IBM

SAP In-Memory Computing on IBM eX5 Systems

IBM Systems Solution for SAP HANA

SAP HANA overview and use cases

Basic In-Memory Computing principles

Gereon Vey
Ilya Krutov

Redpaper

ibm.com/redbooks

International Technical Support Organization

SAP In-Memory Computing on IBM eX5 Systems

January 2012

REDP-4814-00

Note: Before using this information and the product it supports, read the information in “Notices” on page v.

First Edition (January 2012)

This edition applies to IBM System Solution for SAP HANA, an appliance based on IBM System eX5 servers and the SAP HANA offering.

Contents

Notices v
Trademarks vi

Preface vii
The team who wrote this paper vii
Now you can become a published author, too! ix
Comments welcome. ix
Stay connected to IBM Redbooks ix

Chapter 1. History of in-memory computing at SAP 1
1.1 SAP Search and Classification (TREX) 2
1.2 SAP liveCache 2
1.3 SAP NetWeaver Business Warehouse Accelerator 3
1.3.1 SAP BusinessObjects Explorer Accelerated. 4
1.3.2 SAP BusinessObjects Accelerator 4
1.4 SAP HANA 5

Chapter 2. Basic concepts 7
2.1 Keeping data in-memory 8
2.1.1 Using main memory as the data store 8
2.1.2 Data persistence 9
2.2 Minimizing data movement 10
2.2.1 Compression 10
2.2.2 Columnar storage 11
2.2.3 Pushing application logic to the database. 13
2.3 Dividing and conquering 13
2.3.1 Parallelization on multi-core systems 14
2.3.2 Data partitioning and scale-out 14

Chapter 3. SAP HANA 15
3.1 SAP HANA overview 16
3.1.1 SAP HANA architecture 16
3.1.2 SAP HANA appliance 17
3.2 SAP HANA delivery model 17
3.3 SAP HANA use cases 18
3.3.1 SAP HANA in a data mart side-car approach 18
3.3.2 SAP HANA as a database for SAP NetWeaver BW 21
3.3.3 SAP HANA with third-party software. 22
3.4 Data replication methods for SAP HANA 23
3.4.1 Log-based replication with Sybase Replication Server 23
3.4.2 Trigger-based replication with SAP LT Replication. 24
3.4.3 ETL-based replication with SAP BusinessObjects Data Services 24
3.4.4 Comparison of the replication methods. 25
3.5 Sizing SAP HANA 26
3.5.1 The concept of T-shirt sizes for SAP HANA 26
3.5.2 Sizing approach 26
3.6 SAP HANA software licensing 29

Chapter 4. The IBM Systems Solution for SAP HANA 31

4.1 IBM eX5 Systems 32
4.1.1 IBM System x3850 X5 and x3950 X5 32
4.1.2 IBM System x3690 X5 33
4.1.3 Intel Xeon processor E7 family 34
4.1.4 Memory 38
4.1.5 Flash technology storage 40
4.2 General Parallel File System 42
4.3 Custom server models for SAP HANA 44
4.3.1 IBM System x workload-optimized models for SAP HANA 44
4.3.2 SAP HANA T-shirt sizes 46
4.3.3 Scale-up 46
4.4 Scale-out solution for SAP HANA 47
4.4.1 Scale-out solution without high-availability capabilities 48
4.4.2 Scale-out solution with high-availability capabilities 50
4.5 Monitoring SAP HANA 53
4.5.1 Monitoring with SAP HANA Studio 53
4.5.2 Monitoring SAP HANA with Tivoli 53
4.6 Support processwint 54
4.6.1 IBM SAP integrated support 55
4.6.2 The IBM SAP International Competence Center (ISICC) InfoService 56

Chapter 5. Summary 57
5.1 Benefits of in-memory computing 58
5.2 SAP HANA: An innovative analytic appliance 58
5.3 IBM Systems Solution for SAP HANA 59
5.3.1 Workload Optimized Solution 59
5.3.2 Leading performance 59
5.3.3 IBM GPFS enhancing performance, scalability, and reliability 60
5.3.4 Scalability 60
5.3.5 Services to speed deployment 61
5.4 Going beyond infrastructure 61
5.4.1 A trusted service partner 61
5.4.2 IBM and SAP team for long-term business innovation 61

Abbreviations and acronyms 63

Related publications 65
IBM Redbooks 65
Other publications 65
Online resources 65
Help from IBM 65

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX®
BladeCenter®
DB2®
Global Business Services®
Global Technology Services®
GPFS™
IBM®
Intelligent Cluster™
Redbooks®
Redpaper™
Redbooks (logo) ®
System x®
Tivoli®

The following terms are trademarks of other companies:

Intel Xeon, Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Preface

This IBM® Redpaper™ publication describes the new in-memory computing appliances from IBM and SAP that are based on IBM eX5 flagship systems and SAP HANA. We first discuss the history and basic principles of in-memory computing, then we describe the SAP HANA offering, its architecture, sizing methodology, and licensing policy. We also review IBM eX5 hardware offerings from IBM. Then we describe the architecture and components of IBM System Solution for SAP HANA and its delivery, operational, and support aspects. Finally, we discuss the advantages of using IBM infrastructure platforms for running the SAP HANA solution.

These topics are covered:

- The history of in-memory computing
- The basic principles of in-memory computing
- The SAP HANA offering
- The IBM System Solution for SAP HANA
- Benefits of using the IBM infrastructure for SAP HANA

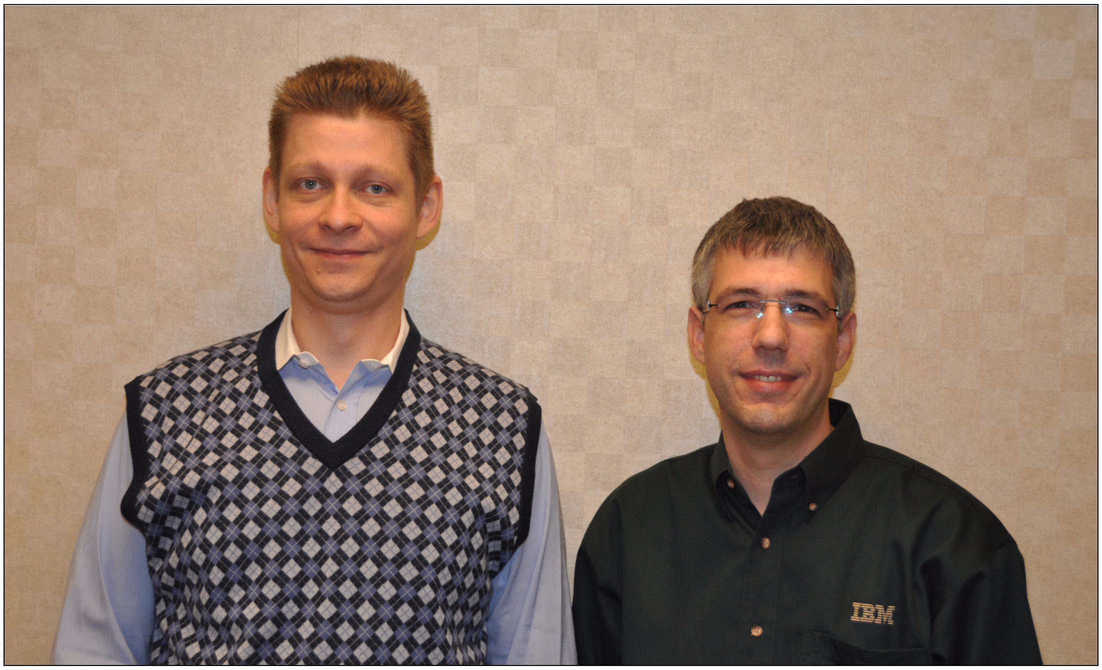
This paper is intended for SAP administrators and technical solution architects. It is also for IBM Business Partners and IBM employees who want to know more about the SAP HANA offering and other available IBM solutions for SAP customers.

The team who wrote this paper

This paper was produced by a team of specialists from around the world working at the International Technical Support Organization, Raleigh Center.

Gereon Vey has been a member of the IBM System x® Team at the IBM SAP International Competence Center (ISICC) in Walldorf, Germany since 2004. He is the Global Subject Matter Expert for the SAP Appliances, such as SAP NetWeaver BW Accelerator and SAP HANA, at the ISICC, and is part of the team developing the IBM Systems Solution for SAP HANA. His other activities include the maintenance of sizing guidelines and capacity data for System x servers and pre-sales support for IBM worldwide. He has worked in the IT industry since 1992. He graduated with a degree in computer science from the University of Applied Sciences in Worms, Germany in 1999.

Ilya Krutov is an Advisory IT Specialist and Project Leader at the International Technical Support Organization, Raleigh Center, and has been with IBM since 1998. Prior roles at IBM include STG Run Rate Team Leader, Brand Manager for IBM System x and BladeCenter®, Field Technical Sales Support (FTSS) Specialist for System x and BladeCenter products, and Instructor at IBM Learning Services Russia (IBM x86 servers, Microsoft NOS, Cisco). He graduated from the Moscow Engineering and Physics Institute, and has a bachelor's degree in computer engineering.



The team (left to right): Ilya and Gereon

Thanks to the following people for their contributions to this project:

From the International Technical Support Organization, Raleigh Center:

- David Watts
- Linda Robinson
- Tamikia Lee
- Erica Wazewski

From IBM:

- Carol Davis
- Volker Fischer
- Irene Hopf
- Oliver Rettig
- Tag Robertson

From SAP:

- Uwe Heinz

From Intel:

- Thomas Willhalm

Portions of this paper were taken from *IBM eX5 Portfolio Overview: IBM System x3850 X5, x3950 X5, x3690 X5, and BladeCenter HX5*, REDP-4650. Thanks to the authors:

- David Watts
- Duncan Furniss
- Scott Haddow
- Jeneea Jervay
- Eric Kern
- Cynthia Knight

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our papers to be as helpful as possible. Send us your comments about this paper or other IBM Redbooks® publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

History of in-memory computing at SAP

In-memory computing has a long history at SAP. This chapter provides a short overview of the history of SAP in-memory computing. It describes the evolution of SAP in-memory computing and gives a short overview of SAP products involved in this process:

- 1.1, “SAP Search and Classification (TREX)” on page 2
- 1.2, “SAP liveCache” on page 2
- 1.3, “SAP NetWeaver Business Warehouse Accelerator” on page 3
- 1.4, “SAP HANA” on page 5

1.1 SAP Search and Classification (TREX)

SAP first made SAP In-Memory Computing available in product form with the introduction of SAP Search and Classification, better known as Text Retrieval and Information Extraction (TREX). TREX is a search engine for both structured and unstructured data. It provides SAP applications with numerous services for searching and classifying large collections of documents (unstructured data) and for searching and aggregating business data (structured data).

TREX offers a flexible architecture that enables a distributed installation, which can be modified to match various requirements. A minimal system consists of a single host that provides all TREX functions. Starting with a single-host system, you can extend TREX to be a distributed system and thus increase its capacity. TREX stores its data, usually referred to as indexes, not in the way traditional databases would do, but merely as flat files in a file system. For a distributed system, the file system must be a clustered or shared file system, which presents all files to all nodes of the distributed system.

For performance reasons, TREX indexes are loaded to working memories. Indexes for structured data are implemented compactly using data compression, and the data can be aggregated in linear time to enable large volumes of data to be processed entirely in memory.

Earlier TREX releases (TREX 7.0 and earlier) are supported on a variety of platforms (such as IBM AIX®, HP-UX, SOLARIS, Linux, and Windows). To optimize the performance of the search and indexing functions provided by the TREX engine, SAP decided to concentrate on the Intel platform to be able to optimally utilize the CPU architecture. Therefore, the newest version of TREX (Version 7.10) is only available on Windows and Linux 64-bit operating systems.

TREX as a search engine component is used as an integral part of various SAP software offerings, such as SAP NetWeaver Enterprise Search. TREX as an SAP NetWeaver standalone engine is a significant part of most search features in SAP applications.

1.2 SAP liveCache

SAP liveCache technology can be characterized by a hybrid main-memory database with intensive usage of database procedure. It is based on MaxDB, a relational database owned by SAP, introducing a combination of in-memory data storage with special object-oriented database technologies supporting the application logic. This hybrid database system can process enormous volumes of information, such as planning data. It significantly increases the speed of the algorithmically complex, data-intensive and runtime-intensive functions of various SAP applications, especially within SAP Supply Chain Management (SCM) and SAP Advanced Planner and Optimizer (APO). The APO/liveCache architecture consists of these major components:

- ABAP code in APO, which deals with APO functionality
- Application functions providing extended database functionality to manipulate business objects
- LiveCache's special MaxDB implementation, providing a memory resident database for fast data processing

From the view of the APO application servers, the liveCache database appears as a second database connection. LiveCache provides a native SQL interface, which also allows the application servers to trigger object-oriented functions at the database level. These functions are provided by means of C++ code running on the liveCache server with extremely fast access to the objects in liveCache memory. This is the functionality, which allows processing load to be passed from the application server to the LiveCache server, rather than just accessing database data. This functionality, referred to as the *COM-Modules* or *liveCache Applications*, supports the manipulation of memory resident objects and datacubes, and significantly increases the speed of the algorithmically complex, data-intensive, and runtime-intensive functions.

SAP APO transfers performance-critical application logic to the SAP liveCache. Data needed for these operations is sent to SAP liveCache and kept in-memory. This ensures that the processing happens where the data is, to deliver the highest possible performance. The object-oriented nature of the application functions enable parallel processing so that modern multi-core architectures can be leveraged.

1.3 SAP NetWeaver Business Warehouse Accelerator

The two primary drivers of the demand for business analytics solutions are increasing data volumes and user populations. These drivers place new performance requirements on existing analytic platforms. To address these requirements, SAP introduced SAP NetWeaver Business Warehouse Accelerator[1] (BW Accelerator) in 2006, deployed as an integrated solution combining software and hardware to increase the performance characteristics of SAP NetWeaver Business Warehouse deployments.

The SAP NetWeaver BW Accelerator is based on TREX technology. SAP used this existing technology and extended it with more functionality to efficiently support the querying of massive amounts of data, and to perform simple operations on the data frequently used in a business analytics environment.

The software's engine decomposes table data vertically into columns that are stored separately. This makes more efficient use of memory space than row-based storage, because the engine needs to load only the data for relevant attributes or characteristics into memory. In general, this is a good idea for analytics, where most users want to see only a selection of data. We discuss the technology and advantages of column-based storage in Chapter 2, "Basic concepts" on page 7, along with other basic in-memory computing principles employed by BW Accelerator.

SAP NetWeaver BW Accelerator is built for a special use case, speeding up queries and reports in SAP NetWeaver Business Warehouse (SAP BW). In a nutshell, after connecting the BW Accelerator to the BW system, InfoCubes can be marked to be indexed in BW Accelerator, and subsequently all database-bound queries (or even parts of queries) that operate on the indexed InfoCubes actually get executed in-memory by the BW Accelerator.

Because of this tight integration with SAP BW and the appliance-like delivery model, BW Accelerator requires minimal configuration and setup. Intel helped develop this solution with SAP, so it is optimized for, and only available on, Intel Xeon processor-based blade technology. SAP partners with several hardware vendors to supply the infrastructure for the BW Accelerator software. Customers acquire the SAP software license from SAP, and the hardware partner delivers a pre-configured and pre-installed solution.

The IBM Systems solution for SAP NetWeaver BW Accelerator helps provide customers with near real-time business intelligence for those companies that need timely answers to vital

[1] Formerly named SAP NetWeaver Business Intelligence Accelerator, SAP changed the software solution name in 2009 to SAP NetWeaver Business Warehouse Accelerator. The solution functions remain the same.

business questions. It allows customers to perform queries in seconds rather than tens of minutes and gives them better visibility into their business.

IBM has significant competitive advantages with our IBM BladeCenter-based implementation:

- Better density
- More reliable cooling
- Fibre storage switching
- Fully redundant enterprise class chassis
- Systems management

BW Accelerator plugs into existing SAP NetWeaver Business Warehouse environments regardless of the server platform used in that environment.

The IBM solution consists of these components:

- IBM BladeCenter chassis with HS22 blade servers with Intel Xeon processors, available in standard configurations scaling from 2 to 28 blades and custom configurations up to 140 blades
- IBM DS3524 with scalable disk
- SUSE Linux Enterprise Server as the operating system
- IBM General Parallel File System (GPFS™)
- IBM Services including Lab Services, GBS, and GTS offerings and Intelligent Cluster™ enablement team services

This intelligent scalable design is based around the IBM General Parallel File System, exclusive from IBM. GPFS is a highly scalable, high-performance shared disk file system, powering many of the world‘s largest supercomputing clusters. Its advanced high-availability and data replication features are a key differentiator for the IBM offering. GPFS not only provides scalability, but also offers exclusive levels of availability that are easy to implement with no manual intervention or scripting.

IBM has shown linear scalability for the BW Accelerator through 140 blades[2]. Unlike all other BWA providers, the IBM BWA solution provides a seamless growth path for customers from two blades to 140 blades with no significant changes in the hardware or software infrastructure.

1.3.1 SAP BusinessObjects Explorer Accelerated

To extend the functionality of BW Accelerator, SAP created a special version of the SAP BusinessObjects Explorer (BO Explorer or BOE), which can connect directly to the BW Accelerator using its proprietary communication protocol. This version, called SAP BusinessObjects Explorer Accelerated, provides an alternative front end to navigate through the data contained in BW Accelerator, with a much simpler, web-based user interface than the SAP BW front ends can provide, broadening the user base towards the less experienced BI users.

1.3.2 SAP BusinessObjects Accelerator

SAP enabled the combination of BW Accelerator and SAP BusinessObjects Data Services to load data into BW Accelerator from virtually any data source, both SAP and non-SAP data

[2] WinterCorp white paper: "Large-Scale Testing of the SAP NetWeaver BW Accelerator on an IBM Platform," available at ftp://ftp.software.ibm.com/common/ssi/sa/wh/n/spw03004usen/SPW03004USEN.PDF.

sources. In combination with BO Explorer as an independent front end, the addition of SAP BusinessObjects Data Services created a solution that is independent of SAP BW.

This combination of BW Accelerator, SAP BusinessObjects Explorer Accelerated, and SAP BusinessObjects Data Services is often referred to as the SAP BusinessObjects Accelerator, or SAP BusinessObjects Explorer Accelerated Wave 2. Additional blades are added to the BW Accelerator configuration to support the BusinessObjects Explorer Accelerated workload, enabling it to be delivered as part of the BW Accelerator solution.

1.4 SAP HANA

SAP HANA is the next logical step in SAP in-memory computing. By combining earlier developed or acquired technologies, such as the BW Accelerator (including TREX technology), MaxDB with its in-memory capabilities originating in SAP liveCache, or P*Time (acquired by SAP in 2005), with recent research results from the Hasso Plattner Institute for Software Systems Engineering[3] (HPI), SAP created an in-memory database appliance for a wide range of applications. Figure 1-1 shows the evolution of SAP in-memory computing.

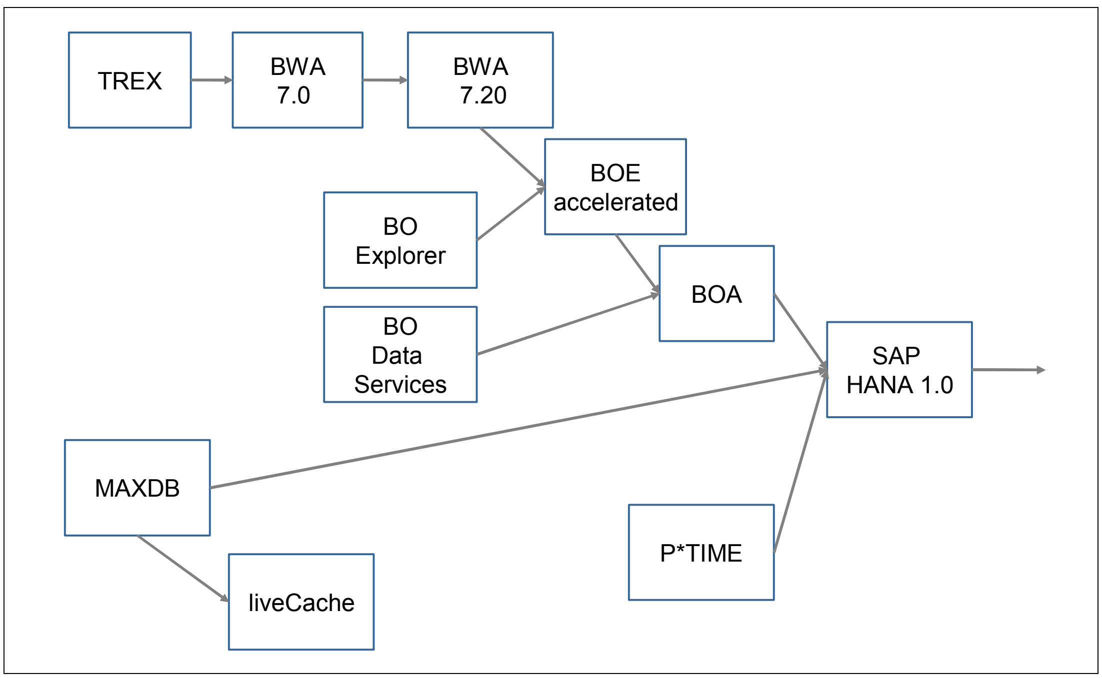


Figure 1-1 Evolution of SAP in-memory computing

Initially targeted at analytical workloads, Hasso Plattner presented (during the announcement of SAP HANA at SapphireNOW 2010) his vision of SAP HANA becoming a database suitable as a base for SAP's entire enterprise software portfolio.

Just as with BW Accelerator, SAP decided to deploy SAP HANA in an appliance-like delivery model. IBM was one of the first hardware partners to work with SAP on an infrastructure solution for SAP HANA.

This IBM Redpaper publication focuses on SAP HANA and the IBM solution for SAP HANA.

[3] Founded in 1998 by Hasso Plattner, one of the founders of SAP AG, chairman of the board until 2003, and currently chairman of the supervisory board of SAP AG

2

Basic concepts

In-memory computing is a technology that allows the processing of massive quantities of data in main memory to provide immediate results from analysis and transaction. The data to be processed is ideally real-time data (that is, data that is available for processing or analysis immediately after it is created).

To achieve the desired performance, in-memory computing follows these basic concepts:

- Keep data in main memory to speed up data access.
- Minimize data movement by leveraging the columnar storage concept, compression, and performing calculations at the database level.
- Divide and conquer. Leverage the multi-core architecture of modern processors and multi-processor servers, or even scale out into a distributed landscape, to be able to grow beyond what can be supplied by a single server.

In this chapter we describe those basic concepts with the help of a few examples. We do not describe the full set of technologies employed with in-memory databases like SAP HANA, but we provide an overview of how in-memory computing is different from traditional concepts.

2.1 Keeping data in-memory

The capacity of main memory in servers has continuously increased over the years, whereas prices have dramatically dropped. Today, a single enterprise class server can hold several terabytes of main memory. At the same time, prices for server main memory dramatically dropped over the last few decades. This increase in capacity and reduction in cost makes it a viable approach to keep huge amounts of business data in memory. This section discusses the benefits and challenges.

2.1.1 Using main memory as the data store

The most obvious reason to use main memory as the data store for a database is because accessing data in main memory is much faster than accessing data on disk. Figure 2-1 compares the access times for data in several locations.

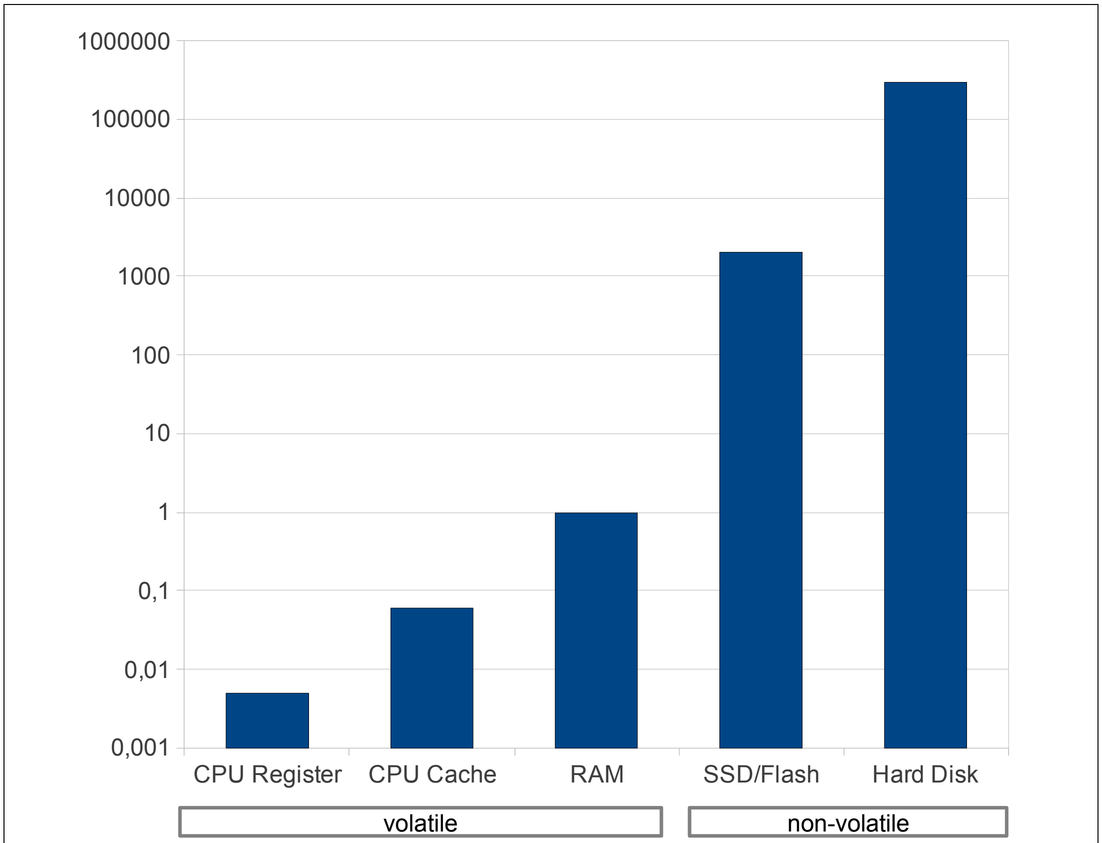


Figure 2-1 Data access times of various storage types, relative to RAM (logarithmic scale)

The main memory (RAM) is the fastest storage type that can hold a significant amount of data. While CPU registers and CPU cache are faster to access, their usage is limited to the actual processing of data. Data in main memory can be accessed more than a hundred thousand times faster than data on a spinning hard disk, and even flash technology storage is about a thousand times slower than main memory. Main memory is connected directly to the processors through a very high-speed bus, whereas hard disks are connected through a chain of buses (QPI, PCIe, SAN) and controllers (I/O hub, RAID controller or SAN adapter, and storage controller).

Compared with keeping data on disk, keeping the data in main memory can dramatically improve database performance just by the advantage in access time.

2.1.2 Data persistence

Keeping data in main memory brings up the question of what will happen in case of a loss of power.

In database technology, atomicity, consistency, isolation, and durability (ACID) is a set of requirements that guarantees that database transactions are processed reliably:

- A transaction has to be atomic. That is, if part of a transaction fails, the entire transaction has to fail and leave the database state unchanged.
- The consistency of a database must be preserved by the transactions that it performs.
- Isolation ensures that no transaction is able to interfere with another transaction.
- Durability means that after a transaction has been committed it will remain committed.

While the first three requirements are not affected by the in-memory concept, durability is a requirement that cannot be met by storing data in main memory alone. Main memory is volatile storage. That is, it looses its content when it is out of electrical power. To make data persistent, it has to reside on non-volatile storage, such as hard drives, SSD, or Flash devices.

The storage used by a database to store data (in this case, main memory) is divided into pages. When a transaction changes data, the corresponding pages are marked and written to non-volatile storage in regular intervals. In addition, a database log captures all changes made by transactions. Each committed transaction generates a log entry that is written to non-volatile storage. This ensures that all transactions are permanent. Figure 2-2 illustrates this using the example of SAP HANA. SAP HANA stores changed pages in savepoints, which are asynchronously written to persistent storage in regular intervals (by default every 5 minutes). The log is written synchronously. That is, a transaction does not return before the corresponding log entry has been written to persistent storage, in order to meet the durability requirement, as described above.

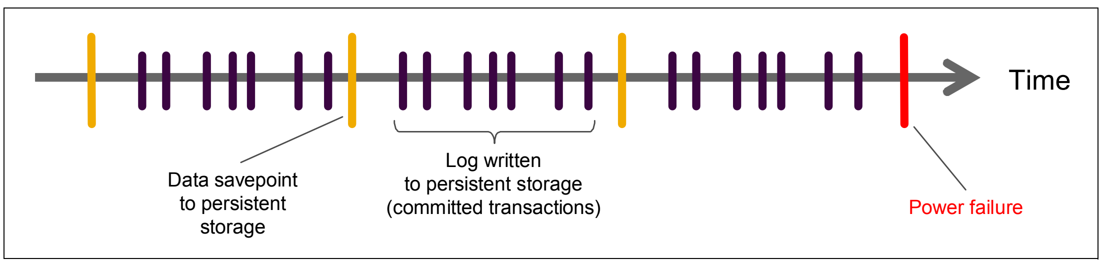


Figure 2-2 Savepoints and logs in SAP HANA

After a power failure, the database can be restarted like a disk-based database. The database pages are restored from the savepoints, and then the database logs are applied (rolled forward) to restore the changes that were not captured in the savepoints. This ensures that the database can be restored in memory to exactly the same state as before the power failure.

2.2 Minimizing data movement

The second key to improving data processing performance is to minimize the movement of data within the database and between the database and the application. This section describes measures to achieve this.

2.2.1 Compression

Even though today's memory capacities allow keeping enormous amounts of data in-memory, compressing the data in-memory is still desirable. The goal is to compress data in a way that does not use up performance gained, while still minimizing data movement from RAM to the processor.

By working with dictionaries to be able to represent text as integer numbers, the database can compress data significantly and thus reduce data movement, while not imposing additional CPU load for decompression, but even adding to the performance[1]. Figure 2-3 illustrates this with a simplified example.

Row ID	Date/ Time	Material	Customer Name	Quantity
1	14:05	Radio	Dubois	1
2	14:11	Laptop	Di Dio	2
3	14:32	Stove	Miller	1
4	14:38	MP3 Player	Newman	2
5	14:48	Radio	Dubois	3
6	14:55	Refrigerator	Miller	1
7	15:01	Stove	Chevrier	1
⋮				

#	Customers
1	Chevrier
2	Di Dio
3	Dubois
4	Miller
5	Newman

#	Material
1	MP3 Player
2	Radio
3	Refrigerator
4	Stove
5	Laptop

Row ID	Date/ Time	Material	Customer Name	Quantity
1	845	2	3	1
2	851	5	2	2
3	872	4	4	1
4	878	1	5	2
5	888	2	3	3
6	895	3	4	1
7	901	4	1	1
⋮				

Figure 2-3 Illustration of dictionary compression

On the left-hand side of Figure 2-3 the original table is shown containing text attributes (that is, material and customer name) in their original representation. The text attribute values are stored in a dictionary (upper right), assigning an integer value to each distinct attribute value. In the table the text is replaced by the corresponding integer value as defined in the dictionary. The date/time attribute has also been converted to an integer representation. Using dictionaries for text attributes reduces the size of the table, because each distinct attribute value has only to be stored once, in the dictionary, so each additional occurrence in the table just needs to be referred to with the corresponding integer value.

[1] See the example in Figure 2-5 on page 12.

The compression factor achieved by this method is highly dependent on data being compressed. Attributes with few distinct values compress very well, whereas attributes with many distinct values do not benefit as much.

While there are other, more effective compression methods that could be employed with in-memory computing, to be useful, they must have the correct balance between compression effectiveness, which gives you more data in your memory, or less data movement (that is, higher performance), resources needed for decompression, and data accessibility (that is, how much unrelated data has to be decompressed to get to the data that you need). As discussed here, dictionary compression combines good compression effectiveness with low decompression resources and high data access flexibility.

2.2.2 Columnar storage

Relational databases organize data in tables, which contain the data records. The difference between row-based and columnar (or column-based) storage is the way in which the table is stored:

- Row-based storage stores a table in a sequence of rows.
- Column-based storage stores a table in a sequence of columns.

Figure 2-4 illustrates this.

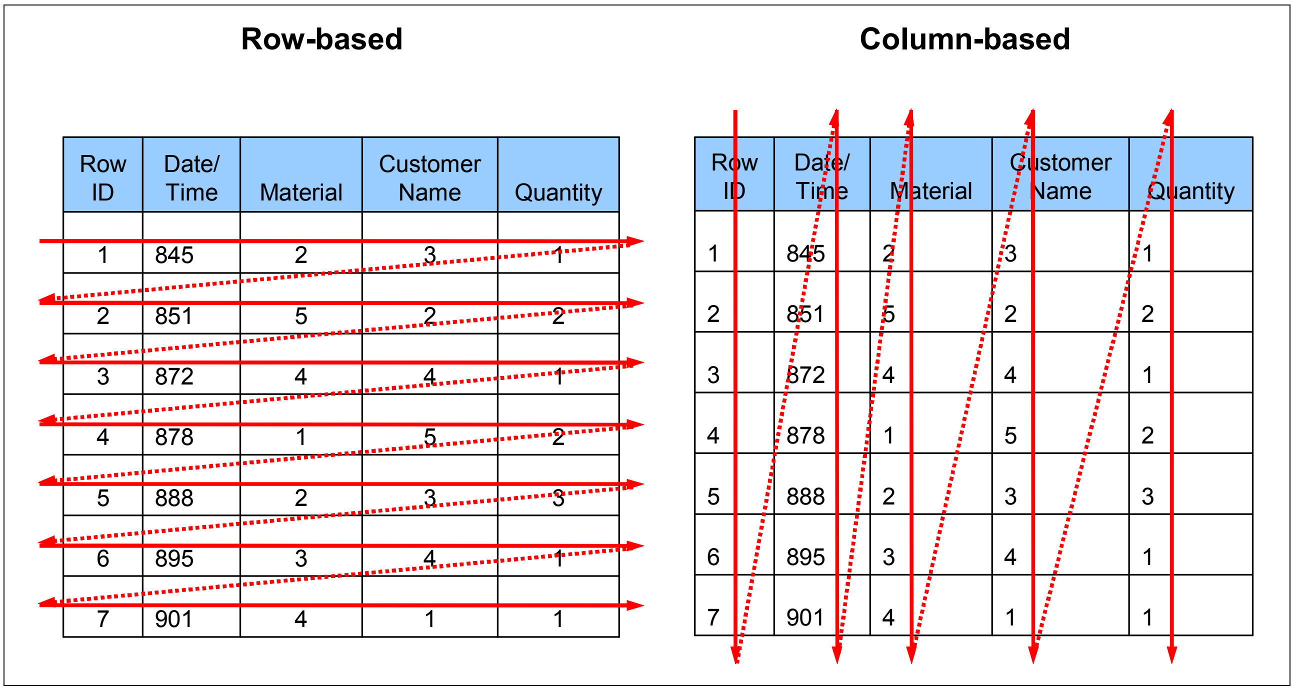


Figure 2-4 Row-based and column-based storage models

Both storage models have benefits and drawbacks (Table 2-1).

Table 2-1 Benefits and drawbacks of row-based and column-based storage

	Row-based storage	Column-based storage
Benefits	► Record data is stored together. ► Easy to insert/update.	► Only affected colums have to be read during the selection process of a query. ► Efficient projections[a]. ► Any column can serve as an index.
Drawbacks	► All data has to be read during selection, even if only a few columns are involved in the selection process.	► After selection, selected rows have to be reconstructed from columns. ► No easy insert/update.

a. Projection: View on the table with a subset of columns

The drawbacks of column-based storage are not as grave as they seem. In most cases, not all attributes (that is, column values) of a row are needed for processing, especially in analytic queries. Also, inserts or updates to the data are less frequent in an analytical environment[2].

To show the benefits of dictionary compression combined with columnar storage, Figure 2-5 shows an example of how a query is executed. Figure 2-5 refers to the table shown in Figure 2-3 on page 10.

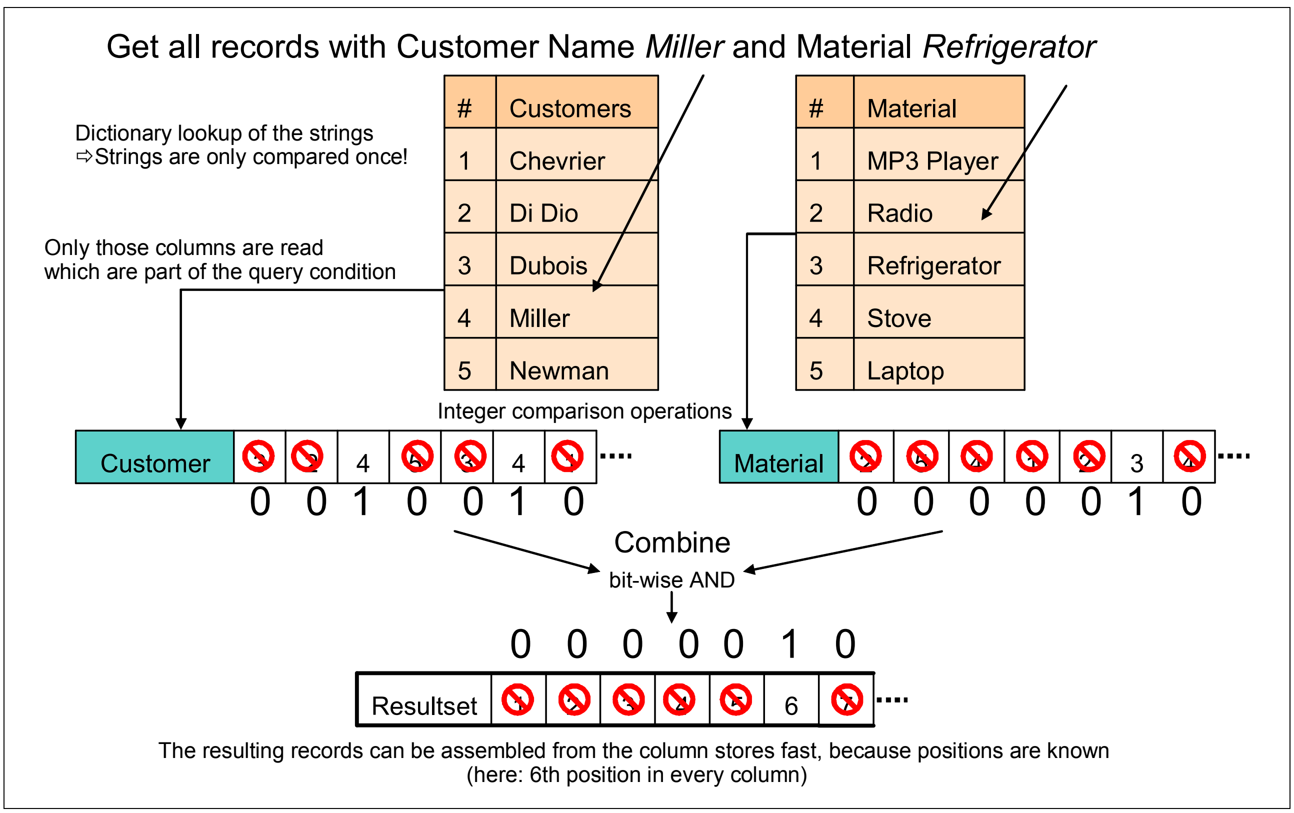


Figure 2-5 Example of a query executed on a table in columnar storage

[2] An exception is bulk loads (for example, when replicating data in the in-memory database, which can be handled differently).

The query asks to get all records with Miller as the customer name and Refrigerator as the material.

First, the strings in the query condition are looked up in the dictionary. Miller is represented as the number 4 in the customer name column. Refrigerator is represented as the number 3 in the material column. Note that this lookup has to be done only once. Subsequent comparison with the values in the table will be based on integer comparisons, which are less resource intensive than string comparisons.

In a second step the columns are read that are part of the query condition (that is, the Customer and Material columns). The other columns of the table are not needed for the selection process. The columns are then scanned for values matching the query condition. That is, in the Customer column all occurrences of 4 are marked as selected, and in the Material column all occurrences of 3 are marked.

These selection marks can be represented as bitmaps, a data structure that allows efficient boolean operations on them, which is used to combine the bitmaps of the individual columns to a bitmap representing the selection or records matching the entire query condition. In our example the record number 6 is the only matching record. Depending on the columns selected for the result, now the additional columns have to be read in order to compile the entire record to return. But because the position within the column is known (record number 6) only the parts of the columns have to be read that contain the data for this record.

This example shows how compression not only can limit the amount of data needed to be read for the selection process, but even simplify the selection itself, while the columnar storage model further reduces the amount of data needed for the selection process. Although the example is simplified, it illustrates the benefits of dictionary compression and columnar storage.

2.2.3 Pushing application logic to the database

Whereas the concepts described above speed up processing within the database, there is still one factor that can significantly slow down the processing of data. An application executing the application logic on the data has to get the data from the database, process it, and possibly send it back to the database to store the results. Sending data back and forth between the database and the application usually involves communication over a network, which introduces communication overhead and latency and is limited by the speed and throughput of the network between the database and the application itself.

To eliminate this factor to increase overall performance, it is beneficial to process the data where it is, at the database. If the database is able to perform calculations and apply application logic, less data needs to be sent back to the application and might even eliminate the need for the exchange of intermediate results between the database and the application. This minimizes the amount of data transfer, and the communication between database and application contributes a less significant amount of time to the overall processing time.

2.3 Dividing and conquering

The phrase "divide and conquer" (derived from the Latin saying *divide et impera*) is typically used when a big problem is divided into a number of smaller, easier-to-solve problems. With regard to performance, processing huge amounts of data is a big problem that can be solved by splitting it up into smaller chunks of data, which can be processed in parallel.

2.3.1 Parallelization on multi-core systems

When chip manufactures reached the physical limits of semiconductor-based microelectronics with their single-core processor designs, they started to increase processor performance by increasing the number of cores, or processing units, within a single processor. This performance gain can only be leveraged through parallel processing because the performance of a single core remained unchanged.

The rows of a table in a relational database are independent of each other, which allows parallel processing. For example, when scanning a database table for attribute values matching a query condition, the table, or the set of attributes (columns) relevant to the query condition, can be divided into subsets and spread across the cores available to parallelize the processing of the query. Compared with processing the query on a single core, this basically reduces the time needed for processing by a factor equivalent to the number of cores working on the query (for example, on a 10-core processor the time needed is one-tenth of the time that a single core would need).

The same principle applies for multi-processor systems. A system with eight 10-core processors can be regarded as an 80-core system that can divide the processing into 80 subsets processed in parallel, as described above.

2.3.2 Data partitioning and scale-out

Even though servers available today can hold terabytes of data in memory and provide up to eight processors per server with up to 10 cores per processor, the amount of data to be stored in an in-memory database or the computing power needed to process such quantities of data might exceed the capacity of a single server. To accommodate the memory and computing power requirements that go beyond the limits of a single server, data can be divided into subsets and placed across a cluster of servers, forming a distributed database (scale-out approach).

The individual database tables can be placed on different servers within the cluster, or tables bigger than what a single server can hold can be split into several partitions, either horizontally (a group of rows per partition) or vertically (a group of columns per partition), with each partition residing on a separate server within the cluster.

3

SAP HANA

In this chapter we describe the SAP HANA offering and its architecture and components, use cases, delivery model, and sizing and licensing aspects.

This chapter contains the following sections:

- 3.1, "SAP HANA overview" on page 16
- 3.2, "SAP HANA delivery model" on page 17
- 3.3, "SAP HANA use cases" on page 18
- 3.4, "Data replication methods for SAP HANA" on page 23
- 3.5, "Sizing SAP HANA" on page 26
- 3.6, "SAP HANA software licensing" on page 29

3.1 SAP HANA overview

This section gives an overview of SAP HANA. When talking about SAP HANA, these terms are used:

- SAP In-Memory Appliance (SAP HANA)

  SAP HANA is a flexible, data source agnostic appliance that allows you to analyze large volumes of data in real time, without the need to materialize aggregations. It is a combination of hardware and software, and it is delivered as an optimized appliance in cooperation with SAP's hardware partners for SAP HANA.

- SAP in-memory database, also referred to as the SAP HANA database

  The SAP in-memory database is a hybrid in-memory database that combines row-based, column-based, and object-based database technology, optimized to exploit the parallel processing capabilities of current hardware. It is the heart of SAP offerings like SAP HANA.

For the sake of simplicity, we use the terms SAP HANA, SAP in-memory database, SAP HANA database, and SAP HANA appliance synonymously in this paper, as we only cover the SAP in-memory database as part of the SAP HANA appliance. Where required, we make sure that the context makes it clear which part we are talking about.

3.1.1 SAP HANA architecture

Figure 3-1 shows the high-level architecture of the SAP HANA appliance. The following sections explain the most important components of this architecture.

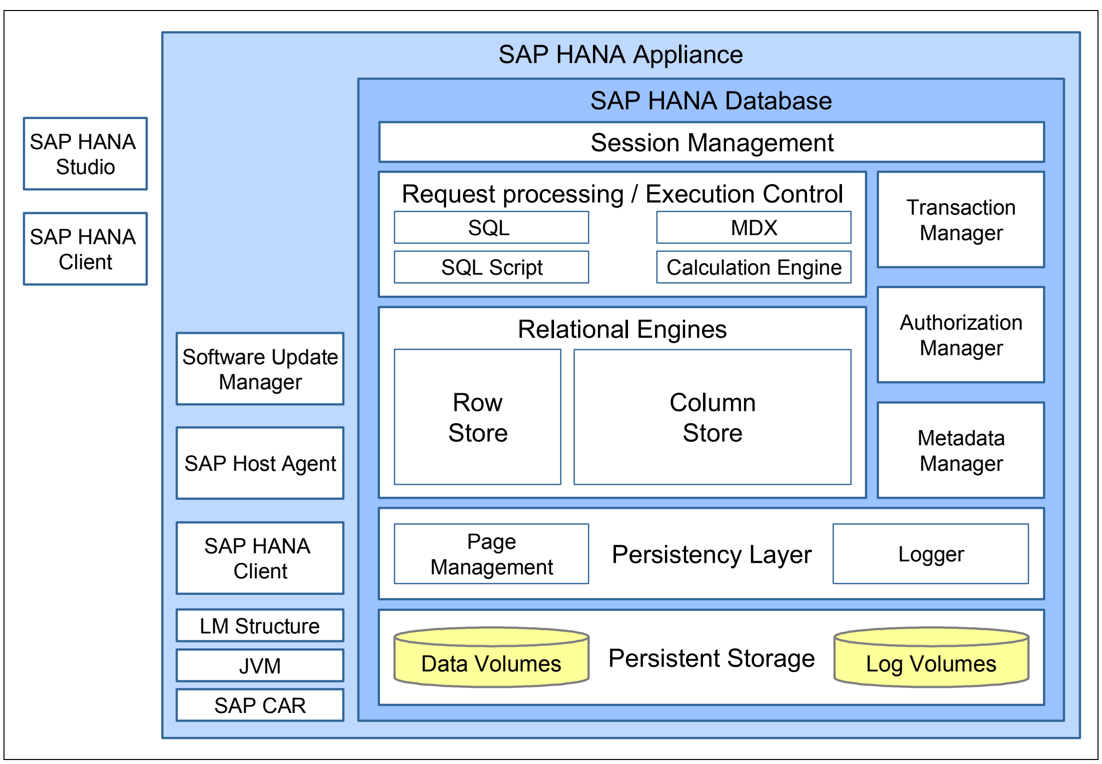


Figure 3-1 SAP HANA architecture

SAP HANA database

The heart of the SAP HANA database is the relational database engines. There are two engines within the SAP HANA database:

- The column-based store, storing relational data in columns, optimized holding tables with huge amounts of data, which are aggregated and used in analytical operations.
- The row-based store, storing relational data in rows, as traditional database systems do. This row store is more optimized for write operation and has a lower compression rate, and query performance is much lower compared to the column-based store.

The engine used to store data can be selected on a per-table basis at the time of creation of a table[1]. Tables in the row-store are loaded at startup time, whereas tables in the column-store can be either loaded at startup or on demand, during normal operation of the SAP HANA database.

Both engines share a common persistency layer, which provides data persistency consistent across both engines. There is page management and logging, much like in traditional databases. Changes to in-memory database pages are persisted through savepoints written to the data volumes on persistent storage, which is usually hard drives. Every transaction committed in the SAP HANA database is persisted by the logger of the persistency layer in a log entry written to the log volumes on persistent storage. The log volumes use flash technology storage for high I/O performance and low latency.

The relational engines can be accessed through a variety of interfaces. The SAP HANA database supports SQL (JDBC/ODBC), MDX (ODBO), and BICS (SQL DBC). The calculation engine allows calculations to be performed in the database, without moving the data into the application layer. It also includes a business functions library that can be called by applications to do business calculations close to the data. The SAP HANA-specific SQL Script language as an extension to SQL that can be used to push down data-intensive application logic into the SAP HANA database.

3.1.2 SAP HANA appliance

The SAP HANA appliance consists of the SAP HANA database, as described above, and adds components needed to work with, administer, and operate the database. It contains the installation files for the SAP HANA Studio, which is an Eclipse-based administration and data-modeling tool for SAP HANA, in addition to the SAP HANA client, a set of libraries required for applications to be able to connect to the SAP HANA database. Both the SAP HANA Studio and the client libraries are usually installed on a client PC or server.

The Software Update Manager (SUM) for SAP HANA is the framework allowing the automatic download and installation of SAP HANA updates from SAP Marketplace and other sources using a host agent. It also allows distribution of the Studio repository to the users.

The Lifecycle Management (LM) Structure for HANA is a description of the current installation and is, for example, used by SUM to perform automatic updates.

3.2 SAP HANA delivery model

SAP decided to deploy SAP HANA as an integrated solution combining software and hardware, frequently referred to as the SAP HANA appliance. As with BW Accelerator, SAP partners with several hardware vendors to provide the infrastructure needed to run the SAP

[1] The storage type can be selected during and changed after creation.

HANA software. IBM was among the first hardware vendors to partner with SAP to provide an integrated solution.

Infrastructure for SAP HANA needs to run through a certification process to ensure that certain performance requirements are met. Only certified configurations are supported by SAP and the respective hardware partner. These configurations have to adhere to certain requirements and restrictions to provide a common platform across all hardware providers:

- Only certain Intel Xeon processors can be used. For the currently available Intel Xeon processor E7 family, the allowed processor models are E7-2870, E7-4870, and E7-8870. The previous CPU generation was limited to Intel Xeon processor X7560.
- All configurations have to provide a certain main memory per core ratio, which is defined by SAP to balance CPU processing power and the amount of data being processed.
- All configurations have to meet minimum performance requirements for various load profiles. SAP tests for these requirements as part of the certification process.
- The capacity of the storage devices used in the configurations have to meet the sizing rules (see 3.5, “Sizing SAP HANA” on page 26).
- The networking capabilities of the configurations must include 10Gb Ethernet for the SAP HANA software.

By imposing these requirements, SAP can rely on the availability of certain features and ensure a well-performing hardware platform for their SAP HANA software. These requirements give the hardware partners enough room to develop an infrastructure architecture for SAP HANA, which adds differentiating features to the solution. The benefits of the IBM solution are described in Chapter 4, “The IBM Systems Solution for SAP HANA” on page 31.

3.3 SAP HANA use cases

SAP HANA can be used in a number of ways. The following sections give an overview of the various types of usage scenarios.

3.3.1 SAP HANA in a data mart side-car approach

The following usage scenarios are often referred to as a data mart or *side-car approach* because SAP HANA sits by the operational system and receives the operational data (usually only an excerpt) from this system by means of replication.

SAP HANA in a side-car approach as a data mart

In a typical SAP-based application landscape today, you will find a number of SAP systems like ERP, CRM, SCM, and other, possibly non-SAP, applications. All these systems contain loads of operational data, which can be used to improve business decision making using business intelligence (BI) technology. Data used for BI purposes can be gathered either on a business unit level using data marts or on an enterprise level with an enterprise data warehouse, like SAP NetWeaver Business Warehouse (SAP BW). ETL processes feed the data from the operational systems into the data marts and the enterprise data warehouse. Figure 3-2 illustrates such a typical landscape.

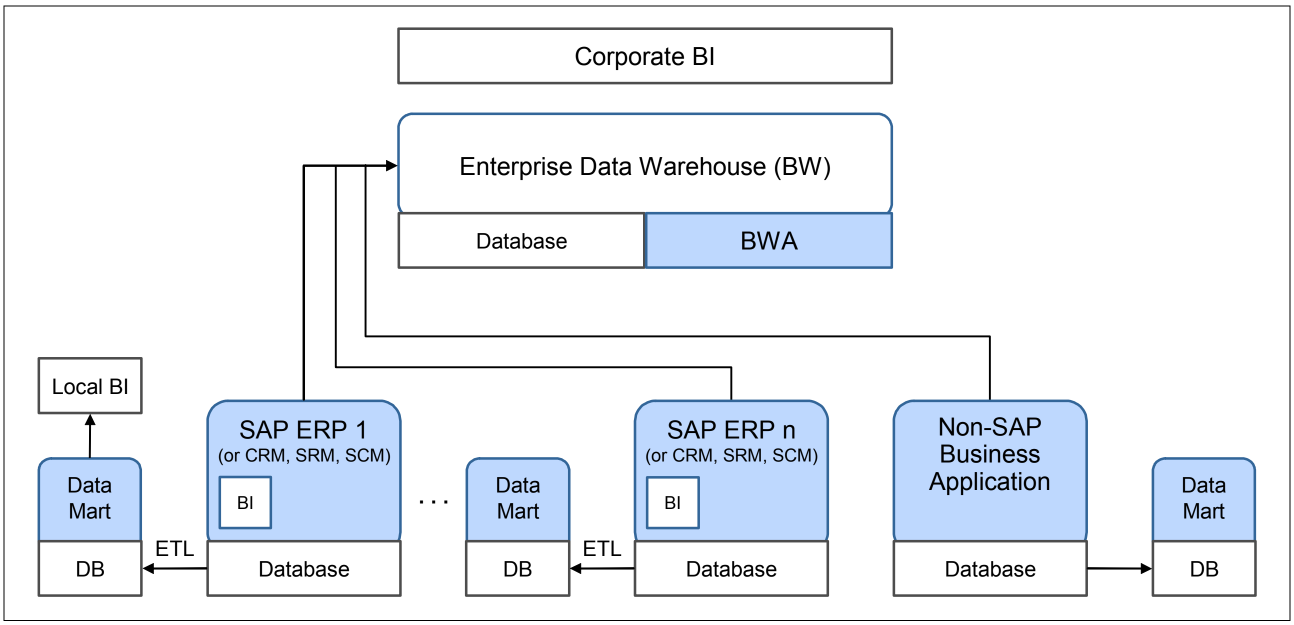


Figure 3-2 Typical view of an SAP-based application landscape today

With the huge amounts of data collected in an enterprise data warehouse, response times of queries for reports or navigation through data can become an issue, generating new requirements to the performance of such an environment. To address these requirements, SAP has introduced SAP NetWeaver Business Warehouse Accelerator (BWA), which is built for this use case by speeding up queries and reports in SAP BW by leveraging in-memory technology. While being a perfect fit for an enterprise data warehouse holding huge amounts of data, the combination of SAP BW and BWA is not always a viable solution for the relatively small data marts.

With the introduction of SAP HANA 1.0, SAP provided an in-memory technology aiming to support BI at a business unit level. SAP HANA combined with BI tools like the SAP BusinessObjects tools, and data replication mechanisms feeding data from the operational system into SAP HANA in real-time, brought in-memory computing to the business unit level. Figure 3-3 shows such a landscape with the local data marts replaced by SAP HANA.

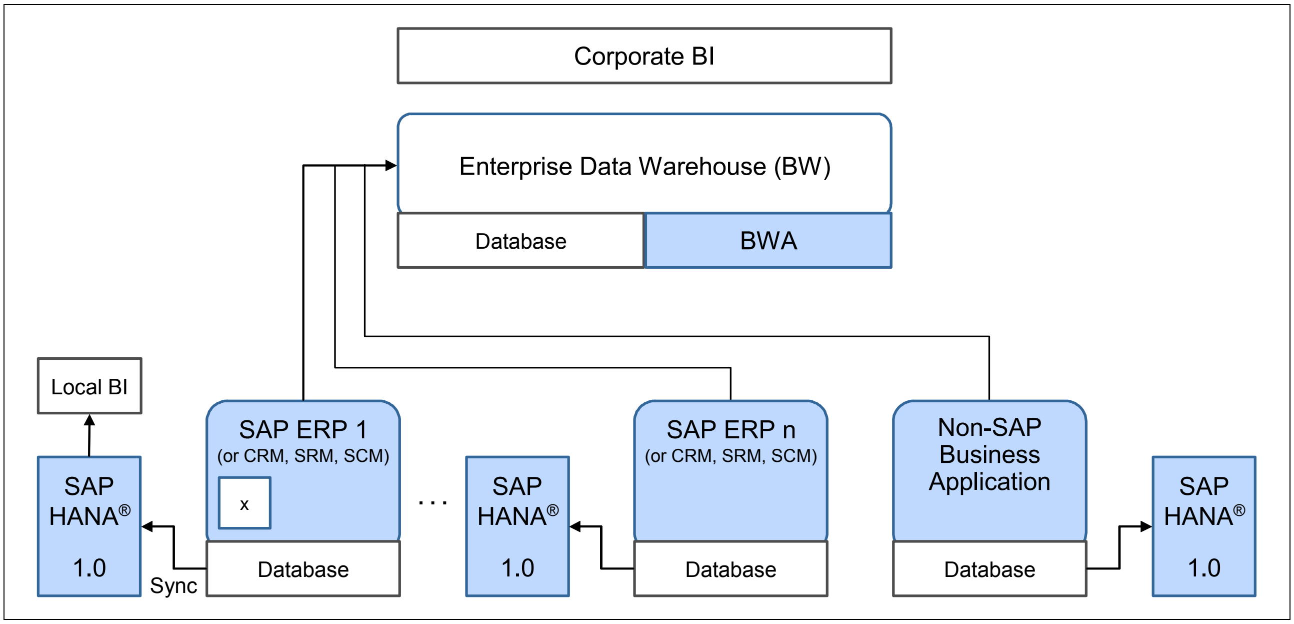


Figure 3-3 SAP vision after the introduction of SAP HANA 1.0

Business intelligence functionality is provided by a SAP BusinessObjects BI tool such as SAP BusinessObjects Explorer communicating with the SAP HANA database through the BI Consumer Services (BICS) interface.

SAP offers SAP Rapid Deployment Solutions (SAP RDS), which provides predefined but customizable SAP HANA data models for SAP BusinessObjects Explorer reports:

- The *SAP ERP rapid-deployment solution for operational reporting with SAP HANA* provides a set of reporting models from business areas such as sales, financial, shipping, purchasing, and master data. The use of SAP HANA ensures optimal performance even with high-data volumes (which cannot currently be provided without in-memory technology). Different SAP BusinessObject clients can be used to visualize the results.
- The *SAP ERP rapid-deployment solution for profitability analysis with SAP HANA* enables customers using SAP ERP Cost and Profitability Analysis (CO-PA) to access and use large sets of profitability data more quickly and efficiently than ever before. SAP HANA is used as a secondary database for CO-PA to accelerate all data read processes for reporting and allocations.
- The *SAP CRM rapid-deployment solution for customer segmentation with SAP HANA* enables marketers to analyze and segment their customers in ways that were impossible to realize without the use of in-memory technology. The power of SAP HANA makes it possible to explore customer data at enormous speeds, making it much easier to develop extremely fine segments of customers with unique combinations of offers and messages.

These SAP Rapid Deployment Solutions enables customers to implement SAP HANA technology quickly.

SAP HANA in a side-car data mart approach for new applications

SAP HANA in a side-car approach for new applications is very similar to a side-car approach for BI purposes. The difference is that the consumer of the data replicated to SAP HANA is not a BI tool, but the source system itself. The source system can use the in-memory capabilities of the SAP HANA database to run analytical queries on the replicated data. This helps applications performing queries on huge amounts of data to run simulations, pattern recognition, planning runs, and so on.

Figure 3-4 illustrates the side-car approach with SAP HANA for such new applications.

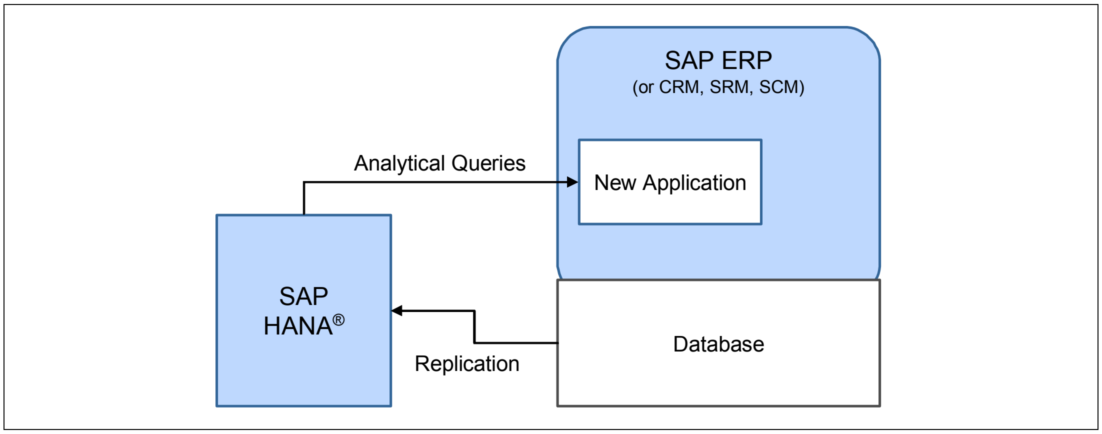


Figure 3-4 SAP HANA in a side-car approach with new applications

Here are examples of these new kinds of applications:

- SAP Smart Meter Analytics, powered by SAP HANA, helping utility companies to instantly turn huge amounts of data into insights that they can use to develop more tailored customer offerings. With this new in-memory offering, utility companies can react quicker to changes in the energy consumption of their customers, or help customers adopt more sustainable energy-use practices.
- SAP COPA Accelerator software improves the speed and depth of working with large volumes of financial data. With the SAP HANA platform, companies can now run complex analyses such as simulations and pattern recognition in seconds.

SAP HANA in a side-car approach allows the creation of new applications on the SAP NetWeaver platform, which leverage the power of in-memory computing, enabling customers to run their business in innovative ways that were not possible before.

3.3.2 SAP HANA as a database for SAP NetWeaver BW

Since November 2011, SAP HANA can be used as the database for an SAP NetWeaver Business Warehouse (SAP BW) installation. In this scenario SAP HANA replaces the traditional database server of a SAP BW installation. The application servers stay the same.

Prerequisites for this scenario are the following software levels:

- SAP NetWeaver BW 7.30 SP5
- SAP HANA 1.0 SPS03

The in-memory performance of SAP HANA dramatically improves query performance and eliminates the need for manual optimizations by materialized aggregates in SAP BW. Figure 3-5 shows SAP HANA as the database for the SAP NetWeaver Business Warehouse.

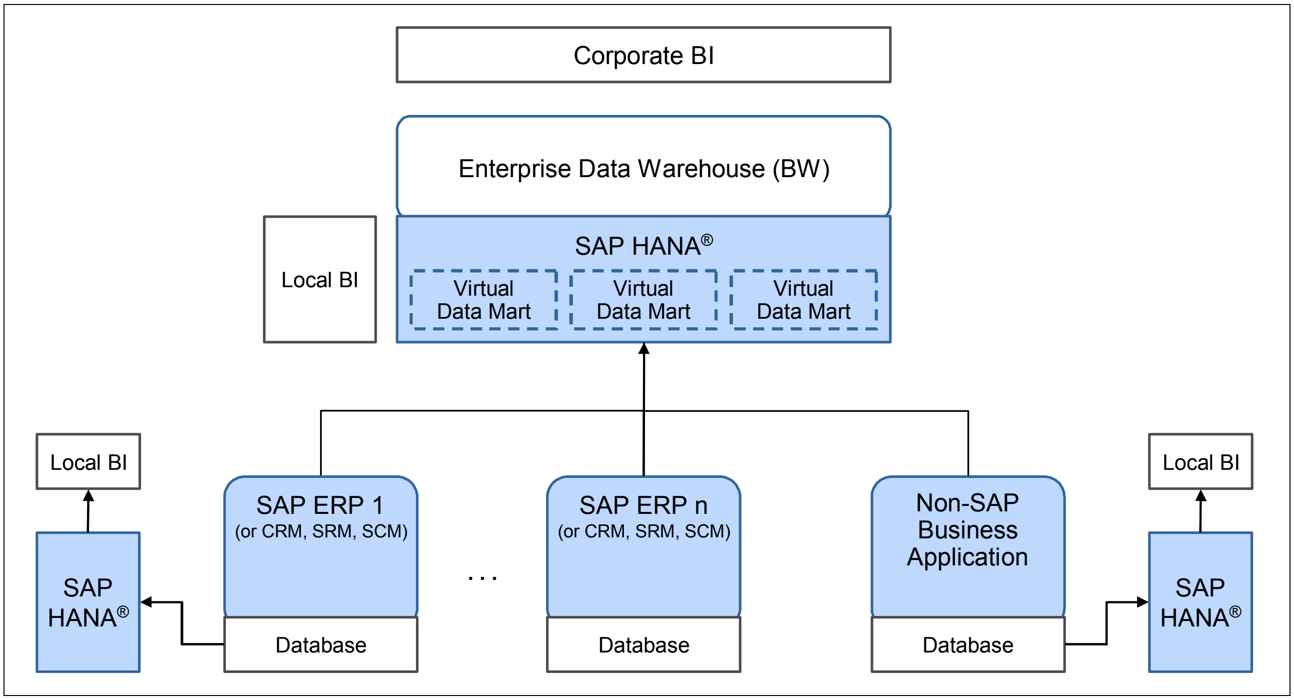


Figure 3-5 SAP vision after the introduction of SAP HANA 1.0 as the database for SAP NetWeaver Business Warehouse

In contrast to a SAP BW system accelerated by the in-memory capabilities of SAP BW Accelerator, a SAP BW system with SAP HANA as the database keeps *all* data in-memory. With SAP BW Accelerator, the customer chooses the data to be accelerated, which is then copied to the SAP BW Accelerator. Here the traditional database server (for example, IBM DB2® or Oracle) still acts as the primary datastore.

3.3.3 SAP HANA with third-party software

While the focus for SAP HANA is within the SAP ecosystem, its openness also enables other usage scenarios. Data replication from third-party applications for BI purposes, or even the use of SAP HANA as a database for analytical applications completely unrelated to SAP, is possible.

3.4 Data replication methods for SAP HANA

When SAP HANA is used in a side-car data mart approach (that is, not as a database server), data has to be fed from the source system into SAP HANA by replication. There are several replication methods available to accomplish this:

- Log-based replication: This method is based on reading the transaction logs from the source database and re-applying them to the SAP HANA database.
- Trigger-based replication: This method is based on database triggers on the source system to record changes to the source database and replicate these changes to the SAP HANA database.
- ETL-based replication: This method employs an Extract, Transform, and Load (ETL) process to extract data from the data source, transform it to meet the business or technical needs, and then load it into the SAP HANA database.

Figure 3-6 illustrates these replication methods.

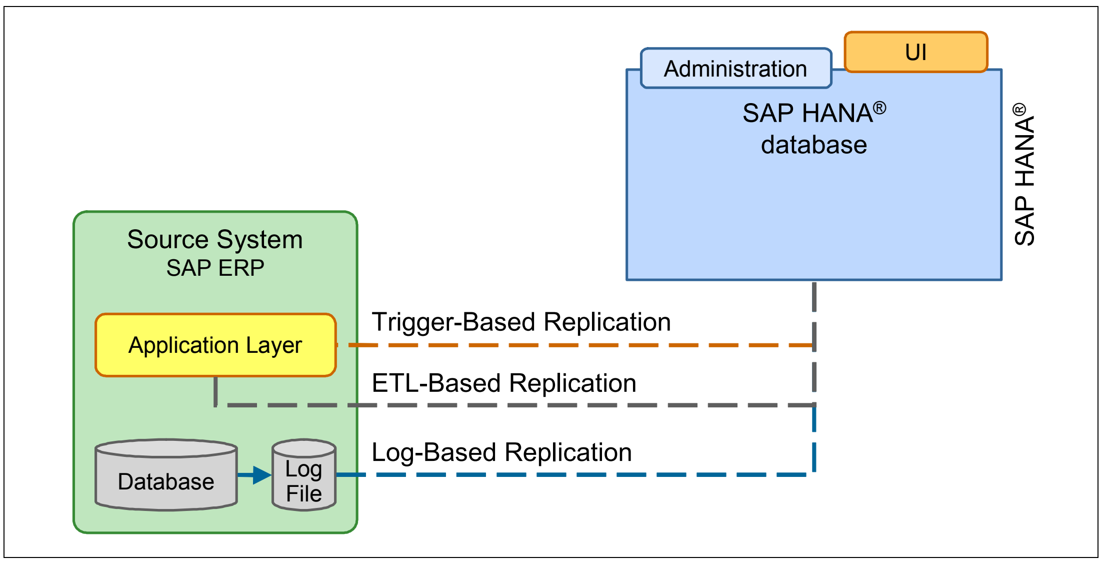


Figure 3-6 Replication methods for SAP HANA

The following sections discuss these replication methods for SAP HANA in more detail.

3.4.1 Log-based replication with Sybase Replication Server

The log-based replication for SAP HANA is realized with Sybase Replication Server. It captures table changes from low-level database log files and transforms them into SQL statements that are in turn executed on the SAP HANA database. This is very similar to what is known as *log shipping* between two database instances.

Replication with the Sybase Replication Server is very fast and consumes little processing power due to its closeness to the database system. However, this mode of operation makes this replication method highly database dependant, and the source database system coverage is limited[2]. It also limits the conversion capabilities, and so replication with Sybase Replication Server only supports Unicode source databases. The Sybase Replication Server cannot convert between code pages, and because SAP HANA works with unicode encoding

[2] Only certain versions of IBM DB2 on AIX, Linux, and HP-UX are supported with this replication method.

internally, the source database has to use unicode encoding as well. Also, certain table types used in SAP systems are unsupported.

To set up replication with the Sybase Replication Server, the tables chosen to be replicated must initially be copied from the source database to the SAP HANA database. This initial load is done with the R3Load program, also used for database imports and exports. During copying the source tables must not be altered, which usually is only achievable with a planned downtime of the system.

This replication method is only recommended for SAP customers who have been invited to use it during the ramp up of SAP HANA 1.0.

SAP recommends using trigger-based data replication using the SAP LT Replicator, which is described in the next section.

3.4.2 Trigger-based replication with SAP LT Replication

SAP Landscape Transformation (LT) Replication is also based on tracking database changes, but on a much higher level. The efficient initialization of data replication is based on database triggers and a delta logging concept. It allows real-time or scheduled data replication of the tables that you choose.

Because it is operating on the application level, the trigger-based replication method benefits from the database abstraction provided by the SAP software stack, which makes it database independent. It also has extended source system release coverage, supporting releases starting from SAP R/3 4.6C up to the newest SAP ERP version.

SAP LT Replication leverages proven System Landscape Optimization (SLO) technologies (such as Near Zero Downtime, Test Data Migration Server (TDMS), and SAP LT) and can handle both unicode and non-unicode source databases. The SAP LT Replication Server provides a flexible and reliable replication process, fully integrates with SAP HANA Studio, and is simple and fast to set up.

he SAP LT Replication Server does not have to be a separate SAP system. It can run on any SAP system with the SAP NetWeaver 7.02 ABAP stack (Kernel 7.20EXT). However, we recommend installing the SAP LT Replication Server on a separate system when high replication load would impact the performance of the base system (that is, the source system).

The SAP LT Replication Server is the ideal solution for all SAP HANA customers who need real-time (or scheduled) data replication sourcing from SAP ERP systems into SAP HANA.

3.4.3 ETL-based replication with SAP BusinessObjects Data Services

An ETL-based replication for SAP HANA can be set up using SAP BusinessObjects Data Services, which is a full-featured ETL tool, giving customers maximum flexibility with regard to the source database system:

- Customers can specify and load the relevant business data in defined periods of time from an ERP system into the HANA database.
- ERP application logic can be reused by reading extractors or utilizing SAP function modules.
- It offers options for the integration of third-party data providers, and supports replication from virtually any data source.

Data transfers are done in batch mode, which limits the real-time capabilities of this replication method.

SAP BusinessObjects Data Services provides several kinds of data quality and data transformation functionality. Due to the rich feature set available, implementation time for the ETL-based replication is longer than for the other replication methods. SAP BusinessObjects Data Services offers integration with SAP HANA. SAP HANA is a available as a predefined data target for the load process.

The ETL-based replication server is the ideal solution for all SAP HANA customers who need data replication from non-SAP data sources.

3.4.4 Comparison of the replication methods

Each of the described data replication methods for SAP HANA has its benefits and weaknesses:

- The log-based replication method with the Sybase Replication Server provides the fastest replication from the source database to SAP HANA. However, it is limited to unicode-encoded source databases, it does not support all table types used in SAP applications, it provides no transformation functionality, and the source database system support is very limited.
- The trigger-based replication method with the SAP LT system provides real-time replication while supporting a wide range of source database systems. It can handle both unicode and non-unicode databases, and makes use of proven data migration technology. It leverages the SAP application layer, which limits it to SAP source systems. Compared to the log-based replication method, it offers a broader support of source systems, while providing almost similar real-time capabilities, and for that reason it is recommended for replication from SAP source systems.
- The ETL-based replication method is the most flexible of all, paying the price for flexibility with only near real-time capabilities. With its variety of possible data sources, advanced transformation, and data quality functionality, it is the ideal choice for replication from non-SAP data sources.

Figure 3-7 shows these in comparison.

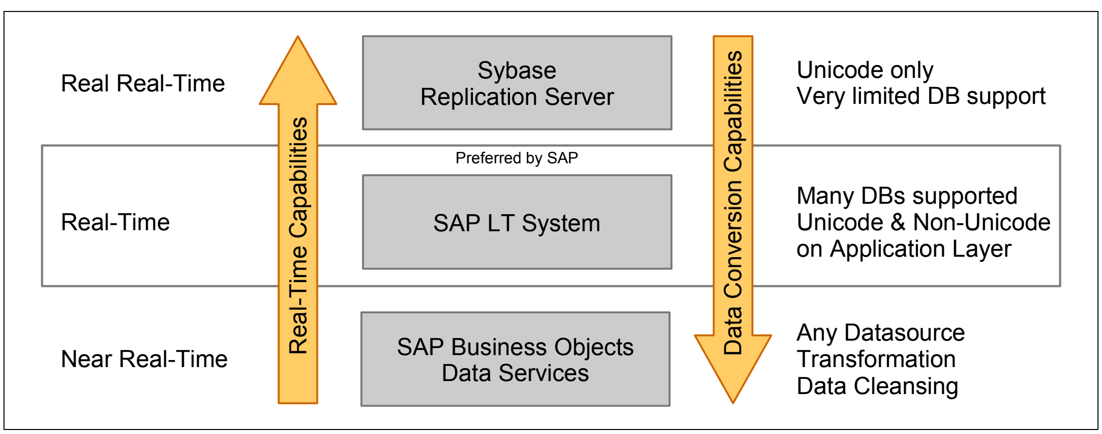


Figure 3-7 Comparison of the replication methods for SAP HANA

3.5 Sizing SAP HANA

This section introduces the concept of T-shirt sizes for SAP HANA and gives a short overview of how to size for an SAP HANA system.

3.5.1 The concept of T-shirt sizes for SAP HANA

SAP defined so-called T-shirt sizes for SAP HANA to both simplify the sizing and to limit the number of hardware configurations to support, thus reducing complexity. SAP's hardware partners provide configurations for SAP HANA according to one or more of these T-shirt sizes. Table 3-1 lists the T-shirt sizes for SAP HANA.

Table 3-1 SAP HANA T-shirt sizes

SAP T-shirt size	XS	S and S+	M and M+	L
Compressed data in memory	64 GB	128 GB	256 GB	512 GB
Server main memory	128 GB	256 GB	512 GB	1024 GB
Number of CPUs	2	2	4	8

In addition to the T-shirt sizes listed in Table 3-1, you might come across the T-shirt size XL, which denotes a scale-out configuration for SAP HANA.

The T-shirt sizes S+ and M+ denote upgradable versions of the S and M sizes:

- S+ delivers capacity equivalent to S, but the hardware is upgradable to an M size.
- M+ delivers capacity equivalent to M, but the hardware is upgradable to an L size.

These T-shirt sizes are used when relevant growth of the data size is expected.

3.5.2 Sizing approach

The sizing of SAP HANA depends on the scenario in which SAP HANA is used. We discuss these scenarios here:

- SAP HANA in a side-car data mart approach, used for business intelligence or new applications, as described in 3.3.1, "SAP HANA in a data mart side-car approach" on page 18
- SAP HANA as the database for a SAP NetWeaver Business Warehouse 7.30 SP5, as described in 3.3.2, "SAP HANA as a database for SAP NetWeaver BW" on page 21

The sizing methodology for SAP HANA is described in detail in SAP Note 1514966[3] (SAP HANA in a side-car approach) and SAP Note 1637145 (SAP HANA as the database for SAP BW). The the following sections provide a brief overview of sizing for SAP HANA in a sidecar approach.

[3] SAP Notes can be accessed via http://service.sap.com/notes. An SAP S-user ID is required.

Sizing the RAM needed

Sizing an SAP HANA system is mainly based on the amount of data to be loaded into the SAP HANA database, because this determines the amount of main memory (or RAM) needed in an SAP HANA system. To size the RAM, the following steps have to be performed:

1. Determine the information that has to be transferred (either by replication or extraction) to the SAP HANA database. Note that typically customers will only select a sub-set of information from their ERP or CRM database, so this has to be done at the table level. The sizing methodology is based on uncompressed source data size, so in case compression is used in the source database, this has to be taken into account as well. The information required for this step can be acquired with database tools. SAP Note 1514966 contains a script supporting this process for several database systems, for example, DB2 LUW and Oracle.

   The current size of all the tables (without DB indexes) storing the required information in the source database is denoted as A.

2. Although the compression ratio achieved by SAP HANA can vary depending on the data distribution, a working assumption is that, in general, a compression factor of 7 can be achieved:

   ```
   B = ( A / 7 )
   ```

   B is the amount of RAM required to store the data in the SAP HANA database.

3. Only 50% of the total RAM should be used for the in-memory database. The other 50% is needed for temporary objects (for example, intermediate results), the operating system, and application code:

   ```
   C = B * 2
   ```

   C is the total amount of RAM required.

The total amount of RAM should be rounded up to the next T-shirt configuration size, as described in 3.5.1, “The concept of T-shirt sizes for SAP HANA” on page 26, to get the correct T-shirt size needed.

Sizing the disks

The capacity of the disks is based on the total amount of RAM.

As described in 2.1.2, “Data persistence” on page 9, there are two types of storage in SAP HANA:

- $Disk_{persistence}$

  The persistence layer writes snapshots of the database in HANA to disk in regular intervals. These are usually written to an array of SAS drives[4]. The capacity for this storage is calculated based on the total amount of RAM:

  $Disk_{persistence} = 4 * C$

- $Disk_{log}$

  This contains the database logs, written to flash technology storage devices, that is, SSDs or PCIe Flash adapters. The capacity for this storage is calculated based on the total amount of RAM:

  $Disk_{log} = 1 * C$

The certified hardware configurations already take these rules into account, so there is no need to perform this disk sizing. However, we still include it here for your understanding.

[4] The SSD building block, as described in 4.3, “Custom server models for SAP HANA” on page 44, combines $Disk_{persistence}$ and $Disk_{log}$ on a single SSD array with sufficient capacity.

Sizing the CPUs

A CPU sizing only has to be performed in addition to the memory sizing if a massive amount of users working on a relatively small amount of data is expected. Choose the T-shirt configuration size that satisfies both the memory and CPU requirements.

The CPU sizing is user-based. The SAP HANA system has to support 300 SAPS for each concurrently active user. The servers used for the IBM Systems Solution for SAP HANA support about 60 - 65 concurrently active users per CPU, depending on the server model.

Selecting a T-shirt size

According to the sizing results, select a SAP HANA T-shirt size that satisfies the sizing requirements in terms of main memory, and possibly CPU capabilities. For example, a sizing result of 400 GB for the main memory (C) suggests a T-shirt size of M.

The sizing methodology described above is valid for SAP HANA in a side-car approach. Other use cases might require another sizing methodology, for example, for SAP HANA as the database for an SAP BW system. Also, SAP HANA is constantly being optimized, which might affect the sizing methodology. Consult SAP documentation regarding other use cases and up-to-date sizing information.

> **Note:** The sizing approach described here is simplified and can only provide a rough idea of the sizing process for the actual sizing for SAP HANA. Consult the SAP sizing documentation for SAP HANA when performing an actual sizing.

In addition to the sizing methodologies described in SAP Notes, SAP provides sizing support for SAP HANA in SAP Quick Sizer. SAP Quick Sizer is an online sizing tool that supports most of the SAP solutions available. For SAP HANA it supports sizing for these:

- Standalone SAP HANA system, implementing the sizing algorithms described in SAP Note 1514966 (which we described above)
- SAP HANA as the database for a SAP BW system, implementing the sizing algorithms described in SAP Note 1637145
- Special sizing support for the SAP HANA Rapid Deployment solutions

The SAP Quick Sizer is accessible online at:

`http://service.sap.com/quicksizer`[5]

[5] SAP S-user ID required

3.6 SAP HANA software licensing

As described in 3.2, "SAP HANA delivery model" on page 17, SAP HANA has an appliance-like delivery model. However, while the hardware partners deliver the infrastructure, including operating system and middleware, the license for the SAP HANA software has to be obtained directly from SAP. Figure 3-8 shows an overview over the licensing structure.

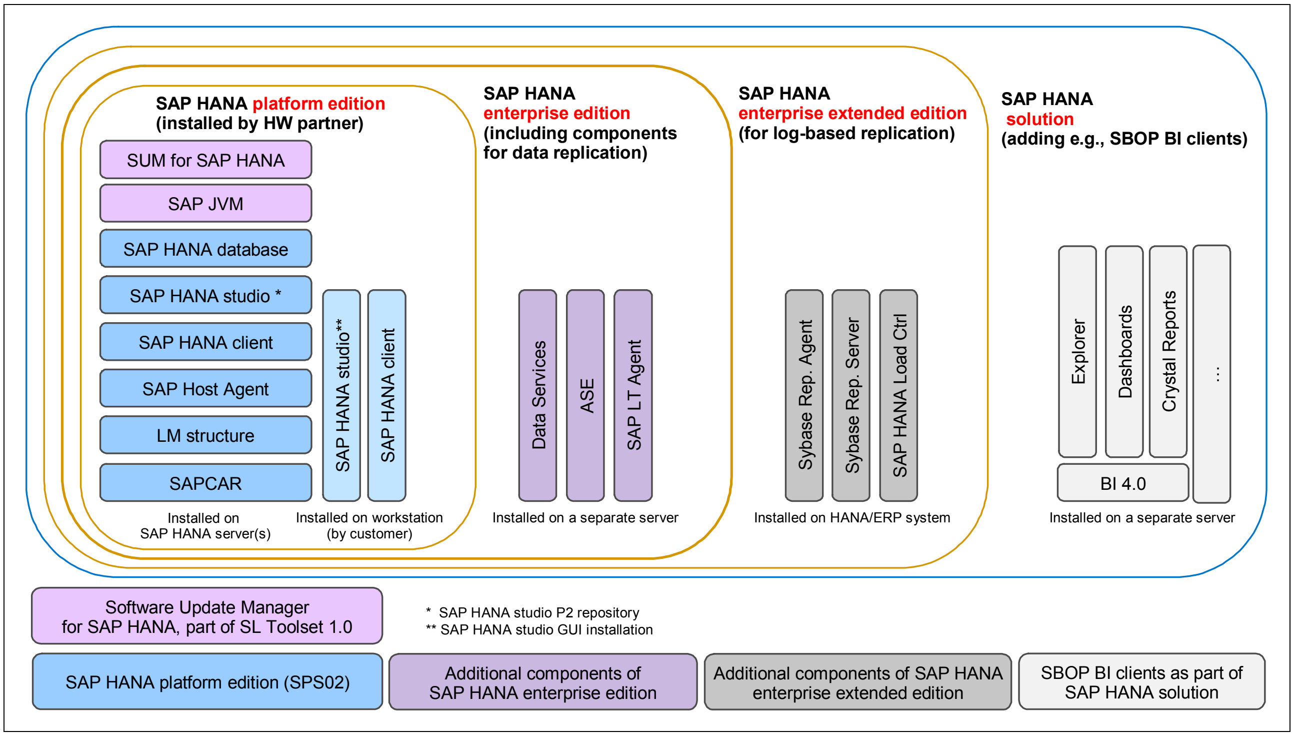


Figure 3-8 SAP HANA editions

The SAP HANA software is available in these editions:

- SAP HANA platform edition

  This is the basic edition containing the software stack needed to use SAP HANA as a database, including the SAP HANA database, SAP HANA Studio for data modeling and administration, the SAP HANA clients, and software infrastructure components. The software stack comes with the hardware provided by the hardware partners, whereas the license has to be obtained from SAP.

- SAP HANA enterprise edition

  The SAP HANA enterprise edition extends the SAP HANA platform edition with the software licenses needed for SAP LT replication or ETL-based replication.

- SAP HANA extended enterprise edition

  SAP HANA extended enterprise edition extends the SAP HANA platform edition with the software licenses needed for log-based replication with the Sybase Replication server.

As Figure 3-8 suggests, additional licenses for SAP BusinesObjects BI tools might be needed to get a complete SAP HANA-based solution.

The SAP HANA licenses are based on the amount of main memory for SAP HANA. The smallest licensable memory size is 64 GB, increasing in steps of 64 GB. The hardware might provide up to double the amount of main memory than licensed. Table 3-2 illustrates this.

Table 3-2 Licensable memory per T-shirt size

T-shirt size	**Server main memory**	**Licensable memory[a]**
XS	128 GB	64 - 128 GB
S	256 GB	128 - 256 GB
M	512 GB	256 - 512 GB
L	1024 GB (= 1 TB)	512 - 1024 GB

a. in steps of 64 GB

As you can see from Table 3-2, the licensing model allows you to have a matching T-shirt size for any licensable memory size between 64 GB and 1024 GB.

4

The IBM Systems Solution for SAP HANA

This chapter discusses the IBM Systems Solution for SAP HANA. We describe the hardware and software components, scale-up and scale-out approaches, workload-optimized models, and monitoring and support processes.

The following topics are covered:

- 4.1, “IBM eX5 Systems” on page 32
- 4.2, “General Parallel File System” on page 42
- 4.3, “Custom server models for SAP HANA” on page 44
- 4.4, “Scale-out solution for SAP HANA” on page 47
- 4.5, “Monitoring SAP HANA” on page 53
- 4.6, “Support process” on page 54

4.1 IBM eX5 Systems

IBM decided to base their offering for SAP HANA on their high-performance, scalable IBM eX5 family of servers. These servers represent the IBM high-end Intel-based enterprise servers. IBM eX5 systems, all based on the eX5 Architecture, are the HX5 blade server, the x3690 X5, the x3850 X5, and the x3950 X5. They have a common set of technical specifications and features:

- The IBM System x3850 X5 is a 4U highly rack-optimized server. The x3850 X5 also forms the basis of the x3950 X5, the new flagship server of the IBM x86 server family. These systems are designed for maximum utilization, reliability, and performance for compute-intensive and memory-intensive workloads such as SAP HANA.
- The IBM System x3690 X5 is a 2U rack-optimized server. This machine brings the eX5 features and performance to the mid tier. It is an ideal match for the smaller, two-CPU configurations for SAP HANA.
- The IBM BladeCenter HX5 is a single wide (30 mm) blade server that follows the same design as all previous IBM blades. The HX5 brings unprecedented levels of capacity to high-density environments.

When compared with other machines in the System x portfolio, these systems represent the upper end of the spectrum and are suited for the most demanding x86 tasks.

For SAP HANA, the x3690 X5 and the x3950 X5 are used, which is why we feature only these systems in this paper.

> **Note:** For the latest information about the eX5 portfolio, see the IBM Redpaper publication *IBM eX5 Portfolio Overview: IBM System x3850 X5, x3950 X5, x3690 X5, and BladeCenter HX5*, REDP-4650, for further eX5 family members and capabilities. This paper is available at the following web page:
>
> http://www.redbooks.ibm.com/abstracts/redp4650.html

4.1.1 IBM System x3850 X5 and x3950 X5

The IBM System x3850 X5 (Figure 4-1) offers improved performance and enhanced features, including MAX5 memory expansion and workload-optimized x3950 X models to maximize memory, minimize costs, and simplify deployment.

Figure 4-1 IBM System x3850 X5 and x3950 X5

The x3850 X5 and the workload-optimized x3950 X5 are the logical successors to the x3850 M2 and x3950 M2, featuring the IBM eX4 chipset. Compared with previous generation servers, the x3850 X5 offers:

- High memory capacity

  Up to 64 DIMMS standard and 96 DIMMs with the MAX5 memory expansion per 4-socket server
- Intel Xeon processor E7 family

  Exceptional scalable performance with advanced reliability for your most data-demanding applications
- Extended SAS capacity with eight HDDs and 900 GB 2.5" SAS drives or 1.6 TB of hot-swappable RAID 5 with eXFlash technology
- Standard dual-port Emulex 10 GB Virtual Fabric adapter
- Ten-core, 8-core, and 6-core processor options with up to 2.4 GHz (10-core), 2.13 GHz (8-core), and 1.86 GHz (6-core) speeds with up to 30 MB L3 cache
- Scalable to a two-node system with eight processor sockets and 128 DIMM sockets
- Seven PCIe x8 high-performance I/O expansion slots to support hot-swap capabilities
- Optional embedded hypervisor

The x3850 X5 and x3950 X5 both scale to four processors and 2 Terabytes (TB) of RAM. With the MAX5 attached, the system can scale to four processors and 3 TB of RAM. Two x3850 X5 servers can be connected together for a single system image with eight processors and 4 TB of RAM.

With their massive memory capacity and computing power, the IBM System x3850 X5 and x3950 X5 rack-mount servers are the ideal platform for high memory demanding, high workload applications such as SAP HANA.

4.1.2 IBM System x3690 X5

The IBM System x3690 X5 (Figure 4-2) is a 2U rack-optimized server, bringing new features and performance to the mid tier.

Figure 4-2 IBM System x3690 X5

This machine is a two-socket, scalable system that offers up to four times the memory capacity of current two-socket servers. It supports the following specifications:

- Up to two sockets for Intel Xeon E7 processors. Depending on the processor model, processors have six, eight, or ten cores.
- Scalable from 32 to 64 DIMMs sockets with the addition of a MAX5 memory expansion unit.
- Advanced networking capabilities with a Broadcom 5709 dual Gb Ethernet controller standard in all models and an Emulex 10 Gb dual-port Ethernet adapter standard on some models, optional on all others.

- Up to 16 hot-swap 2.5-inch SAS HDDs, up to 16 TB of maximum internal storage with RAID 0, 1, or 10 to maximize throughput and ease installation. RAID 5 is optional. The system comes standard with one HDD backplane that can hold four drives. A second and third backplane are optional for an additional 12 drives.
- New eXFlash high-IOPS solid-state storage technology.
- Five PCIe 2.0 slots.
- Integrated Management Module (IMM) for enhanced systems management capabilities.

The x3690 X5 features the IBM eXFlash internal storage using solid state drives to maximize the number of I/O operations per second (IOPS). All configurations for SAP HANA based on x3690 X5 use eXFlash internal storage for high IOPS log storage or for both data and log storage.

The x3690 X5 is an excellent choice for a memory-demanding and performance-demanding business application such as SAP HANA. It provides maximum performance and memory in a dense 2U package.

4.1.3 Intel Xeon processor E7 family

The IBM eX5 portfolio of servers uses CPUs from the Intel Xeon processor E7 family to maximize performance. These processors are the latest in a long line of high-performance processors.

The Intel Xeon processor E7 family CPUs are the latest Intel scalable processors and can be used to scale up to four or more processors. When used in the IBM System x3850 X5 or x3950 X5, these servers can scale up to eight processors.

The Intel Xeon E7 processors have a lot of features that are relevant for the SAP HANA workload. We cover some of these features in the following sections. For more in-depth information about the benefits of the Intel Xeon processor E7 family for SAP HANA see the Intel white paper “Analyzing Business as it Happens,” April 2011, available for download at:

http://www.Intel.com/en_us/ssets/pdf/whitepaper/mc_sap_wp.pdf

Instruction set extensions

SAP HANA makes use of several instruction set extensions of the Intel Xeon E7 processors. For example, these extensions allow you to process multiple data items with one instruction. SAP HANA uses these instructions to speed up compression and decompression of in-memory data and to improve search performance.

Intel Hyper-Threading Technology

Intel Hyper-Threading Technology enables a single physical processor to execute two separate code streams (threads) concurrently on a single processor core. To the operating system, a processor core with Hyper-Threading appears as two logical processors, each of which has its own architectural state. Hyper-Threading Technology is designed to improve server performance by exploiting the multi-threading capability of operating systems and server applications. SAP HANA makes extensive use of Hyper-Threading to parallelize processing.

For more information about Hyper-Threading Technology, see the following web page:

http://www.intel.com/technology/platform-technology/hyper-threading/

Intel Turbo Boost Technology 2.0

Intel Turbo Boost Technology dynamically turns off unused processor cores and increases the clock speed of the cores in use. For example, with six cores active, a 2.4 GHz 10-core processor can run the cores at 2.67 GHz. With only four cores active, the same processor can run those cores at 2.8 GHz. When the cores are needed again, they are dynamically turned back on and the processor frequency is adjusted accordingly. When temperature, power, or current exceed factory-configured limits and the processor is running higher than the base operating frequency, the processor automatically reduces the core frequency to reduce temperature, power, and current.

Turbo Boost Technology is available on a per-processor number basis for the eX5 systems. For ACPI-aware operating systems, no changes are required to take advantage of it. Turbo Boost Technology can be engaged with any number of cores enabled and active, resulting in increased performance of both multi-threaded and single-threaded workloads.

For more information about Turbo Boost Technology, see the following web page:

http://www.intel.com/technology/turboboost/

Quick Path Interconnect (QPI)

Earlier versions of the Intel Xeon processor were connected by a parallel bus to a core chipset, which functions as both a memory and I/O controller. The new Intel Xeon E7 processors implemented in IBM eX5 servers include a separate memory controller to each processor. Processor-to-processor communication is carried over shared-clock or coherent quick path interconnect (QPI) links, and I/O is transported over non-coherent QPI links through I/O hubs (Figure 4-3).

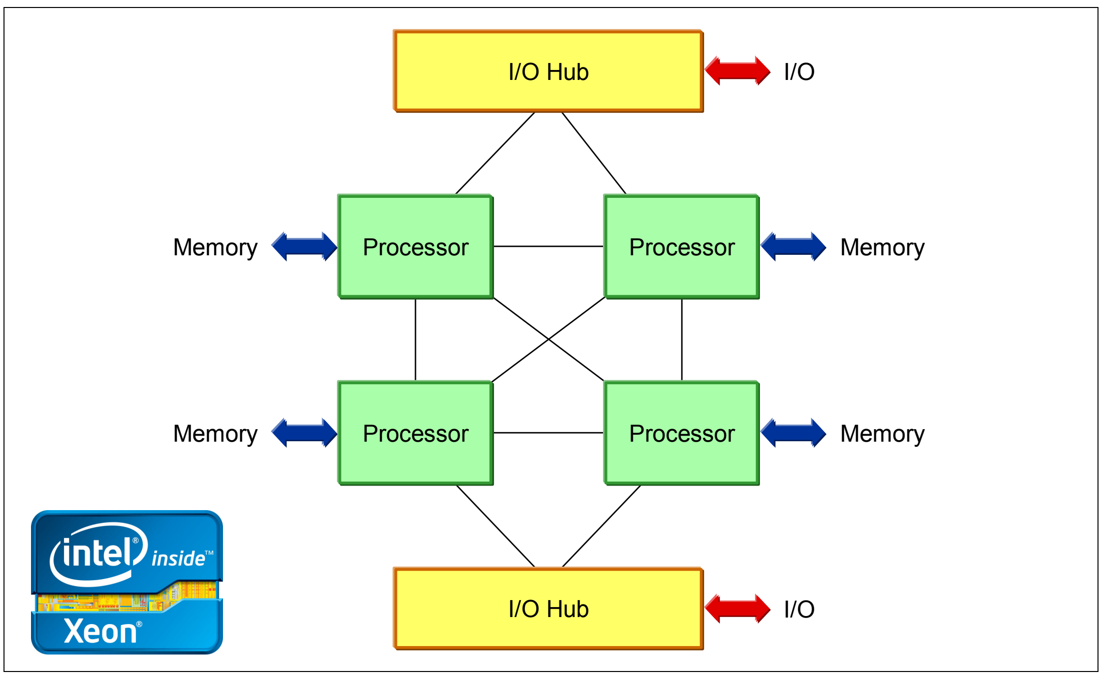


Figure 4-3 Quick path interconnect, as in the eX5 portfolio

In previous designs, the entire range of memory was accessible through the core chipset by each processor. This is called a *shared memory* architecture. This new design creates a non-uniform memory access (NUMA) system in which a portion of the memory is directly connected to the processor where a given thread is running, and the rest must be accessed over a QPI link through another processor. Similarly, I/O can be local to a processor or remote through another processor.

For more information about QPI, see the following web page:

http://www.intel.com/technology/quickpath/

Reliability, availability, and serviceability

Most system errors are handled in hardware by the use of technologies such as error checking and correcting (ECC) memory. The E7 processors have additional reliability, availability, and serviceability (RAS) features due to their architecture:

- Cyclic redundancy checking (CRC) on the QPI links

  The data on the QPI link is checked for errors.

- QPI packet retry

  If a data packet on the QPI link has errors or cannot be read, the receiving processor can request that the sending processor retry sending the packet.

- QPI clock failover

  In the event of a clock failure on a coherent QPI link, the processor on the other end of the link can take over providing the clock. This is not required on the QPI links from processors to I/O hubs, as these links are asynchronous.

- SMI packet retry

  If a memory packet has errors or cannot be read, the processor can request that the packet be resent from the memory buffer.

- Scalable memory interconnect (SMI) retry

  If there is an error on an SMI link, or a memory transfer fails, the command can be retried.

- SMI lane failover

  When an SMI link exceeds the preset error threshold, it is disabled, and memory transfers are routed through the other SMI link to the memory buffer.

All these features help prevent data from being corrupted or lost in memory. This is especially important with an application like SAP HANA, because any failure in the area of memory or inter-CPU communication leads to an outage of the application or even of the complete system. With huge amounts of data loaded into main memory, even a restart of only the application means considerable time required to return to operation.

Machine check architecture

The Intel Xeon processor E7 family also features the machine check architecture (MCA), a RAS feature that enables the handling of system errors that otherwise require the operating system to be halted. For example, if a dead or corrupt memory location is discovered, but it cannot be recovered at the memory subsystem level, and provided that it is not in use by the system or an application, an error can be logged but the operation of the server can continue. If it is in use by a process, the application to which the process belongs can be aborted or informed about the situation.

Implementation of the MCA requires hardware support, firmware support (such as that found in the unified extensible firmware interface (UEFI)), and operating system support. Microsoft, SUSE, Red Hat, VMware, and other operating system vendors have included support for the Intel MCA on the Intel Xeon processors in their latest operating system versions.

SAP HANA is the first application that leverages the MCA to handle system errors to prevent the application from being terminated in case of a system error. Figure 4-4 shows how SAP HANA leverages the Machine Check Architecture.

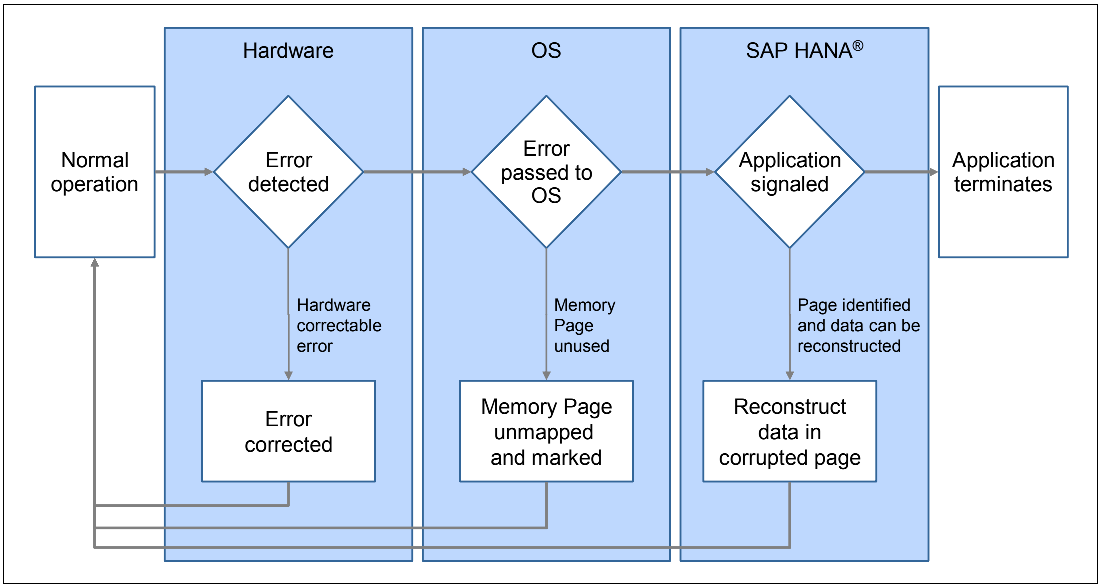


Figure 4-4 Intel Machine Check Architecture (MCA) with SAP HANA

If a memory error is encountered that cannot be corrected by the hardware, the processor sends an MCA recovery signal to the operating system. An operating system supporting MCA, such as SUSE Linux Enterprise Server used in the SAP HANA appliance, now determines whether the affected memory page is in use by an application. If unused, it unmaps the memory page and marks it as bad. If the page is used by an application, traditionally the OS would have to hold that application, or in the worst case stop all processing and halt the system. With SAP HANA being MCA-aware, the operating system can signal the error situation to SAP HANA, giving it the chance to try to repair the effects of the memory error.

Using the knowledge of its internal data structures, SAP HANA can decide what course of action to take. If the corrupted memory space is occupied by one of the SAP in-memory tables, SAP HANA reloads the associated tables. In addition, it analyzes the failure and checks whether it affects other stored or committed data, in which case it uses savepoints and database logs to reconstruct the committed data in a new, unaffected memory location.

With the support of MCA, SAP HANA can takes appropriate action at the level of its own data structures to ensure a smooth return to normal operation, avoiding a time-consuming restart or loss of information.

I/O hubs

The connection to I/O devices (such as keyboard, mouse, and USB) and to I/O adapters (such as hard disk drive controllers, Ethernet network interfaces, and Fibre Channel host bus adapters) is handled by I/O hubs, which then connect to the processors through QPI links. Figure 4-3 on page 35 shows the I/O Hub connectivity. Connections to the I/O devices are fault tolerant, as data can be routed over either of the two QPI links to each I/O hub.

For optimal system performance in the four processor systems (with two I/O hubs), balance high-throughput adapters across the I/O hubs. The configurations used for SAP HANA contain several components that require high throughput I/O:

- Dual-port 10 Gb Ethernet adapters
- ServeRAID controllers to connect the SAS drives
- High IOPS PCIe Adapters

To ensure optimal performance, the placement of these components in the PCIe slots has been optimized according to the I/O architecture outlined above.

4.1.4 Memory

For an in-memory appliance like SAP HANA, a system's main memory, its capacity, and its performance play an important role. The Intel Xeon processor E7 family has a memory architecture that is well suited to the requirements of such an appliance.

The E7 processors have two SMIs. Therefore, memory needs to be installed in matched pairs. For better performance, or for systems connected together, memory has to be installed in sets of four. The memory used in the eX5 systems is DDR3 SDRAM registered DIMMs. All of the memory runs at 1066 MHz or less, depending on the processor.

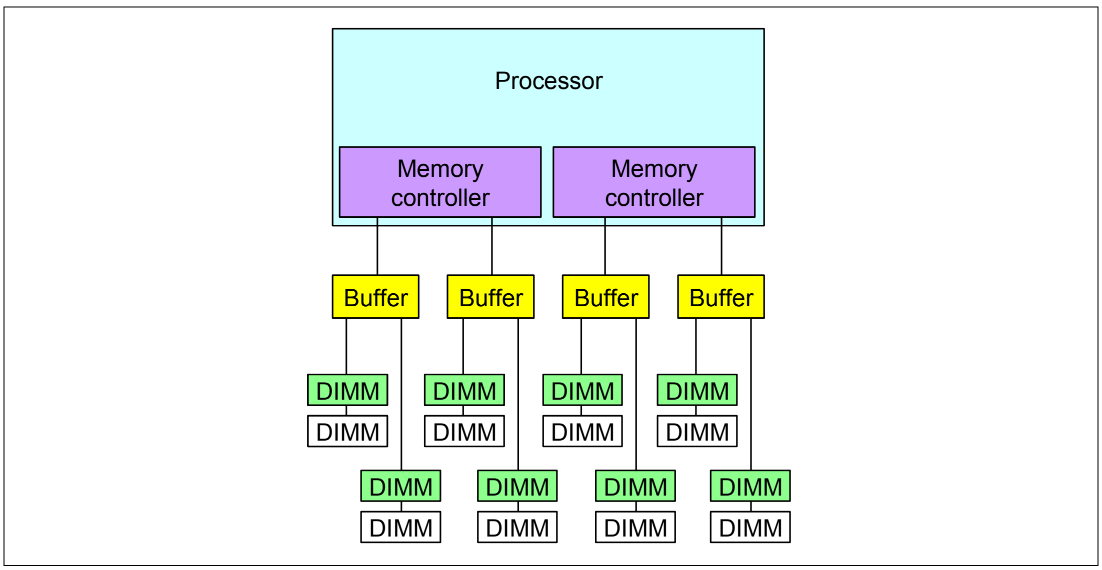


Figure 4-5 Memory architecture with Intel Xeon processor E7 family

Memory DIMM placement

The eX5 servers support a variety of ways to install memory DIMMs. It is important to understand that because of the layout of the SMI links, memory buffers, and memory channels, you must install the DIMMs in the correct locations to maximize performance.

Figure 4-6 shows eight possible memory configurations for the two memory cards and 16 DIMMs connected to each processor socket in an x3850 X5. Similar configurations apply to the x3690 X5 and HX5. Each configuration has a relative performance score. The following key information from this chart is important:

- The best performance is achieved by populating all memory DIMMs in the server (configuration 1 in Figure 4-6).
- Populating only one memory card per socket can result in approximately a 50% performance degradation. (Compare configuration 1 with 5.)
- Memory performance is better if you install DIMMs on all memory channels than if you leave any memory channels empty. (Compare configuration 2 with 3.)
- Two DIMMs per channel result in better performance than one DIMM per channel. (Compare configuration 1 with 2, and compare 5 with 6.)

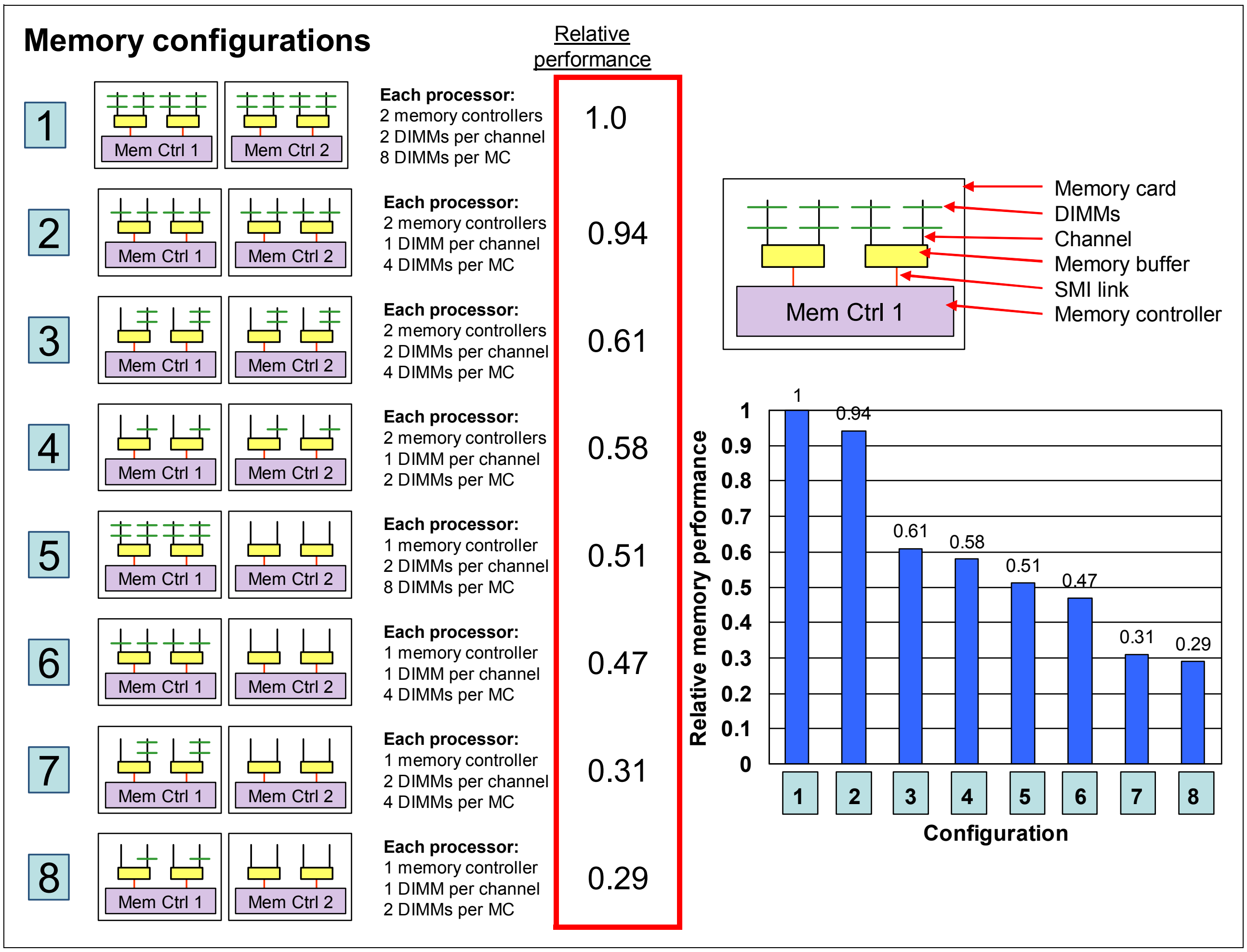


Figure 4-6 Relative memory performance based on DIMM placement (one processor and two memory cards shown)

Nonuniform memory architecture

Nonuniform memory architecture (NUMA) is an important consideration when configuring memory, because a processor can access its own local memory faster than non-local memory. The configurations used for SAP HANA do not use all available DIMM sockets. For configurations like these, another principle to consider when configuring memory is that of *balance*. A balanced configuration has all of the memory cards configured with the same *amount* of memory. This principle helps to keep remote memory access to a minimum.

A server with a NUMA, such as the servers in the eX5 family, has local and remote memory. For a given thread running in a processor core, *local memory* refers to the DIMMs that are directly connected to that particular processor. *Remote memory* refers to the DIMMs that are not connected to the processor where the thread is running currently. Remote memory is attached to another processor in the system and must be accessed through a QPI link (Figure 4-3 on page 35). However, using remote memory adds latency. The more such latencies add up in a server, the more performance can degrade. Starting with a memory configuration where each CPU has the same local RAM capacity is a logical step toward keeping remote memory accesses to a minimum.

In a NUMA system, each processor has fast, direct access to its own memory modules, reducing the latency that arises due to bus-bandwidth contention. SAP HANA is NUMA-aware, and thus benefits from this direct connection.

Hemisphere mode

Hemisphere mode is an important performance optimization of the Intel Xeon processor E7, 6500, and 7500 product families. Hemisphere mode is automatically enabled by the system if the memory configuration allows it. This mode interleaves memory requests between the two memory controllers within each processor, enabling reduced latency and increased throughput. It also allows the processor to optimize its internal buffers to maximize memory throughput.

Hemisphere mode is enabled only when the memory configuration behind each memory controller on a processor is identical. In addition, because eight DIMMs per processor are required for using all memory channels, eight DIMMs per processor need to be installed at a time for optimized memory performance.

4.1.5 Flash technology storage

As discussed in 2.1.2, “Data persistence” on page 9, storage technology providing high IOPS capabilities with low latency is a key component of the infrastructure for SAP HANA. The IBM eX5 systems used for the IBM Systems Solution for SAP HANA feature two kinds of flash technology storage devices:

- eXFlash, as used in the IBM System x3690 X5-based configuration
- High IOPS adapters, as used in the IBM System x3950 X5-based configurations

The following sections provide more information about these options.

eXFlash

IBM eXFlash is the name given to the eight 1.8-inch solid state drives (SSDs), the backplanes, SSD hot swap carriers, and indicator lights that are available for the x3850 X5/x3950 X5 and x3690 X5. Each eXFlash can be put in place of four SAS or SATA disks. The eXFlash units connect to the same types of ServeRAID disk controllers as the SAS/SATA disks. Figure 4-7 shows an eXFlash unit, with the status light assembly on the left side.

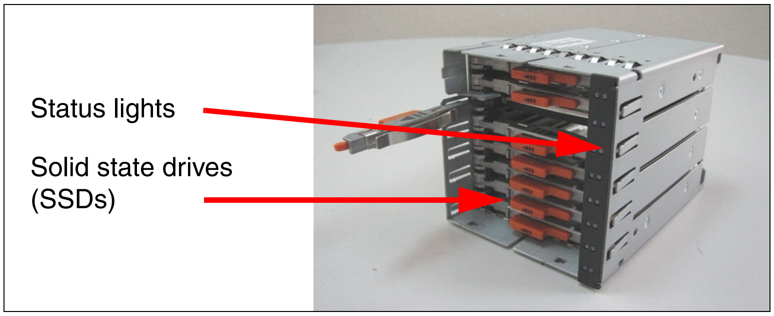


Figure 4-7 IBM eXFlash unit

In addition to using less power than rotating magnetic media, the SSDs are more reliable and can service many more I/O operations per second (IOPS). These attributes make them suited to I/O-intensive applications, such as such as transaction processing, logging, backup and recovery, and business intelligence. Built on enterprise-grade MLC NAND flash memory, the SSD drives used in eXFlash deliver up to 30,000 IOPS per single drive. Combined into an eXFlash unit, these drives can deliver up to 240,000 IOPS and up to 2 GBps of sustained read throughput per eXFlash unit.

In addition to its superior performance, eXFlash offers superior uptime with three times the reliability of mechanical disk drives. SSDs have no moving parts to fail. Each drive has its own backup power circuitry, error correction, data protection, and thermal monitoring circuitry. They use Enterprise Wear-Leveling to extend their use even longer.

A single eXFlash unit accommodates up to eight hot-swap SSDs and can be connected to up to two performance-optimized controllers. The x3690 X5-based models for SAP HANA enable RAID protection for the SSD drives by using two ServeRAID M5015 controllers with the ServeRAID M5000 Performance Accelerator Key for the eXFlash units.

High IOPS adapter

The IBM High IOPS SSD PCIe Adapters provide a new generation of ultra-high-performance storage based on solid state device technology for System x and BladeCenter. These adapters are alternatives to disk drives and are available in several sizes, from 160 GB to 1.2 TB. Designed for high-performance servers and computing appliances, these adapters deliver throughput of up to 300,000 I/O operations per second (IOPS), while providing the added benefits of lower power, cooling, and management overhead and a smaller storage footprint. Based on standard PCIe architecture coupled with silicon-based NAND clustering storage technology, the High IOPS adapters are optimized for System x racks and can be deployed in blades via the PCIe expansion units. They are available in storage capacities ranging from 160 GB to 1.28 TB.

These adapters use NAND flash memory as the basic building block of solid-state storage and contain no moving parts. Thus, they are less sensitive to issues associated with vibration, noise, and mechanical failure. They function as a PCIe storage and controller device (PCI Express 1.1 x4 host interface), and after the appropriate drivers are loaded, the host operating system sees them as block devices. Therefore, these adapters cannot be used as bootable devices.

The IBM High IOPS SSD PCIe Adapters combine high IOPS performance with low latency. As an example, with 4 KB block random reads these adapters can deliver 97,014 IOPS, compared with 420 IOPS for a 15 K RPM 146 GB disk drive. The latency is about 50 microseconds, which is one hundredth of the latency of a 15 K RPM 146 GB disk drive (about 5 ms or 5000 microseconds).

Reliability features include advanced wear-leveling, ECC, and N+1 chip-level redundancy, providing unparalleled reliability and efficiency. Advanced bad-block management algorithms enable taking blocks out of service when their failure rate becomes unacceptable. These reliability features provide a predictable lifetime and up to 25 years of data retention.

The x3950 X5-based models of the IBM Systems Solution for SAP HANA come with IBM High IOPS adapters, either with 320 GB (7143-H1x) or 640 GB storage capacity. Figure 4-8 shows the adapters.

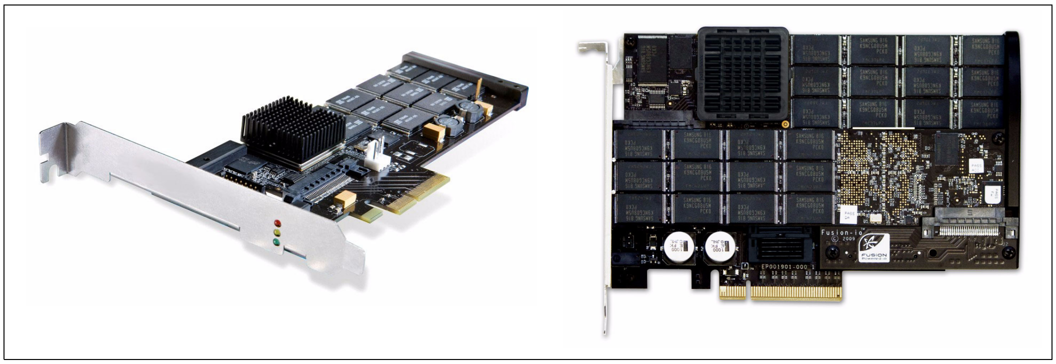


Figure 4-8 BM 320 GB High IOPS SSD PCIe Adapter and 640 GB High IOPS MLC Duo Adapter for IBM System x (left to right)

4.2 General Parallel File System

The IBM General Parallel File System (GPFS) is a key component of the IBM Systems Solution for SAP HANA. It is a high-performance shared-disk file management solution that can provide faster, more reliable access to a common set of file data. It enables a view of distributed data with a single global namespace.

GPFS leverages its cluster architecture to provide quicker access to your file data. File data is automatically spread across multiple storage devices, providing optimal use of your available storage to deliver high performance.

GPFS is designed for high-performance parallel workloads. Data and metadata flow from all the nodes to all the disks in parallel under control of a distributed lock manager. It has a very flexible cluster architecture that enables the design of a data storage solution that not only meets current needs but that can also quickly be adapted to new requirements or technologies. GPFS configurations include direct-attached storage, network block input and output (I/O), or a combination of the two, and multi-site operations with synchronous data mirroring.

GPFS can intelligently prefetch data into its buffer pool, issuing I/O requests in parallel to as many disks as necessary to achieve the peak bandwidth of the underlying storage-hardware infrastructure. GPFS recognizes multiple I/O patterns, including sequential, reverse sequential, and various forms of striped access patterns. In addition, for high-bandwidth environments, GPFS can read or write large blocks of data in a single operation, minimizing the overhead of I/O operations.

Expanding beyond a storage area network (SAN) or locally attached storage, a single GPFS file system can be accessed by nodes using a TCP/IP or InfiniBand connection. Using this block-based network data access, GPFS can outperform network-based sharing technologies like NFS and even local file systems such as the EXT3 journaling file system for Linux or Journaled File System. Network block I/O (also called network shared disk (NSD)) is a software layer that transparently forwards block I/O requests from a GPFS client application node to an NSD server node to perform the disk I/O operation, and then passes the data back to the client. Using a network block I/O configuration can be more cost effective than a full-access SAN.

Storage pools enable you to transparently manage multiple tiers of storage based on performance or reliability. You can use storage pools to transparently provide the appropriate type of storage to multiple applications or different portions of a single application within the same directory. For example, GPFS can be configured to use low-latency disks for index operations and high-capacity disks for data operations of a relational database. You can make these configurations even if all database files are created in the same directory.

For optimal reliability, GPFS can be configured to eliminate single points of failure. The file system can be configured to remain available automatically in the event of a disk or server failure. A GPFS file is designed to transparently fail over token (lock) operations and other GPFS cluster services, which can be distributed throughout the entire cluster to eliminate the need for dedicated metadata servers. GPFS can be configured to automatically recover from node, storage, and other infrastructure failures.

GPFS provides this functionality by supporting these:

- Data replication to increase availability in the event of a storage media failure
- Multiple paths to the data in the event of a communications or server failure
- File system activity logging, enabling consistent fast recovery after system failures

In addition, GPFS supports snapshots to provide a space-efficient image of a file system at a specified time, which allows online backup and can help protect against user error.

GPFS offers time-tested reliability and has been installed on thousands of nodes across industries, from weather research to multimedia, retail, financial industry analytics, and web service providers. GPFS also is the basis of many cloud storage offerings.

The IBM Systems solution for SAP HANA benefits in several ways from the features of GPFS:

- GPFS provides a stable, industry-proven, cluster-capable file system for SAP HANA.
- GPFS adds extra performance to the storage devices by striping data across devices.
- GPFS enables the IBM Systems solution for SAP HANA to grow beyond the capabilities of a single system, into a scale-out solution, without introducing the need for external storage.
- GPFS adds high-availability features to the solution.

This makes GPFS the ideal file system for the IBM Systems solution for SAP HANA.

4.3 Custom server models for SAP HANA

Following the appliance-like delivery model for SAP HANA, IBM created several custom server models for SAP HANA. These workload-optimized models have been designed to match and exceed the performance requirements and the functional requirements as specified by SAP. With a small set of IBM System x workload-optimized models for SAP HANA, all sizes of SAP HANA solutions can be built, from the smallest to large installations.

4.3.1 IBM System x workload-optimized models for SAP HANA

In the first half of 2011, IBM announced a full range of IBM System x workload-optimized models for SAP HANA, covering all SAP HANA T-shirt sizes with the newest generation technology. Because there is no direct relationship between the workload-optimized models and the SAP HANA T-shirt sizes, we refer to these models as building blocks. In some cases there are several building blocks available for one T-shirt size. In some, two-building blocks have to be combined to build a specific T-shirt size. Table 4-1 shows all currently available building blocks and their features.

Table 4-1 IBM System x workload-optimized models for SAP HANA

Building block	Server (MTM)	CPUs	Main memory	Log storage	Data storage	Software preload
XS	x3690 X5 (7147-H1x[a])	2x Intel Xeon E7-2870	128 GB DDR3 (8x 16 GB)	8x 50 GB 1.8” MLC SSD	8x 300 GB 10 K SAS HDD	Yes
S	x3690 X5 (7147-H2x)	2x Intel Xeon E7-2870	256 GB DDR3 (16x 16 GB)	8x 50 GB 1.8” MLC SSD	8x 300 GB 10 K SAS HDD	Yes
SSD	x3690 X5 (7147-H3x)	2x Intel Xeon E7-2870	256 GB DDR3 (16x 16 GB)	10x 200 GB 1.8” MLC SSD (combined log and data)		Yes
S+	x3950 X5 (7143-H1x)	2x Intel Xeon E7-8870	256 GB DDR3 (16x 16 GB)	320 GB High IOPS adapter	8x 600 GB 10 K SAS HDD	Yes
M	x3950 X5 (7143-H2x)	4x Intel Xeon E7-8870	512 GB DDR3 (32x 16 GB)	640 GB High IOPS adapter	8x 600 GB 10 K SAS HDD	Yes
L Option	x3950 X5 (7143-H3x)	4x Intel Xeon E7-8870	512 GB DDR3 (32x 16 GB)	640 GB High IOPS adapter	8x 600 GB 10 K SAS HDD	No

a. x x = Country-specific letter (for example, EMEA MTM is 86774SG, and the US MTM is 86774SU). Contact your IBM representative for regional part numbers.

In addition to the models listed in Table 4-1, there are models specific to a geographic region:

- Models 7147-H7x, -H8x, and -H9x are for Canada only and are the same configurations as H1x, H2x, and H3x, respectively.
- Models 7143-H4x and -H5x are for Canada only and are the same configuration as H1x and H2x, respectively.

All models (except for 7143-H3x) come with a preload comprising SUSE Linux Enterprise Server for SAP Applications (SLES for SAP), IBM GPFS, and the SAP HANA software stack. Licenses and maintenance fees (for three years) for SLES for SAP and GPFS are included. The licenses for the SAP software components have to be acquired separately from SAP.

The L-Option building block (7143-H3x) is intended as an extension to an M building block. When building an L-Size SAP HANA system, one M building block has to be combined with an L-Option building block, leveraging eX5 scalability. Both systems will than act as one single eight-socket, 1 TB server. Therefore, the L-Option building block does not require a software preload.

The building blocks are configured to match the SAP HANA sizing requirements. The main memory sizes match the number of CPUs, to give the correct balance between processing power and data volume. Also, the storage devices in the systems provide the storage capacity required to match the amount of main memory.

All systems come with storage for both the data volume and the log volume (Figure 4-9). Savepoints are stored on a RAID protected array of 10 K SAS hard drives, optimized for data throughput. The SAP HANA database logs are stored on flash technology storage devices:

- RAID-protected, hot swap eXFlash SSD drives on the models based on IBM System x3690 X5
- Flash-based High IOPS PCIe adapters for the models based on IBM System x3950 X5

These flash technology storage devices are optimized for high IOPS performance and low latency to provide the SAP HANA database with a log storage that allows the highest possible performance. Because a transaction in the SAP HANA database can only return after the corresponding log entry has been written to the log storage, high IOPS performance and low latency are key to database performance.

The SSD building block, based on the IBM System x3690 X5, comes with combined data and log storage on an array of RAID-protected, hot-swap eXFlash SSD drives. Optimized for throughput, high IOPS performance, and low latency, this building block gives extra flexibility when dealing with large amounts of log data, savepoint data, or backup data.

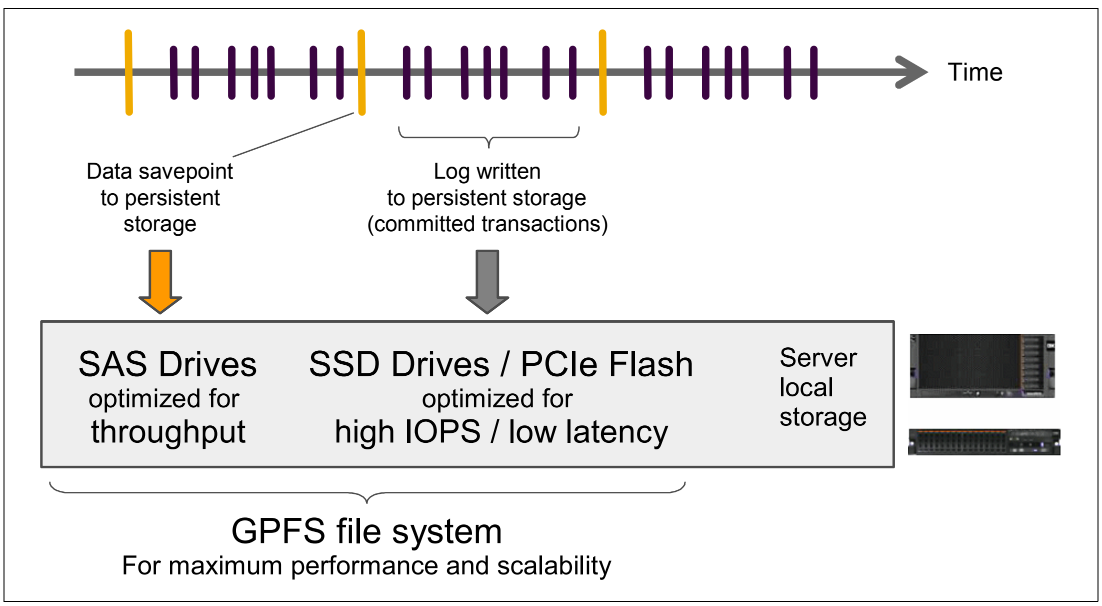


Figure 4-9 SAP HANA data persistency with the internal storage of the workload-optimized systems

4.3.2 SAP HANA T-shirt sizes

This section provides information about how the SAP HANA T-shirt sizes, as described in 3.5.1, "The concept of T-shirt sizes for SAP HANA" on page 26, can be realized using the IBM System x workload-optimized models for SAP HANA[1]:

- For a T-shirt size XS 128 GB SAP HANA system, building block XS (7147-H1x) is the correct choice. This x3690 X5-based building block is the entry-level model of the line of IBM System x workload-optimized systems for SAP HANA.
- A T-shirt size S 256 GB can either be realized with an S building block (7147-H2x) with separate storage for data (SAS drives) and logs (SSD drives), or an SSD building block (7147-H3x) with combined data and log storage on eXFlash SSD drives. Both are based on IBM System x3690 X5.
- For a T-shirt size S 256 GB with upgradability to M (that is, a T-shirt size S+), the S+ building block (7143-H1x) is the perfect choice. Unlike the S and SSD building blocks, it is based on the IBM System x3950 X5 4-socket system to ensure upgradability.
- A T-shirt size M 512 GB can be realized with the M building block (7143-H2x). As it can be upgraded to a T-shirt size L using the L-Option building block, it is also the perfect fit if a T-shirt size M+ is required.
- For a T-shirt size L 1 TB, one M building block (7143-H2x) has to be combined with an L-Option building block (7143-H3x), connected together to form a single server, using eX5 scaling technology.

Table 4-2 gives an overview.

Table 4-2 SAP HANA T-shirt sizes and their relation to the IBM custom models for SAP HANA

SAP T-shirt size	XS	S	S+	M and M+	L
Compressed data in memory	64 GB	128 GB	128 GB	256 GB	512 GB
Server main memory	128 GB	256 GB	256 GB	512 GB	1024 GB
Number of CPUs	2	2	2	4	8
Mapping to building blocks[a]	XS (7147-H1x)	S (7147-H2x) or SSD (7147-H3x)	S+ (7143-H1x)	M (7143-H2x)	Combination of M and L options (7143-H2x + 7143-H3x)

a. For a region-specific equivalent, see 4.3.1, "IBM System x workload-optimized models for SAP HANA" on page 44.

4.3.3 Scale-up

This section talks about upgradability, or scale-up, and shows how IBM custom models for SAP HANA can be upgraded to accommodate the need to grow into bigger T-shirt sizes.

[1] The model numbers given might have be to replaced by a region-specific equivalent by changing the *x* to a region-specific letter identifier. See 4.3.1, "IBM System x workload-optimized models for SAP HANA" on page 44.

To accommodate growth, the IBM Systems Solution for SAP HANA can be scaled in these ways:

- Scale-up approach: Increase the capabilities of a single system by adding more components.
- Scale-out approach: Increase the capabilities of the solution by using multiple systems working together in a cluster.

We discuss the scale-out approach in 4.4, “Scale-out solution for SAP HANA” on page 47.

The building blocks of the IBM Systems Solution for SAP HANA, as described previously in this section, have been designed with extensibility in mind. The following upgrade options exist:

- An XS building block can be upgraded to be an S-size HANA system by adding 128 GB of main memory to the system.
- An S+ building block can be upgraded to be an M-Size HANA system by adding two more CPUs, another 256 GB of main memory, and another 320 GB High IOPS adapter to the system.
- An M building block can be extended with the L option to resemble an L-Size HANA System.
- With the option to upgrade S+ to M, and M to L, IBM can provide an unmatched upgrade path from a T-shirt size S up to a T-shirt size L, without the need to retire a single piece of hardware.

Of course, upgrading server hardware requires system downtime. However, due to GPFS’s capability to add storage capacity to an existing GPFS file system by just adding devices, data residing on the system will remain intact. We nevertheless recommend that you do a backup of the data before changing the system’s configuration.

4.4 Scale-out solution for SAP HANA

Up to now we talked about single-server solutions. Although the scale-up approach gives flexibility to expand the capabilities of a SAP HANA installation, there might be cases where the required data volumes exceed the capabilities of a single server. To meet such requirements, the IBM Systems Solution for SAP HANA supports a scale-out approach (that is, combining a number of systems into a clustered solution, which represents a single SAP HANA installation). A SAP HANA system can span multiple servers, partitioning the data, to be able to hold and process larger amounts of data than a single server would be able to accommodate.

To illustrate this scale-out solution, the following figures show a schematic depiction of such an installation. Figure 4-10 shows a single-node SAP HANA system.

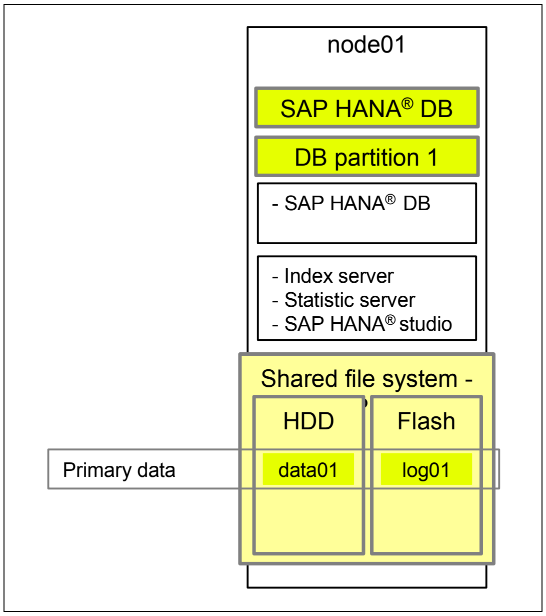


Figure 4-10 Single-node SAP HANA system

This single-node solution has these components:

- The SAP HANA software (SAP HANA database with index server and statistic server)
- The shared file system (GPFS) on the two types of storage
  - The data storage (on SAS disks), here referred to as *HDD*, which holds the savepoints
  - The log storage (on SSD drives or PCIe Flash devices), here referred to as *Flash*, which holds the database logs

This single node represents one single SAP HANA database consisting of one single database partition. Both the savepoints (data01) and the logs (log01) are stored once (that is, they are not replicated), denoted as being *primary data* in Figure 4-10.

4.4.1 Scale-out solution without high-availability capabilities

The first step towards a scale-out solution was to introduce a clustered solution without failover or high-availability (HA) capabilities. IBM was the first hardware partner to validate a scale-out solution for SAP HANA. SAP validated this solution for clusters of up to four nodes, using S or M building blocks in a homogeneous cluster (that is, no mixing of S and M building blocks).

This scale-out solution differs from a single server solution in a number of ways:

- The solution consists of a homogeneous cluster of building blocks, interconnected with two separate 10 Gb Ethernet networks (not shown in Figure 4-11), one for the SAP HANA application and one for the GPFS file system communication.
- The SAP HANA database is split into partitions, forming a single instance of the SAP HANA database.
- Each node of the cluster holds its own savepoints and database logs on the local storage devices of the server.
- The GPFS file system spans all nodes of the cluster, making the data of each node available to all other nodes of the cluster.

Figure 4-11 illustrates this solution, showing a 3-node configuration as an example.

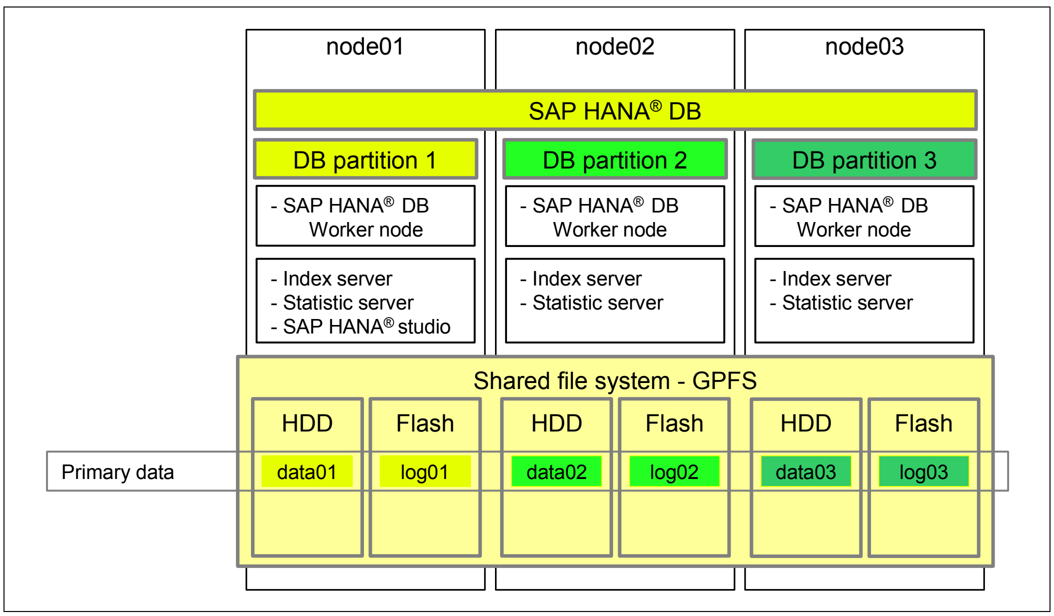


Figure 4-11 A 3-node clustered solution without failover capabilities

To an outside application connecting to the SAP HANA database, this looks like a single instance of SAP HANA. The SAP HANA software distributes the requests internally across the cluster to the individual worker nodes, which process the data and exchange intermediate results, which are then combined and sent back to the requestor. Each node maintains its own set of data, persisting it with savepoints and logging data changes to the database log.

GPFS combines the storage devices of the individual nodes into one big file system, making sure that the SAP HANA software has access to all data regardless of its location in the cluster, while making sure that savepoints and database logs of an individual database partition are stored on the appropriate storage device of the node on which the partition is located. While GPFS provides the SAP HANA software with the functionality of a shared storage system, it ensures maximum performance and minimum latency by using locally attached disks and flash devices. In addition, due to the fact that server-local storage devices are used, the total capacity and performance of the storage within the cluster automatically increases with the addition of nodes, maintaining the same per-node performance characteristics regardless of the size of the cluster. This kind of scalability is not achievable with external storage systems.

The absence of fail-over capabilities represents a major disadvantage of this solution. The cluster acts as a single-node configuration. In case one node becomes unavailable for any reason, the database partition on that node becomes unavailable, and with it the entire SAP HANA database. Loss of the storage of a node means data loss (as with a single-server solution), and the data has to be recovered from a backup. For this very reason this scale-out solution without failover capabilities is an intermediate solution that will go away after all of SAP's hardware partners can provide a solution featuring high-availability capabilities. The IBM version of such a solution is described in the next section.

4.4.2 Scale-out solution with high-availability capabilities

The scale-out solution for SAP HANA with high-availability capabilities enhances the four-node scale-out solution described in the previous section in two major fields:

- Making the SAP HANA application highly available by introducing standby nodes, which can take over from a failed node within the cluster
- Making the data provided through GPFS highly available to the SAP HANA application, even in the event of the loss of one node, including its data on the local storage devices

SAP HANA allows the addition of nodes in the role of a standby node. These nodes run the SAP HANA application, but do not hold any data or take an active part in the processing. In case one of the active nodes fails, a standby node will take over the role of the failed node, including the data (that is, the database partition) of the failed node. This mechanism allows the clustered SAP HANA database to continue operation.

Figure 4-12 illustrates a four-node cluster with the fourth node being a standby node.

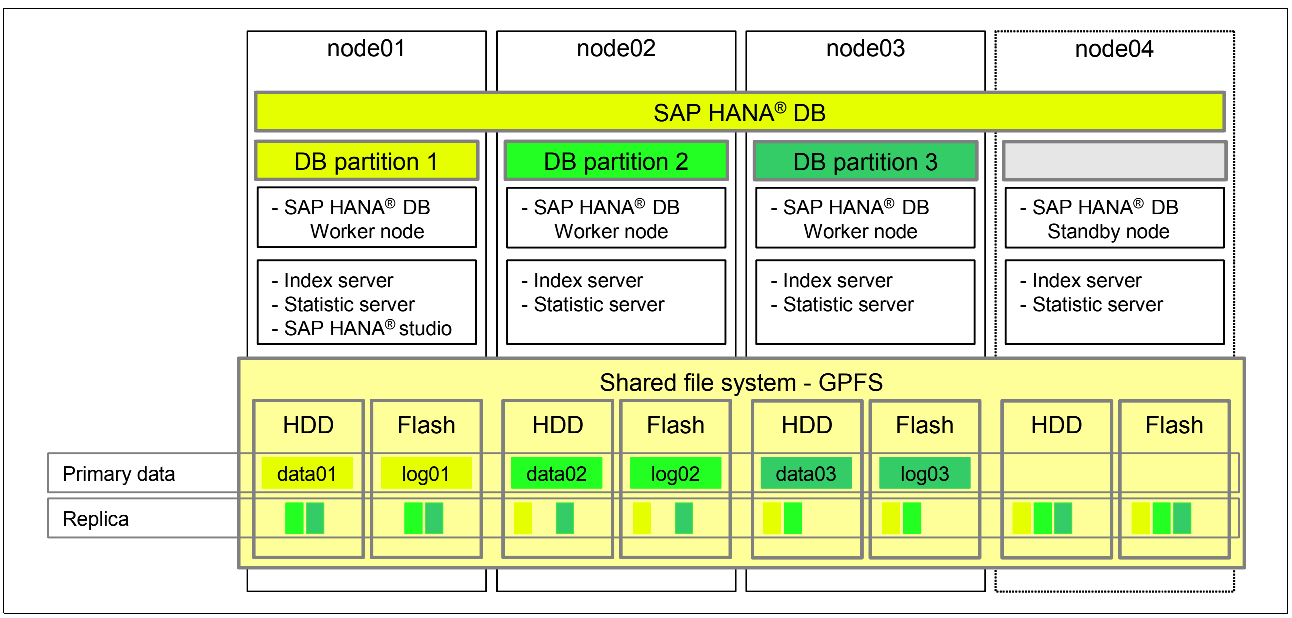


Figure 4-12 A 4-node clustered solution with failover capabilities

To be able to take over the database partition from the failed node, the standby node has to load the savepoints and database logs of the failed node to recover the database partition and resume operation in place of the failed node. This is possible because GPFS provides a global file system across the entire cluster, giving each individual node access to all the data stored on the storage devices managed by GPFS.

In case a node has an unrecoverable hardware error, the storage devices holding the node's data might become unavailable or even destroyed. In contrast to the solution without high-availability capabilities, here the GPFS file system replicates the data of each node to the other nodes, to prevent data loss in case one of the nodes goes down. Replication is done in a striping fashion. That is, every node has a piece of data of all other nodes. In the example illustrated in Figure 4-12 on page 50, the contents of the data storage (that is, the savepoints, here data01) and the log storage (that is, the database logs, here log01) of node01 are replicated to node02, node03, and node04, each holding a part of the data on the matching device (that is, data on HDD, log on flash). The same is true for all nodes carrying data, so that all information is available twice within the GPFS file system, which makes it tolerant to the loss of a single node. The replication is done synchronously. That is, the write operation only finishes when the data has been both written locally and replicated. This ensures consistency of the data at any point in time. Although GPFS replication is done over the network and in a synchronous fashion, this solution still over achieves the performance requirements for validation by SAP.

Using replication, GPFS provides the SAP HANA software with the functionality and fault tolerance of a shared storage system while maintaining its performance characteristics. Again, due to the fact that server-local storage devices are used, the total capacity and performance of the storage within the cluster automatically increases with the addition of nodes, maintaining the same per-node performance characteristics regardless of the size of the cluster. This kind of scalability is not achievable with external storage systems.

Example of a node takeover

To further illustrate the capabilities of this solution, this section provides a node takeover example. In this example we have a 4-node setup, initially configured as illustrated in Figure 4-12 on page 50, with three active nodes and one standby node.

First, node03 experiences a problem and fails unrecoverably. The master node (node01) recognizes this and directs the standby node, node04, to take over from the failed node. Remember that the standby node is running the SAP HANA application and is part of the cluster, but in an inactive role.

To recreate database partition 3 in memory to be able to take over the role of node03 within the cluster, node04 reads the savepoints and database logs of node03 from the GPFS file system, reconstructs the savepoint data in memory, and re-applies the logs so that the partition data in memory is exactly like it was before node03 failed. Node04 is in operation, and the database cluster has recovered.

Figure 4-13 illustrates this scenario.

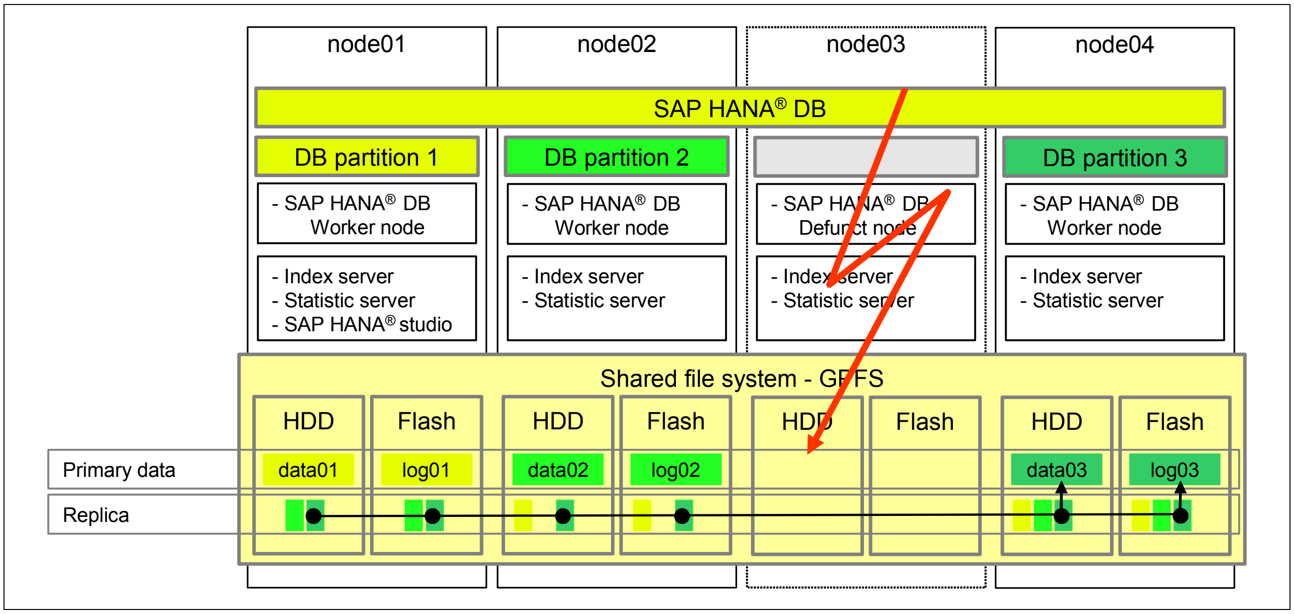


Figure 4-13 Standby node 4 takes over from failed node 3

The data that node04 was reading was the data of node03, which failed, including the local storage devices. For that reason GPFS had to deliver the data to node04 from the replica spread across the cluster via the network. Now when node04 starts writing savepoints and database logs again during the normal course of operations, these are not written over the network, but to the local drives, again with a replica striped across the cluster.

After fixing the cause for the failure of node03, it can be reintegrated into the cluster as the new standby system (Figure 4-14).

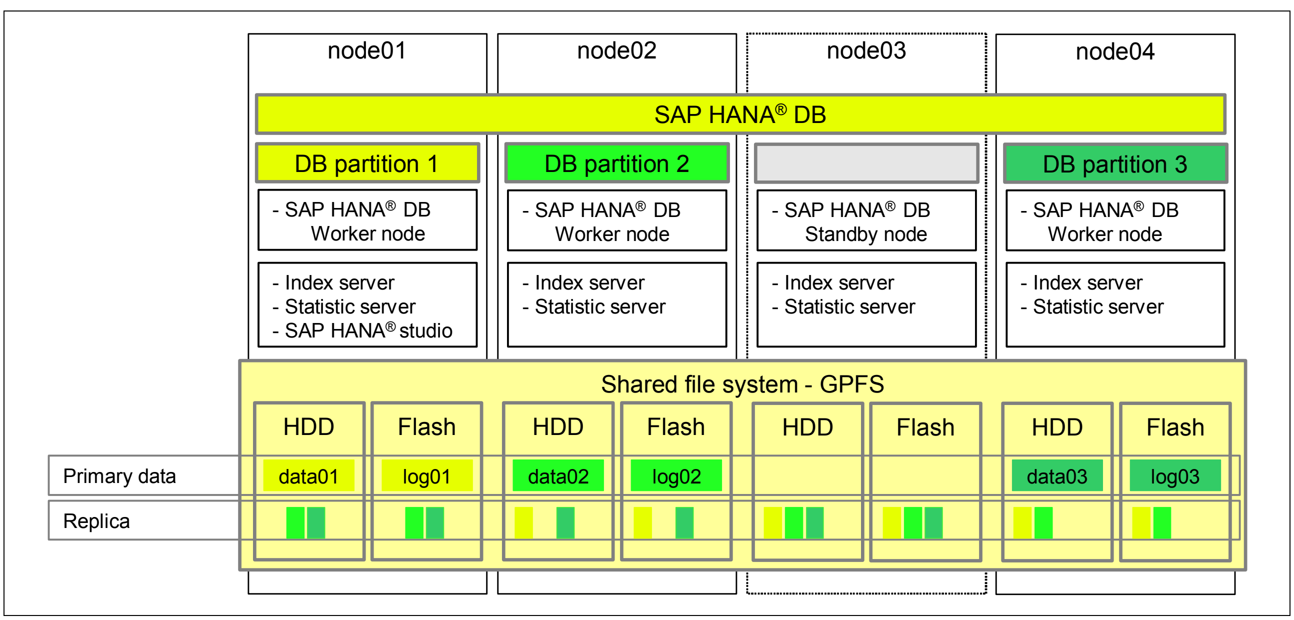


Figure 4-14 Node 3 is reintegrated into the cluster as a standby node

This example illustrates how IBM combines two independently operating high-availability measures (that is, the concept of standby nodes on the SAP HANA application level and the reliability features of GPFS on the infrastructure level), resulting in a highly available and scalable solution.

At the time of writing, clusters of up to 16 nodes using M building blocks are validated by SAP. This means that the cluster has a total main memory of up to 8 TB, or up to 4 TB of compressed data. Depending on compression factor, this accommodates up to 28 TB of source data[2]. IBM plans to extend the scale-out offering for SAP HANA using other building blocks with possibly different maximum cluster sizes.

Note: SAP validated this scale-out solution (with HA), which is documented in the SAP product availability matrix. At the time of the publication of this book, this scale-out solution can be implemented in a proof-of-concept scenario, and after that for a certain time in limited availability followed by general availability. Contact your IBM sales representative for exact dates.

4.5 Monitoring SAP HANA

In a productive environment, administration and monitoring of an SAP HANA appliance play an important role.

4.5.1 Monitoring with SAP HANA Studio

The SAP tool for administration of and monitoring the SAP HANA appliance is the SAP HANA Studio. It allows you to monitor the overall system state:

- General system information (such as software versions).
- A warning section shows the latest warnings generated by the statistics server. Detailed information about these warnings is available as a tooltip.
- Bar views provide an overview of important system resources. The amount of available memory, CPUs, and storage space is displayed, in addition to the used amount of these resources.

In a distributed landscape the amount of available resources is aggregated over all servers.

Note: More information about administration and monitoring of SAP HANA is available in the SAP HANA administration guide, accessible online:

http://help.sap.com/hana

4.5.2 Monitoring SAP HANA with Tivoli

Most of the monitoring data visible in SAP HANA Studio is collected by the statistics server, which is a monitoring tool for the SAP HANA database. It collects statistical and performance information from the database using SQL statements.

[2] Uncompressed source data, compression factor of 7:1

Monitoring data provided by the statistics sever can be used by other monitoring tools also. Figure 4-15 shows an image of this data integrated into IBM Tivoli® monitoring.

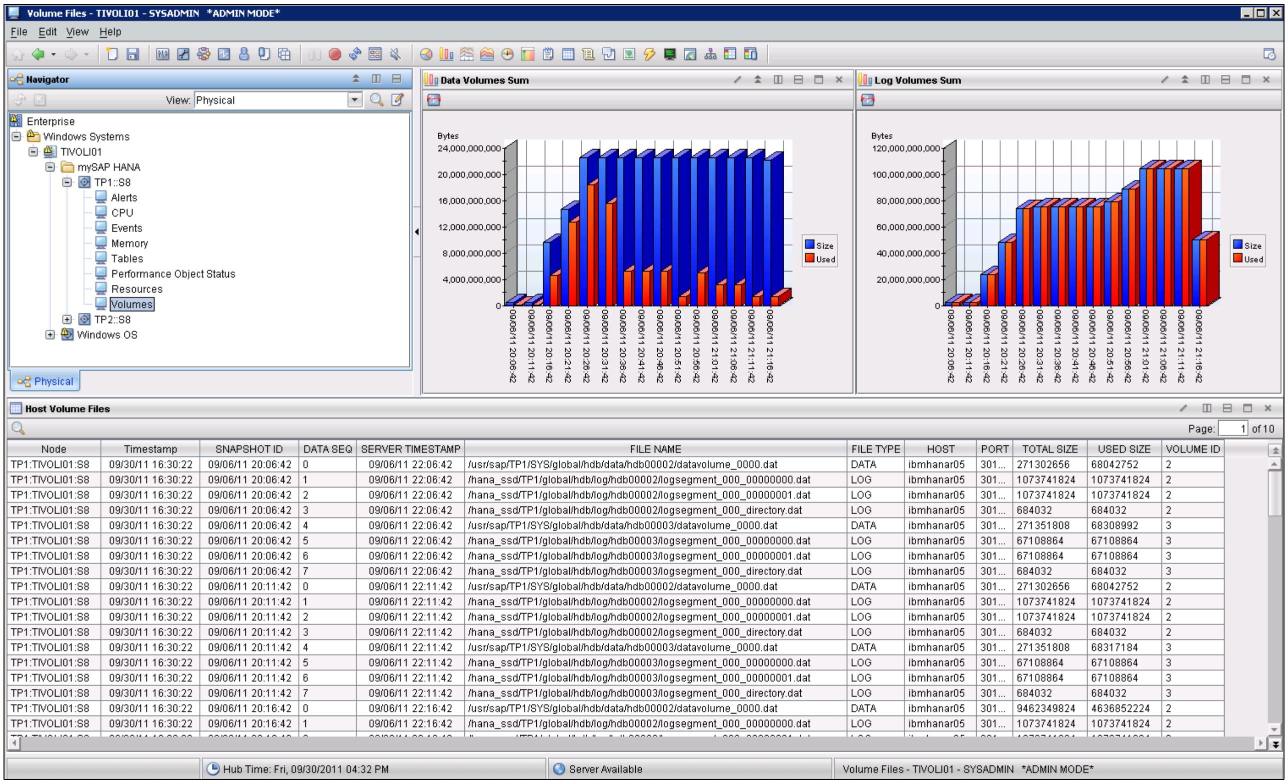


Figure 4-15 Monitoring the SAP HANA database with Tivoli

Tivoli monitoring also provides agents to monitor the operating system of the SAP HANA appliance. Hardware monitoring of the SAP HANA appliance servers can be achieved with IBM Systems Director, which also can be integrated into a Tivoli monitoring landscape.

By integrating the monitoring data collected by the statistics server, the Tivoli Monitoring agent for the operating system, and hardware information provided by IBM director, Tivoli Monitoring can provide a holistic view of the SAP HANA appliance.

Installing monitoring agents on the SAP HANA appliance

Only the software installed by your hardware partners is recommended on the SAP HANA appliance. Do not install additional software on the SAP HANA appliance. Initial tests have shown that some agents could decrease performance or even possibly corrupt the SAP HANA database (for example, virus scanners).

The list of agents that are supported, tolerated, or prohibited for use on the SAP HANA appliance will be published in an SAP Note.

4.6 Support process

The deployment of SAP HANA as an integrated solution combining software and hardware from both IBM and SAP also reflects in the support process for the IBM Systems Solution for SAP HANA.

All SAP HANA models offered by IBM include SLES for SAP Applications with SUSE 3-year priority support and IBM GPFS with 3-year support. The hardware comes with a 3-year limited warranty[3], including a customer replaceable unit (CRU) and on-site support[4].

4.6.1 IBM SAP integrated support

SAP integrates the support process with SUSE and IBM as part of the HANA appliance solution-level support. If you encounter software problems on your SAP HANA system, access the SAP Online Service System (SAP OSS) website (https://service.sap.com) and create a service request ticket using a subcomponent of BC-HAN or BC-DB-HDB as the problem component. IBM support works closely with SAP and SUSE and is dedicated to supporting SAP HANA software and hardware issues. You might see a message from a IBM support person in SAP's response.

Before you contact support, make sure that you have taken these steps to try to solve the problem yourself:

- Check all cables to make sure that they are connected.
- Check the power switches to make sure that the system and any optional devices are turned on.
- Use the troubleshooting information in your system documentation, and use the diagnostic tools that come with your system. Information about diagnostic tools is available in the *Problem Determination and Service Guide* on the IBM Documentation CD that comes with your system.
- Go to the following IBM support website to check for technical information, hints, tips, and new device drivers or to submit a request for information:

  http://www.ibm.com/supportportal/
- For SAP HANA software-related issues you can search the SAP Online Service System (OSS) website for problem resolutions. The OSS website has a knowledge database of known issues and can be accessed here:

  https://service.sap.com/notes

The main SAP HANA information source is available here:

https://help.sap.com/hana

If you have a specific operating system question or issue, contact SUSE regarding SUSE Linux Enterprise Server for SAP Applications. Go to the SUSE website:

http://www.suse.com/products/prioritysupportsap/

Media is available for download here:

http://download.novell.com/index.jsp?search=Search&families=2658&keywords=SAP

> **Note:** Registration is required before you can download software packages from the SUSE website.

3 For information about the IBM Statement of Limited Warranty, see http://www.ibm.com/servers/support/machine_warranties/.

4 IBM sends a technician after attempting to diagnose and resolve the problem remotely.

4.6.2 The IBM SAP International Competence Center (ISICC) InfoService

The ISICC InfoService is the key support function of the IBM and SAP Alliance. It serves as a single point of entry for all SAP-related questions for customers using IBM Systems and Solutions with SAP applications. As a managed question and answer service it has access to a worldwide network of experts on technology topics about IBM products in SAP environments. You can contact the ISICC InfoService via email at infoservice@de.ibm.com.

5

Summary

This chapter summarizes benefits of in-memory computing and advantages of IBM infrastructure for running the SAP HANA solution. It discusses these topics:

- 5.1, “Benefits of in-memory computing” on page 58
- 5.2, “SAP HANA: An innovative analytic appliance” on page 58
- 5.3, “IBM Systems Solution for SAP HANA” on page 59
- 5.4, “Going beyond infrastructure” on page 61

5.1 Benefits of in-memory computing

In today's data-driven culture, tools for business analysis are quickly evolving. Organizations need new ways to leverage critical data on the fly to not only accelerate decision making, but also to gain insights into key trends. The ability to instantly explore, augment, and analyze all data in near real time could deliver the competitive edge that your organization needs to make better decisions faster, as well as to take advantage of favorable market conditions, customer trends, price fluctuations, and other factors that directly influence the bottom line.

Made possible through recent technology advances that combine large, scalable memory, multi-core processing, fast solid-state storage, and data management, in-memory computing leverages these technology innovations to establish a continuous real-time link between insight, foresight, and action to deliver significantly accelerated business performance.

5.2 SAP HANA: An innovative analytic appliance

To support today's information-critical business environment, SAP HANA gives companies the ability to process huge amounts of data faster than ever before. The appliance lets business users instantly access, model, and analyze all of a company's transactional and analytical data from virtually any data source in real time, in a single environment, without impacting existing applications or systems.

The result is accelerated business intelligence (BI), reporting, and analysis capabilities with direct access to the in-memory data models residing in SAP in-memory database software. Advanced analytical workflows and planning functionality directly access operational data from SAP ERP or other sources. SAP HANA provides a high-speed data warehouse environment, with a SAP in-memory database serving as a next-generation, in-memory acceleration engine.

SAP HANA efficiently processes and analyzes massive amounts of data by packaging SAP's use of in-memory technology, columnar database design, data compression, and massive parallel processing together with essential tools and functionality such as data replication and analytic modeling.

Delivered as an optimized hardware appliance based on IBM eX5 enterprise servers, the SAP HANA software includes these:

- High-performance SAP in-memory database and a powerful data calculation engine
- Real-time replication service to access and replicate data from SAP ERP
- Data repository to persist views of business information
- Highly tuned integration with SAP BusinessObjects BI solutions for insight and analytics
- SQL and MDX interfaces for third-party application access
- Unified information-modeling design environment
- SAP BusinessObjects Data services to provide access to virtually any SAP and non-SAP data source

To explore, model, and analyze data in real time without impacting existing applications or systems, SAP HANA can be leveraged as a high-performance "side-by-side" data mart to an existing data warehouse. It can also replace the database server for SAP NetWeaver Business Warehouse, adding in-memory acceleration features.

These components create an excellent environment for business analysis, letting organizations merge large volumes of SAP transactional and analytical information from across the enterprise, and instantly explore, augment, and analyze it in near-real time.

5.3 IBM Systems Solution for SAP HANA

The IBM Systems solution for SAP HANA based on IBM eX5 enterprise servers provides the performance and scalability to run SAP HANA, enabling customers to drive near real-time business decisions and helping organizations stay competitive. IBM eX5 enterprise servers provide a proven, scalable platform for SAP HANA that enables better operational planning, simulation, and forecasting, in addition to optimized storage, search, and ad hoc analysis of today's information. SAP HANA running on powerful IBM eX5 enterprise servers combines the speed and efficiency of in-memory processing with the ability of IBM eX5 enterprise servers to analyze massive amounts of business data.

Based on scalable IBM eX5 technology included in IBM System x3690 X5 and System x3950 X5 servers, SAP HANA running on eX5 enterprise servers offers a solution that can help meet the need to analyze growing amounts of transactional data, delivering significant gains in both performance and scalability in a single, flexible appliance.

5.3.1 Workload Optimized Solution

IBM offers several Workload Optimized Solution models for SAP HANA. These models, based on the 2-socket x3690 X5 and 4-socket x3950 X5, are optimally designed and certified by SAP. They are delivered preconfigured with key software components preinstalled to help speed delivery and deployment of the solution. The x3690 X5-based configurations offer 128 - 256 GB of memory and the choice of only solid-state disk or a combination of spinning disk and solid-state disk. The x3950 X5-based configurations leverage the scalability of eX5 and offer the capability to pay as you grow, starting with a 2-processor, 256 GB configuration and growing to an 8-processor, 1 TB configuration. The x3950 X5-based configurations integrate either the 320 GB High IOPS SD Class SSD PCIe adapter or the High IOPS MLC Duo Adapter. The 8-socket configuration uses a scalability kit that combines the 7143-H2x with the 7143-H3x to create a single 8-socket, 1 TB system.

IBM offers the *appliance in a box* with no need for external storage. With the x3690 X5-based SSD model, IBM has a unique model using SSD only (no moving parts), providing greater reliability and performance versus spinning HDDs.

5.3.2 Leading performance

IBM eX5 enterprise servers offer extreme memory and performance scalability. With improved hardware economics and new technology offerings, IBM is helping SAP realize a real-time enterprise with in-memory business applications. IBM eX5 enterprise servers deliver a long history of leading SAP benchmark performance.

IBM eX5 enterprise servers come equipped with the Intel Xeon processor E7 series. These processors deliver performance that is ideal for your most data-demanding SAP HANA workloads and offer improved scalability along with increased memory and I/O capacity, which is critical for SAP HANA. Advanced reliability and security features work to maintain data integrity, accelerate encrypted transactions, and maximize the availability of SAP HANA applications. In addition, Machine Check Architecture Recovery, a reliability, availability, and serviceability (RAS) feature built into the Intel Xeon processor E7 series, enables the hardware platform to generate machine check exceptions. In many cases, these notifications enable the system to take corrective action that allows the SAP HANA to keep running when an outage would otherwise occur.

IBM eX5 features such as eXFlash solid-state disk technology can yield significant performance improvements in storage access, helping deliver an optimized system solution for SAP HANA. Standard features in the solution, such as the High IOPS MLC Duo Adapter for IBM System x, can also provide fast access to storage.

5.3.3 IBM GPFS enhancing performance, scalability, and reliability

Explosions of data, transactions, and digitally aware devices are straining IT infrastructure and operations, while storage costs and user expectations are increasing. The IBM General Parallel File System (GPFS), with its high-performance enterprise file management, can help you move beyond simply adding storage to optimizing data management for SAP HANA. High-performance enterprise file management using GPFS gives SAP HANA applications these:

- Performance to satisfy the most demanding SAP HANA applications
- Seamless capacity expansion to handle the explosive growth of SAP HANA information
- Scalability to enable support for the largest SAP HANA database requirements
- High reliability and availability to help eliminate production outages and provide disruption-free maintenance and capacity upgrades

Seamless capacity and performance scaling, along with the proven reliability features and flexible architecture of GPFS, help your company foster innovation by simplifying your environment and streamlining data workflows for increased efficiency.

5.3.4 Scalability

IBM offers configurations allowing customers to start with a 2 CPU/256 GB RAM model (S+), which can scale up to a 4 CPU/512 GB RAM model (M), and then to an 8 CPU/1024 GB configuration (L). With the option to upgrade S+ to M, and M+ to L, IBM can provide an unmatched upgrade path from a T-shirt size S up to a T-shirt size L, without the need to retire a single piece of hardware.

If you have large database requirements, you can scale the workload-optimized solutions to multi-server configurations. IBM and SAP have validated configurations of up to sixteen 512 GB nodes with high availability. This scale-out support enables support for databases as large 8 TB, able to hold the equivalent of about 28 TB of uncompressed data. While the IBM solution is certified for up to 16 nodes, its architecture is designed for extreme scalability and can even grow beyond that. The IBM solution does not require external storage for the standalone or for the scale-out solution. The solution is easy to grow by the simple addition of nodes to the network. There is no need to reconfigure a storage area network for failover. That is all covered by GPFS under the hood.

IBM uses the same base building blocks from standalone servers to scale out, providing investment protection for customers who want to grow their SAP HANA solution beyond a single server.

IBM or IBM Business Partners can provide these scale-out configurations preassembled in a rack, helping to speed installation and setup of the SAP HANA appliance.

5.3.5 Services to speed deployment

To help speed deployment and simplify maintenance of your x3690 X5 and x3950 X5, the Workload Optimized Solution for SAP HANA, IBM Lab Services, and IBM Global Technology Services offer quick-start services to help set up and configure the appliance and health-check services to ensure that it continues to run optimally. In addition, IBM also offers skills and enablement services for administration and management of IBM eX5 enterprise servers.

5.4 Going beyond infrastructure

Many clients require more than software and hardware products. IBM as a globally integrated enterprise can provide clients with the full range of hardware, software, infrastructure, and consulting services.

5.4.1 A trusted service partner

Clients need a partner to help them assess their current capabilities, identify areas for improvement, and develop a strategy for moving forward. This is where IBM Global Business Services® provides immeasurable value with thousands of SAP consultants in 80 countries. The SAP Consulting Practice offers a broad range of services for SAP HANA, such a these:

- Discovery and assessment services to maximize business impact
- Architecture assessment and benchmark services
- Proof-of-concept services
- Express deployment offerings, including industry best practices

By drawing on these resources, we can help you take full advantage of SAP HANA running on IBM eX5 enterprise servers.

5.4.2 IBM and SAP team for long-term business innovation

With a unique combination of expertise, experience, and proven methodologies—and a history of shared innovation—IBM can help strengthen and optimize your information infrastructure to support your SAP applications.

IBM and SAP have worked together for nearly 40 years to deliver innovation to their shared customers. Since 2006, IBM has been the market leader for implementing SAP's original in-memory appliance, the SAP NetWeaver Business Warehouse Accelerator (BWA). Hundreds of BWA deployments have been successfully completed in multiple industries. These BWA appliances have been successfully deployed on many of SAP's largest business warehouse implementations, which are based on IBM hardware and DB2, optimized for SAP.

IBM and SAP offer solutions that move business forward and anticipate organizational change by strengthening your business analytics information infrastructure for greater operational efficiency and offering a way to make smarter decisions faster.

Abbreviations and acronyms

ABAP	Advanced Business Application Programming
ACID	Atomicity, Consistency, Isolation, Durability
APO	Advanced Planner and Optimizer
BI	Business Intelligence
BM	Bridge module
BO	BusinessObjects
BOE	BusinessObjects Explorer
BW	Business Warehouse
BWA	Business Warehouse Accelerator
CD	Compact disc
CPU	Central processing unit
CRC	Cyclic redundancy checking
CRM	Customer Relationship Management
CRU	Customer-replaceable unit
DB	Database
DIMM	Dual inline memory module
ECC	Error correction code
ERP	Enterprise resource planning
ETL	Extract, transform, and load
FTSS	Field Technical Sales Support
GB	Gigabyte
GBS	Global Business Services
GPFS	General Parallel File System
GTS	Global Technology Services®
HA	High availability
HDD	Hard disk drive
HPI	Hasso Plattner Institute
I/O	Input/output
IBM	International Business Machines
ID	Identifier
IMM	Integrated Management Module
IOPS	I/O operations per second
ISICC	IBM SAP International Competence Center
ITSO	International Technical Support Organization
LM	Lifecycle Management
LT	Landscape Transformation
LUW	Logical unit of work
MB	Megabyte
MCA	Machine Check Architecture
NOS	Notes object services
NSD	Network shared disk
NUMA	Non-Uniform Memory Access
OS	Operating system
OSS	Online service system
PC	Personal computer
PCI	Peripheral Component Interconnect
QPI	QuickPath Interconnect
RAID	Redundant Array of Independent Disks
RAM	Random access memory
RAS	Remote access services; row address strobe
RDS	Rapid Deployment Solutions
RPM	Revolutions per minute
SAN	Storage area network
SAPS	SAP Application Benchmark Performance Standard
SAS	Serial Attached SCSI
SATA	Serial ATA
SCM	Software configuration management
SD	Sales and distribution
SDRAM	Synchronous dynamic random access memory
SLES	SUSE Linux Enterprise Server
SLO	System Landscape Optimization
SMI	Synchronous memory interface
SQL	Structured Query Language
SSD	Solid state drive
STG	Server & Technology Group
TCO	Total cost of ownership
TCP/IP	Transmission Control Protocol/Internet Protocol
TDMS	Test Data Migration Server
TREX	Text Retrieval and information EXtraction
UEFI	Unified Extensible Firmware Interface
USB	Universal Serial Bus

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this paper.

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- *The Benefits of Running SAP Solutions on IBM eX5 Systems*, REDP-4234
- *IBM eX5 Portfolio Overview: IBM System x3850 X5, x3950 X5, x3690 X5, and BladeCenter HX5*, REDP-4650

You can search for, view, download, or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Other publications

This publication is also relevant as a further information source:

- Prof. Hasso Plattner, Dr. Alexander Zeier, “In-Memory Data Management,” Springer, 2011

Online resources

These websites are also relevant as further information sources:

- IBM Systems Solution for SAP HANA

  http://www.ibm.com/systems/x/solutions/sap/hana/

- IBM Systems and Services for SAP HANA

  http://www.ibm-sap.com/hana

- IBM and SAP: Business Warehouse Accelerator

  http://www.ibm-sap.com/bwa

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

IBM®

SAP In-Memory Computing on IBM eX5 Systems

Redpaper™

IBM Systems Solution for SAP HANA

SAP HANA overview and use cases

Basic In-Memory Computing principles

This IBM® Redpaper™ publication describes the new in-memory computing appliances from IBM and SAP that are based on IBM eX5 flagship systems and SAP HANA. We first discuss the history and basic principles of in-memory computing, then we describe the SAP HANA offering, its architecture, sizing methodology, and licensing policy. We also review IBM eX5 hardware offerings from IBM. We describe the architecture and components of IBM System Solution for SAP HANA and its delivery, operational, and support aspects. Finally, we discuss the advantages of using IBM infrastructure platforms for running the SAP HANA solution.

These topics are covered:

- The history of in-memory computing
- The basic principles of in-memory computing
- The SAP HANA offering
- The IBM System Solution for SAP HANA
- Benefits of using the IBM infrastructure for SAP HANA

This paper is for SAP administrators and technical solution architects. It is also for IBM Business Partners and IBM employees who want to know more about the SAP HANA offering and other available IBM solutions for SAP customers.

REDP-4814-00