



Data Services Getting Started Guide

Copyright

© 2008 Business Objects, an SAP company. All rights reserved. Business Objects owns the following U.S. patents, which may cover products that are offered and licensed by Business Objects: 5,295,243; 5,339,390; 5,555,403; 5,590,250; 5,619,632; 5,632,009; 5,857,205; 5,880,742; 5,883,635; 6,085,202; 6,108,698; 6,247,008; 6,289,352; 6,300,957; 6,377,259; 6,490,593; 6,578,027; 6,581,068; 6,628,312; 6,654,761; 6,768,986; 6,772,409; 6,831,668; 6,882,998; 6,892,189; 6,901,555; 7,089,238; 7,107,266; 7,139,766; 7,178,099; 7,181,435; 7,181,440; 7,194,465; 7,222,130; 7,299,419; 7,320,122 and 7,356,779. Business Objects and its logos, BusinessObjects, Business Objects Crystal Vision, Business Process On Demand, BusinessQuery, Cartesis, Crystal Analysis, Crystal Applications, Crystal Decisions, Crystal Enterprise, Crystal Insider, Crystal Reports, Crystal Vision, Desktop Intelligence, Inxight and its logos, LinguistX, Star Tree, Table Lens, ThingFinder, Timewall, Let There Be Light, Metify, NSite, Rapid Marts, RapidMarts, the Spectrum Design, Web Intelligence, Workmail and Xcelsius are trademarks or registered trademarks in the United States and/or other countries of Business Objects and/or affiliated companies. SAP is the trademark or registered trademark of SAP AG in Germany and in several other countries. All other names mentioned herein may be trademarks of their respective owners.

Third-party Contributors

Business Objects products in this release may contain redistributions of software licensed from third-party contributors. Some of these individual components may also be available under alternative licenses. A partial listing of third-party contributors that have requested or permitted acknowledgments, as well as required notices, can be found at: <http://www.businessobjects.com/thirdparty>

2008-08-26



Contents

Chapter 1	Overview of Data Services	5
	Data Services and the Business Objects Product Suite.....	6
	Data Services product benefits.....	7
	Unification with the platform.....	7
	Ease of use and high productivity.....	8
	High availability and performance.....	8
	Data Services associated products.....	8
	BusinessObjects Composer.....	9
	BusinessObjects Metadata Manager.....	9
	Data Services interfaces.....	10
Chapter 2	Data Services Architecture	11
	Standard Data Services components.....	12
	Data Services Designer.....	14
	Data Services repository.....	14
	Data Services Job Server.....	15
	Data Services engine.....	15
	Data Services Access Server.....	16
	Data Services Address Server.....	16
	Data Services Administrator.....	16
	Data Services Metadata Reports applications.....	17
	Data Services Metadata Integrator.....	19
	Data Services Service.....	20
	Data Services SNMP Agent.....	21
	Data Services Adapter SDK.....	21
	Optional Data Services components.....	21

Contents

Data Services Multi-user.....	21
Data Services management tools.....	22
License Manager.....	22
Repository Manager.....	22
Server Manager.....	22
Data Services operating system platforms.....	23
Data Services distributed architecture.....	23
Host names and port numbers.....	25

Index

27



Overview of Data Services



1

chapter



About this section

This section introduces Data Services and explains its place in the Business Objects product suite.

Related Topics

- [Data Services and the Business Objects Product Suite](#) on page 6
- [Data Services product benefits](#) on page 7
- [Data Services interfaces](#) on page 10

Data Services and the Business Objects Product Suite

The Business Objects product suite delivers extreme insight through specialized end-user tools on a single, trusted business intelligence platform. This entire platform is supported by BusinessObjects™ Data Services. On top of Data Services, Business Objects layers the most reliable, scalable, flexible, and manageable BI platform which supports the industry's best integrated end-user interfaces: reporting, query and analysis, and performance management dashboards, scorecards, and applications.

True data integration blends batch extraction, transformation, and loading (ETL) technology with real-time bi-directional data flow across multiple applications for the extended enterprise.

By building a relational data store and intelligently blending direct real-time and batch data-access methods to access data from enterprise resource planning (ERP) systems and other sources, Business Objects has created a powerful, high-performance data integration product that allows you to fully leverage your ERP and enterprise application infrastructure for multiple uses.

Business Objects provides a batch and real-time data integration system to drive today's new generation of analytic and supply-chain management applications. Using the highly scalable data integration solution provided by Business Objects, your enterprise can maintain a real-time, on-line dialogue with customers, suppliers, employees, and partners, providing them with the critical information they need for transactions and business analysis.

Data Services product benefits

Use Data Services to develop enterprise data integration for batch and real-time uses. With Data Services:

- You can create a single infrastructure for batch and real-time data movement to enable faster and lower cost implementation.
- Your enterprise can manage data as a corporate asset independent of any single system. Integrate data across many systems and reuse that data for many purposes.
- You have the option of using pre-packaged data solutions for fast deployment and quick ROI. These solutions extract historical and daily data from operational systems and cache this data in open relational databases.

Data Services customizes and manages data access and uniquely combines industry-leading, patent-pending technologies for delivering data to analytic, supply-chain management, customer relationship management, and Web applications.

Unification with the platform

Data Services provides several points of platform unification:

- Get end-to-end data lineage and impact analysis
- Create the semantic layer (universe) and manage change within the ETL design environment

Business Objects deeply integrates the entire ETL process with the business intelligence platform so you benefit from:

- Easy metadata management
- Simplified and unified administration
- Life cycle management
- Trusted information

Ease of use and high productivity

Data Services combines both batch and real-time data movement and management to provide a single data integration platform for information management from any information source, for any information use.

Using Data Services, you can:

- Stage data in an operational datastore, data warehouse, or data mart.
- Update staged data in batch or real-time modes.
- Create a single graphical development environment for developing, testing, and deploying the entire data integration platform.
- Manage a single metadata repository to capture the relationships between different extraction and access methods and provide integrated lineage and impact analysis.

High availability and performance

Data Services' high-performance engine and proven data movement and management capabilities include:

- Scalable, multi-instance data-movement for fast execution
- Load balancing
- Changed-data capture
- Parallel processing

Data Services associated products

Choose from several Business Objects product options to further support and enhance the power of your Data Services product.

BusinessObjects Composer

Even before you start a data warehouse implementation, Business Objects knows that you need to collect and organize critical data from information users to "compose" a blueprint from which you can start to build your Data Services jobs.

The BusinessObjects Composer is a stand-alone, web-based application for designing extraction, transformation, and loading (ETL) projects.

Developing higher-level designs for your ETL jobs can make the implementation process more efficient and improve the quality of resulting jobs. Composer complements other ETL tools, providing a platform for creating your design even before you start to build your data warehouse. Use Composer to identify sources of data, implement transformations, and document your designs.

For more information on BusinessObjects Composer, contact your Business Objects sales representative.

BusinessObjects Metadata Manager

Business Objects™ Metadata Manager provides an integrated view of metadata and its multiple relationships for an complete Business Intelligence project spanning some or all of the Business Objects stack of products. Use Metadata Manager to:

- View metadata about Business Objects Reports, Documents, and data sources from a single repository.
- Analyze lineage to determine data sources of Business Objects Documents and Reports.
- Analyze the impact of changing a source table, column, element, or field on existing Business Objects Documents and Reports.
- Track different versions (changes) to each object over time.
- View operational metadata (such as the number of rows processed and CPU utilization) as historical data with a datetime.
- View metadata in different languages.

For more information on Business Objects Metadata Manager, contact your Business Objects sales representative.

Data Services interfaces

Data Services provides many types of interface components. Your version of Data Services may provide some or all of them.

You can use the Data Services Interface Development Kit to develop adapters that read from and/or write to other applications.

In addition to the interfaces listed above, the Data Services Nested Relational Data Model (NRDM) allows you to apply the full power of SQL transforms to manipulate, process, and enrich hierarchical business documents.

For a detailed list of supported environments and hardware requirements, see the Platforms Availability Report available on the Business Objects support site: http://support.businessobjects.com/documentation/support_ed_platforms. This document includes specific version and patch-level requirements for databases, applications, web application servers, web browsers, and operating systems.

Related Topics

- [Designer Guide: Nested Data](#)



Data Services Architecture



2

chapter



This section describes Data Services components and their distribution on your network.

This section contains the following topics:

- Standard Data Services components
- Optional Data Services components
- Data Services management tools
- Data Services operating system platforms
- Data Services distributed architecture

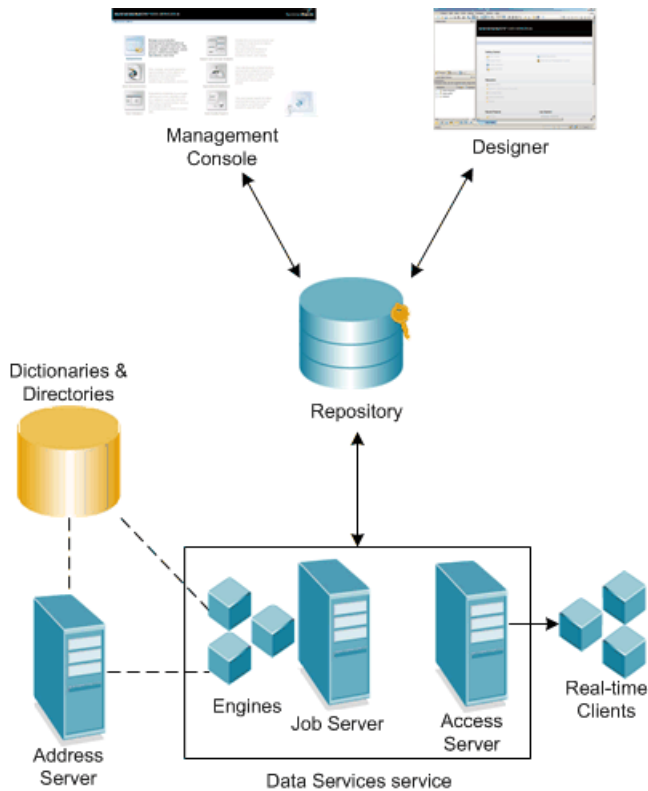
The Data Services architecture is layered to allow data integration to occur over a variety of open, industry-standard APIs for optimal data and metadata management.

Related Topics

- [Standard Data Services components](#) on page 12
- [Optional Data Services components](#) on page 21
- [Data Services management tools](#) on page 22
- [Data Services operating system platforms](#) on page 23
- [Data Services distributed architecture](#) on page 23

Standard Data Services components

The following diagram summarizes the relationships among Data Services components.



For a detailed list of supported environments and hardware requirements, see the Platforms Availability Report available on the Business Objects support site: http://support.businessobjects.com/documentation/supported_platforms. This document includes specific version and patch-level requirements for databases, applications, web application servers, web browsers, and operating systems.

Related Topics

- [Data Services Designer](#) on page 14
- [Data Services repository](#) on page 14
- [Data Services Job Server](#) on page 15
- [Data Services engine](#) on page 15
- [Data Services Access Server](#) on page 16
- [Data Services Address Server](#) on page 16
- [Data Services Administrator](#) on page 16

- [Data Services Metadata Reports applications](#) on page 17
- [Data Services Service](#) on page 20
- [Data Services SNMP Agent](#) on page 21
- [Data Services Adapter SDK](#) on page 21

Data Services Designer

The Designer is a development tool with an easy-to-use graphical user interface. It enables developers to define data management applications that consist of data mappings, transformations, and control logic.

Use the Designer to create applications containing work flows (job execution definitions) and data flows (data transformation definitions).

To use the Designer, create objects, then drag, drop, and configure them by selecting icons in flow diagrams, table layouts, and nested workspace pages. The objects in the Designer represent metadata. The Designer interface allows you to manage metadata stored in a Data Services repository. From the Designer, you can also trigger the Data Services Job Server to run your jobs for initial application testing.

Related Topics

- [Data Services repository](#) on page 14
- [Data Services Job Server](#) on page 15

Data Services repository

The Data Services repository is a set of tables that hold user-created and predefined system objects, source and target metadata, and transformation rules. Set up repositories on an open client/server platform to facilitate sharing metadata with other enterprise tools. Store each repository on an existing RDBMS.

Each repository is associated with one or more Data Services Job Servers which run the jobs you create. There are two types of repositories:

- A local repository is used by an application designer to store definitions of Data Services objects (like projects, jobs, work flows, and data flows) and source/target metadata.

- A central repository is an optional component that can be used to support multi-user development. The central repository provides a shared object library allowing developers to check objects in and out of their local repositories.

Data Services Job Server

The Data Services Job Server starts the data movement engine that integrates data from multiple heterogeneous sources, performs complex data transformations, and manages extractions and transactions from ERP systems and other sources. The Data Services Job Server can move data in either batch or real-time mode and uses distributed query optimization, multi-threading, in-memory caching, in-memory data transformations, and parallel processing to deliver high data throughput and scalability.

While designing a job, you can run it from the Designer which tells the Job Server to run the job. The Job Server gets the job from its associated repository, then starts a Data Services engine to process the job. In your production environment, the Job Server runs jobs triggered by a scheduler or by a real-time service managed by the Data Services Access Server. In production environments, you can balance job loads by creating a Job Server Group (multiple Job Servers) which executes jobs according to overall system load.

Related Topics

- [Data Services engine](#) on page 15
- [Data Services Access Server](#) on page 16

Data Services engine

When Data Services jobs are executed, the Job Server starts Data Services engine processes to perform data extraction, transformation, and movement. Data Services engine processes use parallel processing and in-memory data transformations to deliver high data throughput and scalability.

Data Services Access Server

The Access Server is a real-time, request-reply message broker that collects message requests, routes them to a real-time service, and delivers a message reply within a user-specified time frame. The Access Server queues messages and sends them to the next available real-time service across any number of computing resources. This approach provides automatic scalability because the Access Server can initiate additional real-time services on additional computing resources if traffic for a given real-time service is high. You can configure multiple Access Servers.

Data Services Address Server

The Data Services Address Server provides address validation and correction for the Global Address Cleanse EMEA engine and Global Suggestion Lists. The Address Server must be started prior to processing data flows that contain the Global Suggestion List transform or the Global Address Cleanse transform with the EMEA engine enabled.

Data Services Administrator

The Administrator provides browser-based administration of Data Services resources including:

- Scheduling, monitoring, and executing batch jobs
- Configuring, starting, and stopping real-time services
- Configuring Job Server, Access Server, and repository usage
- Configuring and managing adapters
- Managing users
- Publishing batch jobs and real-time services via Web services

Data Services Metadata Reports applications

The Metadata Reports applications provides browser-based analysis and reporting capabilities on metadata that is associated with:

- your Data Services jobs
- other Business Objects applications associated with Data Services

Metadata Reports provide four applications for exploring your metadata:

- Impact and lineage analysis
- Operational dashboards
- Auto documentation
- Data validation

Impact and Lineage Analysis reports

Impact and Lineage Analysis reports include:

- Datastore Analysis — For each datastore connection, view overview, table, function, and hierarchy reports. Data Services users can determine:
 - What data sources populate their tables
 - What target tables their tables populate
 - Whether one or more of the following Business Objects reports uses data from their tables:
 - Business Views
 - Crystal Reports
 - Universes
 - Web Intelligence documents
 - Desktop Intelligence documents
- Universe analysis — View Universe, class, and object lineage. Universe users can determine what data sources populate their Universes and what reports use their Universes.

- **Business View analysis** — View the data sources for Business Views in the Central Management Server (CMS). You can view business element and business field lineage reports for each Business View. Crystal Business View users can determine what data sources populate their Business Views and what reports use their views.
- **Report analysis** — View data sources for reports in the Central Management Server (CMS). You can view table and column lineage reports for each Crystal Report and Web Intelligence Document managed by CMS. Report writers can determine what data sources populate their reports.nic
- **Dependency analysis** — Search for specific objects in your repository and understand how those objects impact or are impacted by other Data Services or Business Objects Universe objects and reports. Metadata search results provide links back into associated reports.

To view impact and lineage analysis for Business Objects applications, you must configure the Metadata Integrator.

Related Topics

- [Installation Guide: Installing and Configuring the Metadata Integrator](#)

Operational Dashboard reports

Operational dashboard reports provide graphical depictions of Data Services job execution statistics. This feedback allows you to view at a glance the status and performance of your job executions for one or more repositories over a given time period. You can then use this information to streamline and monitor your job scheduling and management for maximizing overall efficiency and performance.

Auto Documentation reports

Auto documentation reports provide a convenient and comprehensive way to create printed documentation for all of the objects you create in Data Services. Auto documentation reports capture critical information for understanding your Data Services jobs so you can see at a glance the entire ETL process.

After creating a project, you can use Auto documentation reports to quickly create a PDF or Microsoft Word file that captures a selection of job, work flow, and/or data flow information including graphical representations and key mapping details.

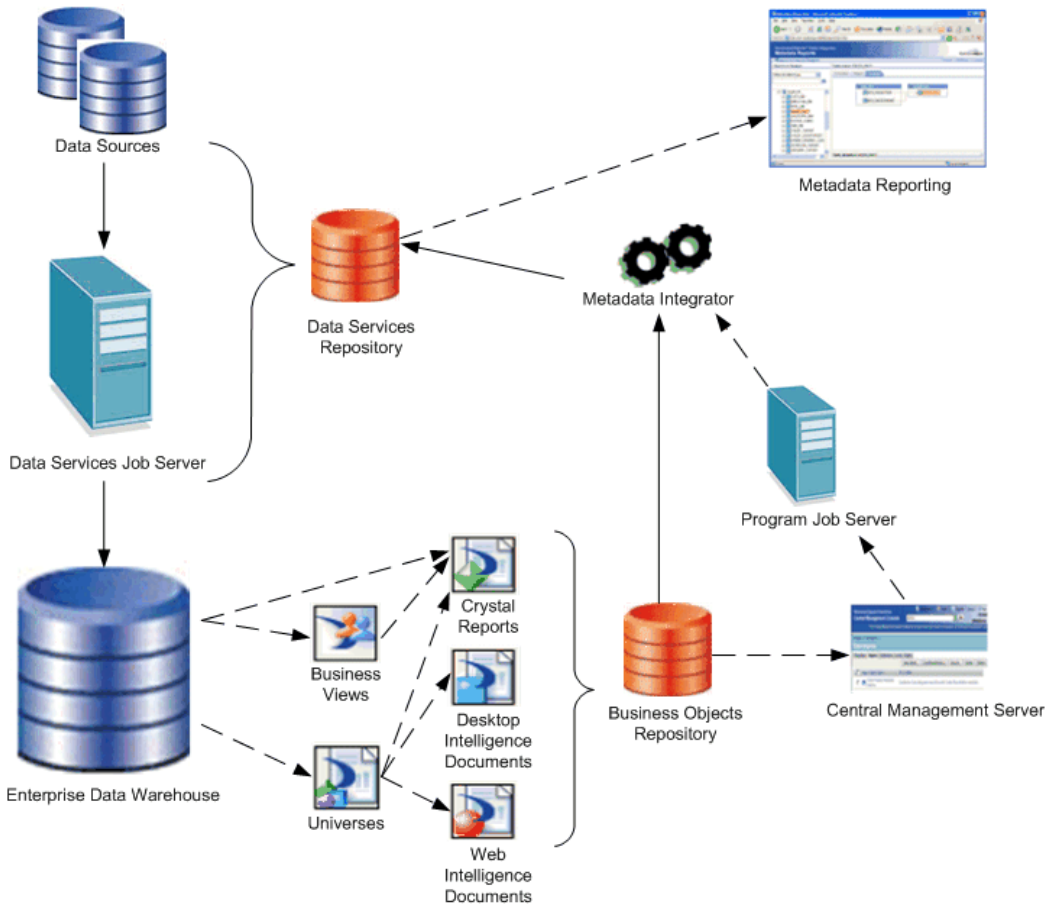
Data Validation dashboard

Data Validation dashboard reports provide graphical depictions that let you evaluate the reliability of your target data based on the validation rules you created in your Data Services batch jobs. This feedback allows business users to quickly review, assess, and identify potential inconsistencies or errors in source data.

Data Services Metadata Integrator

The Metadata Integrator allows Data Services to seamlessly share metadata with Business Objects Intelligence products. Run the Metadata Integrator to collect metadata into the Data Services repository for Business Views and Universes used by Crystal Reports, Desktop Intelligence documents, and Web Intelligence documents.

2 | Data Services Architecture Standard Data Services components



Data Services Service

The Data Services Service is installed when Data Services Job and Access Servers are installed. The Data Services Service starts Job Servers and Access Servers when you restart your system. The Windows service name is Data Services Service. The UNIX equivalent is a daemon named AL_JobService.

Data Services SNMP Agent

Data Services error events can be communicated using applications supported by simple network management protocol (SNMP) for better error monitoring. Install a Data Services SNMP agent on any computer running a Job Server. The Data Services SNMP agent monitors and records information about the Job Servers and jobs running on the computer where the agent is installed. You can configure network management software (NMS) applications to communicate with the Data Services SNMP agent. Thus, you can use your NMS application to monitor the status of Data Services jobs.

Data Services Adapter SDK

The Data Services Adapter SDK provides a Java platform for rapid development of adapters to other applications and middleware products such as EAI systems. Data Services adapters use industry-standard XML and Java technology to ease the learning curve. Adapters provide all necessary styles of interaction including:

- reading, writing, and request-reply from Data Services to other systems
- request-reply from other systems to Data Services

For detailed information, see the *Data Services Adapter SDK User's Guide*.

Optional Data Services components

Data Services Multi-user

Data Services Multi-user is an advanced optional component that enables your development team to work together on interdependent parts of an application through all phases of development. While each user works on applications in a unique local repository, the team uses a central repository to store the master copy of the entire project. The central repository preserves all versions of an application's objects, so you can revert to a previous version if needed.

Multi-user development includes other advanced features such as labeling and filtering to provide you with more flexibility and control in managing application objects.

For more details, see the *Data Services Management Console: Administrator Guide* and the *Data Services Advanced Development Guide*.

Data Services management tools

Data Services has several management tools to assist you in managing your Data Services components.

License Manager

The License Manager displays the Data Services components for which you currently have a license.

Repository Manager

The Repository Manager allows you to create, upgrade, and check the versions of local and central repositories.

Server Manager

The Server Manager allows you to add, delete, or edit the properties of Job Servers and Access Servers. It is automatically installed on each computer on which you install a Job Server or Access Server.

Use the Server Manager to define links between Job Servers and repositories. You can link multiple Job Servers on different machines to a single repository (for load balancing) or each Job Server to multiple repositories (with one default) to support individual repositories (separating test from production, for example).

You can also specify a Job Server as SNMP-enabled.

The Server Manager is also where you specify SMTP server settings for the smtp_to email function..

Related Topics

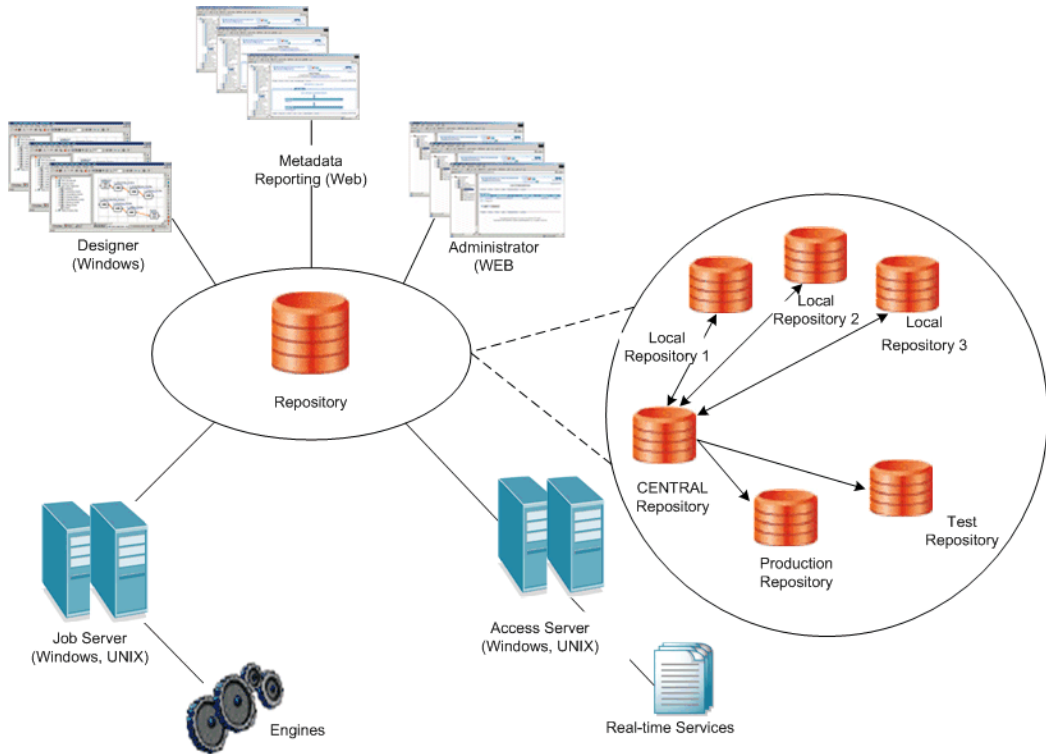
- *Designer Guide: Monitoring Jobs, SNMP support*
- *Reference Guide: To define and enable the smtp_to function*

Data Services operating system platforms

For a detailed list of supported environments and hardware requirements, see the Platforms Availability Report available on the Business Objects support site: http://support.businessobjects.com/documentation/supported_platforms. This document includes specific version and patch-level requirements for databases, applications, web application servers, web browsers, and operating systems.

Data Services distributed architecture

Data Services has a distributed architecture. An Access Server can serve multiple Job Servers and repositories. The multi-user licensed extension allows multiple Designers to work from a central repository. The following diagram illustrates both of these features.



You can distribute Data Services components across multiple computers, subject to the following rules:

- Engine processes run on the same computer as the Job Server that spawns them
- Adapters require a local Job Server

Distribute Data Services components across a number of computers to best support the traffic and connectivity requirements of your network. You can create a minimally distributed system, designed for developing and testing or a highly distributed system designed to scale with the demands of a production environment.

Host names and port numbers

Communication between a Web application, the Data Services Access Server, the Data Services Job Server, and real-time services occurs through TCP/IP connections specified by IP addresses (or host names) and port numbers.

If your network does not use static addresses, use the name of the computer as the host name. If connecting to a computer that uses a static IP address, use that number as the host name for Access Server and Job Server configurations.

To allow for a highly scalable system, each component maintains its own list of connections. You define these connections through the Server Manager, the Data Services Administrator, Repository Manager, and the Message Client library calls (from Web client).

Related Topics

- [Installation Guide: Preparing to Install Data Services, Check port assignments](#)

Index

A

- Access Server
 - description 16
- Adapter SDK 21
- Address Server, Data Services 16
- Administrator
 - description 16
- Auto Documentation reports 18

C

- central repository 14
- components
 - description 12

D

- Data Services
 - Adapter SDK 21
 - components, standard 12
 - management tools 22
 - utilities 22
- Data Services Metadata Integrator
 - description 19
- Designer
 - description 14
- distributed architecture 23
- distributing components across network 23

E

- engine, Data Services 15

H

- host names using IP address 25

I

- Impact and Lineage Analysis reports 17
- IP addresses
 - host name, using for 25
 - specifying connection 25

J

- Job Server
 - description 15

L

- License Manager 22
- license-controlled features
 - ABAP 25
 - BAPI 25
 - IDoc 25
 - JD Edwards interface 25
 - Oracle Applications interface 25
 - PeopleSoft interface 25
 - SAP BW 25
 - SNMP 25
- local repository 14

M

- management tools 22

Index

metadata, reporting tool 17

N

network, models of distribution 23

O

operating systems supported 23

Operational Dashboard reports 18

P

ports
 requirement for 25

R

repository
 central 14

repository (*continued*)

 description 14

 local 14

Repository Manager 22

S

scalability 25

Server Manager 22

T

TCP/IP

 connections required 25

 connections, defining 25

U

utilities 22