication 15.7*. Trois fichiers de script sont générés.
2. Sélectionnez le fichier de script contenant la commande de procédure stockée `sp_setreptable`.
3. Vérifiez le script pour vous assurer que le paramètre `owner_on` est défini pour les éléments dont le nom de table principale est qualifié par le propriétaire.

L'extension SAP Data Services CDC 14.2.0 active l'attribut de propriétaires multiples pour toutes les tables principales qualifiées par le propriétaire.

4. Déployez ces scripts à l'aide de l'outil isql et de l'outil de base de données source correspondant pour créer des droits et les octroyer aux utilisateurs.

L'utilisateur de la base de données qui déploie les scripts doit être un superutilisateur. Saisissez le code :

```
isql -Usa -Ppwd -Spdb -Dpdb -i Primary_database_Runtime script.sql -o  
Primary_database_Runtime script.out
```

où :

- **<sa>** est le nom de connexion de l'administrateur système sur le serveur de données intermédiaire CDM Data Services.
- **<pwd>** est le mot de passe correspondant au nom de connexion de l'administrateur système.
- **<pds>** est le nom du serveur de données intermédiaire CDM Data Services.
- **<pdb>** est le nom de la base de données intermédiaire CDM Data Services.

Les tables principales sont marquées comme répliquables et les données modifiées peuvent être capturées par l'agent de réplication et propagées par le serveur de réplication.

34.6.2.4 Configuration de SAP BusinessObjects Data Services

Pour utiliser Data Services en vue de lire et charger les données modifiées à partir du serveur de réplication SAP Sybase à l'aide de la modélisation PowerAMC, appliquez les procédures suivantes dans Designer.

- Créez une banque de données CDM.
- Importez les métadonnées pour les tables de SAP Replication Server.
- Créez et définissez un flux de données.
- Configurez une source CDM.

34.6.2.4.1 Création d'une banque de données CDM et importation des métadonnées de table

Cette procédure décrit comment créer une banque de données CDM à l'aide de SAP Replication Server. Pour accéder à la source, créez une banque de données CDM à l'aide de Designer. Une banque de données CDM est une banque de données en lecture seule qui peut uniquement accéder aux tables. Après avoir créé la banque de données CDM, vous l'utilisez pour importer les métadonnées de table CDM.

1. Dans l'onglet *Banques de données* de la bibliothèque d'objets, ouvrez l'éditeur de banque de données.
2. Saisissez un nom pour la banque de données.
3. Pour le *Type de banque de données*, sélectionnez *CDM RepServer*.
4. Pour le *Type de base de données*, sélectionnez *SAP ASE*.
5. Sélectionnez la *Version de la base de données*.
6. Saisissez le *Nom du serveur de la base de données*.
7. Saisissez un *Nom de base de données*.
8. Saisissez un *Nom d'utilisateur* et un *Mot de passe* de base de données.

9. Pour créer plusieurs configurations pour cette banque de données, cliquez sur [Appliquer](#), puis sur [Modifier](#) et entrez les détails de configuration.

i Remarque

Il est impossible de modifier le type de base de données, la version ou la méthode CDM.

10. Cliquez sur [OK](#).
11. Pour importer les tables, cliquez avec le bouton droit de la souris sur le nom de la banque de données dans la bibliothèque d'objets et sélectionnez [Ouvrir](#), [Importer par nom](#) ou [Rechercher](#).
Seules les tables CDM sélectionnées lors de la configuration de SAP Replication Server pour le CDM sont visibles lorsque vous parcourez les métadonnées externes. Vous devez créer une table CDM dans SAP ASE pour chaque table que vous voulez lire afin de pouvoir importer cette table CDM à l'aide de SAP Data Services.

Informations associées

[Définition d'une banque de données d'une base de données \[page 85\]](#)

34.6.2.4.2 Création et définition du flux de données

Cette procédure décrit comment créer un flux de données pour lire les données modifiées à l'aide de SAP Replication Server et de la modélisation SAP PowerAMC.

1. Dans la bibliothèque d'objets, faites glisser la table CDM, une transformation Query et une transformation Map_CDC_Operation vers l'espace de travail.

i Remarque

- Les tables CDM multiples du serveur de réplication sont autorisées dans le même flux de données.
- La table CDM du serveur de réplication stocke les enregistrements de données extraits du journal de la base de données source. Cependant, ces enregistrements de données stockés ne sont pas équivalents aux enregistrements disponibles dans la table source correspondante. De plus, les enregistrements de données modifiées capturés antérieurs à la période de rétention sont également supprimés de la table CDM du serveur de réplication. De ce fait, joindre deux ou plusieurs tables CDM de ce type risque de ne pas produire le même résultat que de joindre leurs tables de base de données source correspondantes.

2. Configurez la table CDM comme indiqué dans la section suivante.
3. Ajoutez la table cible appropriée et reliez les objets.
4. Ouvrez l'Editeur de requêtes.
5. Mappez les colonnes de contrôle Data Services et les colonnes de la table source que vous voulez dans la cible.

La transformation Map_CDC_Operation utilise les valeurs de la colonne dans la zone [Colonne de l'opération sur la ligne](#) pour exécuter l'opération appropriée sur la ligne source pour la table cible. Pour la table source CDM, la colonne DI_OPERATION_TYPE est automatiquement sélectionnée comme [Colonne de l'opération sur la ligne](#).

Les opérations peuvent être INSERT, DELETE ou UPDATE. Par exemple, si l'opération est DELETE, la ligne correspondante est supprimée de la table cible.

Informations associées

[Configuration d'une table source CDM \[page 835\]](#)

34.6.2.4.3 Configuration d'une table source CDM

Cette procédure décrit comment configurer une table source CDM. Data Services peut appliquer des points de contrôle sur toutes les tables (dans une configuration de banque de données particulière) pour tous les flux de données d'un job afin d'assurer la cohérence des données.

1. Faites glisser une table de banque de données CDM dans un flux de données.
2. Cliquez sur le nom de cet objet source pour ouvrir son éditeur.
3. Cliquez sur l'onglet *Options CDM*.
4. Vous pouvez sélectionner l'option facultative *Activer le point de contrôle*.

Une fois qu'un point de contrôle est placé, à la prochaine exécution du job CDM, celui-ci lit uniquement les lignes insérées dans la table CDM depuis le dernier point de contrôle.

Si les points de contrôle ne sont pas activés, le job CDM lit toutes les lignes de la source de données CDM (avec un temps de traitement augmenté).

34.7 Utilisation de la CDM avec les bases de données Microsoft SQL Server

Capture de données modifiées sur des bases de données Microsoft SQL Server

SAP Data Services peut capturer les données modifiées dans les bases de données Microsoft SQL Server et les appliquer aux systèmes cible à l'aide des méthodes suivantes :

- Capture des données modifiées
- Suivi des modifications
- Serveur de réplication

Les trois méthodes utilisent le concept de points de contrôle pour lire le set de modifications le plus récent et marquer la fin de la lecture. Si les points de contrôle ne sont pas activés, SAP BusinessObjects Data Services lit l'ensemble des modifications. En outre, si vous activez les points de contrôle et exécutez votre job CDM en mode de récupération, le job récupéré commence à examiner la source de données CDM au dernier point de contrôle.

Pour capturer les données modifiées :

- Activez la CDM sur la base de données Microsoft SQL Server
- Activez la CDM sur les tables source
- Configurez les banques de données, les jobs et les sources de SAP BusinessObjects Data Services

Consultez la documentation Microsoft pour obtenir des détails sur toutes les méthodes. Le tableau suivant compare ces méthodes.

Fonctionnalité	Capture des données modifiées	Suivi des modifications	Serveur de réplication
Version de Microsoft SQL Server prise en charge	2008 et ultérieure	2008 et ultérieure	2000, 2005, 2008
Synchrone (immédiate ; latence si asynchrone)	Non, les modifications sont capturées par le job CDM de SQL Server.	Oui, suit les modifications en ligne avec les opérations INSERT, UPDATE et DELETE, les modifications sont donc disponibles immédiatement.	Non, les modifications sont disponibles après transfert des opérations par l'agent lecteur de journal dans la base de données de distribution.
Requiert un agent SQL	Oui, un job de capture remplit les tables CDM.	Non, le moteur de base de données renseigne les clés primaires dans les tables de modifications durant les opérations DML.	Oui, un agent lecteur du journal de réplication gère la réplication.
Processus automatique de nettoyage	Oui, périodiquement	Oui, périodiquement	Oui, périodiquement
Requiert des tables distinctes pour stocker les données de suivi	Oui, stocké dans un table de captures Le stockage dépend du nombre de colonnes capturées pour la CDM.	Oui. Utilise une tables de suivi des modifications interne par table source. Utilise une table de transactions par base de données. Le stockage dépend du nombre de colonnes de clés primaires de la table source.	Oui, stocké dans une base de données de distribution.
Données historiques disponibles	Oui	Non	Oui
Requiert une clé primaire	Non	Oui	Oui
Pré-image disponible pour l'opération UPDATE dans SAP BusinessObjects Data Services	Oui. SAP BusinessObjects Data Services lit automatiquement la pré-image et la met à disposition pour utilisation ultérieure dans le flux de données.	Non	Oui (facultatif)
Recommandation	Utiliser si les données historiques ou la cohérence entre plusieurs tables sont requises.	Utiliser pour une synchronisation rapide.	Pris en charge, mais obsolète.

34.7.1 Restrictions

Les restrictions suivantes s'appliquent à toutes les méthodes de capture des données modifiées dans Microsoft SQL Server :

- Il est impossible d'utiliser les transformations et fonctions suivantes pour une table source importée avec une banque de données CDM en raison de l'existence des colonnes générées par SAP BusinessObjects Data Services pour les tables CDM. Le progiciel ne peut pas comparer ou rechercher ces colonnes.
 - Transformations Table_Comparison, Key_Generation, SQL, et Data_Transfer. La transformation History_Preserving n'est pas prise en charge par la méthode de suivi des modifications.
 - Toutes les fonctions de base de données, y compris lookup, lookup_ext, lookup_seq, search_replace, pushdown_sql, truncate_table, key_generation, sql et total_rows.
- Une seule table de source CDM peut être créée dans un flux de données.
- Les tables CDM ne sont pas autorisées dans les jobs en temps réel.
- Le profilage n'est pas disponible pour les tables CDM.
- L'affichage du SQL optimisé n'est pas disponible pour les flux de données qui lisent les tables CDM.
- L'exportation ou l'importation de jobs avec les tables CDM ne retient pas les informations des points de contrôle enregistrés dans le référentiel de SAP BusinessObjects Data Services.

Pour les méthodes du suivi des modifications et du serveur de réplication, n'utilisez pas le même nom d'abonnement pour lire la même table CDM dans des flux de données parallèles, parce que le point de contrôle est identifié par une combinaison du nom d'abonnement et du nom de table. Autrement dit, pour éviter les problèmes de corruption de données, ne réutilisez pas les flux de données qui utilisent les banques de données CDM parce que chaque fois qu'une table source extrait des données, elle utilise le même nom d'abonnement. C'est-à-dire que des jobs identiques, selon le moment où ils sont exécutés, peuvent obtenir des résultats différents et laisser des points de contrôle à différents endroits du fichier.

Pour la méthode de suivi des modifications, comme Microsoft SQL Server ne peut pas suivre l'opération TRUNCATE sur une table source, SAP BusinessObjects Data Services ne peut pas propager ces résultats vers la cible.

34.7.2 Colonnes de SAP BusinessObjects Data Services

Lorsque vous importez des métadonnées dans une banque de données CDM à l'aide d'une des méthodes, SAP BusinessObjects Data Services ajoute les colonnes suivantes à la table d'origine :

- DI_SEQUENCE_NUMBER
- DI_OPERATION_TYPE

34.7.2.1 Colonne DI_SEQUENCE_NUMBER

La colonne DI_SEQUENCE_NUMBER commence par zéro au début de chaque extraction. Ce champ est incrémenté d'un chaque fois que le progiciel lit une ligne, sauf s'il rencontre une paire de pré et post-images. La pré-image et la post-image reçoivent le même numéro de séquence. Cette colonne de séquençement fournit un moyen de rassembler les paires d'images si elles sont séparées en raison de la conception du flux de données.

Lors du déroulement du flux de données, si les pré et post-images sont séparées ou réparties en plusieurs lignes (par exemple, en utilisant des clauses GROUP BY ou ORDER BY dans une requête), vous pouvez perdre l'ordre des lignes.

La transformation Map_CDC_Operation permet de restaurer l'ordre d'origine des paires d'images en utilisant la colonne DI_SEQUENCE_NUMBER comme *Colonne de séquençement*.

Pour la méthode du serveur de réplication, vous pouvez configurer le serveur afin de récupérer les pré-images des lignes UPDATE avant que SAP BusinessObjects Data Services n'applique l'opération UPDATE à la cible. Notez que si vous ne configurez pas le serveur de réplication pour capturer les pré-images dans la base de données, seules les post-images sont capturées par défaut. Pour en savoir plus sur les cas où il faut envisager l'utilisation des pré-images, voir *Utilisation des pré-images* [page 795].

34.7.2.2 Colonne DI_OPERATION_TYPE

SAP BusinessObjects Data Services génère des valeurs dans la colonne DI_OPERATION_TYPE. Les valeurs valides pour cette colonne sont :

- I pour INSERT
- D pour DELETE
- B pour pré-image d'une opération UPDATE (sauf pour la méthode de suivi des modifications, qui n'utilise pas de pré-image)
- U pour une post-image d'une opération UPDATE

34.7.3 Méthode de capture des données modifiées (CDM)

Avec la méthode de capture des données modifiées (CDM), SAP BusinessObjects Data Services applique des points de contrôle sur toutes les tables (dans une configuration de banque de données particulière) pour tous les flux de données d'un job afin d'assurer la cohérence des données. Activez d'abord la CDM dans l'Editeur de la banque de données, puis sélectionnez l'option *Activer le point de contrôle* dans l'éditeur de tables source du flux de données. Si l'option *Activer le point de contrôle* n'est pas sélectionnée, SAP BusinessObjects Data Services récupère toutes les données disponibles à ce moment pour cette table. Cette méthode de CDM n'est pas disponible avec les versions antérieures à Microsoft SQL Server 2008.

34.7.3.1 Ajout d'une banque de données CDM

Cette procédure décrit comment créer une connexion de banque de données pour Microsoft SQL Server et activer la méthode de capture des données modifiées (CDM).

1. Ouvrez l'éditeur de banque de données.
2. Saisissez un nom pour la banque de données.
3. Dans la zone *Type de banque de données*, sélectionnez *Base de données*.
4. Dans la zone *Type de base de données*, sélectionnez *Microsoft SQL Server*.

5. Cochez la case [Activer la CDM](#).
6. Pour [Version de la base de données](#), sélectionnez [Microsoft SQL Server 2008](#) ou une version ultérieure.
7. Dans la zone déroulante sous l'option [Activer la CDM](#), sélectionnez la méthode [CDM](#).
8. Saisissez un [Nom de serveur de base de données](#).
9. Saisissez un [Nom de base de données](#).
10. Saisissez un [Nom d'utilisateur](#) et un [Mot de passe](#) de base de données.
11. Pour créer plusieurs configurations pour cette banque de données, cliquez sur [Appliquer](#), puis sur [Modifier](#) et entrez les détails de configuration. Notez que vous ne pouvez pas modifier le type de base de données, la version ou la méthode CDM.
12. Cliquez sur [OK](#).

Vous pouvez maintenant utiliser la nouvelle connexion de banque de données pour importer les métadonnées d'une instance de capture dans le référentiel actuel.

Informations associées

[Définition d'une banque de données d'une base de données \[page 85\]](#)

34.7.3.2 Importation de métadonnées CDM

Après avoir créé une banque de données CDM, vous pouvez l'utiliser pour importer des métadonnées d'instance de capture.

Si la CDM est activée sur une table, Microsoft SQL Server crée une instance de capture. Une table peut avoir jusqu'à deux instances de capture avec des noms différents. Dans la bibliothèque d'objets, cliquez avec le bouton droit sur le nom de la banque de données et sélectionnez [Ouvrir](#), [Importer par nom](#) ou [Rechercher](#). SAP BusinessObjects Data Services affiche le nom de l'instance de capture (au lieu du nom de la table sous-jacente). Par conséquent, si vous effectuez une importation par nom, ce dernier doit être le nom de l'instance de capture (et non celui de la table).

L'opération d'importation ajoute les colonnes suivantes à la table d'origine :

Nom de colonne	Type de données	Source de la colonne
DI_SEQUENCE_NUMBER	integer	Généré par SAP BusinessObjects Data Services
DI_OPERATION_TYPE	varchar(1)	Généré par SAP BusinessObjects Data Services
MSSQL_TRAN_SEQNO	varchar(256)	Fourni par Microsoft SQL Server
MSSQL_TRAN_TIMESTAMP	timestamp	Fourni par Microsoft SQL Server
MSSQL_COLUMN_UPDATE_MASK	varchar(258)	Fourni par Microsoft SQL Server

34.7.3.3 Configuration d'une table source utilisant la CDM

Cette procédure décrit comment configurer une table source CDM en employant la méthode CDM (Capture des Données Modifiées). Pour en savoir plus, consultez le *Guide de référence*.

1. Faites glisser une table de banque de données CDM dans un flux de données.
La table devient automatiquement un objet source.
2. Cliquez sur le nom de cet objet source pour ouvrir son éditeur de table source.
3. Cliquez sur l'onglet *Options CDM*.
4. Facultativement, sélectionnez *Activer le point de contrôle*.
Une fois qu'un point de contrôle est placé, à la prochaine exécution du job CDM, celui-ci lit uniquement les lignes insérées dans la table CDM depuis le dernier point de contrôle.
5. Facultativement, sélectionnez *Supprimer automatiquement les lignes après la lecture*.

34.7.3.4 Utilisation de CDM pour les flux de données dans une boucle WHILE

Pour la méthode CDM, les points de contrôle s'appliquent au niveau du job. Si vous avez des flux de données s'exécutant dans une boucle WHILE pour récupérer les modifications, utilisez la fonction `set_cdc_checkpoint()` pour chaque itération de la boucle. Cette fonction donne au lecteur source l'instruction de définir des points de contrôle afin que la prochaine itération prélève les dernières modifications. Appelez cette fonction pour toutes les banques de données utilisées dans l'ensemble des flux de données du job. Pour en savoir plus, consultez le *Guide de référence*.

```
set_cdc_checkpoint (<Banque de données>)
```

La fonction renvoie 1 si le point de contrôle a été correctement défini, sinon elle renvoie 0. La valeur `<banque de données>` est le nom de la banque de données contenant les tables CDM.

❖ Exemple

```
set_cdc_checkpoint('MyCdcSource');
```

34.7.4 Méthode de suivi des modifications

La méthode de suivi des modifications identifie les lignes d'une table ayant été modifiées, mais ignore le nombre de fois où la ligne a été modifiée ou les valeurs de toute modification intermédiaire. SAP BusinessObjects Data Services extrait uniquement les dernières données disponibles. Le suivi des modifications est donc limité dans les réponses aux questions d'historique par rapport à la méthode CDM (Capture des données modifiées). Toutefois, il y a nettement moins de données inutiles à stocker parce que les données modifiées ne sont pas capturées. En outre, un mécanisme de suivi synchrone utilisé pour suivre les modifications implique un minimum de données inutiles dans les opérations.

Le suivi des modifications doit d'abord être activé pour la base de données Microsoft SQL Server, puis pour les tables à suivre dans cette base de données. Les informations de suivi des modifications sont enregistrées pour

les lignes modifiées. Les valeurs de la colonne de clé primaire de la table suivie sont enregistrées avec les informations de modification afin d'identifier les lignes ayant été modifiées. Pour obtenir les dernières données de ces lignes, SAP BusinessObjects Data Services utilise les valeurs de la colonne de clé primaire pour relier la table source et la table suivie. Les informations sur les modifications peuvent inclure le type d'opération ayant entraîné la modification (INSERT, UPDATE ou DELETE) ou les colonnes ayant été modifiées dans le cadre d'une opération UPDATE, par exemple.

Cette méthode n'est pas disponible avec les versions antérieures à Microsoft SQL Server 2008.

34.7.4.1 Ajout d'une banque de données de suivi des modifications

Cette procédure décrit comment créer une connexion de banque de données à Microsoft SQL Server et activer la méthode de suivi des modifications.

1. Ouvrez l'éditeur de banque de données.
2. Saisissez un nom pour la banque de données.
3. Dans la zone *Type de banque de données*, sélectionnez *Base de données*.
4. Dans la zone *Type de base de données*, sélectionnez *Microsoft SQL Server*.
5. Cochez la case *Activer la CDM*.
6. Pour *Version de la base de données*, sélectionnez *Microsoft SQL Server 2008* ou une version ultérieure.
7. Dans la zone déroulante sous l'option *Activer la CDM*, sélectionnez la méthode *Suivi des modifications*.
8. Saisissez un nom de *serveur de base de données*.
9. Saisissez un *Nom de base de données*.
10. Saisissez un *Nom d'utilisateur* et un *Mot de passe* de base de données.
11. Pour créer plusieurs configurations pour cette banque de données, cliquez sur *Appliquer*, puis sur *Modifier* et entrez les détails de configuration. Notez que vous ne pouvez pas modifier le type de base de données, la version ou la méthode CDM.
12. Cliquez sur *OK*.

Vous pouvez maintenant utiliser la nouvelle connexion de banque de données pour importer les métadonnées de table dans le référentiel actuel.

Informations associées

[Définition d'une banque de données d'une base de données \[page 85\]](#)

34.7.4.2 Importation des métadonnées de suivi des modifications

Après avoir créé une banque de données de suivi des modifications, vous pouvez l'utiliser pour importer les métadonnées des tables.

Dans la bibliothèque d'objets, cliquez avec le bouton droit sur le nom de la banque de données et sélectionnez [Ouvrir](#), [Importer par nom](#) ou [Rechercher](#). Seules les tables CDM que vous sélectionnez en configurant Microsoft SQL Server pour le suivi des modifications sont visibles lorsque vous parcourez les métadonnées externes.

L'opération d'importation ajoute les colonnes suivantes à la table d'origine :

Nom de colonne	Type de données	Source de la colonne
DI_SEQUENCE_NUMBER	integer	Généré par SAP BusinessObjects Data Services
DI_OPERATION_TYPE	varchar(1)	Généré par SAP BusinessObjects Data Services
MSSQL_SYS_CHANGE_VERSION	decimal(19,0)	Fourni par Microsoft SQL Server
MSSQL_SYS_CHANGE_CREATION_VERSION	decimal(19,0)	Fourni par Microsoft SQL Server
MSSQL_SYS_CHANGE_CONTEXT_MASK	varchar(256)	Fourni par Microsoft SQL Server

34.7.4.3 Configuration d'une table source pour utiliser le suivi des modifications

Cette procédure décrit comment configurer une table source CDM pour activer le suivi des modifications.

1. Faites glisser une table de banque de données CDM dans un flux de données.
La table devient automatiquement un objet source.
2. Cliquez sur le nom de cet objet source pour ouvrir son éditeur de table source.
3. Cliquez sur l'onglet [Options CDM](#).
4. Indiquez une valeur pour le [nom d'abonnement CDM](#)
SAP BusinessObjects Data Services utilise ce nom pour suivre le dernier point de contrôle en interne dans le référentiel.
5. Sélectionnez [Activer le point de contrôle](#).
Une fois qu'un point de contrôle est placé, à la prochaine exécution du job CDM, celui-ci lit uniquement les lignes insérées dans la table CDM depuis le dernier point de contrôle.

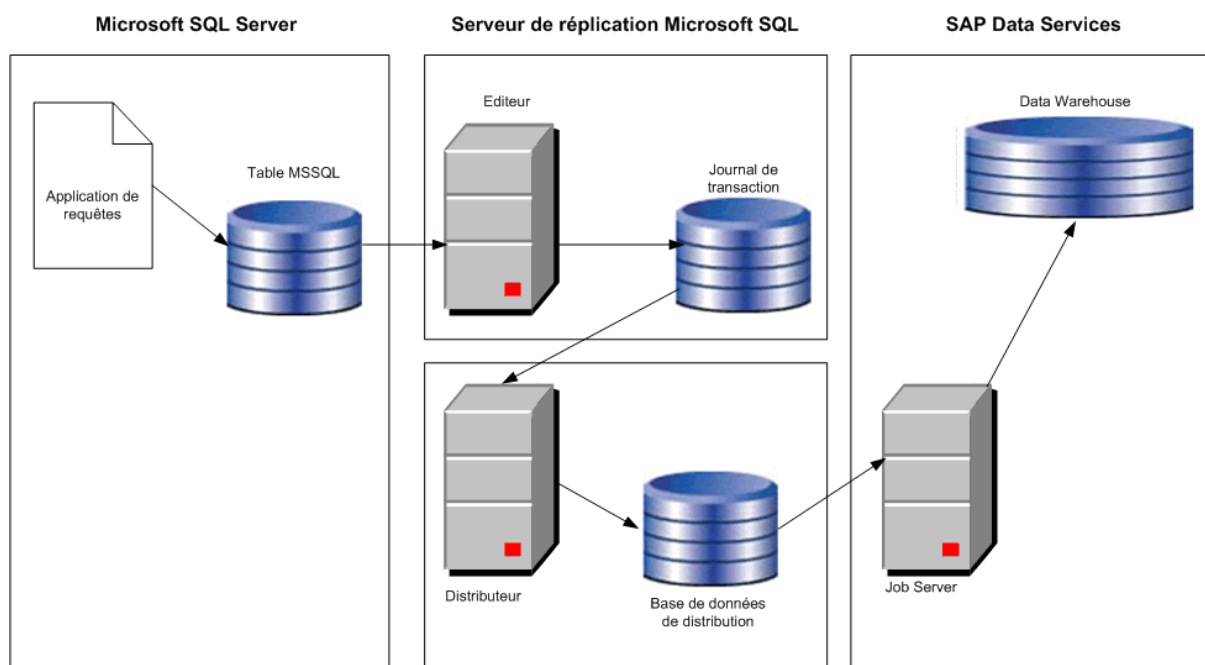
34.7.5 Méthode du serveur de réplication

Microsoft utilise les termes suivants pour le serveur de réplication Microsoft SQL :

- Article : un article est une table, une partition ou un objet de base de données que l'administrateur de base de données indique pour la réplication. Un article peut être un des éléments suivants :

- Une table entière
- Certaines colonnes (à l'aide d'un filtre vertical)
- Certaines lignes (à l'aide d'un filtre horizontal)
- Une procédure stockée ou une définition de vue
- L'exécution d'une procédure stockée
- Une vue
- Une vue indexée
- Une fonction définie par l'utilisateur
- Distributeur : le distributeur est un serveur qui stocke des métadonnées, des données historiques et des transactions dans la base de données de distribution. Le progiciel lit la base de données de distribution pour obtenir les données modifiées.
- Publication : une publication est une collection d'un ou plusieurs articles d'une base de données. Une publication facilite la spécification d'un jeu de données et d'objets de base de données liés logiquement, que vous désirez répliquer ensemble.
- Editeur : l'Editeur est un serveur qui met les données à disposition pour la réplcation sur les autres serveurs.
- Abonné : un Abonné est un serveur qui reçoit les données répliquées. Les Abonnés s'abonnent aux publications et non aux articles individuels d'une publication. Ils s'abonnent uniquement aux publications dont ils ont besoin et non à l'ensemble des publications disponibles sur un Editeur.

Le progiciel obtient les données modifiées de la base de données de distribution dans le serveur de réplcation Microsoft SQL. Le diagramme suivant illustre le flux des données modifiées du serveur de réplcation vers SAP BusinessObjects Data Services.



- Une application apporte des modifications à une base de données et L'Editeur hébergé dans le serveur de réplcation capture ces modifications dans un journal de transactions.
- L'agent lecteur de journaux du distributeur lit le journal de transactions de l'Editeur et enregistre les données modifiées dans la base de données de distribution.

- Le progiciel lit les données de la table de commandes dans la base de données de distribution, applique les filtres adéquats et crée des lignes d'entrée pour une table de l'entrepôt de données cible.

Le progiciel accède aux tables suivantes dans la base de données de distribution :

- MSarticles : contient une ligne pour chaque article répliqué par un Editeur.
- MSpublications : contient une ligne pour chaque publication répliquée par un Editeur.
- MSPublisher_databases : contient une ligne pour chaque paire Editeur et base de données d'Editeur desservie par le distributeur local.
- MSrepl_commands : contient les lignes des commandes répliquées (modifications des données).

Lorsque vous activez une base de données pour la réplication, le serveur de réplication crée des tables dans la base de données source. Une de ces tables est Sysarticles, qui contient une ligne pour chaque article défini dans cette base de données spécifique. Une des colonnes de la table Sysarticles indique quelles colonnes d'une table source sont publiées.

34.7.5.1 Configuration de la base de données de distribution

Si le progiciel se connecte à Microsoft SQL Server pour extraire des données, vous devez configurer la base de données de distribution dans le serveur de réplication pour capturer les modifications de ces tables.

34.7.5.1.1 Microsoft SQL Server 2000

Les étapes suivantes décrivent la procédure de configuration du serveur de réplication pour les bases de données Microsoft SQL Server 2000.

1. Sur le nœud de réplication de Microsoft SQL Enterprise Manager, sélectionnez l'option [Configurer la publication, les abonnés et la distribution](#). Suivez les étapes de l'assistant pour créer les bases de données de distribution et du serveur de distribution.

Les étapes suivantes décrivent la procédure de configuration du serveur de réplication SQL pour les bases de données Microsoft SQL Server 2000.

L'assistant génère les composants suivants que vous devez spécifier dans l'Editeur de banque de données lors de la définition d'une banque de données CDM Microsoft SQL Server :

- Nom du serveur de distribution MSSQL
 - Nom de la base de données de distribution MSSQL
 - Nom d'utilisateur de distribution MSSQL
 - Mot de passe de distribution MSSQL
2. Sélectionnez l'option [Nouvelles publications](#) sur le nœud de réplication de Microsoft SQL Enterprise Manager pour créer de nouvelles publications indiquant les tables que vous souhaitez publier. Le progiciel nécessite les paramètres suivants dans les Options avancées :
 - Sélectionnez l'option [Publication transactionnelle](#) dans la fenêtre Sélectionner le type de publication. Ce type de publication met à jour les données de l'Editeur et envoie les modifications de manière incrémentielle à l'Abonné.
 - Dans l'onglet Commandes de la fenêtre Propriétés de l'article de table :

- Si vous souhaitez des pré-images pour les commandes UPDATE et DELETE, sélectionnez XCALL. Autrement, sélectionnez CALL.
- Effacez les options *Créer les procédures stockées lors de la synchronisation initiale des abonnements* et *Envoyer les paramètres au format binaire* car le progiciel n'utilise pas de procédure de stockage et dispose de son propre format interne.
- Dans l'onglet Instantané de la fenêtre Propriétés de l'article de table :
 - Sélectionnez *Conserver la table existante inchangée* car le progiciel traite la table en tant que journal.
 - Effacez l'option *Index cluster* car le progiciel traite la table en tant que journal et lit dans celle-ci de manière séquentielle.
- Spécifiez un nom et une description pour la publication. Vous indiquez le nom de cette publication dans l'Editeur de banque de données lors de la définition d'une banque de données CDM MSSQL.
- Sélectionnez l'option *Oui, autoriser les abonnements anonymes* pour enregistrer toutes les transactions dans la base de données de distribution.

Pour en savoir plus, voir l'aide en ligne de Microsoft SQL Enterprise Manager.

34.7.5.1.2 Microsoft SQL Server 2005 et 2008

La procédure suivante décrit la configuration de publications pour les bases de données Microsoft SQL Server 2005 et 2008 pour CDM.

1. Démarrez Microsoft SQL Server Management Studio.
2. Connectez-vous et accédez au nœud *Réplication* dans l'Explorateur d'objets.
3. Si vous configurez la distribution de ce serveur pour la première fois, cliquez avec le bouton droit de la souris sur le nœud *Réplication* et sélectionnez *Configurer la distribution* dans le menu de contexte. Vous pouvez configurer le serveur de distribution, le dossier d'instantanés, la base de données de distribution et les utilisateurs de ce serveur de distribution.
4. Cliquez de nouveau avec le bouton droit de la souris sur le nœud *Réplication* et sélectionnez ► *Nouveau* ► *Publication* ►. L'Assistant Nouvelle publication s'ouvre.
5. Dans l'Assistant Nouvelle publication, cliquez sur *Suivant*.
6. Sélectionnez la base de données que vous souhaitez publier, puis cliquez sur *Suivant*.
7. Sous Type de publication, sélectionnez *Publication transactionnelle*, puis cliquez sur *Suivant*.
8. Cliquez pour sélectionner les tables et colonnes à publier en tant qu'articles. Ouvrez ensuite la fenêtre *Propriétés de l'article*. Pour chaque table sélectionnée, cliquez sur *Définir les propriétés de l'article de table en surbrillance*. La fenêtre Propriétés de l'article s'ouvre :
 - a. Définissez *Copier l'index cluster*, *Copier les procédures stockées INSERT, UPDATE et DELETE* et *Créer des schémas sur l'Abonné* sur FALSE.
 - b. Définissez *Action si le nom est déjà utilisé* sur *Conserver la table existante inchangée*.
 - c. Définissez *Format de remise UPDATE* et *Format de remise DELETE* sur *XCALL <procédure stockée>*.
 - d. Cliquez sur *OK* pour enregistrer les propriétés de l'article.
9. Cliquez sur *Suivant*. Vous pouvez cliquer sur *Ajouter* pour ajouter des filtres de lignes. Cliquez sur *Suivant* si le filtrage de données n'est pas nécessaire dans la publication.
10. Configurez la Sécurité de l'agent et spécifiez le paramètre de connexion du compte.

- a. Pour l'Agent d'instantané, cliquez sur [Paramètres de sécurité](#). Spécifiez le compte d'exécution de l'agent d'instantané. Configurez le compte avec des droits d'administration système. Spécifiez le compte se connectant à l'Editeur et cliquez sur [OK](#).
 - b. Pour l'Agent de lecture du journal, l'option [Utiliser les paramètres de sécurité de l'Agent de capture instantanée](#) est sélectionnée par défaut. Pour utiliser des paramètres différents, effacez cette option et cliquez sur [Paramètres de sécurité](#). Notez qu'un identifiant octroyant des droits d'administration système est nécessaire.
11. Dans la fenêtre Actions de l'assistant, sélectionnez [Créer la publication](#) puis cliquez sur [Suivant](#).
 12. Pour terminer l'assistant, saisissez un nom de publication et cliquez sur [Terminer](#).
- Pour en savoir plus, voir la documentation relative à Microsoft SQL Enterprise Manager.

34.7.5.2 Configuration de SAP BusinessObjects Data Services

Pour utiliser SAP BusinessObjects Data Services afin de lire et charger les données modifiées à partir de bases de données SQL Server à l'aide du serveur de réplication, appliquez les procédures suivantes dans Designer :

- Création d'une banque de données CDM
- Importation de métadonnées pour les tables Microsoft SQL Server
- Configuration d'une source CDM

34.7.5.2.1 Ajout de la banque de données CDM

Cette procédure décrit comment créer une banque de données CDM à l'aide de la méthode de réplication.

1. Ouvrez l'Editeur de banque de données.
2. Saisissez un nom pour la banque de données.
3. Dans la zone [Type de banque de données](#), sélectionnez [Base de données](#).
4. Dans la zone [Type de base de données](#), sélectionnez [Microsoft SQL Server](#).
5. Cochez la case [Activer la CDM](#) pour activer la fonctionnalité de CDM.
6. Sélectionnez une [Version de base de données](#).
7. Saisissez un [Nom de base de données](#) (utilisez le nom du serveur de réplication).
8. Saisissez un [Nom d'utilisateur](#) et un [Mot de passe](#) de base de données.
9. Dans la section CDM, entrez les noms que vous avez créés pour cette banque de données lorsque vous avez configuré le Distributeur et l'Editeur dans le serveur de réplication :
 - Nom du serveur de distribution MSSQL
 - Nom de la base de données de distribution MSSQL
 - Nom de publication MSSQL
 - Nom d'utilisateur de distribution MSSQL
 - Mot de passe de distribution MSSQL
10. Pour créer plusieurs configurations pour cette banque de données, cliquez sur [Appliquer](#), puis sur [Modifier](#) et suivez à nouveau l'étape 9 pour toute configuration supplémentaire.

11. Cliquez sur [OK](#).

Vous pouvez maintenant utiliser la nouvelle connexion de banque de données pour importer les métadonnées de table dans le référentiel actuel.

Informations associées

[Définition d'une banque de données d'une base de données \[page 85\]](#)

34.7.5.2.2 Importation de métadonnées CDM

Après avoir créé une banque de données CDM, vous pouvez l'utiliser pour importer les métadonnées des tables CDM. Dans la bibliothèque d'objets, cliquez avec le bouton droit sur le nom de la banque de données et sélectionnez [Ouvrir](#), [Importer par nom](#) ou [Rechercher](#). Seules les tables CDM que vous sélectionnez en configurant Microsoft SQL Server pour le CDM sont visibles lorsque vous parcourez les métadonnées externes. SAP BusinessObjects Data Services utilise les tables MSpublishings et MSArticles dans la base de données de distribution du serveur de réplication SQL pour créer une liste de tables publiées.

Lorsque vous importez chaque table CDM, le progiciel utilise la table Sysarticles de la base de données de l'Editeur du serveur de réplication SQL pour afficher uniquement les colonnes publiées.

L'opération d'importation ajoute les colonnes suivantes à la table d'origine :

Nom de colonne	Type de données	Source de la colonne
DI_SEQUENCE_NUMBER	integer	Généré par SAP BusinessObjects Data Services
DI_OPERATION_TYPE	varchar(1)	Généré par SAP BusinessObjects Data Services
MSSQL_TRAN_SEQNO	varchar(256)	Fourni par le serveur de réplication
MSSQL_TRAN_TIMESTAMP	timestamp	Fourni par le serveur de réplication

Informations associées

[Configuration de la base de données de distribution \[page 830\]](#)

34.7.5.2.3 Configuration d'une table source utilisant la réplication

Cette procédure décrit comment configurer une table source CDM à l'aide de la méthode de réplication. Pour en savoir plus, consultez le *Guide de référence*.

1. Faites glisser une table de banque de données CDM dans un flux de données.
La table devient automatiquement un objet source.
2. Cliquez sur le nom de cet objet source pour ouvrir son éditeur de table source.
3. Cliquez sur l'onglet *Options CDM*.
4. Indiquez une valeur pour le *nom d'abonnement CDM*
SAP BusinessObjects Data Services utilise ce nom pour suivre le dernier point de contrôle en interne dans le référentiel.

Informations associées

[Utilisation des points de contrôle de l'ordinateur central \[page 804\]](#)

34.7.5.2.3.1 Utilisation des points de contrôle avec les serveurs de réplication

Un agent lecteur de journaux du serveur de réplication Microsoft SQL lit le journal de transactions de l'Editeur et enregistre les données modifiées dans la base de données de distribution, que SAP BusinessObjects Data Services utilise comme source de données CDM. Les lignes de données modifiées s'ajoutent au précédent chargement dans la source de données CDM.

Lorsque vous activez des points de contrôle, un job CDM utilise le nom d'abonnement pour lire le plus récent ensemble de lignes ajoutées et pour marquer la fin de la lecture. Si les points de contrôle ne sont pas activés, le job CDM lit toutes les lignes de la source de données CDM (avec un temps de traitement augmenté).

Pour utiliser les points de contrôle, dans l'Editeur de table source, saisissez le *Nom de l'abonnement CDM* et sélectionnez l'option *Activer le point de contrôle*.

34.7.5.2.3.2 Utilisation des pré-images des sources Microsoft SQL Server

Si vous devez capturer des pré-images de lignes UPDATE :

1. Vérifiez que le serveur de réplication est configuré pour récupérer des pré-images complètes.
2. Lorsque vous créez une publication dans le serveur de réplication, indiquez XCALL pour les commandes UPDATE et DELETE afin d'obtenir les pré-images.

3. Sélectionnez l'option *Obtenir les pré-images pour chaque ligne de mise à jour* dans l'éditeur source de la table CDM.

Après avoir coché l'option *Obtenir les pré-images pour chaque ligne de mise à jour*, le progiciel traite deux lignes pour chaque mise à jour. Outre l'impact sur les performances de cette augmentation du volume de données, les paires de pré et post-images pourraient être séparées ou perdues selon la conception de votre flux de données, ce qui provoquerait des problèmes d'intégrité des données.

i Remarque

La transformation Map_CDC_Operation peut résoudre les problèmes, mais des résultats indésirables peuvent toujours se produire en raison d'erreurs de programmation. Lors de l'utilisation de fonctions et transformations qui réordonnent, redirigent, éliminent et multiplient le nombre de lignes dans un flux de données, soyez conscients de l'impact possible sur les cibles.

Informations associées

[Utilisation des pré-images \[page 795\]](#)

34.7.5.2.3.3 Configuration d'une table source CDM

Cette procédure décrit comment configurer une table source CDM. Data Services peut appliquer des points de contrôle sur toutes les tables (dans une configuration de banque de données particulière) pour tous les flux de données d'un job afin d'assurer la cohérence des données.

1. Faites glisser une table de banque de données CDM dans un flux de données.
2. Cliquez sur le nom de cet objet source pour ouvrir son éditeur.
3. Cliquez sur l'onglet *Options CDM*.
4. Vous pouvez sélectionner l'option facultative *Activer le point de contrôle*.

Une fois qu'un point de contrôle est placé, à la prochaine exécution du job CDM, celui-ci lit uniquement les lignes insérées dans la table CDM depuis le dernier point de contrôle.

Si les points de contrôle ne sont pas activés, le job CDM lit toutes les lignes de la source de données CDM (avec un temps de traitement augmenté).

34.8 Utilisation de la CDM avec des sources basées sur l'horodatage

Utilisez la CDM basée sur l'horodatage pour suivre les modifications si vous utilisez des sources autres que Oracle 9i, DB2 8.2, des ordinateurs centraux avec accès via IBM II Classic Federation ou via Attunity et si les conditions suivantes sont respectées.

- Il existe des champs date et heure dans les tables mises à jour
- Vous mettez à jour une table de grande taille ayant un faible pourcentage de modifications entre les extractions et un index sur les champs date et heure.
- Vous n'êtes pas concerné par la capture des résultats intermédiaires de chaque transaction entre les extractions (par exemple, si un client change de région deux fois dans la même journée).

Vous ne devez pas utiliser la CDM basée sur l'horodatage dans les cas suivants :

- Vous avez une table de grande taille, un pourcentage important de modifications entre les extractions et aucun index sur les horodatages.
- Vous avez besoin de capturer les suppressions physiques de lignes.
- Vous avez besoin de capturer plusieurs événements se produisant sur la même ligne entre les extractions.

Généralement, le terme horodatage fait référence aux valeurs de date, heure ou date et heure. Dans le cas de table CDM basée sur l'horodatage, la table source comporte des horodatages de `CREATION` ou `MISE À JOUR` pour chaque ligne.

Les horodatages peuvent indiquer si une ligne a été créée ou mise à jour. Certaines tables ont à la fois des horodatages de création et de mise à jour, certaines tables n'en ont qu'un. Dans le cas de CDM basée sur l'horodatage, les tables contiennent au moins un horodatage de mise à jour.

Certains systèmes ont des horodatages avec dates et heures, certains ont seulement la date et certains des nombres croissants générés de manière monotonique. Dans le cas de CDM basée sur l'horodatage, les dates et les nombres générés sont traités de la même façon.

i Remarque

Pour les horodatages basés sur le temps réel, les fuseaux horaires peuvent devenir importants. Si vous effectuez le suivi des horodatages utilisant la nomenclature du système source (c'est-à-dire utilisant l'heure de la source ou le nombre généré par la source), vous pouvez traiter de la même manière les horodatages temporels (heure spécifique) et logiques (heure relative par rapport à une autre heure ou un événement).

Informations associées

[Traitement des horodatages \[page 836\]](#)

[Chevauchements \[page 838\]](#)

[Types d'horodatages \[page 843\]](#)

34.8.1 Traitement des horodatages

La technique de base pour utiliser les horodatages afin de déterminer les modifications consiste à enregistrer l'horodatage ayant la valeur la plus élevée chargé dans un job donné et lancer le job suivant avec cet horodatage.

Pour ce faire, créez une table de statut qui suit les horodatages des lignes chargées dans un job. A la fin d'un job, `METTEZ À JOUR` cette table avec le dernier horodatage chargé. Le job suivant lit alors l'horodatage dans la table de statut et sélectionne uniquement les lignes de la source pour lesquelles l'horodatage est ultérieur à celui de la table de statut.

L'exemple suivant illustre cette technique. Supposons que le dernier chargement s'est produit à 14 h 00 le 1er janvier 1998. A cette heure, la table source comportait une seule ligne (clé=1) ayant un horodatage antérieur au chargement précédent. SAP Data Services charge cette ligne dans la table cible et met à jour la table de statut avec l'horodatage le plus élevé chargé : 13 h 10 le 1er janvier 1998. Après 14 h 00, le progiciel ajoute des lignes à la table source :

Table source

Clé	Données	Update_Timestamp
1	Alvarez	01/01/98 13:10
2	Tanaka	01/01/98 14:12
3	Lani	01/01/98 14:39

Table cible

Clé	Données	Update_Timestamp
1	Alvarez	01/01/98 13:10

Table de statut

Last_Timestamp
01/01/98 13:10

Le 1er janvier 1998 à 15 h 00, le job s'exécute à nouveau. Cette fois, le job effectue les actions suivantes :

1. Il lit le champ Last_Timestamp de la table de statut (01/01/98 13:10).
2. Il sélectionne les lignes de la table source dont les horodatages sont ultérieurs à la valeur du champ Last_Timestamp. La commande SQL pour sélectionner ces lignes est :

```
SELECT * FROM Source
WHERE 'Update_Timestamp' > '01/01/98 01:10 pm'
```

Cette opération renvoie les deuxième et troisième lignes (clé=2 et clé=3).

3. Il charge les nouvelles lignes dans la table cible.
4. Il met à jour la table de statut avec le dernier horodatage de la table cible (01/01/98 14:39) grâce à l'instruction SQL suivante :

```
UPDATE STATUS SET 'Last_Timestamp' = SELECT MAX('Update_Timestamp') FROM
target_table
```

La cible affiche les nouvelles données :

Table source

Clé	Données	Update_Timestamp
1	Alvarez	01/01/98 13:10

Clé	Données	Update_Timestamp
2	Tanaka	01/01/98 14:12
3	Lani	01/01/98 14:39

Table cible

Clé	Données	Update_Timestamp
2	Tanaka	01/01/98 14:12
3	Lani	01/01/98 14:39
1	Alvarez	01/01/98 13:10

Table de statut

Dernier_horodatage
01/01/98 14:39

Pour indiquer ces opérations, un flux de données a besoin des objets suivants (et suppose que toutes les métadonnées requises pour les tables source et cible ont été importées) :

- Un flux de données pour extraire les données modifiées de la table source et les charger dans la table cible, par exemple Source > Requête > Cible.
La requête sélectionne les lignes de SOURCE_TABLE à charger dans TARGET_TABLE.
Par exemple :

```
SOURCE.UPDATE_TIMESTAMP > $Last_Update
```

La requête comprend une clause WHERE pour filtrer les lignes entre les horodatages.

- Un workflow pour effectuer les actions suivantes :
 1. Lire la table de statut
 2. Définir la valeur d'une variable pour le dernier horodatage
 3. Appeler le flux de données avec la variable qui lui est passée comme paramètre
 4. Mettre à jour la table de statut avec le nouvel horodatage
- Un job pour exécuter le workflow

34.8.2 Chevauchements

Sauf si les données source sont rigoureusement isolées durant le processus d'extraction (ce qui n'est habituellement pas réalisable), il existe une fenêtre de temps où les modifications peuvent être perdues entre deux cycles d'extraction. Cette période de chevauchement affecte la capture des données modifiées basée sur l'horodatage parce que ce type de capture de données repose sur un horodatage statique pour déterminer les données modifiées.

Supposons par exemple qu'une table comporte 1 000 lignes (classées de 1 à 1 000). Le job commence avec l'horodatage 3:00 et extrait chaque ligne. Pendant que le job s'exécute, il met à jour deux lignes (1 et 1 000)

ayant respectivement les horodatages 3:01 et 3:02. Le job extrait la ligne 200 pendant que quelqu'un met à jour la ligne 1. Lorsque le job extrait la ligne 300, il met à jour la ligne 1 000. Lorsqu'il est terminé, le job extrait le dernier horodatage (3:02) de la ligne 1 000, mais omet la mise à jour de la ligne 1.

Voilà les données de la table :

Numéro de ligne	Colonne A
1	...
2	...
3	...
...	...
200	...
...	...
600	...
...	...
1 000	...

Il existe trois techniques pour gérer cette situation :

- Evitement des chevauchements
- Rapprochement des chevauchements
- Pré-échantillonnage

Les sections suivantes décrivent ces techniques et leur implémentation dans SAP Data Services. Cette section se poursuit avec l'hypothèse qu'il y a au moins un horodatage de mise à jour.

Informations associées

[Evitement des chevauchements \[page 839\]](#)

[Rapprochement des chevauchements \[page 840\]](#)

[Pré-échantillonnage \[page 840\]](#)

[Types d'horodatages \[page 843\]](#)

34.8.2.1 Evitement des chevauchements

Dans certains cas, il est possible de configurer un système où il n'existe pas de possibilité de chevauchement. Vous pouvez éviter les chevauchements s'il existe un intervalle de traitement où aucune mise à jour ne se produit dans le système cible.

Par exemple, si vous pouvez garantir que l'extraction de données du système source ne dure pas plus d'une heure, vous pouvez exécuter à 1 h 00 chaque nuit un job qui sélectionne uniquement les données mises à jour avant minuit le jour précédent. Comme ce job régulier ne donne pas de mises à jour minute par minute, il garantit de ne jamais avoir de chevauchement et simplifie nettement la gestion de l'horodatage.

34.8.2.2 Rapprochement des chevauchements

Le rapprochement des chevauchements requiert un processus d'extraction spécial qui réapplique les modifications pouvant s'être produites pendant la période de chevauchement. Cette extraction peut être exécutée séparément de l'extraction normale. Par exemple, si l'horodatage le plus élevé chargé par le job précédent était 01/01/98 22:30 et que la période de chevauchement dure une heure, le rapprochement des chevauchements réapplique les données mises à jour entre 21 h 30 et 22 h 30 le 1er janvier 1998.

La période de chevauchement est habituellement égale à la plus longue durée possible de l'extraction. Si l'extraction des données du système source peut durer jusqu'à `<n>` heures, une période de chevauchement de `<n>` heures (ou `<n>` plus un petit incrément) est conseillée. Par exemple, si l'exécution du job dure au plus deux heures, une période de chevauchement d'au moins deux heures est recommandée.

La création d'un flux de données de chevauchement distinct présente un avantage. Un flux de données "normal" peut supposer que toutes les modifications sont nouvelles et faire des hypothèses pour simplifier les calculs et améliorer les performances. Par exemple, les lignes marquées comme INSEREES sont souvent chargées dans une table de faits, mais les lignes marquées comme MISES A JOUR le sont rarement. Donc, le flux de données normal sélectionne les nouvelles lignes de la source, génère de nouvelles clés pour elles et utilise le chargeur de base de données pour ajouter les nouveaux faits à la base de données cible. Comme le flux de données de chevauchement est susceptible d'appliquer à nouveau les mêmes lignes, il ne peut pas les charger par lots aveuglément, sinon il crée des doublons. Par conséquent, le flux de données de chevauchement doit contrôler si les lignes existent dans la cible et insérer uniquement celles qui manquent. Cette recherche affecte les performances, donc exécutez-la pour le plus petit nombre de lignes possible.

Si le volume de données est assez faible, vous pouvez charger l'ensemble du nouveau jeu de données à l'aide de cette technique de contrôle avant chargement, ce qui évite d'avoir à créer deux flux de données distincts.

34.8.2.3 Pré-échantillonnage

Le pré-échantillonnage élimine le chevauchement en identifiant d'abord l'horodatage le plus récent du système, en l'enregistrant, puis en extrayant les lignes jusqu'à cet horodatage.

Cette technique est une extension du traitement d'horodatage simple. Principale différence : la table de statut contient désormais un horodatage de début et de fin. L'horodatage de début est le dernier extrait par le job précédent. L'horodatage de fin est celui sélectionné par le job en cours.

Revenons à l'exemple : le dernier job d'extraction a chargé des données de la table source dans la table cible et mis à jour la table de statut avec le dernier horodatage chargé :

Table source

Clé	Données	Update_Timestamp
1	Alvarez	01/01/98 13:10
2	Tanaka	01/01/98 14:12
3	Lani	01/01/98 14:39

Table cible

Clé	Données	Update_Timestamp
1	Alvarez	01/01/98 13:10

Table de statut

Start_Timestamp	End_Timestamp
01/01/98 13:10	NULL

Il est maintenant 15 h 00 le 1er janvier 1998 et le job suivant s'exécute, il effectue les actions suivantes :

1. Il sélectionne l'horodatage le plus récent de la table source et l'insère dans la table de statut comme horodatage de fin.

La commande SQL pour sélectionner une ligne est :

```
SELECT MAX(Update_Timestamp) FROM source table
```

La table de statut devient :

Table de statut

Start_Timestamp	End_Timestamp
01/01/98 13:10	01/01/98 14:39

1. Il sélectionne les lignes de la table source dont les horodatages sont supérieurs à l'horodatage de début, mais inférieurs ou égaux à l'horodatage de fin. La commande SQL pour sélectionner ces lignes est :

```
SELECT *
FROM source table
WHERE Update_Timestamp > '1/1/98 1:10pm'
AND Update_Timestamp <= '1/1/98 2:39pm'
```

Cette opération renvoie les deuxième et troisième lignes (clé=2 et clé=3)

2. Il charge les nouvelles lignes dans la table cible.
3. Met à jour la table de statut en définissant l'horodatage de début comme étant le précédent horodatage de fin et en donnant à l'horodatage de fin la valeur NULL.

Au final, les valeurs des tables sont les suivantes :

Table source

Clé	Données	Update_Timestamp
1	Alvarez	01/01/98 13:10
2	Tanaka	01/01/98 14:12
3	Lani	01/01/98 14:39

Table cible

Clé	Données	Update_Timestamp
1	Alvarez	01/01/98 13:10
2	Tanaka	01/01/98 14:12
3	Lani	01/01/98 14:39

Table de statut

Start_Timestamp	End_Timestamp
01/01/98 14:39	NULL

Pour développer l'exemple précédent afin de prendre en compte le temps de chevauchement, il faut apporter les modifications suivantes au workflow :

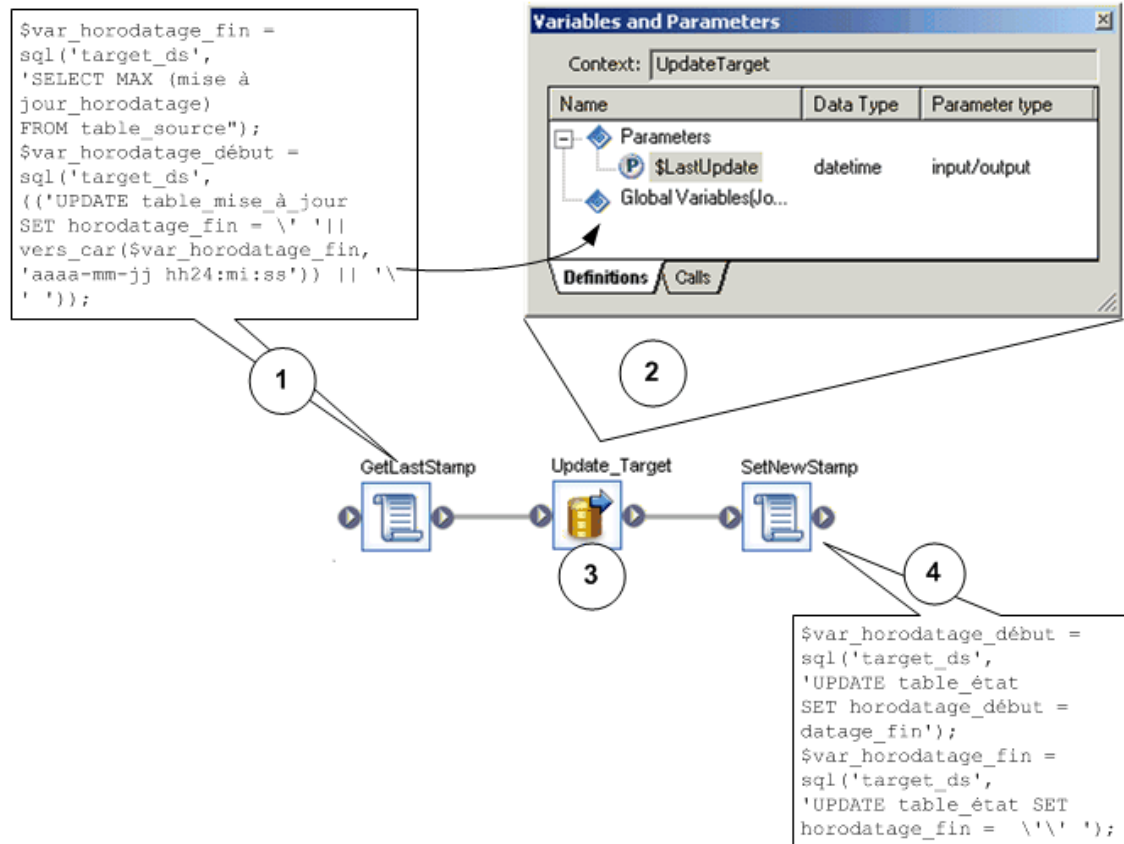
- Un flux de données pour extraire les modifications depuis la dernière mise à jour et avant l'horodatage le plus récent, par exemple Source > Requête > Cible.
La requête sélectionne les lignes de SOURCE_TABLE à charger dans TARGET_TABLE.
Par exemple :

```
SOURCE.UPDATE_TIMESTAMP > $start_last_update and
SOURCE.UPDATE_TIMESTAMP < $end_last_update
```

La requête comprend une clause WHERE pour filtrer les lignes entre les horodatages.

- Un workflow pour effectuer les actions suivantes :
 1. Lire la table source pour rechercher l'horodatage le plus récent.
 2. Définir la valeur de deux variables respectivement sur le début et la fin du temps de chevauchement.
 3. Appeler le flux de données avec les variables qui lui sont passées comme paramètres.
 4. Mettre à jour l'horodatage de début avec la valeur de l'horodatage de fin et donner à l'horodatage de fin la valeur NULL.

Workflow : Données modifiées avec horodatages



Informations associées

[Traitement des horodatages \[page 836\]](#)

34.8.3 Types d'horodatages

Certains systèmes ont des horodatages qui s'enregistrent uniquement lors de la création de lignes. D'autres ont des horodatages qui s'enregistrent uniquement lors de la mise à jour de lignes. (Habituellement, les systèmes avec horodatage à la mise à jour seulement activent l'horodatage de mise à jour lorsque la ligne est créée ou mise à jour.) Enfin, il existe des systèmes qui gèrent des horodatages distincts qui enregistrent la création et la mise à jour des lignes.

Informations associées

[Horodatages à la création seulement \[page 844\]](#)

[Horodatages à la mise à jour seulement \[page 844\]](#)

[Horodatages de création et mise à jour \[page 844\]](#)

34.8.3.1 Horodatages à la création seulement

Si le système source fournit des horodatages à la mise à jour seulement, vous disposez de ces options :

- Si la table est suffisamment petite, vous pouvez la traiter en intégralité pour identifier les modifications. La section [Utilisation de la CDM pour les cibles \[page 851\]](#) décrit comment identifier les modifications.
- Si la table n'est jamais mise à jour, vous pouvez extraire uniquement les nouvelles lignes.
- Si la table est importante et mise à jour, vous pouvez combiner les deux techniques suivantes :
 - Extraction périodique (par exemple, journalière) des nouvelles lignes uniquement.
 - Extraction moins fréquente (par exemple, hebdomadaire) des lignes mises à jour en traitant la totalité de la table.

34.8.3.2 Horodatages à la mise à jour seulement

L'utilisation d'un horodatage à la mise à jour seulement aide à minimiser l'impact sur les systèmes source, mais cela complique le chargement des systèmes cible. Si le système propose un horodatage à la mise à jour seulement et qu'il n'existe pas de moyen d'indiquer les nouvelles lignes par rapport aux lignes mises à jour, votre job doit rapprocher le nouveau jeu de données des données existantes à l'aide des techniques décrites dans la section [Utilisation de la CDM pour les cibles \[page 851\]](#).

34.8.3.3 Horodatages de création et mise à jour

Les deux horodatages permettent de séparer facilement les nouvelles données des mises à jour de données existantes. Le job extrait toutes les lignes modifiées, puis filtre les lignes indésirables à l'aide de leurs horodatages.

Accomplissez ces extractions dans le progiciel en ajoutant la clause WHERE des commandes SQL suivantes à une transformation Query appropriée :

- Recherche de nouvelles lignes :

```
SELECT * FROM source_table
WHERE Create_Timestamp > $Last_Timestamp
```

- Recherche de lignes mises à jour :

```
SELECT * FROM source_table
WHERE Create_Timestamp <= $Last_Timestamp AND
Update_Timestamp > $Last_Timestamp)
```

De là, les nouvelles lignes sont soumises au processus de génération de clés et sont insérées dans la cible. Les lignes mises à jour passent par le processus de recherche de clé et sont mises à jour dans la cible.

Pour des raisons de performances, vous pouvez éventuellement isoler l'extraction des nouvelles lignes dans un flux de données distinct pour tirer parti du chargement par lots dans la cible. Les lignes mises à jour ne peuvent pas être chargées par lots dans la même cible au même moment.

34.8.4 Timestamp-based CDC examples

34.8.4.1 Conservation des clés générées

Pour des raisons de performances, la plupart des tables de dimension des entrepôts de données utilisent des clés générées pour effectuer la jointure avec la table des faits. Par exemple, le client ABC a la clé générée 123 dans la table de la dimension Client. Tous les faits du client ABC ont 123 comme clé client. Même si la dimension Client est de petite taille, vous ne pouvez pas simplement la recharger chaque fois qu'un enregistrement est modifié : à moins d'affecter la clé générée 123 au client ABC, la table de la dimension Client et la table des faits ne sont pas en corrélation.

Vous pouvez conserver les clés générées en utilisant la fonction de recherche ou en comparant les tables.

Informations associées

[Utilisation de la fonction lookup \[page 845\]](#)

[Comparaison de tables \[page 847\]](#)

34.8.4.1.1 Utilisation de la fonction lookup

Si la conservation de l'historique ne présente pas de problème et que le seul objectif est de générer les bonnes clés pour les lignes existantes, la technique la plus simple consiste à rechercher la clé de toutes les lignes à l'aide de la fonction `lookup` dans une requête. Si vous ne recherchez pas la clé, générez-en une nouvelle.

Dans l'exemple suivant, la table de la dimension Client contient des clés générées. Lorsque vous exécutez un job pour mettre à jour cette table, les lignes source des clients doivent correspondre aux clés existantes.

Table source des clients

Nom de société	ID client
ABC	001
DEF	002
GHI	003
JKL	004

Table de dimension cible

Gen_Key	Nom de société	ID client
123	ABC	001
124	DEF	002
125	GHI	003

Cet exemple de flux de données effectue les actions suivantes :

1. Il extrait les lignes source.
2. Il récupère les clés existantes à l'aide d'une fonction `lookup` dans le mappage d'une nouvelle colonne dans une requête.
3. Il charge le résultat dans un fichier (pour pouvoir tester cette phase du flux de données avant d'ajouter les étapes suivantes).

La fonction `lookup` compare les lignes source à la cible. Les arguments de la fonction sont les suivants :

arguments de la fonction de recherche	Description
<code>target_ds.owner.customer</code>	Nom qualifié complet de la table cible contenant les clés générées.
GKey	Nom de la colonne de la table cible contenant les clés générées.
NULL	Valeur NULL à insérer dans la colonne des clés si aucune clé existante n'est trouvée.
CACHE_DE_PRECHARGEMENT	Option de mise en cache pour optimiser les performances de la recherche.
Customer_ID	Colonne de la table cible contenant la valeur à utiliser pour faire correspondre les lignes.
Customer_ID	Colonne de la table source contenant les valeurs à utiliser pour faire correspondre les lignes.

Le jeu de données résultant contient toutes les lignes de la source comportant les clés générées si elles sont disponibles.

Jeu de données du résultat

Gen_Key	Nom de société	ID client
123	ABC	001
124	DEF	002
125	GHI	003

Gen_Key	Nom de société	ID client
NULL	JKL	004

L'ajout d'une clé générée aux nouveaux enregistrements requiert le filtrage des nouvelles lignes à partir des lignes existantes et mises à jour. Dans le flux de données, les étapes suivantes sont nécessaires : une requête pour sélectionner les lignes avec des clés générées de valeur NULL. Une transformation Key_Generation pour déterminer la clé appropriée à ajouter. Un cible pour charger les nouvelles lignes dans la table de la dimension Client.

Ce flux de données gère les nouvelles lignes. Cependant, les lignes de la source dont les clés ont été trouvées dans la table cible peuvent contenir des données mises à jour. Comme cet exemple suppose que la conservation de l'historique n'est pas une exigence, le progiciel charge toutes les lignes de la source dans la cible.

34.8.4.1.1.1 Gestion des lignes mises à jour dans le flux de données

Le flux de données requiert de nouvelles étapes pour gérer les lignes mises à jour, comme suit :

1. Une nouvelle ligne quittant la requête qui a recherché les clés existantes.
2. Une requête pour filtrer les lignes comportant des clés existantes des lignes sans clés.
3. Un cible pour charger les lignes dans la table de la dimension Client.

34.8.4.1.2 Comparaison de tables

Inconvénient de la méthode des clés générées, même si la ligne n'a pas été modifiée, elle génère une opération UPDATE et elle est chargée dans la cible. Si le nombre de lignes est important, une transformation Table_Comparison présente une meilleure alternative en permettant au flux de données de charger uniquement les lignes modifiées.

La transformation Table_Comparison examine toutes les lignes sources et effectue les opérations suivantes :

- Elle génère une ligne INSERT pour toute nouvelle ligne absente de la table cible.
- Elle génère une ligne UPDATE pour toute ligne de la table cible ayant été modifiée.
- Elle ignore toute ligne de la table cible n'ayant pas été modifiée.
- Elle remplit les clés générées pour les lignes mises à jour.

Vous pouvez ensuite soumettre le résultat à une transformation Key Generation pour affecter une nouvelle clé à chaque ligne INSERT. C'est le jeu de données chargé dans la table cible par le progiciel.

Le flux de données qui applique cette transformation comprend les étapes suivantes :

1. Une source pour extraire les lignes des tables source.
2. Une requête pour mapper les colonnes de la source.
3. Une transformation Table_Comparison pour générer les lignes INSERT et UPDATE et renseigner les clés existantes.

4. Une transformation Key Generation pour générer les nouvelles clés.
5. Un cible pour charger les lignes dans la table de la dimension Client.

34.8.4.2 Conservation de l'historique

La transformation History Preserving permet à l'entrepôt de données ou au datamart de gérer l'historique des données pour pouvoir les analyser au fil du temps. Vous appliquerez probablement la conservation de l'historique aux tables de dimension.

Par exemple, si un client déménage d'une région de vente dans une autre, la simple mise à jour de la fiche client pour tenir compte de la nouvelle région donne des résultats trompeurs dans une analyse par région dans le temps, parce que les achats effectués par ce client avant son déménagement seront attribués de manière erronée à la nouvelle région.

SAP Data Services fournit une transformation spéciale qui conserve l'historique des données pour éviter ce genre de situation. La transformation History_Preserving ignore tout, sauf les lignes marquées comme UPDATE. Pour ces lignes, elle compare les valeurs des colonnes indiquées et marque la ligne comme INSERT si les valeurs ont été modifiées. Cela génère une deuxième ligne dans la cible au lieu de remplacer la première ligne.

Pour développer la manière dont le progiciel gère l'exemple du client qui déménage d'une région à une autre :

- Si `Région` est une colonne marquée pour la comparaison, la transformation History_Preserving génère une nouvelle ligne pour ce client.
- Une transformation History_Preserving attribue à la nouvelle ligne une clé générée et la charge dans la table de la dimension Client.
- La ligne d'origine qui décrit le client reste dans la table de la dimension Client avec une clé unique générée.

Dans l'exemple suivant, un client a déménagé de la région Est dans la région Ouest et le numéro de téléphone d'un autre client a changé.

Table Client source

Client	Région	Téléphone
Café Chez Fred	Est	(212) 123-4567
Les beignets de Jane	Ouest	(650) 222-1212
Confiserie Sandy	Centre	(115) 231-1233

Table Client cible

GKey	Client	Région	Téléphone
1	Café Chez Fred	Est	(212) 123-4567
2	Les beignets de Jane	Est	(201) 777-1717
3	Confiserie Sandy	Centre	(115) 454-8000

Dans cet exemple, le flux de données conserve l'historique pour la colonne `Région`, mais pas pour la colonne `Téléphone`. Le flux de données comporte les étapes suivantes :

1. Une source pour extraire les lignes des tables source.
2. Une requête pour mapper les colonnes de la source.
3. Une transformation Table_Comparison pour générer les lignes INSERT et UPDATE et renseigner les clés existantes.
4. Une transformation History_Preserving pour convertir certaines lignes UPDATE en lignes INSERT.
5. Une transformation Key Generation pour générer de nouvelles clés pour les lignes mises à jour, maintenant marquées comme INSERT.
6. Un cible pour charger les lignes dans la table de la dimension Client.

Il y a maintenant deux lignes pour Les beignets de Jane, les corrélations entre la table de dimension et la tables des faits doivent utiliser la valeur de clé la plus élevée.

Notez que les mises à jour des colonnes sans conservation de l'historique mettent à jour toutes les versions de la ligne si la mise à jour est effectuée sur la clé naturelle (par exemple, `Client`), et actualisent uniquement la dernière version si la mise à jour porte sur la clé générée (par exemple, `GKey`). Vous pouvez contrôler la clé à utiliser pour la mise à jour en configurant correctement les options de chargement dans l'éditeur de la cible.

34.8.4.2.1 date de début_validité et date de fin_validité

Pour prendre en charge des requêtes temporelles telles que "Quelle était l'adresse de facturation du client au 24 mai 1998", SAP Data Services prend en charge les colonnes de date de *Début de validité* et *Fin de validité*.

Dans les techniques de conservation de l'historique, plusieurs enregistrements sont définis dans la table cible avec les mêmes valeurs clés primaires source. Un enregistrement de la table source est considéré valide dans la table de dimensions pour toutes les valeurs de date t , de sorte que la date de *Début de validité* soit antérieure ou identique à t , qui est antérieur à la date de *Fin de validité*. (Valide dans ce sens signifie que la valeur clé d'enregistrement générée est utilisée pour charger la table de faits au cours de cet intervalle de temps.)

Lorsque vous indiquez les entrées *Début de validité* et *Fin de validité*, la transformation History_Preserving génère une mise à jour d'enregistrement et une instruction INSERT pour des raisons de conservation de l'historique (elle convertit une instruction UPDATE en une instruction INSERT). La mise à jour d'enregistrement activera la colonne de date *Fin de validité* pour l'enregistrement en cours (l'une avec la même clé primaire comme instruction INSERT) sur la valeur figurant dans la colonne de date *Début de validité* dans l'enregistrement INSERT.

34.8.4.2.2 Indicateur de mise à jour

Pour prendre en charge les techniques de dimensions évoluant lentement, Data Services permet d'activer un indicateur de mise à jour pour marquer l'enregistrement actuel dans une table de dimensions.

La valeur *Définir la valeur* dans la colonne *Colonne* identifie l'enregistrement valide actuel dans la table cible pour une clé primaire de table source donnée.

Lorsque vous indiquez *Colonne*, la transformation History_Preserving génère une mise à jour de l'enregistrement avant de générer une instruction INSERT.

Cette mise à jour de l'enregistrement active la valeur *Colonne* sur *Réinitialiser la valeur* dans l'enregistrement de table cible avec la même clé primaire source comme instruction INSERT.

Dans l'instruction INSERT, la *colonne* sera activée sur *Définir la valeur*.

Lorsque vous indiquez des entrées dans les deux groupes, la transformation History_Preserving génère une seule instruction UPDATE supplémentaire pour chaque instruction INSERT créée. Cette instruction UPDATE met à jour la valeur *Fin de validité*.

34.8.5 Autres astuces de conception de job

Lors de la conception d'un job destiné à implémenter la capture des données modifiées (CDM), vous devez tenir compte de :

- Synchronisation de l'en-tête et des détails
- Capture des suppressions physiques

34.8.5.1 Synchronisation de l'en-tête et des détails

Habituellement, les systèmes source effectuent le suivi des modifications des informations de l'en-tête et des détails de manière indépendante. Par exemple, si le statut d'un poste individuel est modifié, sa colonne "Date de dernière modification" se met à jour, contrairement à la même colonne au niveau de l'en-tête de la commande. À l'inverse, une modification de l'adresse par défaut du destinataire dans l'en-tête de la commande peut n'impacter aucun des postes individuels existants.

Toutefois, dans certains cas, le système source peut ne pas mettre à jour ces colonnes de suivi de manière cohérente, ou bien vous pouvez ne pas avoir accès à ces informations (par exemple, si les lignes sont physiquement supprimées). Dans ces cas, vous pouvez choisir d'extraire toutes les informations de l'en-tête et des détails, que les modifications soient au niveau de l'en-tête ou d'un poste individuel.

Pour extraire toutes les lignes de l'en-tête et des détails si un de ces éléments a été modifié, utilisez un calcul similaire à cette instruction SQL :

```
SELECT ... FROM HEADER, DETAIL WHERE HEADER.ID = DETAIL.ID AND (HEADER.LAST_MODIFIED  
BETWEEN $G_SDATE AND $G_EDATE OR DETAIL.LAST_MODIFIED BETWEEN $G_SDATE AND $G_EDATE)
```

Pour certaines bases de données, cette clause WHERE n'est pas bien optimisée et peut provoquer une sérieuse dégradation des performances. Vous pouvez opter pour une simplification de cette clause en supprimant une des bornes supérieures, comme dans :

```
... WHERE HEADER.ID = DETAIL.ID AND (HEADER.LAST_MODIFIED BETWEEN $G_SDATE AND $G_EDATE OR  
DETAIL.LAST_MODIFIED >= $G_SDATE) ...
```

Le nombre de lignes extraites peut être un peu supérieur à ce qui était prévu à l'origine, mais les performances finales du système peuvent être améliorées tout en ne modifiant pas le résultat de la base de données cible.

34.8.5.2 Capture des suppressions physiques

Lorsque le système source permet de supprimer physiquement des lignes, votre job doit comprendre un calcul pour mettre à jour la base de données cible en conséquence. Plusieurs méthodes le permettent :

- Lecture d'un journal des opérations — Si votre système journalise les transactions dans un format lisible ou si vous pouvez modifier le système pour qu'il génère un tel journal, vous pouvez lire ce journal pour identifier les lignes à supprimer.
- Exécution d'une actualisation complète — Rechargez simplement l'ensemble des données, ce qui synchronise entièrement le système source et la base de données cible.
- Exécution d'une actualisation partielle sur la base d'une fenêtre temporelle pilotée par les données — Par exemple, supposons que le système source permet uniquement la suppression physique des commandes n'ayant pas été clôturées. Si la première commande non clôturée de la table source a eu lieu il y a six mois, l'actualisation des six derniers mois de données garantit la réussite de la synchronisation.
- Exécution d'une actualisation partielle sur la base d'une fenêtre temporelle pilotée par l'entreprise — Par exemple, supposons que l'entreprise prise en charge par le job supprime habituellement les commandes peu de temps après les avoir créées. Dans ce cas, l'actualisation du dernier mois de commandes est appropriée pour conserver l'intégrité.
- Contrôle de chaque commande pouvant éventuellement être supprimée — Vous devez vérifier si toute commande non clôturée a été supprimée. Pour être efficace, cette technique requiert de conserver un enregistrement des clés primaires pour chaque objet candidat à la suppression.

Lorsque les suppressions physiques des informations détaillées d'une relation en-tête/détails sont possibles (par exemple, suppression des postes individuels d'une commande existante), vous devez capturer ces suppressions physiques lors de la synchronisation des informations de l'en-tête et des détails.

34.9 Utilisation de la CDM pour les cibles

La capture des données modifiées basée sur la source est pratiquement toujours préférable à la capture basée sur la cible pour des raisons de performances.

Cependant, certains systèmes source ne fournissent pas assez d'informations pour utiliser les techniques de capture des données modifiées basée sur la source. La capture des données modifiées basée sur la cible permet d'utiliser cette technique lorsque les informations de modification basées sur la source sont limitées.

35 Data Services data types

SAP Data Services supports several specific data types.

The following table provides a list of the data types that Data Services supports with a brief description. For complete information about the data types, see the *Reference Guide*.

Supported data types

Data type	Description
blob	Stores binary data. Use for multimedia data such as images, audio, and video. Do not use columns that contain blob data types in comparisons, calculations, or data type conversions. Blob data type has many of the same limitations as long data type.
date	Defines calendar dates. Specify date formats by generating strings using certain date format codes. For example, DD = 2-digit day of the month. MM = 2-digit month number.
datetime	Defines calendar dates and times. Behaves like a concatenation of date and time data items. Specify date rules for the date part and time rules for the time part. For example: dd:mm:yy hh24:mi:ss.
decimal	Defines exact decimal numbers. Specify the precision and scale when you specify a decimal data type: • Precision: Total number of digits in the value Use the following relation: $1 \leq \text{precision} \leq 96$ • Scale: Number of digits to the right of the decimal point. Use the following relation: $0 \leq \text{scale} \leq \text{precision}$
double	Defines an 8-byte floating point value. The 8-byte floating point value has radix, exponent range, and precision of the platform on which Data Services is running.

Data type	Description
int	Defines a 4-byte signed binary integer. Can have a plus or minus sign that indicates a positive or negative value. Sign can appear before or after the value with any number of blanks between the value and the sign. When the number does not have a plus or minus sign, the software considers it as positive.
interval	Defines differences between dates or times. Value is in days unless you specify another unit. Use the following conversion functions to make interval values accessible: <code>interval_to_char</code> <code>num_to_interval</code>
long	Stores character data of variable length up to 2 gigabytes. Use to represent character-based large objects or "clob". Converts several other database-specific large object types such as longvarchar and text, to the long data type. Supports mapping from all databases. Long data type has many of the same limitations as blob data types.
numeric	Identical to the decimal data type.
real	Defines a 4-byte floating point value with radix, exponent range, and precision of the platform on which Data Services is running. Can have a plus or minus sign that indicates a positive or negative value. Sign can appear before or after the value with any number of blanks between the value and the sign. When the number does not have a plus or minus sign, the software considers it as positive. Databases store real values as a 32-bit approximation of the number. The approximation makes comparison results unpredictable when you use a real value in an equality or inequality comparison. Therefore, we recommend that you do not use a real value in a <code>WHERE</code> clause.

Data type	Description
time	Defines times of the day with no calendar date. Manages time operations in the format used by the database manager. Conversion operations to or from time require you to specify the format of the time value. Specify a time format by generating a string using codes and literal strings or punctuation. For example HH24 = 2-digit hour of the day.
timestamp	Supports the Oracle 9i timestamp data type with no zone information. Incorporates up to a 9-digit subsecond. Add or subtract timestamp values. The resulting data type from addition or subtraction operations depend on the operation and data types involved.
varchar	Can contain letters and numbers. Indicate the length when specifying a varchar data type.

36 Conversion de types de données

Cette section explique comment Data Service traite les conversions de différents types de données lors des opérations arithmétiques et entre types de données.

[Calcul de date \[page 855\]](#)

[Types de données non pris en charge \[page 856\]](#)

[Conversion de/en types de données internes \[page 856\]](#)

[Conversion de types de données dans des expressions \[page 883\]](#)

[Conversion parmi des types de données Number \[page 884\]](#)

[Conversions entre types de données explicites \[page 888\]](#)

[Conversion entre des types de données natifs \[page 888\]](#)

36.1 Calcul de date

Data Service effectue des conversions implicites de types de données sur des valeurs de date, heure, date/heure, horodatage et intervalle lors du calcul de date. Le tableau suivant décrit ces conversions :

Opération	Type de renvoi
DATE + INTERVALLE	DATE
HEURE + INTERVALLE	TIME
DATE/HEURE + INTERVALLE	DATE/HEURE
INTERVALLE + INTERVALLE	INTERVAL
DATE - DATE	INTERVAL
DATE - INTERVALLE	DATE
HEURE - HEURE	INTERVAL
HEURE - INTERVALLE	TIME
INTERVALLE - INTERVALLE	INTERVAL
DATE/HEURE - DATE	INTERVAL
DATE/HEURE - HEURE	INTERVAL
DATE/HEURE - DATE/HEURE	INTERVAL
DATE/HEURE - INTERVALLE	DATETIME
HORODATAGE + HORODATAGE	INTERVAL
HORODATAGE - HORODATAGE	INTERVAL

Opération	Type de renvoi
HORODATAGE + INTERVALLE	TIMESTAMP
HORODATAGE - INTERVALLE	TIMESTAMP

36.2 Types de données non pris en charge

Data Services peut lire, charger et appeler des procédures stockées faisant intervenir des types de données inconnus, à condition que vos serveurs de base de données effectuent la conversion de VARCHAR vers le type de données (inconnu) natif et du type de données (inconnu) natif vers VARCHAR. Data Services peut rencontrer un problème en chargeant VARCHAR dans une colonne CLOB physique si la base de données native ne prend pas en charge cette conversion (le chargement par lots ou la correction automatique du chargement peut échouer, par exemple).

Lorsque le logiciel détecte une colonne affectée à un type de données non pris en charge, il n'importe pas les métadonnées pour la colonne et indique une erreur. Le fichier `errorlog.txt` contient une entrée indiquant la colonne ignorée. Pour inclure la colonne dans votre job, convertissez le type de données dans un type pris en charge par le logiciel avant d'importer les métadonnées pour la table.

i Remarque

Utilisez la fonction `varchar_to_long` pour convertir un type de données VARCHAR en un type de données LONG avant de charger une colonne CLOB physique. Si vous utilisez une colonne VARCHAR dans le schéma physique pour le chargement suite à une installation antérieure, cela continuera de fonctionner.

36.3 Conversion de/en types de données internes

Data Services convertit les types de données lorsqu'il importe des métadonnées de sources ou cibles externes dans le référentiel et lorsqu'il charge des données dans une table ou un fichier externe. Le logiciel utilise ses propres fonctions de conversion et non celles spécifiques à la base de données ou à l'application source des données.

De plus, si vous utilisez un modèle de table ou une table `Data_Transfer` comme cible, le logiciel convertit les types de données internes vers les types du système de gestion de bases de données respectif.

[Flux Attunity \[page 857\]](#)

[Copybook Cobol \[page 858\]](#)

[Data Federator \[page 860\]](#)

Le logiciel convertit les types de données de Data Federator lorsque vous importez des tables de Data Federator.

[IBM DB2 \[page 860\]](#)

[Informix \[page 863\]](#)

[Microsoft Excel \[page 864\]](#)

[Microsoft SQL Server \[page 865\]](#)

[MySQL \[page 868\]](#)

[Netezza \[page 870\]](#)

[ODBC \[page 871\]](#)

[Oracle \[page 873\]](#)

[SQL Anywhere \[page 875\]](#)

[SAP ASE \[page 877\]](#)

[SAP Sybase IQ \[page 879\]](#)

[Teradata \[page 881\]](#)

36.3.1 Flux Attunity

Le tableau suivant illustre la conversion de types de données Attunity Streams en types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible Attunity Streams dans le référentiel.

Type de données Attunity Streams	Convertit en type de données Data Services.
bigint	double
i Remarque Comme le type de données int (entier) contient seulement quatre octets, les données sont perdues lors de la conversion.	
double	double
float	double
long varchar	Converti en Long uniquement si importé à l'aide d'une banque de données ODBC.
real	real
date	datetime
decimal	decimal
integer	int
smallint	int
time	time
timestamp	datetime
varchar	varchar

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Copybook Cobol \[page 858\]](#)

[Data Federator \[page 860\]](#)

[IBM DB2 \[page 860\]](#)

[Informix \[page 863\]](#)

[Microsoft Excel \[page 864\]](#)

[Microsoft SQL Server \[page 865\]](#)

[MySQL \[page 868\]](#)

[Netezza \[page 870\]](#)

[ODBC \[page 871\]](#)

[Oracle \[page 873\]](#)

[SQL Anywhere \[page 875\]](#)

[SAP ASE \[page 877\]](#)

[SAP Sybase IQ \[page 879\]](#)

[Teradata \[page 881\]](#)

36.3.2 Copybook Cobol

Le tableau suivant illustre la conversion de types de données de copybook Cobol en types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible de copybook Cobol dans le référentiel.

Type de données de copybook Cobol	Convertit en type de données Data Services.
computational	decimal
comp	decimal
computational-1	float
comp-1	float
computational-2	double
comp-2	double
computational-3	decimal
comp-3	decimal
computational-4	decimal
comp-4	decimal
computational-5	decimal
comp-5	decimal
computational-6	decimal
comp-6	decimal

Type de données de copybook Cobol	Convertit en type de données Data Services.
computational-x	decimal
comp-x	decimal
binary	decimal
packed-decimal	decimal
float	float
double	double
signed-short	int
unsigned-short	int
signed-int	int
unsigned-int	int
signed-long	int
unsigned-long	int
integer	int
DISPLAY, PICTURE contient A ou X (données de type Character)	varchar
Virgule flottante comme +9.9(5)E+99	double

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)

[Data Federator \[page 860\]](#)

[IBM DB2 \[page 860\]](#)

[Informix \[page 863\]](#)

[Microsoft Excel \[page 864\]](#)

[Microsoft SQL Server \[page 865\]](#)

[MySQL \[page 868\]](#)

[Netezza \[page 870\]](#)

[ODBC \[page 871\]](#)

[Oracle \[page 873\]](#)

[SQL Anywhere \[page 875\]](#)

[SAP ASE \[page 877\]](#)

[SAP Sybase IQ \[page 879\]](#)

[Teradata \[page 881\]](#)

36.3.3 Data Federator

Le logiciel convertit les types de données de Data Federator lorsque vous importez des tables de Data Federator.

Conversions de types de données

Data Federator	Data Services
Décimal	Précision et échelle décimale (28,6)
VarChar	VarChar (1024)

Il est possible de modifier la précision ou l'échelle décimale et la taille varchar dans le logiciel après l'importation depuis la source de données Data Federator.

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)
[Copybook Cobol \[page 858\]](#)
[IBM DB2 \[page 860\]](#)
[Informix \[page 863\]](#)
[Microsoft Excel \[page 864\]](#)
[Microsoft SQL Server \[page 865\]](#)
[MySQL \[page 868\]](#)
[Netezza \[page 870\]](#)
[ODBC \[page 871\]](#)
[Oracle \[page 873\]](#)
[SQL Anywhere \[page 875\]](#)
[SAP ASE \[page 877\]](#)
[SAP Sybase IQ \[page 879\]](#)
[Teradata \[page 881\]](#)

36.3.4 IBM DB2

Le tableau suivant illustre la conversion entre des types de données DB2 et des types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible DB2 dans le référentiel et lorsqu'il charge des données dans une table ou un fichier externe.

Type de données DB2	Convertit de/en type de données Data Services.
bigint	int i Remarque Les données sont perdues pendant la conversion en type de données Data Services, car int ne fait que quatre octets.
blob	blob i Remarque Ce type de données n'est pas pris en charge lors de l'importation de données dans des modèles de table stockés par colonne.
caractère	varchar
clob	long i Remarque Ce type de données n'est pas pris en charge lors de l'importation de données dans des modèles de table stockés par colonne.
date	date
dbclob	long i Remarque Ce type de données n'est pas pris en charge lors de l'importation de données dans des modèles de table stockés par colonne.
decimal	decimal
double	double
float	double
graphic	varchar
integer	int
long varchar	long
long vargraphic	long
real	real
smallint	int
time	time
timestamp	datetime
varchar	varchar
vargraphic	varchar

Le tableau suivant illustre la conversion de types de données internes en types de données DB2 dans des modèles de tables ou des tables de transformation Data_Transfer.

Type de données Data Services	Type de données DB2 dans un modèle de table
blob	blob
date	date
datetime	date
decimal	decimal
double	double
int	int
interval	real
long	clob
numeric	caractère
real	real
time	time
varchar	varchar
timestamp	date

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)

[Copybook Cobol \[page 858\]](#)

[Data Federator \[page 860\]](#)

[Informix \[page 863\]](#)

[Microsoft Excel \[page 864\]](#)

[Microsoft SQL Server \[page 865\]](#)

[MySQL \[page 868\]](#)

[Netezza \[page 870\]](#)

[ODBC \[page 871\]](#)

[Oracle \[page 873\]](#)

[SQL Anywhere \[page 875\]](#)

[SAP ASE \[page 877\]](#)

[SAP Sybase IQ \[page 879\]](#)

[Teradata \[page 881\]](#)

36.3.5 Informix

Le tableau suivant illustre la conversion entre des types de données Informix et des types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible Informix dans le référentiel et lorsqu'il charge des données dans une table ou un fichier externe.

Type de données Informix	Convertit de/en type de données Data Services.
blob	blob
byte	blob
char	varchar
caractère	varchar
character varying	non pris en charge
clob	long
date	date
datetime	datetime
dec	decimal
decimal	decimal
double	double
float	double
int	integer
integer	integer
money	decimal
numeric	decimal
real	real
serial	non pris en charge
smallfloat	double
smallint	integer
texte	long
varchar	varchar

Le tableau suivante illustre la conversion de types de données internes en types de données Informix dans des modèles de tables ou des tables de transformation Data_Transfer.

Type de données Data Services	Type de données Informix dans un modèle de table
blob	blob
date	date
datetime	date
decimal	decimal

Type de données Data Services	Type de données Informix dans un modèle de table
double	float
int	int
interval	int
long	clob
numeric	decimal
real	real
time	date
varchar	varchar
timestamp	date

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)
[Copybook Cobol \[page 858\]](#)
[Data Federator \[page 860\]](#)
[IBM DB2 \[page 860\]](#)
[Microsoft Excel \[page 864\]](#)
[Microsoft SQL Server \[page 865\]](#)
[MySQL \[page 868\]](#)
[Netezza \[page 870\]](#)
[ODBC \[page 871\]](#)
[Oracle \[page 873\]](#)
[SQL Anywhere \[page 875\]](#)
[SAP ASE \[page 877\]](#)
[SAP Sybase IQ \[page 879\]](#)
[Teradata \[page 881\]](#)

36.3.6 Microsoft Excel

Microsoft ActiveX Data Objects (ADO) permet de mettre en forme et de convertir des sources de données Excel. Le tableau suivant illustre la conversion de types de données ADO en types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible Excel dans le référentiel.

Type de données ADO	Converti en type de données Data Services.
adDouble	double
adCurrency	double
adBoolean	varchar
adDate	timestamp
adDBTimestamp	timestamp
ad...Char	varchar

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)

[Copybook Cobol \[page 858\]](#)

[Data Federator \[page 860\]](#)

[IBM DB2 \[page 860\]](#)

[Informix \[page 863\]](#)

[Microsoft SQL Server \[page 865\]](#)

[MySQL \[page 868\]](#)

[Netezza \[page 870\]](#)

[ODBC \[page 871\]](#)

[Oracle \[page 873\]](#)

[SQL Anywhere \[page 875\]](#)

[SAP ASE \[page 877\]](#)

[SAP Sybase IQ \[page 879\]](#)

[Teradata \[page 881\]](#)

36.3.7 Microsoft SQL Server

Le tableau suivant illustre la conversion entre des types de données Microsoft SQL Server et des types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible Microsoft SQL Server dans le référentiel et lorsqu'il charge des données dans une table ou un fichier externe.

Type de données Microsoft SQL Server	Convertit de/en type de données Data Services.
binary	non pris en charge
bigint	decimal

Type de données Microsoft SQL Server

Convertit de/en type de données Data Services.

bit	int
char	varchar
date (SQL Server 2008 et version supérieure uniquement)	date
datetime	datetime
datetime2 (SQL Server 2008 et version supérieure uniquement)	datetime
decimal	decimal
float	double
image	blob
int	int
money/smallmoney	decimal
nchar	varchar
ntext	long
numeric	decimal
nvarchar	varchar
nvarchar(max)	long
real	real
smalldatetime	datetime
smallint	int
text	long
time (SQL Server 2008 et version supérieure uniquement)	time
i Remarque Sur UNIX, les sous-secondes (HH:MI:SS.<ff> ou ff correspond aux sous-secondes) pour le type de données Time sont toujours définies sur zéro (0) Pour obtenir une heure plus précise, utilisez le type de données datetime ou datetime2.	
timestamp	non pris en charge
tinyint	int
uniqueidentifier	varchar
varbinary	non pris en charge
varbinary(max)	blob
varchar	varchar
varchar(max)	long
xml	long

Le tableau suivant illustre la conversion de types de données internes en types de données Microsoft SQL Server dans des modèles de tables ou des tables de transformation Data_Transfer.

Type de données Data Services	Type de données MS SQL Server dans un modèle de table
blob	image
date	datetime
	date (SQL Server 2008 et version supérieure uniquement)
datetime	datetime
	datetime2 (SQL Server 2008 et version supérieure uniquement)
decimal	decimal
double	float
int	int
interval	real
long	texte
numeric	decimal
real	real
time	datetime
	time (SQL Server 2008 et version supérieure uniquement)
varchar	varchar
timestamp	datetime
	datetime2 (SQL Server 2008 et version supérieure uniquement)

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)
[Copybook Cobol \[page 858\]](#)
[Data Federator \[page 860\]](#)
[IBM DB2 \[page 860\]](#)
[Informix \[page 863\]](#)
[Microsoft Excel \[page 864\]](#)
[MySQL \[page 868\]](#)
[Netezza \[page 870\]](#)
[ODBC \[page 871\]](#)
[Oracle \[page 873\]](#)
[SQL Anywhere \[page 875\]](#)
[SAP ASE \[page 877\]](#)

36.3.8 MySQL

Le tableau suivant illustre la conversion entre des types de données MySQL et des types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible MySQL dans le référentiel et lorsqu'il charge des données dans une table ou un fichier externe.

Type de données MySQL	Convertit de/en type de données Data Services.
bigint	decimal
decimal	decimal
dec	decimal
bit	int
tinyint	int
bool	int
smallint	int
mediumint	int
int	int
integer	int
year	int
float	real
double	double
datetime	datetime
timestamp	datetime
date	date
time	time
varchar	varchar
nvarchar	varchar
nchar	varchar
char	varchar
enum	varchar
set	varchar
tinytext	long
texte	long
mediumtext	long

Type de données MySQL	Convertit de/en type de données Data Services.
longtext	long
tinyblob	blob
blob	blob
mediumblob	blob
longblob	blob

Le tableau suivant illustre la conversion de types de données internes en types de données MySQL dans des modèles de tables ou des tables de transformation Data_Transfer.

Type de données Data Services	Type de données MySQL dans un modèle de table
blob	blob
date	date
datetime	timestamp
decimal	decimal
double	double
int	int
interval	float
long	texte
numeric	decimal
real	float
time	time
varchar	varchar
timestamp	timestamp

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)

[Copybook Cobol \[page 858\]](#)

[Data Federator \[page 860\]](#)

[IBM DB2 \[page 860\]](#)

[Informix \[page 863\]](#)

[Microsoft Excel \[page 864\]](#)

[Microsoft SQL Server \[page 865\]](#)

[Netezza \[page 870\]](#)

[ODBC \[page 871\]](#)

[Oracle \[page 873\]](#)

[SQL Anywhere \[page 875\]](#)

[SAP ASE \[page 877\]](#)

[SAP Sybase IQ \[page 879\]](#)

[Teradata \[page 881\]](#)

36.3.9 Netezza

Le tableau suivant illustre la conversion entre des types de données Netezza et des types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible Netezza dans le référentiel et lorsqu'il charge des données dans une table ou un fichier externe.

Type de données Netezza	Convertit de/en type de données Data Services.
bigint	decimal
boolean	int
byteint	int
char	varchar
date	date
double precision	double
float	double
integer	int
interval	varchar
nchar	varchar
nvarchar	varchar
numeric	decimal
real	real
smallint	int
time	time
time with time zone	varchar
timestamp	datetime
varchar	varchar

Le tableau suivant illustre la conversion de types de données internes en types de données Netezza dans des modèles de tables ou des tables de transformation Data_Transfer.

Type de données Data Services	Type de données Netezza dans un modèle de table
date	date
datetime	timestamp

Type de données Data Services	Type de données Netezza dans un modèle de table
decimal	numeric
double	double precision
int	integer
interval	real
numeric	numeric
real	real
time	time
timestamp	timestamp
varchar	varchar

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)
[Copybook Cobol \[page 858\]](#)
[Data Federator \[page 860\]](#)
[IBM DB2 \[page 860\]](#)
[Informix \[page 863\]](#)
[Microsoft Excel \[page 864\]](#)
[Microsoft SQL Server \[page 865\]](#)
[MySQL \[page 868\]](#)
[ODBC \[page 871\]](#)
[Oracle \[page 873\]](#)
[SQL Anywhere \[page 875\]](#)
[SAP ASE \[page 877\]](#)
[SAP Sybase IQ \[page 879\]](#)
[Teradata \[page 881\]](#)

36.3.10 ODBC

Le tableau suivant illustre la conversion entre des types de données ODBC et des types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible ODBC dans le référentiel et lorsqu'il charge des données dans une table ou un fichier externe.

Type de données ODBC	Convertit de/en type de données Data Services.
bigint	decimal
char	varchar
datalink	non pris en charge
date	date
decimal	decimal
double	double
float	double
graphic	non pris en charge
int	int
nclob	non pris en charge
numeric	decimal
real	real
sql_longvarchar	long
sql_longvarbinary	blob
sql_wlongvarchar	long
time	time
timestamp	datetime
tinyint	int
user-defined	non pris en charge
varchar	varchar

Le tableau suivant illustre la conversion de types de données internes en types de données ODBC dans des modèles de tables ou des tables de transformation Data_Transfer.

Type de données Data Services	Type de données ODBC dans un modèle de table
blob	sql_long varbinary
date	date
datetime	timestamp
decimal	decimal
double	double
int	int
interval	real
long	sql_long varchar
numeric	decimal
real	real
time	time

Type de données Data Services	Type de données ODBC dans un modèle de table
varchar	varchar
timestamp	timestamp

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)
[Copybook Cobol \[page 858\]](#)
[Data Federator \[page 860\]](#)
[IBM DB2 \[page 860\]](#)
[Informix \[page 863\]](#)
[Microsoft Excel \[page 864\]](#)
[Microsoft SQL Server \[page 865\]](#)
[MySQL \[page 868\]](#)
[Netezza \[page 870\]](#)
[Oracle \[page 873\]](#)
[SQL Anywhere \[page 875\]](#)
[SAP ASE \[page 877\]](#)
[SAP Sybase IQ \[page 879\]](#)
[Teradata \[page 881\]](#)

36.3.11 Oracle

La table suivante illustre la conversion entre des types de données Oracle et des types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible Oracle dans le référentiel et lorsqu'il charge des données dans une table ou un fichier externe.

Type de données Oracle	Convertit de/en type de données SAP Data Service
char	varchar
blob	blob
clob	long
date	datetime
decimal	decimal
doubleprecision	double
float	double

Type de données Oracle	Convertit de/en type de données SAP Data Service
label	non pris en charge
long	long
long raw	blob
nchar	varchar
nclob	long
Numéro	int : si échelle = 0 et précision < 9 decimal : Tous les autres
nvarchar2	varchar
real	double
row	non pris en charge
rowid	non pris en charge
timestamp	timestamp
varchar	varchar
varchar2	varchar

Le tableau suivant illustre la conversion de types de données internes en types de données Oracle dans des modèles de tables ou des tables de transformation Data_Transfer.

Type de données Data Services	Type de données Oracle dans un modèle de table
blob	blob
date	date
datetime	date
decimal	decimal
double	double
int	int
interval	real
long	clob
numeric	Numéro
real	double
time	date
varchar	varchar2
timestamp	timestamp

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)
[Copybook Cobol \[page 858\]](#)
[Data Federator \[page 860\]](#)
[IBM DB2 \[page 860\]](#)
[Informix \[page 863\]](#)
[Microsoft Excel \[page 864\]](#)
[Microsoft SQL Server \[page 865\]](#)
[MySQL \[page 868\]](#)
[Netezza \[page 870\]](#)
[ODBC \[page 871\]](#)
[SQL Anywhere \[page 875\]](#)
[SAP ASE \[page 877\]](#)
[SAP Sybase IQ \[page 879\]](#)
[Teradata \[page 881\]](#)

36.3.12 SQL Anywhere

Le tableau suivant illustre la conversion entre des types de données SAP Sybase SQL Anywhere et des types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible SQL Anywhere dans le référentiel et lorsqu'il charge des données dans une table ou un fichier externe.

Type de données SQL Anywhere	Convertit de/en type de données Data Services
BIT	int
[UNSIGNED] TINYINT	int
[UNSIGNED] SMALLINT	int
INT ou INTEGER	int
UNSIGNED INT ou INTEGER	decimal(10,0)
BIGINT	decimal(19,0)
UNSIGNED BIGINT	decimal(20,0)
DEC ou DECIMAL	decimal
NUMERIC	decimal
FLOAT(1-24) ou FLOAT	real
FLOAT(25-53)	double
REAL	real
DOUBLE [PRECISION]	double
SMALLMONEY	decimal(10,4)

Type de données SQL Anywhere	Convertit de/en type de données Data Services
MONEY	decimal(19,4)
CHAR	varchar
NCHAR	varchar
VARCHAR	varchar
NVARCHAR	varchar
LONG VARCHAR	long
LONG NVARCHAR	long
TEXT (LONG VARCHAR)	long
NTEXT (LONG NVARCHAR)	long
XML	long
UNIQUEIDENTIFIERSTR (CHAR(26))	varchar(36)
DATE	date
TIME	time
SMALLDATETIME	datetime
DATETIME	datetime
TIMESTAMP	datetime
BINARY	blob
IMAGE	blob
LONG BINARY	blob
VARBINARY	blob
UNIQUEIDENTIFIER	blob

Le tableau suivant illustre la conversion de types de données internes en types de données SQL Anywhere dans des modèles de tables ou des tables de transformation Data_Transfer.

Type de données Data Services	Type de données SQL Anywhere du modèle de table
blob	LONG BINARY
date	DATE
datetime	DATETIME
decimal	DECIMAL
double	DOUBLE
int	INTEGER
interval	REAL
long	LONG VARCHAR/LONG NVARCHAR
real	REAL
time	TIME

Type de données Data Services	Type de données SQL Anywhere du modèle de table
timestamp	DATETIME
varchar	VARCHAR/NVARCHAR

i Remarque

Lors de l'utilisation de SQL Anywhere :

- TIMESTAMP prend en charge des sous-secondes à six chiffres (c'est-à-dire des millisecondes)
- DECIMAL prend en charge une précision jusqu'à 96 grâce à ACTADECIMAL_MAXPRECISION (SQL Anywhere prend en charge une précision jusqu'à 127)
- Les colonnes VARCHAR prennent en charge une longueur maximale de 32 767 caractères
- Les colonnes LONG prennent en charge une longueur maximale de 2 Go de caractères

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)

[Copybook Cobol \[page 858\]](#)

[Data Federator \[page 860\]](#)

[IBM DB2 \[page 860\]](#)

[Informix \[page 863\]](#)

[Microsoft Excel \[page 864\]](#)

[Microsoft SQL Server \[page 865\]](#)

[MySQL \[page 868\]](#)

[Netezza \[page 870\]](#)

[ODBC \[page 871\]](#)

[Oracle \[page 873\]](#)

[SAP ASE \[page 877\]](#)

[SAP Sybase IQ \[page 879\]](#)

[Teradata \[page 881\]](#)

36.3.13 SAP ASE

Le tableau suivant illustre la conversion entre des types de données SAP ASE et des types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible SAP ASE dans le référentiel et lorsqu'il charge des données dans une table ou un fichier externe.

Type de données SAP ASE	Convertit de/en type de données Data Services.
binary	non pris en charge
bit	int
char	varchar
datetime	datetime
decimal	decimal
double	double
float	double
image	blob
int	int
money	decimal(20,4)
numeric	decimal
real	real
smalldatetime	datetime
smallint	int
smallmoney	decimal(12,4)
texte	long
timestamp	non pris en charge
tinyint	int
varbinary	non pris en charge
varchar	varchar

Le tableau suivant illustre la conversion de types de données internes en types de données SAP ASE dans des modèles de tables ou des tables de transformation Data_Transfer.

Type de données Data Services	Type de données SAP ASE dans un modèle de table
blob	image
date	datetime
datetime	datetime
decimal	decimal
double	float
int	int
interval	real
long	texte
numeric	decimal
real	real
time	datetime

Type de données Data Services	Type de données SAP ASE dans un modèle de table
varchar	varchar
timestamp	datetime

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)
[Copybook Cobol \[page 858\]](#)
[Data Federator \[page 860\]](#)
[IBM DB2 \[page 860\]](#)
[Informix \[page 863\]](#)
[Microsoft Excel \[page 864\]](#)
[Microsoft SQL Server \[page 865\]](#)
[MySQL \[page 868\]](#)
[Netezza \[page 870\]](#)
[ODBC \[page 871\]](#)
[Oracle \[page 873\]](#)
[SQL Anywhere \[page 875\]](#)
[SAP Sybase IQ \[page 879\]](#)
[Teradata \[page 881\]](#)

36.3.14 SAP Sybase IQ

Le tableau suivant illustre la conversion entre des types de données SAP Sybase IQ et des types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible Sybase IQ dans le référentiel et lorsqu'il charge des données dans une table ou un fichier externe.

Type de données SAP Sybase IQ	Convertit de/en type de données Data Services.
bigint	decimal
binary	non pris en charge
bit	int
blob	blob
char	varchar
clob	long
date	date

Type de données SAP Sybase IQ	Convertit de/en type de données Data Services.
datetime	datetime
decimal	decimal
double	double
float	real
int	int
long binary	blob
long varchar	long
money	decimal(19,4)
numeric	decimal
real	real
rowid	decimal
smalldatetime	datetime
smallint	int
smallmoney	decimal(10,4)
time	time
timestamp	datetime
tinyint	int
unsigned bigint	double
unsigned int	int
varbinary	non pris en charge
varchar	varchar

Le tableau suivant illustre la conversion de types de données internes en types de données SAP Sybase IQ dans des modèles de tables ou des tables de transformation Data_Transfer.

Type de données Data Services	Type de données SAP Sybase IQ dans un modèle de table
blob	longbinary
date	date
datetime	timestamp
decimal	decimal
double	double
int	int
interval	float
long	N/A
numeric	N/A
real	real

Type de données Data Services	Type de données SAP Sybase IQ dans un modèle de table
time	time
varchar	varchar
timestamp	timestamp

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)
[Copybook Cobol \[page 858\]](#)
[Data Federator \[page 860\]](#)
[IBM DB2 \[page 860\]](#)
[Informix \[page 863\]](#)
[Microsoft Excel \[page 864\]](#)
[Microsoft SQL Server \[page 865\]](#)
[MySQL \[page 868\]](#)
[Netezza \[page 870\]](#)
[ODBC \[page 871\]](#)
[Oracle \[page 873\]](#)
[SQL Anywhere \[page 875\]](#)
[SAP ASE \[page 877\]](#)
[Teradata \[page 881\]](#)

36.3.15 Teradata

Le tableau suivant illustre la conversion entre des types de données Teradata et des types de données Data Services lorsque Data Services importe des métadonnées depuis une source ou une cible Teradata dans le référentiel et lorsqu'il charge des données dans une table ou un fichier externe.

Types de données Teradata	Convertit de/en type de données Data Services.
bigint	decimal
blob	blob
byteint	int
clob	long
char varying (n)	varchar(n)
char [(n)]	varchar(n)

Types de données Teradata	Convertit de/en type de données Data Services.
date	date
decimal	decimal
double precision	float
float	float
int	int
long varchar	long
long vargraphic	long
numeric	decimal
real	float
smallint	int
time	time
timestamp	datetime
varchar < 32000	varchar
varchar >= 32000	long
varbyte < 32000	non pris en charge
varbyte >= 32000	blob

Le tableau suivant illustre la conversion de types de données internes en types de données Teradata dans des modèles de tables ou des tables de transformation Data_Transfer.

Type de données Data Services	Type de données Teradata dans un modèle de table
blob	blob
date	date
datetime	timestamp
decimal	decimal
double	N/A
int	int
interval	N/A
long	long varchar
numeric	N/A
real	N/A
time	time
varchar	varchar
timestamp	N/A

Rubrique parent : [Conversion de/en types de données internes \[page 856\]](#)

Informations associées

[Flux Attunity \[page 857\]](#)
[Copybook Cobol \[page 858\]](#)
[Data Federator \[page 860\]](#)
[IBM DB2 \[page 860\]](#)
[Informix \[page 863\]](#)
[Microsoft Excel \[page 864\]](#)
[Microsoft SQL Server \[page 865\]](#)
[MySQL \[page 868\]](#)
[Netezza \[page 870\]](#)
[ODBC \[page 871\]](#)
[Oracle \[page 873\]](#)
[SQL Anywhere \[page 875\]](#)
[SAP ASE \[page 877\]](#)
[SAP Sybase IQ \[page 879\]](#)

36.4 Conversion de types de données dans des expressions

Dans la mesure du possible, SAP Data Services optimise les flux de données en poussant les expressions vers un gestionnaire de base de données sous-jacent. Dans une transaction unique, le logiciel peut pousser vers le bas les expressions afin qu'elles soient exécutées par le gestionnaire de base de données sous-jacent. Toutefois, lorsque le logiciel évalue une expression contenant des opérandes de plusieurs types de données, il tente d'abord de convertir les opérandes dans le même type de données. (À l'exception des types de données de jeu de caractères national, qui peuvent être poussés vers le bas contrairement à d'autres dans une expression. Pour en savoir plus sur les types de données de jeu de caractères national, voir).

Lorsqu'une conversion est requise, le logiciel génère un message au moment de la validation.

Si la conversion n'est pas autorisée, le logiciel renvoie une erreur et vous devez supprimer la non-concordance avant d'exécuter le job.

Si la conversion est autorisée, le logiciel renvoie un avertissement indiquant que l'exécution du job ne sera pas interrompue.

i Remarque

Lorsque le logiciel convertit un type de données pour évaluer une expression, les résultats ne sont pas toujours ceux attendus. Pour éviter les conversions autorisées mais incorrectes, validez toujours avant l'exécution et lisez attentivement les avertissements concernant les conversions de types de données.

36.5 Conversion parmi des types de données Number

SAP Data Services utilise un algorithme de promotion de type pour évaluer les expressions contenant plusieurs types de données Number. Les types de données Number sont classés dans l'ordre décroissant comme suit :

- decimal
- double
- real
- int

Si le logiciel détecte des expressions contenant plusieurs types de données Number parmi les opérandes, il convertit tous les opérandes dans le type de données de l'opérande ayant le type de classement le plus élevé.

Par exemple, si A est un type `int` et B un type `double`, l'expression `A+B` est évaluée en convertissant A en `double`, puis en ajoutant les deux valeurs `double`. Le résultat est un type `double`.

Si en plus de A et B, vous multipliez le résultat par un nombre `decimal` C, `(A+B) * C` est évaluée en convertissant `(A+B)` dans le type `decimal`, puis en effectuant les opérations indiquées sur les deux valeurs `decimal`. Le résultat est un type `decimal`.

Pour les additions, soustractions et multiplications, le résultat de l'opération sera égal au plus élevé des deux opérandes. Par exemple :

```
int + double = double
```

L'algorithme suivant est utilisé pour les divisions :

Type de données Numerator	Type de données Denominator			
	int	real	double	decimal(p,s1)
int	double	double	double	decimal(p,s1)
real	double	double	double	decimal(p,s1)
double	double	double	double	decimal(p,s1)
decimal(p,s2)	decimal(p,s2)	decimal(p,s2)	decimal(p,s2)	decimal(p,max(10,s1,s2))

Conversion parmi des décimales d'échelle ou de précision différentes

Si des décimales de deux échelles différentes figurent dans une seule expression, le logiciel utilise la plus élevée des deux échelles. Par exemple :

```
decimal(5,4) * decimal(7,2) = decimal(7,2)
```

Une perte de précision est à prévoir avec deux décimales de valeurs d'échelle différentes. Par exemple, si `decimal(28, 26)` est ajouté à `decimal(28, 1)`, la valeur décimale qui en résulte a la plus faible des deux valeurs d'échelle :

```
400000.5 + 40.00005 = 400040.5
```

L'échelle minimale pour les divisions impliquant une décimale est de 10.

Conversions entre chaînes et nombres

Lorsque le logiciel détecte une chaîne dans laquelle un nombre est généralement attendu (par exemple, dans des opérations ou des fonctions mathématiques attendant des arguments numériques), il tente de convertir la chaîne en nombre.

Pour les multiplications et les divisions, les opérandes sont convertis en nombres. D'autres algorithmes de promotion sont indiqués dans le tableau suivant.

Addition

Type de données fourni	Type de données requis pour évaluer l'expression		
	Number	Date/time	Chaîne
Number	OK (promu)	Number en interval	String en number
Date/time	Number en interval	Non autorisé	String en interval
Chaîne	String en number	String en interval	String en real

Soustraction

Type de données fourni	Type de données requis pour évaluer l'expression			
	Number	Date/time	Chaîne	Interval
Number	OK (promu)	Non autorisé	String en number	Interval en number
Date/time	Number en interval	OK	String en interval	OK
Chaîne	String en number	String en datetime	String en real	String en number
Interval	Number en interval	Non autorisé	String en interval	OK

Comparaison

Type de données fourni	Type de données requis pour évaluer l'expression			
	Number	Date/time	Chaîne	Interval
Number	OK (promu)	Non autorisé	String en number	OK
Date/time	Non autorisé	OK	Non autorisé	Non autorisé
Chaîne	Non autorisé	String en datetime	OK	String en interval
Interval	OK	Non autorisé	String en interval	OK

Conversions entre chaînes et dates

Pour Oracle, si vous chargez des données datetime de SAP Data Services dans un champ de caractères d'une table Oracle, Oracle convertit les données dans son format datetime par défaut, qui inclut uniquement des valeurs de date et perd l'heure de la valeur.

36.5.1 Division algorithm

SAP Data Services uses a special algorithm when it divides number data types.

Division algorithm

Numerator data type	Denominator data type			
	int	real	double	decimal(p,s1)
int	double	double	double	decimal(p,s1)
real	double	double	double	decimal(p,s1)
double	double	double	double	decimal(p,s1)
decimal(p,s2)	decimal(p,s2)	decimal(p,s2)	decimal(p,s2)	decimal(p,max(10,s1,s2))

36.5.2 Conversion of decimals with different scale and precision

Data Services follows a specific process when it converts decimal data types that have different scale and precision settings.

If decimals of two different scales are included in a single expression, the software uses the higher of the two scales.

❖ Exemple

For example:

```
decimal(5,4) * decimal(7,2) = decimal(7,2)
```

When Data Services operates on two `decimals` of different scale values, it loses the precision.

❖ Exemple

For example, when adding a `decimal(28, 26)` to a `decimal(28,1)`, the resulting decimal value has the lower of the two scale values:

```
400000.5 + 40.00005 = 400040.5
```

The lowest scale for division that involves a decimal is 10.

36.5.3 Conversion between string and number data types

Data Services attempts to convert a string to a number data type when it encounters a string where it expects a number.

For example, in mathematical operations or functions where Data Services expects numeric arguments, Data Services attempts to convert a string to a number.

For multiplication and division operations, Data Services converts operands to numbers.

The following tables show promotion algorithms for addition, subtraction, and comparison.

Addition promotion algorithm

Provided data type	Data type required to evaluate the expression		
	Number	Date/time	String
Number	OK (promoted)	Number to interval	String to number
Date/time	Number to interval	Illegal	String to interval
String	String to number	String to interval	String to real

Subtraction promotion algorithm

Provided data type	Data type required to evaluate the expression			
	Number	Date/time	String	Interval
Number	OK (promoted)	Illegal	String to number	Interval to number
Date/time	Number to interval	OK	String to interval	OK
String	String to number	String to datetime	String to real	String to number
Interval	Number to interval	Illegal	String to interval	OK

Comparison promotion algorithm

Provided data type	Data type required to evaluate the expression			
	Number	Date/time	String	Interval
Number	OK (promoted)	Illegal	String to number	OK
Date/time	Illegal	OK	Illegal	Illegal
String	Illegal	String to datetime	OK	String to interval
Interval	OK	Illegal	String to interval	OK

36.5.4 Conversion between Oracle string and date data types

Oracle changes data to a default datetime format in certain circumstances.

If you load `datetime` data from SAP Data Services into a `char` field in an Oracle table, Oracle puts the data in its default datetime format. The default datetime format includes only date values so the conversion causes the data to lose the time value.

36.6 Conversions entre types de données explicites

Vous pouvez utiliser des fonctions pour convertir les données d'un type dans un autre type :

-
-
-
-
-
-

Vous pouvez aussi importer des fonctions spécifiques à la base de données pour effectuer des conversions de types de données.

36.7 Conversion entre des types de données natifs

Une transformation Data Quality peut obtenir et définir des données de champ dans un format autre que le type de données déclaré. Par exemple, si le champ d'entrée est varchar, il peut être mappé vers un champ dont le type de données est un nombre entier à condition que le champ varchar contienne tous les chiffres. Cependant, certaines conversions ne sont pas prises en charge, selon le type de données et le contenu du champ.

i Remarque

Lorsqu'un type de données est mappé à un champ d'entrée ou de sortie dont le type de données n'est pas valide, la transformation affiche une erreur de vérification.

❖ Exemple

La transformation USA Regulatory Address Cleanse inclut un champ d'entrée de type, nommé Code_postal_complet. Les types de champ Varchar peuvent écrire toute sorte de types de données à condition que celles-ci soient mises en forme correctement et contiennent tous les chiffres. Le champ Code_postal_complet peut être un nombre entier car le champ contient des nombres. Cependant, le champ Code_postal_complet ne peut être de type date car il n'est pas conforme au format date.

La partie restante de cette section répertorie tous les types de champ et pour chacun, les types de données applicables ou non valides.

[date \[page 889\]](#)

[datetime \[page 889\]](#)

[decimal \[page 890\]](#)

[double \[page 891\]](#)

[int \(entier\) \[page 892\]](#)

[caractères \(varchar\) \[page 892\]](#)

36.7.1 date

Entrée : Une transformation peut lire dans un champ d'entrée de date et écrire dans les types de données suivants :

- date
- caractère
- datetime

Sortie : Une transformation peut écrire dans les types de données suivants à partir d'un type de données date :

- date
- caractère
- datetime

Types de données d'entrée et de sortie non valides pour la date :

- nombre entier
- double
- decimal

Rubrique parent : [Conversion entre des types de données natifs \[page 888\]](#)

Informations associées

[datetime \[page 889\]](#)

[decimal \[page 890\]](#)

[double \[page 891\]](#)

[int \(entier\) \[page 892\]](#)

[caractères \(varchar\) \[page 892\]](#)

36.7.2 datetime

Entrée : Une transformation peut lire dans un champ d'entrée de date-heure et écrire dans les types de données suivants :

- datetime
- character
- date (avec troncature possible)
- time (avec troncature possible)

Sortie : Une transformation peut écrire dans les types de données suivants à partir d'un type de données datetime :

- datetime
- character

- date (avec troncature possible)
- time (avec troncature possible)

Types de données d'entrée et de sortie non valides pour la date/heure :

- nombre entier
- double
- decimal

Rubrique parent : [Conversion entre des types de données natifs \[page 888\]](#)

Informations associées

[date \[page 889\]](#)

[decimal \[page 890\]](#)

[double \[page 891\]](#)

[int \(entier\) \[page 892\]](#)

[caractères \(varchar\) \[page 892\]](#)

36.7.3 decimal

Entrée : Une transformation peut lire dans un champ d'entrée de nombre décimal et écrire dans les types de données suivants :

- decimal
- caractère
- nombre entier (les données peuvent être tronquées)
- double (les données peuvent être tronquées)

Sortie : Une transformation peut écrire dans les types de données suivants à partir d'un type de données decimal :

- decimal
- caractère
- nombre entier (avec troncature possible)
- double (avec troncature possible)

Types de données d'entrée et de sortie non valides pour le nombre décimal :

- date
- time
- datetime

Rubrique parent : [Conversion entre des types de données natifs \[page 888\]](#)

Informations associées

[date \[page 889\]](#)

[datetime \[page 889\]](#)

[double \[page 891\]](#)

[int \(entier\) \[page 892\]](#)

[caractères \(varchar\) \[page 892\]](#)

36.7.4 double

Entrée : Une transformation peut lire dans un champ d'entrée double et écrire dans les types de données suivants :

- double
- character
- decimal (avec troncature possible)
- nombre entier (avec troncature possible)

Sortie : Une transformation peut écrire dans les types de données suivants à partir d'un type de données double :

- double
- character
- decimal (avec troncature possible)
- nombre entier (avec troncature possible)

Types de données d'entrée et de sortie non valides pour le type de données double :

- date
- time
- datetime

Rubrique parent : [Conversion entre des types de données natifs \[page 888\]](#)

Informations associées

[date \[page 889\]](#)

[datetime \[page 889\]](#)

[decimal \[page 890\]](#)

[int \(entier\) \[page 892\]](#)

[caractères \(varchar\) \[page 892\]](#)

36.7.5 int (entier)

Entrée : Une transformation peut lire dans un champ d'entrée de nombre entier et écrire dans les types de données suivants :

- nombre entier
- caractère
- decimal
- double

Sortie : Une transformation peut écrire dans les types de données suivants à partir d'un type de données nombre entier :

- nombre entier
- caractère
- double
- decimal

Types de données d'entrée et de sortie non valides pour le nombre entier :

- date
- time
- datetime

Rubrique parent : [Conversion entre des types de données natifs \[page 888\]](#)

Informations associées

[date \[page 889\]](#)

[datetime \[page 889\]](#)

[decimal \[page 890\]](#)

[double \[page 891\]](#)

[caractères \(varchar\) \[page 892\]](#)

36.7.6 caractères (varchar)

Entrée : Une transformation peut lire des données d'entrée varchar et les écrire dans n'importe quel type de données pris en charge. Cependant, si les données varchar ne sont pas mises en forme correctement pour le type de données, les résultats ne sont pas définis. Par exemple, si un type de données varchar est converti en nombre entier, il doit contenir tous les chiffres pour que la conversion soit correcte.

Sortie : Une transformation peut écrire tout type de données pris en charge et les écrire dans un type de données varchar. La transformation convertit automatiquement le contenu du champ en données varchar.

i Remarque

Les données peuvent être tronquées si le champ de sortie n'est pas suffisamment long.

Rubrique parent : [Conversion entre des types de données natifs \[page 888\]](#)

Informations associées

[date \[page 889\]](#)

[datetime \[page 889\]](#)

[decimal \[page 890\]](#)

[double \[page 891\]](#)

[int \(entier\) \[page 892\]](#)

37 Surveillance des Jobs

37.1 Administrateur

L'application Administrateur dans la Console de gestion constitue la principale ressource de surveillance pour tous les jobs conçus dans Designer. Pour en savoir plus, voir la section « Administrateur » du guide *Management Console Guide*.

38 Développement multi-utilisateurs

Travailler dans un environnement de développement multi-utilisateurs dans SAP Data Services à l'aide d'un référentiel central.

A propos des utilisateurs multiples

SAP Data Services prend en charge un environnement de développement multi-utilisateurs. Une équipe peut travailler en commun sur une application durant la phase de développement, de test ou de production. Différentes équipes peuvent également travailler sur différentes phases au même moment.

Chaque développeur travaille sur une application dans son référentiel local unique. Chaque équipe utilise un référentiel central pour stocker la copie principale de son application. Le référentiel central conserve toutes les versions de tous les objets de l'application, les utilisateurs peuvent donc revenir à une version précédente si nécessaire.

Informations associées

[Bibliothèque d'objets centrale \[page 25\]](#)

[Utilisateurs multiples \[page 897\]](#)

38.1 Référentiel central et référentiel local

Un référentiel central archive la copie d'équipe d'un objet SAP Data Services et constitue un contrôle de source pour environnements multi-utilisateurs.

Un référentiel central contient toutes les informations dans un référentiel local, tels que des définitions d'objets, des banques de données, etc. Le référentiel central est un emplacement de stockage de ces informations. Pour modifier l'objet, travaillez dans un référentiel local.

Un référentiel local contient des copies d'objets dans le référentiel central. Il existe deux manières de mettre à jour votre référentiel local avec le contenu du référentiel central :

- Vous pouvez copier dans votre référentiel local les objets du référentiel central.
- Vous pouvez extraire des objets du référentiel central vers votre référentiel local.

Copie d'objets

Lorsque vous copiez un objet du référentiel central, le logiciel place la copie dans votre référentiel local si vous n'avez pas déjà l'objet. Si vous disposez déjà de cet objet dans le référentiel local, le logiciel remplace votre copie locale avec la version du référentiel central.

Lorsque vous modifiez votre copie locale :

- L'objet dans le référentiel central n'est pas modifié.
- Vous ne pouvez pas télécharger les objets modifiés localement dans le référentiel central pour remplacer la version actuelle étant donné que vous avez uniquement copié et pas extrait l'objet.

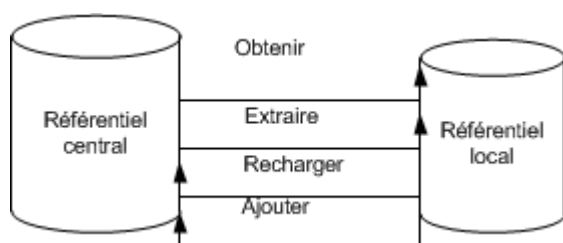
Extraction d'objets

Lorsque vous extrayez un objet du référentiel central, il se produit la même chose dans votre référentiel local que lorsque vous obtenez une copie : la copie est placée dans votre référentiel local si vous ne l'avez pas déjà, ou la copie que vous avez déjà est écrasée par la version du référentiel central. En outre, l'objet dans le référentiel central est bloqué pour vous permettre d'apporter des modifications à cet objet que vous téléchargez plus tard dans le référentiel central. Lorsque vous avez un objet extrait, aucun autre utilisateur ne peut extraire l'objet. Les utilisateurs peuvent copier l'objet, mais les modifications apportées à leur copie locale restent locales.

Lorsque vous modifiez votre copie locale :

- L'objet dans le référentiel central n'est pas modifié tant que vous contrôlez votre copie locale dans le référentiel central.
- Après avoir contrôlé vos modifications dans le référentiel central, un utilisateur autorisé peut extraire l'objet pour le modifier ultérieurement.
- Les modifications que vous avez effectuées localement sur un objet et contrôlées dans le référentiel central ne sont pas dispersées vers une copie locale de cet objet, sauf si celui-ci est copié à partir du référentiel central ou extrait.

L'image suivante montre toutes les façons d'accéder à des objets du référentiel central et de placer des objets dans le référentiel central.

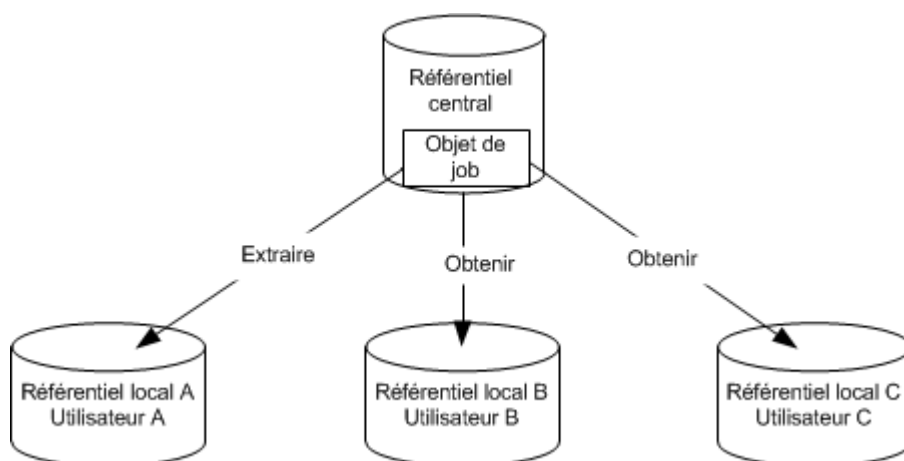


Historique des objets

Le référentiel central effectue le suivi de toutes les versions d'un objet ainsi que des dates et heures auxquelles le logiciel a enregistré cette version de l'objet. La bibliothèque d'objets centrale dans Designer affiche une brève histoire d'un élément dans la colonne [Version la plus récente](#). Pour plus d'informations, sélectionnez l'icône du menu [Afficher l'historique](#) dans la bibliothèque d'objets centrale.

Fonctionnement

Plusieurs utilisateurs travaillant à partir de référentiels locaux uniques peuvent se connecter au même référentiel central. Ces utilisateurs peuvent travailler sur la même application et partager leur travail. Cependant, à un moment donné, un seul utilisateur peut extraire et modifier un objet particulier. Lorsqu'un objet est extrait pour un utilisateur, les autres utilisateurs peuvent obtenir la dernière version de l'objet ou la copier. Cependant, les autres utilisateurs ne peuvent pas modifier l'objet dans le référentiel central.



Le référentiel central conserve chaque version d'un objet. La rétention de version vous permet de rétablir un objet à sa version précédente lorsque les modifications que vous avez apportées ne fonctionnent pas ou sont erronées.

Le référentiel central et le référentiel local doivent utiliser la même version logicielle. Vous pouvez, par exemple, exécuter la version X.2 de Designer avec une version X.1 de référentiel central et local. Toutefois, vous ne pouvez pas exécuter Designer version X.2 avec un référentiel central version X.1 et un référentiel local version X.2.

Informations associées

[Bibliothèque d'objets centrale \[page 25\]](#)

[Affichage et modification des autorisations \[page 908\]](#)

38.2 Utilisateurs multiples

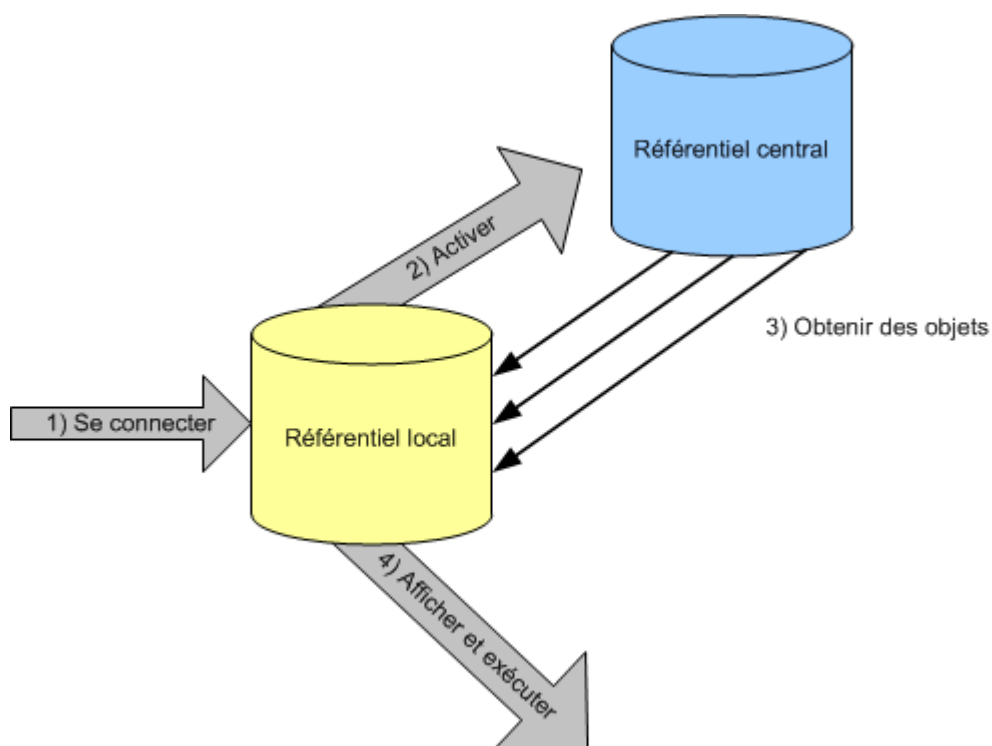
Un environnement multi-utilisateurs affecte la manière dont vous utilisez SAP Data Services et dont vous gérez les différentes phases d'une application. Pour une utilisation adéquate d'un environnement multi-utilisateurs, vous devez maintenir la cohérence entre votre référentiel local et le référentiel central.

Les termes suivants s'appliquent aux environnements logiciel et multi-utilisateurs :

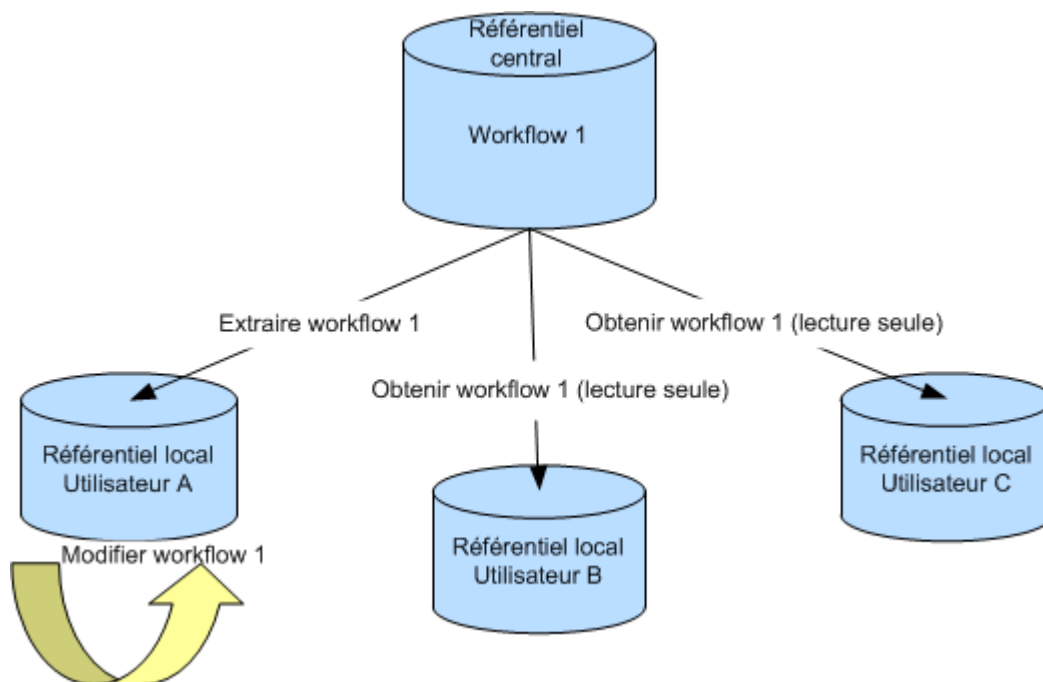
Terme	Définition
Objet de niveau supérieur	L'objet de niveau supérieur ne dépend d'aucun autre dans la hiérarchie des objets. Par exemple, si Job 1 comprend Workflow 1 et Flux de données 1, Job 1 est l'objet de niveau supérieur.

Terme	Définition
Objets dépendants	Les objets dépendants sont des objets associés en dessous de l'objet de niveau supérieur dans la hiérarchie. Par exemple, si Job 1 comprend Workflow 1, qui contient Flux de données 1, alors Workflow 1 et Flux de données 1 sont dépendants de Job 1. En outre, Flux de données 1 est dépendant de Workflow 1.
Version d'objet	Une version d'objet est une instance d'un objet. Chaque fois que vous ajoutez ou rechargez un objet dans le référentiel central, le logiciel crée une nouvelle version de l'objet. La dernière version d'un objet est la version la plus récemment créée.

En travaillant dans un environnement multi-utilisateurs, vous activez le lien entre votre référentiel local et le référentiel central chaque fois que vous vous connectez. Pour garantir l'actualisation de votre référentiel, vous pouvez obtenir (copier) la dernière version de chaque objet du référentiel central. Après avoir obtenu une application dans votre référentiel local, vous pouvez l'afficher et l'exécuter depuis Designer.



Toutefois, si vous prévoyez de modifier des objets dans l'application, vous devez les extraire. Après extraction d'un objet, aucun autre utilisateur ne peut le modifier. Principalement, vous verrouillez la version du référentiel central et vous seul pouvez la modifier. Les autres utilisateurs peuvent uniquement obtenir et afficher l'objet.



Lorsque vous apportez des modifications à un objet, enregistrez-les dans le référentiel local et rechargez l'objet dans le référentiel central. Le progiciel enregistre les objets modifiés dans le référentiel central et les rend disponibles pour une extraction par les autres utilisateurs. Le progiciel gère toutes les versions des objets enregistrés dans le référentiel central. Ultérieurement, vous pouvez donc copier une ancienne version d'un objet enregistré, même après l'avoir remplacé dans votre référentiel local par une nouvelle version.

A tout moment, vous pouvez étiqueter un objet ou un groupe d'objets. Une étiquette d'objet représente un mécanisme pratique pour identifier les objets par la suite. Par exemple, vous pouvez trouver utile d'étiqueter des objets par fonctionnalité. Ensuite, si vous décidez d'éliminer une fonctionnalité récemment ajoutée, vous pouvez obtenir tous les objets ayant une étiquette sans cette fonctionnalité.

Vous pouvez également comparer deux objets, par exemple deux versions d'objet différentes dans le référentiel central ou un objet de votre référentiel local avec un objet du référentiel central. En comparant deux objets, vous pouvez déterminer quelles parties d'un objet ont été modifiées et décider de revenir à une ancienne version d'un objet.

Informations associées

[Guide de Designer : Conception et débogage, Comparaison d'objets \[page 755\]](#)

38.3 Sécurité et référentiel central

Il existe plusieurs options pour sécuriser le référentiel central. Utilisez ces options pour contrôler l'accès et fournir un suivi des objets dans le référentiel central. Ces options de sécurité s'appliquent uniquement aux référentiels centraux et comprennent :

- Authentification : autorise uniquement les utilisateurs valides à se connecter à un référentiel central.
- Autorisation : accorde différents niveaux d'autorisations sur les objets.
- Audit : gère un historique des modifications apportées à un objet, y compris les noms d'utilisateur.

Mettez en place la sécurité d'un référentiel central en établissant une structure de groupes et utilisateurs associés à l'aide de l'Administrateur.

Informations associées

[Implémentation de la sécurité du référentiel central \[page 904\]](#)

38.4 Configuration d'environnement multi-utilisateurs

Présentation de la configuration multi-utilisateurs

Pour prendre en charge plusieurs développeurs, configurez un environnement multi-utilisateurs et plusieurs référentiels. Notamment, vous devez :

- Créer un référentiel local pour chaque développeur.
- Créer un référentiel central.
- Définir une connexion au référentiel central à partir de chaque référentiel local.
- Activer la connexion à un référentiel central.

Informations associées

[Création d'un référentiel central non sécurisé \[page 901\]](#)

[Définition d'une connexion à un référentiel central non sécurisé \[page 901\]](#)

[Activation d'un référentiel central \[page 902\]](#)

38.4.1 Création d'un référentiel central non sécurisé

Pour prendre en charge plusieurs utilisateurs dans un seul environnement de développement, nous vous recommandons d'utiliser un référentiel central. Le référentiel central archive des informations de base pour l'environnement de développement.

Cette procédure s'applique uniquement à des référentiels non sécurisés.

1. Créez une base de données à utiliser pour le référentiel central à l'aide de votre système de gestion de base de données.
2. Sélectionnez ► [Démarrer](#) ► [Programmes](#) ► [SAP Data Services 4.2](#) ► [Gestionnaire de référentiel Data Services](#) ►.
3. Dans la fenêtre Gestionnaire de référentiel, cliquez sur le bouton [Central](#) dans le champ [Type de référentiel](#) et saisissez les informations de connexion au référentiel central.
4. Cliquez sur [Créer](#).

Les tables de référentiel sont créées dans la base de données que vous avez identifiée.

Informations associées

[Implémentation de la sécurité du référentiel central \[page 904\]](#)

38.4.2 Définition d'une connexion à un référentiel central non sécurisé

Une équipe travaillant sur une application a besoin d'un seul référentiel central. Cependant chaque membre de l'équipe a besoin d'un référentiel local. De plus, chaque référentiel local a besoin d'informations de connexion à tout référentiel central auquel il doit accéder.

Cette procédure s'applique uniquement à des référentiels non sécurisés.

i Remarque

La version du référentiel central doit correspondre à la version du référentiel local.

1. Lancez Designer et connectez-vous au référentiel local.
2. Sélectionnez ► [Outils](#) ► [Référentiels centraux](#) ► pour ouvrir la fenêtre Options.
L'option Connexions au référentiel central est sélectionnée dans la liste des options de Designer
3. Cliquez avec le bouton droit de la souris dans la case [Connexions au référentiel central](#) et sélectionnez [Ajouter](#).
La fenêtre [Mot de passe du référentiel](#) s'ouvre.
4. Saisissez le mot de passe pour le référentiel central.
Le référentiel apparaît dans la zone [Connexions au référentiel central](#).

5. Cliquez sur [Activer](#).
6. Saisissez à nouveau le mot de passe pour le référentiel central.
7. Cliquez sur [OK](#).

Informations associées

[Implémentation de la sécurité du référentiel central \[page 904\]](#)

38.4.3 Activation d'un référentiel central

Pour connecter un référentiel central, vous devez activer le lien entre votre référentiel local et un référentiel central spécifique.

i Remarque

Lorsque vous démarrez Designer, connectez-vous toujours à un référentiel local. Ne vous connectez jamais à un référentiel central. Si vous le faites, le référentiel central agit comme un référentiel local. Vous augmentez alors le risque de corruption des informations de version. Si vous essayez de vous connecter au référentiel central, un message d'avertissement s'affichera. Vous devez vous déconnecter immédiatement et vous connecter à un référentiel local.

Votre référentiel local donne un aperçu des objets du référentiel central actif. Chaque fois que vous accédez à des objets ou les extrayez, vous les copiez du référentiel central actif. Chaque fois que vous rechargez des objets, vous sauvegardez la version de votre référentiel local dans votre référentiel central actif.

Vous devez activer le référentiel central correct chaque fois que vous vous connectez. Lorsque vous activez un référentiel central, la bibliothèque centrale d'objets s'ouvre et présente tous les objets du référentiel central et le statut d'extraction de chaque objet.

38.4.3.1 Activation d'un référentiel central

1. Sélectionnez **► Outils ► Référentiels centraux ►** pour ouvrir la fenêtre Options.
L'option Connexions au référentiel central est sélectionnée dans la liste des options de Designer
2. Dans la liste [Connexions au référentiel central](#), déterminez un référentiel central à activer.
3. Cochez [Réactiver automatiquement](#) pour réactiver le référentiel central actif lors de votre prochaine connexion à ce référentiel local.
4. Cliquez avec le bouton droit de la souris sur le référentiel central et sélectionnez [Activer](#).

La bibliothèque d'objets centrale s'ouvre. La fenêtre Options indique que le référentiel central sélectionné est actif et se ferme automatiquement.

38.4.3.2 Ouverture de la bibliothèque d'objets centrale

Ouvrez la bibliothèque d'objets centrale directement à partir de SAP Data Service Designer.

Designer dispose d'une icône qui facilite l'ouverture de la bibliothèque d'objets centrale.

1. Cliquez sur l'icône [Bibliothèque d'objets centrale](#)  dans la barre d'outils, ou bien sélectionnez ► [Outils](#) ► [Bibliothèque d'objets centrale](#) ►.

La bibliothèque d'objets centrale ressemble à la bibliothèque d'objets locale. Elle affiche tous les objets dans le référentiel, regroupés dans les onglets appropriés. Cependant, la bibliothèque centrale d'objets contient les colonnes supplémentaires [Référentiel d'extraction](#) et [Dernière version](#) pour le suivi des objets et l'affichage de l'historique. Les colonnes supplémentaires indiquent les autorisations de groupe et le nom de l'utilisateur qui a extrait l'objet.

i Remarque

Si la bibliothèque centrale est sécurisée, d'autres colonnes sont présentes : [Utilisateur à l'origine de l'extraction](#) et [Autorisation](#).

La boîte de dialogue [Bibliothèque d'objets centrale](#) s'ouvre en mode flottant.

2. Faites glisser la boîte de dialogue vers la zone de projet pour l'ancrer avec vos autres outils.
3. Cliquez avec le bouton droit sur la barre d'outils [Bibliothèque d'objets centrale](#), puis sélectionnez [Flottant](#) dans la liste déroulante pour détacher la boîte de dialogue.

Vous pouvez également modifier les informations de connexion dans la boîte de dialogue [Bibliothèque d'objets centrale](#).

Informations associées

[Définition d'une connexion à un référentiel central non sécurisé \[page 901\]](#)

[Bibliothèque d'objets centrale \[page 25\]](#)

[Implémentation de la sécurité du référentiel central \[page 904\]](#)

38.4.3.3 Modifier le référentiel central actif

Sélectionnez un référentiel central dans la liste en haut de la bibliothèque d'objets centrale.

SAP Data Services active la bibliothèque centrale sélectionnée, les objets de ce référentiel s'affichent dans la bibliothèque d'objets centrale. Les informations de connexion relatives à ce référentiel s'affichent en haut à droite de la bibliothèque d'objets centrale.

38.4.3.4 Modifier les connexions au référentiel central

1. Cliquez sur le bouton [Modifier la connexion au référentiel central](#) en haut de la bibliothèque d'objets centrale.

La fenêtre Options s'ouvre avec l'option Connexions au référentiel central sélectionnée dans la liste des options de Designer.

Vous pouvez également ouvrir la fenêtre Options en sélectionnant ► [Outils](#) ► [référentiels centraux](#) ►.

2. Sélectionnez un référentiel central dans le cadre [Connexions au référentiel central](#), cliquez avec le bouton droit de la souris et sélectionnez [Modifier](#).

Lorsque la fenêtre Administrateur de banque de données s'ouvre :

- Pour vous déconnecter du référentiel central actuellement actif, cliquez avec le bouton droit de la souris sur le référentiel central dans le cadre [Banques de données du référentiel central](#) et sélectionnez [Désactiver](#)
- Pour supprimer des informations de connexion pour un référentiel central, cliquez avec le bouton droit de la souris sur le cadre [Banques de données de référentiel central](#) et sélectionnez [Supprimer](#). Une fois que vous avez confirmé votre sélection, les informations de connexion de ce référentiel local sont supprimées. Vous ne pouvez plus vous connecter à ce référentiel central à partir de ce référentiel local.

i Remarque

Vous ne supprimez pas le référentiel central, vous supprimez uniquement les informations de connexion entre votre référentiel local et ce référentiel central.

- Pour activer un autre référentiel du référentiel central, cliquez avec le bouton droit de la souris sur le cadre [Banques de données de référentiel central](#) et sélectionnez [Activer](#).

38.5 Implémentation de la sécurité du référentiel central

A propos de cette section

Cette section décrit comment implémenter des fonctionnalités de sécurité facultatives pour les référentiels centraux.

38.5.1 Présentation

SAP Data Services propose des options pour la gestion de l'accès sécurisé et le suivi des objets dans les référentiels centraux. Les mécanismes de gestion de la sécurité du référentiel central comprennent :

- Authentification : autorise les utilisateurs valides à se connecter à un référentiel central.
- Autorisation : accorde différents niveaux d'autorisations sur les objets.

- **Audit** : gère un historique des modifications apportées à un objet, y compris les noms d'utilisateur.

Notez que ces mécanismes et procédures de sécurité s'appliquent uniquement aux référentiels centraux.

38.5.1.1 Autorisations basées sur le groupe

Vous implémentez la sécurité d'un référentiel central en établissant une structure de groupes et d'utilisateurs associés à l'aide de l'Administrateur et de la CMC (Central Management Console).

Les autorisations d'accès aux objets s'appliquent au niveau du groupe. Plusieurs groupes peuvent disposer de la même autorisation sur le même objet à un moment donné. Les groupes sont spécifiques à un référentiel et ne sont visibles dans aucun autre référentiel local ou central.

Par conséquent, les utilisateurs n'obtiennent pas d'autorisations individuelles. Dans Designer, les utilisateurs sélectionnent parmi les groupes auxquels ils appartiennent, et le groupe sélectionné (actuel) impose son accès à cet objet. Chaque utilisateur doit avoir un groupe par défaut, mais peut appartenir à plusieurs groupes. Lorsqu'un utilisateur ajoute un objet à un référentiel central sécurisé, le groupe actuel de l'utilisateur dispose d'autorisations complètes sur cet objet.

Une authentification par nom d'utilisateur et mot de passe est requise pour chaque connexion à un référentiel central sécurisé. Les utilisateurs peuvent modifier leur mot de passe à tout moment dans la CMC (Central Management Console).

38.5.1.2 Niveaux d'autorisation

Chaque objet d'un référentiel central sécurisé peut disposer d'un des niveaux d'autorisation suivants :

- **Complète** : c'est le niveau d'autorisation le plus élevé. Le groupe peut exécuter toutes les actions possibles, y compris rechargement, extraction et suppression de l'objet. Vous pouvez par exemple affecter ce type d'accès aux développeurs.
- **Lecture** : les utilisateurs peuvent uniquement obtenir une copie de l'objet depuis le référentiel central ou comparer les objets entre leurs bibliothèques d'objets locale et centrale. Vous pouvez par exemple affecter ce type d'accès à l'Assurance-Qualité.
- **Aucune** : les utilisateurs ne peuvent pas obtenir de copies de l'objet, mais peuvent l'afficher avec ses propriétés.

Lorsqu'un utilisateur authentifié ajoute un objet à un référentiel central sécurisé, le groupe actuel de l'utilisateur dispose d'autorisations complètes sur l'objet. Tous les autres groupes reçoivent l'autorisation Lecture. Les membres du groupe disposant des autorisations complètes peuvent modifier les autorisations des autres groupes pour cet objet.

38.5.1.3 Synthèse du processus

Vous implémentez la sécurité pour un référentiel central en :

1. Utilisant le Gestionnaire de référentiel pour ajouter un référentiel central sécurisé ou mettre à niveau un référentiel central non sécurisé existant.

2. Utilisant la CMC (Central Management Console) pour ajouter des utilisateurs.
3. Utilisant l'Administrateur pour ajouter des utilisateurs aux groupes des référentiels centraux.
4. Définissant la connexion à partir de Designer.
5. Ajoutant des objets au référentiel central ainsi que des autorisations d'affichage et modification des objets.

Informations associées

[Création d'un référentiel central sécurisé \[page 906\]](#)

[Ajout d'un administrateur multi-utilisateurs \(facultatif\) \[page 907\]](#)

[Configuration des groupes et utilisateurs \[page 907\]](#)

[Définition d'une connexion vers un référentiel central sécurisé \[page 908\]](#)

[Utilisation d'objets dans un référentiel central sécurisé \[page 908\]](#)

38.5.2 Création d'un référentiel central sécurisé

La première étape de l'établissement de mesures de sécurité pour le développement multi-utilisateurs consiste à créer un référentiel central sécurisé ou à mettre à niveau un référentiel central non sécurisé existant.

i Remarque

Ces procédures s'appliquent uniquement aux référentiels sécurisés.

Informations associées

[Configuration d'environnement multi-utilisateurs \[page 900\]](#)

38.5.2.1 Création d'un référentiel central sécurisé

1. Créez une base de données à utiliser pour le référentiel central à l'aide de votre système de gestion de base de données.
2. Sélectionnez ► [Démarrer](#) ► [Programmes](#) ► [SAP Data Services 4.2](#) ► [Gestionnaire de référentiel Data Services](#) ►.
3. Dans la fenêtre Gestionnaire de référentiel, cliquez sur le bouton [Central](#) dans le champ Type de référentiel et saisissez les informations de connexion au référentiel central.
4. Cochez la case [Activer la sécurité](#).
5. Cliquez sur [Créer](#).

Le progiciel crée les tables du référentiel dans la base de données que vous avez identifiée.

38.5.2.2 Mise à niveau d'un référentiel central, de non sécurisé à sécurisé

Vous pouvez modifier un référentiel central existant pour le sécuriser, mais cette modification ne peut pas être annulée.

1. Ouvrez le Gestionnaire de référentiel.
2. Dans la fenêtre Gestionnaire de référentiel, cliquez sur le bouton [Central](#) dans le champ Type de référentiel et saisissez les informations de connexion à la base de données du référentiel central à modifier.
3. Cochez la case [Activer la sécurité](#).
4. Cliquez sur [Mettre à niveau](#).

Le progiciel met à niveau les tables du référentiel dans la base de données que vous avez identifiée.

i Remarque

Lorsque vous mettez à niveau un référentiel central non sécurisé existant pour qu'il devienne sécurisé, un nouveau groupe, DIGroup, est automatiquement créé pour vous et affiché dans ► [Console de gestion](#) ► [Administrateur](#) ► [Référentiels centraux](#) ►. Pour accéder au référentiel, ajoutez des utilisateurs existants au groupe dans l'Administrateur.

38.5.3 Ajout d'un administrateur multi-utilisateurs (facultatif)

Dans la Central Management Console (CMC), vous avez la possibilité d'ajouter un utilisateur ayant le rôle d'Administrateur multi-utilisateurs. Ce rôle est limité à la gestion des référentiels centraux sécurisés, c'est donc un sous-ensemble du rôle d'Administrateur. Par exemple, les administrateurs multi-utilisateurs ne peuvent pas ajouter un référentiel local ou un référentiel central non sécurisé.

Les administrateurs multi-utilisateurs peuvent :

- ajouter et supprimer des référentiels centraux sécurisés ;
- gérer les utilisateurs et les groupes ;
- afficher les rapports des référentiels centraux sécurisés.

38.5.4 Configuration des groupes et utilisateurs

L'étape suivante de l'implémentation de la sécurité d'un référentiel central consiste à ajouter et configurer les groupes et utilisateurs avec la Central Management Console (CMC) et l'Administrateur.

38.5.5 Définition d'une connexion vers un référentiel central sécurisé

L'étape suivante de l'implémentation de la sécurité d'un référentiel central est la définition d'une connexion vers le répertoire dans Designer.

Cette procédure s'applique uniquement aux référentiels centraux sécurisés.

1. Lancez Designer et connectez-vous au référentiel local.
2. Dans le menu *Outils*, cliquez sur *Référentiels centraux* pour ouvrir la fenêtre Options.
L'option Connexions au référentiel central doit être sélectionnée dans la liste Designer.
3. Cliquez sur *Ajouter*.
4. Saisissez vos informations de connexion au CMS (Central Management Server) et cliquez sur *Se connecter*.
5. Sélectionnez le référentiel central sécurisé auquel vous désirez vous connecter.
6. Cliquez sur *OK*.
La liste de connexions au référentiel central comprend désormais le référentiel central venant d'être connecté et il est identifié comme sécurisé.
7. Le référentiel étant sélectionné, cliquez sur *Activer*.
8. Cliquez sur *OK*.

Informations associées

[Configuration d'environnement multi-utilisateurs \[page 900\]](#)

[Activation d'un référentiel central \[page 902\]](#)

38.5.6 Utilisation d'objets dans un référentiel central sécurisé

Informations associées

[Ajout d'objets au référentiel central \[page 911\]](#)

[Affichage et modification des autorisations \[page 908\]](#)

38.5.6.1 Affichage et modification des autorisations

Après avoir terminé toutes les tâches de configuration et l'ajout d'objets au référentiel central sécurisé, utilisez la bibliothèque d'objets centrale pour afficher et modifier les autorisations de groupe des objets.

38.5.6.1.1 Afficher les autorisations d'un objet

1. Lancez Designer et connectez-vous au référentiel local.
2. Ouvrez la bibliothèque d'objets centrale sécurisée.

Votre groupe par défaut s'affiche dans la liste déroulante en haut de la fenêtre et il est marqué d'un astérisque. La colonne Autorisations affiche le niveau d'accès du groupe actuel pour chaque objet. Si vous ajoutez un nouvel objet à la bibliothèque centrale, le groupe actuel obtient les autorisations COMPLETES et tous les autres groupes obtiennent l'autorisation LECTURE.

38.5.6.1.2 Modifier les autorisations des objets pour les autres groupes.

Vous devez disposer d'autorisations complètes pour modifier l'accès aux objets pour les autres groupes.

1. Dans la bibliothèque d'objets centrale, cliquez avec le bouton droit sur l'objet, puis cliquez sur [► Autorisation ► Configuration de l'adaptateur du CDM ► Objet](#) ou sur [► Autorisation ► Objet et dépendants](#).
2. La boîte de dialogue [Autorisation](#) s'ouvre, elle affiche une liste des groupes disponibles et le niveau d'accès du groupe pour les objets.
3. Cliquez dans la colonne [Autorisation](#), puis sélectionnez dans la liste déroulante un niveau d'autorisation pour le groupe.
4. Cliquez sur [Appliquer](#) ou [OK](#).

38.5.6.1.3 Modification du groupe actuel ou du groupe par défaut

1. Pour modifier le groupe actuel, sélectionnez un groupe dans la zone déroulante de la bibliothèque d'objets centrale.
 2. Pour modifier votre groupe par défaut, sélectionnez le groupe voulu dans la liste déroulante et cliquez sur l'icône Enregistrer.
- Le progiciel marque le groupe par défaut d'un astérisque.

38.6 Travail dans un environnement multi-utilisateurs

Pour des résultats optimum du développement au sein d'un environnement multi-utilisateurs, certains processus sont recommandés, tels que le rechargement et l'extraction d'objets modifiés, ainsi que l'établissement d'un ensemble de conventions respectées par votre équipe, comme l'étiquetage des objets.

[Filtrage \[page 910\]](#)

[Ajout d'objets au référentiel central \[page 911\]](#)

[Extraction d'objets \[page 913\]](#)

[Annulation de l'extraction \[page 916\]](#)

[Rechargement d'objets \[page 917\]](#)

[Étiquetage d'objets \[page 919\]](#)

[Obtention d'objets \[page 922\]](#)

[Comparaison d'objets \[page 924\]](#)

[Affichage de l'historique de l'objet \[page 924\]](#)

L'une des caractéristiques principales du référentiel central est qu'il conserve l'historique des objets à des fins d'information et de reporting.

[Suppression d'objets \[page 927\]](#)

38.6.1 Filtrage

SAP Data Services permet de personnaliser des informations spécifiques à l'environnement dans les définitions d'objets en les filtrant (en les modifiant par une sélection). Les objets d'application peuvent contenir des informations spécifiques aux référentiels. Par exemple, les banques et bases de données peuvent faire référence à une connexion à la base de données spécifique à un utilisateur ou à une phase du développement. Lorsque plusieurs utilisateurs travaillent sur une application, ils peuvent modifier les informations spécifiques aux référentiels.

Le filtrage permet notamment de :

- Modifier les informations de connexion d'une banque ou d'une base de données
- Modifier le répertoire racine de fichiers associés à un format de fichier particulier
- Sélectionner ou effacer certains objets dépendants

Le processus de filtrage est accessible lors de l'ajout, du rechargement, de l'extraction ou de l'obtention des derniers objets ou des objets étiquetés dans un référentiel central.

Lors de la sélection d'une commande utilisant l'option de filtrage :

1. La fenêtre *Confirmation du contrôle de version* affiche les objets sélectionnés et tous les objets dépendants. Vous pouvez exclure des objets en les sélectionnant puis en remplaçant la valeur du *statut cible créer* par *exclure*.
2. La fenêtre *Options de la banque de données* affiche toutes les banques de données utilisées par l'objet. Cette fenêtre ne s'ouvre que si les objets ajoutés, rechargés ou extraits incluent une banque de données.

Rubrique parent : [Travail dans un environnement multi-utilisateurs \[page 909\]](#)

Informations associées

[Ajout d'objets au référentiel central \[page 911\]](#)

[Extraction d'objets \[page 913\]](#)

[Annulation de l'extraction \[page 916\]](#)
[Rechargement d'objets \[page 917\]](#)
[Etiquetage d'objets \[page 919\]](#)
[Obtention d'objets \[page 922\]](#)
[Comparaison d'objets \[page 924\]](#)
[Affichage de l'historique de l'objet \[page 924\]](#)
[Suppression d'objets \[page 927\]](#)

38.6.2 Ajout d'objets au référentiel central

Une fois le référentiel central créé, connecté au référentiel local puis activé, vous pouvez ajouter des objets du référentiel local au référentiel central. Souvenez-vous : l'ensemble de la conception (création des jobs, des workflows et des flux de données) s'effectue dans un référentiel local. Vous utilisez ainsi un référentiel local pour la création initiale de tout objet dans une application. Après la création initiale d'un objet, vous l'ajoutez au référentiel central. Une fois dans le référentiel central, l'objet est soumis au contrôle de version et peut être partagé entre plusieurs utilisateurs.

Vous pouvez ajouter un seul objet ou un objet et tous ses objets dépendants au référentiel central. Lors de l'ajout d'un objet unique tel qu'un flux de données, seul cet objet est ajouté. Aucun objet dépendant n'est ajouté.

Vous pouvez ajouter des objets au référentiel central à tout moment. Toutefois, il est impossible d'ajouter un objet déjà existant dans le référentiel central.

Il est impossible d'ajouter une configuration de transformation en lecture seule au référentiel. Vous pouvez cependant répliquer une configuration de transformation puis ajouter la réplique au référentiel.

Rubrique parent : [Travail dans un environnement multi-utilisateurs \[page 909\]](#)

Informations associées

[Filtrage \[page 910\]](#)
[Extraction d'objets \[page 913\]](#)
[Annulation de l'extraction \[page 916\]](#)
[Rechargement d'objets \[page 917\]](#)
[Etiquetage d'objets \[page 919\]](#)
[Obtention d'objets \[page 922\]](#)
[Comparaison d'objets \[page 924\]](#)
[Affichage de l'historique de l'objet \[page 924\]](#)
[Suppression d'objets \[page 927\]](#)

38.6.2.1 Ajouter un objet unique au référentiel central

1. Ouvrez la bibliothèque d'objets locale.
2. Cliquez avec le bouton droit de la souris sur l'objet et sélectionnez ► *Ajouter au référentiel central* ► *Objet* ►.
3. La fenêtre Commentaires s'ouvre. Saisissez un commentaire dans le champ *Commentaires* puis cliquez sur *OK*.

Le progiciel ajoute l'objet au référentiel central actif.

i Remarque

La commande *Ajouter au référentiel central* n'est pas disponible si l'objet existe déjà dans le référentiel central.

38.6.2.2 Ajouter un objet et ses objets dépendants au référentiel central

1. Ouvrez la bibliothèque d'objets locale.
2. Cliquez avec le bouton droit de la souris sur l'objet puis sélectionnez ► *Ajouter au référentiel central* ► *Objet et dépendants* ► ou ► *Ajouter au référentiel central* ► *Avec filtrage* ► (si le filtrage est requis).
3. La fenêtre Commentaires s'ouvre. Saisissez un commentaire dans le champ *Commentaires* puis cliquez sur *OK*.
4. Si vous avez sélectionné *Avec filtrage*, renseignez les fenêtres de filtrage.
5. Cliquez sur *Terminer* pour ajouter les objets sélectionnés.

Il est également possible de sélectionner l'objet et de le faire glisser vers la bibliothèque d'objets centrale pour ajouter l'objet et ses objets dépendants au référentiel central. Les fenêtres de filtrage s'affichent.

i Remarque

La commande *Ajouter au référentiel central* n'est pas disponible si l'objet existe déjà dans le référentiel central. Toutefois, la commande *Ajouter au référentiel central* est disponible si les dépendants de l'objet existent déjà dans le référentiel central et non l'objet lui-même.

i Remarque

Il est impossible d'ajouter une configuration de transformation en lecture seule au référentiel. Pour exécuter cette action, il est nécessaire de créer un nouveau référentiel, de mettre à niveau le référentiel existant ou d'importer l'ATL contenant une nouvelle version des configurations de transformation en lecture seule.

Informations associées

[Filtrage \[page 910\]](#)

38.6.3 Extraction d'objets

Lorsque vous envisagez de modifier un ou plusieurs objets d'une application, il est conseillé d'extraire les objets concernés. Lors de l'extraction, l'objet n'est plus disponible pour les autres utilisateurs. Ils peuvent afficher l'objet mais ne peuvent pas le modifier. L'extraction d'un objet permet d'éviter toute modification conflictuelle d'objet effectuée simultanément par deux utilisateurs.

Data Services modifie les icônes de l'objet dans les bibliothèques d'objets locale et centrale pour indiquer que l'objet est extrait.

Lorsque vous avez extrait un objet, la bibliothèque d'objets centrale affiche le référentiel central qui a effectué l'extraction. En fonction du nom du référentiel, il est possible de déterminer l'utilisateur travaillant avec cet objet.

Pour afficher les modifications périodiques, actualisez la bibliothèque d'objets centrale en cliquant sur le bouton [Actualiser la bibliothèque d'objets centrale](#) dans la barre d'outils de la bibliothèque d'objets centrale.

Sélectionnez une commande d'extraction selon l'action souhaitée sur l'objet.

Rubrique parent : [Travail dans un environnement multi-utilisateurs \[page 909\]](#)

Informations associées

[Filtrage \[page 910\]](#)

[Ajout d'objets au référentiel central \[page 911\]](#)

[Annulation de l'extraction \[page 916\]](#)

[Rechargement d'objets \[page 917\]](#)

[Etiquetage d'objets \[page 919\]](#)

[Obtention d'objets \[page 922\]](#)

[Comparaison d'objets \[page 924\]](#)

[Affichage de l'historique de l'objet \[page 924\]](#)

[Suppression d'objets \[page 927\]](#)

38.6.3.1 Extraire des objets uniques ou des objets avec leurs dépendants

Les dépendants sont des objets utilisés par d'autres objets, par exemple des flux de données appelés à partir d'un workflow. Vous pouvez extraire un seul objet ou un objet et tous ses objets dépendants (tels que calculés

dans le référentiel central). Il est par exemple possible d'extraire un workflow uniquement. Dans ce cas, vous pouvez modifier ce workflow, en lui ajoutant un nouveau script par exemple. Toutefois, vous ne pouvez pas modifier d'objets dépendants, comme les flux de données, dans le workflow et conserver les modifications dans le référentiel central. Les modifications effectuées sur les objets dépendants seront uniquement conservées dans le référentiel local. Il est également possible d'extraire le workflow et tous ses dépendants. Dans ce cas, vous pouvez modifier le workflow ou ses dépendants, puis conserver ces modifications dans les référentiels central et local.

Il est en général plus sûr d'extraire un objet avec l'ensemble de ses objets dépendants. Vous empêchez ainsi d'autres utilisateurs de modifier des objets dépendants par inadvertance.

38.6.3.1.1 Extraction d'un objet unique

1. Ouvrez la bibliothèque d'objets centrale.
2. Cliquez avec le bouton droit de la souris sur l'objet que vous souhaitez extraire.
3. Sélectionnez ► *Extraire* ► *Objet* ►.

Il est également possible de sélectionner l'objet dans la bibliothèque d'objets centrale, puis de cliquer sur le bouton *Extraire l'objet* situé en haut de la bibliothèque d'objets centrale.

Le progiciel copie la dernière version de l'objet sélectionné du référentiel central vers votre référentiel local, puis marque l'objet comme extrait.

38.6.3.1.2 Extraction d'un objet et de ses objets dépendants

1. Ouvrez la bibliothèque d'objets centrale.
2. Cliquez avec le bouton droit de la souris sur l'objet que vous souhaitez extraire.
3. Sélectionnez ► *Extraire* ► *Objet et dépendants* ►.

Il est également possible de sélectionner l'objet dans la bibliothèque d'objets centrale, puis de cliquer sur le bouton *Extraire l'objet et les dépendants* situé en haut de la bibliothèque d'objets centrale.

SAP Data Services copie la version la plus récente de l'objet sélectionné et de tous ses objets dépendants depuis le référentiel central, puis marque ces objets comme extraits.

Si vous ou un autre utilisateur extrayez un objet dépendant, le progiciel vous le signale à l'aide d'une fenêtre Alerte d'extraction et vous invite à obtenir la dernière version de l'objet extrait.

Informations associées

[Obtention d'objets \[page 922\]](#)

38.6.3.2 Extraire des objets uniques ou des objets avec leurs dépendants sans remplacement

Lors de l'extraction d'un objet, il est possible de remplacer l'objet de votre référentiel local par la dernière version de celui-ci dans le référentiel central. Vous pouvez aussi laisser sa version actuelle intacte dans votre référentiel local.

Lorsque vous extrayez un objet, SAP Data Services copie sa définition à partir du référentiel central et remplace toute définition d'objet présente dans votre référentiel local.

Vous pouvez extraire des objets sans les remplacer dans votre référentiel local. Supposons que vous travailliez dans votre référentiel local et que vous modifiez un objet non extrait. Si vous jugez que la modification optimise la conception ou les performances de l'application, vous souhaitez inclure cette modification dans le référentiel central.

Pour ce faire, il convient d'extraire l'objet sans remplacer l'objet présent dans le référentiel local, celui que vous avez déjà optimisé par une modification. Rechargez ensuite l'objet modifié dans le référentiel central.

i Remarque

Soyez vigilant lors de l'extraction d'objets sans remplacement de leur version dans le référentiel local. Si vous ne remplacez pas les versions d'objets dans le référentiel local, vous pouvez perdre les modifications intégrées par d'autres utilisateurs.

38.6.3.2.1 Extraire un objet ou un objet et ses objets dépendants sans remplacement

1. Ouvrez la bibliothèque d'objets centrale.
2. Cliquez avec le bouton droit de la souris sur l'objet que vous souhaitez extraire, puis sélectionnez **► Extraire ► Objet ► sans remplacement** pour extraire l'objet unique, ou sélectionnez **► Extraire ► Objet et dépendants sans remplacement** pour extraire l'objet et tous ses objets dépendants.

SAP Data Services marque tous les objets appropriés comme extraits (dans la bibliothèque d'objets et dans l'espace de travail), mais ne copie aucun objet du référentiel central vers le référentiel local.

38.6.3.3 Extraire des objets avec filtrage

Lors de l'extraction d'un objet avec filtrage, l'objet et tous ses objets dépendants sont extraits.

i Remarque

L'extraction remplace alors toujours les versions locales par les objets filtrés du référentiel central.

38.6.3.3.1 Extraction d'un objet et de ses objets dépendants avec filtrage

1. Ouvrez la bibliothèque d'objets centrale.
2. Cliquez avec le bouton droit de la souris sur l'objet que vous souhaitez extraire, puis sélectionnez .
3. Renseignez les fenêtres de filtrage.
4. Cliquez sur *Terminer* pour extraire les objets sélectionnés.

Informations associées

[Filtrage \[page 910\]](#)

38.6.4 Annulation de l'extraction

Vous pouvez parfois juger l'extraction d'un objet non nécessaire car vous n'aviez effectué aucune modification. Ou vous pouvez estimer inutiles les modifications effectuées sur un objet extrait. Vous préférez ainsi laisser la copie principale de l'objet telle quelle. Dans ces cas, vous pouvez annuler l'extraction.

Lors de l'annulation d'une extraction :

- L'objet dans le référentiel central reste tel qu'il était avant l'extraction. Aucune modification n'est effectuée et aucune autre version n'est enregistrée dans le référentiel central. Seul le statut de l'objet change : il n'est plus extrait, mais disponible.
- La version locale de l'objet conserve les modifications effectuées. Si vous souhaitez que l'objet local soit une copie exacte de l'objet central, appliquez l'opération *Obtention des derniers éléments* à cet objet.

Après l'annulation d'une extraction, les autres utilisateurs peuvent extraire et modifier l'objet.

Rubrique parent : [Travail dans un environnement multi-utilisateurs \[page 909\]](#)

Informations associées

[Filtrage \[page 910\]](#)

[Ajout d'objets au référentiel central \[page 911\]](#)

[Extraction d'objets \[page 913\]](#)

[Rechargement d'objets \[page 917\]](#)

[Etiquetage d'objets \[page 919\]](#)

[Obtention d'objets \[page 922\]](#)

[Comparaison d'objets \[page 924\]](#)

[Affichage de l'historique de l'objet \[page 924\]](#)

38.6.4.1 Annuler l'extraction d'un objet unique

1. Ouvrez la bibliothèque d'objets centrale.
2. Sélectionnez un objet extrait
3. Cliquez sur le bouton [Annuler l'extraction de l'objet](#). Vous pouvez également cliquer avec le bouton droit de la souris sur l'objet, puis sélectionner ► [Annuler l'extraction](#) ► [Objet](#) ►.

SAP Data Services supprime le symbole d'extraction et ne modifie pas l'objet dans le référentiel central. Tout objet dépendant extrait reste extrait.

38.6.4.2 Annuler l'extraction d'un objet et de ses dépendants

1. Ouvrez la bibliothèque d'objets centrale.
2. Sélectionnez l'objet extrait de niveau le plus élevé pour lequel vous souhaitez annuler l'extraction.
3. Cliquez sur le bouton [Annuler l'extraction de l'objet et des dépendants](#). Vous pouvez également cliquer avec le bouton droit de la souris sur l'objet dans la bibliothèque d'objets centrale, puis sélectionner ► [Annuler l'extraction](#) ► [Objet et dépendants](#) ►.

SAP Data Services supprime les symboles d'extraction de l'objet et de tous les objets dépendants également extraits. Ces objets ne sont pas modifiés dans le référentiel central.

38.6.5 Rechargement d'objets

Une fois les objets extraits modifiés, vous devez les recharger dans le référentiel central. Le rechargement d'objets crée une nouvelle version dans le référentiel central et permet aux autres utilisateurs d'obtenir les modifications effectuées. Une copie des modifications est également conservée à des fins de contrôle de révision. Vous pourrez obtenir ultérieurement une version donnée d'un objet rechargé puis la comparer aux modifications suivantes ou même rétablir la version précédente.

Rechargez un objet lorsque vos modifications sont terminées, lorsque d'autres utilisateurs ont besoin de l'objet contenant vos modifications ou lorsque vous souhaitez conserver une copie de l'objet en son statut actuel.

Sélectionnez une commande de rechargement selon l'action souhaitée sur l'objet.

Rubrique parent : [Travail dans un environnement multi-utilisateurs \[page 909\]](#)

Informations associées

[Filtrage \[page 910\]](#)

[Ajout d'objets au référentiel central \[page 911\]](#)
[Extraction d'objets \[page 913\]](#)
[Annulation de l'extraction \[page 916\]](#)
[Etiquetage d'objets \[page 919\]](#)
[Obtention d'objets \[page 922\]](#)
[Comparaison d'objets \[page 924\]](#)
[Affichage de l'historique de l'objet \[page 924\]](#)
[Suppression d'objets \[page 927\]](#)
[Rechargement d'objets uniques et d'objets avec leurs dépendants \[page 918\]](#)
[Rechargement d'un objet avec filtrage \[page 919\]](#)

38.6.5.1 Rechargement d'objets uniques et d'objets avec leurs dépendants

Tout comme vous pouvez extraire un objet unique ou un objet avec l'ensemble de ses objets dépendants, vous pouvez recharger un objet unique ou un objet avec tous ses objets dépendants extraits (tels que calculés dans le référentiel local).

38.6.5.1.1 Rechargement d'un objet unique

1. Ouvrez la bibliothèque d'objets centrale.
2. Sélectionnez l'objet que vous souhaitez recharger.
3. Cliquez sur le bouton [Recharger l'objet](#) situé en haut de la bibliothèque d'objets centrale.

Vous pouvez également cliquer avec le bouton droit de la souris sur l'objet dans la bibliothèque d'objets centrale, puis sélectionner ► [Recharger](#) ► [Objet](#) ►.

4. La fenêtre Recharger s'ouvre avec une zone [Commentaire](#) dans laquelle vous pouvez saisir des commentaires. Une fois les commentaires saisis, cliquez sur [OK](#).

SAP Data Services copie l'objet du référentiel local vers le référentiel central, puis supprime la marque d'extraction.

38.6.5.1.2 Rechargement d'un objet et de ses objets dépendants

1. Ouvrez la bibliothèque d'objets centrale.
2. Sélectionnez l'objet de niveau le plus élevé que vous souhaitez recharger.
3. Cliquez sur le bouton [Recharger l'objet et les dépendants](#) situé en haut de la bibliothèque d'objets centrale.

Vous pouvez également cliquer avec le bouton droit de la souris sur l'objet dans la bibliothèque d'objets centrale, puis sélectionner ► [Recharger](#) ► [Objet et dépendants](#) ►.

4. La fenêtre Recharger s'ouvre avec une zone [Commentaire](#) dans laquelle vous pouvez saisir des commentaires. Une fois les commentaires saisis, cliquez sur [OK](#).

SAP Data Services copie l'objet sélectionné et tous ses objets dépendants depuis votre référentiel vers le référentiel central, puis supprime la marque d'extraction.

38.6.5.2 Rechargement d'un objet avec filtrage

Tout comme vous pouvez extraire des objets avec filtrage, vous pouvez recharger des objets avec filtrage. Lors du rechargement d'un objet avec filtrage, l'objet et tous ses objets dépendants sont rechargés.

38.6.5.2.1 Rechargement d'un objet avec filtrage

1. Ouvrez la bibliothèque d'objets centrale.
2. Cliquez avec le bouton droit de la souris sur l'objet que vous souhaitez recharger, puis sélectionnez [Recharger](#) > [Avec filtrage](#).

3. La fenêtre Recharger s'ouvre avec une zone [Commentaire](#) dans laquelle vous pouvez saisir des commentaires. Une fois les commentaires saisis, cliquez sur [OK](#).

SAP Data Services vous avertit que vous êtes sur le point de créer une nouvelle version de l'objet dans le référentiel central.

4. Cliquez sur [Oui](#) pour poursuivre le rechargement.
5. Renseignez les fenêtres de filtrage.
6. Cliquez sur [Terminer](#) pour recharger les objets sélectionnés.

Informations associées

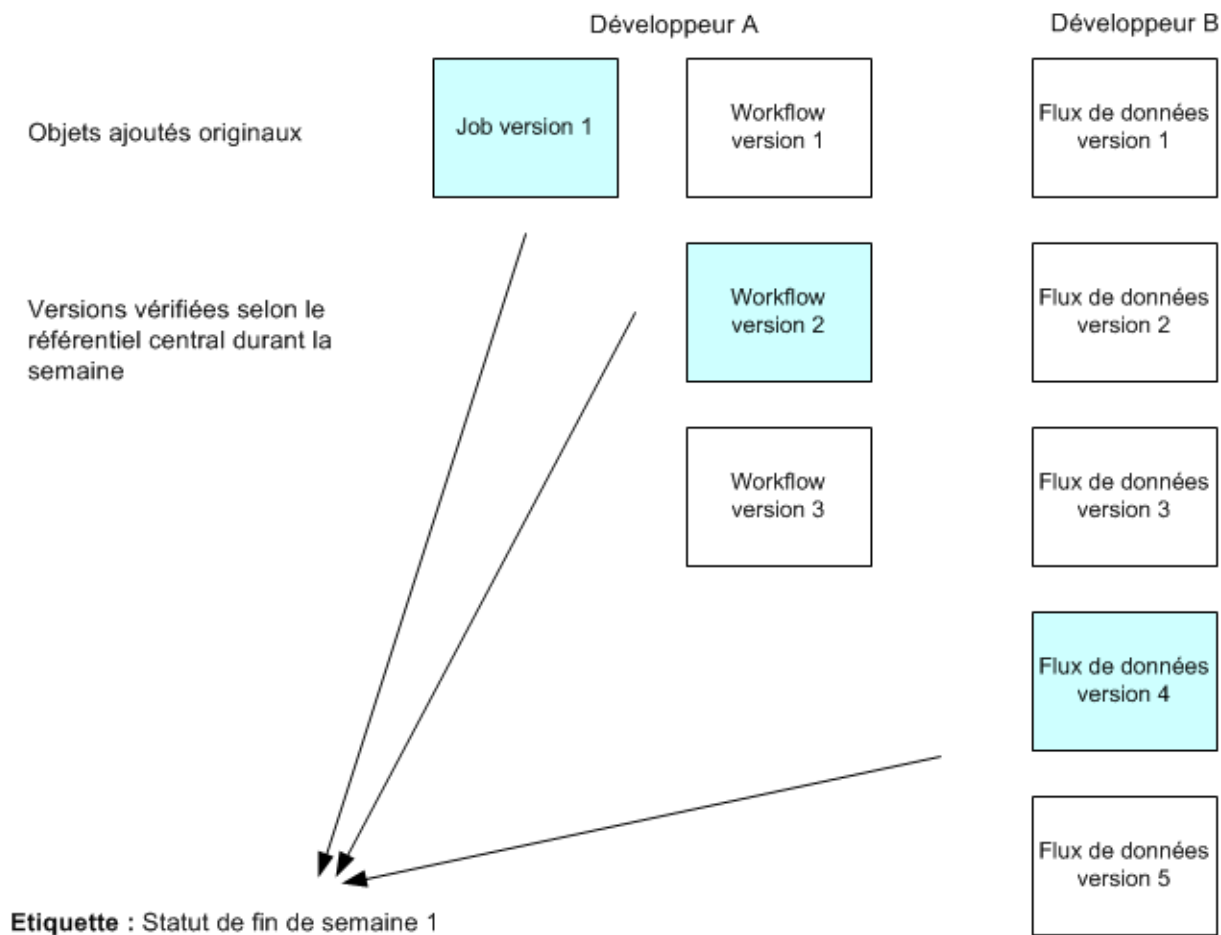
[Filtrage \[page 910\]](#)

38.6.6 Etiquetage d'objets

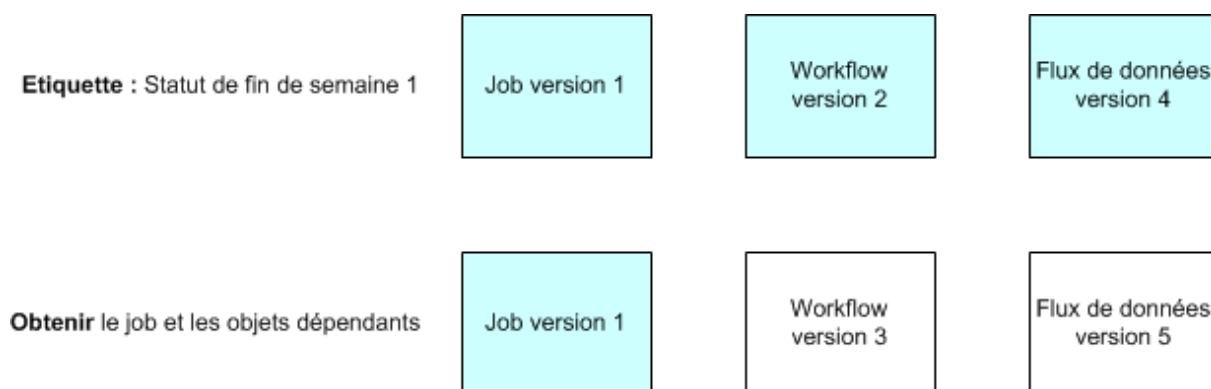
Pour vous aider à organiser et à suivre le statut des objets dans l'application, il est possible d'étiqueter les objets. Vous pouvez choisir d'étiqueter un objet ou un objet et tous ses objets dépendants. Une étiquette décrit non seulement un objet, mais elle permet également de conserver une relation entre les différentes versions d'un objet.

Supposons que le développeur A ajoute un job au référentiel central et travaille sur un workflow dans ce job. Le développeur B travaille simultanément sur un flux de données dans le même job. A la fin de la semaine, une fois que le développeur A a rechargé les deux versions du workflow et le développeur B les quatre versions du flux de données dans le référentiel central, le job est étiqueté "Statut Fin de semaine 1". Cette étiquette contient la

version 1 du job, la version 2 du workflow et la version 4 du flux de données. Les deux développeurs peuvent poursuivre les modifications dans leurs workflow et flux de données respectifs.



Si vous souhaitez obtenir ultérieurement le job avec la version du flux de données et cette étiquette, il convient d'obtenir le job par son étiquette et non d'extraire le job et ses dépendants.



L'étiquette "Statut Fin de semaine 1" permet de collecter les versions du workflow et du flux de données rechargées à la fin de la semaine. Sans cette étiquette, il faudrait obtenir une version particulière de chaque objet afin de rassembler la collection d'objets portant l'étiquette "Statut Fin de semaine 1".

Rubrique parent : [Travail dans un environnement multi-utilisateurs \[page 909\]](#)

Informations associées

[Filtrage \[page 910\]](#)

[Ajout d'objets au référentiel central \[page 911\]](#)

[Extraction d'objets \[page 913\]](#)

[Annulation de l'extraction \[page 916\]](#)

[Rechargement d'objets \[page 917\]](#)

[Obtention d'objets \[page 922\]](#)

[Comparaison d'objets \[page 924\]](#)

[Affichage de l'historique de l'objet \[page 924\]](#)

[Suppression d'objets \[page 927\]](#)

[Obtention d'objets \[page 922\]](#)

38.6.6.1 Etiqueter un objet et ses dépendants

1. Ouvrez la bibliothèque d'objets centrale.
2. Cliquez avec le bouton droit de la souris sur l'objet que vous souhaitez étiqueter, puis sélectionnez  *Dernière version d'étiquette*  *Objet*  pour étiqueter uniquement l'objet mis en surbrillance ou sélectionnez *Objet et dépendants* pour étiqueter l'objet et tous ses objets associés.

La fenêtre *Dernière version d'étiquette* s'ouvre.

3. Dans la zone *Etiquette*, saisissez un texte décrivant le statut actuel de l'objet, puis cliquez sur *OK*.

L'étiquette est insérée dans l'historique de l'objet et des dépendants.

Informations associées

[Affichage de l'historique de l'objet \[page 924\]](#)

38.6.6.2 Obtenir un objet étiqueté avec filtrage

L'option de filtrage associée à l'opération Obtention par étiquette vous permet de filtrer (modification sélective) des informations spécifiques à l'environnement dans les définitions d'objet lorsque vous travaillez dans un environnement multi-utilisateur.

1. Ouvrez la bibliothèque d'objets centrale.
2. Sélectionnez l'objet de niveau le plus élevé que vous souhaitez obtenir.
3. Cliquez avec le bouton droit sur l'objet dans la bibliothèque d'objets centrale, puis sélectionnez  *Obtention par étiquette*  *Avec filtrage* .
4. Renseignez la fenêtre de filtrage.
5. Cliquez sur *Terminer* pour obtenir les objets sélectionnés.

Informations associées

[Filtrage \[page 910\]](#)

38.6.7 Obtention d'objets

Pour vous assurer que le référentiel est à jour, vous "obtenez" des objets. Lorsque vous obtenez un objet, vous copiez la dernière version de cet objet dans la bibliothèque d'objets centrale et la collez dans le référentiel local, remplaçant ainsi la version du référentiel local. L'obtention d'un objet n'engendre pas son extraction. L'objet reste disponible pour les autres utilisateurs ; ils peuvent l'extraire et le modifier.

Vous pouvez obtenir un objet avec ou sans ses objets dépendants et avec ou sans filtrage.

Rubrique parent : [Travail dans un environnement multi-utilisateurs \[page 909\]](#)

Informations associées

[Filtrage \[page 910\]](#)

[Ajout d'objets au référentiel central \[page 911\]](#)

[Extraction d'objets \[page 913\]](#)

[Annulation de l'extraction \[page 916\]](#)

[Rechargement d'objets \[page 917\]](#)

[Étiquetage d'objets \[page 919\]](#)

[Comparaison d'objets \[page 924\]](#)

[Affichage de l'historique de l'objet \[page 924\]](#)

[Suppression d'objets \[page 927\]](#)

[Affichage de l'historique de l'objet \[page 924\]](#)

38.6.7.1 Obtention d'un objet unique

1. Ouvrez la bibliothèque d'objets centrale.
2. Sélectionnez l'objet que vous souhaitez obtenir.
3. Cliquez sur le bouton *Obtention de la dernière version de l'objet* situé en haut de la bibliothèque centrale. Vous pouvez également cliquer avec le bouton droit de la souris sur l'objet dans la bibliothèque d'objets centrale, puis sélectionner ► *Obtention de dernière version* ► *Objet* ►.

La version la plus récente de l'objet dans le référentiel central est copiée vers le référentiel local.

38.6.7.2 Obtenir un objet et ses objets dépendants

1. Ouvrez la bibliothèque d'objets centrale.
2. Sélectionnez l'objet de niveau le plus élevé que vous souhaitez obtenir.
3. Cliquez sur le bouton *Obtention de la dernière version de l'objet avec dépendants* situé en haut de la bibliothèque centrale. Vous pouvez également cliquer avec le bouton droit de la souris sur l'objet dans la bibliothèque d'objets centrale, puis sélectionner ► *Obtention de dernière version* ► *Objet et dépendants* ►.

La version la plus récente de l'objet sélectionné et tous les objets dépendants du référentiel central sont copiés vers votre référentiel local.

38.6.7.3 Obtenir un objet et ses objets dépendants avec filtrage

1. Ouvrez la bibliothèque d'objets centrale.
2. Sélectionnez l'objet de niveau le plus élevé que vous souhaitez obtenir.
3. Cliquez avec le bouton droit de la souris sur l'objet dans la bibliothèque d'objets centrale, puis sélectionnez ► *Obtention de dernière version* ► *Avec filtrage* ►.
4. Renseignez les fenêtres de filtrage.
5. Cliquez sur *Terminer* pour obtenir les objets sélectionnés.

Informations associées

[Filtrage \[page 910\]](#)

38.6.8 Comparaison d'objets

SAP Data Services permet de comparer deux objets des référentiels local et central pour déterminer leurs différences.

Rubrique parent : [Travail dans un environnement multi-utilisateurs \[page 909\]](#)

Informations associées

[Filtrage \[page 910\]](#)

[Ajout d'objets au référentiel central \[page 911\]](#)

[Extraction d'objets \[page 913\]](#)

[Annulation de l'extraction \[page 916\]](#)

[Rechargement d'objets \[page 917\]](#)

[Etiquetage d'objets \[page 919\]](#)

[Obtention d'objets \[page 922\]](#)

[Affichage de l'historique de l'objet \[page 924\]](#)

[Suppression d'objets \[page 927\]](#)

38.6.9 Affichage de l'historique de l'objet

L'une des caractéristiques principales du référentiel central est qu'il conserve l'historique des objets à des fins d'information et de reporting.

Le référentiel central conserve un historique de toutes les modifications d'objets effectuées dans le référentiel central. Utilisez cet historique pour gérer et contrôler le développement de l'application.

Lorsque vous affichez des objets dans la bibliothèque d'objets centrale dans Designer, vous pouvez rapidement afficher la version, la date et les informations d'horodatage. Affichez des informations plus détaillées en sélectionnant l'icône [Afficher historique](#).

Rubrique parent : [Travail dans un environnement multi-utilisateurs \[page 909\]](#)

Informations associées

[Filtrage \[page 910\]](#)

[Ajout d'objets au référentiel central \[page 911\]](#)

[Extraction d'objets \[page 913\]](#)

[Annulation de l'extraction \[page 916\]](#)

[Rechargement d'objets \[page 917\]](#)

[Etiquetage d'objets \[page 919\]](#)
[Obtention d'objets \[page 922\]](#)
[Comparaison d'objets \[page 924\]](#)
[Suppression d'objets \[page 927\]](#)

38.6.9.1 Examiner l'historique d'un objet

Les utilisateurs et les administrateurs peuvent suivre les modifications apportées à chaque version d'un objet en examinant l'historique de l'objet.

1. Sélectionnez **Outils** > **Bibliothèque d'objets centrale** dans SAP Data Services Designer.

La bibliothèque d'objets centrale s'ouvre en mode flottant.

2. Cliquez sur un objet pour le sélectionner.

Analysez les informations de l'objet dans la bibliothèque d'objets centrale, pour savoir si :

- un objet est extrait. Un objet est extrait lorsqu'une croix rouge apparaît en regard du nom d'objet. Un objet est disponible pour extraction si aucune croix rouge n'apparaît en regard du nom d'objet.
- Emplacement de l'objet extrait. Lorsqu'un objet est extrait, le nom du référentiel local auquel l'objet appartient, est extrait.
- Informations relatives à la dernière modification : pour chaque objet référencé, le numéro de version actuelle, la date et l'heure auxquelles l'objet a été contrôlé pour la dernière fois dans le référentiel central et une description de la dernière modification.

3. Cliquez sur le bouton **Afficher l'historique** situé en haut de la bibliothèque d'objets centrale.

Vous pouvez également cliquer avec le bouton droit de la souris sur l'objet dans la bibliothèque d'objets centrale, puis sélectionner **Afficher l'historique**. La boîte de dialogue **Historique** affiche des informations sur chaque révision de l'objet.

Le tableau suivant compare les informations disponibles sur la page d'historique et dans la bibliothèque d'objets centrale.

Colonne	Page Historique	Bibliothèque d'objets centrale
<i>Version</i>	Numéro de révision de l'objet. Dès qu'un utilisateur enregistre l'objet, le progiciel en crée une nouvelle version.	Identique
<i>Étiquette</i>	Texte saisi par l'utilisateur pour décrire le statut de l'objet à un moment donné.	N/A
<i>Référentiel</i>	Informations sur le référentiel local à partir duquel le progiciel a enregistré cette version de l'objet et le nom d'utilisateur.	Si l'objet est actuellement extrait, le nom du référentiel local à l'objet a été extrait.

Colonne	Page Historique	Bibliothèque d'objets centrale
<i>Date</i>	Date et heure de l'enregistrement de cette version de l'objet par le progiciel.	Identique
<i>Action</i>	Type de modification effectuée par l'utilisateur sur l'objet. Cette table enregistre des actions telles que : Rechargé : l'utilisateur a rechargé l'objet	N/A
<i>Commentaire</i>	Commentaires saisis par l'utilisateur lors de l'ajout d'un objet ou de son rechargement dans un référentiel central.	N/A
<i>Description</i>	N/A	Description de l'objet saisi lors de la création de l'objet.

Lorsque le référentiel central sécurisé, les utilisateurs peuvent voir le nom d'utilisateur de la personne qui a extrait l'objet et son autorisation.

38.6.9.2 Obtenir une version précédente d'un objet

1. Sélectionnez un objet.
2. Cliquez sur le bouton *Afficher l'historique* situé en haut de la bibliothèque d'objets centrale.
3. Cliquez sur la version de l'objet souhaitée.
4. Cliquez sur le bouton *Obtention de l'objet par version*.

i Remarque

Lors de l'obtention d'une version précédente d'un objet, vous n'obtenez que l'objet lui-même, sans ses objets dépendants.

38.6.9.3 Obtenir un objet avec une étiquette donnée

1. Sélectionnez un objet.
2. Cliquez sur le bouton *Afficher l'historique* situé en haut de la bibliothèque d'objets centrale.
3. Cliquez sur la version de l'objet avec l'étiquette particulière souhaitée.
4. Cliquez sur le bouton *Obtention par étiquette*.

38.6.10 Suppression d'objets

Vous pouvez supprimer des objets du référentiel central ou d'un référentiel local. Pour supprimer un objet du référentiel central, cliquez avec le bouton droit de la souris sur l'objet dans la bibliothèque centrale d'objets puis sélectionnez [Supprimer](#). Pour supprimer un objet du référentiel local, cliquez avec le bouton droit de la souris sur l'objet dans la bibliothèque d'objets puis sélectionnez [Supprimer](#).

Lorsque vous supprimez un objet d'un référentiel local, vous ne le supprimez pas automatiquement du référentiel central actif. Vous pouvez en fait obtenir cet objet à partir du référentiel central pour l'insérer à nouveau.

De même, lors de la suppression d'un objet d'un référentiel central, vous ne le supprimez pas automatiquement des référentiels locaux connectés. A moins que vous ne supprimiez l'objet du référentiel local, vous pouvez ajouter de nouveau cet objet au référentiel central.

Lorsque vous supprimez des objets du référentiel central, vous ne supprimez que l'objet sélectionné et l'ensemble de ses versions. Vous ne supprimez aucun objet dépendant.

Rubrique parent : [Travail dans un environnement multi-utilisateurs \[page 909\]](#)

Informations associées

[Filtrage \[page 910\]](#)

[Ajout d'objets au référentiel central \[page 911\]](#)

[Extraction d'objets \[page 913\]](#)

[Annulation de l'extraction \[page 916\]](#)

[Rechargement d'objets \[page 917\]](#)

[Etiquetage d'objets \[page 919\]](#)

[Obtention d'objets \[page 922\]](#)

[Comparaison d'objets \[page 924\]](#)

[Affichage de l'historique de l'objet \[page 924\]](#)

38.7 Migration de jobs multi-utilisateurs

Vue d'ensemble de la migration de jobs multi-utilisateurs

La migration de jobs s'applique à SAP Data Services sur plusieurs niveaux : au niveau des applications, de la gestion des référentiels et de la mise à niveau des produits. La migration d'application est beaucoup plus flexible au sein d'un environnement multi-utilisateurs, permettant de conserver non seulement plusieurs versions des objets lors du développement, mais aussi lors des phases de test et de production le cas échéant.

Informations associées

[Gestion des phases des applications \[page 928\]](#)

[Copie de contenu entre référentiels centraux \[page 929\]](#)

[Migration du référentiel central \[page 930\]](#)

38.7.1 Gestion des phases des applications

Les applications traversent en général différentes phases, du développement à la production. Par exemple, une application peut traverser trois phases :

- Les développeurs créent une application.
- Les testeurs valident l'application.
- Les administrateurs exécutent l'application.

Un référentiel central unique peut prendre en charge l'application pendant l'ensemble des phases. Utilisez l'étiquetage de job et les projets pour conserver les composants d'application de manière indépendante pour chaque phase. Par exemple, si le développeur souhaite rendre une certaine version d'une application prête pour la phase de test, il peut l'étiqueter « APPL_V1 ». Les testeurs peuvent ensuite obtenir cette version précise de l'application à l'aide de l'étiquette puis procéder au test. Si le test réussit, un administrateur peut permettre à l'application de s'exécuter dans l'environnement de production. En outre, les configurations de banque de données et les emplacements du fichier permettent de configurer l'application afin qu'elle s'exécute dans chaque environnement local.

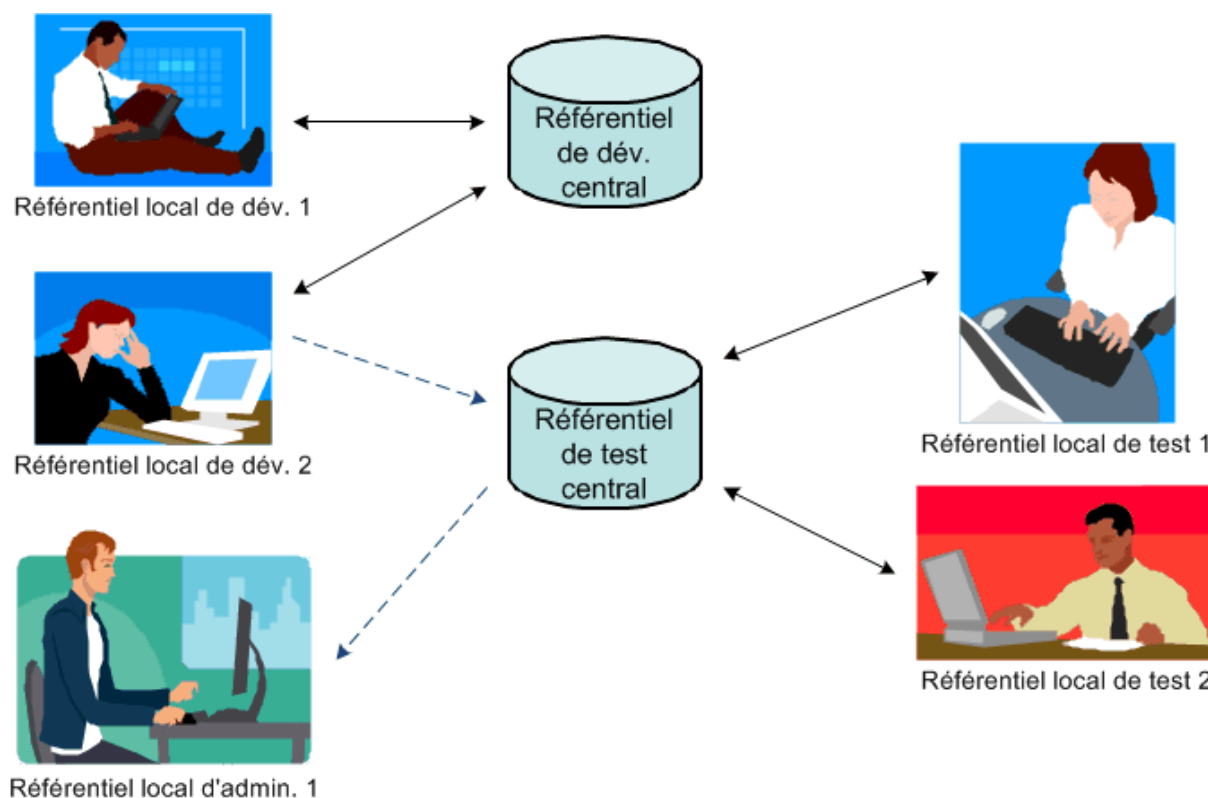
Dans certaines situations, plusieurs référentiels centraux peuvent être nécessaires pour la gestion des phases des applications. Si vous choisissez de prendre en charge plusieurs référentiels centraux, utilisez un référentiel local unique en tant qu'emplacement de transfert pour la transition.

Dans certaines situations, plusieurs référentiels centraux peuvent être nécessaires pour la gestion des phases des applications. En suivant l'exemple ci-dessus, une fois que les développeurs ont créé une version d'application prête pour la phase de test en l'étiquetant, un testeur peut l'obtenir à partir du référentiel central de développement, la tester puis la recharger dans un référentiel central de test.

Ce référentiel central de test contient toutes les versions testées au fil du temps, offrant aux testeurs la flexibilité de revenir à n'importe quelle version sans se baser sur l'environnement de développement. Lorsqu'une version d'application réussit le test, un administrateur peut l'obtenir à partir du référentiel de test et la rendre disponible en production. Encore une fois, si vous devez conserver des versions précédentes d'une application déjà en production, vous devez créer un autre référentiel central.

Avec ce modèle, un développeur n'interférera jamais avec l'environnement de test, ni un testeur avec l'environnement de production, créant ainsi un processus de migration extrêmement sûr.

Notez que si vous choisissez de prendre en charge plusieurs référentiels centraux, il convient d'utiliser un référentiel local unique en tant qu'emplacement de transfert pour la transition de fichiers.



38.7.2 Copie de contenu entre référentiels centraux

Vous ne pouvez pas copier directement le contenu d'un référentiel central vers un autre référentiel central. Vous devez plutôt utiliser le référentiel local en tant que référentiel intermédiaire.

38.7.2.1 Copier le contenu d'un référentiel central vers un autre référentiel central

1. Activez le référentiel central duquel vous souhaitez copier le contenu.
2. Obtenez la dernière version de tous les objets dans ce référentiel central actif pour qu'ils existent dans votre référentiel local.
3. Activez le référentiel central vers lequel vous souhaitez copier le contenu.
4. Lors de la première copie de contenu, ajoutez les objets du référentiel local à ce référentiel central.

Toutefois, si vous devez copier à nouveau le contenu d'un référentiel central vers un autre (par exemple lors de la phase de test, une partie d'un job a été réaffectée à la phase de développement pour une reconception), le processus est légèrement plus complexe :

- a. Extrayez tout d'abord certains objets sans remplacement du second référentiel central.

- b. A partir du référentiel local, obtenez la dernière version des objets du premier référentiel central (par exemple lors du développement).
- c. Ensuite, au lieu de les ajouter, rechargez les objets mis à jour du référentiel local vers le second référentiel central (par exemple lors du test).

Informations associées

[Activation d'un référentiel central \[page 902\]](#)

[Obtention d'objets \[page 922\]](#)

[Ajout d'objets au référentiel central \[page 911\]](#)

[Extraction d'objets \[page 913\]](#)

[Rechargement d'objets \[page 917\]](#)

38.7.3 Migration du référentiel central

Lors de la mise à niveau de la version de SAP Data Services, vous devez migrer le référentiel central vers la nouvelle version. Il est recommandé de suivre les instructions suivantes lors de la migration d'un référentiel central vers une nouvelle version du progiciel.

1. Sauvegardez toutes les tables de bases de données du référentiel central (et du référentiel local) et les données associées avant d'effectuer la mise à niveau.
2. Conservez un référentiel central distinct pour chaque version de SAP Data Services afin de conserver l'historique des objets.
Pour conserver la version actuelle et l'historique des objets dans le référentiel central, créez un nouveau référentiel central de la version actuelle du progiciel et copiez le contenu du référentiel central d'origine vers le nouveau référentiel. Ainsi, le second référentiel central agit en tant que sauvegarde de vos objets et des informations d'historique associées de la version précédente du progiciel.
Lorsque vous installez la nouvelle version du progiciel, mettez à jour le nouveau référentiel central vers la dernière version du progiciel.
3. Coordonnez vos efforts pour mettre à niveau vos référentiels centraux et locaux de manière simultanée. Différentes versions de vos référentiels central et local peuvent ne pas fonctionner simultanément. Vous ne pouvez pas effectuer d'opération multi-utilisateurs entre des référentiels local et central de différentes versions du progiciel.
4. Rechargez tous les objets (ou annulez les extractions si les objets n'ont pas été modifiés après leur extraction) avant de migrer les référentiels centraux.
Si vous ne pouvez pas mettre à niveau vos référentiels local et central simultanément, vous devez recharger tous les objets (ou annuler les extractions si les objets n'ont pas été modifiés lors de leur extraction) et en particulier les objets extraits vers un référentiel local que vous ne mettrez pas à niveau immédiatement. Après la mise à niveau du référentiel central vers la nouvelle version, vous ne pourrez pas recharger les objets du référentiel central de la version précédente du progiciel.

Informations associées

[Copie de contenu entre référentiels centraux \[page 929\]](#)

Clauses de non-responsabilité importantes et informations juridiques

Liens hypertexte

Certains liens affichent une icône et/ou du texte contextuel. Ils fournissent des informations complémentaires.

Explication des icônes :

- Liens accompagnés de l'icône  : vous accédez à un site Web non hébergé par SAP. En utilisant de tels liens, vous acceptez (sauf indication contraire expresse dans vos contrats avec SAP) ce qui suit :
 - Le contenu du site vers lequel redirige le lien n'est pas de la documentation SAP. Vous ne pouvez émettre aucune réclamation produit auprès de SAP sur la base de ces informations.
 - SAP n'accepte pas ou désapprouve le contenu affiché sur le site vers lequel vous êtes redirigé, ni ne garantit la disponibilité et l'exactitude dudit contenu. SAP ne saurait être tenue responsable des dommages causés par l'utilisation dudit contenu sauf si de tels dommages étaient causés par une négligence grave ou une faute intentionnelle de SAP.
- Liens accompagnés de l'icône  : vous quittez la documentation associée à un produit ou service SAP en particulier et accédez à un site Web hébergé par SAP. En utilisant lesdits liens, vous convenez (sauf indication contraire expresse dans vos contrats avec SAP) que vous ne pourrez pas émettre de réclamation produit auprès de SAP sur la base de ces informations.

Fonctionnalités Beta et expérimentales

Les fonctionnalités expérimentales ne font pas partie des éléments officiellement fournis par SAP et garantis pour les versions à venir. Cela signifie que les fonctionnalités expérimentales peuvent être modifiées par SAP à tout moment pour quelle que raison que ce soit, sans préavis. Les fonctionnalités expérimentales ne sont pas conçues pour être utilisées en production. Vous ne pouvez pas faire la démonstration, tester, examiner, évaluer ou utiliser d'une quelconque autre manière les fonctionnalités expérimentales dans un environnement productif ou avec des données n'ayant pas été suffisamment sauvegardées.

Le but des fonctionnalités expérimentales est d'obtenir rapidement des avis afin que les clients et partenaires puissent influencer le produit futur. En partageant votre avis (par exemple sur SAP Community), vous acceptez que les droits de propriété intellectuelle des contributions ou œuvres dérivées constituent la propriété exclusive de SAP.

Exemple de code

Les codes et/ou fragments de code ne sont que des exemples. Ils ne sont pas destinés à une utilisation en production. L'exemple de code est utilisé uniquement pour mieux expliquer et visualiser les règles de syntaxe. SAP ne garantit pas l'exactitude ni l'exhaustivité de l'exemple de code. SAP ne saurait être tenue responsable des erreurs ou dommages causés par l'utilisation dudit exemple de code sauf si de tels dommages étaient causés par une négligence grave ou une faute intentionnelle de SAP.

Formulations neutres

Nous essayons d'utiliser des formulations neutres et inclusives. Si le contexte ou la lisibilité l'exige, SAP peut être amenée à utiliser des formulations utilisant le masculin pour faire référence aux deux sexes.

© 2019 SAP SE ou société affiliée SAP. Tous droits réservés.

Toute reproduction ou communication de la présente publication, même partielle, par quelque procédé et à quelque fin que ce soit, est interdite sans l'autorisation expresse et préalable de SAP SE ou d'une société affiliée SAP. Les informations du présent document sont susceptibles d'être modifiées sans préavis.

Certains logiciels commercialisés par SAP SE et ses distributeurs contiennent des composants logiciels qui sont la propriété d'éditeurs tiers. Les spécifications des produits peuvent varier d'un pays à l'autre.

Les informations du présent document sont fournies par SAP SE ou par une société affiliée SAP uniquement à titre informatif, sans engagement ni garantie d'aucune sorte. SAP SE ou ses sociétés affiliées ne pourront en aucun cas être tenues responsables des erreurs ou omissions relatives à ces informations. Les seules garanties fournies pour les produits et les services de SAP SE ou d'une société affiliée SAP sont celles énoncées expressément à titre de garantie accompagnant, le cas échéant, lesdits produits et services. Aucune des informations contenues dans le présent document ne saurait constituer une garantie supplémentaire.

SAP et tous les autres produits et services SAP mentionnés dans ce document, ainsi que leurs logos respectifs, sont des marques commerciales ou des marques déposées de SAP SE (ou d'une société affiliée SAP) en Allemagne ainsi que dans d'autres pays. Tous les autres noms de produit et service mentionnés sont des marques commerciales de leurs sociétés respectives.

Veuillez consulter <https://www.sap.com/france/about/legal/trademark.html> pour plus d'informations sur les marques déposées.