Sizing Guide CUSTOMER

Document Version: 1.5 – 2016-06-10

Sizing SAP Landscape Transformation Replication Server for SAP HANA, SPS 12



Disclaimer

Some components of this product are based on Java $^{\text{\tiny{M}}}$. Any code change in these components may cause unpredictable and severe malfunctions and is therefore expressly prohibited, as is any decompilation of these components.

Documentation in the SAP Service Marketplace

You can find this documentation at the following address: http://service.sap.com/sizing.

Typographic Conventions

Type Style	Description
Example	Words or characters quoted from the screen. These include field names, screen titles,
	pushbuttons labels, menu names, menu paths, and menu options.
	Textual cross-references to other documents.
Example	Emphasized words or expressions.
EXAMPLE	Technical names of system objects. These include report names, program names, transaction codes, table names, and key concepts of a programming language when they are surrounded by body text, for example, SELECT and INCLUDE.
Example	Output on the screen. This includes file and directory names and their paths, messages, names of variables and parameters, source text, and names of installation, upgrade and database tools.
Example	Exact user entry. These are words or characters that you enter in the system exactly as they appear in the documentation.
<example></example>	Variable user entry. Angle brackets indicate that you replace these words and characters with appropriate entries to make entries in the system.
EXAMPLE	Keys on the keyboard, for example, F2 or ENTER.

Document History

Version	Date	Change
1.0	2010-12-06	Initial Version
1.1	2014-01-29	Revised published version for SAP HANA Platform SPS07 and DMIS 2011 SP06
1.2	2014-03-10	Updates to reflect application component change from BC-HAN-LTR to HAN-DP-LTR.
1.3	2015-01-21	Revised published version for SAP HANA Platform SPS09 and DMIS 2011 SP08
1.4	2015-12-28	Revised published version for SAP HANA Platform SPS11 and DMIS 2011 SP10
1.5	2015-06-10	Revised published version for SAP HANA Platform SPS11 and DMIS 2011 SP11

Table of Contents

Introduction	6
Functions of SAP LT Replication Server for HANA	
Architecture of SAP LT Replication Server Solution	6
Factors that influence the performance	
Sizing Fundamentals and Terminology	12
Initial Sizing for SAP LT Replication Server for HANA	
Assumptions	13
Template based sizing approach	15
Sizing guideline	16
Miscellaneous	20
Comments and Feedback	21

Introduction

SAP Landscape Transformation (SAP LT) Replication Server for SAP HANA offers scheduled and real-time data replication from SAP ERP systems or non-SAP systems to an SAP HANA system.

SAP LT Replication Server for SAP HANA is part of SAP HANA software shipment and is fully integrated into the SAP HANA modeler UI.

Functions of SAP LT Replication Server for HANA

The main function of SAP LT Replication Server is to replicate data from selected tables from SAP systems or non-SAP systems to SAP HANA systems.

Some adjustments of data from source systems to the SAP HANA database format such as conversion to UNICODE format happen automatically during the replication process.

It is possible to define transformation rules for filtering of data or for more advanced transformation requirements such as the scrambling of data. Transformation rules must be defined for the relevant tables in the SAP LT Replication Server system before the data replication starts.

Architecture of SAP LT Replication Server Solution

SAP LT Replication Server uses the following jobs:

Monitoring Job

The monitoring job checks the HANA system to determine whether there are new tasks, and triggers the master control jobs. It also deletes all recorder of processed tasks from table RS_ORDER in the corresponding HANA schema, and updates table RS_STATUS with statistical information.

Master Control Job

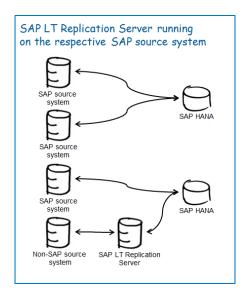
This job is started on demand by the monitor job and is responsible for creating database triggers and logging tables in the source system, creating tables and synonyms in the SAP HANA system, and for generating either the load object or the replication objects.

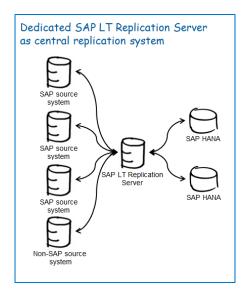
Data Load Job

This job is responsible for loading and replicating data.

Landscape Installation and Configuration

SAP LT Replication Server is flexible with regards to landscape installation and configuration. It can be installed as an add-on on any SAP source system, and also on a separate host. SAP LT Replication Server can also be used as a central component which serves multiple source and multiple target systems.





For SAP source systems, the data is extracted by means of remote function calls and pushed to HANA by means of direct SQL calls.

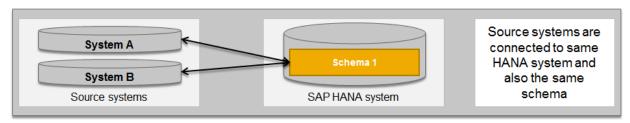


For non-SAP systems, the data is extracted with direct SQL calls through a database connection.

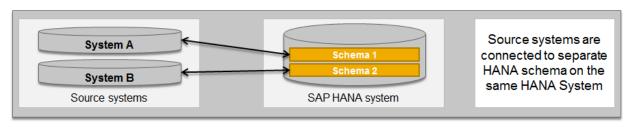


There are four different replication options; the preferred installation option depends mainly on the intended analytical use case.

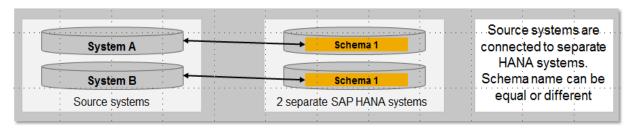
a) Multiple source systems are replicated to a single HANA schema



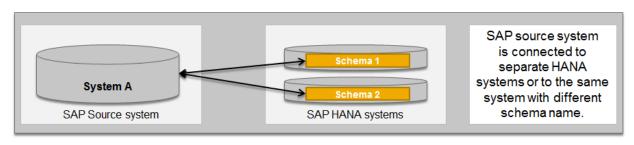
b) Multiple source systems are replicated to their own separate HANA schemas in the same SAP HANA system.



c) Multiple source systems are replicated to their own separate SAP HANA system.



d) One source system is replicated to up to four separate schemas in more than one SAP HANA system.



Factors that Influence Performance

There are various factors which influence the performance of the replication. These include the following:

a) Version of SAP LT Replication Server and HANA Software

A major prerequisite for good SAP LT Replication Server performance is to have all SAP LT Replication Server related SAP Notes and recommendations outlined in the Installation Guide and in the SAP LT Replication Server central SAP Note applied to the system. For more information, see chapter 4 "Miscellaneous".

b) Customer-Specific Configuration

The use of advanced replication settings, for example complex transformation rules for filtering data, advanced data mapping, or conversions requirements cause additional resource consumption and influence the response time. These scenarios are not covered in this sizing guide - we recommend that you engage the services of an SAP LT Replication Server consultant to support you when calculating these sizing requirements.

c) Network Factors

The performance of SAP LT Replication Server solution benefits from a network bandwidth of 1 GB or higher. In environments with frequent network failures, the replication will be repeated until the replication is successfully completed, therefore resulting in increased resource consumption and longer execution times.

In WAN environments (for example as part of a Cloud infrastructure), you can increase the speed of the replication by using the **Accelerated Application Delivery (AccAD) for SAP NetWeaver**. For more information, see:

http://www.sdn.sap.com/irj/scn/go/portal/prtroot/docs/library/uuid/b08443bc-75fd-2a10-3da0-8ef1d4f7c622?QuickLink=index&overridelayout=true&47197395724152

d) Choice of data reading type

SAP LT Landscape Transformation Server supports the following reading types:

Reading Type	Advantages	Disadvantages
Type 1 – Access Plan Calculation	 Fast data load if index exists Parallel data load possible 	 An additional index may be required. Requires a key field which is sufficiently selective. Calculation required before load
Type 3 (default) – DB_SETGET	 No separate index required Parallel data load possible with DMIS_2010 SP07 or higher 	Additional consumption of database buffer
Type 4 & 5 – Index Cluster	 Very fast data load after data is extracted to table DMC_INDXCL Minimal usage of database buffer 	Additional tablespace temporarily required in the source system

For initial sizing, this sizing guide only covers reading type 3, which is the default option for SAP LT Replication Server

If you want to use another reading type in order to optimize performance, we recommend that you contact SAP for support.

e) Appropriate Configuration for Source system, SAP LT Replication Server and HANA

There is direct relationship between the number of replication jobs in the SAP LT Replication Server system and the number of background work processes for the ERP source system.

The source system and SAP LT Replication Server should have a sufficient number of background work processes to ensure parallel execution and full allocation of the hardware configuration which is sized for the expected replication. An insufficient number of background work processes would result in a sequential execution of the replication, and therefore result in delays while replicating data changes to the SAP HANA system.

For each SAP LT Replication Server configuration, the parameter *Data Transfer Jobs* restricts the maximum number of data load jobs which can be started for one mass transfer ID (MT_ID). In total, one mass transfer ID requires that the following background work processes are available in the SAP LT Replication Server:

- 1 monitoring job (IUUC_MONITOR_<MT_ID>)
- 1 master controller job (IUUC_REPLIC_CNTR_<MT_ID>)
- 1 job for either defining the migration objects (IUUC_DEF_MIG_OBJ_<2digits>), calculating the access plan (ACC_PLAN_CALC_<MT_ID>_<2digits>), or for changing configuration settings.
- N data transfer jobs (DTL_MT_DATA_LOAD_<MT_ID>_<2digits>)

In the source system, the number of available dialog work processes which are reserved for the replication should be equal to the number of data transfer jobs running in the SAP LT Replication Server system.

i Note

The above requirements only cover the replication. We recommend that you ensure that the source system and he SAP LT Replication Server have enough additional dialog and background work processes to handle their own tasks.

Sizing Fundamentals and Terminology

SAP provides general sizing information on the SAP Service Marketplace. For the purpose of this guide, we assume that you are familiar with sizing fundamentals. You can find more information at http://service.sap.com/sizing Sizing Guidelines General Sizing Procedures.

This section explains the most important sizing terms, as these terms are used extensively in this document.

Sizing

Sizing means determining the hardware requirements of an SAP application, such as the network bandwidth, physical memory, CPU processing power, and I/O capacity. The size of the hardware and database is influenced by both business aspects and technological aspects. This means that the number of users using the various application components and the data load they put on the server must be taken into account.

Benchmarking

Sizing information can be determined using SAP Standard Application Benchmarks and scalability tests (www.sap.com/benchmark). Released for technology partners, benchmarks provide basic sizing recommendations to customers by placing a substantial load upon a system during the testing of new hardware, system software components, and relational database management systems (RDBMS). All performance data relevant to the system, user, and business applications are monitored during a benchmark run and can be used to compare platforms.

SAPS

The SAP Application Performance Standard (SAPS) is a hardware-independent unit that describes the performance of a system configuration in the SAP environment. It is derived from the Sales and Distribution (SD) Benchmark, where 100 SAPS is defined as the computing power to handle 2,000 fully business processed order line items per hour. (For more information about SAPS, see http://www.sap.com/benchmark \rightarrow Measuring in SAPS).

Initial Sizing

Initial sizing refers to the sizing approach that provides statements about platform-independent requirements of the hardware resources necessary for representative, standard delivery SAP applications. The initial sizing guidelines assume optimal system parameter settings, standard business scenarios, and so on.

Expert Sizing

This term refers to a sizing exercise where customer-specific data is being analyzed and used to put more detail on the sizing result. The main objective is to determine the resource consumption of customized content and applications (not SAP standard delivery) by comprehensive measurements. For more information, see http://service.sap.com/sizing \rightarrow Sizing Guidelines \rightarrow General Sizing Procedures \rightarrow Expert Sizing.

Configuration and System Landscaping

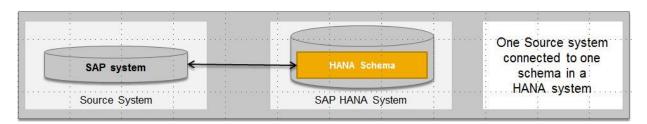
Hardware resource and optimal system configuration greatly depend on the requirements of the customer-specific project. This includes the implementation of distribution, security, and high availability solutions by different approaches using various third-party tools. In the case of high availability through redundant resources, for example, the final resource requirements must be adjusted accordingly.

There are some "best practices" which may be valid for a specific combination of operating system and database. To provide guidance, SAP created the NetWeaver configuration guides (http://service.sap.com/instguides → SAP NetWeaver).

Initial Sizing for SAP LT Replication Server for HANA

Assumptions

This sizing approach focuses on the simplest use case where data from one source system is replicated to a single HANA schema without any complex data transformation.



All more advanced variants, such as replicating data from multiple source systems to one or different HANA schemas can be calculated based on this sizing information. You can therefore you sum up the requirements of the single SLT configurations to calculate the overall expected initial load capacity or the expected maximum replication throughput rate as described in this chapter.

Categorization of Replication Tables

As input for sizing SAP LT Replication Server for an SAP HANA scenario, you need to analyze tables which will be replicated, and classify them into categories. Determine the following information for all tables (or only for the most frequently modified (inserted, updated and deleted)):

- The weighted average number of table columns (one value)
- The weighted average record length (one value)

Based on the analysis, determine the appropriate category for the volume of data, either small (S), medium (M), large (L), or extra large (XL).

	Up to 150 columns	151 to 250 columns	More than 250 columns
< 1500 bytes per record	S	М	L
> 1500 bytes per record	М	L	XL

Example 1: Categorization of individual replication-relevant tables in the relevant ERP system based their characteristics:

Table Name	# Columns	Length	Category
COBK	41	510	S
BKPF	111	1350	S
ANLA	136	2130	М
MSEG	180	2170	L
BSEG	312	3250	XL

Example 2: Weighted categorization of replication-relevant tables in the relevant ERP system based their characteristics and modification rate:

Table Name	# Columns	Length	Modification rate (per hour)	% of all modifications
СОВК	41	510	50.000	3.5%
BKPF	111	1350	1.000.000	66%
ANLA	136	2130	50.000	3.5%
MSEG	180	2170	200.000	13.5%
BSEG	312	3250	200.000	13.5%
Total			1.500.000	100%
Weighted avg. number of columns (WC)	146			
Weighted avg. bytes per record (WL)		1715		
Weighted Category = M				

Weighted average number of columns =

3.5*41 + 66*111 + 3.5*136 + 13.5*180 + 13.5*312 = 143 + 7326 + 476 + 2430 + 4212 = 14587/100 = ~146 columns

Weighted average number of bytes per record =

3.5*510 + 66*1350 + 3.5*2130 + 13.5*2170 + 13.5*3250 = 1785 + 89100 + 7455 + 29295 + 43875 = 171510/100 = ~1715 bytes

Template-Based Sizing Approach

In order to provide a sizing estimate, SAP offers a simplified approach with three different SAP LT Replication Server scenarios:

Server scenarios			1.1505
	SMALL	MEDIUM	LARGE
Use Case	A small scenario with: Typically one LTR configuration Approximately up to 50 tables A weighted table size category of S-M An overall expected throughput of less than 1,000,000 records/hour	 A midrange scenario with: Approximately 3 different source systems (equivalent to 3 SAP LT Replication Server configurations), And/or up to 200 tables in total A weighted table size category of M-L An overall expected throughput of less than 10,000,000 records/hour 	An upper mid-range scenario with: Up to 10 different source systems (equivalent to 10 SAP LT Replication Server configurations), And/or up to 500 tables in total A weighted table size category of M-XL an overall expected throughput of up to 50,000,000 records/hour
SLT Server	1 SAP LT Replication Server configuration with 2 data transfer jobs Hardware: 2-4 CPU cores, 16-32 GB of main memory	 10 data transfer jobs in total (sum of all configurations) Hardware: 4-6 CPU cores, 32-64 GB of main memory 	 25 data transfer jobs in total (sum of all configurations) Hardware: 8-10 CPU cores, 48-128 GB of main memory
Source System(s)	The number of dialog work processes in the source system corresponds 1:1 with the number of data transfer jobs in the SAP LT Replication Server system Reserve 2 background work processes for access plan calculation, ensure 2 free dialog work processes for data load/replication Additional hardware requirements: ~ 1 CPU core	The number of dialog work processes in the source system corresponds 1:1 with the number of data transfer jobs in the SAP LT Replication Server system In total, for all source systems: Reserve 2-4 BTC work processes for ACL (Access plan calculation), Ensure 10 free dialog work processes for data load/replication Additional hardware requirements: ~ 5 CPU cores in total	The number of dialog work processes in the source system corresponds 1:1 with the number of data transfer jobs in the SAP LT Replication Server system In total, for all source systems: Reserve 4-8 BTC work processes for ACL (Access plan calculation), Ensure 25 free dialog work processes for data load/replication Additional Hardware requirements: ~ 12 CPU cores in total

	SMALL	MEDIUM	LARGE
SAP HANA System	~ 1 additional CPU core	~ 3-4 additional CPU cores	~ 8 additional CPU cores

Sizing Guideline

Sizing input for replication load

For "initial" and "scheduled" data replication, you should calculate:

- The total number of replication-relevant records (NR) in all replication-relevant tables
- The required speed of load, that is the target duration (D) of the initial or scheduled load in hours

The number of replication-relevant records (NR) for the *initial load* is equal to the total number of records of all replication-relevant tables.

The number of replication-relevant records (NR) for the *scheduled load* is equal to the number of modified (inserted, updated, deleted) records of all replication-relevant tables for the period since last replication completed.

You can determine the overall target throughput (T) by using the following formula:

T[records/second] = NR / (3600*D)

For "real-time" data replication, you can specify the total number of modified (inserted, updated, deleted) records per hour T[records/second] for all replication-relevant tables.

The sizing result refers to the **net resource requirement for each of the SAP LT Replication Server replication-relevant components**. This net resource capacity **should be available in addition to any other resource capacity** required for running other functions and tasks in that system in parallel.

The resource consumption depends on the number of parallel jobs required to achieve the expected overall throughput. We assume a standardized average throughput rate of 1000 records per second, per data transfer job, so that for the default reading type 3 (portions, ordered by primary index) the number of required data transfer jobs (NJ) can be determined by using the following formula::

NJ = T[records/second] /1000

If you replicate cluster tables, you can estimate the resource consumption more precisely if you calculate the throughput per second for the cluster tables separately, and then calculate the number of required parallel jobs for clustered tables by using the following formula:

NJ = T[records/second] /600

You can then calculate the sizing result for cluster tables and non-cluster tables separately, and then add the results together to get a final sizing result for the entire load.

CPU Sizing (SAPS)

The table below provides information about how to calculate the additional CPU resource consumption caused by the active data replication scenario on top of the underlying system.

The minimum requirements for SAP LT Replication Server are 2-4 CPU cores, and approximately 16GB of main memory for the SAP NetWeaver AS ABAP instance.

For a standalone SAP LT Replication Server based on a SAP NetWeaver AS ABAP instance, you have to implement the minimum requirements and, in addition, the calculated SAPS.

Table Category	SAPS SAP Source System	SAPS SAP Source System DB	SAPS HANA DB	SAPS SLT Server& DB
S	60 * NJ	200 * NJ	200* NJ	250 * NJ
М	90 * NJ	350 * NJ	300* NJ	370 * NJ
L	150 * NJ	480 * NJ	480* NJ	600 * NJ
XL	180 * NJ	600 * NJ	600 * NJ	750 * NJ

i Note

If the sizing result is less than 1000 SAPS, then one available core (1000 SAPS is equivalent to one core) would be sufficient to fulfill the replication requirements. For a higher number of SAPS, and by extension the requirement to add additional CPU cores, we recommend that you contact your hardware provider for hardware upgrade advice.

The following table details two calculation examples:

Example 1	Example 2	
Weighted table category: L	Weighted Table Category: M	
Overall change rate = 50,000,000 [records/hour]	Overall Change Rate = 10, 000,000 [records/hour]	
T= 13889 records/second	T= 2778 records/second	
Nj = T[records/second] /1000 = 13,9	Nj = T[records/second] /1000 = 2,78	
→ ~ 14 Data Transfer Jobs required	→ ~ 3 Data Transfer Jobs required	
→ SAPS HANA DB = 6672	→ SAPS HANA DB = 834	
→ SAPS SLT Server = 8340	→ SAPS SLT Server = 1029	
→ SAPS Source System = 8757	→ SAPS Source System = 1223	

Memory Sizing

The additional memory requirements for replication on top of the existing memory utilization (that is, in the source system or the HANA system) can be calculated by using the following table:

Table Category	SAP Source System [MB]	SLT [MB]	HANA DB [MB]
S	40 * NJ	40 * NJ	40 * NJ
М	60 * NJ	60 * NJ	60 * NJ
L	80 * NJ	80 * NJ	80 * NJ
XL	100 * NJ	100 * NJ	100 * NJ



Besides the memory sizing for the replication detailed in the table above, the following general recommendations also apply for the underlying SAP HANA instance:

The memory requirements for HANA are the runtime requirements for the processing of data.

HANA is an in-memory database, which supports compression. For this reason the table size in HANA database may differ from the size of same on the source system. The overall HANA memory sizing for table content http://service.sap.com/quicksizer -> SAP In-Memory Computing -> Standalone HANA should be applied.

Note that if the table size (in the SAP HANA database) exceeds 2 billion records, you must split the table by using the available partitioning features. For more information, see the administration guide SAP HANA Database - Partitioning and Distribution of Large Tables on the SAP Help Portal at http://help.sap.com/hana -> SAP HANA Appliance Software.



The memory sizing for the underlying databases of the SAP source system and SAP LT Replication Server should be performed according to the sizing procedures of the database vendors.

Disk Sizing

There are no disk space requirements for the SAP source system and for the SAP LT Replication Server system because there is no growth of business data in these systems. However, temporary disk space allocation in the source system (required for the logging tables) should be handled appropriately.

In normal replication conditions, SAP LT Replication Server logging tables in the source system are usually empty (or close to empty). However, these logging tables may increase in size during the initial load if business processes change data in the source system. They can also increase in size if the replication is suspended for a long period of time, if there is a system outage, or if the SAP LT Replication Server system or the SAP HANA system is unavailable.

We therefore recommend that you create a separate tablespace for the SAP LT Replication Server logging tables in the source system. You can do this when you create the data transfer configuration in transaction LTR. We recommend including a safety buffer for unexpected incidents, and therefore creating the tablespace with an initial size of 20 to 40 GB that has an automatic extend feature.

Data replication from SAP source systems is managed by RFC connections. We recommend that you monitor the size of related RFC log files.

The HANA disk sizing should be performed using the Quick Sizer tool. For more information, see http://service.sap.com/quicksizer -> SAP In-Memory Computing -> Standalone HANA.

Network Sizing

Replicating data is a business process with significant network requirements that depends on the actual throughput and the weighted average length (WL) of record in bytes.

You can calculate the network bandwidth by using the following formula:

Network Bandwidth [Mbit/s] = 8 * WL * T[records/second] / 1000000

Miscellaneous

The following information is also available for SAP LT Replication Server:

Installation Guide - Trigger-based Data Replication Using SAP LT Replication Server (https://service.sap.com/~sapidb/011000358700000604912011)

Technical Operations Guide – SAP HANA Appliance Software (http://help.sap.com/hana/hana1_tom_en.pdf)

SAP Note 1605140 (Central SAP Note for SAP LT Replication Server)

(https://css.wdf.sap.corp/sap/bc/bsp/spn/sapnotes/index2.htm?numm=0001605140)

Comments and Feedback

Both are very welcome! Please use component HAN-DP-LTR to address questions, comments and feedback to SAP.



www.sap.com/contactsap

Material Number

© 2015 SAP AG. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP AG. The information contained herein may be changed without prior notice.

Some software products marketed by SAP AG and its distributors contain proprietary software components of other software vendors.

Microsoft, Windows, Excel, Outlook, and PowerPoint are registered trademarks of Microsoft Corporation.

IBM, DB2, DB2 Universal Database, System ads, System i5, System p, System p5, System x, System z, System z10, System z9, z10, z9, iSeries, pSeries, xSeries, zSeries, eServer, z/VM, z/OS, i5/OS, S/390, OS/390, OS/400, AS/400, S/390 Parallel Enterprise Server, PowerVM, Power Architecture, POWER6+, POWER6, POWER5+, POWER5, POWER, OpenPower, PowerPC, BatchPipes, BladeCenter, System Storage, GPFS, HACMP, RETAIN, DB2 Connect, RACF, Redbooks, OS/2, Parallel Sysplex, MVS/ESA, AIX, Intelligent Miner, WebSphere, Netfinity, Tivoli and Informix are trademarks or registered trademarks of IBM Corporation. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries.

Adobe, the Adobe logo, Acrobat, PostScript, and Reader are either trademarks or registered trademarks of Adobe Systems Incorporated in the United States and/or other countries. Oracle is a registered trademark of Oracle Corporation. UNIX, X/Open, OSF/1, and Motif are registered trademarks of the Open Group.

Citrix, ICA, Program Neighborhood, MetaFrame, WinFrame, VideoFrame, and MultiWin are trademarks or registered trademarks of Citrix Systems, Inc.

HTML, XML, XHTML and W3C are trademarks or registered trademarks of W3C*, World Wide Web Consortium, Massachusetts Institute of Technology.

Java is a registered trademark of Sun Microsystems, Inc. JavaScript is a registered trademark of Sun Microsystems, Inc., used under license for technology invented and implemented by Netscape.

SAP, R/3, xApps, xApp, SAP NetWeaver, Duet, PartnerEdge, ByDesign, SAP Business ByDesign, and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP AG in Germany and in several other countries all over the world. All other product and service names mentioned are the trademarks of their respective companies. Data contained in this document serves informational purposes only. National product specifications may vary. These materials are subject to change without notice. These materials are provided by SAP AG and its affiliated companies ("SAP Group") for informational purposes only, without representation or warranty of any kind, and SAP Group shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP Group products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty