Lenovo

In-memory Computing with SAP HANA on Lenovo Systems

Introduces the Lenovo Solution for SAP HANA

Explores the SAP HANA features and use cases

Explains the delivery models

Describes business continuity and operational disciplines

Irene Hopf

LENOVO PRESS

SAP HANA on Lenovo Systems

August 2019

SG24-8086-05

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

Last update on August 2019

This edition applies to Lenovo ThinkSystem and Lenovo storage and networking offerings.

Contents

Notices

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

Lenovo 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:

Lenovo (United States), Inc.
1009 Think Place - Building One
Morrisville, NC 27560
U.S.A.
Attention: Lenovo Director of Licensing

LENOVO 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 jurisdictions 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. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary.

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

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

Any performance data contained herein was determined in a controlled environment. Therefore, the result obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Trademarks

Lenovo, the Lenovo logo, and For Those Who Do are trademarks or registered trademarks of Lenovo in the United States, other countries, or both. These and other Lenovo 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 Lenovo at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of Lenovo trademarks is available on the Web at http://www.lenovo.com/legal/copytrade.html.

The following terms are trademarks of Lenovo in the United States, other countries, or both:

AnyBay™
Lenovo®
Lenovo XClarity™
RackSwitch™
Lenovo(logo)®
ServeRAID™
System x®
ThinkAgile™
ThinkSystem™
TruDDR4™

The following terms are trademarks of other companies:

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

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

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

SAP, SAP HANA, SAP S/4HANA, SAP BW/4HANA, SAP NetWeaver, and other SAP products and services are trademarks or registered trademarks of SAP SE or an SAP affiliate company in Germany and other countries.

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

Preface

The sixth edition of this Lenovo^{TM} Press publication describes in-memory computing systems from Lenovo® and SAP that are based on Lenovo systems and SAP HANA. It covers the basic principles of in-memory computing, describes the Lenovo hardware offerings, and explains the corresponding SAP HANA IT landscapes that use these offerings.

The focus of this new edition is the introduction of persistent memory in Intel Xeon SP Gen 2 systems to be used by SAP HANA based applications. Also the change of a more open platform to design specific configurations according to the customers workload requirements.

This book also describes the architecture and components of the Lenovo systems solution for SAP HANA. The following SAP HANA operational disciplines are explained: Scalability options; high availability and disaster recovery; backup and restore; and virtualization possibilities for SAP HANA systems.

The following topics also are covered:

- Basic principles of in-memory computing
- SAP HANA overview
- Software components and replication methods
- SAP HANA use cases and integration scenarios
- The Lenovo solution for SAP HANA
- SAP IT landscapes that use the System solution for SAP HANA
- Options for business continuity (high availability, disaster recovery, and backup and restore)
- SAP HANA operations

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

> **Note:** This book has not yet been edited for grammar and readability. Please excuse any errors you might find!

Author

Irene Hopf leads the Center of Competence for SAP Solutions at Lenovo globally. She has more than 25 years of experience in the SAP customer space with infrastructure solutions.

In this context she is recognized as a thought leader and gives trusted advise for large scale customer implementations for the integration of services, hardware and software supporting SAP implementations.

She has written many publications for the SAP ecosystem.

Irene has graduated in material science engineering and information technology and is based in Walldorf, Germany.

Thanks to the authors of the previous edition of this book:

- Martin Bachmaier
- Ilya Krutov

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

From Lenovo:

- Gereon Vey
- Vijayakumar Kulageri
- Tag Robertson
- David Watts

From SAP:

- Dr. Richard Bremer
- Dr. Thomas Becker

Note: Thank you to our SAP colleagues for their input and contribution to this book. Their insights and perspectives help provide deeper insights into the SAP HANA database technology. Please note that the sources of information referenced in this book can be found on www.sap.com and help.sap.com.

1

Basic concepts of in-memory computing

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

To achieve the preferred performance, in-memory computing adheres to the following basic concepts:

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

This chapter describes these basic concepts and provides some examples. It does not describe the full set of technologies that are used with in-memory databases, such as SAP HANA, but it does provide an overview of how in-memory computing is different from traditional concepts.

In the sixth edition of this book, this chapter got enhanced with the usage of persistent memory in SAP HANA, the principle and benefits

This chapter includes the following topics:

1.1 Keeping data in-memory

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 describes the benefits and challenges.

1.1.1 Using main memory as the data store

The most obvious reason to use main memory (RAM) as the data store for a database is that accessing data in main memory is much faster than accessing data on disk. The compared access times for data in several levels are shown in Figure 1-1.

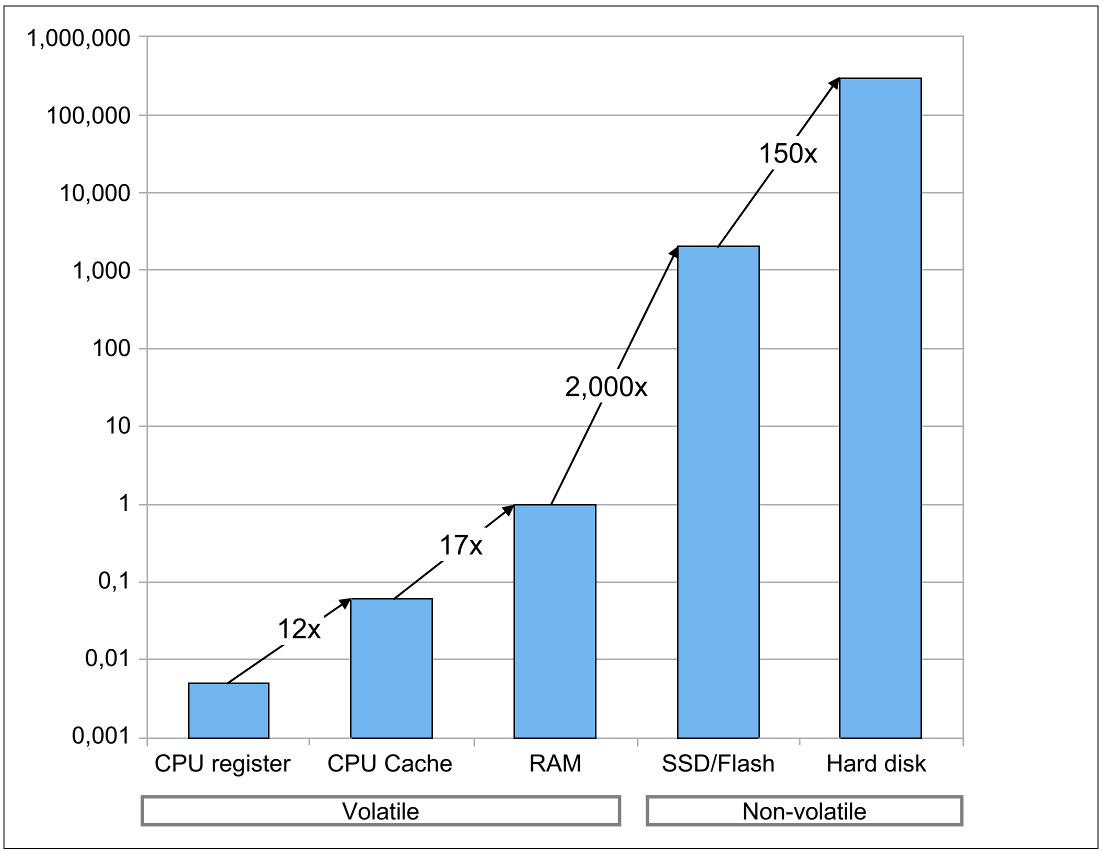


Figure 1-1 Data access times of various storage types relative to RAM (logarithmic scale)

The main memory is the fastest storage type that can hold a significant amount of data. Although CPU registers and CPU cache are faster to access, their usage is limited to the actual data processing. Data in main memory can be accessed more than a hundred thousand times faster than data on a spinning hard disk drive (HDD), and even flash technology storage is approximately a thousand times slower than main memory. Main memory is connected directly to the processors through a high-speed bus, and hard disks are connected through a chain of buses (QPI, PCIe, and 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 improve database performance through the advantage in access time.

1.1.2 Data persistence

Keeping data in main memory brings up the issue of what happens if there is a loss of power.

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

- A transaction must be atomic. If part of a transaction fails, the entire transaction must 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 interferes with another transaction.
- Durability means that after a transaction is committed, it remains committed.

Although 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. It loses its content when it is out of electrical power. To make data persistent, it must be on non-volatile storage, such as HDDs, solid-state drives (SSDs), or flash devices.

The storage that is 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 that are made by transactions. Each committed transaction generates a log entry that is written to non-volatile storage, which ensures that all transactions are permanent.

Figure 1-2 shows this setup by 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. A transaction does not return before the corresponding log entry is written to persistent storage to meet the durability requirement.

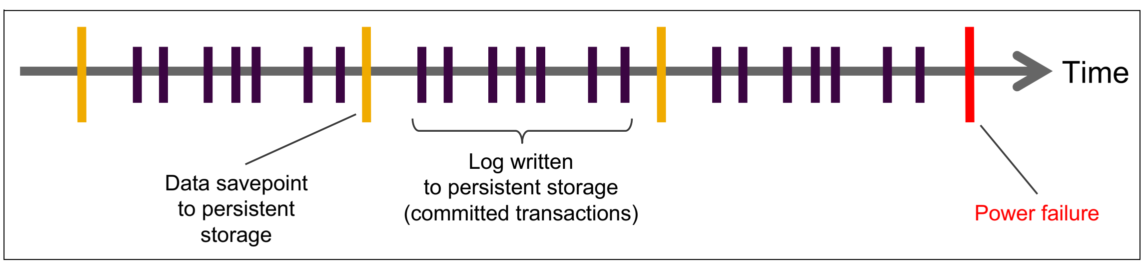


Figure 1-2 Savepoints and logs in SAP HANA

After a power failure, the database can be restarted much 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 action ensures that the database can be restored in memory to the same state as before the power failure. Even faster ways of restarting the applications to improve business continuity are described in chapter 1.4, “Principles for persistent memory” on page 9.

1.2 Minimizing data movement

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

1.2.1 Compression

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

By working with dictionaries to represent text as integer numbers, the database can compress data significantly and thus reduce data movement while not imposing more CPU load for decompression; in fact, it can add to the performance, as shown in Figure 1-5 on page 6. This situation with a simplified example is shown in Figure 1-3.

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 1-3 Dictionary compression

The original table is shown on the left side of Figure 1-3, and it contains text attributes (that is, material and customer name) in their original representation. The text attribute values are stored in a dictionary (upper right), and an integer value is assigned to each distinct attribute value. In the table, the text is replaced by the corresponding integer value as defined in the dictionary. The date and time attribute also are converted to an integer representation. The use of dictionaries for text attributes reduces the size of the table because each distinct attribute value must be stored only once in the dictionary; therefore, each additional occurrence in the table must be referred to by the corresponding integer value.

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

There are other, more effective compression methods that can be used with in-memory computing. However, for these methods 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 that are needed for decompression, and data accessibility (that is, how much unrelated data must be decompressed to get to the data that you need). Dictionary compression combines good compression effectiveness with low decompression resources and high data access flexibility.

1.2.2 Columnar storage

Relational databases organize data in tables that contain the data records. The difference between row-based and columnar (or column-based) storage is how 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. The row-based and column-based models are shown in Figure 1-4.

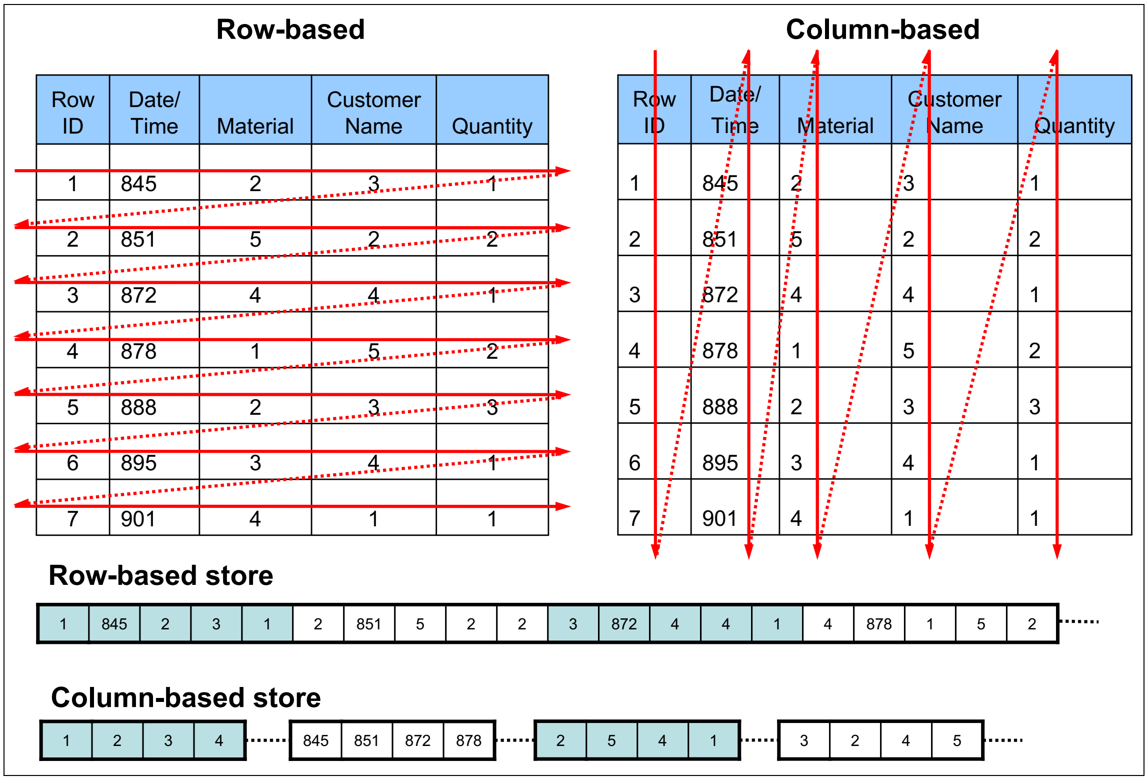


Figure 1-4 Row-based and column-based storage models

Both storage models have benefits and drawbacks, which are listed in Table 1-1.

Table 1-1 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 columns must be read during the selection process of a query. ► Efficient projections.[a] ► Any column can serve as an index.
Drawbacks	All data must be read during selection, even if only a few columns are involved in the selection process.	► After selection, selected rows must be reconstructed from columns. ► No easy insert/update.

a. Projection refers to the 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.[1] SAP HANA implements a row-based storage and a column-based storage; however, its performance originates in the usage of column-based storage in memory. The following sections describe how column-based storage is beneficial to query performance and how SAP HANA handles the drawbacks of column-based storage.

Efficient query execution

To show the benefits of dictionary compression that is combined with columnar storage, Figure 1-5 shows an example of how a query is run. (Figure 1-5 refers to the table that is shown in Figure 1-3 on page 4.)

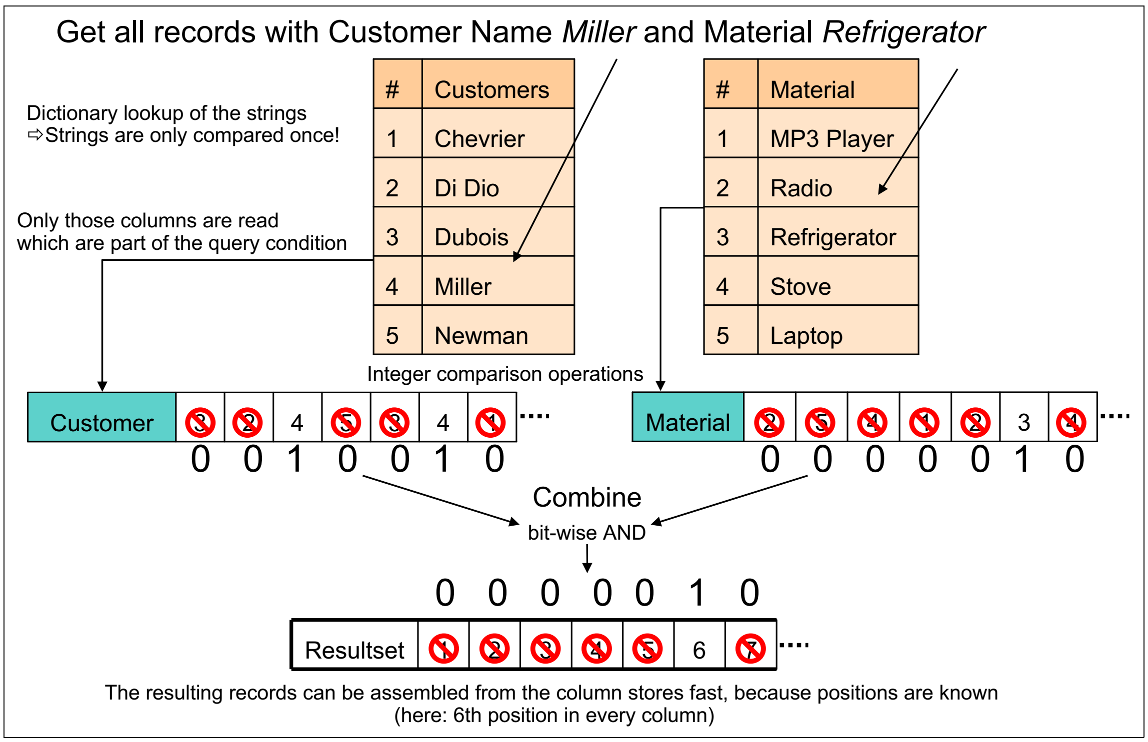


Figure 1-5 Example of a query that is run on a table in columnar storage

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 by the number 4 in the customer name column. Refrigerator is represented by the number 3 in the material column. This lookup must be done only once. Subsequent comparisons with the values in the table are 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 that match 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.

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

These selection marks can be represented as bitmaps, which are data structures that allow efficient Boolean operations on them that are used to combine the bitmaps of the individual columns to a bitmap that represents the selection or records that match the entire query condition. In this example, the record number 6 is the only matching record. Depending on the columns that are selected for the result, the extra columns must be read to compile the entire record to return it. However, because the position within the column is known (record number 6), only the parts of the columns that contain the data for this record must be read.

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

Delta merge and bulk inserts

To overcome the drawback of inserts or updates that affect the performance of the column-based storage, SAP implemented a lifecycle management for database records.[2]

The following types of storage for a table are available:

- L1 Delta store is optimized for fast write operations. The update is performed by inserting a new entry into the delta storage. The data is stored in records, as with a traditional row-based approach. This action ensures high performance for write, update, and delete operations on records that are stored in the L1 Delta Storage.
- L2 Delta store is an intermediate step. Although organized in columns, the dictionary is not as optimized as in the main storage because it appends new dictionary entries to the end of the dictionary. This action results in easier inserts, but has drawbacks regarding search operations on the dictionary because it is not sorted.
- Main store contains the compressed data for fast read with a search optimized dictionary.

All write operations on a table work on the L1 Delta store. Bulk inserts bypass L1 Delta store and write directly into L2 Delta store. Read operations on a table always read from all stores for that table. The result set is merged to provide a unified view of all data records in the table.

During the lifecycle of a record, it is moved from L1 Delta store to L2 Delta store and finally to the Main store. The process of moving changes to a table from one store to the next one is known as *Delta Merge*, which is an asynchronous process. During the merge operations, the columnar table is still available for read and write operations.

Moving records from L1 Delta store to L2 Delta store involves reorganizing the record in a columnar fashion and compressing it, as shown in Figure 1-3 on page 4. If a value is not yet in the dictionary, a new entry is appended to the dictionary. Appending to the dictionary is faster than inserting, but results in an unsorted dictionary, which affects the data retrieval performance.

Eventually, the data in the L2 Delta store must be moved to the Main store. To accomplish that task, the L2 Delta store must be locked and a new L2 Delta store must be opened to accept further additions. Then, a new Main store is created from the old Main store and the locked L2 Delta store. This task is resource-intensive and must be scheduled carefully.

[2] *Efficient Transaction Processing in SAP HANA Database - The End of a Column Store Myth*, which is available at this website:
`http://dl.acm.org/citation.cfm?id=2213946`.

1.2.3 Pushing application logic to the database

Although the concepts that are described in 1.2.2, "Columnar storage" on page 5 speedup processing within the database, there is still one factor that can slow down the processing of data. An application that is running the application logic on the data must 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 an effect on communication and latency and is limited by the speed and throughput of the network between the database and the application.

To eliminate this factor and increase overall performance, it is beneficial to process the data where it is (in the database.) If the database can perform calculations and apply application logic, less data must 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 action minimizes the amount of data transfer, and the communication between database and application adds less time to the overall processing time.

1.3 Divide and conquer

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

1.3.1 Parallelization on multi-core systems

When chip manufacturers 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 be used through parallel processing only because the performance of a single core remains unchanged.

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

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 that are processed in parallel.

1.3.2 Data partitioning and scale-out

Although servers can hold terabytes of data in memory and provide multiple processors per server with up to 28 cores per processor, the amount of data that is stored in an in-memory database or the computing power that is 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, which forms a distributed database (scale-out approach).

The individual database tables can be placed on different servers within the cluster. Tables bigger than what a single server can hold can be split into several partitions horizontally (a group of rows per partition) with each partition on a separate server within the cluster. This is applicable to analytical as well as transactional scenarios.

1.4 Principles for persistent memory

We know that SAP HANA stores and processes the application data in memory. What if the server is being powered off? The data would be lost. Of course there is still all the data persisted on disk, but it would take potentially a lot of time to bring it back into memory. This depends on the amount of data and subsequently the system topology, whether it is a single node (scale-up) system, or a scale-out cluster as well as on the business use case.

What if the data can be kept in memory despite of a power loss? What if, after a reboot, the application can continue faster to work, because the loadtime from the persistency layer on disk to memory can be omitted?

This is possible with the technology called Intel Optane DC Persistent Memory (DCPMM) starting in Xeon SP Gen 2 (formerly codenamed Cascade Lake) servers. It is supported with SAP HANA 2.0 SPS 03, Revision 35 (2.00.035), HANA 2.0 SPS 04 any revision and onwards.

The business benefit can be threefold:

- Higher maximum capacity of main memory
- Better business continuity with less downtime and shorter startup times
- Additional use cases reducing the total cost of ownership

Early measurements of business continuity showed an improvement for the SAP HANA restart time from 50 minutes on a traditional system to 4 minutes with persistent memory in a 6 TB SAP HANA configuration. That is a 12.5x improvement in restart time when compared to the system with the previous generation processor and traditional DRAM[3]. This measurement does not include the system and OS boot times.

1.4.1 Architecture

As depicted in Figure 1-6 on page 11 at the top, the volatile data structure remains in DRAM, in the middle the column store main moves to persistent memory and at the bottom, there are no changes to the persistence layer. Please refer to 1.2.2, “Columnar storage” on page 5 for more details on the structure of the column store.

There are three modes for persistent memory as described in 5.1.7, “Intel Optane DC Persistent Memory” on page 68. For SAP HANA only the App Direct Mode is applicable.

SAP HANA controls what is placed in DCPMM and what remains in DRAM. The column store main is heavily optimized in terms of compression, leading to a very stable – non-volatile – data structure. The main store typically contains well over 90% of the data footprint in most SAP HANA databases, which means it offers a lot of potential. Furthermore, it is reconstructed rarely during the delta merge. A process that is only triggered after a certain threshold of changes to the database table was reached. For most tables, a delta merge does not happen more than once a day.

The main store is read-optimzed and that leads to a perfect fit to place this in persistent memory. The delta store is bound to be write-optimized and hence fitting better into DRAM.

[3] Results have been estimated based on tests conducted on pre-production systems, and provided for informational purposes. Any differences in your system hardware, software or configuration may affect the actual performance. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark* and MobileMark*, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information go to intel.com/benchmarks.

SAP HANA simulated workload for SAP BW edition for SAP HANA Standard Application Benchmark Version 2 as of 30 May 2018. Baseline configuration with traditional DRAM: Lenovo ThinkSystem™ SR950* server with 8 x Intel Xeon Platinum 8176M processors (28 cores, 165 watt, 2.1 GHz). Total memory consists of 48 x 16 GB TruDDR4™* 2,666 MHz RDIMMS and 5 x ThinkSystem* 2.5” PM1633a 3.84 TB capacity SAS 12 Gb hot-swap solid-state drives (SSDs) for SAP HANA storage. The operating system is SUSE* Linux* Enterprise Server 12 SP3 and uses SAP HANA 2.0 SPS 03 with a 6 TB dataset. Average start time for all data finished after table preload for 10 iterations: 50 minutes.

New configuration with a combination of DRAM and Intel Optane DC persistent memory: Lenovo ThinkSystem SR950* server with 8 x Intel Xeon Platinum 8176M processors (28 cores, 165 watt, 2.1 GHz). Total memory consists of 48 x 16 GB TruDDR4* 2,666 MHz RDIMMs and 48 x 128 GB Intel Optane DC persistent memory modules (PMMs), and 5 x ThinkSystem* 2.5” PM1633a 3.84 TB capacity SAS 12 Gb hot-swap solid-state drives (SSDs) for SAP HANA* storage. The operating system is SUSE* Linux* Enterprise Server 12 SP3 and uses SAP HANA 2.0 SPS 03 with a 6 TB dataset. Average start time for all data finished after table preload for 10 iterations: 4 minutes (12.5x improvement).

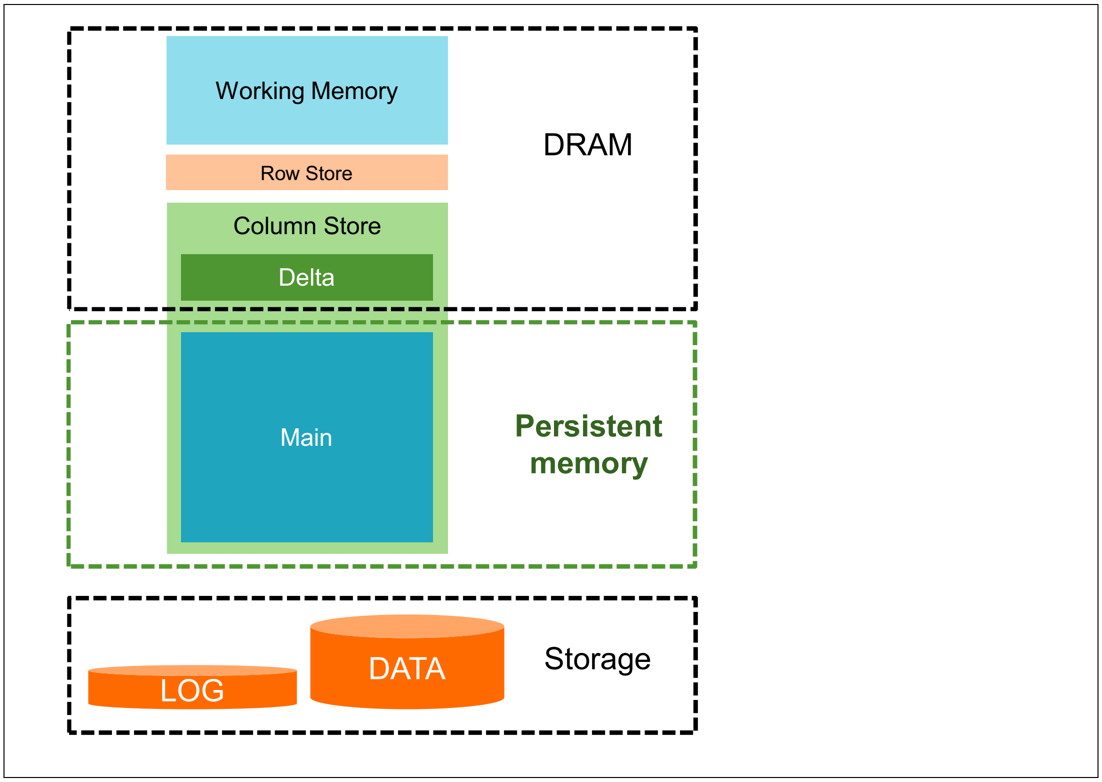
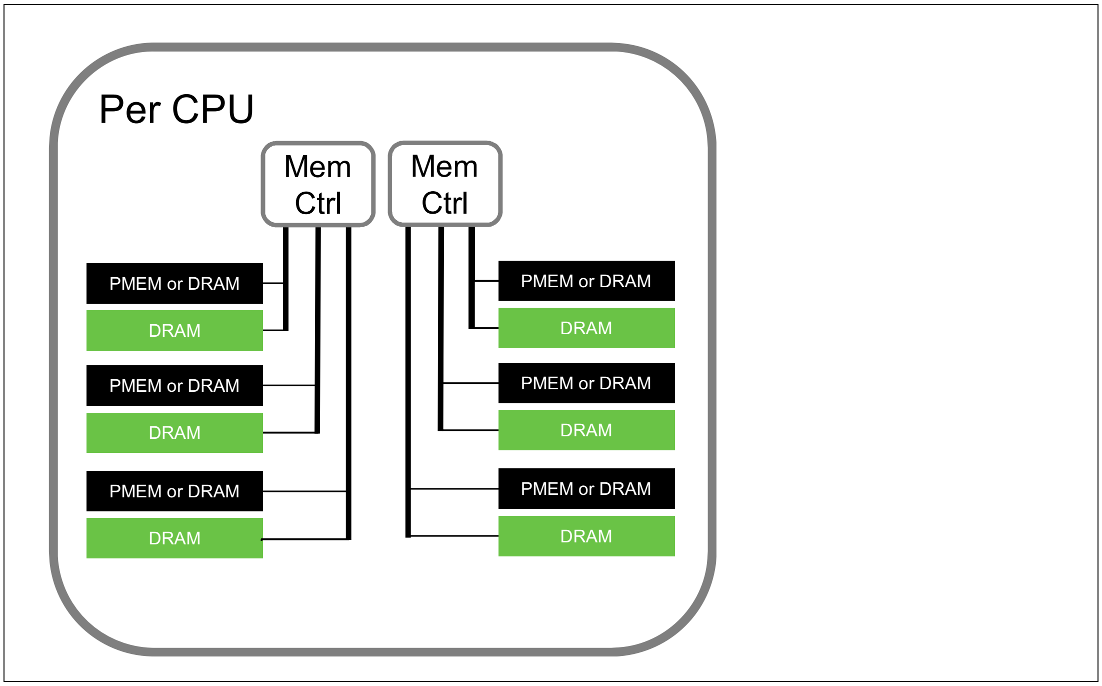


Figure 1-6 Utilization of Persistent Memory by SAP HANA

1.4.2 Memory capacity

There will still be standard DDR4 memory DIMMs referred to as DRAM and Persistent memory DIMMs referred to as DCPMM or PMEM in a given server of Intel Xeon SP Gen 2 for SAP HANA. The CPU has up to 28 cores and supports two memory controllers per CPU, three memory channels per controller and two DIMMs per channel depicted in Figure 1-7 on page 12. This results in a total of 12 memory slots per CPU, so there are 24 memory slots in a 2 socket server, and so on, scaling linearly to 4 and 8 socket systems. The ratio of DRAM to persistent memory can vary dependent on the business workload requirements.



Memory architecture

SAP requires a uniform memory configuration across all memory channels in a system.

At the time of writing possible DRAM sizes are 16GB, 32GB, 64GB, 128GB, 256GB per DIMM and possible DCPMM sizes are 128GB, 256GB or 512GB per DIMM.

Figure 1-7 on page 12 illustrates an example of a 2 socket server. On the left side a configuration with DRAM and PMEM is depicted with the largest possible memory DIMM sizes and a ratio of 1:4. On the left side you see the configuration with DRAM only.

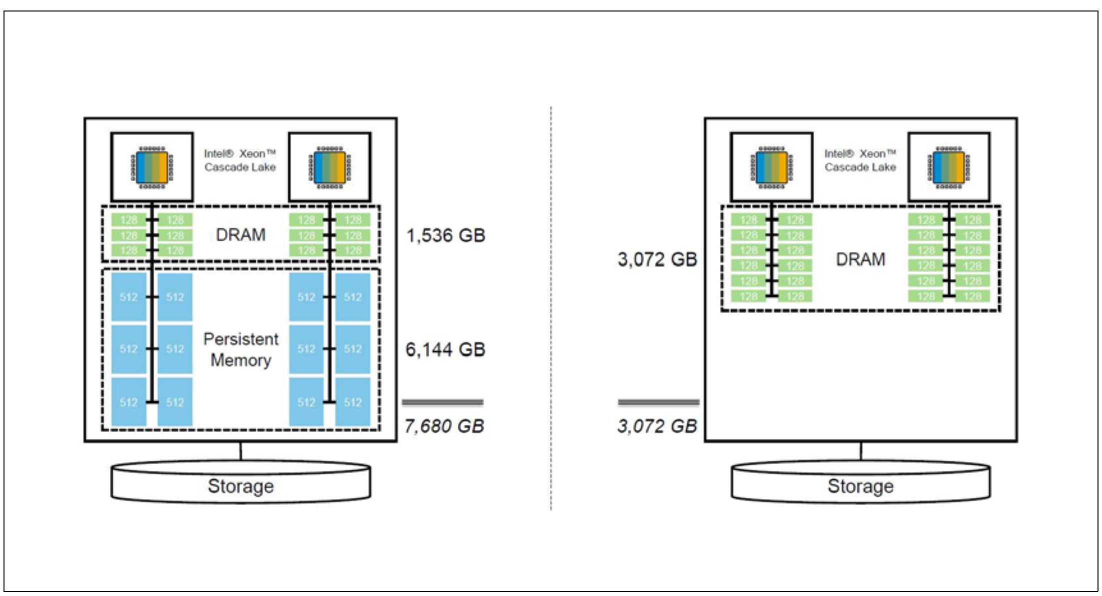


Figure 1-7 Example of memory configurations

1.4.3 Ratio of DRAM to PMEM

An important aspect to consider, when designing an appropriate configuration for a given workload is the capacity ratio between DRAM and PMEM.

SAP investigated various workloads, transactional and analytical, both from customer examples as well as inhouse lab environments. Compared were the total memory in use and the distribution of the data consumption in DRAM without the main store and the main store in persistent memory. The various systems showed very different PMEM to DRAM ratios.

It varied from **less than half of the capacity to almost 9 times the capacity of DRAM**! To predict the ratio for a given customer workload in a sizing exercise is critical and SAP provides help via several tools. Please refer to chapter "Intel Optane DC Persistent Memory and sizing considerations" on page 26.

SAP has complied a Frequently Asked Questions document in

```
SAP Note 2700084 - FAQ: SAP HANA Persistent Memory
```

2

SAP HANA and SAP S/4HANA overview

This chapter describes the SAP HANA offering, including its architecture, components, use cases, delivery model, and sizing and licensing aspects.

This chapter includes the following topics:

2.1 SAP HANA overview

This section gives an overview of SAP HANA. The following terms are used regarding SAP HANA:

- SAP HANA database

  The SAP HANA database (which is also referred to as the *SAP in-memory database*) is a hybrid in-memory database that combines row-based, column-based, and object-based database technology that is optimized to use the parallel processing capabilities of current hardware. It is the heart of SAP offerings, such as SAP HANA.

- SAP HANA appliance (SAP HANA)

  SAP HANA is a flexible, data source-neutral appliance with which you can analyze large volumes of data in real time without needing to materialize aggregations. It is a combination of hardware and software and is delivered as an integrated system with SAP's hardware partners for SAP HANA.

For the sake of simplicity, this book uses the terms SAP HANA, SAP in-memory database, SAP HANA database, SAP HANA system and SAP HANA appliance synonymously. The focus is the in-memory database. Where required, we ensure that the context makes it clear which part is being described.

2.1.1 SAP HANA architecture

The high-level architecture of the SAP HANA system is depicted in Figure 2-1. The most important software components of the SAP HANA database are described in 3.1, "SAP HANA software components" on page 32.

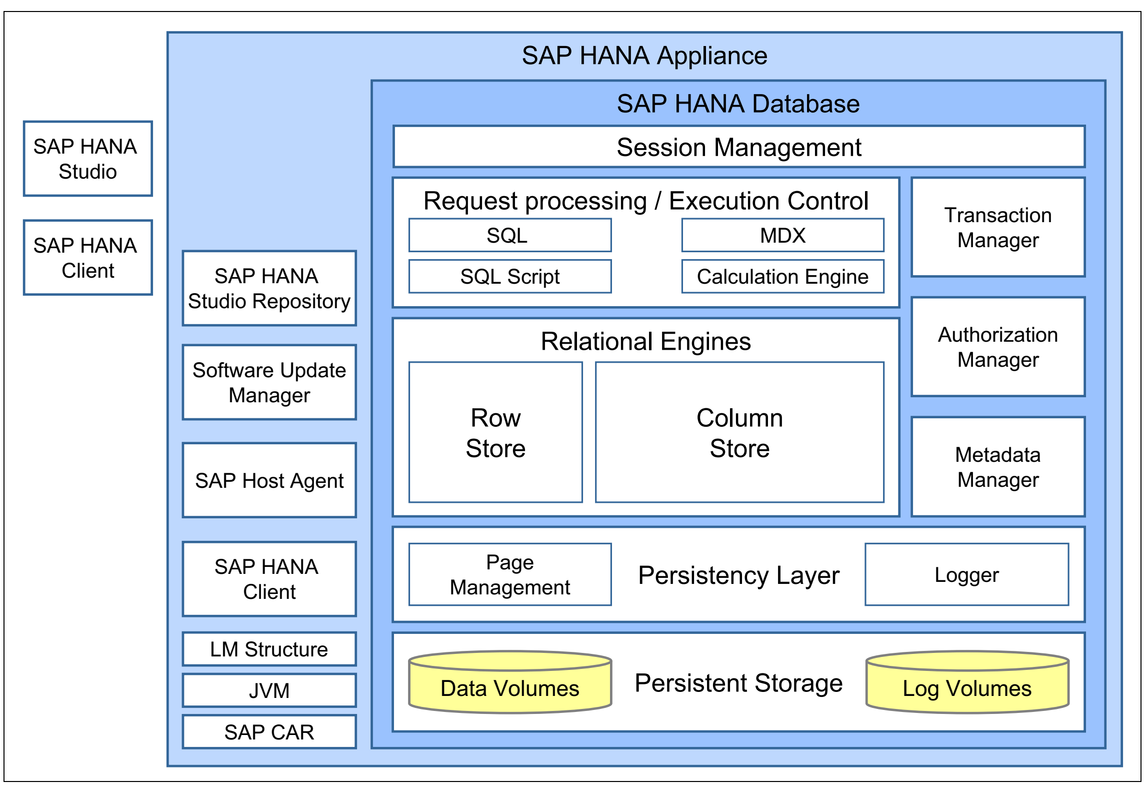


Figure 2-1 SAP HANA architecture

SAP HANA database

The heart of the SAP HANA database is the relational database engines. The SAP HANA database features the following engines:

- The column-based store

  Stores relational data in columns, which are optimized holding tables with large amounts of data that can be aggregated in real time and used in analytical operations.

- The row-based store

  Stores relational data in rows, as traditional database systems do. This row store is more optimized for row operations, such as frequent inserts and updates. It has a lower compression rate and query performance is much lower compared to the column-based store.

The engine that is used to store data can be selected on a per-table basis when the table is created. A table can be converted from one type to another type. Tables in the row-store are loaded at start time, but tables in the column-store can be loaded at start or on demand during normal operation of the SAP HANA database.

Both engines share a common persistency layer, which provides data persistency that is consistent across both engines. There is page management and logging, as with traditional databases. Changes to in-memory database pages are persisted through savepoints that are written to the data volumes on persistent storage, which often is hard disk drives (HDDs). Every transaction that is committed in the SAP HANA database is persisted by the logger of the persistency layer in a log entry that is 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 various 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 perform business calculations close to the data. The SAP HANA-specific SQL Script language is an extension to SQL that can be used to push down data-intensive application logic into the SAP HANA database.

Native Storage Extension

As a new feature starting with SAP HANA 2.0 SPS04 SAP introduced Native Storage Extension (NSE). There are multiple options to reduce the memory footprint and keep less frequently accessed data on disk in the same system or a separate layer on a different physical server.

All of those solutions need to be evaluated carefully, dependent on what kind of application scenario is being used (e.g. analytical, a native data warehouse or transactional, SAP S/4HANA, single nodes or a scale-out cluster).

More information about NSE is available in the SAP Administration Guide and SAP Note: 2771956

2.1.2 SAP HANA system

SAP HANA consists of the SAP HANA database and adds components that are needed to work with, administer, and operate the database. It contains the repository files for the SAP HANA studio respectively the SAP HANA cockpit, which is an Eclipse-based administration and data-modeling tool for SAP HANA. It also includes the SAP HANA client, which is a set of libraries that is required for applications to connect to the SAP HANA database. The SAP HANA studio or the cockpit and the client libraries often are installed on a client PC or server.

To support the lifecycle management, SAP offers the SAP HANA database lifecycle manager: HDBLCM. This tool is for example used to apply downloaded software updates.

For more information about software components, see 3.1, “SAP HANA software components” on page 32.

2.2 SAP S/4HANA overview

In 2015 SAP released SAP S/4HANA which is a Business Suite application built natively on the in-memory database SAP HANA to serve as a cornerstone of the digital enterprise. This fourth generation of business suite software from SAP is only available with the SAP HANA database underneath.

The tight integration of the application components with the database allows to integrate all mission-critical processes of an enterprise within one system and provides instant insight into a business.

SAP uses SAP Fiori as the user interface platform to ensure a simple and consistent experience for SAP users across products and end-user devices including tablets and mobile phones. SAP Fiori UX is designed to replace SAP UI and supports HTML5 and JavaScript.

SAP S/4HANA can be installed on-premises or consumed as a cloud-based offering. The on-premise edition is simply called SAP S/4HANA. The releases are named by the year and month they are generally available. The releases so far were 1511, 1610, 1709, 1809 and the next release will be 1909 (expected in September of 2019). The public cloud offerings are all named SAP S/4HANA Cloud.

As the time of writing, in July 2019, the following components are available with the SAP S/4HANA on-premise edition (release SAP S/4HANA 1809):

- SAP S/4HANA Enterprise Management
- SAP S/4HANA LoB Products (to enhance core functions of SAP S/4HANA Enterprise Management for specific lines of business)
- SAP S/4HANA LoB Products for specific industries (to enhance SAP S/4HANA Enterprise Management to provide industry specific benefits for certain lines of business)
- SAP S/4HANA Compatibility Packs

The following offerings can be integrated with the standard business functionality in S/4HANA Cloud (1905):

- SAP Analytics Cloud
- SAP Ariba Integration
- SAP Business Planning and Consolidation, version for SAP BW/4HANA
- SAP BusinessObjects Planning and Consolidation, version for SAP NetWeaver
- and many more

The on-premise edition follows a yearly release cycle while the cloud edition follows a quarterly release cycle.

More documentation on available features can be found:

https://help.sap.com/viewer/product/SAP_S4HANA_ON-PREMISE/1809.002/en-US

2.3 SAP HANA and SAP S/4HANA delivery models

This section provides an introduction of the various delivery models of SAP HANA, predefined (aka “appliance”), as a Tailored Datacenter Integration model and in the cloud.

SAP deployed SAP HANA that combines software and hardware, which is frequently referred to as the *SAP HANA appliance*. Just like with the SAP NetWeaver Business Warehouse Accelerator SAP partners with several hardware vendors to provide the infrastructure that is needed to run the SAP HANA software. Lenovo partners with SAP to provide an integrated solution.

Over the last several years, SAP gained more experience with running SAP HANA in production environments, so an additional delivery model has been implemented, which is known as *Tailored Data Center Integration* (TDI). TDI aims to integrate clients’ hardware from different vendors. Both approaches are described briefly in this chapter. All further chapters of this book do not make a distinction between “appliance” or TDI. Customers should receive a mature, robust solution to support the necessary business applications based on SAP HANA.

Please be aware that TDI is not a solution with a lower quality. The solutions are both fully supported by SAP and the hardware vendor.

SAP S/4HANA uses SAP HANA as the underlying database and uses the same infrastructure building blocks used with other SAP HANA deployments. In the following section we only use the term SAP HANA for simplicity.

SAP HANA Hardware Directory:

SAP lists supported configurations on this website:

https://www.sap.com/dmc/exp/2014-09-02-hana-hardware/enEN/index.html

Please make sure, when you use this website, that you start with the tabs **Overview** and **Details** on the Home screen as depicted in Figure 2-2 on page 20.

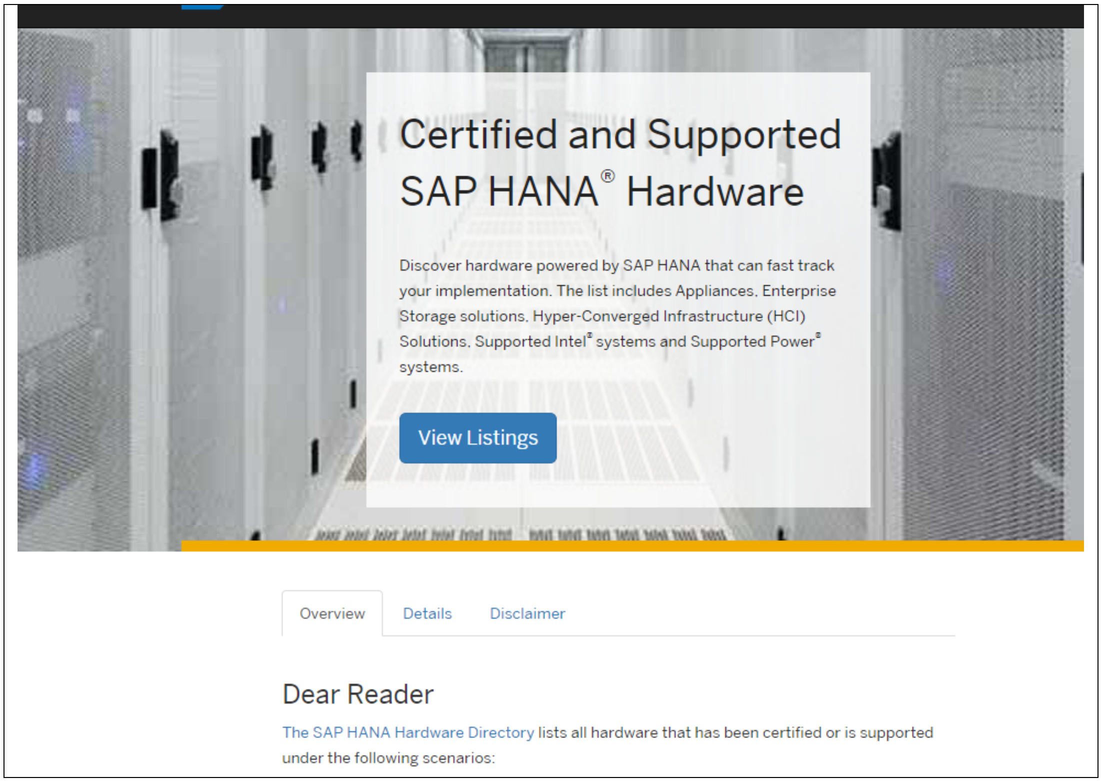


Figure 2-2 SAP HANA HW Directory home page

The individual listings are here (when you click “View Listings”):

`https://www.sap.com/dmc/exp/2014-09-02-hana-hardware/enEN/intel-systems.html`

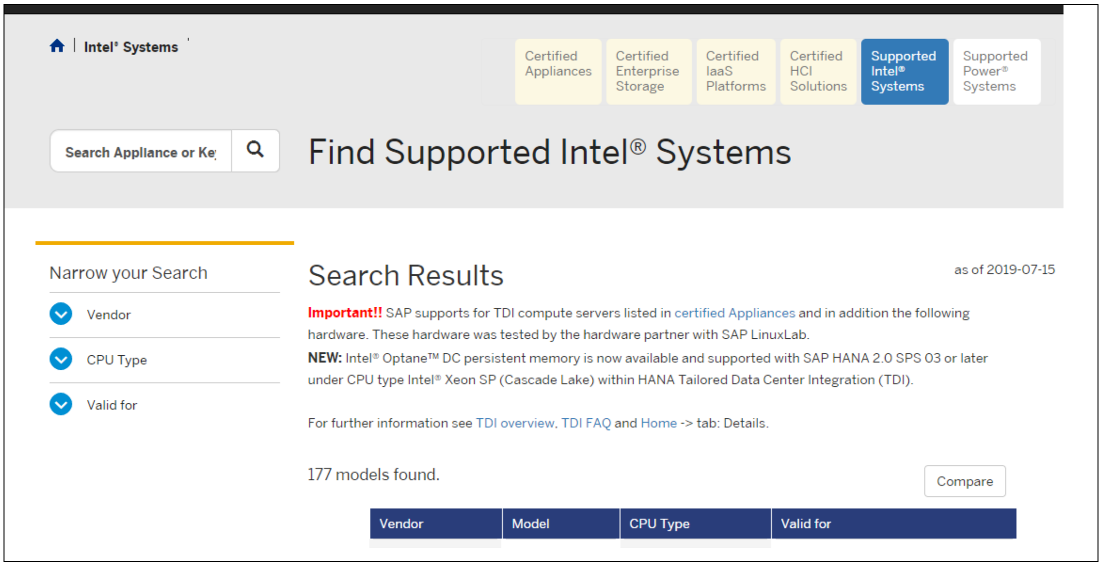


Figure 2-3 Individual Listings in the SAP HANA HW Directory

The **Listings** page has currently six tabs at the top with the following titles:

- Certified Appliances
- Certified Enterprise Storage
- Certified IaaS Platforms
- Certified HCI Solutions
- Supported Intel Systems
- Supported Power Systems

> **Note:** There is no difference in quality of performance or reliability for the tab called “Appliance” and “Supported Intel Systems”. Both are fully supported by SAP and Lenovo.

2.3.1 SAP HANA as an “appliance”

In the early days of SAP HANA it made a lot of sense to define simple T-Shirt size like bundles, which were easy to order and easy to deploy in customers data centers. However to speak of an “appliance” is raising expectations along the lines of “black box - don’t touch it”. SAP HANA is still a system with HW and SW, which needs to be implemented, monitored, kept alive and up to date, protected, secured etc., just like any other system in the datacenter.

Several hardware partners of SAP as well as hosting partners or cloud providers still offer predefined, pretested configurations to deploy SAP HANA.

Infrastructure for SAP HANA runs through a quality assurance process to ensure that certain performance requirements are met. These certified configurations are supported by SAP and the respective hardware partner. The configurations adhere to the following requirements and restrictions to provide a common platform across all hardware providers:

- Only certain Intel Xeon processors can be used.
- All configurations follow a certain main memory per core ratio, which is defined by SAP to balance CPU processing power and the amount of data that is processed.
- All configurations meet minimum redundancy and performance requirements for various load profiles. SAP tests for these requirements as part of the certification process.

- The capacity of the storage devices that are used in the configurations must meet certain criteria that are defined by SAP.
- The networking capabilities of the configurations must include 10 Gb 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. The hardware partners develop an infrastructure architecture for SAP HANA, which adds differentiating features to the solution. For more information about the benefits of the Lenovo solution, see Chapter 5, “Components of the Lenovo Solution for SAP HANA” on page 59.

2.3.2 SAP HANA Tailored Data Center Integration

To allow for an existing infrastructure to be integrated and reused when SAP HANA is deployed, clients can follow the Tailored Data Center Integration (TDI) delivery model. Storage and networks that fulfill certain criteria can be used to run SAP HANA. Among others, these criteria include storage and network performance. TDI evolved in multiple phases over the years.

With respect to the memory to core ratio, SAP loosened the restrictions, which were given in the early days of SAP HANA. This is what is frequently referred to as TDI Phase 5 (workload driven sizing). This applies for the Intel processor types Intel Xeon EX E7 systems and the Intel Xeon Scalable Processor family onwards. This includes also the systems, which use persistent memory technology (DCPMM).

To leverage the possibilities, it is required that a thorough sizing exercise is conducted for the individual workload. This is true for both green field and brown field situations. The sizing process is described in Chapter 2.4.1, “Sizing process” on page 24.

Details:

- System memory and processor sizing are fine-tuned for the specific customer workload
- SAP HANA hardware partners translate the sizing requirements (SAPS for CPU, RAM, disk I/O and disk capacity) into customer- tailored system configurations using a wide range of CPUs (including lower-end CPUs, not only high-end CPUs)
- The resulting HANA TDI configurations will extend the choice of HANA system sizes; and customers with less CPU-intensive workloads may have bigger main memory capacity.

This leads to the following considerations for the Lenovo solution.

CPU:

Any CPU with a minimum of 8 cores from the CPU families Intel Xeon E5 v2/v3/v4 based 2-socket systems, Intel Xeon E7 v4 or Intel Xeon Platinum/Gold/Silver Generation 1 (aka under the codename Skylake) or 2 (aka under the codename Cascade Lake) can be used in the new solution configurations (see the Supported Intel Systems description on the ‘Details’ tab of the SAP HANA Hardware directory). This applies to legacy systems which you might already have in house or new Skylake or Cascade Lake based configurations.

Memory:

SAP requires a homogenous symmetric assembly of DIMMs and a maximum utilization of all DDR memory channels per processor. Memory modules come in different technologies, which are not compatible to each other. Mixing RDIMMs and LRDIMMs is not supported. Mixing 3DS RDIMMs with either RDIMMs or LRDIMMs is also not supported.

Figure 2-4 shows examples of possible memory configurations for SAP HANA in Xeon SP Gen 1 based systems.

Controller #	0						1						Capacity (GB) per CPU	Capacity (GB) with #CPUs			
Channel #	0		1		2		0		1		2						
DIMM slot #	1	2	3	4	5	6	7	8	9	10	11	12		2	4	6	8
16GB	16		16		16		16		16		16		192	192	384	576	768
16GB	16	16	16	16	16	16	16	16	16	16	16	16	192	384	768	1152	1536
16/32GB	32	16	32	16	32	16	32	16	32	16	32	16	288	576	1152	1728	2304
32GB	32		32		32		32		32		32		192	384	768	1152	1536
32GB	32	32	32	32	32	32	32	32	32	32	32	32	384	768	1536	2304	3072
64GB	64		64		64		64		64		64		384	768	1536	2304	3072
32/64GB	64	32	64	32	64	32	64	32	64	32	64	32	576	1152	2304	3456	4608
64GB	64	64	64	64	64	64	64	64	64	64	64	64	768	1536	3072	4608	6144
128GB	128		128		128		128		128		128		768	1536	3072	4608	6144
64/128GB	128	64	128	64	128	64	128	64	128	64	128	64	1152	2304	4608	6910	9216
128GB	128	128	128	128	128	128	128	128	128	128	128	128	1536	3072	6144	9216	12288

Figure 2-4 Example of possible memory configurations

For support of more than 768 GB of main memory per CPU socket, it is required to have the Intel M CPUs installed. When ordering the system please keep potential upgrades and higher requirements in mind.

Storage

SAP HANA storage requirements are described in detail in this document:

https://www.sap.com/documents/2015/03/74cdb554-5a7c-0010-82c7-eda71af511fa.html

The solutions are fully supported both by SAP and the hardware partner.

Implementing SAP HANA by following the TDI model requires close collaboration between the client, SAP, and the vendor of the infrastructure element that is integrated. For more information about this SAP delivery model, see this website:

https://www.sap.com/documents/2017/09/e6519450-d47c-0010-82c7-eda71af511fa.html

Only certain components are eligible for integration. For more information about the list of certified enterprise storage for SAP HANA, see this website:

https://www.sap.com/dmc/exp/2014-09-02-hana-hardware/enEN/enterprise-storage.html

2.3.3 SAP HANA in the cloud

A third option is to deploy SAP HANA in the cloud. Several cloud providers have Infrastructure-as-a-Service (Iaas) or Platform-as-a-Service (Paas) offerings around the SAP HANA in-memory database.

> **Note:** The SAP application server needs to be close to the database server, just like with classic SAP architectures.

The SAP S/4HANA software stack is available as an on-premise solution or a cloud based offering hosted by SAP or other hyperscale vendors like AWS, Microsoft, Google or Alibaba.

This option is not covered further in this publication.

2.4 Identifying the right Lenovo solution for SAP HANA

In this section, we explain the methodology behind sizing Lenovo solutions for SAP use cases. Then, we introduce the concept of shortnames for the Lenovo solution for SAP HANA.

2.4.1 Sizing process

The following section explains how to best approach the difficult topic of sizing Lenovo servers for SAP HANA.

Sizing methodology

SAP sizing is translating business requirements and the according workload into hardware configurations. This must be an iterative process during the implementation project and relies on correct input values. There are multiple approaches and tools. The key results of a sizing exercise are resources like CPU capacity (frequently expressed in SAPS, which stands for SAP Application Performance Standard and describes system performance in a hardware-independent way), amount of main memory, and disk capacity and I/O performance needed.

Dependent on the customer scenario the sizing methodology for SAP HANA follows different approaches. SAP has published an SAP HANA sizing decision tree which is available online:

https://www.sap.com/documents/2016/10/aaa93b78-8f7c-0010-82c7-eda71af511fa.html

Table 2-1 SAP HANA sizing for different scenarios

Current situation	Sizing Method	Further information
Greenfield	SAP Quicksizer for HANA	Available online: http://www.sap.com/sizing
Brownfield OLTP (Business Suite on HANA and S/4HANA)	SAP report on the source system	Note 1872170 - *Business Suite on HANA and S/4HANA sizing report*
Brownfield OLAP	SAP report on the source system	Note 2296290 - *Sizing Report for BW on HANA*

The majority of the customers today already have an SAP landscape established so in most customer engagements one of the brown field approaches has to be used. In some documents “brownfield sizing” is also called “productive sizing”. This means the same.

Greenfield using SAP Quicksizer

The SAP Quicksizer tool offers the possibility to estimate the workload of a customer scenario. You either provide the number of expected business users or business transactions and the tool calculates a result (SAPS for the database and application layer, amount of main memory, I/O requirements, and necessary disk space). This result needs to be compared with the capacity of server models and architecture guidelines by Lenovo. The Lenovo Expert Technical Sales (LETS) teams worldwide will help in determining the necessary data center infrastructure.

It is important to involve either SAP or consulting partners working on the customer project to define the input to the Quicksizer as precise as possible. A good understanding of the business processes and the application behavior is required for a successful sizing exercise.

Brownfield for transactional applications running on SAP HANA

Follow the instructions in Note 1872170, *Business Suite on HANA and S/4HANA sizing report,* and run the attached report. It can run on any database and operating system combination and the result is the memory footprint required for the SAP HANA system. See Example 2-1 for the sizing report output.

Example 2-1 Example of an S/4HANA sizing report for transactional workload

```
-------------------------------------------------------------------------------
| SIZING RESULTS IN GB                                                        |
|-----------------------------------------------------------------------------|
| Based on the selected table(s), the anticipated maximum requirement are     |
|                                                                             |
| for S/4HANA:                                                                |
| - Memory requirement for the initial installation                 3,727.8  |
| - Net data size on disk for the initial installation              2,118.6  |
|                                                                             |
| - Estimated memory requirement after data clean-up                2,639.7  |
| - Estimated net data size on disk after data clean-up             1,725.9  |
|                                                                             |
| Other possible additional memory requirement:                               |
| - during the transition to S/4HANA (See FAQ)                         36.0  |
| - for an upgrade shadow instance                                     69.7  |
-------------------------------------------------------------------------------
 Check the FAQ document attached to SAP Note 1872170 for explanations on how
 to interpret the sizing terms and calculations.

 Sizing report:                                                /SDF/HDB_SIZING
....
```

Since 2018, the Quicksizer result provides also information on CPU requirements in SAPS. This leads to more custom made configurations specifically designed for the real customer workload. In a brownfield scenario, an expert analysis determines the SAPS capacity as well as average and peak CPU utilization of the source system over an extended period of time. With this information it is possible to derive a value for SAPS consumed on any database and any operating system. SAP provides a rule of thumb to multiply this figure by 3 to get the result in SAPS required for the SAP HANA system. See Example 2-2 below for all steps required.

Example 2-2 SAP ECC on a legacy database and operating system combination

CPU utilization: 40% average, 60% peak

SAPS capacity of the legacy server: 75.000 SAPS

SAPS consumed (peak): 75.000 x 0.6 = 45.000 SAPS

Target CPU utilization on SAP HANA: 65% → 45.000 / 0.65 = 69.230 SAPS

Required SAPS for ECC on HANA: 69.230 x 3 = 207.700 SAPS

Those calculations are valid as long as the workload of the customer does not change with the move from the legacy environment to SAP HANA. If the customer adds functionality or plans to use for example the functionality of embedded analytics on an SAP S/4HANA system the hardware configuration will increase accordingly.

Brownfield for analytical applications running on SAP HANA

If the application area is an SAP Business Warehouse then the document attached to Note 2296290, *Sizing Report for BW on HANA,* needs to be used. The report is designed to run on the legacy system planned to migrate onto SAP HANA in the future. It works in a similar way as described for transactional systems and provides an estimation of necessary memory, a disk space recommendation as well as CPU requirements in SAPS. The SAP Note also contains an attachment with detailed instructions how to deal with the result.

Regarding processing capacity, the output of the report also talks about servers “Class L, M and S”.

In the past SAP required a fixed memory to core ratio for analytical systems. With the experience from real live systems and workloads at customer implementations, this could be adjusted. Dependent on the workload scenario, the result of the sizing report recommends the class of CPU power.

- CLASS L (100% CPU required)
  - SAPS close to the CPU power of a machine with standard configuration (aka: core:memory ratio)
- CLASS M (50% CPU required)
  - Half of SAPS for CLASS L
- CLASS S (25% CPU
  - Half of SAPS for CLASS M

Intel Optane DC Persistent Memory and sizing considerations

DCPMM is only designed for SAP HANA under the TDI umbrella. This does not influence the level of support or the quality of the final solutions supporting the SAP HANA based applications. For the basics around Intel DCPMM please refer to 1.4, “Principles for persistent memory” on page 9.

A detailed determination of the workload is required with either the SAP Quicksizer or the sizing reports from SAP as described above. Especially, because the boundary of the main store and the rest of the data is fixed with the usage of PMEM and DRAM. In DRAM only systems this boundary can float.

Additionally SAP provides an SQL statement, which can be executed only on SAP applications, which are based on SAP HANA 2.0 already. This is not applicable for SAP systems which are based on classic DB/OS combinations.

The query provides a snapshot view. It is recommended to run the query multiple times with diverse load situations.

```
/*======SANPSHOT VIEW=========*/
select ' ' as host, 0 as NR , 'Run this Query in Systems without Persistent
Memory only' as Category, ''as "TYPE", '0' as FIGURE from (select top 1 *
from "PUBLIC"."M_INIFILE_CONTENTS" where section = 'persistence' and key =
'basepath_persistent_memory_volumes' and key != '')
union all
/*======TOTAL PHYSICAL MEMORY VIEW=========*/
select host, 1 as NR , 'TOTAL Physical Memory in HOST (GB)' as Category,
```

```
'MEMORY' as "TYPE" ,round(value/1024/1024/1024,2) as FIGURE from
m_host_information where key = 'mem_phys'
union all
/*======POTENTIAL PERSISTENT_MEMORY VIEW=========*/
select host, 2 as NR,
'Used Persistent Memory Part (GB)' as Category, 'PM'as TYPE,
round(sum(exclusive_size_in_use) /1024/1024/1024, 2) as "FIGURE"
from m_heap_memory
where
(category like 'Pool/ColumnStore%/Main/Compressed/Cluster' or
category like 'Pool/ColumnStore%/Main/Compressed/Indirect' or
category like 'Pool/ColumnStore%/Main/Compressed/Lrle' or
category like 'Pool/ColumnStore%/Main/Compressed/Prefix' or
category like 'Pool/ColumnStore%/Main/Compressed/Rle' or
category like 'Pool/ColumnStore%/Main/Compressed/Sparse' or
category like 'Pool/ColumnStore%/Main/Dictionary/RoDict' or
category like 'Pool/ColumnStore%/Main/Dictionary/RoDictCompletePages' or
category like 'Pool/ColumnStore%/Main/Dictionary/ValueDict' or
category like 'Pool/ColumnStore%/Main/Index/Block' or
category like 'Pool/ColumnStore%/Main/Index/MinMax' or
category like 'Pool/ColumnStore%/Main/Index/MinMax' or
category like 'Pool/ColumnStore%/Main/Index/RangedBlock' or
category like 'Pool/ColumnStore%/Main/Index/ReverseHash' or
category like 'Pool/ColumnStore%/Main/Index/Single' or
category like 'Pool/ColumnStore%/Main/HashKey' or
category like 'Pool/ColumnStore%/Main/MultiValue' or
category like 'Pool/ColumnStore%/Main/Rowid' or
category like 'Pool/ColumnStore%/Main/Uncompressed' or
category like 'Pool/Text/AEText/termset_container' or
category like 'Pool/SerializedObject')
group by host
/*======DRAM_MEMORY =========*/
UNION ALL
select host, 3 as NR,
'Used DRAM Part (GB)' as Category, 'DRAM'as MEM_TYPE,
round(sum(exclusive_size_in_use) /1024/1024/1024, 2) as FIGURE
from m_heap_memory
where not
(category like 'Pool/ColumnStore%/Main/Compressed/Cluster' or
category like 'Pool/ColumnStore%/Main/Compressed/Indirect' or
category like 'Pool/ColumnStore%/Main/Compressed/Lrle' or
category like 'Pool/ColumnStore%/Main/Compressed/Prefix' or
category like 'Pool/ColumnStore%/Main/Compressed/Rle' or
category like 'Pool/ColumnStore%/Main/Compressed/Sparse' or
category like 'Pool/ColumnStore%/Main/Dictionary/RoDict' or
category like 'Pool/ColumnStore%/Main/Dictionary/RoDictCompletePages' or
category like 'Pool/ColumnStore%/Main/Dictionary/ValueDict' or
```

```
category like 'Pool/ColumnStore%/Main/Index/Block' or
category like 'Pool/ColumnStore%/Main/Index/MinMax' or
category like 'Pool/ColumnStore%/Main/Index/MinMax' or
category like 'Pool/ColumnStore%/Main/Index/RangedBlock' or
category like 'Pool/ColumnStore%/Main/Index/ReverseHash' or
category like 'Pool/ColumnStore%/Main/Index/Single' or
category like 'Pool/ColumnStore%/Main/HashKey' or
category like 'Pool/ColumnStore%/Main/MultiValue' or
category like 'Pool/ColumnStore%/Main/Rowid' or
category like 'Pool/ColumnStore%/Main/Uncompressed' or
category like 'Pool/Text/AEText/termset_container' or
category like 'Pool/SerializedObject')
group by host
union all
/*======SHARED_MEMORY =========*/
select host, 4 as NR, 'Allocated Shared Memory__must be added to DRAM (GB)'
as Category, 'DRAM'as MEM_TYPE,
round(sum(shared_memory_allocated_size)/1024/1024/1024, 2) as FIGURE
from M_SERVICE_MEMORY group by host
union all
/*====== CPU VIEW=========*/
select host, 5 as NR, 'Number of CPU Sockets' as Category, 'CPU' as
TYPE,value as FIGURE from m_host_information where key = 'cpu_sockets'
union all
select host, 6 as NR, 'Number of CPU Cores' as Category, 'CPU' as
TYPE,value as FIGURE from m_host_information where key = 'cpu_cores'

order by 1,2
```

Example output of the sql query:

HOST	NR	CATEGORY	TYPE	FIGURE
server1	1	TOTAL Physical Memory in HOST (GB)	MEMORY	1.007,78
server1	2	Used Persistent Memory Part (GB)	PMEM	681,33
server1	3	Used DRAM Part (GB)	DRAM	197,26
server1	4	Allocated Shared Memory__must be added to DRAM (GB)	DRAM	1,86
server1	5	Number of CPU Sockets	CPU	4
server1	6	Number of CPU Cores	CPU	72
server2	1	TOTAL Physical Memory in HOST (GB)	MEMORY	1.007,78
server2	2	Used Persistent Memory Part (GB)	PMEM	680,33
server2	3	Used DRAM Part (GB)	DRAM	60,12
server2	4	Allocated Shared Memory__must be added to DRAM (GB)	DRAM	10,28
server2	5	Number of CPU Sockets	CPU	4
server2	6	Number of CPU Cores	CPU	72

Some CS table column types have a fix amount of DRAM overhead (helper data structures). The query therefore is not 100% accurate. The “Used Persistent Memory Part” is generally lower than the actual required PMEM at a healthy buffer of additional PMEM shall be added to make sure that it is not to tightly sized. Also, fragmentation effects must be considered as the PM namespaces are partitioned and bin packing of the files should be considered. Additionally, with some head room available per PMEM namespace, the CPU and the PM modules have more flexibility to cope with bad blocks without creating PMEM pressure.

More information can be found in SAP Note

```
2786237 - Sizing SAP HANA with Persistent Memory
```

Specific rules for SAP BW/4HANA

1. If the memory requirements per server node (as per Sizing Report) are less than 3 TB, a DRAM only configuration is recommended.

   Explanation: If server nodes with less than 3TB RAM use persistent memory, the risk of underestimating either DRAM or PMEM is very high, compared to server nodes with larger amounts of memory. Therefore, PMEM is not recommended for such small configurations.

2. If the CPU requirements analysis of the Sizing Report shows a "Class L" result, a DRAM only configuration is recommended (the term “Class L” is explained in “Brownfield for analytical applications running on SAP HANA” on page 26).

   Explanation: For "Class L" CPU requirements servers should have at most 768GB memory per CPU socket. A PMEM configuration with the smallest DIMM modules, however, results in at least 960GB per socket.

3. In a scale-out landscape, a configuration with the same amount of DRAM and PMEM (i.e. a 1:1 share) is recommended.

   Explanation: SAP HANA uses persistent memory to store the main indexes of column store tables, while delta indexes, row store tables and all kinds of temporary memory objects go to DRAM. In a scale-out landscape, the master node usually requires less persistent, but more volatile memory than the worker nodes. Since HANA relies on a symmetric scale-out landscape architecture where all nodes have exactly the same configuration, different memory configurations for master and worker nodes are not possible. Only a DRAM:PMEM ratio of 1:1 can satisfy the memory requirements for both master and worker.

   Please note that for scale-out landscapes, the first two rules must also be fulfilled, i.e. each server node must have at least 3TB, and the overall system must have CPU requirements below "Class L".

4. Single node (scale-up) systems with at least 20% "warm" data are recommended to use a 1:2 ratio of DRAM and PMEM.

Please check SAP Note 2813454 for any updates over time.

Application servers

In most cases SAP HANA is deployed in a 3-tier client server architecture which means the database layer is sized using above approaches. In a greenfield scenario, the Quicksizer will also provide recommendations for the application server layer. In the brownfield or migration scenario, either the existing application servers can be kept as-is and continued to be used or a straight forward calculation of SAPS and memory requirements from the existing system to the target platform needs to be done.

There is one exception to this rule. If an ABAP application instance is deployed on a physical system together with the SAP HANA database then an additive sizing approach is required. Details can be found in SAP Note 1953429.

Transfer the sizing result from SAP tools into a Lenovo configuration

Lenovo maintains a SAPS capacity ranking for the full portfolio of servers that are supported for usage in an SAP deployment. There is experienced personnel from Lenovo available to help in a sizing exercise in each geography of the world. Dependent on the functional and non-functional requirements of SAP customers, which can vary case by case quite significantly, the specific use case determines the individual hardware configuration.

> **Note:** Sizing is an iterative process. Over the course of an implementation project the input typically gets more and more precise. When a sizing is carried out for budgetary planning purposes it might look different than later on when the implementation has progressed and the business processes and volumes are more predictable.

3

Software components and data replication methods

This chapter describes the purpose of individual software components of the SAP HANA solution and introduces available replication technologies.

This chapter includes the following topics:

3.1 SAP HANA software components

The SAP HANA solution is composed of the main software components that are described in the following sections:

- 3.1.1, “SAP HANA database”
- 3.1.2, “SAP HANA client” on page 33
- 3.1.3, “SAP HANA landscape management structure” on page 34
- 3.1.4, “SAP host agent” on page 34
- 3.1.5, “Solution Manager Diagnostics agent” on page 34

The locations of these components are shown in Figure 3-1.

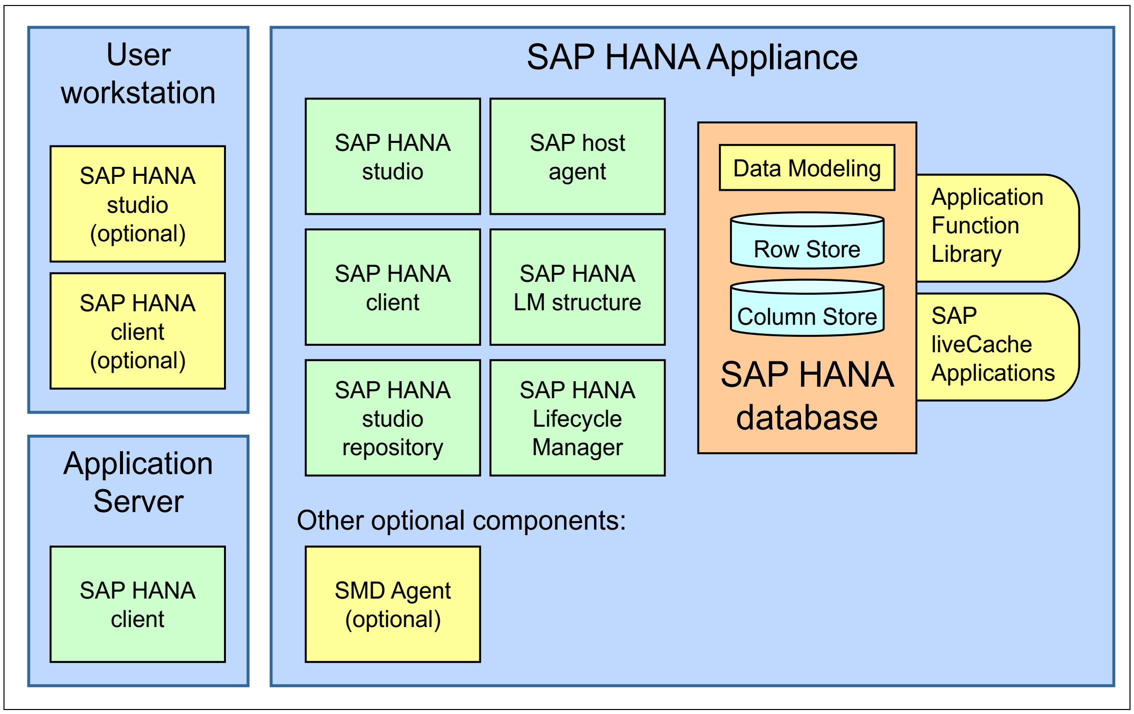


Figure 3-1 Distribution of software components that are related to SAP HANA

3.1.1 SAP HANA database

The SAP HANA database is the heart of the SAP HANA offering and the most important software component that runs on the SAP HANA appliance.

SAP HANA is an in-memory database that combines row-based and column-based database technology. All standard features that are available in other relational databases are supported (for example, tables, views, indexes, triggers, and SQL interface).

In addition to these standard functions, the SAP HANA database offers modeling capabilities that with which you can define in-memory transformation of relational tables into analytic views. These views are not materialized; therefore, all queries are providing real-time results that are based on the content of the underlying tables.

Another feature that extends the capabilities of the SAP HANA database is the SQLscript programming language, with which you can capture transformations that might not be easy to define by using simple modeling.

The SAP HANA database also can be integrated with external applications, such as an SAP application environment (for example, ERP). By using these possibilities, customers can extend their models by implementing existing statistical and analytical functions that were, for example, developed in ABAP.

For more information about the internal structures of the SAP HANA database, see Chapter 2, "SAP HANA and SAP S/4HANA overview" on page 15.

3.1.2 SAP HANA client

The SAP HANA client is a set of libraries that are used by external applications to connect to the SAP HANA database.

The following interfaces are available after the SAP HANA client libraries are installed:

- SQLDBC

  An SAP native database SDK that can be used to develop new custom applications that are working with the SAP HANA database.

- OLE DB for OLAP (ODBO) (available for Windows only)

  ODBO is a Microsoft driven industry standard for multi-dimensional data processing. The query language that is used with ODBO is the Multidimensional Expressions (MDX) language.

- Open Database Connectivity (ODBC)

  The ODBC interface is a standard for accessing database systems, which was originally developed by Microsoft.

- Java Database Connectivity (JDBC)

  JDBC is a Java based interface for accessing database systems.

The SAP HANA client libraries are delivered in 32-bit and 64-bit editions. It is important always to use the correct edition that is based on the architecture of the application that uses this client. The 32-bit applications cannot use 64-bit client libraries and vice versa.

To access the SAP HANA database from Microsoft Excel, you also can use a special 32-bit edition of the SAP HANA client that is called SAP HANA client package for Microsoft Excel.

The SAP HANA client is compatible with earlier versions; that is, the revision of the client must be the same or higher than the revision of the SAP HANA database.

The SAP HANA client libraries must be installed on every machine where connectivity to the SAP HANA database is required, including all servers and user workstations that are hosting applications that are directly connecting to the SAP HANA database (for example, SAP BusinessObjects Client Tools or Microsoft Excel).

Whenever the SAP HANA database is updated to a more recent revision, all clients that are associated with this database also must be upgraded. For more information about how to install the SAP HANA client, see the official SAP guide *SAP HANA Database - Client Installation Guide*, which is available at this website:

http://help.sap.com/hana_platform

3.1.3 SAP HANA landscape management structure

The SAP HANA landscape management (LM) structure (`lm_structure`) is an XML file that describes the software components that are installed on a server. The information in this file contains the following items:

- System ID (SID) of the SAP HANA system and host name
- A stack description, including the edition (depending on the license schema)
- Information about the SAP HANA database, including the installation directory
- Information about the SAP HANA studio repository, including its location
- Information about the SAP HANA client, including its location
- Information about the host controller

The LM structure description also contains revisions of individual components; therefore, it must be upgraded when the SAP HANA database is upgraded. Information that is contained in this file is used by the System Landscape Directory (SLD) data suppliers and by the SAP HANA HLM.

3.1.4 SAP host agent

The SAP host agent is a standard part of every SAP installation. In an SAP HANA environment, it is important in the following situations:

- Automatic update by using SAP HANA LM
- Remote starting and stopping of SAP HANA database instances

3.1.5 Solution Manager Diagnostics agent

SAP HANA can be connected to SAP Solution Manager 7.1, SP03 or higher.[1] The SMD provides a set of tools to monitor and analyze SAP systems, including SAP HANA. It also provides a centralized method to trace problems in all systems that are connected to an SAP Solution Manager system.

The SMD agent is an optional component, which can be installed on the SAP HANA appliance. It enables diagnostic tests of the SAP HANA appliance through SAP Solution Manager. The SMD agent provides access to the database logs and the file system, and collects information about the system's CPU and memory consumption through the SAP host agent.

For more information about how to deploy SMD agent, see the official SAP guide, *SAP HANA Update and Configuration Guide*, which is available at this website:

http://help.sap.com/hana_platform

[1] With monitor content update and more SAP notes for SP02

3.2 Data replication methods for SAP HANA

Data can be written to the SAP HANA database directly by a source application or replicated by using replication technologies.

The following replication methods are available for use with the SAP HANA database:

- Trigger-based replication

  This method is based on database triggers that are created in the source system to record all changes to monitored tables. These changes are then replicated to the SAP HANA database by using the SAP Landscape Transformation system.

- ETL-based replication

  This method uses an Extract, Transform, and Load (ETL) process to extract data from the data source, transform it to meet the business or technical needs, and load it into the SAP HANA database. The SAP BusinessObject Data Services application is used as part of this replication scenario.

- Extractor-based replication

  This approach uses the embedded SAP NetWeaver Business Warehouse (SAP NetWeaver BW) that is available on every SAP NetWeaver based system. SAP NetWeaver BW starts an extraction process by using available extractors and then redirects the write operation to the SAP HANA database instead of the local Persistent Staging Area (PSA).

- Log-based replication

  This method is based on reading the transaction logs from the source database and reapplying them to the SAP HANA database.

These replication methods are shown in Figure 3-2.

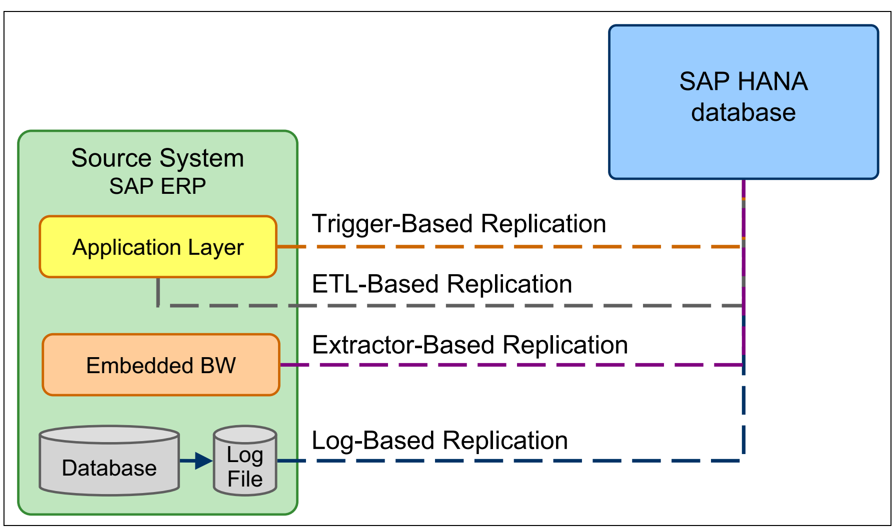


Figure 3-2 Available replication methods for SAP HANA

The following sections describe these replication methods for SAP HANA.

3.2.1 Trigger-based replication with SAP Landscape Transformation

SAP Landscape Transformation (SLT) replication is based on tracking database changes by using database triggers. All modifications are stored in logging tables in the source database, which ensures that every change is captured, even when the SLT system is not available.

The SLT system reads changes from source systems and updates the SAP HANA database. The replication process can be configured as real time (continuous replication) or scheduled replication in predefined intervals.

The SLT operates on the application level; therefore, the trigger-based replication method benefits from the database abstraction that is provided by the SAP software stack, which makes it database-independent. It also features extended source system release coverage, where supported releases start from SAP R/3 4.6C up to the newest SAP Business Suite releases.

The SLT also supports direct replication from database systems that are supported by the SAP NetWeaver platform. In this case, the database must be connected to the SLT system directly (as another database) and the SLT plays the role of the source system.

The replication process can be customized by creating ABAP routines and configuring their execution during the replication process. This feature allows the SLT system to replicate more calculated columns and to scramble existing data or filter-replicated data that is based on defined criteria.

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

The SLT 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, it is preferable to install the SLT Replication Server on a separate system to avoid a high replication load that affects the base system performance.

The SLT Replication Server is the ideal solution for all SAP HANA customers who need real-time (or scheduled) data replication from SAP NetWeaver based systems or databases that are supported by SAP NetWeaver.

3.2.2 Smart Data Integration and Smart Data Quality

SAP HANA Smart Data Integration and SAP HANA Smart Data Quality load data, in batch or real-time, into HANA (on premise or in the cloud) from a variety of sources using pre-built and custom adapters.

You deploy this method by installing a Data Provisioning Agent to house adapters and connect the source system with the Data Provisioning server, housed in the HANA system. You then create replication tasks, using WebIDE, to replicate data, or flowgraphs, using Application Function Modeler nodes, to transform and cleanse the data on its way to HANA. For more information about deploying the smart data integration, see the SAP HANA Smart Data Integration and SAP HANA Smart Data Quality Master Guide on the SAP Help Portal.

https://help.sap.com/viewer/d60a5abb34d246cdb4ab7a4f6b9e3c93/2.0_SPS04/en-US

3.2.3 ETL-based replication with SAP BusinessObjects Data Services

An ETL-based replication for SAP HANA can be set up by using SAP BusinessObjects Data Services, which is a full-featured ETL tool that gives customers the following maximum flexibility regarding the source database system:

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

SAP BusinessObjects Data Services provide several kinds of data quality and data transformation functions. Because of the rich feature set that is available, implementation time for the ETL-based replication is longer than for the other replication methods. SAP BusinessObjects Data Services offer integration with SAP HANA. SAP HANA is available as a predefined data target for the load process.

3.2.4 SAP HANA Smart Data Access

SAP HANA Smart Data Access (SDA) allows you to access remote data as if the data were stored in local tables in SAP HANA, without copying the data into SAP HANA.

This capability provides operational and cost benefits and supports the development and deployment of next-generation analytical applications requiring the ability to access, synthesize, and integrate data from multiple systems in real time.

In SAP HANA, you use linked databases or create virtual tables, which point to remote tables in different data sources, and then write SQL queries in SAP HANA that use these virtual tables. The SAP HANA query processor optimizes these queries by executing the relevant part of the query in the target database, returning the results of the query to SAP HANA, and then completing the operation. Physical data movement is not supported by SAP HANA SDA.

For a list of supported remote source databases and versions, see SAP Note 2600176 - Smart Data Access - Supported Remote Source Databases and Versions.

3.2.5 Extractor-based replication with Direct Extractor Connection

Extractor-based replication for SAP HANA is based on application logic that is available in every SAP NetWeaver system. The SAP NetWeaver BW package that is a standard part of the SAP NetWeaver platform can be used to run an extraction process and store the extracted data in the SAP HANA database.

This function requires some corrections and configuration changes to the SAP HANA database (import of delivery unit and parameterization) and on the SAP NetWeaver BW system as part of the SAP NetWeaver platform (implementing corrections by using an SAP Note or installing a support package and parameterization). Corrections in the SAP NetWeaver BW system ensure that extracted data is not stored in local Persistent Staging Area (PSA), but diverted to the external SAP HANA database.

The use of native extractors instead of the replication of underlying tables can bring certain benefits. Extractors offer the same transformations that are used by SAP NetWeaver BW systems, which can decrease the complexity of modeling tasks in the SAP HANA database.

This type of replication is not real time, and only the available features and transformation capabilities that are provided by a specific extractor can be used.

Replication by using Direct Extractor Connection (DXC) can be achieved in the following basic scenarios:

- By using the embedded SAP NetWeaver BW function in the source system

  SAP NetWeaver BW functions in the source system often are not used. After the implementation of the required corrections, the source system calls its own extractors and pushes data into the external SAP HANA database.

  The source system must be based on SAP NetWeaver 7.0 or higher. Because the function of a specific extractor is diverted into SAP HANA database, this extractor must not be in use by the embedded SAP NetWeaver BW component for any other purpose.

- By using an SAP NetWeaver BW to drive replication

  An SAP NetWeaver BW can be used to extract data from the source system and to write the result to the SAP HANA system.

  The release of the SAP NetWeaver BW system that is used must be at least SAP NetWeaver 7.0, and the specific extractor must not be in use for this particular source system.

- By using a dedicated SAP NetWeaver BW to drive replication

  The last option is to install a dedicated SAP NetWeaver system to extract data from the source system and store the result in the SAP HANA database. This option has a minimal effect on functions because no system is changed in any way. However, a new system is required for this purpose.

The current implementation of this replication technology allows for only one database schema in the SAP HANA database. The use of one system for controlling the replication of multiple source systems can lead to collisions because all source systems use the same database schema in the SAP HANA database.

3.2.6 Log-based replication with SAP Replication Server

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

Replication with the SAP Replication Server is fast and uses little processing power because of its closeness to the database system. However, this mode of operation makes this replication method highly database-dependent, and the source database system coverage is limited.[2] It also limits the conversion capabilities; therefore, replication with the SAP Replication Server supports only Unicode source databases. The SAP Replication Server cannot convert between code pages, and because SAP HANA works with Unicode encoding internally, the source database also must use Unicode encoding. Also, certain table types that are used in SAP systems are unsupported.

To set up replication with the SAP Replication Server, the definition and content of tables that are chosen to be replicated must be copied initially from the source database to the SAP HANA database. This initial load is done with the R3Load program, which is also used for database imports and exports. Changes in tables during initial copy operation are captured by the SAP Replication Server; therefore, no system downtime is required.

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

SAP recommends that you use trigger-based data replication by using the SAP Landscape Transformation Replicator.

3.2.7 Comparing the replication methods

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

- The trigger-based replication method with the SLT system provides real-time replication while supporting a wide range of source database systems. It can handle Unicode and non-Unicode databases and use proven data migration technology. It uses 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. For this reason, it is preferred for replication from SAP source systems.
- The ETL-based replication method is the most flexible; however, its flexibility comes at a cost because it has only near real-time capabilities. With its various possible data sources, advanced transformation, and data quality functions, it is the ideal choice for replication from non-SAP data sources.
- The extractor-based replication method offers reuse of transformation capabilities that are available in every SAP NetWeaver based system. This method can decrease the required implementation effort. However, this type of replication is not real time and is limited to capabilities that are provided by the available extractors in the source system.
- The log-based replication method with the SAP Replication Server provides the fastest replication from the source database to SAP HANA. However, it is limited to Unicode-encoded source databases and it does not support all table types that are used in SAP applications. It provides no transformation function, and the source database system support is limited.

These replication methods are compared in Figure 3-3.

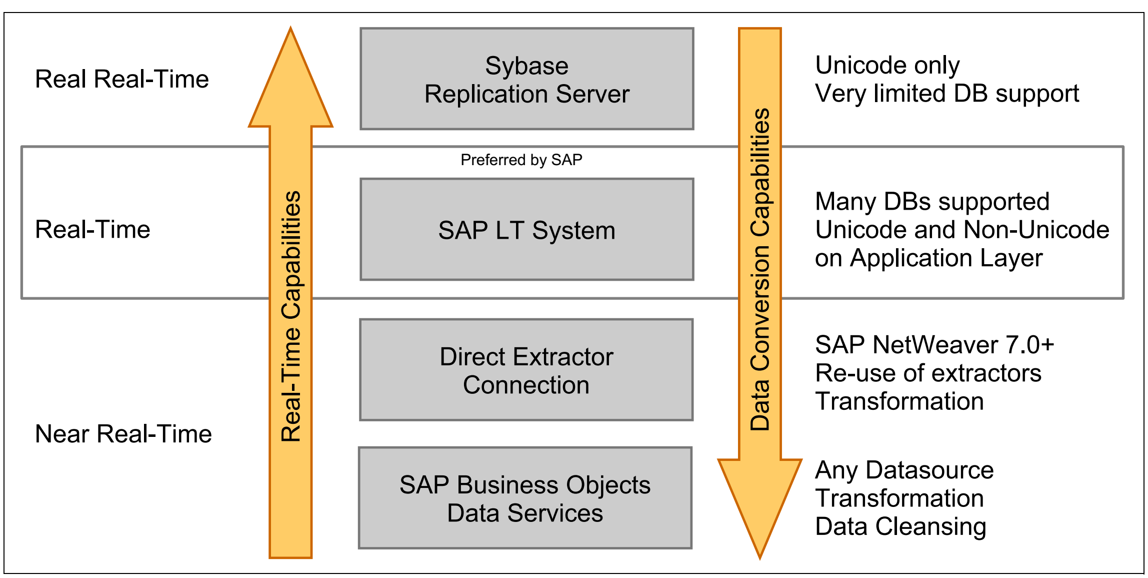


Figure 3-3 Comparison of the replication methods for SAP HANA

The replication method that you choose depends on the requirements. When real-time replication is needed to provide benefit to the business and the replication source is an SAP system, the trigger-based replication is the best choice.

Extractor-based replication might keep project costs down by reusing existing transformations.

ETL-based replication provides the most flexibility regarding data source, data transformation, and data cleansing options, but does not provide real-time replication.

4

SAP HANA integration scenarios

This chapter describes the different ways that SAP HANA can be implemented in client landscapes and highlights various aspects of such an integration. Whenever possible, real-world examples and related offerings are included.

This chapter includes the following topics:

4.1 Basic use case scenarios

The following classification of use cases was presented in the "EIM205 Applications powered by SAP HANA" session, already a long time ago, during the SAP TechEd 2011 event:

- Technology platform
- Operational reporting
- Accelerator
- In-memory products
- Next generation applications

These use case scenarios are shown in Figure 4-1.

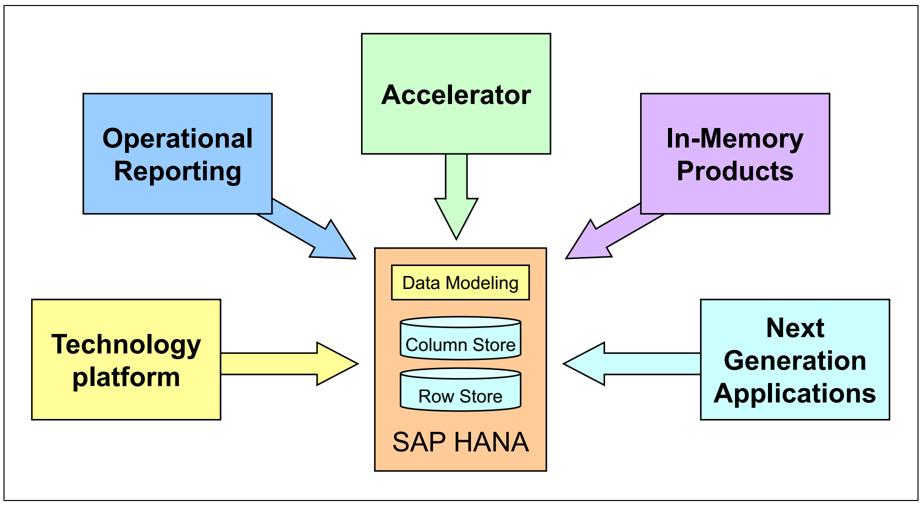


Figure 4-1 Basic use case scenarios that are defined by SAP in session EIM205

These five basic use case scenarios describe the ways that SAP HANA can be integrated. Each of these use case scenarios is described in this chapter.

SAP maintains a "SAP HANA Use Case Repository" with specific examples for how SAP HANA can be integrated. This repository is available at this website:

`http://www.experiencesaphana.com/community/resources/use-cases`

The use cases in this repository are divided into categories that are based on their relevance to a specific industry sector. It is a good idea to review this repository to find inspiration about how SAP HANA can be used in various scenarios.

4.2 SAP HANA as a technology platform

SAP HANA can be used even in non-SAP environments. The client can use structured and unstructured data that is derived from non-SAP application systems to use SAP HANA as a database platform. SAP HANA can be used to accelerate functions or to provide new functions that were, until now, unrealistic.

SAP HANA as a technology platform is shown in Figure 4-2.

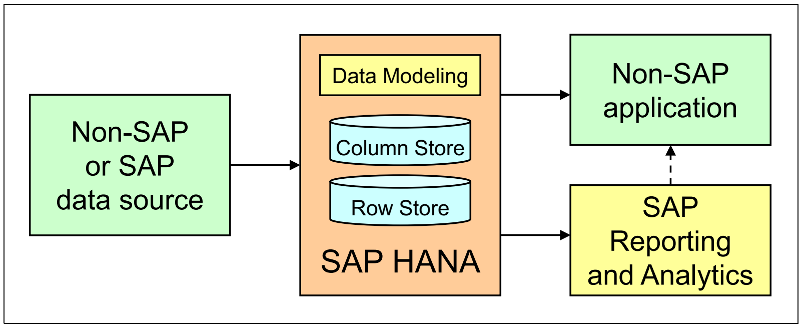


Figure 4-2 SAP HANA as technology platform

SAP HANA is not technologically dependent on other SAP products and can be used independently as the only one SAP component in the client's information technology (IT) landscape. However, SAP HANA can be easily integrated with other SAP products, such as SAP BusinessObjects BI platform for reporting or SAP BusinessObjects Data Services for Extract, Transform, and Load (ETL) replication, which gives clients the possibility to use only the components that are needed.

There are many ways that SAP HANA can be integrated into a client landscape, and it is not possible to describe all combinations. Software components around the SAP HANA offering can be seen as building blocks, and every solution must be assembled from the blocks that are needed in a particular situation. This approach is versatile and the number of possible combinations is growing because SAP constantly adds components to their SAP HANA-related portfolio.

Lenovo offers consulting services that help clients to choose the correct solution for their business needs.

4.2.1 SAP HANA data acquisition

There are multiple ways that data can flow into SAP HANA. This section describes the various options that are available, as shown in Figure 4-3.

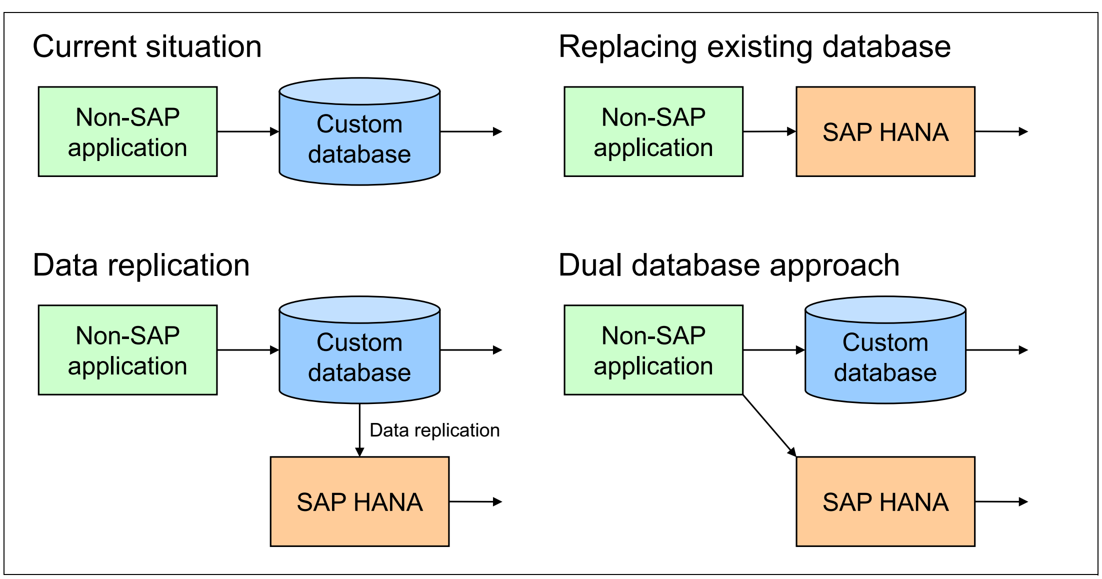


Figure 4-3 Examples of SAP HANA deployment options regarding data acquisition

The initial situation is displayed schematically in the upper left of Figure 4-3 on page 43. In this example, a client-specific non-SAP application writes data to a custom database that is slow and is not meeting client needs.

The other three examples in Figure 4-3 on page 43 show that SAP HANA can be deployed in such a scenario. These examples show that there is no single solution that is best for every client, but that each situation must be considered independently.

Each of these three solutions has the following advantages and disadvantages, which are highlighted to show the aspects of a specific solution that might need further consideration:

- Replacing the existing database with SAP HANA

  The advantage of this solution is that the overall architecture is not going to be significantly changed. The solution remains simple without the need to include more components. Customers might also save on license costs for the original database.

  A disadvantage to this solution is that the custom application must be adjusted to work with the SAP HANA database. If ODBC or JDBS is used for database access, this problem is insignificant. Also, the entire setup must be tested properly. Because the original database is being replaced, a certain amount of downtime is inevitable.

  Clients that are considering this approach must be familiar with the features and characteristics of SAP HANA, especially when certain requirements must be met by the database that is used (for example, special purpose databases).

- Populating SAP HANA with data replicated from the existing database

  This option integrates SAP HANA as a “side-car” database to the primary database and replicates required data by using one of the available replication techniques.

  An advantage of this approach is that the original solution is not touched and no downtime is required. Also, only the required subset of data must be replicated from the source database, which might allow customers to minimize acquisition costs because SAP HANA acquisition costs are linked directly to the volume of stored data.

  The need for implementing replication technology can be seen as the only disadvantage of this solution. Because data is delivered only into SAP HANA through replication, this component is a vital part of the whole solution. Customers that are considering this approach must be familiar with various replication technologies (including their advantages and disadvantages) as described in 3.2, “Data replication methods for SAP HANA” on page 35.

  Clients must also be aware that replication might cause extra load on the database because modified records must be extracted and then transported to the SAP HANA database. This aspect is highly dependent on the specific situation and can be addressed by choosing the proper replication technology.

- Adding SAP HANA as a second database in parallel to the existing database

  This option keeps the existing database in place while adding SAP HANA as a secondary database. The custom application then stores data in the original database and in the SAP HANA database.

  This option balances the advantages and disadvantages of the other options. A main prerequisite is the ability of the source application to work with multiple databases and the ability to control where data is stored. This prerequisite can be easily achieved if the source application was developed by the client and can be changed, or if the source application is going to be developed as part of this solution. If this prerequisite cannot be met, this option is not viable.

An advantage of this approach is that no replication is required because data is stored directly in SAP HANA as required. Customers also can store some of the records in both databases.

If data that is stored in the original database is not going to be changed and SAP HANA data is stored in both databases simultaneously, customers might achieve only minimal disruption to the existing solution.

A main disadvantage is the prerequisite that the application must work with multiple databases and store data according to the customer's expectations.

Customers considering this option must be aware of the abilities that are provided by the application that is delivering data into the existing database. Also, disaster recovery plans must be carefully adjusted, especially when consistency between both databases is seen as a critical requirement.

These examples must not be seen as a complete list of integration options for an SAP HANA implementation, but rather as a demonstration of how to develop a solution that matches client needs.

The SAP HANA database can be populated with data from different sources, such as SAP or non-SAP applications, and custom databases. These sources can feed data into SAP HANA independently, each by using a different approach or in a synchronized manner by using the SAP BusinessObjects Data Services.

4.2.2 SAP HANA as a source for other applications

The second part of integrating SAP HANA is to connect existing or new applications to run queries against the SAP HANA database, as shown in Figure 4-4.

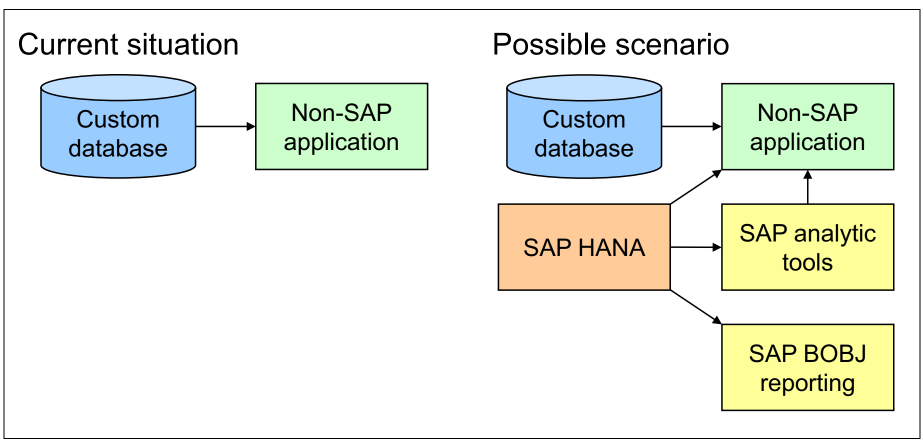


Figure 4-4 Example of SAP HANA as a source for other applications

The initial situation is shown on the left side of Figure 4-4. A customer-specific application runs queries against a custom database, which is a function that must be preserved.

A potential solution is shown on the right side of Figure 4-4. A customer-specific application runs problematic queries against the SAP HANA database. If the database is still part of the solution, specific queries that do not need acceleration can still be run against the original database.

Specialized analytic tools, such as SAP BusinessObjects Predictive Analysis, can be used to run statistical analysis on data that is stored in the SAP HANA database. This tool can run analysis directly inside the SAP HANA database, which helps avoid expensive transfers of massive volumes of data between the application and the database. The result of this analysis can be stored in SAP HANA, and the custom application can use these results for further processing; for example, to facilitate decision making.

SAP HANA can be easily integrated with products from the SAP BusinessObjects family. Therefore, these products can be part of the solution, which are responsible for reporting, monitoring critical key performance indicators (KPIs) that use dashboards, or for data analysis.

These tools can also be used without SAP HANA; however, SAP HANA is enabling these tools to process much larger volumes of data and still provide results in reasonable time.

4.3 SAP HANA for operational reporting

Operational reporting is playing a more important role. In today's economic environment, companies must understand how various events in the globally integrated world affect their business to make proper adjustments to counter the effects of these events. Therefore, the pressure to minimize the delay in reporting is becoming higher. An ideal situation is to use a real-time snapshot of current situations within seconds of a request.

Concurrently, the amount of data that is being captured grows every year. More information is collected and stored at more detailed levels. All of these issues make operational reporting more challenging because huge amounts of data must be processed quickly to produce the preferred result.

SAP HANA is a perfect fit for this task. Required information can be replicated from transactional systems into the SAP HANA database and then processed faster than directly on the source systems.

The following use case is often referred to as a data mart or *side-car approach* because SAP HANA sits by the operational system and receives the operational data (often only an excerpt) from this system by using replication.

You find many systems in a typical SAP-based application landscape today, such as SAP ERP, SAP CRM, SAP SCM, and other, possibly non-SAP, applications. All of these systems contain loads of operational data, which can be used to improve business decision making by using business intelligence technology. Data that is used for business intelligence purposes can be gathered on a business unit level by using data marts or on an enterprise level with an enterprise data warehouse, such as the SAP Business Warehouse (SAP BW). ETL processes feed the data from the operational systems into the data marts and the enterprise data warehouse.

A typical landscape is shown in Figure 4-5.

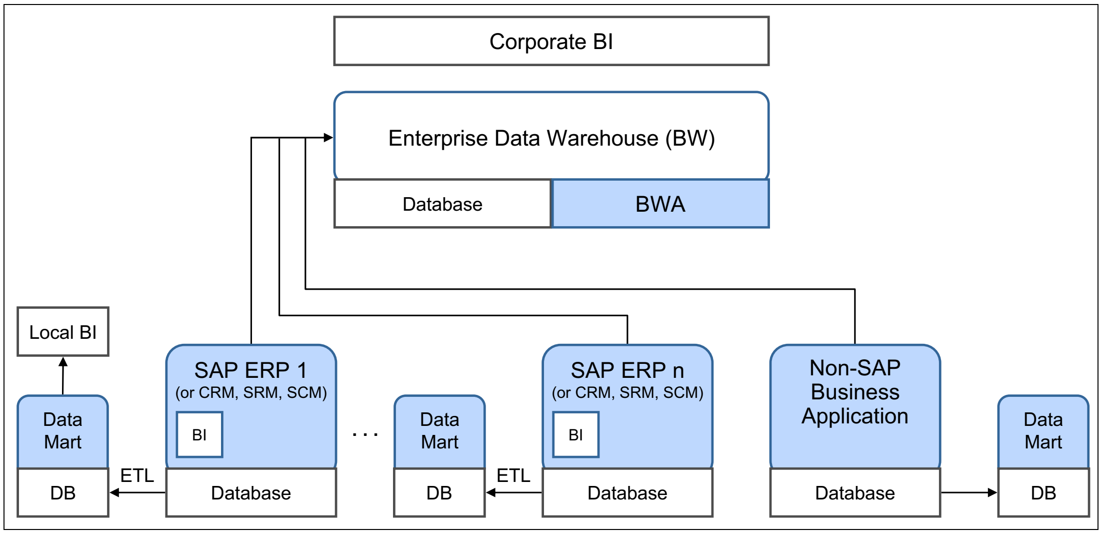


Figure 4-5   Typical view of an SAP-based application landscape

With the huge amount of data that is collected in an enterprise data warehouse, response times of queries for reports or navigation through data can become an issue, which generates new requirements for the performance of such an environment. To address these requirements, SAP introduced the SAP Business Warehouse Accelerator (SAP BW Accelerator), which speeds up queries and reports in the SAP BW by using in-memory technology. Although being a perfect fit for an enterprise data warehouse holding huge amounts of data, the combination of SAP BW and SAP BW Accelerator is not always a viable solution for relatively small data marts.

With the introduction of SAP HANA 1.0, SAP provided an in-memory technology that supports Business Intelligence (BI) at a business unit level. SAP HANA combined with business intelligence tools, such as 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 4-6 shows such a landscape with the local data marts replaced by SAP HANA.

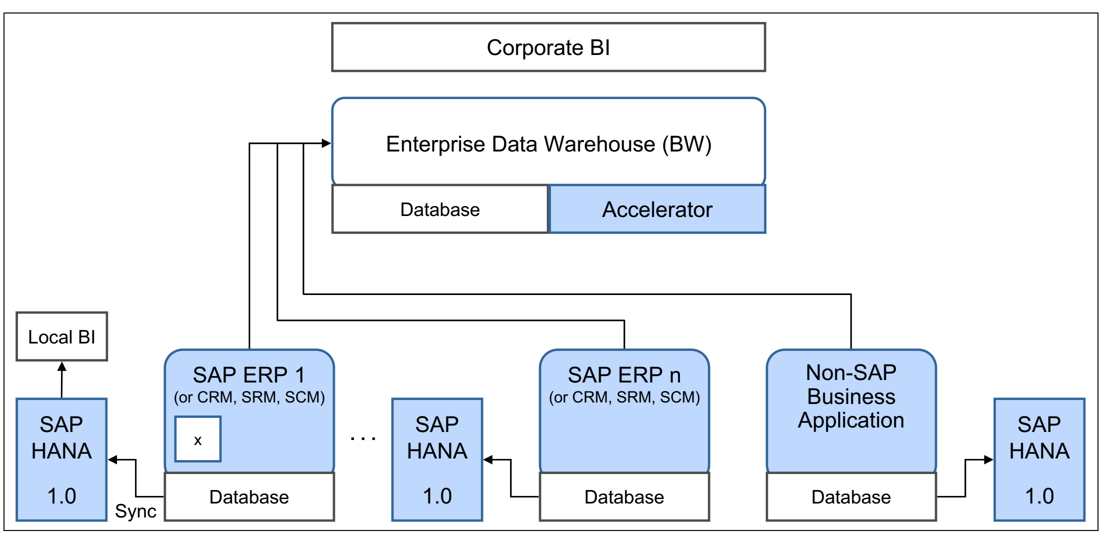


Figure 4-6   SAP vision after the introduction of SAP HANA

BI functionality is provided by an SAP BusinessObjects BI tool, such as the SAP BusinessObjects Explorer, which communicates with the SAP HANA database through the BI Consumer Services (BICS) interface.

This use case scenario is for products from the SAP Business Suite, where SAP HANA acts as a foundation for reporting on large volumes of data.

The role of SAP HANA in an operational reporting use case scenario is shown in Figure 4-7.

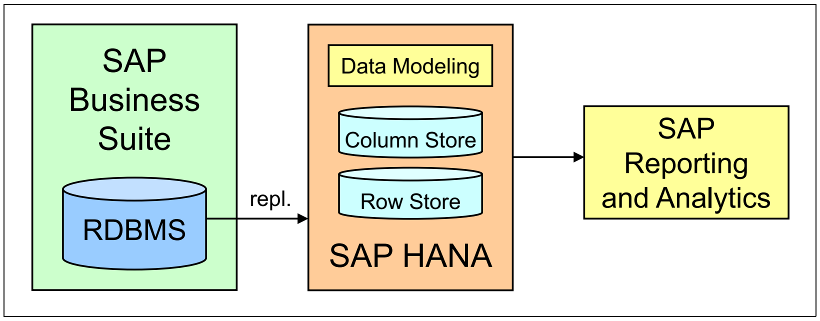


Figure 4-7 SAP HANA for operational reporting

The first step in the process is the replication of data into the SAP HANA database, which often originates from the SAP Business Suite. However, some solution packages are also built for non-SAP data sources.

There are instances in which source systems must be adjusted by implementing modifications or by performing specific configuration changes.

Data often is replicated by using the SAP Landscape Transformation replication; however, other options, such as replication by using SAP Smart Data Access (SDA), SAP BusinessObjects Data Services or SAP HANA Direct Extractor Connection (DXC), also are possible. The replication technology often is chosen as part of the package design and cannot be changed easily during implementation.

A list of tables to replicate (for SAP Landscape Transformation replication) or transformation models (for replication by using Data Services) are part of the package.

SAP HANA is loaded with models (views) that are static (designed by SAP and packaged) or automatically generated based on customized criteria. These models describe the transformation of source data into the resulting column views. These views are then used by SAP BusinessObjects BI 4.0 reports or dashboards that are delivered as final products or pre-made templates that can be finished as part of implementation process.

Some solution packages are based on more components (for example, SAP BusinessObjects Event Insight). If required, extra content that is specific to these components can also be part of the solution package.

Individual use cases, required software components, prerequisites, and configuration changes (including overall implementation processes) are properly documented and attached as part of the delivery.

Solution packages can contain the following items:

- SAP BusinessObjects Data Services Content (data transformation models)
- SAP HANA Content (exported models with attribute views and analytic views)
- SAP BusinessObjects BI Content (prepared reports and dashboards)
- Transports and ABAP reports (adjusted code to be implemented in source system)
- Content for other software components, such as SAP BusinessObjects Event Insight and Sybase Unwired Platform
- Documentation

4.4 SAP HANA as an accelerator

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

SAP HANA also can be used to accelerate processes in SAP Business Suite systems, even for those systems that are not yet released to be running directly on the SAP HANA database.

Some SAP systems are processing large amounts of records that must be filtered or aggregated based on specific criteria. Results are then used as inputs for all dependent activities in a specific system.

In the case of large data volumes, the running time can be unacceptable (in number of hours). Such workloads can easily run several hours, which can cause unnecessary delays. Currently, these tasks often are processed overnight as batch jobs.

SAP HANA as an accelerator can help decrease this running time.

This use case scenario is shown in Figure 4-8.

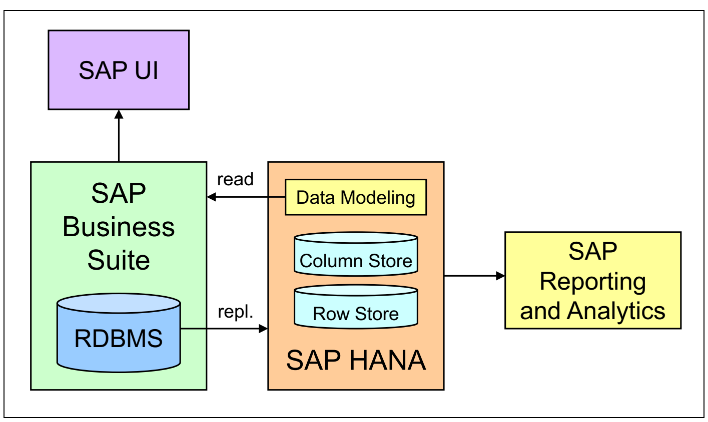


Figure 4-8 SAP HANA as an accelerator

The accelerated SAP system must meet specific prerequisites. Before this solution can be implemented, installation of specific support packages or implementation of SAP Notes might be required, which introduces the necessary code changes in the source system.

The SAP HANA client must be installed on a specific server, and the SAP kernel must be adjusted to support direct connectivity to the SAP HANA database.

As a next step, replication of data from the source system is configured. Each specific use case features a defined replication method and a list of tables that must be replicated. The most common method is the SAP Landscape Transformation replication. However, some solutions offer alternatives. For example, for the SAP CO-PA Accelerator, replication can also be performed by an SAP CO-PA Accelerator-specific ABAP report in the source system.

The source system is configured to have direct connectivity into SAP HANA as the secondary database. The required scenario is configured according to the specifications and then activated. During activation, the source system automatically deploys the required column views into SAP HANA and activates new ABAP code that was installed in the source system as the solution prerequisite. This new code can run and then use queries against the SAP HANA database, which leads to shorter execution times.

Because SAP HANA is populated with valuable data, it is easy to extend the accelerator use case by adding operational reporting functions. More (often optional) content is delivered for SAP HANA and for SAP BusinessObjects BI 4.0 client tools, such as reports or dashboards.

4.5 SAP products that are running on SAP HANA

Another way that SAP HANA can be deployed is to use SAP HANA as the primary database for selected products.

SAP BW that is running on SAP HANA was available since April 2012. The SAP ERP Central Component (SAP ECC) that is running on HANA was announced in early 2013, and the remaining products of the SAP Business Suite family became available in the second half of 2013. In 2015 SAP S/4HANA was announced, which is the new SAP Business Suite 4 SAP HANA and optimized to run on SAP HANA only.

One significant advantage of running products to use SAP HANA as the primary database is the minimal disruption to the existing system. Almost all functions, customizations, and (with SAP BW) client-specific modeling, are preserved because application logic that is written in ABAP is not changed. From a technical perspective, the SAP HANA conversion is similar to any other database migration.

In 2017 SAP introduced the new, rewritten SAP Business Warehouse for HANA (SAP BW/4HANA) built entirely on SAP HANA.

In parallel SAP S/4HANA is the go-to platform by SAP which serves as the cornerstone of the digital core concept by SAP. The application combines both transactional, classic business applications based solely on SAP HANA and incorporating certain analytical disciplines like real time reporting.

Figure 4-9 on page 51 shows an overview of the different solutions.

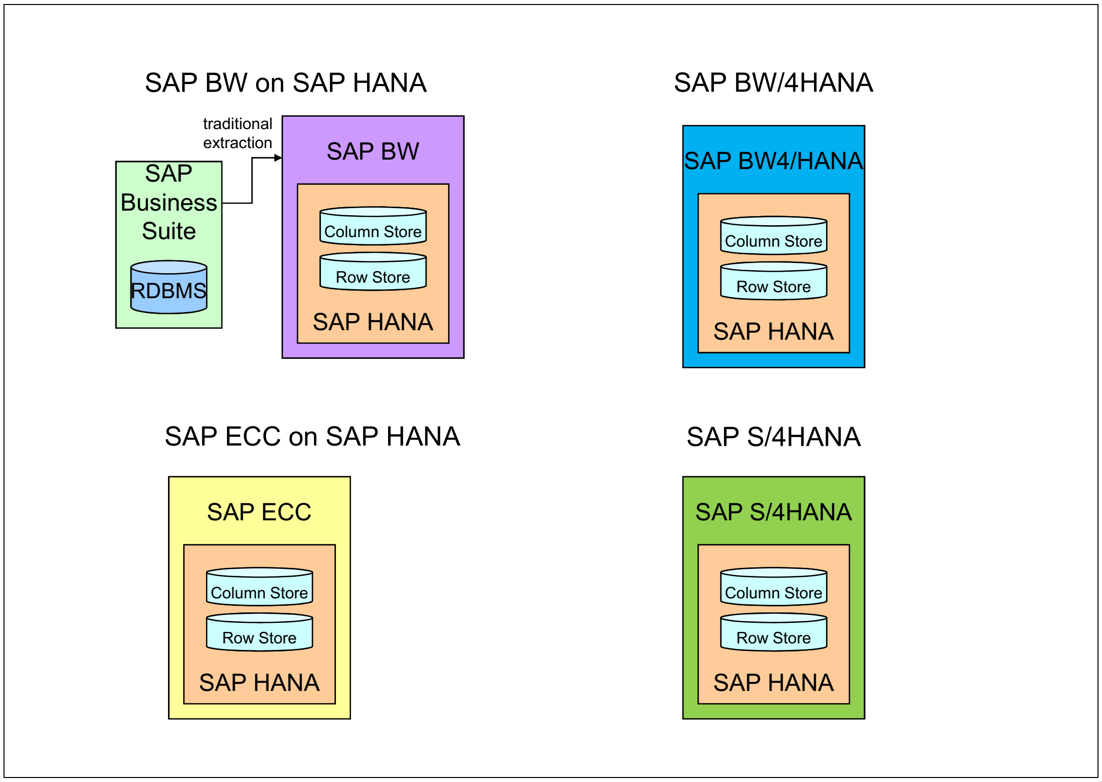


Figure 4-9 SAP Applications using SAP HANA as the database

4.5.1 SAP Business Warehouse on SAP HANA

SAP HANA can be used as the database for an SAP BW installation. In this scenario, SAP HANA replaces the traditional database server of an SAP BW installation. The application servers stay the same.

The in-memory performance of SAP HANA improves query performance and eliminates the need for manual optimizations by materialized aggregates in SAP BW. Figure 4-10 shows SAP HANA as the database for the SAP BW.

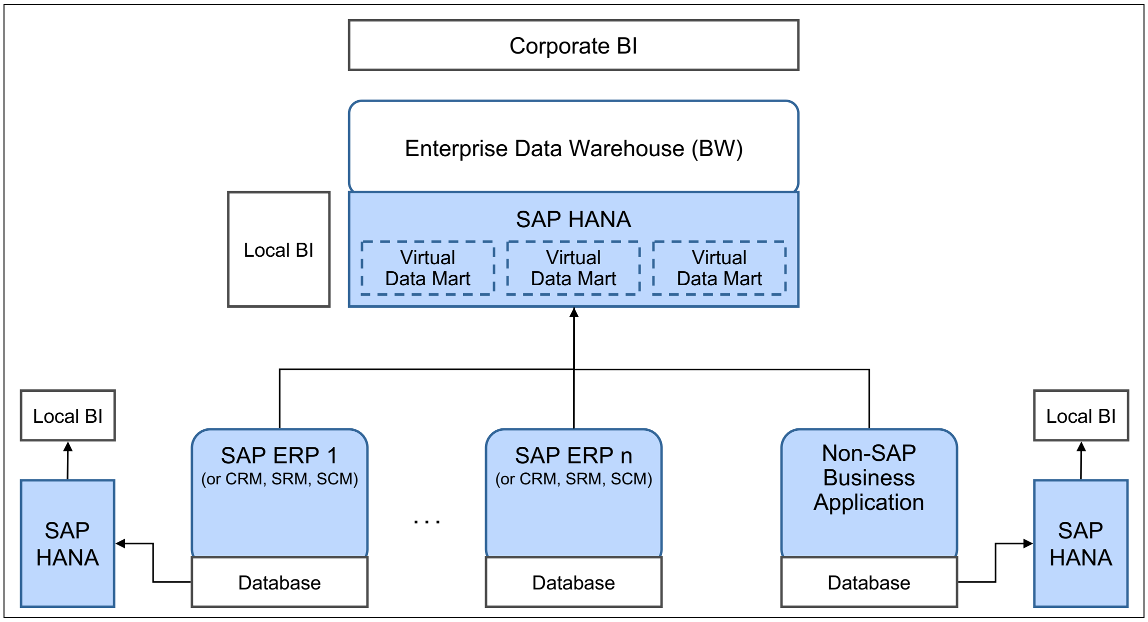


Figure 4-10 SAP HANA as the database for SAP BW

In contrast to an SAP BW system that is accelerated by the in-memory capabilities of SAP BW Accelerator, an SAP BW system with SAP HANA as the database keeps *all* data in-memory. With SAP BW Accelerator, the client 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 data store.

SAP BW on HANA often is the starting point to adopt SAP HANA, which achieves performance improvements with relatively small efforts.

The underlying database is replaced by the SAP HANA database, which improves data loading times and query run times. Because the application logic that is written in ABAP is not affected by this change, all investments in developing BW models are preserved. The transition to SAP HANA is a transparent process that requires minimal effort to adjust existing modeling.

In-memory optimized InfoCubes

InfoCubes in SAP BW that is running on traditional database use the so-called *Enhanced Star Schema*. This schema was designed to optimize different performance aspects of working with multidimensional models on database systems.

The Enhanced Star Schema in SAP BW with an example is shown in Figure 4-11.

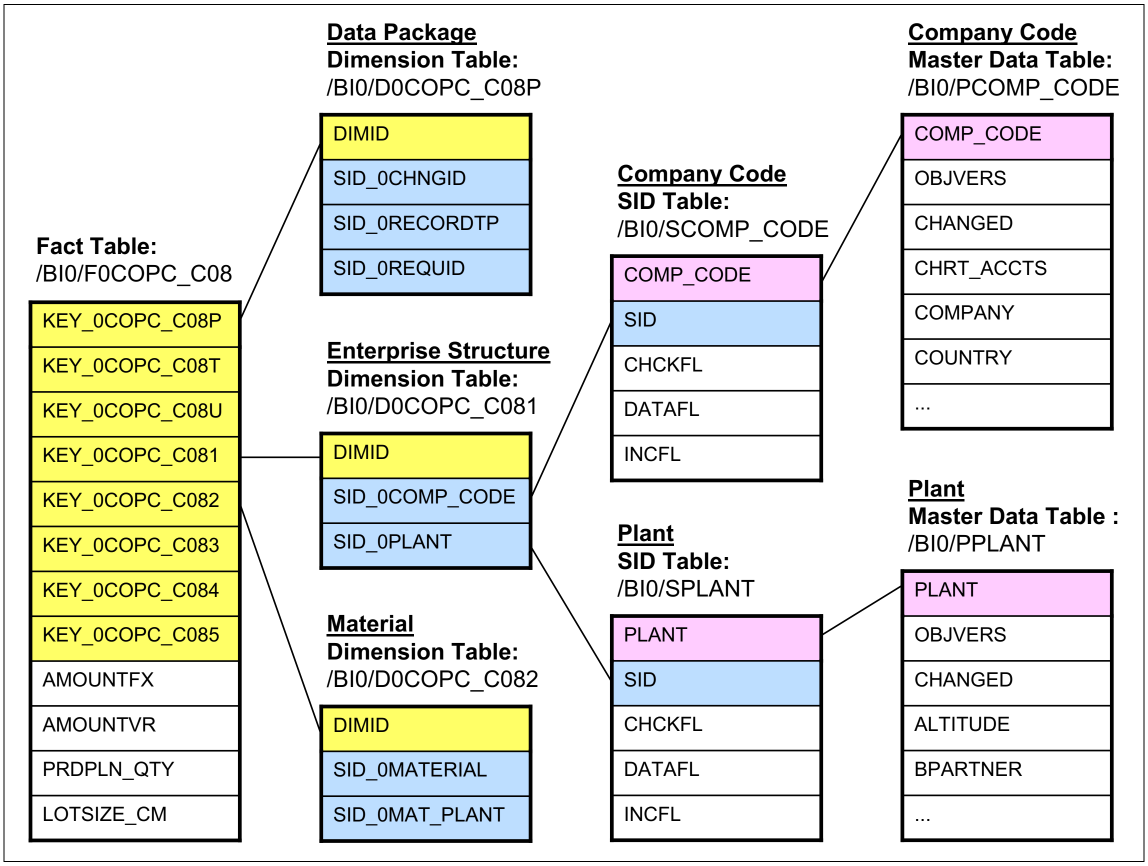


Figure 4-11 Enhanced Star Schema in SAP BW

The core part of every InfoCube is the fact table. This table contains dimension identifiers (IDs) and corresponding key figures (measures). This table is surrounded by dimension tables that are linked to fact tables by using the dimension IDs.

Dimension tables are small tables that group *logically connected* combinations of characteristics, usually representing master data. Logically connected means that the characteristics are highly related to each other; for example, company code and plant. Combining unrelated characteristics leads to many possible combinations, which can have a negative effect on the performance.

Because master data records are in separate tables outside of the InfoCube, another table is required to connect these master data records to dimensions. These extra tables contain a mapping of auto-generated Surrogate IDs (SIDs) to the real master data.

This complex structure is required on classical databases; however, with SAP HANA, this requirement is obsolete. Therefore, SAP introduced the SAP HANA Optimized Star Schema, as shown in Figure 4-12.

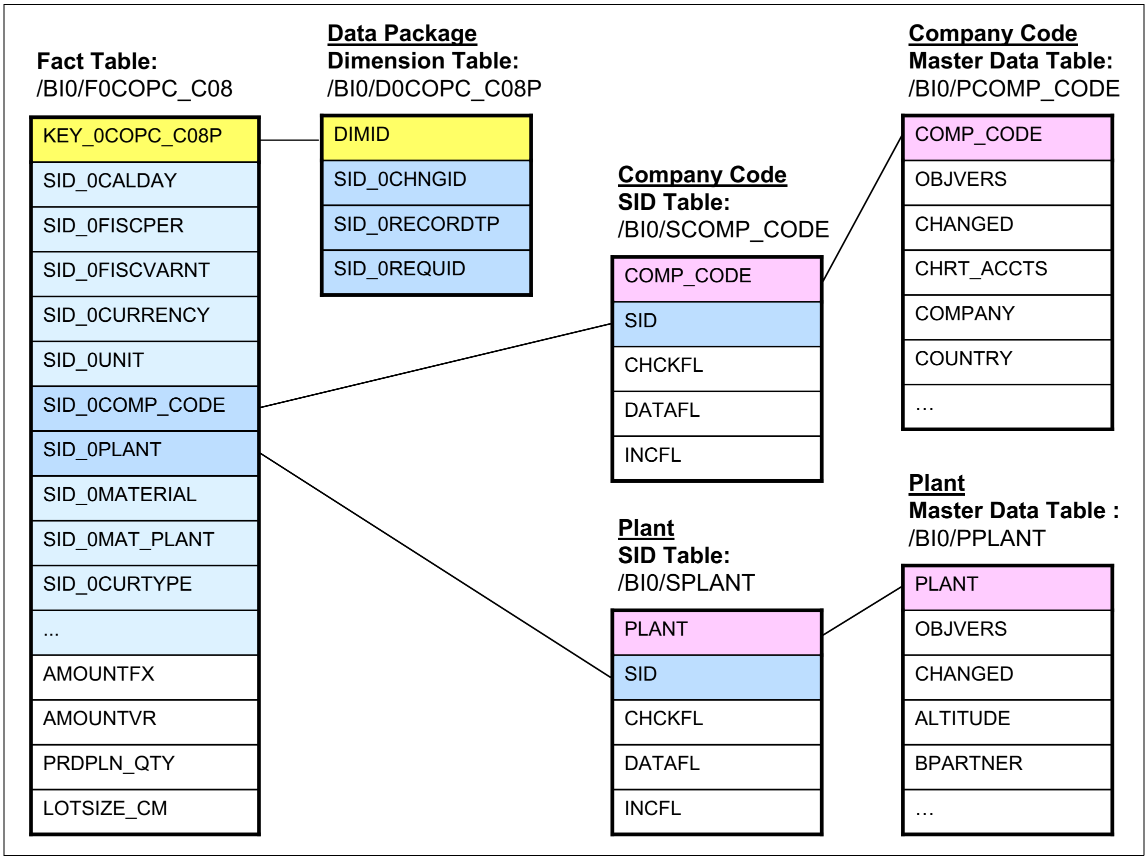


Figure 4-12 SAP HANA Optimized Star Schema in an SAP BW system

The content of all dimensions (except for the Data Package dimension) is incorporated into the fact table. This modification brings the following advantages:

- Simplified modeling

  Poorly designed dimensions (wrong combinations of characteristics) no longer can affect performance. Moving characteristics from one dimension to another is not a physical operation; instead, it is a metadata update.

- Faster loading

  Because dimension tables do not exist, all overhead workload that is related to the identification of combinations or creating combinations in the dimension tables is no longer required. Instead, the required SID values are inserted directly into the fact table.

The SAP HANA Optimized Star Schema is used automatically for all created InfoCubes on the SAP BW system that are running on the SAP HANA database.

InfoCubes are not automatically converted to this new schema during the SAP HANA conversion of the SAP BW system. The conversion of standard InfoCubes to in-memory optimized InfoCubes must be done manually as a follow-up task after the database migration.

SAP HANA acceleration areas

The SAP HANA database can bring performance benefits; however, it is important to set the expectations correctly. SAP HANA can improve loading and query times, but certain limits cannot be overcome.

The migration of SAP BW to run on SAP HANA does not improve extraction processes because extraction occurs in the source system. Therefore, it is important to understand how much of the overall load time is taken by extraction from the source system. This information is needed to correctly estimate the potential performance improvement for the load process.

Other parts of the load process are improved. The new Optimized Star Schema removes unnecessary activities from the loading process.

Some of the calculations and application logic can be pushed to the SAP HANA database, which ensures that data-intensive activities are being done at the SAP HANA database level instead of at the application level. This activity increases the performance because the amount and volume of data that is exchanged between the database and the application are reduced.

SAP HANA can calculate all aggregations in real time. Therefore, aggregates are no longer required, and roll-up activity that is related to aggregate updates is obsolete, which also reduces the overall run time of update operations.

If SAP BW Accelerator is used, the update of its indexes is also no longer needed. Because SAP HANA is based on technology that is similar to SAP BW Accelerator, all queries are accelerated. Query performance with SAP HANA can be compared to situations where all cubes are indexed by the SAP BW Accelerator. In reality, query performance can be even faster than with SAP BW Accelerator because extra features are available for SAP BW that is running on SAP HANA; for example, the possibility of removing an InfoCube and instead running reports against in-memory optimized DataStore Objects (DSOs) and advanced DataStore Objects (aDSOs).

4.5.2 SAP BW/4HANA

SAP BW4HANA is a data warehouse solution which is highly optimized for the SAP HANA platform. It offers a managed approach to data warehousing. This means that prefabricated templates (building blocks) are offered for building a data warehouse in a standardized way. The use case illustrates how you can use your SAP BW4HANA implementation. The main use case of SAP BW4HANA is Data Warehousing.

SAP BW4HANA provides a simplified Data Warehouse, with agile and flexible data modeling, SAP HANA-optimized processes and state of the art user interfaces. The core functionality of SAP BW is preserved. In SAP BW4HANA, objects for data modeling, as well as processes and user interfaces, are especially primed for use with an SAP HANA database. Data modeling is restricted to the small number of objects that are well suited for modeling the layer architecture of a data warehouse on SAP HANA (LSA++). In SAP BW4HANA, data warehouse models can be flexibly combined with SAP HANA views.

SAP BW/4HANA 2.0 requires Unicode and SAP HANA 2.0 SPS03 or higher. The system architecture allows both scale-up as well as scale-out configuration, just like for SAP BW on HANA in the past.

More information is collected in an FAQ document about SAP BW/4HANA:

https://assets.cdn.sap.com/sapcom/docs/2016/08/c4458a08-877c-0010-82c7-eda71af511fa.pdf

4.5.3 SAP Business Suite on SAP HANA

This solution is generally available since 2013.

SAP HANA can be used as the database for an SAP Business Suite installation. In this scenario, SAP HANA replaces the traditional database server of an SAP Business Suite installation. The application servers stay the same, and can run on any platform that supports the SAP HANA database client. As of May 2016, the following applications of SAP Business Suite are supported by SAP HANA as their primary database:

- Enterprise Resource Planning (ERP)
- Customer Relationship Management (CRM)
- Supply Chain Management (SCM)
- Supplier Relationship Management (SRM)
- several others

SAP Business Suite on SAP HANA does not induce any functional changes. Configuration, customization, the ABAP Workbench, connectivity, security, transports, and monitoring stay unchanged. For modifications, the same upgrade requirements as with any other upgrade apply. Customized code can stay unchanged, or can be adjusted to use extra performance.

SAP Business Suite applications can benefit in the following ways from the in-memory technology of SAP HANA:

- Running dialog processes instead of batch
- Integration of unstructured data and machine-to-machine data (M2M) with ERP processes
- Integration of predictive analysis with ERP processes
- Running operational reports in real time, directly on the source data
- Removing the need for operational data stores
- Eliminating the need for data replication or transfers to improve operational report performance

Restrictions

From an architectural perspective, SAP Business Suite with SAP HANA should be installed on a scale-up, single node system.

High availability (HA) scenarios for SAP Business Suite with SAP HANA are supported, but are restricted to the simplest case of two servers, one being the worker node and one acting as a standby node. In this case, the database is not partitioned, but the entire database is on a single node. This configuration is sometimes also referred to as a *single-node HA* configuration. For more information about the available configurations that are dedicated to SAP Business Suite (which is powered by SAP HANA), see Chapter 6, “SAP HANA IT landscapes with Lenovo solutions” on page 99.

4.5.4 SAP S/4HANA

In 2015 SAP released SAP Business Suite 4 SAP HANA, the business suite offering that is tightly integrated with SAP HANA as the underlying database.

SAP S/4HANA is the next step of innovation towards an IT platform that provides real-time business insights. SAP S/4HANA is based on new concepts in several different areas:

- Simplified data model
- New user experience that is consistent across all access methods
- Advanced processing
- Instant Insight
- Support for Internet-of-Things (IoT)
- Connectivity to third-party systems

SAP S/4HANA only runs on SAP HANA. Traditional database management systems are not supported with SAP S/4HANA.

Transition to SAP S/4HANA

There are three different ways to transition to S/4HANA:

- New Implementation

  This method is applicable for all new customers and for customers who currently run their ERP software on non-SAP products. Existing data from a legacy system can be loaded into the new implementation with tools like SAP Data Services (SAP DS). This approach is also called greenfield implementation.

  New implementations can be done either on-premise or in the cloud.

- System Conversion

  This method covers the conversion of existing customers running their ERP on SAP already; for example, all customers running SAP Business Suite. System conversion works on all layers of the software stack, the database layer, SAP Netweaver, and the application. Ultimately, also a conversion of the user interface to SAP Fiori UX is required.

  System conversions require an on-premise installation.

- Landscape Transformation

  Landscape Transformation provides a more granular approach of converting an SAP Business Suite environment into an SAP S/4HANA environment. It allows to consolidate multiple different source ERP systems into one SAP S/4HANA system, or to only selectively transform source data into an SAP S/4HANA system. This allows for a phased migration with those components that provide the highest return on investment of a migration. SAP Landscape Transformation (SLT) is the migration tool that supports this transition method.

  Landscape Transformation projects work with on-premise or cloud systems as the destination environment.

System Conversion and Landscape Transformation are often referred to as “Brownfield” scenarios.

More product information about SAP S/4HANA can be found at the following website:

http://www.sap.com/s4hana

System topology for SAP S/4HANA

SAP S/4HANA installations benefit from a single node implementation. It is heavily recommended to go through a detailed sizing exercise.

Before going to SAP S/4HANA (and any other SAP HANA based application for that matter), customers should look at the database content and size of the source system.

Invest in housekeeping, data reduction and archiving as strongly as possible. This eases operational tasks like e.g. backup / restore times and business continuity disciplines.

Usually, SAP objects like basis tables containing IDOCs, workflow documents, application logs can be deleted or at least considered to be placed on disk in HANA, not in memory. This could be an ideal candidate for native storage extension as described in “Native Storage Extension” on page 17.

While it is recommended to go scale-up first, scale-out clusters are generally supported for SAP S/4HANA systems. If a scale-out configuration is necessary due to the database size, a few boundary conditions apply:

- Use as few nodes as possible to reduce network traffic between the nodes
- Use the largest possible node size
- Carefully determine the data distribution

It is essential in a scale-out configuration, which can very well only consist of 2 nodes, to avoid cross-node joins for the transactions sent to the database. These can involve significant degradation in performance. SAP provides advise and tools for the table distribution in a scale-out configuration.

More details can be found in SAP Note 2447004 - Table Grouping Report for S/4HANA in scale-out systems.

4.6 Programming techniques that use SAP HANA

The last use case scenario is based on recent developments from SAP where applications can be built directly against the SAP HANA database by using all its features, such as the embedded application server (XS Engine) or stored procedures, which allows logic to be directly processed inside the SAP HANA database. They are sometimes referred to as “next generation applications”.

A new software component can be integrated with SAP HANA directly or it can be built on top of the SAP NetWeaver stack, which can work with the SAP HANA database by using client libraries.

Because of its breadth and depth, this use case scenario is not described in detail as part of this publication.

5

Components of the Lenovo Solution for SAP HANA

This chapter describes the components that make up the Lenovo solutions for SAP HANA. It covers the solution based on the ThinkSystem SR950 servers in detail as well as SR850 and SR650 on a more high level. It describes the configurations of the workload-optimized solutions for all generations of Intel processors. It then describes the different storage subsystems including the file systems and the features they provide to the solution. The final section describes the networking switches that are part of the solution.

This chapter includes the following topics:

5.1 Lenovo ThinkSystem Servers

In June 2017, Lenovo announced a new generation of servers under the new ThinkSystem brand. They are the designated successor to the well-known Lenovo System x® brand and the high-end flagship enterprise server is the Lenovo ThinkSystem SR950. The following sections describe the compute, storage and I/O components.

The components of SR850 is available in 5.1.9, “Lenovo ThinkSystem SR850” on page 73 and the details about SR650 can be found in 5.1.10, “Lenovo ThinkSystem SR650” on page 74.

Table 5-1 compares the previous generation of the Intel Xeon processors to the new Scalable Family processors that are supported in SR950 systems.

Table 5-1 Processor comparisons across CPU generations

Feature	**X6 family, Xeon E7 v3**	**X6 family, Xeon E7 v4**	**Intel Xeon SP Gen 1**	**Intel Xeon SP Gen 2**
Processor family	Intel Xeon E7-8800 v3 Intel Xeon E7-4800 v3	Intel Xeon E7-8800 v4 Intel Xeon E7-4800 v4	Intel Xeon 8100 Intel Xeon 6100 Intel Xeon 5100	Intel Xeon 8200 Intel Xeon 6200 Intel Xeon 5200
Processor codenames	Haswell EX	Broadwell EX	Skylake SP	Cascade Lake SP
Cores per CPU	Up to 18 cores	Up to 24 cores	Up to 28 cores	Up to 28 cores
Last level cache	Up to 45 MB L3 cache	Up to 60 MB L3 cache	Up to 1.375 MB per core Up to 38.5 MB per processor	Up to 1.375 MB per core Up to 38.5 MB per processor
QPI / UPI Data rate	QPI: 3 links 9.6 GT/s max	QPI: 3 links 9.6 GT/s max	UPI: 2 or 3 links 10.4 GT/s max	UPI: 2 or 3 links 10.4 GT/s max
CPU TDP rating	Up to 165 W	Up to 165 W	Up to 205 W	Up to 205 W
DIMM sockets	24 DDR3 DIMMs per CPU 24 DDR4 DIMMs per CPU	24 DDR4 DIMMs per CPU	12 DDR4 DIMMs per CPU	12 DDR4 DIMMs per CPU
Maximum memory speeds	2133 MHz	2400 MHz	2666 MHz	2933 MHz
PCIe technology	PCIe 3.0 (8 GTps)	PCIe 3.0 (8 GTps)	PCIe 3.0 (8 GTps)	PCIe 3.0 (8 GTps)

The SR950 uses the Intel Scalable processor family with Platinum (81xx, up to 8 sockets) and Gold (61xx and 51xx, up to 4 sockets) processors. Figure 5-1 shows the ThinkSystem SR950 server.

Figure 5-1 Lenovo ThinkSystem SR950

There are three types of trays that make up the SR950:

- **Compute tray:** The processors and memory are located on the compute trays, the Upper Compute Tray and the Lower Compute Tray. Each compute tray holds up to four processors and 24 DIMMs on two system boards. The compute trays are accessible from the front of the server once the front bezel is removed.
- **Storage tray:** For some four-socket storage-rich configurations, a storage tray can be used in the upper tray area instead of a compute tray to add additional storage to the chassis. The storage tray is similarly accessible from the front of the server.
- **I/O tray:** The I/O tray is accessible from the rear of the server and houses all the PCIe slots. The I/O tray can be configured with multiple types of riser cards to suit the type of I/O cards being installed in the server.

These trays allow the server to be configured in multiple ways to achieve the desired level of processing power, I/O, storage and memory capacities.

The following subsections discuss each of these trays in detail.

5.1.1 Compute trays

The ThinkSystem SR950 can scale from 2-sockets to 4-sockets, all the way up to 8-sockets without the need to replace the servers enclosure or upgrade to a physically larger design. This is achieved through adding compute trays and system boards to the 4U chassis.

Following we cover the version of SR950 which contains the processor Xeon SP Gen 1 (also referred to as the codename Skylake). The only difference in SR950 with Xeon SP Gen 2 is the CPU itself and the memory DIMMS in the same chassis.

The server supports up to two Compute Trays. The compute trays are accessible from the front of the server. Figure 5-2 shows the top compute tray being removed from the chassis.

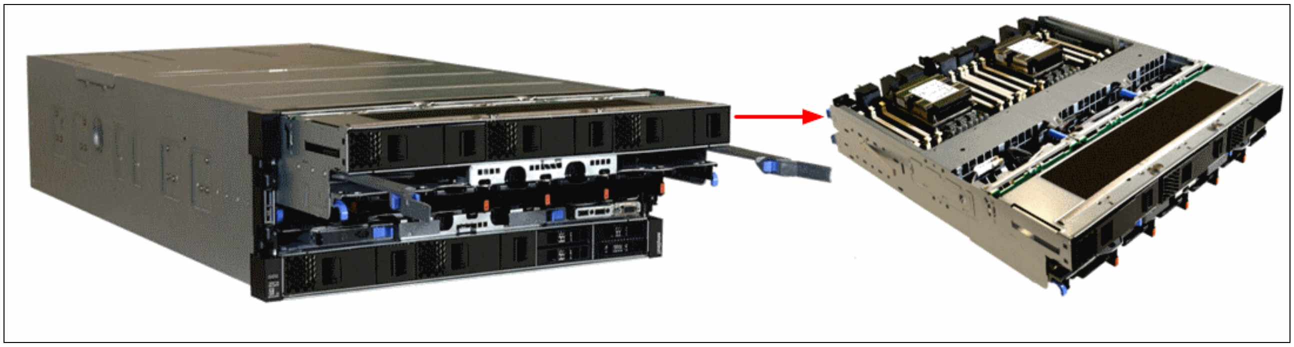


Figure 5-2   ThinkSystem SR950 compute tray being removed (upper)

Figure 5-3 shows a top view of the lower and upper compute trays. The processors and memory are located towards the rear of the compute tray. The front of the compute tray holds the storage and long sliding tabs for pulling the out hot swap fans.

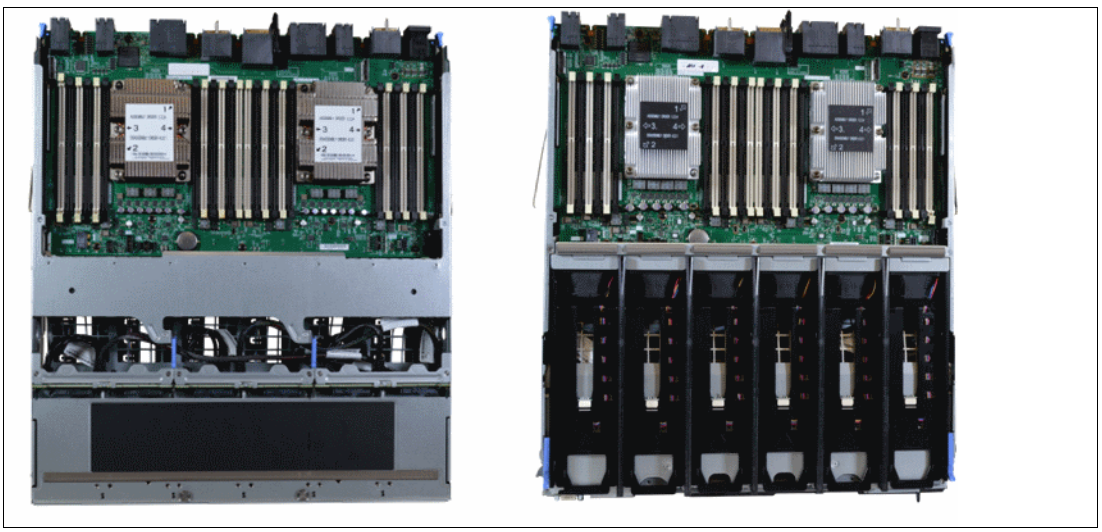


Figure 5-3   Upper (left) and lower (right) compute trays

Each compute tray contains:

- One or two compute system boards, each comprising:
  - Two processors
  - 24 DIMMs
- Six hot-swap fans, accessible from the front of the server even when the compute tray is installed
- 12x 2.5-inch hot-swap drive bays
- One dedicated PCIe 3.0 x8 slot reserved for a internal RAID adapter or SAS HBA for internal SAS/SATA drives

Figure 5-4 shows the rear of the compute tray with two compute system boards, where one of the system boards is being removed.

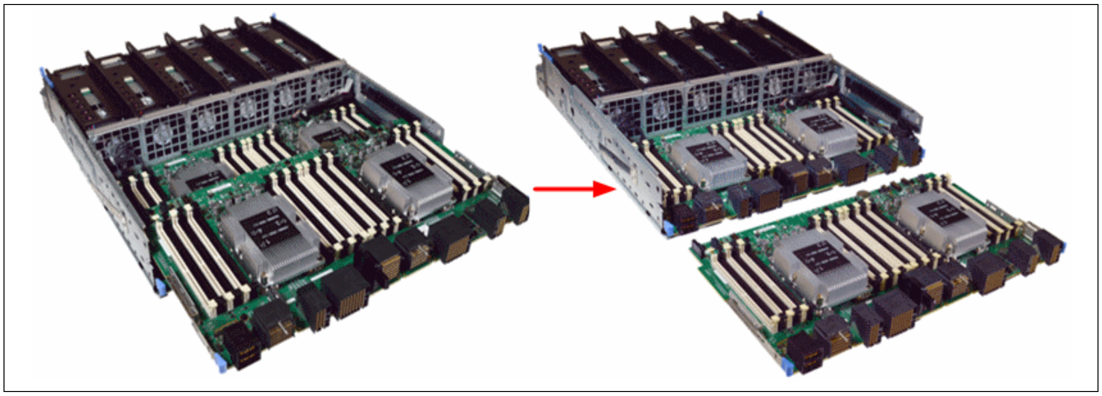


Figure 5-4 Compute tray with two system boards, one being removed

The compute system boards on its own is shown in Figure 5-5 when not installed in a Compute tray. Each system board holds two processors and 24 DIMM sockets, 12 per processor. The system board also has connectors for the NVMe ports used to connect to the PCIe NVMe drive bays at the front of the server.

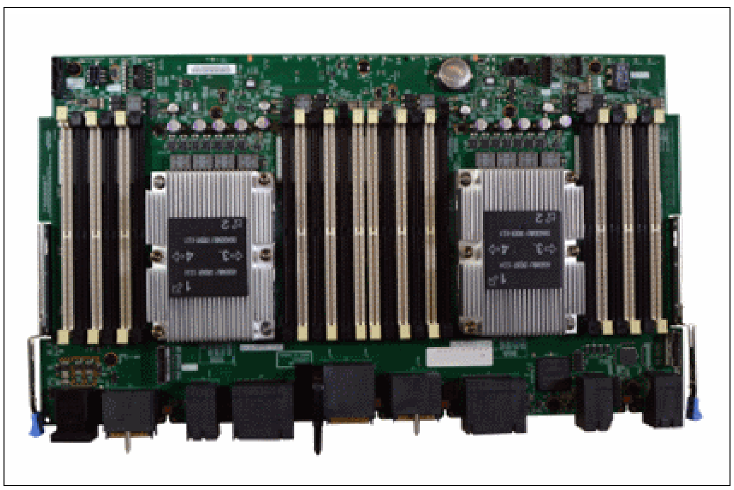


Figure 5-5 Compute system board

The dedicated PCIe 3.0 x8 slot used with the RAID adapter or SAS HBA that connects to the drive bays in the compute tray is located between the fan chutes and the drive bays, as shown in Figure 5-6.

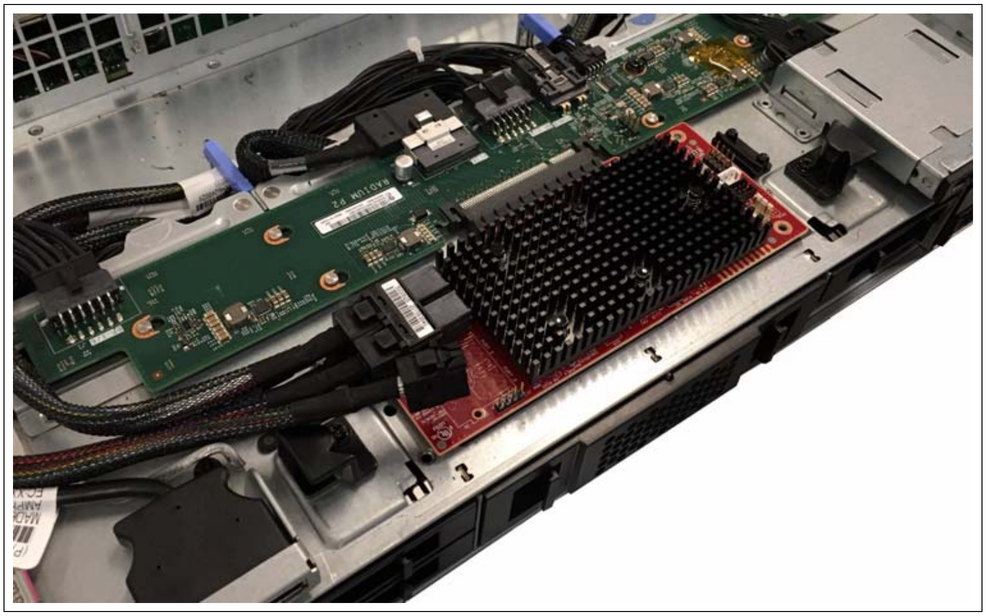


Figure 5-6 Internal RAID adapter slot (lower compute tray)

5.1.2 Storage tray

The SR950 chassis can support up to 24 drives internally when both compute trays are installed (12 drives in each tray). However, for some four-socket storage-rich configurations with only one compute tray installed, a storage tray can be used in the upper tray area instead of an additional compute tray to provide the storage. The storage tray has no memory or processors on it.

Only one storage tray can be installed and it must be installed in the upper tray area. The lower tray area requires a compute tray to be installed for the processor(s) and memory.

A storage tray contains:

- NVMe ports (for connecting NVMe drives in the tray to the processors in the lower compute tray)
- Six hot-swap fans, accessible from the front of the server even when the storage tray is installed
- 12x 2.5-inch hot-swap drive bays (6 NVMe)
- One PCIe slot reserved for a RAID adapter for internal SAS/SATA drives (located under the drives)

This applies to both the first and the second generation of Intel Xeon Scalable Processors.

Figure 5-7 shows the storage tray pulled out of the system. The rear of the tray lacks the system board that contains the CPU and memory DIMMs

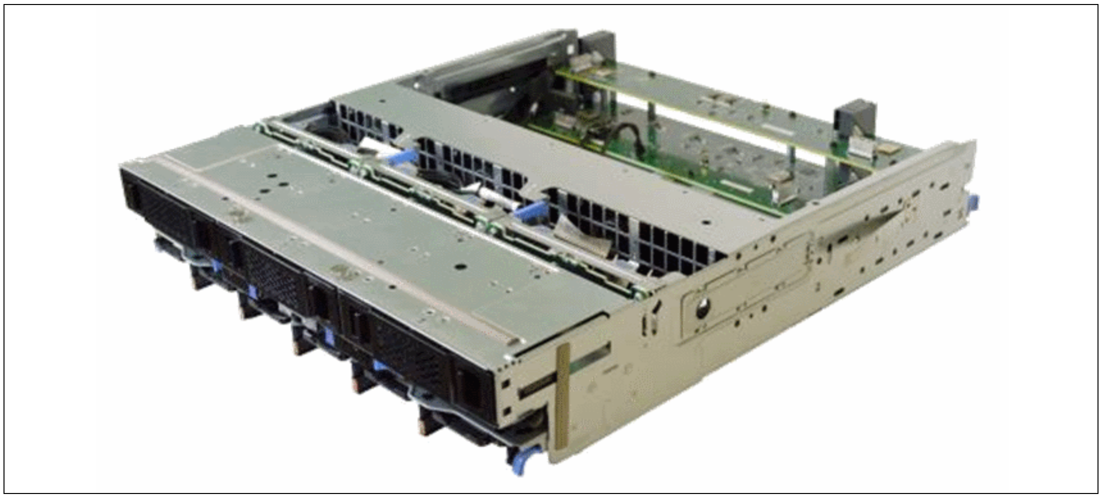


Figure 5-7 Storage tray

5.1.3 I/O tray

At the rear of the server is the I/O tray. Figure 5-8 displays the locations of the slots at the rear of the server that make up the I/O tray.

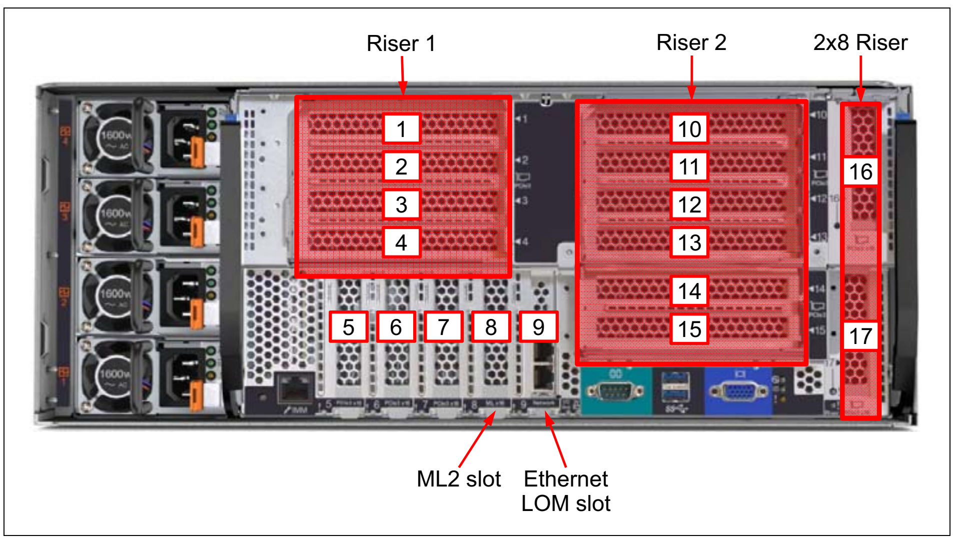


Figure 5-8 PCIe 3.0 slots at the rear of the SR950 server

The I/O tray can contain the following:

- 17 PCIe 3.0 slots depending on the riser cards installed in the tray. These are all accessible from the rear of the server
- One internal dedicated M.2 slot for an M.2 adapter located internally on the I/O tray

This applies to both the first and the second generation of Intel Xeon Scalable Processors.

Figure 5-9 shows the rear of the server where the I/O tray is being removed.

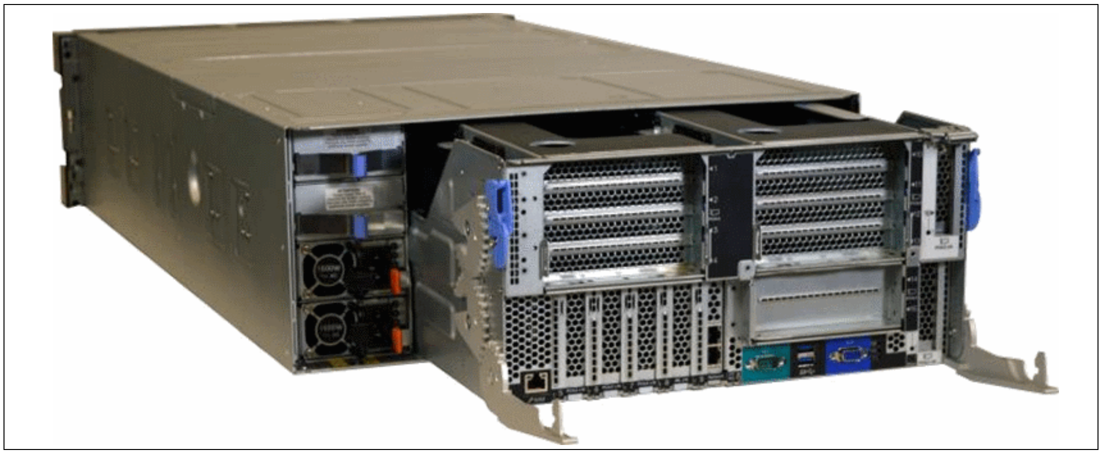


Figure 5-9 I/O tray being removed from the rear of the server

5.1.4 Intel Xeon Scalable Processor (SP) family

The new Intel Xeon Processor Scalable Family (formerly codenamed “Skylake-SP” and “Cascade Lake SP”) processors have been grouped into four functional levels or *shelves*:

- Platinum
- Gold
- Silver
- Bronze

The SR950 and SR850 supports the Gold (51xx and 61xx) and Platinum (81xx) level processors, this includes the "top bin" highest performing 205 W processors. These processors feature the new Intel microarchitecture and provide a higher core count, higher frequencies, faster UPI links (the follow-on to QPI), more UPI links, faster DDR4 bus speeds, faster AVX-512 (instruction set extension) and advanced RAS features. The same applies to Xeon SP Gen 2 supports where only the processors are Gold (52xx and 62xx) and Platinum (82xx).

The SR650 is supported for SAP HANA as well and offers Silver level processors (41xx and 42xx) additionally.

5.1.5 Gold and Platinum product family

The following groups of the Intel Xeon Scalable Family are used in the SR950 server:

- The Intel Xeon processor Gold product family. This includes processor models 51xx / 52xx and 61xx / 62xx. This family supports four-processor configurations only. Processor models 51xx have two UPI links and models 61xx have three UPI links.
- The Intel Xeon processor Platinum product family with processor models 81xx / 82xx. This family supports both four-processor and eight-processor configurations and have three UPI links.

The Gold and Platinum Scalable Family of processors offers the following key features:

- Up to 2933MHz DDR4
- 14nm process technology
- Up to 28 Cores per processor

- UPI links at speeds up to 11.2GT/s
- Intel Hyper-Threading Technology (2 threads per core)
- Intel AVX-512 (AVX2 in previous generations)
- DDR4 memory interface support, which brings greater performance and power efficiency
- Rebalanced Cache Hierarchy: Increased MLC 1.375 MB Last Level Cache per core
- Optional integrated Fabric: Intel Omni-Path Architecture
- Integrated PCIe 3.0 controller with 48 lanes per processor at 8 GT/s
- 1.5x memory bandwidth increase (6 channels vs. 4 in previous generation)
- Memory Technology: 6xDDR4 channels running at up to 2666 MHz
- Intel Virtualization Technology (VT-x and VT-d)
- Intel Turbo Boost Technology
- Improved performance for integer and floating point operations
- Virtualization improvements with regards to posted jnterrupts, page modification logging, and VM enter/exit latency reduction
- Intel AES-New instructions for accelerating of encryption
- Advanced UPI and memory reliability, availability, and serviceability (RAS) features
- Machine Check Architecture (MCA) recovery (non-running and running paths)
- Resource director technology (RDT): Cache monitoring technology, cache allocation technology, memory bandwidth monitoring

Security technologies include:

- Intel BIOS Guard 2.0
- Intel Boot Guard
- Secure Key
- Intel Trusted Execution Technology (TXT)
- Intel QuickAssist Technology (QAT
- Intel Platform Trust Technology (PTT)
- Intel Memory Protection Extensions (MPX)
- Mode based execution control (XU/XS bits)

5.1.6 Memory subsystem on SR950

The SR950 with Xeon SP Gen 1 uses Lenovo TruDDR4 memory operating at up to 2666 MHz. The SR950 with Xeon SP Gen 2 uses Lenovo TruDDR4 memory operating at up to 2933 MHz. Both support 12 DIMMs per processor, which corresponds to 48 DIMMs with four processors installed and 96 DIMMs when eight processors are installed. Each processor has six memory channels with two DIMMs per channel. With 128 GB 3DS RDIMMs installed, an 8-socket server supports a total of 12 TB of system memory.

The SR950 memory interface support Lenovo TruDDR4 memory modules, which are tested and tuned to maximize performance and reliability. Lenovo TruDDR4 DIMMs can operate at greater speeds and have higher performance than DIMMs that meet industry standards.

TruDDR4 memory types have ECC protection and support Chipkill technologies as well as redundant bit steering.

Each processor has two Integrated Memory Controllers (IMC) that allows the SR950 to support the following memory configuration:

- Two IMC per processor

- Three channels per IMC
- Six channels per processor
- Each channel consists of 64bit data and 8bit ECC
- Each channel supports the following DDR4 memory types:
  - Registered DIMMs (RDIMMs)
  - Load Reduced DIMMs (LRDIMMs)
  - 3DS DIMMs
- Up to two DIMMs per channel
- Up to 12 DIMMs per socket
- Up to 24 DIMMs per system board
- Up to 48 DIMMs per compute tray
- Up to 96 DIMMs per system

Figure 5-10 shows the IMC with attached memory. The allow memory configurations and the placement order is part of the definition of the Lenovo solution for SAP HANA. It is described in 5.2, “Storage subsystem”

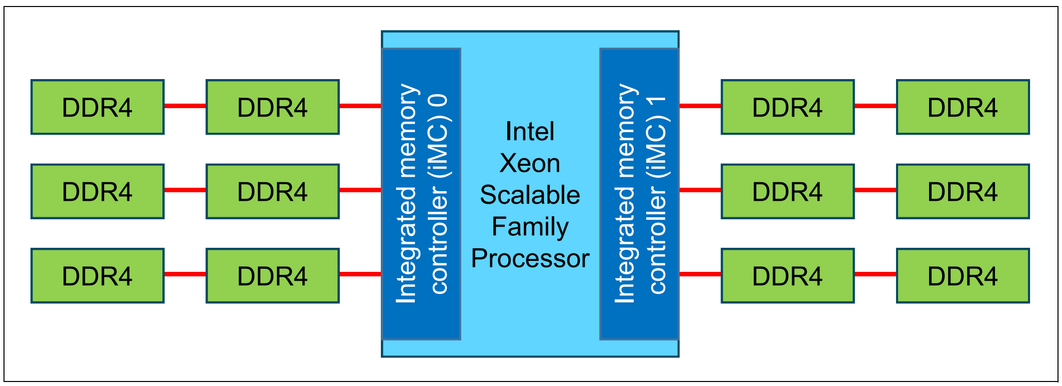


Figure 5-10 Integrated Memory Controllers with attached DIMMs

All DIMMs in Intel Xeon SP Gen 1 systems operate at a speed of 2666 MHz, both at 1 DIMM per channel and 2 DIMMs per channel. However, if the processor selected has a lower memory bus speed (e.g. 2400 MHz), then all DIMMs will operate at that lower speed.

In the second generation processors based systems, the DIMMs operate as follows:

- 2666 MHz DIMMs, that operate at 2666 MHz both at 1 DIMM per channel and 2 DIMMs per channel
- 2933 MHz DIMMs, that operate at 2933 MHz at 1 DIMM per channel, and at 2666 MHz at 2 DIMMs per channel
- 2933 MHz Performance+ DIMMs, that operate at 2933 MHz both at 1 DIMM per channel and 2 DIMMs per channel

5.1.7 Intel Optane DC Persistent Memory

In the SR950 with Xeon SP Gen 2, Intel Optane DC Persistent Memory represents a new class of memory and storage technology explicitly architected for data center usage.

Using Lenovo ThinkSystem servers running applications that are tuned for Intel Optane DC Persistent Memory will result in lower data latency compared to solid-state drive technology. Intel Optane DC Persistent Memory Modules (DCPMM) became generally available in April 2019 and have the form factor of a DDR4 DIMM as depicted in Figure 5-11 on page 69, but the persistence and capacity of data storage of a solid-state drive. This means the DCPMMs have performance characteristics similar to that of TruDDR4 DIMMs, the storage capacity of an SSD, and the ability to stay active after a power cycle or reboot of the server. These features open up a new way of performing data I/O to application developers and new levels of server performance to customers.

Figure 5-11 Intel Optane DC Persistent Memory Module (top) and Lenovo TruDDR4 DIMM (bottom)

DCPMM modes

Intel Optane DC Persistent Memory operates in one of three modes:

- Memory Mode

  In this mode, the DCPMMs act as large capacity DDR4 memory modules. In such a configuration, the memory that the operating system recognizes is the DCPMMs; the installed TruDDR4 DIMMs are hidden from the operating system and act as a caching layer for the DCPMMs. In this mode, the persistence feature of the DCPMMs is disabled. This mode does not require the application to be DCPMM-aware.

- App Direct Mode

  In this mode, the DCPMMs provide all persistence features to the operating system and applications that support them. The operating system presents both TruDDR4 DIMMs and DCPMMs to the applications, as system memory and persistent storage respectively.

  Depending on the configuration in UEFI and the operating system, the DCPMMs appear as one of two types of namespaces:

  – Direct access (DAX): byte-addressable storage accessible via an API. The applications must be DCPMM-aware and use the published APIs to implement the DCPMM features.

  – Block storage: the persistent memory presented to applications is seen as a block storage device, similar to an SSD. The operating system needs to be DCPMM-aware, however the applications do not.

- Mixed Memory Mode

  Mixed Memory Mode is a combination of Memory Mode and App Direct Mode, where a portion of the capacity of the DCPMMs is used for the Memory Mode operations, and the remaining capacity of the DCPMMs is used for the App Direct Mode operations. In this mode, all installed TruDDR4 DIMMs are hidden from the operating system and act as a caching layer for portion of the DCPMMs in Memory Mode.

In App Direct Mode (and the persistent portion of Mixed Memory Mode), the persistent memory can be configured in one of two ways:

- Interleaved, where all DCPMMs are seen as one single monolithic space. This is similar in concept to RAID-0 in storage.
- Non-interleaved, where each DCPMM is seen as a separate space. This is similar in concept to JBOD in storage.

For SAP HANA systems, **only the App Direct Mode and interleaved** is applicable and used. The biggest possible configuration in an 8-socket server offers a total of 24 TB of system memory.

More information is available here:

https://lenovopress.com/lp1066-intel-optane-dc-persistent-memory

5.1.8 Storage subsystem

In this section, we give an overview of the following technologies that are used in the SR950 system and relevant for the Lenovo solution for SAP HANA:

- 2.5-inch SAS/SATA slots
- Lenovo AnyBay™: NVMe technology
- M.2 adapter

SAS/SATA SSD technology

The SR950 offers up to 24x 2.5-inch drive bays, depending on the server configuration selected. These can be populated with HDDs or SSDs. An all-flash solutions can provide the following benefits:

- Significantly lower implementation cost of high performance I/O-intensive storage systems.
- Significantly higher performance and better response time of storage-intensive applications.
- Significant savings in power and cooling with high performance-per-watt ratio.

SSDs are optimized for a heavy mix of random read and write operations, such as transaction processing, data mining, business intelligence, and decision support, and other random I/O-intensive applications.

SSD units require the following components:

- Backplane(s). There are two different backplanes available for the SR950:
  - SAS/SATA backplane: Supports 4 SAS or SATA SSD drives
  - AnyBay backplane: Supports 4 SAS or SATA SSD drives (or 2 NVMe drives)
- 2.5-inch SSDs
- Supported RAID or HBA adapter

Figure 5-12 shows the two available backplanes. The AnyBay backplane is shown at the top of the photo. The SAS/SATA backplane is shown at the bottom of the photo.

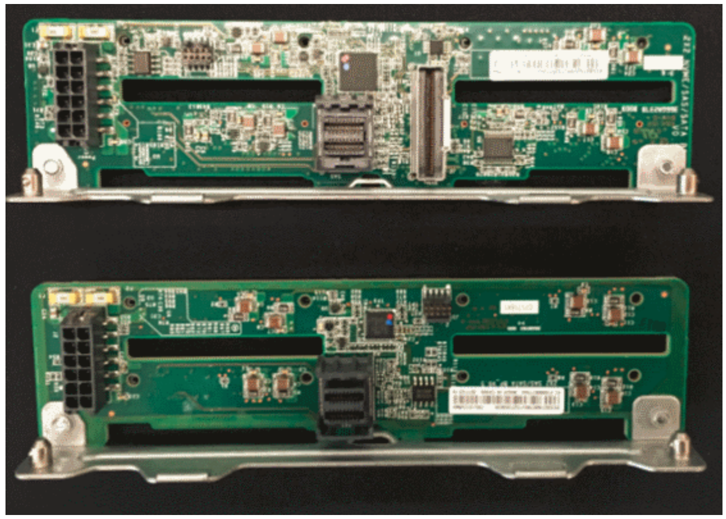


Figure 5-12 SR950 backplane options

For configurations with local storage, the Lenovo solution for SAP HANA has a pre-selected storage subsystem configuration that is validated and certified with SAP.

Lenovo AnyBay: NVMe technology

Non-Volatile Memory Express (NVMe) is new PCIe 3.0 high performance solid-state drive (SSD) technology that provides high I/O throughput and low latency. NVMe interfaces remove SAS/SATA bottlenecks and enable all of the capabilities of contemporary NAND flash memory.

Figure 5-13 on page 71 shows the PCIe NVMe SSDs of three different vendors: Toshiba, Intel and Samsung.

Figure 5-13 NVMe PCIe SSDs: (l-r): Toshiba, Intel and Samsung

Each NVMe SSD has direct PCIe 3.0 x4 connection, which provides at least 2x more bandwidth and 2x lower latency than SATA/SAS-based SSD solutions. NVMe drives are also optimized for heavy multi-threaded workloads by using internal parallelism and many other improvements, such as enlarged I/O queues.

NVMe technology has the following key characteristics:

- PCIe 3.0 connection. There is a PCIe 3.0 x4 connection for each NVMe drive with up to 4 GBps overall throughput
- Low I/O latency. For example, the average read/write latency for the Intel P4800X Optane drives is 10 µs
- High sequential throughput. For example, Toshiba drives offer up to 3100 MBps sequential read speed with 128 KB blocks, and up to 2350 MBps sequential write speed with 128 KB blocks per drive
- High I/O operations per second. For example the Toshiba drives support up to 666,000 IOPS of random read with 4 KB blocks, and up to 105,000 IOPS of random writes with 4 KB blocks
- A total of 65,536 I/O queues supported and 65,536 commands per queue supported, which provides great performance on heavily multithreaded workloads with combined sequential and random access
- High endurance: The Intel P4800X Optane drives, for example, include features which combine NAND silicon enhancements and SSD NAND management techniques to extend SSD write endurance up to 30 drive writes per day (DWPD) for 5 years
- Support for software RAID under operating system management
- Hot add and hot remove features are available on specific servers with supported operating systems
- Most operating systems have native support of NVMe drives or provide support through software drivers, such as
  - RHEL 6.5 and later
  - SLES 11 SP3 and later
  - Windows Server 2008 R2 and later
  - VMware ESXi 5.5 and later
- NVMe drives can be used as boot drives

Please find the usage of NVMe in SAP HANA configurations at 6.7, “Consolidation of SAP HANA instances” on page 116.

M.2 drives

The server (no difference for Xeon SP Gen 1 or 2) supports one or two M.2 form-factor SATA drives for use as an operating system boot solution. With two M.2 drives configured, the drives are configured by default as a RAID-1 mirrored pair for redundancy.

The M.2 drives install into an M.2 adapter which in turn is installed in a dedicated slot on the system board.

There are two M.2 adapters supported:

- Single M.2 Boot Adapter, which supports one M.2 drive; available as ThinkSystem M.2 Enablement Kit, 7Y37A01092
- Dual M.2 Boot Adapter, which supports one or two M.2 drives; available as ThinkSystem M.2 with Mirroring Enablement Kit, 7Y37A01093

A Dual M.2 Boot Adapter with one 128GB M.2 drive partially inserted is shown in Figure 5-14. The second M.2 drive is installed on the other side of the adapter.

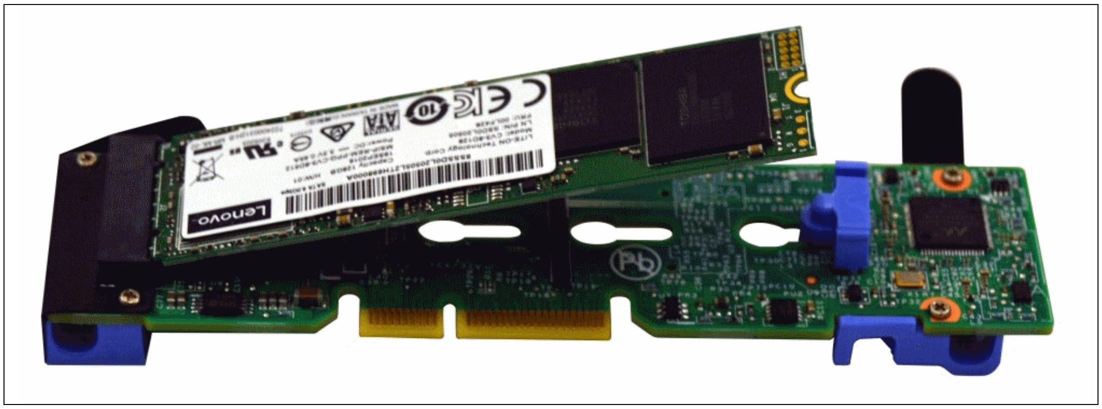


Figure 5-14 Dual M.2 Boot Adapter and a 128 GB M.2 drive

Features of the Dual M.2 Boot Adapter are listed below:

- PCIe 2.0 x2 host interface (connects to the PCH)
- Based on the Marvell 88SE9230 6 Gbps SATA controller
- Supports two 6 Gbps SATA M.2 drives (it is not supported to have only one drive installed)
- RAID functionality provided by the M.2 adapter
- RAID 1 by default; also supports RAID 0 and JBOD
- UEFI-based settings to enable/disable RAID mode and to review inventory
- Adapter and drive firmware update using Lenovo firmware tools
- Management via I2C interface

Features of the Single M.2 Boot Adapter:

- 6 Gbps SATA host interface (connects to the PCH)
- Supports one 6 Gbps SATA M.2 drive
- Drive firmware update using Lenovo firmware tools
- Management via I2C interface
- VPD reporting of adapter inventory

The Single M.2 Boot Adapter is shown in Figure 5-15, with the 32GB M.2 drive installed.

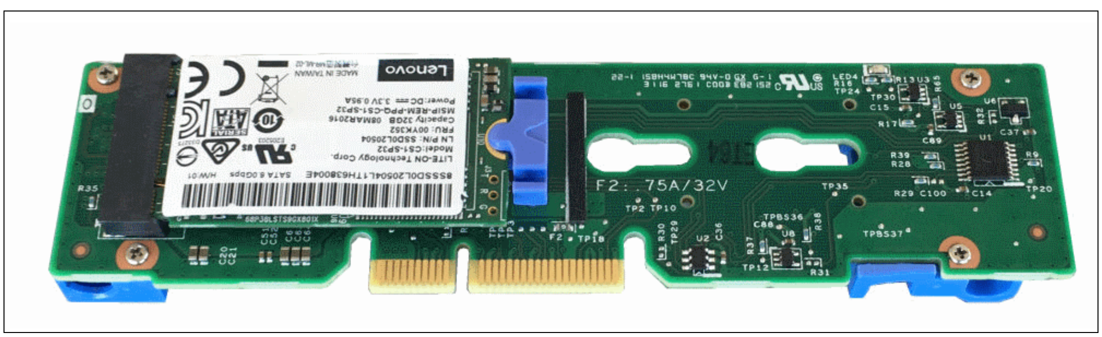


Figure 5-15 Single M.2 Boot Adapter and a 32 GB M.2 drive

5.1.9 Lenovo ThinkSystem SR850

The Lenovo ThinkSystem SR850 as depicted in Figure 5-16 is a 4-socket server that features a streamlined 2U rack design which is optimized for price and performance, with best-in-class flexibility and expandability. Models of the SR850 are powered by up to four Intel Xeon Processor Scalable Family processors, each with up to 28 cores, for an efficient 4-socket solution & memory support up to 6TB. The ThinkSystem SR850's agile design provides rapid upgrades for processors and memory, and its large, flexible storage capacity helps to keep pace with data growth.

Figure 5-16 ThinkSystem SR850

Suggested uses: General business consolidation, data analytics, virtualization, database, dense computing and scientific applications.

The flexible ThinkSystem SR850 server can scale from two to four Intel Xeon Gold or Platinum processors, delivering significantly faster performance than the previous generation. Built for standard workloads like general business applications and server consolidation, it can also accommodate high-growth areas such as databases and virtualization. The

With the capability to support up to 48 DIMMs, four sockets, mix-and-match internal storage with up to 16 drives, and a dedicated slot for Gigabit or 10 GbE networking, the SR850 provides unmatched features and capabilities in a dense 2U rack-mount design.

More details are available here:

https://lenovopress.com/lp0645-thinksystem-sr850-server-xeon-sp-gen-1 and
https://lenovopress.com/lp1052-thinksystem-sr850-server-xeon-sp-gen-2

5.1.10 Lenovo ThinkSystem SR650

Lenovo ThinkSystem SR650 as depicted in Figure 5-17 is an ideal 2-socket 2U rack server for small businesses up to large enterprises that need industry-leading reliability, management, and security, as well as maximizing performance and flexibility for future growth. The SR650 server is designed to handle a wide range of workloads, such as databases, virtualization and cloud computing, virtual desktop infrastructure (VDI), enterprise applications, collaboration/email, and business analytics and big data.

Featuring the Intel Xeon Processor Scalable Family, the SR650 server offers scalable performance, storage capacity, and I/O expansion. The SR650 server supports up to two processors, up to 3 TB of 2666 MHz TruDDR4 memory, up to 24x 2.5-inch or 14x 3.5-inch drive bays with an extensive choice of NVMe PCIe SSDs, SAS/SATA SSDs, and SAS/SATA HDDs, and flexible I/O expansion options with the LOM slot, the dedicated storage controller slot, and up to 6x PCIe slots.

Figure 5-17 ThinkSystem SR650

The SR650 server offers basic or advanced hardware RAID protection and a wide range of networking options, including selectable LOM, ML2, and PCIe network adapters. The next-generation Lenovo XClarity™ Controller, which is built into the SR650 server, provides advanced service processor control, monitoring, and alerting functions.

More details are available here:

`https://lenovopress.com/lp0644-thinksystem-sr650-server-xeon-sp-gen1` and
`https://lenovopress.com/lp1050-thinksystem-sr650-server-xeon-sp-gen2`

5.1.11 XClarity Controller (XCC)

The SR950 server includes an xClarity Controller (XCC, which is renamed from IMM) to monitor server availability and perform remote management. The XCC is based on the Pilot4 XE401 baseboard management controller (BMC) using a dual-core ARM Cortex A9 service processor.

XCC Enterprise is included as standard, which enables remote KVM and remote media files (ISO and IMG image files), boot capture, and power capping.

The XCC monitors the following components:

- System voltages
- System temperatures
- Fan speed control
- Fan tachometer monitor
- Good Power signal monitor
- System ID and planar version detection
- System power and reset control
- Non-maskable interrupt (NMI) detection (system interrupts)
- Serial port text console redirection
- System LED control (power, HDD, activity, alerts, and heartbeat)

XCC collects inventory on the following components:

- CPU
- DIMM
- Disk
- Power supply units
- Fan
- PCI cards
- System board
- Compute board 1 and 2
- System LOM
- System firmware

XCC provides the following features:

- An embedded web server, which gives you remote control from any standard web browser
- Shared or dedicated Ethernet port. Support for LAN over USB for in-band communications to the XCC.
- A command-line interface (CLI), which the administrator can use from a Telnet or SSH session.
- Secure Sockets Layer (SSL) and Lightweight Directory Access Protocol (LDAP).
- Built-in LAN and serial connectivity that support virtually any network infrastructure.
- Multiple alerting functions with access to VPD/PFA to warn systems administrators of potential problems through email, IPMI platform event traps (PETs), and Simple Network Management Protocol (SNMP).

- Event logs that are time stamped that can be attached to e-mail alerts
- Boot Capture
- Remote mounting of ISO and IMG files
- Virtual console collaboration
- Power capping

XCC in the SR950 server supports remote management though the following interfaces:

- Intelligent Platform Management Interface (IPMI) Version 2.0
- Simple Network Management Protocol (SNMP) Version 3
- Common Information Model (CIM-XML)
- Representational State Transfer (REST) support
- Redfish support (DMTF compliant)
- Web browser - HTML 5-based browser interface (Java and ActiveX not required) using a responsive design (content optimized for device being used - laptop, tablet, phone) with NLS support

5.2 Storage subsystem

This section describes the different choices that are available for the underlying storage architecture of the Lenovo solution for SAP HANA. It explains the options customers have as part of the solution offering including support for the full solution stack that comprises all components.

Table 5-2 lists the available options as part of the Lenovo Solution for SAP HANA offerings.

Table 5-2 Storage file systems supported with the Lenovo Solution for SAP HANA

	SUSE		**RedHat**	
	Single Node	**Scale-out**	**Single Node**	**Scale-out**
XFS	Yes	No[a]	No	No
SUSE Enterprise Storage (DSS-C)	Yes	Yes	No	No
Spectrum Scale	Yes	Yes	Yes	Yes

a. if an external SAN or NAS is used (TDI), then scale-out with the XFS filesystem is also possible

The following sections first explain how additional storage is made available local on the server nodes or via SAN technology. It then continues to the three different file system choices and describes how a Lenovo solution with each file system is built.

5.2.1 Lenovo Storage D1224 Drive Enclosure for SAP HANA

Lenovo Storage D1224 Drive Enclosure models for SAP HANA provide additional required disk storage capacity for certain SAP HANA systems with large memory configurations. They are designed specifically for SAP HANA requirements and are supported on all SAP HANA systems.

On ThinkSystem SR950 systems the D1224 drive enclosure is attached to host a non-production SAP HANA instance. There is no need to attach the D1224 drive enclosure for production instance data because the internal SSDs are high-capacity devices and provide already enough storage capacity for SAP HANA needs.

The D1224 for SAP HANA is a 2U rack-mount, 12 Gbps SAS expansion enclosure that has 24 SFF hot-swap drive bays and is attached to SAP HANA systems via the Lenovo ServeRAID™ M5225 RAID adapter. Connectivity to the SAP HANA SR950 system is made via the ThinkSystem 930-8e RAID adapter. All D1224 models for SAP HANA contain single Environmental Service Module (ESM) with 3x 12 Gbps SAS x4 ports for direct-attach host connectivity.

The D1224 for SAP HANA is shown in Figure 5-18.

Figure 5-18 Lenovo Storage D1224 Disk Expansion Enclosure for SAP HANA

Table 5-3 lists the models of the D1224 drive enclosure for SAP HANA.

Table 5-3 Lenovo Storage D1224 models for SAP HANA

Description	Part number
Lenovo Storage D1224 Single ESM Drive Enclosure for SAP HANA (US English documentation)	4587E41[a]
Lenovo Storage D1224 Single ESM Drive Enclosure for SAP HANA (Simplified Chinese documentation)	4587E4C[b]
Lenovo Storage D1224 Single ESM Drive Enclosure for SAP HANA (Japanese documentation)	4587E4J[c]

a. Available worldwide (except China and Japan).
b. Available only in China.
c. Available only in Japan.

All D1224 models for SAP HANA contain one environmental service module (ESM) with three 12 Gb SAS x4 ports (Mini-SAS HD SFF-8644 connectors) for direct-attach host connectivity (daisy-chaining is not supported).

The D1224 models for SAP HANA have two redundant hot-swap 580 W AC power supplies, and they include two 1.5m, 10A/100-250V, C13 to IEC 320-C14 rack power cables.

The D1224 for SAP HANA is designed to support the SAP HANA data storage requirements with the following 2.5-inch SAS drives:

- High performance solid-state drives (3 DWD): 400 GB
- Performance-optimized, enterprise class disk drives: 1.2 TB 10K rpm

The number of drives needed depends on the selected SAP HANA configuration, as described in Chapter 6, “SAP HANA IT landscapes with Lenovo solutions” on page 99.

Table 5-4 list ordering part numbers for the drives.

Table 5-4 Lenovo Storage D1224 for SAP HANA drive options

Description	Part number
Lenovo Storage 1.2TB 10K 2.5" SAS HDD	01DC407
Lenovo Storage 400GB 3 DWD 2.5" SAS SSD	01DC482

Table 5-5 lists ordering part numbers for the external SAS connectivity cables.

Table 5-5 Lenovo Storage D1224 for SAP HANA SAS connectivity cables

Description	Part number
External MiniSAS HD 8644/MiniSAS HD 8644 2M Cable	00YL849

5.2.2 XFS technology

XFS is a high-performance journaling 64-bit file system initially developed by SGI in 1994. Since 2000 it has been licensed under the GPL. It is natively included in both SUSE and RedHat enterprise distributions and is included at no additional cost to the customer.

XFS is approved by SAP for usage with SAP HANA and provides the following key features:

- 64-bit journaling file system
- Online growth (no support for shrinking)
- Online defragmentation
- User and group quota
- Extended attributes
- Variable block sizes
- Quick restart period after unexpected interruption, regardless of the number of files
- Delayed allocation (also called allocate-on-flush)
- High scalability through allocation groups that allow parallel IO to the file system

XFS is available with all single-node SAP HANA solutions by Lenovo.

5.2.3 Lenovo ThinkSystem DM and DE series storage arrays

Lenovo provides unified, all flash storage systems that are designed to provide performance, simplicity, capacity, security, and high availability for medium (DM5000F) and large (DM70000F) enterprises. Powered by the ONTAP software, both systems deliver enterprise-class storage management capabilities with a wide choice of host connectivity options and enhanced data management features.

The DM family can be used as external SAN storage devices for SAP HANA based application landscapes. The architecture and the usage is described in detail in this book:

`https://lenovopress.com/lp1166-designing-sap-hana-solutions-using-dm-series-storage`

At the same time, Lenovo provides SAN solutions with the Lenovo ThinkSystem DE series. The technology provides scalability and advantages regarding price/performance. The models suitable for SAP HANA are described in “Lenovo ThinkSystem DE hybrid and all flash storage arrays” on page 80.

Lenovo ThinkSystem DM5000F

ThinkSystem DM5000F models as depicted in Figure 5-19 on page 79 are 2U rack-mount controller enclosures that include two controllers, 64 GB RAM and 8 GB battery-backed NVRAM (32 GB RAM and 4 GB NVRAM per controller), and 24 SFF hot-swap drive bays (2U24 form factor). Controllers provide universal 1/10 GbE NAS/iSCSI or 8/16 Gb Fibre Channel (FC) ports for host connectivity.

Figure 5-19 ThinkSystem DM5000F

A single ThinkSystem DM5000F Storage Array scales up to 144 solid-state drives (SSDs) with the attachment of Lenovo ThinkSystem DM240S 2U24 SFF Expansion Enclosures. It scales up to 2.2 PB of raw storage capacity. A cluster of the DM5000F storage systems scales up to 26.5 PB for NAS or up to 13.2 PB for SAN environments.

The ThinkSystem DM5000F offers unified file and block storage connectivity with support for 1 GbE or 10 GbE NAS and iSCSI, and 8 Gb or 16 Gb Fibre Channel protocols at the same time.

The ThinkSystem DM5000F supports the 2.5-inch 960 GB, 3.84 TB, 7.68 TB, and 15.36 TB capacity-optimized SAS SSDs. All drives are dual-port and hot-swappable.

Lenovo ThinkSystem DM7000F

ThinkSystem DM7000F models as depicted in Figure 5-20 are 3U rack-mount controller enclosures that include two controllers, and 256 GB RAM and 16 GB battery-backed NVRAM (128 GB RAM and 8 GB NVRAM per controller). Universal 1/10 GbE NAS/iSCSI or 4/8/16 Gb Fibre Channel (FC) ports and 1/10 GbE RJ-45 ports provide base host connectivity, with an option for additional 1/10 GbE or 40 GbE NAS/iSCSI, or 8/16/32 Gb FC connections with the adapter cards.

Figure 5-20 ThinkSystem DM7000F

A single ThinkSystem DM7000F Storage Array scales up to 384 SFF solid-state drives (SSDs) with the attachment of Lenovo ThinkSystem DM240S 2U24 SFF Expansion Enclosures.

Up to 12 DM7000F Storage Arrays can be combined into a clustered system in a NAS environment, or up to 6 DM7000F Storage Arrays can be combined into a clustered system in a SAN environment.

A single ThinkSystem DM7000F scales up to 5.89 PB of raw storage capacity. A cluster of the DM7000F storage systems scales up to 70.7 PB for NAS or up to 35.3 PB for SAN environments.

The ThinkSystem DM7000F offers unified file and block storage connectivity with support for 1 GbE, 10 GbE, and 40 GbE NAS and iSCSI, and 4 Gb, 8 Gb, 16 Gb, and 32 Gb Fibre Channel protocols at the same time.

Lenovo ThinkSystem DE hybrid and all flash storage arrays

Several models are supported for SAP landscapes. For SAP HANA systems, this includes Lenovo DE4000F/H & DE6000F/H. All offer block storage connectivity.

The models with the suffix F are all-flash storage systems, H stands for hybrid. DE4000 are entry level models and DE6000 are suitable for midrange use cases.

Lenovo ThinkSystem DE4000F All Flash Storage Array

ThinkSystem DE4000F models (as depicted in Figure 5-21) are available in a 2U rack form-factor and scale up to 2.94 PB of raw storage capacity. It offers block storage connectivity with support for 1/10 Gb iSCSI or 4/8/16 Gb FC, and 12 Gb SAS, 10/25 Gb iSCSI, or 8/16/32 Gb FC at the same time.

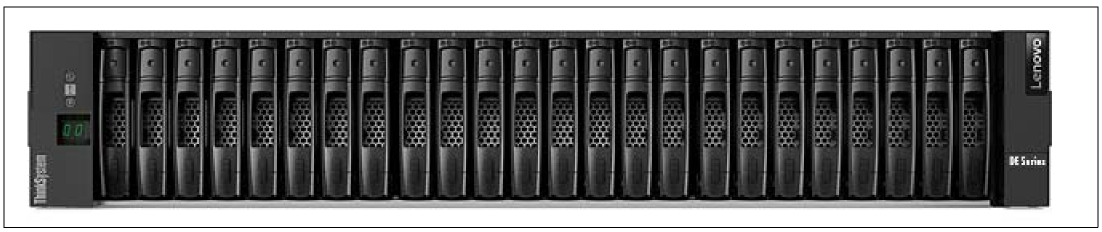


Figure 5-21 Lenovo ThinkSystem DE4000F 2U24 SFF enclosure

More information can be found here:

https://lenovopress.com/lp0909-lenovo-thinksystem-de4000f-all-flash-storage-array

Lenovo ThinkSystem DE4000H Hybrid Storage Array

ThinkSystem DE4000H models (as depicted in Figure 5-22 on page 81)are available in a 2U rack form-factor and scale up to 2.3 PB of raw storage capacity. It offers block storage connectivity with support for 1/10 Gb iSCSI or 4/8/16 Gb FC, and 12 Gb SAS, 10/25 Gb iSCSI, or 8/16/32 Gb FC at the same time.

The ThinkSystem DE4000H Storage Array scales up to 192 drives with the attachment of Lenovo ThinkSystem DE120S 2U12, DE240S 2U24 SFF, and DE600S 4U60 LFF Expansion Enclosures. It also offers flexible drive configurations with the choice of 2.5-inch (SFF) and 3.5-inch (LFF) form factors, 10 K rpm SAS and 7.2 K rpm NL SAS hard disk drives (HDDs), and SAS solid-state drives (SSDs).

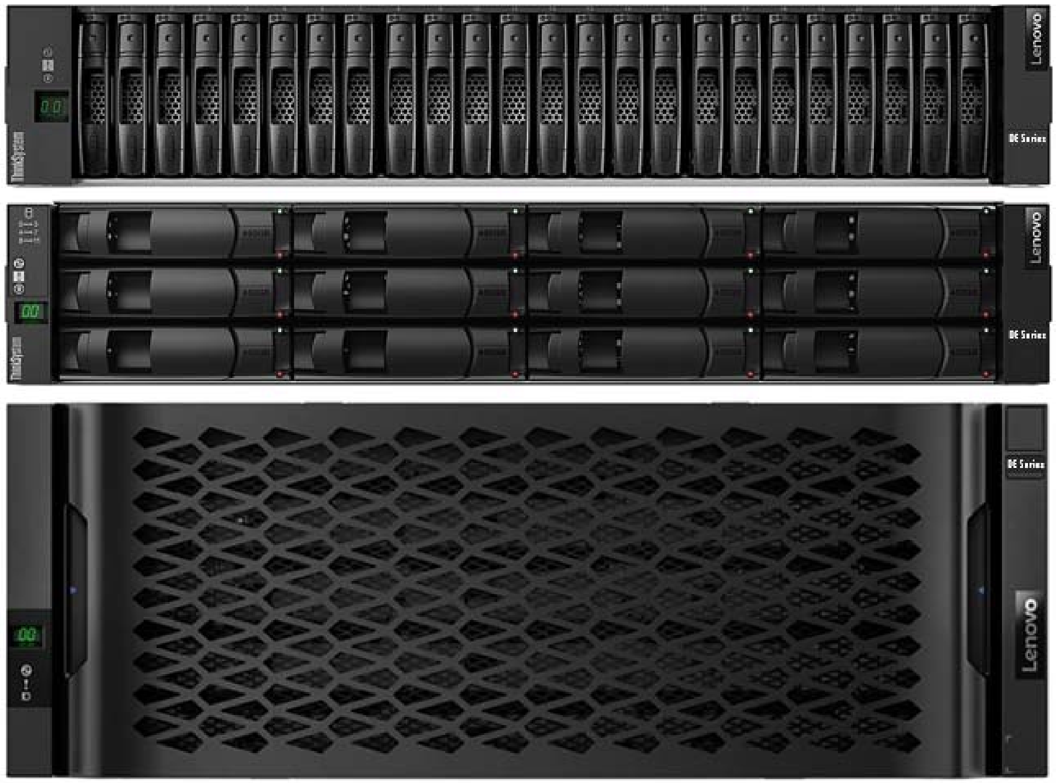


Figure 5-22 Lenovo ThinkSystem DE4000H 2U24 SFF (top), 2U12 LFF (middle), and 4U60 LFF (bottom)

More information can be found here:

`https://lenovopress.com/lp0882-lenovo-thinksystem-de4000h-hybrid-storage-array`

Lenovo ThinkSystem DE6000F All Flash Storage Array

Lenovo ThinkSystem DE6000F is a scalable, all flash mid-range storage system that is designed to provide high performance, simplicity, capacity, security, and high availability for medium to large businesses. The ThinkSystem DE6000F delivers enterprise-class storage management capabilities in a performance-optimized system with a wide choice of host connectivity options and enhanced data management features. The ThinkSystem DE6000F is a perfect fit for a wide range of enterprise workloads, including big data and analytics, video surveillance, technical computing, and other storage I/O-intensive applications.

ThinkSystem DE6000F (as depicted in Figure 5-23) models are available in a 2U rack form-factor and scale up to 2.94 PB of raw storage capacity.

Figure 5-23 Lenovo ThinkSystem DE6000F 2U24 SFF enclosure

More information can be found here:

https://lenovopress.com/lp0910-lenovo-thinksystem-de6000f-all-flash-storage-array

Lenovo ThinkSystem DE6000H Hybrid Storage Array

The ThinkSystem DE6000H (as depicted in Figure 5-24) models are available in a 2U rack form-factor. It scales up to 2.88 PB of raw storage capacity in the base configuration or up to 5.76 PB with the optional Features on Demand upgrade.

It scales up to 240 (base configuration) or 480 (optional upgrade) drives with the attachment of Lenovo ThinkSystem DE240S 2U24 SFF and DE600S 4U60 LFF Expansion Enclosures. It also offers flexible drive configurations with the choice of 2.5-inch (SFF) and 3.5-inch (LFF) form factors, 10 K rpm SAS and 7.2 K rpm NL SAS hard disk drives (HDDs), and SAS solid-state drives (SSDs).

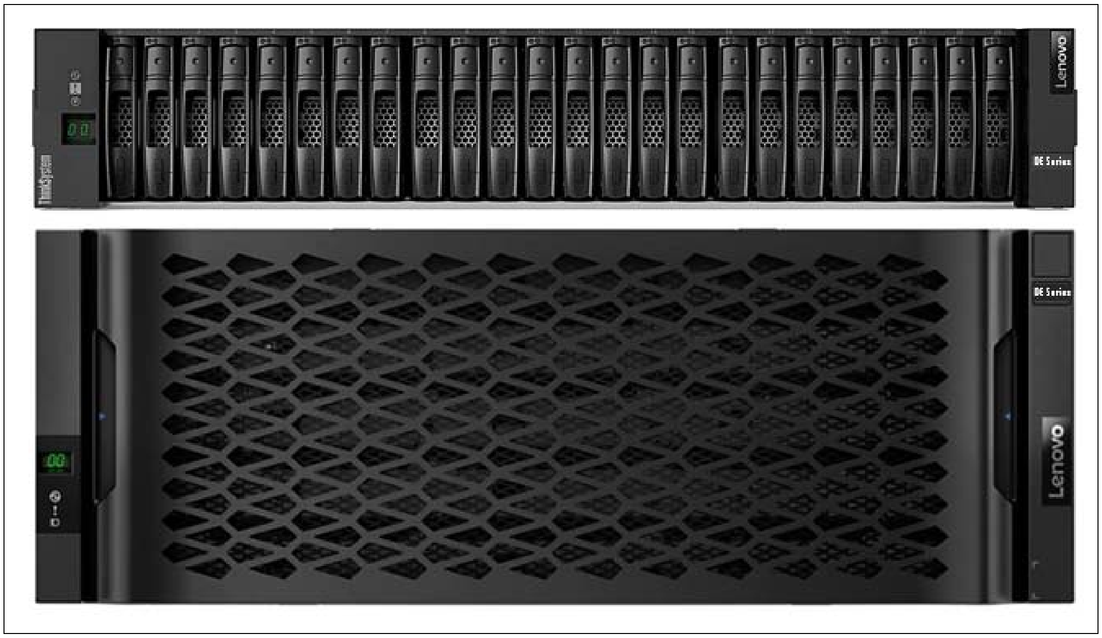


Figure 5-24 Lenovo ThinkSystem DE6000H 2U24 SFF (top) and 4U60 LFF (bottom) enclosures

More information can be found here:

https://lenovopress.com/lp0883-lenovo-thinksystem-de6000h-hybrid-storage-array

5.2.4 Software-defined storage using SUSE Enterprise Storage

In May 2016, Lenovo was first to market with a software-defined storage solution that is certified for use as enterprise storage with SAP HANA. All solutions are based on an offering from SUSE called SUSE Enterprise Storage. This solution is used as one storage option for the Lenovo solution for SAP HANA.

SUSE Enterprise Storage (SES) is a highly scalable and resilient software-defined storage offering based on the open source Ceph technology and it runs on standard off-the-shelf rack servers equipped with local storage devices like hard drives and flash storage. SUSE Linux Enterprise Server is used as the basis for the SUSE Enterprise Storage add-on. Running SUSE Enterprise Storage software on the servers turns them into a storage cluster that can be accessed through different protocols. Those servers are also called *OSD nodes*, which stands for Object Storage Device, to denote their designation as SES servers holding storage devices.

The Lenovo portfolio of solutions is called Distributed Storage Solution for Ceph (DSS-C) and there are three basic building blocks to choose from as a starting point:

- DSS-C 100, using 2.5-inch drives for standard installations
- DSS-C 101, using all-flash for highest performance
- DSS-C 103, using 3.5-inch drives for high-capacity demands

To meet the requirements for performance and availability, a minimum of three DSS-C buildings blocks are required for a production SAP HANA installation. More blocks can be added to provide more performance and support additional SAP HANA nodes or to provide more capacity for additional needs like backup.

It is also possible to enhance the SES cluster with additional components (DSS-C building blocks and storage devices per building block) to provide space for non-SAP HANA components.

Architecture for SAP HANA

The storage environment running SES software is interconnected via a fully redundant Ethernet connection. The SAP HANA servers access the storage via another fully redundant Ethernet connection. This is shown in Figure 5-25.

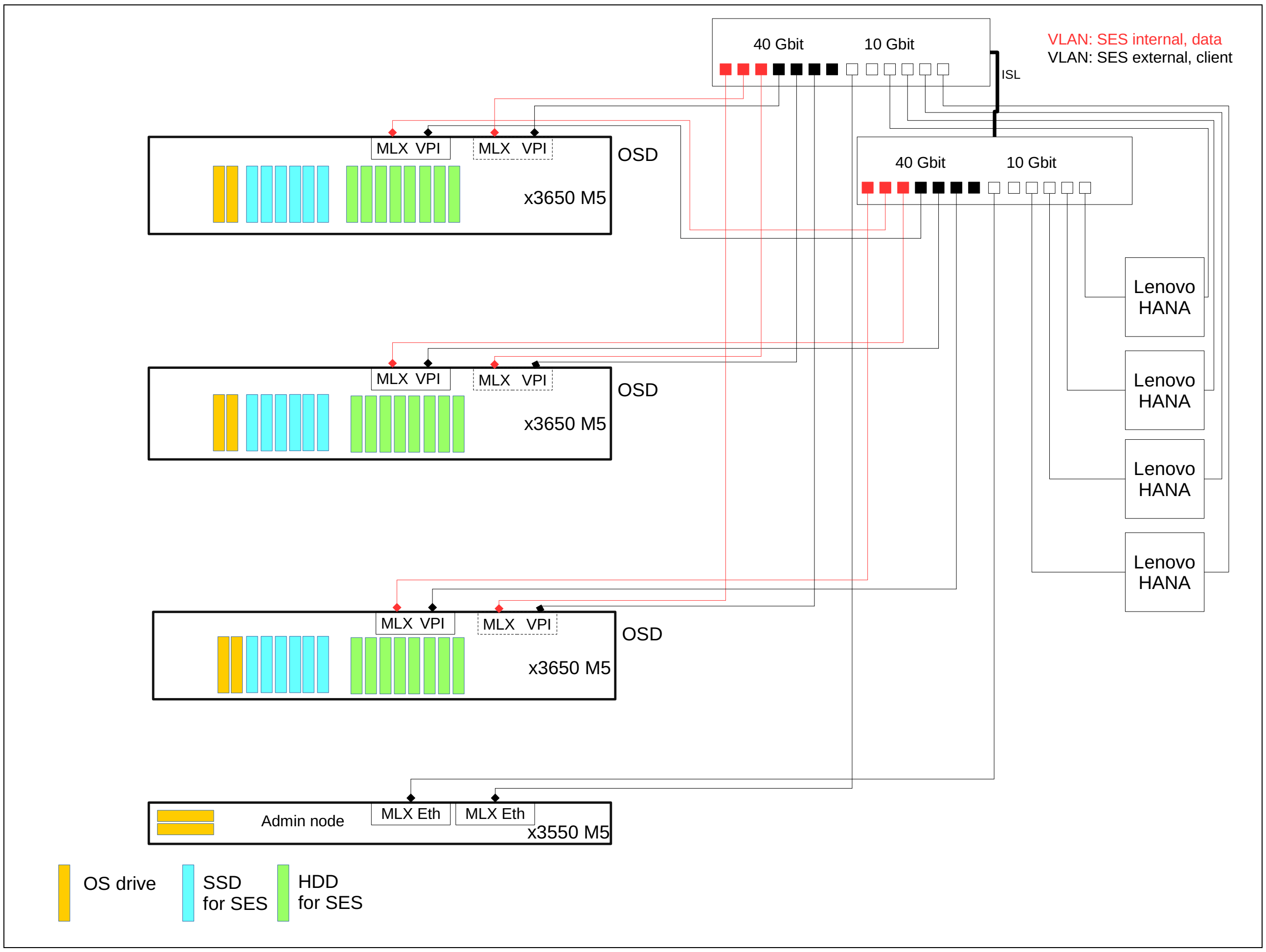


Figure 5-25 High-level architecture of a Lenovo Distributed Storage Solution for Ceph (DSS-C)

The connection from the DSS-C building block to the Ethernet switches has to be implemented with either 25 Gbit or 40 Gbit Ethernet to avoid the links becoming the bottleneck in a multi-node scale-out SAP HANA environment. The downlinks from the switches to the Lenovo servers running SAP HANA continues to be 10 Gbit Ethernet. Higher Ethernet speeds can be implemented on specific request.

For the Lenovo DSS-C building blocks the minimum amount of hard drives and SSDs is determined by Lenovo development and its configuration is described in detail in a Lenovo Implementation Guide to ensure that the final customer setup is identical to the setup that has been certified with SAP. Lenovo Professional Services implements the solution on-site together with the SAP HANA servers.

It is possible to extend a SES cluster with additional elements, either before going into production or while being in production:

- Additional DSS-C building blocks in the form of servers
- Additional networking ports on the DSS-C building blocks
- Additional storage devices on the DSS-C building blocks

One example for enhancing the predefined storage building blocks is adding 3.5-inch drives that provide extra backup space. A separate storage pool has to be created containing these extra drives.

Integration with SAP HANA

All Lenovo SAP HANA servers connect via standard TCP/IP Ethernet to the DSS-C building blocks running SUSE Enterprise Storage. The storage in the SES is grouped into a storage pool with certain features. This pool is comparable to a pool known from legacy SANs. Each SAP HANA server is provided exclusive access to two block devices in that pool and given access to a shared file system that holds its data also in the SES cluster. This is shown in Figure 5-26.

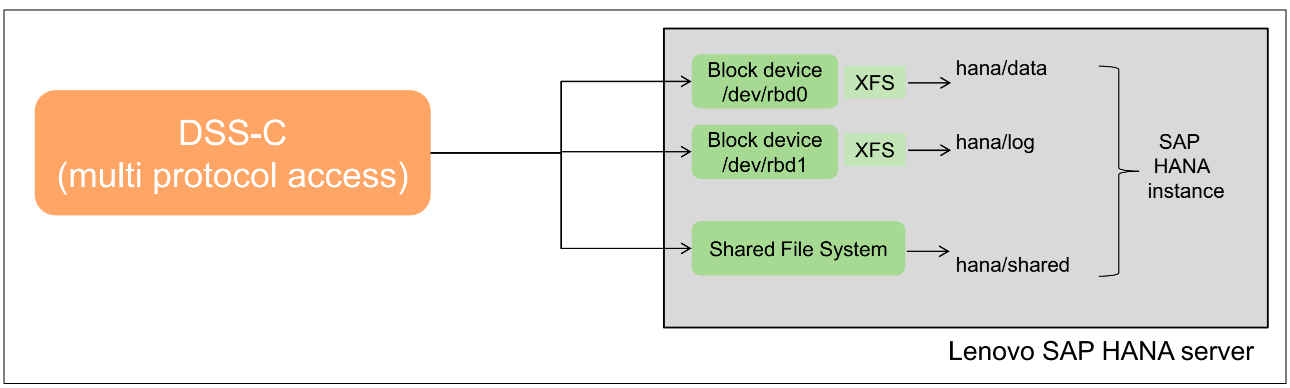


Figure 5-26 Integration of Lenovo SAP HANA with DSS-C storage

The block devices provided by the SES cluster can be compared to virtual LUNs from a storage subsystem. They are made accessible through a Linux kernel module and are called rbd0 and rbd1. They can be used like any other block device. Lenovo creates a standard XFS file system on top of these block devices and assigns them for SAP HANA data and log respectively.

The shared files of SAP HANA are made accessible through a shared file system. This can either be the integrated CephFS, which provides a POSIX-compliant access, or it can be through an additional block device that is exported via the Network File System (NFS) protocol.

Additional SAP HANA servers can access to the same shared file system but each server gets its own block devices. This enables to implement high-availability by having a standby SAP HANA server ready to take over the block devices from a failed node.

Storage features

The current version of SUSE Enterprise Storage is release 4.0 with release 5.0 being scheduled for October 2017. This update brings additional features and a major overhaul of how metadata is stored internally to further accelerate I/O performance.

With SES 4.0 the following features are included with no additional license cost:

- Data replication, including the possibility to influence onto which server in which rack the data is stored
- Geo replication (sync and async)
- Integrated cache tiering
- Snapshot support
- Thin provisioning
- Self-healing, from disk and server outages
- Encryption of data-at-rest
- Data compression
- Erasure coding
- End-to-end data checksum calculation with background scrubbing
- Support for rolling SUSE updates and SUSE kernel live patching
- Support for NFS and CIFS/Samba

Graphical interface

An integral part of SUSE Enterprise Storage is a web-based graphical user interface (GUI) that is based on the open source project called OpenAttic. It is used to manage the storage cluster and to monitors its performance. The monitoring elements is based on the open source Grafana toolset. The SUSE Enterprise Storage GUI allows to:

- Create and manage storage devices (OSDs) that are installed in the respective DSS-C building blocks
- Create and manage storage pools and its assigned features
- Manage iSCSI connectivity for iSCSI initiators
- Monitor the health of the storage cluster
- Monitor capacity usage
- Monitor performance of the storage cluster
- Monitor performance of individual storage devices (OSDs)

Figure 5-27 show the main dashboard after logging in.

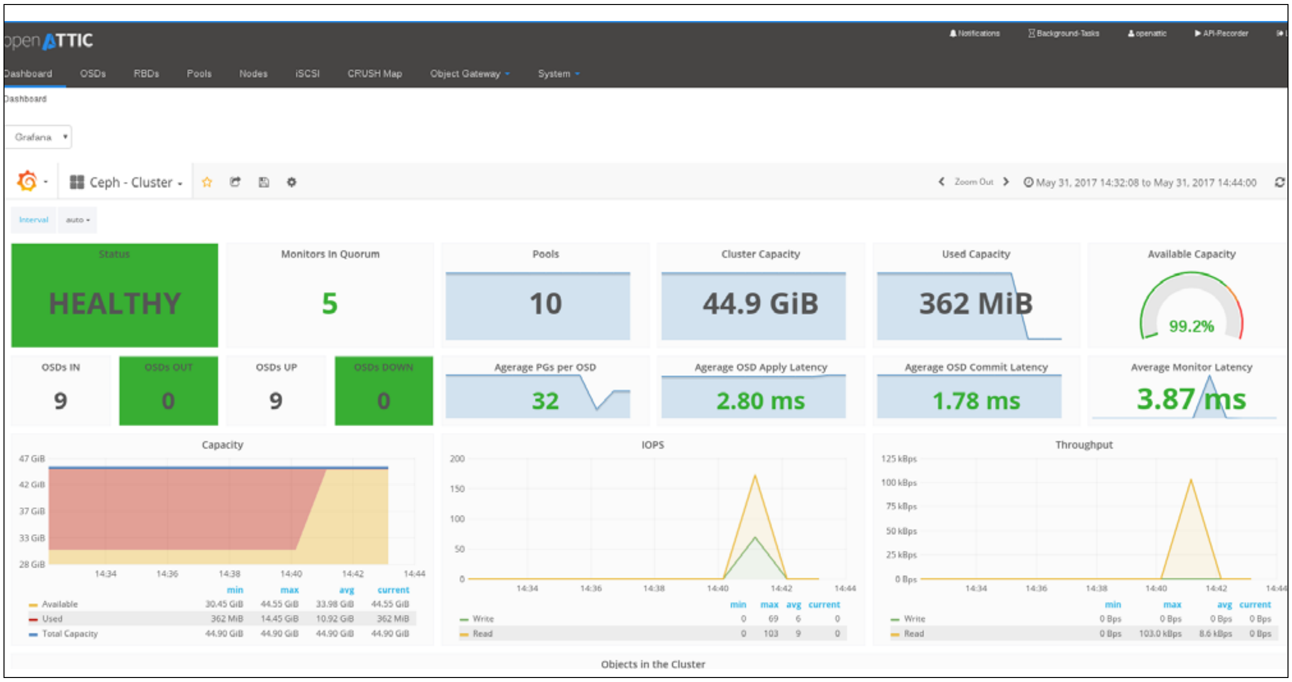


Figure 5-27 Main dashboard of the SUSE Enterprise Storage GUI

The dashboard presents a quick overview of the cluster health and its current usage and performance.

Detailed statistics can be drilled into via additional views. One example is given below in Figure 5-28 which shows the details of a single storage pool within the storage cluster.

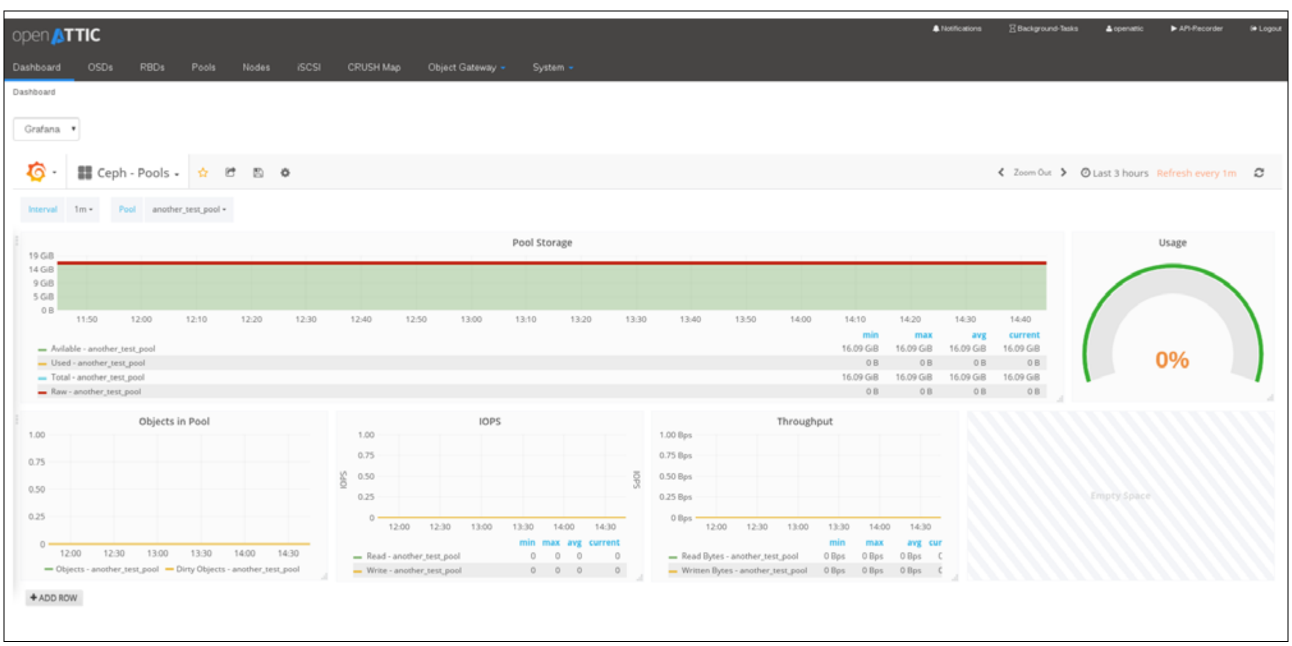


Figure 5-28 Storage pool details in the SUSE Enterprise Storage GUI

Other interfaces allow to create and manage the OSDs and to prepare virtual LUNs for access from Lenovo SAP HANA servers or additional machines.

5.2.5 IBM Spectrum Scale technology

The IBM Spectrum Scale (formerly called IBM GPFS) is a high-performance, shared-disk file management solution that provides very fast, reliable access to a common set of file data. It enables a view of distributed data with a single global namespace. It is used as one option for a scale-out Lenovo solution for SAP HANA with a shared-nothing scale-out architecture.

Common Spectrum Scale features

Spectrum Scale uses its cluster architecture to provide quicker access to your file data. File data automatically is spread across multiple storage devices, which provides optimal use of your available storage to deliver high performance.

Spectrum Scale is designed for high-performance parallel workloads. Data and metadata flow from all the nodes to all of the disks in parallel under the control of a distributed lock manager. Spectrum Scale configurations include direct-attached storage, network block I/O (or a combination of the two), and multisite operations with synchronous data mirroring.

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

Expanding beyond a storage area network (SAN) or locally attached storage, a single Spectrum Scale file system can be accessed by nodes via a TCP/IP or InfiniBand connection. Network block I/O, also called network shared disk (NSD), is a software layer that transparently forwards block I/O requests from a Spectrum Scale client application node to an NSD server node to perform the disk I/O operation and then passes the data back to the client. By using a network block I/O, this configuration can be more cost-effective than a full-access SAN.

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

Spectrum Scale provides this function by supporting the following functions:

- Data replication to increase availability if there is a storage media failure
- Multiple paths to the data if there is a communications or server failure
- File system activity logging, which enables consistent fast recovery after system failures

In addition, Spectrum Scale supports snapshots to provide a space-efficient image of a file system at a specified time, which enables online backup and can help protect against user error. How snapshots can be used to create backups of SAP HANA databases is described in 7.4.5, “Database backups by using Spectrum Scale snapshots” on page 143.

Spectrum Scale extensions for shared-nothing architectures

IBM added several features to Spectrum Scale that support the design of shared-nothing architectures. A single shared storage is not necessarily the best approach when dozens, hundreds, or even thousands of servers must access the same set of data. Shared storage can impose a single point of failure unless designed in a fully redundant way by using storage mirroring to a secondary storage building block.

Spectrum Scale File Placement Optimizer (Spectrum Scale FPO) is a set of features to support big data applications on shared-nothing architectures. In such scenarios, hundreds or even thousands of commodity servers compute certain problems. They do not include shared storage to hold the data. The internal disks of the nodes are used to store all data, which requires a new way of thinking to run a cluster file system on top of a shared-nothing architecture.

The following features were introduced with Spectrum Scale FPO and are relevant for the Lenovo SAP HANA solution:

- Write affinity: Provides control over the placement of new data. It can be written to the local node or wide striped across multiple nodes.
- Locality awareness: The ability to obtain on which node certain data chunks are stored. This ability allows jobs to be scheduled on the node that is holding the data, which avoids costly transfer of data across the network.
- Pipelined replication: Makes the most effective use of the node interconnect bandwidth. Data that is written on node A sends data to node B, which in turn sends data to node C. In contrast to pipelined replication, the other replication schema is star replication, where node A sends data to both node B and node C. For bandwidth-intense operations or for servers with limited network bandwidth, the outgoing link of node A can limit replication performance in such a scenario. Choosing the correct replication schema is important when a shared-nothing architecture is run because this process almost always involves replicating data over the network.
- Fast recovery: An intelligent way to minimize recovery efforts after the cluster is healthy again. After an error, Spectrum Scale tracks the updates that are missing through the failed drives. In addition, the load to recover the data is distributed across multiple nodes. Spectrum Scale also allows two different recovery policies. After a drive fails, data can be rebuilt when the drive is replaced or it can immediately be rebuilt by using other nodes or disks to hold the data.

The Lenovo solution for SAP HANA benefits in the following ways from the features of Spectrum Scale:

- Spectrum Scale provides a stable, industry-proven, cluster-capable file system for SAP HANA.
- Spectrum Scale transparently works with multiple replicas (that is, copies) of a single file to protect from drive failures.
- Spectrum Scale adds extra performance to the storage devices by striping data across devices.
- With the FPO extensions, Spectrum Scale enables the Lenovo solution for SAP HANA to grow beyond the capabilities of a single system, into a scale-out solution, without the need for external storage.
- Spectrum Scale adds high-availability and disaster recovery features to the solution.

5.2.6 Single node SAP HANA using XFS

Lenovo offers single node SAP HANA solutions based on the XFS file system. Further drive and file system setup is handled automatically by the installation wizard from Lenovo.

Figure 5-29 on page 89 shows the storage architecture on a single node ThinkSystem solution running XFS. Linux-internal tools manage the different sizes of the block devices and balances I/O operations to maximize the use of both devices.

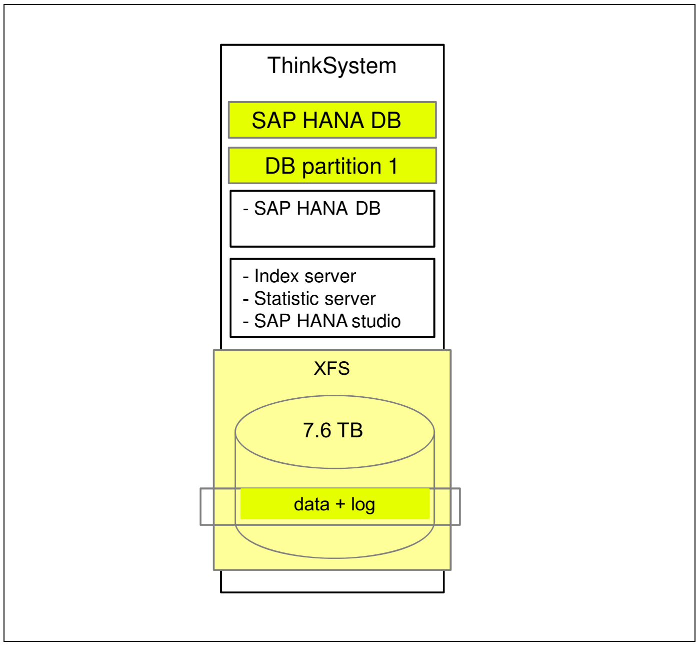


Figure 5-29 Storage Architecture for a single node ThinkSystem with XFS

Single node installations based on XFS require a reinstallation of SAP HANA when turned into scale-out due to the change of file system to Spectrum Scale.

5.2.7 Scaling out SAP HANA using Spectrum Scale

By scaling up a single node SAP HANA appliance, you can expand the capabilities of an SAP HANA installation up to a certain point when the physical limit is reached. To allow for further growth, the Lenovo System Solution for SAP HANA supports a scale-out approach. An SAP HANA system can span multiple servers, partitioning the data to hold and process larger amounts of data than a single server can accommodate.

All scale-out solutions are based on the same building blocks, as described in section 5.1, “Lenovo ThinkSystem Servers” on page 60. All Lenovo scale-out solutions for SAP HANA have the following properties:

- The scale-out solution is a cluster of servers, which are interconnected with two separate 10 Gb Ethernet networks, one for the SAP HANA application and one for the shared Spectrum Scale file system communication. Both networks are redundant.
- The SAP HANA database is split into partitions on each cluster node, which forms 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 Spectrum Scale file system is a shared file system. Because Spectrum Scale spans all nodes of the cluster, it makes the data of each node available to all other nodes in the cluster despite using only local storage devices (for more information about this technology, see, “Spectrum Scale extensions for shared-nothing architectures” on page 87).

To an outside application that is connecting to the SAP HANA database, this configuration appears to be 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 requester. Each node maintains its own set of data, persisting it with savepoints and logging data changes to the database log that are stored on local storage.

Spectrum Scale combines the storage devices of the individual nodes into one large file system, ensuring that the SAP HANA software has access to all data regardless of its location in the cluster. Spectrum Scale also ensures 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. This feature is called *locality*.

Although Spectrum Scale provides the SAP HANA software with the functionality of a shared storage system, it ensures maximum performance and minimum latency by using locally attached drives and flash devices.

In addition, because server-local storage devices are used, the total capacity and performance of the storage within the cluster automatically increases with the addition of nodes, which maintains the same per-node performance characteristics regardless of the size of the cluster. This kind of scalability is not achievable with external storage systems.

Maximum scalability:

- With SR950 nodes, SAP validated the Lenovo scale-out solution for up to 94 nodes in a cluster.
- The building block approach of Lenovo makes the solution virtually scalable without any known limitations. Clients that require scale-out configurations beyond the generally available 94 nodes can work with Lenovo and SAP to jointly validate such large clusters at the client site.
- Lenovo has shown scalability for up to 224 nodes in a single SAP HANA scale-out cluster.

Scaling out a Lenovo SAP HANA solution creates a cluster of nodes. SAP HANA designates nodes in a scale-out configuration with a certain role. The node can be a worker node or a standby node. Worker nodes actively process workload. Standby nodes are part of the cluster only and do not process workload while the cluster remains in a healthy state. Standby nodes take over the role of a worker node when it fails. Standby nodes are required for scale-out clusters with high availability.

Figure 5-30 on page 91 shows the networking architecture of a four-node scale-out solution using Spectrum Scale. The node designation has no effect on the network connectivity of a node. All nodes are considered equal.

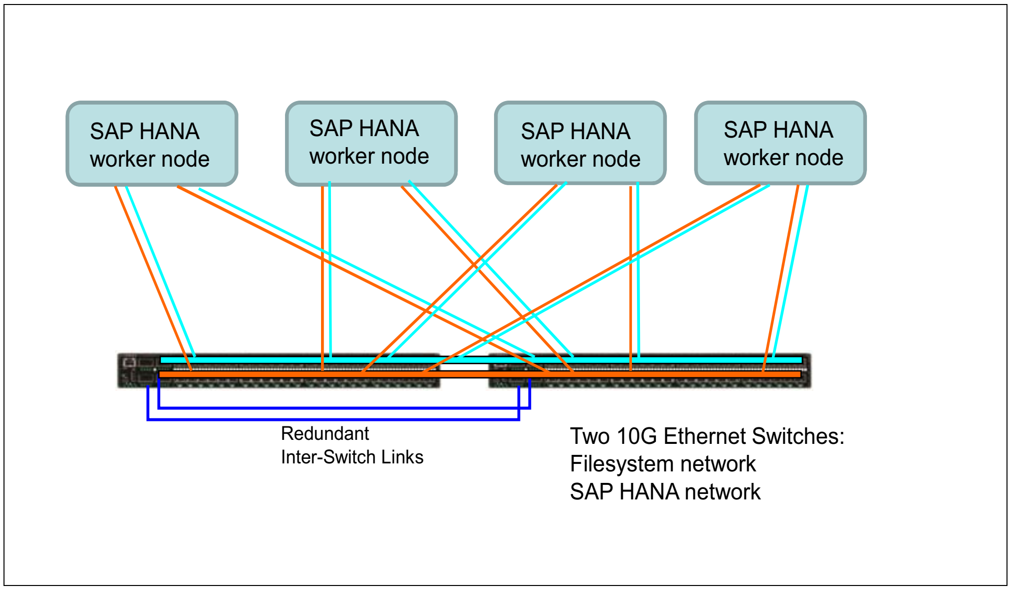


Figure 5-30 Network architecture of a four-node scale-out solution using Spectrum Scale FPO

The following networks span the redundant Ethernet switches:

- Spectrum Scale network, for communication and data transfer between the nodes
- SAP HANA network for database communication

Every node has redundant connectivity to each of the two networks, which leads to four 10 Gbps Ethernet ports that are required per node in scale-out environments. If the SAP HANA database instance that is running on those nodes grows, clients can add nodes to extend the overall main memory of the cluster. This addition is possible without affecting any of the existing nodes, so the cluster does not have to be taken down for this operation.

Scale-out environments with only SAP HANA worker nodes have no support for HA because no standby nodes are part of the cluster.

The absence of failover capabilities represents a major disadvantage of this solution. The cluster acts as a single-node configuration. If one node becomes unavailable for any reason, the database partition on that node becomes unavailable, and with it the entire SAP HANA database. To cover the risk of a node failure, a standby node must be added to the cluster. This solution is described in the next section.

Scale-out solution with high-availability capabilities

The scale-out solution for SAP HANA with high-availability capabilities enhances the scale-out solution in the following ways:

- Making the SAP HANA database services highly available by introducing SAP HANA standby nodes, which can take over from a failed node within the cluster.
- Making the data that is provided through Spectrum Scale highly available to the SAP HANA application, including its data on the local storage devices. This availability allows you to tolerate the loss of a node.

SAP HANA allows the addition of nodes in the role of a standby node. These nodes run the SAP HANA DB service, but do not hold any data in memory or take an active part in the processing. If one of the active worker nodes fails, a standby node takes 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 operating.

To take over the database partition from the failed node, the standby node must 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 process is possible because Spectrum Scale provides a shared file system across the entire cluster, which gives each individual node access to all the data that is stored on the storage devices that are managed by Spectrum Scale.

A four-node cluster with the fourth node in a standby node is shown in Figure 5-31.

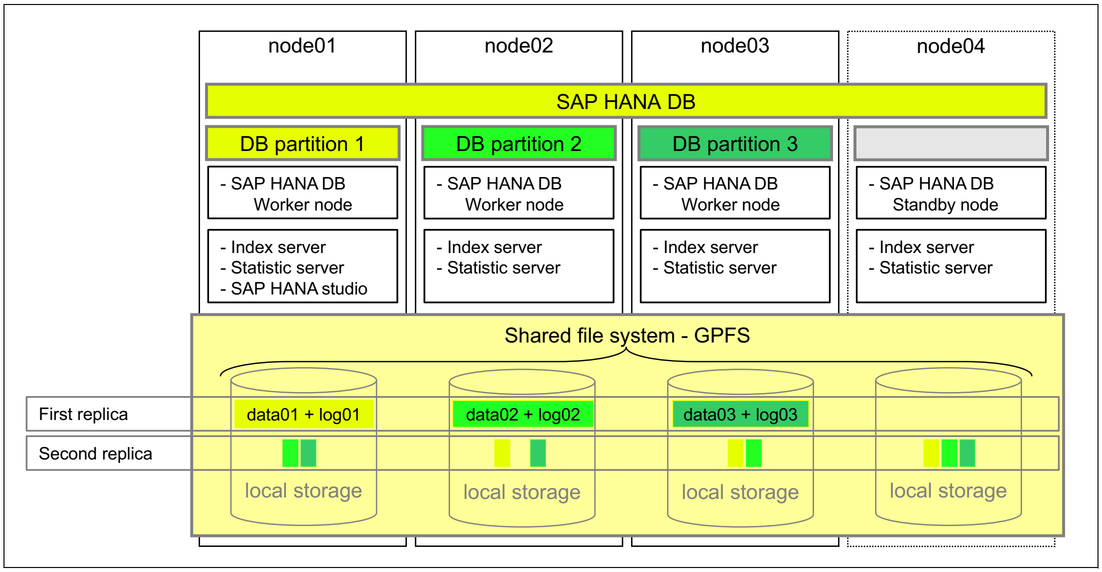


Figure 5-31 Four-node clustered solution with failover capabilities

If a node has an unrecoverable hardware error, the storage devices that are holding the node's data might become unavailable or destroyed. With high-availability features in place, the Spectrum Scale file system replicates the data of each node to the other nodes, which creates a second replica to prevent data loss if one of the nodes fails.

Replication is done in a striping fashion. Every node has a piece of data of all other nodes. In Figure 5-31, 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.

Replication happens for all nodes that are generating data so that all information is available twice within the Spectrum Scale file system, which makes it tolerant to the loss of a single node. Replication occurs synchronously. The write operation finishes only when the data is written locally and on a remote node. This configuration ensures consistency of the data at any point. Although Spectrum Scale replication is done over the network and in a synchronous fashion, this solution still overachieves the performance requirements for validation by SAP.

The File Placement Optimizer (FPO), which is part of Spectrum Scale, ensures that the first replica always is stored local to the node that is generating the data. If SAP HANA data must be read from disk (for example, for backups or restore activity), FPO always prefers the replica that is available locally. This configuration ensures the best read performance of the cluster. By using replication, Spectrum Scale provides the SAP HANA software with the functionality and fault tolerance of a shared storage subsystem while maintaining its performance characteristics.

Example of a node takeover

To further show the capabilities of this solution, this section provides a node takeover example. In this example, we have a four-node setup (initially configured as shown in Figure 5-31 on page 92) with three active nodes and one standby node.

First, node03 experiences a problem and fails unrecoverably. Data that is stored on this node is no longer available. The SAP HANA master node (node01) recognizes this fact and directs the standby node (node04) to take over from the failed node. The standby node is running the SAP HANA application and is part of the cluster, but in an inactive role.

To re-create database partition 3 in memory to take over the role of node03 within the cluster, node04 reads the savepoints and database logs of node03 from the Spectrum Scale, reconstructs the savepoint data in memory, and reapplies the logs so that the partition data in memory is as it was before node03 failed. Node04 is operating and the database cluster recovered. This scenario is shown in Figure 5-32.

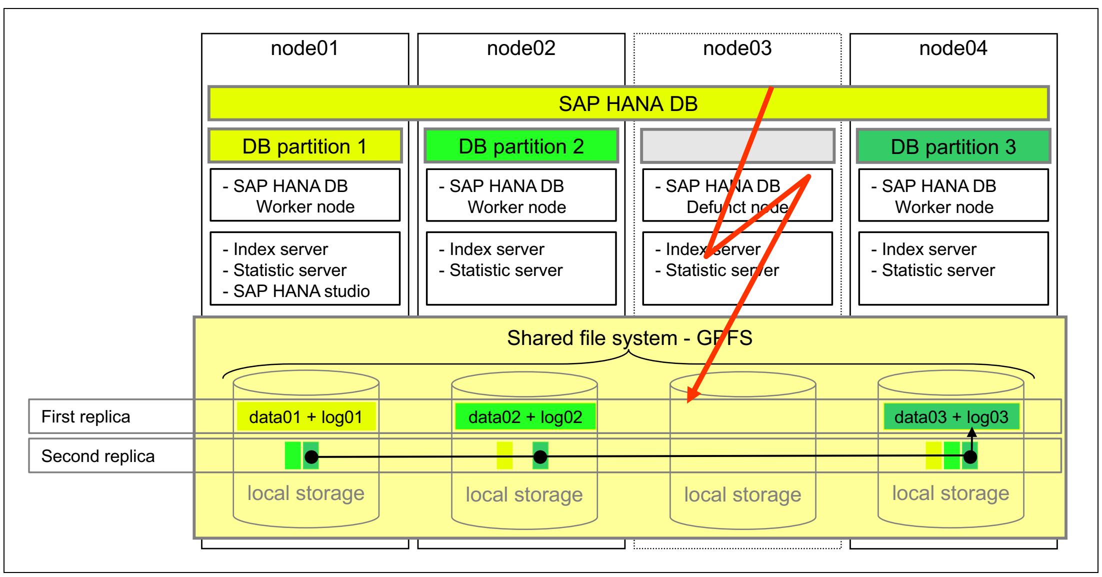


Figure 5-32 Standby node 4 takes over from failed node 3

The data that node04 must load into memory is the data of node03 (which failed), including its local storage devices. For that reason, Spectrum Scale had to deliver the data to node04 from the second replica, which is spread across the cluster. Spectrum Scale handles this process transparently so that the application does not recognize from which node the data was read. If data is available locally, Spectrum Scale prefers to read from node04 and avoid going over the network.

Now, when node04 starts writing savepoints and database logs again during the normal course of operations, these savepoints are not written over the network. Instead, the savepoints are written local drives with a second replica striped across the other cluster nodes.

After fixing the cause for the failure of node03, it can be reintegrated into the cluster as the new standby system, as shown in Figure 5-33.

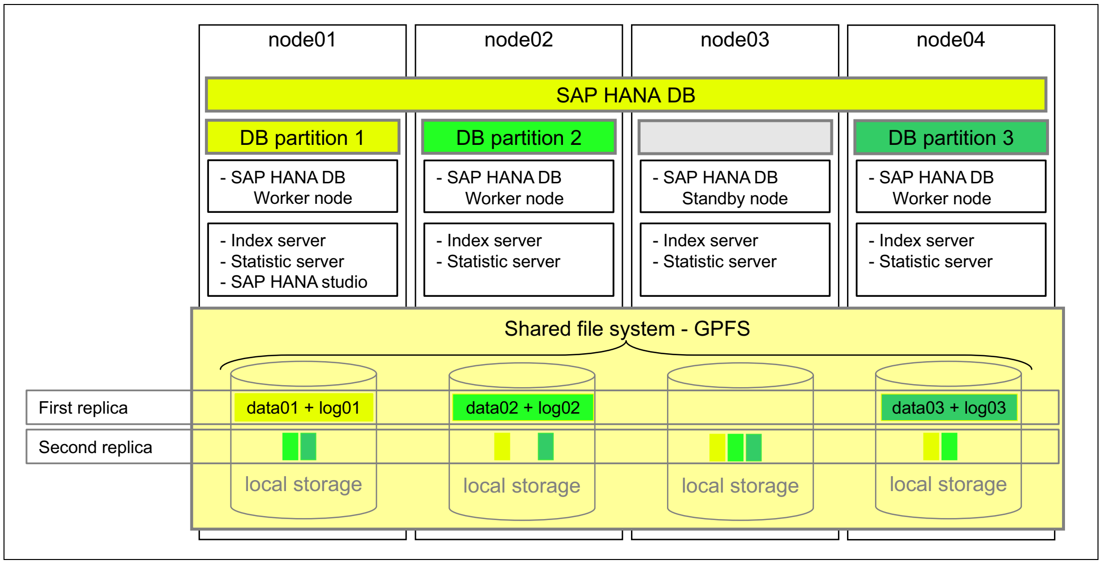


Figure 5-33 Node 3 is reintegrated into the cluster as a standby node

This example shows how Lenovo combines two independently operating high-availability measures (that is, the concept of standby nodes at the SAP HANA application level and the reliability features of Spectrum Scale at the infrastructure level), which results in a scalable solution that provides fully automated high availability with no administrative intervention required.

5.3 Networking options

Larger SAP HANA implementations scale beyond the limits of a single server. In those environments, the database is split into several partitions, with each partition on a separate server within the cluster. Nodes in a cluster communicate with each other through a high-speed interconnect. Network switches are crucial in such scale-out solutions.

All Lenovo solutions for SAP HANA use network switches that meet these requirements. The following top-of-rack Ethernet switches are part of the scale-out solution of Lenovo:

- Lenovo RackSwitch™ G8296

  The G8296 switch (in a 2U chassis) is a 10/40 Gb Ethernet switch with up to 94 10 Gb SFP+ ports. This switch is used in scale-out solutions to provide internal cluster communication for Spectrum Scale and SAP HANA networks for maximum solution scalability of up to 94 nodes.

- Lenovo RackSwitch G8272

  The G8272 switch is a 10/40 Gb Ethernet switch with up to 72 10 Gb SFP+ ports in a 1U rack form factor. This switch is used in scale-out solutions to provide internal cluster communication for Spectrum Scale and SAP HANA networks for maximum solution scalability of up to 64 nodes.

- Lenovo ThinkSystem NE2572 RackSwitch

  The NE2572 RackSwitch has 48x SFP28/SFP+ ports in a 1U rack form factor that support 10 GbE SFP+ and 25 GbE SFP28 optical transceivers, active optical cables (AOCs), and direct attach copper (DAC) cables. This switch is the Lenovo recommendation for deployments of SAP HANA scale-out cluster. Detailed testing has been performed and all exact setup and configuration is part of the Lenovo SAP HANA deployment collateral.

- Lenovo RackSwitch NE10032

  The NE10032 RackSwitch has 32x QSFP+/QSFP28 ports in a 1U rack form factor that support 40 GbE and 100 GbE optical transceivers, active optical cables (AOCs), and direct attach copper (DAC) cables. This switch is the Lenovo recommendation for large scale-out deployments, or for installations requiring maximum throughput (ideal for S/4 deployments when used with 100 Gb Ethernet connectivity).

- Lenovo Switch NE0152T

  The NE0152T RackSwitch has 48x RJ-45 Gigabit Ethernet fixed ports and 4x SFP+ ports in a 1U rack form factor that support 1 GbE and 10 GbE optical transceivers, active optical cables (AOCs), and direct attach copper (DAC) cables. Featuring a small, 1U footprint, this switch is designed for management network and access-layer deployments in data center infrastructures. This switch is the Lenovo recommendation for administration and management connectivity of Lenovo SAP HANA deployments.

More information on all of the networking components is available under

https://lenovopress.com/

Go to the tab called Networking and select “Top of Rack Connectivity” with the according class like 25 Gbit or 100 GBit for example.

5.4 Workload Optimized Solution for SAP HANA

The Lenovo workload-optimized solutions based on the ThinkSystem SR950 server cover the full range of configurations from two sockets to four sockets and eight sockets. Customers can start with a smaller configuration and then scale-up over time when the database size grows by adding only new components.

Continuing its previous X6 solution architecture and portfolio, Lenovo offers all-flash configurations for SAP HANA solutions based on the SR950 server. A minimum Lenovo SR950 workload-optimized solution for SAP HANA has the following components:

- Two, four, or eight Intel Xeon SP Platinum 8176 / 8276 CPUs
- 192 GB, 384 GB, or 768 GB of DDR4 main memory
- Two 400 GB SAS SSDs in a RAID 1 for the operating system (SLES) and local housekeeping
- Three 3.84 TB Capacity SAS SSDs in a RAID 5 configuration holding database data
- One internal storage controller 930-16i with 4 GB flash attached
- Two dual-port Mellanox ConnectX-4 Lx adapter, running at 10 Gbps speed but capable of 25 Gbps Ethernet
- One quad-port Intel I350-T4 1 Gbps RJ45 Ethernet adapter

When your database size grows and you must scale-up your SR950 solution you can add any of the following components:

CPU and memory

SAP supports the usage to the following CPU options (exemplary for Gen1) for SAP HANA:

- Platinum 8176, 28 cores at 2.1 GHz and support of 768 GB memory per socket
- Platinum 8176M, 28 cores at 2.1 GHz and support of 1.5 TB memory per socket
- Platinum 8180, 28 cores at 2.5 GHz and support of 768 GB memory per socket
- Platinum 8180M, 28 cores at 2.5 GHz and support of 1.5 TB memory per socket

For support of more than 768 GB of main memory per CPU socket it is required to have the Intel *M* CPUs installed. When ordering the system the upgrade path must already be kept in mind.

The following main memory modules (exemplary for Gen1) are supported in the Lenovo solution for SAP HANA:

- 16 GB TruDDR4 2666 MHz (1Rx4 1.2V) RDIMM
- 16 GB TruDDR4 2666 MHz (1Rx4 1.2V) RDIMM
- 32 GB TruDDR4 2666 MHz (2Rx4 1.2V) RDIMM
- 64 GB TruDDR4 2666 MHz (4Rx4 1.2V) LRDIMM
- 128 GB TruDDR4 2666 MHz (8Rx4 1.2V) 3DS RDIMM

Mixing RDIMMs and LRDIMMs is not supported. Mixing 3DS RDIMMs with either RDIMMs or LRDIMMs is also not supported.

Storage

SAP HANA data is persisted on a storage layer, which is inside the SAP HANA server or in an external SAN or in software defined storage.

Internal

The following storage devices are used as part of the Lenovo solution for SAP HANA:

- 400 GB Mainstream SAS 12 Gb Hot Swap SSD
- 800 GB Mainstream SAS 12 Gb Hot Swap SSD
- 3.84 TB Capacity SAS 12 Gb Hot Swap SSD
- 7.68 TB Capacity SAS 12 Gb Hot Swap SSD

Additional storage devices can only be installed in accordance with the guidelines of the Lenovo solution for SAP HANA. Configuration details can be found in the tables listing all models in 6.1, "Lenovo SR950 (Gen 1)" on page 100.

The operating system is installed on the 400 GB SSDs and SAP HANA data and logs are stored on either the 3.84 TB or the 7.68 TB options. The devices are installed starting from the lower left bays of the server and connect internally to a 930-16i RAID adapter with 4 GB of flash cache.

Models with more memory have a requirement for more capacity. Additional drives are added to the RAID 5 (or RAID 6, respectively) in order to support this demand. The maximum is nine drives in a single RAID array for models with 3 TB, 6 TB, or 12 TB of main memory.

The 7.68 TB drives are required for the biggest scale-up configurations to provide enough internal storage capacity. They can also be used as an option in lower-end configurations. For details, see 6.1, "Lenovo SR950 (Gen 1)" on page 100. Extra hot-spare drives are supported. They can be added without affecting the overall performance of the solution.

External storage is supported to hold data for non-productive SAP HANA instances. It uses the D1224 storage enclosure connected via a quad-lane 12 Gbps SAS link to a 930-8e RAID controller. The storage enclosure can hold either the 3.84 TB or 7.68 TB SSDs. The RAID controller is installed in PCIe slot 7 of Figure 5-34 on page 98.

Lenovo DM storage array

The solution uses the NFS protocol and consist of at least one Lenovo DM5000F/7000F series all-flash storage array for SAP HANA data. More capacity and performance can be added with additional storage arrays or with the attachment of a storage expansion enclosure (called Lenovo DM240S).

Lenovo SAP HANA servers have their OS installed on local disks. For the SAP HANA persistency layer they are provided with NFS shared over Ethernet. There are dedicated NFS volumes for SAP HANA data and log directories of each node in a scale-out cluster. The HANA share directory is also a dedicated volume but it is the same volume across all cluster nodes.

Network

Lenovo solutions for SAP HANA have several different network interfaces that can be grouped into the following modes:

- Internal communication (HANA communication and cluster file system communication). Redundancy is required.
- External communication (SAP data management, SAP client access, data replication, appliance management, and others, depending on the customer landscape). Redundancy is optional.

Internal communication remains internal within scale-out SAP HANA solutions. They have no connection to the customer network. The networking switches for these networks are part of the solution and cannot be substituted with other than the validated switch models.

On the SR950 solutions, Lenovo uses two dual-port Mellanox ConnectX-4 Lx adapters that are running in 10 Gbit Ethernet mode for both internal communication networks. Each of the two networks requires its own connection on two physical interfaces to allow for redundancy if there is a hardware failure in the network. The Lenovo System Networking RackSwitch G8272 is used as the switch for internal communication. One switch handles traffic, HANA, and cluster file system communication. To allow for a switch failure, a second G8272 is required in scale-out solutions. For more information about the switches, see 5.3, “Networking options” on page 94.

The Mellanox ConnectX-4 Lx adapter is capable to run 25 Gbit Ethernet. To use the additional speed it is only required to replace the transceivers in the network adapter to support a higher speed.

For the uplink into the client network, an Intel quad-port 1 Gigabit Ethernet adapter is included as the default adapter. Additional adapters can be added if more ports are required (for access to more networks or for redundancy reasons) or if, for example, 10 Gigabit is also required towards the customer networks.

Figure 5-34 shows the back side of an SR950 workload-optimized solution for SAP HANA.

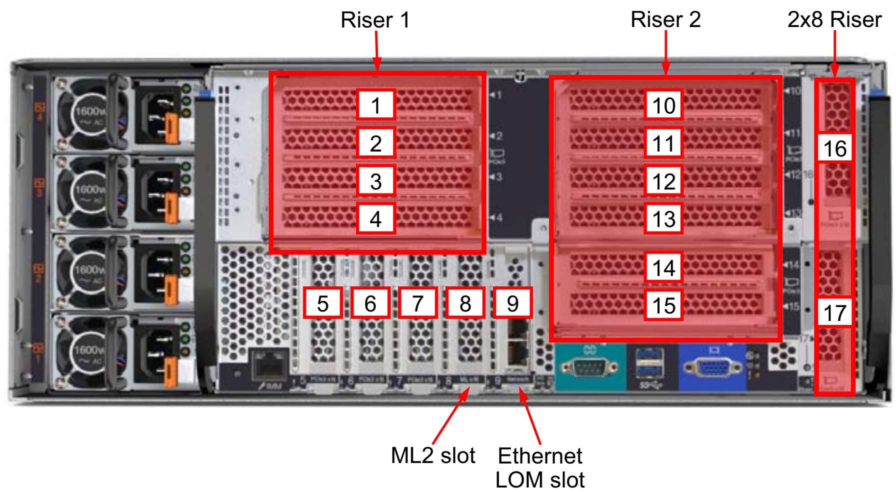


Figure 5-34 Rear view of ThinkSystem SR950 with PCIe slots numbered

The Lenovo solution has a Mellanox ConnectX-4 Lx dual-port 25 Gbit adapter in slot 5 and slot 6 of the machine. This network adapter runs in 10 Gbit mode per guidance of SAP but if additional network capacity is required then this can be easily enabled by replacing 10 Gbit transceivers (or DAC cables) with 25 Gbit transceivers (or DAC cables, respectively).

One port of each adapter is used to form a redundant, bonded network channel. One of the virtual ports carries SAP HANA internal communication traffic and the other virtual port is used for cluster file system communication.

The ML2 slot (slot 8) holds a quad-port 1 Gbit Intel I350-T4 network adapter for administrative access, backup, or other traffic. Additional adapters can be installed according to specific needs.

Virtualization

SAP HANA can run in a virtualized setup on Intel Xeon SP Gen 1 and 2 based architectures.

The details need to be carefully evaluated and a good starting point for that is here:

https://www.vmware.com/solutions/business-critical-apps/sap-virtualization.html

and in the SAP Notes:

2315348 - SAP HANA VM on VMware vSphere 6 in production

2393917 - SAP HANA VM on VMware vSphere 6.5, 6.7 and 6.7U2 in production

For more details please refer to chapter 6.7.3, “SAP HANA on VMware vSphere” on page 119.

6

SAP HANA IT landscapes with Lenovo solutions

This chapter describes IT landscapes in which SAP HANA can be deployed and shows the corresponding Lenovo workload-optimized solution that is based on the building blocks that are introduced in Chapter 5, “Components of the Lenovo Solution for SAP HANA” on page 59.

This chapter also describes implementations based on the latest ThinkSystem SR950 servers.

This chapter also describes solutions how to consolidate multiple database instances on one SAP HANA system and introduces data encryption for SAP HANA solutions.

This chapter includes the following topics:

Concept of short names for the Lenovo solutions for SAP HANA

To simplify the communication and avoid confusion, Lenovo introduced a generic naming schema for servers, which are used for SAP HANA. This was continued and enhanced from the previous processor generations.

The schema consists of five elements that are separated by dashes (can also be left out) leading to a structure of A-BBB-CCCCC-DDD-E, where:

- A is the number of installed processors (or sockets)
- BBB is the processor architecture (at the time of writing this is SKL or CSL)
- CCCCC is the memory amount in gigabytes

- DDD denotes the storage configuration (e.g. SSD)
- E denotes the file system, this can be single node using XFS (S), Spectrum Scale / formerly GPFS (G), or DSS-C (C)

For example, model 8-SKL-12288-SSD-S represents a single node with eight sockets with Intel Xeon SP first generation CPUs and 12 TB of main memory. The system uses flash devices (SSDs) to store data locally on a node and runs the XFS file system.

A model 4-CSL-1536-SSD-S represents a 4-socket cluster node with Intel Xeon SP second generation CPUs, 1.5 TB of main memory and an all-flash storage configuration using the XFS file system.

6.1 Lenovo SR950 (Gen 1)

The latest generation of Intel Xeon processors was announced in July 2017 and are part of a new family called Intel Xeon Scalable Processor (SP). A bin system is used to distinguish certain processors. The listed Lenovo configurations follow the requirements of SAP to use only the highest bin called Platinum which is the only bin that provides eight-way capable CPUs.

Figure 6-1 shows the Lenovo portfolio of workload-optimized solutions for SAP HANA based on the ThinkSystem SR950 server (Generation 1) including potential scalability options.

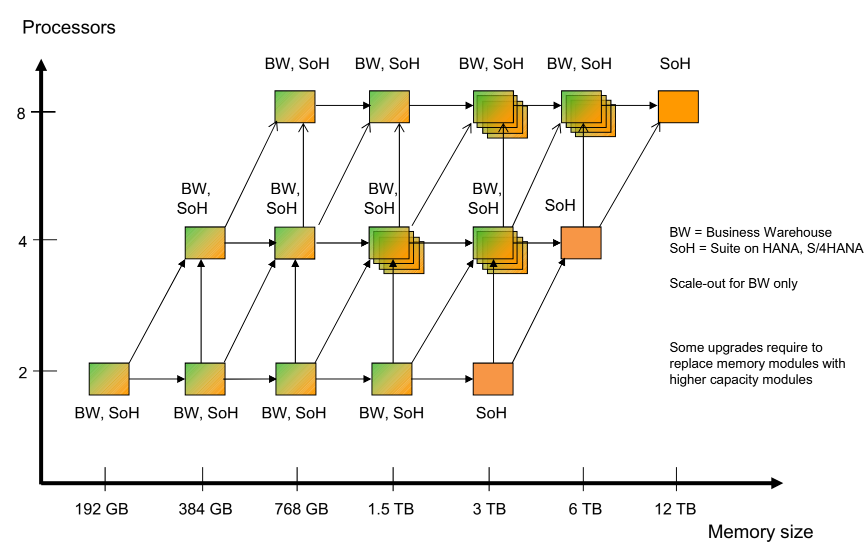


Figure 6-1 SR950 (Gen 1) based building blocks for SAP HANA

Notes:

According to the options of TDI Phase 5, it is possible to have more main memory in a server, or select a different CPU type (Gold or Silver, when we consider SR650 for example) for SAP HANA systems. That is fully supported as well, as long as a reasonable sizing exercise was done.

6.1.1 Single-node SR950 solution for analytical and transactional use cases

Single node SAP HANA solutions become more and more popular because of the increased memory capacity that can be deployed before the need for scaling out arises. This advancement in technology is reflected in the next three tables below covering single node 2, 4 and 8 socket configurations with an all flash disk setup for the persistence layer.

- 2 socket: see Figure 6-2 on page 101,
- 4 socket: see Figure 6-3 on page 102 and
- 8 socket: see Figure 6-4 on page 103.

It starts as small as 192 GB and can grow up to 12 TB in a single server. In all tables of this chapter, the processor type can be exchanged to Intel SP Gen 1 8180 or 8180M for improved performance.

Memory Size	192GB	384GB	576GB mixed DIMM sizes	768GB	1536GB	2304GB* mixed DIMM sizes	3072GB*
Server	SR950	SR950	SR950	SR950	SR950	SR950	SR950
MachineType	7X13	7X13	7X13	7X13	7X13	7X13	7X13
Configuration Identifier	2SKL192SSDS	2SKL384SSDS	2SKL576SSDS	2SKL768SSDS	2SKL1536SSDS	2SKL2304SSDS	2SKL3072SSDS
Skylake CPU	2x 8176	2x 8176	2x 8176	2x 8176	2x 8176	**2x 8176M**	**2x 8176M**
DIMM configuration	192GB DDR4 (12x 16GB)	384GB DDR4 (24x 16GB or 12x 32GB)	576GB DDR4 (12x 16GB **and** 12x 32GB)	768GB DDR4 (24x 32GB or 12x 64GB)	1536GB DDR4 (24x 64GB or 12x 128GB)	2304GB DDR4 (12x 64GB **and** 12x 128GB)	3072GB DDR4 (24x 128GB)
Storage Setup	2x 400GB SSD 3x 3.84TB SSD	2x 400GB SSD 3x 3.84TB SSD	2x 400GB SSD 3x 3.84TB SSD	2x 400GB SSD 3x 3.84TB SSD	2x 400GB SSD 3x 3.84TB SSD	2x 400GB SSD 4x 3.84TB SSD	2x 400GB SSD 5x 3.84TB SSD
RAID Adapters	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i
Storage Capacity	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	11.5 TB RAID5 data/log	15.3 TB RAID5 data/log
Network Connectivity	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE
Upgrade Options	192 → 384	384 → 768 2s → 4s	2s → 4s	768 → 1536 2s → 4s	1536 → 3072 2s → 4s	2s → 4s	2s → 4s

* For transactional applications only

Figure 6-2 SR950 Xeon SP Gen 1 all-flash 2 socket single node configurations

Memory Size	384GB	768GB	1152GB mixed DIMM sizes	1536GB	3072GB	4608GB* mixed DIMM sizes	6144GB*
Server	SR950	SR950	SR950	SR950	SR950	SR950	SR950
MachineType	7X13	7X13	7X13	7X13	7X13	7X13	7X13
Configuration Identifier	4SKL384SSDS	4SKL768SSDS	4SKL1152SSDS	4SKL1536SSDS	4SKL3072SSDS	4SKL4608SSDS	4SKL6144SSDS
Skylake CPU	4x 8176	4x 8176	4x 8176	4x 8176	4x 8176	**4x 8176M**	**4x 8176M**
DIMM configuration	384GB DDR4 (24x 16GB)	768GB DDR4 (48x 16GB or 24x 32GB)	1152GB DDR4 (24x 16GB **and** 24x 32GB)	1536GB DDR4 (48x 32GB or 24x 64GB)	3072GB DDR4 (48x 64GB or 24x 128GB)	4608GB DDR4 (24x 64GB **and** 24x 128GB)	6144GB DDR4 (48x 128GB)
Storage Setup	2x 400GB SSD 3x 3.84TB SSD	2x 400GB SSD 3x 3.84TB SSD	2x 400GB SSD 3x 3.84TB SSD	2x 400GB SSD 3x 3.84TB SSD	2x 400GB SSD 5x 3.84TB SSD	2x 400GB SSD 7x 3.84TB SSD	2x 400GB SSD 9x 3.84TB SSD
RAID Adapters	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i
Storage Capacity	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	15.3 TB RAID5 data/log	23.0 TB RAID5 data/log	30.7 TB RAID5 data/log
Network Connectivity	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE
Upgrade Options	384 → 768	768 → 1536 4s → 8s	4s → 8s	1536 → 3072 4s → 8s	3072 → 6144 4s → 8s	4s → 8s	4s → 8s

* For transactional applications only

Figure 6-3 SR950 Xeon SP Gen 1 all-flash 4 socket single node configurations

Memory Size	768GB	1536GB	2304GB mixed DIMM sizes	3072GB	6144GB	9216GB* mixed DIMM sizes	12288GB*
Server	SR950	SR950	SR950	SR950	SR950	SR950	SR950
MachineType	7X13	7X13	7X13	7X13	7X13	7X13	7X13
Configuration Identifier	8SKL768SSDS	8SKL1536SSDS	8SKL2304SSDS	8SKL3072SSDS	8SKL6144SSDS	8SKL9216SSDS	8SKL12288SSDS
Skylake CPU	8x 8176	8x 8176	8x 8176	8x 8176	8x 8176	**8x 8176M**	**8x 8176M**
DIMM configuration	768GB DDR4 (48x 16GB)	1536GB DDR4 (96x 16GB or 48x 32GB)	2304GB DDR4 (48x 16GB **and** 48x 32GB)	3072GB DDR4 (96x 32GB or 48x 64GB)	6144GB DDR4 (96x 64GB or 48x 128GB)	9216GB DDR4 (48x 64GB **and** 48x 128GB)	12288GB DDR4 (96x 128GB)
Storage Setup	2x 400GB SSD 3x 3.84TB SSD	2x 400GB SSD 3x 3.84TB SSD	2x 400GB SSD 3x 3.84TB SSD	2x 400GB SSD 5x 3.84TB SSD	2x 400GB SSD 9x 3.84TB SSD	2x 400GB SSD 14x 3.84TB SSD or 7x 7.68TB SSD	2x 400GB SSD 18x 3.84TB SSD or 9x 7.68TB SSD
RAID Adapters	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	2x 930-16i (or 1x ; in case 7.68TB SSDs are used)	2x 930-16i (or 1x ; in case 7.68TB SSDs are used)
Storage Capacity	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	15.3 TB RAID5 data/log	30.7 TB RAID5 data/log	46.0 TB RAID5 data/log	61.4 TB RAID5 data/log
Network Connectivity	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE
Upgrade Options	768 → 1536	1536 → 3072		3072 → 6144	6144 → 12288		

* For transactional applications only

Figure 6-4 SR950 Xeon SP Gen 1 all-flash 8 socket single node configurations

You can also grow into a clustered scale-out environment without changing any of the hardware of the single-node solution. The upgrade to scale-out is supported for memory configurations of 1.5 TB (4-socket only), 3TB, and 6 TB (8-socket only) as denoted by the multiple boxes in Figure 6-1 on page 100. Scale-out of other memory sizes is not supported because adding memory first (to scale-up) to reach one of the supported memory sizes gives better performance than adding nodes.

Upgrading the server requires downtime of the system. However, because of the capability of the underlying storage subsystem (XFS or Spectrum Scale) to add capacity online simply by adding only devices, the data on the system remains intact. It is recommended that data is backed up before the system's configuration is changed.

Support for business continuity

All of the solutions that are listed in 6.1.1, "Single-node SR950 solution for analytical and transactional use cases" on page 101 support the following features for business continuity:

- HA within a single data center
- HA across data centers (metro distance)
- HA within a single data center plus DR to a secondary site
- HA across data centers (metro distance) plus DR to a third site
- DR to a secondary site

All high availability and disaster recovery solutions can be implemented by using SAP HANA System Replication. You also can use the standby nodes in HA and DR solutions to run a second SAP HANA instance for non-productive purposes. For more information about these architectures, see 7.2, "HA and DR for single-node SAP HANA" on page 133.

6.1.2 Scale-out SR950 solution for analytical use cases

SAP customers with landscapes that do not fit into the memory provided by a single SR950 system need to implement a scale-out architecture. Single node building blocks are coupled together on a high-speed network for inter-node communication. The models which are approved by SAP for scale-out operation are listed in Figure 6-5 for 4 socket systems as well as Figure 6-6 on page 105 for 8 socket systems.

Memory Size	1536GB	3072GB
Server	SR950	SR950
MachineType	7X13	7X13
Configuration Identifier	4SKL1536SSDG	4SKL3072SSDG
Scaleout Nodes	Scaleout to 94 nodes	Scaleout to 94 nodes
Skylake CPU	4x 8176	4x 8176
DIMM configuration	1536GB DDR4 (48x 32GB or 24x 64GB)	3072GB DDR4 (48x 64GB or 24x 128GB)
Storage Setup	2x 400GB SSD 5x 3.84TB SSD	2x 400GB SSD 9x 3.84TB SSD
RAID Adapters	1x 930-16i	1x 930-16i
Storage Capacity	15.3 TB RAID5 data/log	30.7 TB RAID5 data/log
Network Connectivity	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE
Upgrade Options	1536 → 3072 4s → 8s	4s → 8s

Figure 6-5 SR950 Xeon SP Gen 1 all-flash 4socket scale-out configurations

Memory Size	3072GB	6144GB**	9216GB* mixed DIMM sizes	12288GB*
Server	SR950	SR950	SR950	SR950
MachineType	7X13	7X13	7X13	7X13
Configuration Identifier	8SKL3072SSDG	8SKL6144SSDG	8SKL9216SSDG	8SKL12288SSDG
Scaleout Nodes	Scaleout to 94 nodes	Scaleout to 94 nodes	**Scaleout to 4+1 node**	**Scaleout to 4+1 node**
Skylake CPU	8x 8176	8x 8176	**8x 8176M**	**8x 8176M**
DIMM configuration	3072GB DDR4 (96x 32GB or 48x 64GB)	6144GB DDR4 (96x 64GB or 48x 128GB)	9216GB DDR4 (48x 64GB **and** 48x 128GB)	12288GB DDR4 (96x 128GB)
Storage Setup	2x 400GB SSD 9x 3.84TB SSD	2x 400GB SSD 18x 3.84TB SSD or 9x 7.68TB SSD	2x 400GB SSD 14x 7.68TB SSD	2x 400GB SSD 18x 7.68TB SSD
RAID Adapters	1x 930-16i	2x 930-16i (or 1x ; in case 7.68TB SSDs are used)	2x 930-16i	2x 930-16i
Storage Capacity	30.7 TB RAID5 data/log	61.4 TB RAID5 data/log	92.1 TB RAID5 data/log	122.8 TB RAID5 data/log
Network Connectivity	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE	4x 10GbE 4x 1GigE
Upgrade Options	3072 → 6144			

Figure 6-6 SR950 Xeon SP Gen 1 all-flash 8 socket scale-out configurations

Lenovo has certified scale-out environments up to 94 nodes using IBM Spectrum Scale as the file system. The usage of an external SAN is also possible.

** The 6 TB and larger configurations can be used for SAP S/4HANA scale-out scenarios, described directly here after.

6.1.3 Scale-out SR950 solution for SAP S/4HANA

SAP has announced general availability of scale-out environments for transactional workload with SAP S/4HANA. It is essential to go through a thorough sizing exercise and select the largest possible node size and as few nodes as possible which calls for an 8 socket system with 9 or 12 TB of main memory as depicted in Figure 6-6.

The biggest cluster size is limited to 4(+1) nodes at this time. Please read also “System topology for SAP S/4HANA” on page 58.

Business continuity is essential for the nature of the data which is treated in SAP S/4HANA systems, which typically carry business critical or mission critical applications and data. Failure and takeover scenarios need to be discussed with the customer. Realistically, it is recommended to have an identical stand-by cluster using “synchronous mode in memory” to take over the workload in a short amount of time. When a failover happens in a 4+1 scenario the standby node has to load data from disk into main memory because it doesn’t know which node’s data it needs to take over so it cannot preload. This can extend the failover time significantly.

Please carefully read through SAP notes 2408419 and 2446942 which contain more detailed information and references the conditions when operating an SAP S/4HANA scale-out cluster.

Lenovo is a customer of SAP and runs large instances of SAP applications. The migration of the central ECC system was tested to run on SAP S/4HANA on a scale out configuration. The key performance indicator was to find out the table distribution and the functional behavior of important business transactions. The results are documented and available in a whitepaper, which can be found here:

https://en.resources.lenovo.com/whitepapers/lenovo-sap-s-4hana-scale-out

6.2 Lenovo SR650 (Gen1)

For smaller requirements, 2u configuration SR650 systems are perfectly suitable as Figure 6-8 shows. The scalability options within the 2 socket configurations regarding main memory is depicted in Figure 6-7. Some upgrades require to replace memory modules with higher capacity modules.

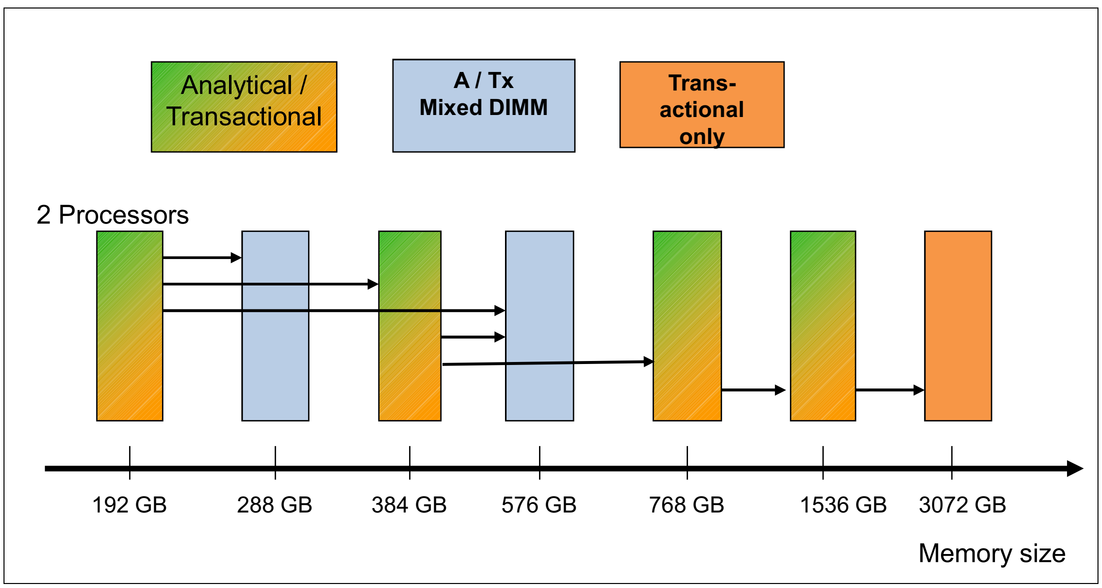


Figure 6-7 SR650 scalability options with Intel Xeon SP Gen1

As in the SR950 systems, the processor can be exchanged with Intel Xeon SP 1 (formerly codename Skylake) 8180 or 8180M for improved performance.

Memory Size	192GB	288GB mixed DIMM sizes	384GB	576GB mixed DIMM sizes	768GB	1536GB	3072GB*
Server	SR650	SR650	SR650	SR650	SR650	SR650	SR650
MachineType	7X06	7X06	7X06	7X06	7X06	7X06	7X06
Configuration Identifier	2S2U_2SKL192 SSDS	2S2U_2SKL288 SSDS	2S2U_2SKL384 SSDS	2S2U_2SKL576S SDS	2S2U_2SKL76 8SSDS	2S2U_2SKL1536SS DS	2S2U_2SKL3072SS DS
Skylake CPU	2 x 8176	2 x 8176	2 x 8176	2 x 8176	2 x 8176	2 x 8176	**2 x 8176M**
DIMM configuration	192GB DDR4 (24x 8GB or 12x 16GB)	288GB DDR4 (12x 8GB **and** 12x 16GB)	384GB DDR4 (24x 16GB or 12x 32GB)	576GB DDR4 (12x 16GB **and** 12x 32GB)	768GB DDR4 (24x 32GB or 12x 64GB)	1536GB DDR4 (24x 64GB or 12x 128GB)	3072GB DDR4 (24x 128GB)
Storage Setup	2x 240GB SATA 3x 800GB SSD	2x 240GB SATA 4x 800GB SSD	2x 240GB SATA 4x 800GB SSD	2x 240GB SATA 6x 800GB SSD or 3x 1.6TB SSD	2x 240GB SATA 6x 800GB or 4x 1.6TB SSD	2x 240GB SATA 3x 3.84TB SSD or 6x 1.6TB SSD	2x 240GB SATA 5x 3.84TB SSD
RAID Adapters	1x 930-8i	1x 930-8i	1x 930-8i	1x 930-8i	1x 930-8i	1x 930-8i	1x 930-8i
Storage Capacity	1.6TB RAID5 data/log	2.4TB RAID5 data/log	2.4 TB RAID5 data/log	4.0 TB or 3.2TB RAID5 data/log	4.0TB or 4.8TB RAID5 data/log	7.6TB or 8.0TB RAID5 data/log	15.3TB RAID5 data/log
Network Connectivity	4x 10GbE or 4x 1GigE	4x 10GbE or 4x 1GigE	4x 10GbE or 4x 1GigE	4x 10GbE or 4x 1GigE	4x 10GbE or 4x 1GigE	4x 10GbE or 4x 1GigE	4x 10GbE or 4x 1GigE
Upgrade Options	192 -> 288 192 -> 384 192 -> 576		384 -> 576 384 -> 768		768 -> 1536	1536 -> 3072	

* For transactional applications only

Figure 6-8 SR650 Xeon SP Gen 1 all-flash 2 socket configurations

6.3 Lenovo SR 950 (Gen 2) all-flash models

The newest Intel Xeon SP processor family can be used for SAP HANA application either in the classic way with an all-flash persistence layer or by leveraging he usage of persistent memory described in 6.4, “Lenovo SR 950 (Gen 2) and Optane DC Persistent Memory” on page 110.

In Figure 6-9 on page 108 you can see the scalability options for the Lenovo solution for SAP HANA.

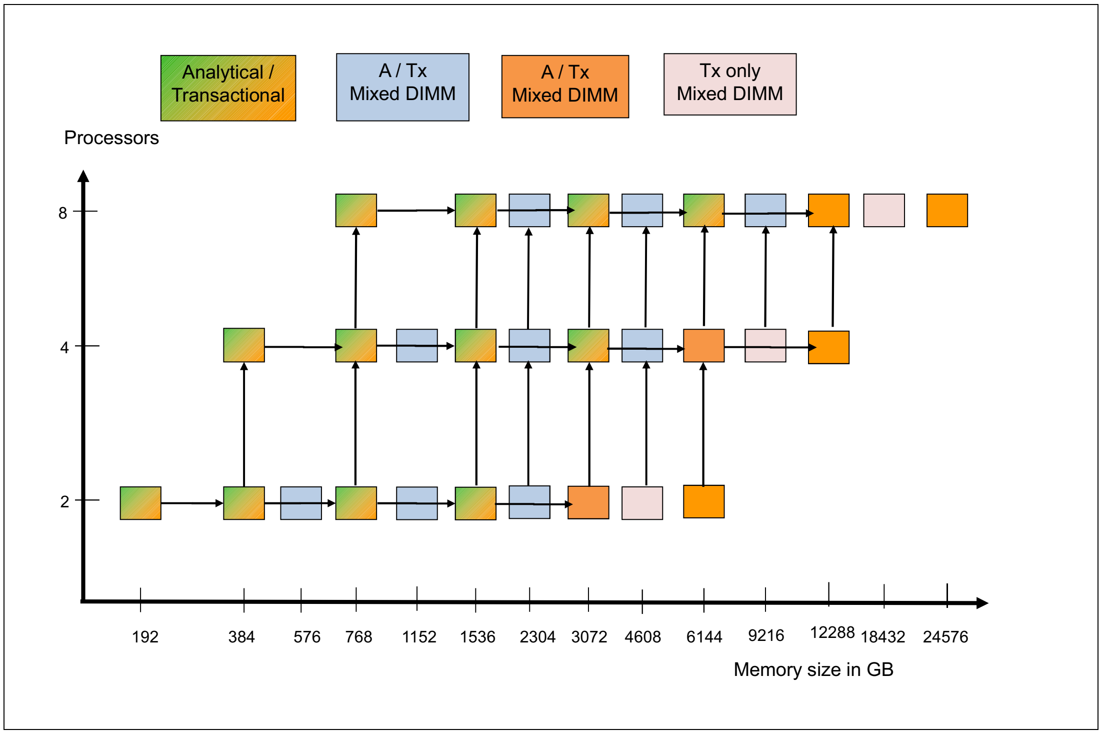


Figure 6-9 SR950 (Gen 2) based building blocks for SAP HANA

Following you can find the single node configurations with 2, 4 and 8 socket SR950 servers in Figure 6-10 on page 108, Figure 6-11 on page 109 and Figure 6-12 on page 109.

Memory Size	192GB	384GB	576GB mixed DIMM sizes	768GB	1152GB mixed DIMM sizes	1536GB	2304GB* mixed DIMM sizes	3072GB*	4608GB* mixed DIMM sizes	6144GB*
Server	SR950	SR950	SR950	SR950	SR950	SR950	SR950	SR950	SR950	SR950
MachineType	7X13	7X13	7X13	7X13	7X13	7X13	7X13	7X13	7X13	7X13
Configuration Identifier	2CSL192SSDS	2CSL384SSDS	2CSL576SSDS	2CSL768SSDS	2CSL1152SSDS	2CSL1536SSDS	2CSL2304SSDS	2CSL3072SSDS	2CSL4608SSDS	2CSL6144SSDS
Cascade Lake CPU	2x 8276	2x 8276	2x 8276	2x 8276	2x 8276	2x 8276	**2x 8276M**	**2x 8276M**	**2x 8276L**	**2x 8276L**
DIMM configuration	192GB DDR4 (12x 16GB)	384GB DDR4 (24x 16GB or 12x 32GB)	576GB DDR4 (12x 16GB **and** 12x 32GB)	768GB DDR4 (24x 32GB or 12x 64GB)	1152GB DDR4 (12x 32GB **and** 12x 64GB)	1536GB DDR4 (24x 64GB or 12x 128GB)	2304GB DDR4 (12x 64GB **and** 12x 128GB)	3072GB DDR4 (24x 128GB or 12x 256GB)	4608GB DDR4 (12x 128GB **and** 12x 256GB)	6144GB DDR4 (24x 256GB)
Storage Setup	2x 800GB SSD 3x 3.84TB SSD	2x 800GB SSD 3x 3.84TB SSD	2x 800GB SSD 3x 3.84TB SSD	2x 800GB SSD 3x 3.84TB SSD	2x 800GB SSD 3x 3.84TB SSD	2x 800GB SSD 3x 3.84TB SSD	2x 800GB SSD 4x 3.84TB SSD	2x 800GB SSD 5x 3.84TB SSD	2x 800GB SSD 7x 3.84TB SSD	2x 800GB SSD 9x 3.84TB SSD
RAID Adapters	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i
Storage Capacity	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	11.5 TB RAID5 data/log	15.3 TB RAID5 data/log	23.0 TB RAID5 data/log	30.7 TB RAID5 data/log
Network Connectivity	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE
Upgrade Options	192 → 384	384 → 768 2s → 4s	2s → 4s	768 → 1536 2s → 4s	2s → 4s	1536 → 3072 2s → 4s	2s → 4s	2s → 4s	2s → 4s	2s → 4s

Figure 6-10 SR950 Xeon SP Gen 2 all-flash 2 socket configurations

Memory Size	384GB	768GB	1152GB mixed DIMM sizes	1536GB	2304GB mixed DIMM sizes	3072GB	4608GB* mixed DIMM sizes	6144GB*	9216GB* mixed DIMM sizes	12288GB*
Server	SR950	SR950	SR950	SR950	SR950	SR950	SR950	SR950	SR950	SR950
MachineType	7X13	7X13	7X13	7X13	7X13	7X13	7X13	7X13	7X13	7X13
Configuration Identifier	4CSL384SSDS	4CSL768SSDS	4CSL1152SSDS	4CSL1536SSDS	4CSL2304SSDS	4CSL3072SSDS	4CSL4608SSDS	4CSL6144SSDS	4CSL9216SSDS	4CSL12288SSDS
Cascade Lake CPU	4x 8276	4x 8276	4x 8276	4x 8276	4x 8276	4x 8276	**4x 8276M**	**4x 8276M**	**4x 8276L**	**4x 8276L**
DIMM configuration	384GB DDR4 (24x 16GB)	768GB DDR4 (48x 16GB or 24x 32GB)	1152GB DDR4 (24x 16GB **and** 24x 32GB)	1536GB DDR4 (48x 32GB or 24x 64GB)	2304GB DDR4 (24x 32GB **and** 24x 64GB)	3072GB DDR4 (48x 64GB or 24x 128GB)	4608GB DDR4 (24x 64GB **and** 24x 128GB)	6144GB DDR4 (48x 128GB or 24x 256GB)	9216GB DDR4 (24x 128GB **and** 24x 256GB)	6144GB DDR4 (48x 256GB)
Storage Setup	2x 800GB SSD 3x 3.84TB SSD	2x 800GB SSD 3x 3.84TB SSD	2x 800GB SSD 3x 3.84TB SSD	2x 800GB SSD 3x 3.84TB SSD	2x 800GB SSD 4x 3.84TB SSD	2x 800GB SSD 5x 3.84TB SSD	2x 800GB SSD 7x 3.84TB SSD	2x 800GB SSD 9x 3.84TB SSD	2x 800GB SSD 14x 3.84TB SSD or 7x 7.68TB SSD	2x 800GB SSD 18x 3.84TB SSD or 9x 7.68TB SSD
RAID Adapters	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	2x 930-16i (or 1x ; in case 7.68TB SSDs are used)	2x 930-16i (or 1x ; in case 7.68TB SSDs are used)
Storage Capacity	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	11.5 TB RAID5 data/log	15.3 TB RAID5 data/log	23.0 TB RAID5 data/log	30.7 TB RAID5 data/log	46.0 TB RAID5 data/log	61.4 TB RAID5 data/log
Network Connectivity	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE
Upgrade Options	384 → 768	768 → 1536 4s → 8s	4s → 8s	1536 → 3072 4s → 8s	4s → 8s	3072 → 6144 4s → 8s	4s → 8s	6144 → 12288 4s → 8s	4s → 8s	4s → 8s

Figure 6-11 SR950 Xeon SP Gen 2 all-flash 4 socket configurations

Memory Size	768GB	1536GB	2304GB mixed DIMM sizes	3072GB	4608GB mixed DIMM sizes	6144GB	9216GB* mixed DIMM sizes	12288GB*	18432GB* mixed DIMM sizes	24576GB*
Server	SR950	SR950	SR950	SR950	SR950	SR950	SR950	SR950	SR950	SR950
MachineType	7X13	7X13	7X13	7X13	7X13	7X13	7X13	7X13	7X13	7X13
Configuration Identifier	8CSL768SSDS	8CSL1536SSDS	8CSL2304SSDS	8CSL3072SSDS	8CSL4608SSDS	8CSL6144SSDS	8CSL9216SSDS	8CSL12288SSDS	8CSL18432SSDS	8CSL24576SSDS
Cascade Lake CPU	8x 8276	8x 8276	8x 8276	8x 8276	8x 8276	8x 8276	**8x 8276M**	**8x 8276M**	**8x 8276L**	**8x 8276L**
DIMM configuration	768GB DDR4 (48x 16GB)	1536GB DDR4 (96x 16GB or 48x 32GB)	2304GB DDR4 (48x 16GB **and** 48x 32GB)	3072GB DDR4 (96x 32GB or 48x 64GB)	4608GB DDR4 (48x 32GB **and** 48x 64GB)	6144GB DDR4 (96x 64GB or 48x 128GB)	9216GB DDR4 (48x 64GB **and** 48x 128GB)	12288GB DDR4 (96x 128GB or 48x 256GB)	18432GB DDR4 (48x 128GB **and** 48x 256GB)	24576GB DDR4 (96x 256GB)
Storage Setup	2x 800GB SSD 3x 3.84TB SSD	2x 800GB SSD 3x 3.84TB SSD	2x 800GB SSD 3x 3.84TB SSD	2x 800GB SSD 5x 3.84TB SSD	2x 800GB SSD 7x 3.84TB SSD	2x 800GB SSD 9x 3.84TB SSD	2x 800GB SSD 14x 3.84TB SSD or 7x 7.68TB SSD	2x 800GB SSD 18x 3.84TB SSD or 9x 7.68TB SSD	2x 800GB SSD 14x 7.68TB SSD	2x 800GB SSD 16x 7.68TB SSD
RAID Adapters	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	2x 930-16i (or 1x ; in case 7.68TB SSDs are used)	2x 930-16i (or 1x ; in case 7.68TB SSDs are used)	2x 930-16i	2x 930-16i
Storage Capacity	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	7.6 TB RAID5 data/log	15.3 TB RAID5 data/log	23.0 TB RAID5 data/log	30.7 TB RAID5 data/log	46.0 TB RAID5 data/log	61.4 TB RAID5 data/log	92.2 TB RAID5 data/log	107.5 TB RAID5 data/log
Network Connectivity	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE
Upgrade Options	768 → 1536	1536 → 3072		3072 → 6144		6144 → 12288				

Figure 6-12 SR950 Xeon SP Gen 2 all-flash 8 socket configurations

In all 2, 4 or 8 socket configurations, the CPU can be replaced with Intel Xeon SP Gen 2 processors 8280 / 8280M / 8280L for improved performance.

6.4 Lenovo SR 950 (Gen 2) and Optane DC Persistent Memory

This technology enables a maximum of 24 TB in an 8 socket SR950 system. Dependent on an expert sizing, this is fully supported by SAP and Lenovo for running production systems on SAP HANA. The SAP HANA version needs to be 2.0 SPS 03 (revision 35) or higher and the persistent memory modules are used in the App Direct Mode.

The DIMM placement in the system needs to be done homogeneous and symmetrically in all DIMM slots. The supported DRAM to PMEM ratios are 1:1, 1:2 and 1:4.

SR950 CLX Memory Configuration

Controller #	0						1						CPU Type	Capacity (GB) per CPU	Capacity (GB) with #CPUs				DRAM PMEM ratio 1:x
Channel #	0		1		2		0		1		2								
DIMM slot #	1	2	3	4	5	6	7	8	9	10	11	12			2	4	6	8	
2DPC PMEM 128GB + 16GB DRAM	128	16	128	16	128	16	128	16	128	16	128	16	base	864	1728	3456	5184	6912	8
2DPC PMEM 128GB + 32GB DRAM	128	32	128	32	128	32	128	32	128	32	128	32	base	960	1920	3840	5760	7680	4
2DPC PMEM 128GB + 64GB DRAM	128	64	128	64	128	64	128	64	128	64	128	64	M	1152	2304	4608	6912	9216	2
2DPC PMEM 128GB + 128GB DRAM	128	128	128	128	128	128	128	128	128	128	128	128	M	1536	3072	6144	9216	12288	1
2DPC PMEM 256GB + 32GB DRAM	256	32	256	32	256	32	256	32	256	32	256	32	M	1728	3456	6912	10368	13824	8
2DPC PMEM 256GB + 64GB DRAM	256	64	256	64	256	64	256	64	256	64	256	64	M	1920	3840	7680	11520	15360	4
2DPC PMEM 128GB + 256GB DRAM	128	256	128	256	128	256	128	256	128	256	128	256	L	2304	4608	9216	13824	18432	0,5
2DPC PMEM 256GB + 128GB DRAM	256	128	256	128	256	128	256	128	256	128	256	128	L	2304	4608	9216	13824	18432	2
2DPC PMEM 256GB + 256GB DRAM	256	256	256	256	256	256	256	256	256	256	256	256	L	3072	6144	12288	18432	24576	1
2DPC PMEM 512GB + 64GB DRAM	512	64	512	64	512	64	512	64	512	64	512	64	L	3456	6912	13824	20736	27648	8
2DPC PMEM 512GB + 128GB DRAM	512	128	512	128	512	128	512	128	512	128	512	128	L	3840	7680	15360	23040	30720	4
2DPC PMEM 512GB + 256GB DRAM	512	256	512	256	512	256	512	256	512	256	512	256	L	4608	9216	18432	27648	36864	2

<green> … allowed memory setup

<red> … exceeds memory limit

<gray> … exceeds DRAM/PMEM ratio

Figure 6-13 Technically possible memory ratios in SR950 Xeon SP Gen 2

SAP Quicksizer example

A workload estimation (greenfield) of 3000 medium CS users and 3000 medium SD users was entered in the SAP Quicksizer and delivered the result depicted in Figure 6-14 on page 110.

All | SAPS and SCU class | Memory and disk | Data IO | Log IO

Concurrent Users + Throughput-only : Results for Software Components

SW component	CPU cat.	SAPS (total, 2-tier)	App. SAPS (ABAP)	DB SAPS	Memory Cat.	Memory (total, 2-tier, MB)	App. Mem. (ABAP)	DB Memory	Persistent DB Memory (optional)	DB Disk cat.	DB Disk (GB, total)	Data IO category	Log IO category	SCU Class for App. Server
S/4 SERVER	M	110,000	78,000	26,000	L	2,337,792	169,984	2,168,832	1,012,736	S	1,287	XL	S	A

All | SAPS and SCU class | Memory and disk | Data IO | Log IO

Throughput + Concurrent Users-only : Results for Software Components

SW component	CPU cat.	SAPS (total, 2-tier)	App. SAPS (ABAP)	DB SAPS	Memory Cat.	Memory (total, 2-tier, MB)	App. Mem. (ABAP)	DB Memory	Persistent DB Memory (optional)	DB Disk cat.	DB Disk (GB, total)	Data IO category	Log IO category	SCU Class for App. Server
S/4 SERVER	S	80,000	60,000	21,000	L	2,042,880	139,264	1,904,640	887,808	S	960	L	S	A

Figure 6-14 Quicksizer Example

Memory total (the higher of the two boxes): 2,337,792 MB = 2283 GB

Optional persistent DB memory: 1,012,736 = 989 GB

Total memory - persistent memory = 2283 - 989 = 1294 GB

DRAM = 1294 GB

the next possible configuration respecting the guidelines results in 1536 GB

Hence looking at the table below in Figure 6-15 on page 113 a 2 socket SR950 (8276 M) with 3 TB memory in total (1:1) is suitable for this requirement. This is a conservative approach to ensure that intermediate results of queries fit in DRAM not leading to a halt of the system. Remember, the boundary between DRAM and persistent memory is fixed.

Brownfield example

For SAP S/4HANA the sizing report was executed on a source system at a customer. Following you can find parts of the sizing report.

```
SIZING RESULTS IN GB

Based on the selected table(s), the anticipated maximum
requirement are for S/4HANA:

- Memory requirement for the initial installation 942.7

- Net data size on disk for the initial installation 591.1

- Estimated memory requirement after data clean-up 688.3

- Estimated net data size on disk after data clean-up 524.1

.....

Other possible additional memory requirement:

- during the transition to S/4HANA (See FAQ) 35.3

Check the FAQ document attached to SAP Note 1872170 for
explanations on how to interpret the sizing terms and
calculations.

Sizing report: /SDF/HDB_SIZING

Version of the report: 77

Date of analysis: 07.06.2019

Selected sample size: M

Data aging retention in days for technical objects: 015

Number of work processes used: 06

Duration of the analysis in seconds: 3,931

HANA version sized: 2.0

SID XXX

NW release: 731 SP 22

Kernel version 721_EXT_REL

Operating system on AS NT 6.1 7601 S x86

Type of analyzed database: ORACLE

Database version: 11.2.0.3.0
```

```
Unicode system: Yes
Data used size on disk of the analysed database in GB: 1,302
Number of tables successfully analyzed: 95,443
Number of tables with error: 0
```

and further in the report:

```
.....
MEMORY SIZING FOR PERSISTENT MEMORY (NVRAM) HANA SIZE IN GB
Column store data 404.2
+ Changes in FI tables and columns 22.5
+ Changes in MM/SD tables and columns 8.6
= Anticipated initial requirement for the Persistent Memory 435.3
Row store data 3.6
+ Cached Hybrid LOB (20%) 14.8
+ Work space 438.9
+ Fixed size for code, stack and other services 50.0
= Anticipated initial requirement for the DRAM 507.4
Check SAP Note 2618154 for more information on Persistent Memory.
.....
```

Here also the ratio of DRAM to PMEM proposed is 1:1 with 435 GB for persistent memory and 507 GB for DRAM. Looking at the table in Figure 6-13 on page 110 above the smallest memory configuration for a 1:1 ratio would be a 3TB system.
For a total memory requirement of 943 GB, this seems not a very reasonable choice, if we only look at capacity. If the business continuity requirements are mission critical, then it might still be a viable option. Additionally, the sizing report is a snapshot on the current situation. Considering that the customer might have data growth over time, a larger system to start with is logical.

The non-functional requirements and the boundary conditions need to be discussed with the customer.

Predefined systems

Following you can find the 2, 4 and 8 socket single node configurations, which have been predefined. See Figure 6-15 on page 113 for 2 socket systems, Figure 6-16 on page 113 for 4 socket systems and Figure 6-17 on page 114 for 8 socket configurations.

Memory Size	1920GB	2304GB	3072GB	3840GB	4608GB	6144GB	7680GB	9216GB
Server◇	SR950	SR950	SR950	SR950	SR950	SR950	SR950	SR950
MachineType	7X13	7X13	7X13	7X13	7X13	7X13	7X13	7X13
Configuration Identifier	2CSL1920PMSSDS	2CSL2304PMSSDS	2CSL3072PMSSDS	2CSL3840PMSSDS	2CSL4608PMSSDS	2CSL6144PMSSDS	2CSL7680PMSSDS	2CSL9216PMSSDS
Cascade Lake CPU*	2x 8276	**2x 8276M**	**2x 8276M**	**2x 8276M**	**2x 8276L**	**2x 8276L**	**2x 8276L**	**2x 8276L**
DIMM configuration *[DRAM : PMM ratio]*	1536GB PMM (12x 128GB) and 382GB DRAM (12x 32GB) *[1:4]*	1536GB PMM (12x 128GB) and 768GB DRAM (12x 64GB) *[1:2]*	1536GB PMM (12x 128GB) and 1536GB DRAM (12x 128GB) *[1:1]*	3072GB PMM (12x 256GB) and 768GB DRAM (12x 64GB) *[1:4]*	3072GB PMM (12x 256GB) and 1536GB DRAM (12x 128GB) *[1:2]*	3072GB PMM (12x 256GB) and 3072GB DRAM (12x 256GB) *[1:1]*	6144GB PMM (12x 512GB) and 1536GB DRAM (12x 128GB) *[1:4]*	6144GB PMM (12x 512GB) and 3072GB DRAM (12x 256GB) *[1:2]*
Storage Setup	2x 800GB SSD 4x 3.84TB SSD	2x 800GB SSD 4x 3.84TB SSD	2x 800GB SSD 5x 3.84TB SSD	2x 800GB SSD 5x 3.84TB SSD	2x 800GB SSD 7x 3.84TB SSD	2x 800GB SSD 9x 3.84TB SSD	2x 800GB SSD 12x 3.84TB SSD or 6x 7.68TB SSD	2x 800GB SSD 14x 3.84TB SSD or 7x 7.68TB SSD
RAID Adapters	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	2x 930-16i (or 1x ; in case 7.68TB SSDs are used)
Storage Capacity	11.5 TB RAID5 data/log	11.5 TB RAID5 data/log	15.3 TB RAID5 data/log	15.3 TB RAID5 data/log	23.4 TB RAID5 data/log	30.7 TB RAID5 data/log	38.4 TB RAID5 data/log	46.0 TB RAID5 data/log
Network Connectivity	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE
Upgrade Options	2s → 4s	2s → 4s	2s → 4s	2s → 4s	2s → 4s	2s → 4s	2s → 4s	2s → 4s

Figure 6-15 SR950 Xeon SP Gen 2, 2 socket single node with DCPMM configurations

Memory Size	3840GB	4608GB	6144GB	7680GB	9216GB	12288GB	15360GB	18432GB
Server◇	SR950	SR950	SR950	SR950	SR950	SR950	SR950	SR950
MachineType	7X13	7X13	7X13	7X13	7X13	7X13	7X13	7X13
Configuration Identifier	4CSL3840PMSSDS	4CSL4608PMSSDS	4CSL6144PMSSDS	4CSL7680PMSSDS	4CSL9216PMSSDS	4CSL12288PMSSDS	4CSL15360PMSSDS	4CSL18432PMSSDS
Cascade Lake CPU*	4x 8276	**4x 8276M**	**4x 8276M**	**4x 8276M**	**4x 8276L**	**4x 8276L**	**4x 8276L**	**4x 8276L**
DIMM configuration *[DRAM : PMM ratio]*	3072GB PMM (24x 128GB) and 768GB DRAM (24x 32GB) *[1:4]*	3072GB PMM (24x 128GB) and 1536GB DRAM (24x 64GB) *[1:2]*	3072GB PMM (24x 128GB) and 3072GB DRAM (24x 128GB) *[1:1]*	6144GB PMM (24x 256GB) and 1536GB DRAM (24x 64GB) *[1:4]*	6144GB PMM (24x 256GB) and 3072GB DRAM (24x 128GB) *[1:2]*	6144GB PMM (24x 256GB) and 6144GB DRAM (24x 256GB) *[1:1]*	12288GB PMM (24x 512GB) and 3072GB DRAM (24x 128GB) *[1:4]*	12288GB PMM (24x 512GB) and 6144GB DRAM (24x 256GB) *[1:2]*
Storage Setup	2x 800GB SSD 5x 3.84TB SSD	2x 800GB SSD 7x 3.84TB SSD	2x 800GB SSD 9x 3.84TB SSD	2x 800GB SSD 12x 3.84TB SSD or 6x 7.68TB SSD	2x 800GB SSD 14x 3.84TB SSD or 7x 7.68TB SSD	2x 800GB SSD 9x 7.68TB SSD	2x 800GB SSD 10x 7.68TB SSD	2x 800GB SSD 14x 7.68TB SSD
RAID Adapters	1x 930-16i	1x 930-16i	1x 930-16i	1x 930-16i	2x 930-16i (or 1x ; in case 7.68TB SSDs are used)	1x 930-16i	1x 930-16i	2x 930-16i
Storage Capacity	15.3 TB RAID5 data/log	23.4 TB RAID5 data/log	30.7 TB RAID5 data/log	38.4 TB RAID5 data/log	46.0 TB RAID5 data/log	61.4 TB RAID5 data/log	69.1 TB RAID5 data/log	92,2 TB RAID5 data/log
Network Connectivity	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE
Upgrade Options		4s → 8s	4s → 8s	4s → 8s	4s → 8s	4s → 8s	4s → 8s	4s → 8s

Figure 6-16 SR950 Xeon SP Gen 2, 4 socket single node with DCPMM configurations

Memory Size	7680GB	9216GB	12288GB	15360GB	18432GB	24576GB
Server	SR950	SR950	SR950	SR950	SR950	SR950
MachineType	7X13	7X13	7X13	7X13	7X13	7X13
Configuration Identifier	8CSL7680PMSSDS	8CSL9216PMSSDS	8CSL12288PMSSDS	8CSL15360PMSSDS	8CSL18432PMSSDS	8CSL24576PMSSDS
Cascade Lake CPU*	8x 8276	**8x 8276M**	**8x 8276M**	**8x 8276M**	**8x 8276L**	**8x 8276L**
DIMM configuration *[DRAM : PMM ratio]*	6144GB PMM (48x 128GB) and 1536GB DRAM (48x 32GB) *[1:4]*	6144GB PMM (48x 128GB) and 3072GB DRAM (48x 64GB) *[1:2]*	6144GB PMM (48x 128GB) and 6144GB DRAM (48x 128GB) *[1:1]*	12288GB PMM (48x 256GB) and 3072GB DRAM (48x 64GB) *[1:4]*	12288GB PMM (48x 256GB) and 6144GB DRAM (48x 128GB) *[1:2]*	12288GB PMM (48x 256GB) and 12288GB DRAM (48x 256GB) *[1:1]*
Storage Setup	2x 800GB SSD 12x 3.84TB SSD or 6x 7.68TB SSD	2x 800GB SSD 14x 3.84TB SSD or 7x 7.68TB SSD	2x 800GB SSD 9x 7.68TB SSD	2x 800GB SSD 10x 7.68TB SSD	2x 800GB SSD 14x 7.68TB SSD	2x 800GB SSD 16x 7.68TB SSD
RAID Adapters	1x 930-16i	2x 930-16i (or 1x ; in case 7.68TB SSDs are used)	1x 930-16i	1x 930-16i	2x 930-16i	2x 930-16i
Storage Capacity	38.4 TB RAID5 data/log	46.0 TB RAID5 data/log	61.4 TB RAID5 data/log	69.1 TB RAID5 data/log	92,2 TB RAID5 data/log	107.5 TB RAID5 data/log
Network Connectivity	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE	4x 10 or 25GbE 4x 1GigE
Upgrade Options						

Figure 6-17 SR950 Xeon SP Gen 2, 8 socket single node with DCPMM configurations

In all 2, 4 or 8 socket configurations using persistent memory, the CPU can be replaced with Intel Xeon SP Gen 2 processors 8280 / 8280M / 8280L for improved performance or any Gold or Platinum processors dependent on the application workload.

6.5 SAP HANA models using SAN storage

The previous chapters covered example configurations using internal, all-flash storage. Of course, other architecture options exist. Dependent on the customer requirements for non-functional requirements like consolidation applications based on SAP HANA and non-HANA, a central SAN array is requested. This applies also to demands regarding backup/restore, cloning and business continuity, which should be same across the complete datacenter.

For this the storage components listed above in the tables can be replaced with SAN storage. Please find the technology description in 5.2.3, “Lenovo ThinkSystem DM and DE series storage arrays” on page 78.

Lenovo DM storage delivers the included features on ONTAP, like efficient SAP backup/restore even for the largest databases in seconds, efficient SAP system copies and clones, cloud connected storage enabling easy true hybrid cloud deployments, seamless integration into higher level SAP management tools such as SAP Landscape Management, and so on.

Lenovo DE storage provides a block-level (no NFS) access to the data on the persistence layer. This is especially interesting, if the focus is on the requirements for consolidation of HANA and non-HANA based workloads, as described above.

Lenovo provides guidelines regarding the architecture and sizing for both options. Please contact SAPsolutions@lenovo.com.

6.6 SAP HANA models using NVMe storage

Additional to the above described options, Lenovo offers SAP HANA systems, where the local storage layer for the persistence of data and log consists of NVMe devices.

This can be used in all of the above described solutions. No matter which system (SR650, SR850 or SR950) or processor generation, no matter whether persistent memory is used or not, the configurations are fully supported under the TDI umbrella.

Measurements of the throughput between memory and storage with hwcct (Hardware Configuration Check Tool) by SAP perfectly achieve or exceed the KPIs for SAP HANA systems. The database reads are performed in more than 5 GB/s.

Linux software RAID provides the redundancy of the contents and is stable and proven for many years. This comes for free with any Linux distribution and is subsequently also supported by RedHat and SUSE.

When starting with a Linux software RAID it is most likely either a RAID-5 or a RAID-6. Remember that you can create a RAID-5 with two devices already. This allows a seamless expansion with more devices when growing the SAP HANA system (just add more NVMe to the array).

In terms of what NVMe devices are to use, from a performance perspective all current (as of 1Q19) Lenovo NVMe devices can be chosen. This means a single device is enough to fulfill SAP HANA TDI KPIs already.

It is recommended to start with two identical devices for fulfill redundancy and stay upgradable for larger capacity needs.

For the operating system it is possible to chose either also NVMe devices or go with two SAS/SATA SSDs of choice and connect them to an entry-level 530-8i controller for straight-forward RAID-1 redundancy. Of course, higher RAID controllers work as well.

As examples you can find configurations using NVMe devices in Figure 6-18 on page 115 for 2 socket and Figure 6-19 on page 116 for 4 socket systems. This will work in SR650 or SR850 accordingly.

Memory Size	192GB	384GB	768GB	1536GB	3072GB
Server	SR950	SR950	SR950	SR950	SR950
Machine Type	7X13	7X13	7X13	7X13	7X13
Cascade Lake CPU	2x 8276	2x 8276	2x 8276	2x 8276	2x 8276M
Storage Setup NVMe	2x 1.92TB	2x 1.92TB	3x 1.92TB	5x 1.92TB or 3x 3.84TB	5x 3.84TB
Storage Capacity	1.92 TB RAID-5 data/log	1.92 TB RAID-5 data/log	3.84 TB RAID-5 data/log	7.68 TB RAID-5 data/log	15.3 TB RAID-5 data/log

Figure 6-18 SR950 Gen 2, 2 socket with NVMe storage

Memory Size	384GB	768GB	1536GB	3072GB	6144GB
Server	SR950	SR950	SR950	SR950	SR950
Machine Type	7X13	7X13	7X13	7X13	7X13
Skylake CPU	4x 8276	4x 8276	4x 8276	4x 8276	4x 8276M
Storage Setup NVMe	2x 1.92TB	3x 1.92TB	5x 1.92TB or 3x 3.84TB	5x 3.84TB	9x 3.84TB or 5x 7.68TB
Storage Capacity	1.92 TB RAID-5 data/log	3.84 TB RAID-5 data/log	7.68 TB RAID-5 data/log	15.3 TB RAID-5 data/log	30.7 TB RAID-5 data/log

Figure 6-19 SR950 Gen 2, 4 socket with NVMe storage

6.7 Consolidation of SAP HANA instances

Deployment of SAP HANA databases on dedicated hardware can lead to many SAP HANA appliances in the data center, such as production, disaster recovery, quality assurance (QA), test, and sandbox systems, and possibly for multiple application scenarios, regions, or lines of business. Therefore, the consolidation of SAP HANA instances seems desirable. The following sections discuss different approaches for how to run more than one SAP HANA database instance per physical machine.

6.7.1 Multitenancy built into SAP HANA (MDC)

With the release of SAP HANA SPS09 in 2014 SAP introduced support for multiple tenant databases within one single SAP HANA system with a new feature called multitenant database containers (MDC).

Note: Since SAP HANA 2.0 this is the default installation option.

SAP HANA supports multiple isolated databases in a single SAP HANA system. These are referred to as tenant databases. An SAP HANA system is capable of containing more than one tenant database.

A system always has exactly one system database, used for central system administration, and any number of tenant databases (including zero). An SAP HANA system is identified by a single system ID (SID). Databases are identified by a SID and a database name. From the administration perspective, there is a distinction between tasks performed at system level and those performed at database level. Database clients, such as the SAP HANA cockpit, connect to specific databases.

All the databases share the same installation of database system software, the same computing resources, and the same system administration. However, each database is self-contained and fully isolated with its own:

- Set of database users
- Database catalog
- Repository

- Persistence
- Backups
- Traces and logs

Although database objects such as schemas, tables, views, procedures, and so on are local to the database, cross-database SELECT queries are possible. This supports cross-application reporting, for example.

SAP HANA is installed natively on the operating system of a server (or multiple servers in case of a scale-out installation) and runs without a hypervisor or other virtualization software. All system resources are assigned to this SAP HANA system. Inside SAP HANA then are multiple isolated containers each comprising a tenant database.

MDC uses only one single system ID (SID) and each tenant database is identified with a dedicated TCP port to which applications can connect.

It is supported to run SAP Business Suite or SAP S/4HANA and SAP Business Warehouse in one SAP HANA system using MDC technology. Cross-database queries are supported, which enable cross-application reporting from within the same physical server.

If MDC is enabled on a scale-out SAP HANA installation then it is supported to create tenant databases spanning more than one physical node.

Backup and restore procedures can be implemented on a per-tenant level. SAP HANA System Replication, however, can only be defined on a per-system level (this *may* change in the future with newer SAP HANA revisions). Keep in mind that all tenants run on the SAP HANA revision as they use the same binaries. If different SAP HANA revisions are required, for example as a test environment for SAP HANA updates, then MDC is not the right feature to use.

Latest information about supported SAP HANA features with MDC enabled can be found in the following SAP notes:

- 2096000 SAP HANA Multitenant Database Containers – Additional Information
- 2101244 FAQ: SAP HANA Multitenant Database Containers
- 2423367 Multitenant database containers will become the standard and only operation mode

6.7.2 Consolidation options of SAP HANA

Another way of consolidating (pre SAP HANA 1.0, SPS09 release) is to install more than one instance of SAP HANA database onto one SAP HANA system. However, there are major drawbacks when multiple SAP HANA instances are consolidated on one appliance; therefore, support for production instances needs to be checked.

For non-production systems, the support status depends on the following scenarios:

- Multiple Components on One System (MCOS)

  Having multiple SAP HANA instances on one system, which is also referred to as *Multiple Components on One System* (MCOS), is not recommended because this poses conflicts between different SAP HANA databases on a single server; for example, common data and log volumes, possible performance degradations, and interference of the systems with each other. SAP and Lenovo support this scenario under certain conditions (see SAP Note 1681092). However, if issues arise, SAP or Lenovo might ask you to stop all but one of the instances as part of the troubleshooting process to see whether the issue persists.

  Consider using SAP HANA multi-tenancy database containers (MDC) instead as explained in 6.7.1, “Multitenancy built into SAP HANA (MDC)” on page 116.

- Multiple Components on One Cluster (MCOC)

  Running multiple SAP HANA instances on one scale-out cluster (which is also referred to as *Multiple Components on One Cluster* (MCOC), is supported if each node of the cluster runs only one SAP HANA instance. A development and a QA instance can run on one cluster, but with dedicated nodes for each of the two SAP HANA instances, for example, each of the nodes runs either the development instance, or the QA instance, but not both. Only the Spectrum Scale file system is shared across the cluster.

- Multiple Components in One Database (MCOD)

  Having one SAP HANA instance containing multiple components, schemas, or application scenarios, which is also referred to as *Multiple Components in One Database* (MCOD), is supported. To have all data within a single database, which is also maintained as a single database, can lead to limitations in operations, database maintenance, backup and recovery, and so on. For example, bringing down the SAP HANA database affects all of the scenarios. It is impossible to bring it down for only one scenario. For more information about the implications, see SAP Note 1661202.

Consider the following factors when SAP HANA instances are consolidated on one SAP HANA appliance:

- An instance filling up the log volume causes all other instances on the system to stop working correctly. This situation can be addressed by monitoring the system closely.
- Installation of another instance might fail when there are other instances that are installed and active on the system. The installation procedures check the available space on the storage and refuse to install when there is less free space than expected. This situation also might occur when you are trying to reinstall an installed instance.
- Installing a new SAP HANA revision for one instance might affect other instances that are installed on the system. For example, new library versions with the new installation might break the installed instances.
- The performance of the SAP HANA system becomes unpredictable because the individual instances on the system are sharing resources, such as memory and CPU.

When asking for support for such a system, you might be asked to remove the extra instances and to recreate the issue on a single instance system.

6.7.3 SAP HANA on VMware vSphere

Another way of consolidating multiple SAP HANA instances on one system[1] is virtualization. SAP HANA is supported on VMware vSphere under several conditions.

Whether the SAP HANA instance is installed on bare metal server by Lenovo or into a VM is transparent. This means, the installable image delivered by Lenovo is the same.

A good starting point for finding out the facts are the following three VMware owned websites including the support matrixes for both SAP HANA on Intel Xeon SP Gen 1 and 2:

https://www.vmware.com/solutions/business-critical-apps/sap-virtualization.html

https://blogs.vmware.com/apps/2018/09/sap-hana-on-intel-skylake-and-broadwell-support-for-vmware-vsphere-6-5-and-6-7.html

https://blogs.vmware.com/apps/2019/05/sap-grants-support-for-sap-hana-2-0-on-vmware-vsphere-6-7u2-running-on-intel-cascade-lake-based-servers.html

Additionally, there is an SAP Wiki page available covering the subject with detailed support tables as well as the relevant SAP notes:

https://wiki.scn.sap.com/wiki/display/VIRTUALIZATION/SAP+HANA+on+VMware+vSphere

At the time of writing (July 2019), Intel Optane DC Persistent Memory (DCPMM) is not supported for SAP HANA in a virtualized environment.

Sizing

The sizing is the same for virtualized and non-virtualized SAP HANA deployments. Although there is a small performance effect because of the virtualization, the database size and the required memory size are not affected. Beware that memory overcommitment is not supported.

Support

As with any other deployment type of SAP HANA, clients are asked to open an SAP support ticket by using the integrated support model that is described in 8.7, "Lenovo and SAP Integrated support process" on page 151. Any non-SAP related issue is routed to VMware first, and it eventually is forwarded to the hardware partner. In certain but rare situations, SAP or its partners might need to reproduce the workload on bare metal.

6.7.4 SAP HANA on hyperconverged infrastructure solutions

Nutanix

Lenovo provides the ThinkAgile™ HX Series which is combined with hyperconvergence software from Nutanix and Lenovo enterprise platforms that feature Intel Xeon Processor Scalable family.

The ThinkAgile HX7820 for SAP HANA is a 4U rack-mount appliance that supports four processors, up to 3 TB of 2666 MHz TruDDR4 memory, up to 76.8 TB of all flash storage capacity with NVMe PCIe cache acceleration, and network connectivity with 1/10 GbE RJ-45, 10 GbE SFP+, and 10/25 GbE SFP28 ports. The maximum size of an SAP HANA database in a virtual machine can be 2.3 TB.

[1] One SAP HANA system, as referred to in this section, can consist of one single server or multiple servers in a clustered configuration.

It is listed on the SAP HANA hardware directory:

https://www.sap.com/dmc/exp/2014-09-02-hana-hardware/enEN/hci.html#categories=certified%23Lenovo&recordid=2059

More information on this solution can be found at:

https://lenovopress.com/lp0992-lenovo-thinkagile-hx7820-appliance-for-sap-hana

VMware VSAN

Lenovo provides the ThinkAgile VX Series which combine hyperconverged infrastructure software from VMware with Lenovo enterprise platforms that feature the first and second generation of the Intel Xeon Processor Scalable family.

The ThinkAgile VX 2U node for SAP HANA is a 2U rack-mount system that supports up to two processors, up to 3 TB of TruDDR4 memory with 2666 MHz memory speeds, 24x 2.5-inch hot-swap drive bays with an extensive choice of NVMe PCIe SSDs, SAS SSDs, and flexible network connectivity options with 1/10 GbE RJ-45, 10 GbE SFP+, and 10/25 GbE SFP28 ports. The maximum size of an SAP HANA database in a virtual machine is theoretically 3 TB, but a buffer for the hypervisor of at least 10% is recommended. Hence a VM can have a maximum size of roughly 2.7 TB.

It is listed on the SAP HANA hardware directory:

https://www.sap.com/dmc/exp/2014-09-02-hana-hardware/enEN/hci.html#categories=certified%23Lenovo&recordid=2081

More information on this solution can be found at:

https://lenovopress.com/lp1074-reference-architecture-for-sap-applications-using-lenovo-thinkagile-vx-series

6.8 Security and encryption of an SAP HANA system

SAP environments contain a company's most critical data and must be protected from unauthorized access physically and virtually. Although restricting physical access is obvious and relatively easy to implement in today's data centers, virtual access is more difficult to control.

Some industries must also obey to certain standards that define rules and processes for handling certain data sets. Three of the more well-known standards are the Payment Card Industry Data Security Standard (PCI DSS), the Sarbanes-Oxley Act (SOX), and the Health Insurance Portability and Accountability Act (HIPAA). These standards regulate banks, credit card companies, insurance companies, pharmaceutical companies, and many more.

They describe in different levels of detail how IT processing must be implemented and what safety measures must be in place to cater for maximum data privacy and to prevent data leaks. Some standards also enforce data encryption to ensure safety and to prevent unauthorized data access. The portfolio of Lenovo System x Solutions for SAP HANA provides several levels of security, including data encryption, communication channels encryption, and data access security.

6.8.1 Encrypting SAP HANA data

SAP HANA systems work with data on the following technologies:

- Volatile storage (main memory that uses DRAM technology)
- Intel Optane DC persistent memory (DCPMM)
- Non-volatile storage (disks or flash storage)

These technologies feature different characteristics and require different levels of precaution to ensure the safety of the data that is on them.

Volatile storage

Volatile storage requires a constant supply of power to continuously refresh the content of the DRAM cells. When power is lost, the data in main memory also is lost. If you turn off your System x server, all data in main memory is gone. This fact is a good enough safety measure for situations in which data must be prevented from being physically stolen. No encryption must be applied in this case.

Preventing main memory data from being virtually stolen (that is, someone can login to the SAP HANA system and work on it) is almost impossible with today's technology. Data is encrypted in main memory and is available as encrypted data without ever decrypting it.

Although it is technically feasible to encrypt data in main memory (however, this encryption imposes a large performance penalty), the data must be decrypted at some point during processing. Every security measure is only as good as its weakest link. Therefore, encrypting main memory is not sufficient if data still must be decrypted before being processed by the CPU. The performance penalty of encrypting main memory defeats the purpose of main-memory computing.

DCPMM

Intel Optane DC persistent memory has 256-AES hardware encryption so the data is more secure. In App Direct Mode (which is the only one used for SAP HANA) data is encrypted using a key on the module. Intel Optane DC persistent memory is locked at power loss and a passphrase is needed to unlock and access the data. The encryption key is stored in a security metadata region on the module and is only accessible by the Intel Optane DC persistent memory controller. If repurposing or discarding the module, a secure cryptographic erase and DIMM over-write is utilized to keep data from being accessed.

The Passphrase can be set though uEFI setup, OneCLI. The passphrase is stored securely and if enabled uEFI can auto unlock the DCPMM.

Non-volatile storage

Non-volatile storage can and should be protected whenever possible because threats are real and safety breaches can easily lead to data leaks. SAP HANA regularly writes out data from main memory to non-volatile storage, such as disk drives or flash drives. This write is required to reload data whenever the system is restarted. This restart can be the result of any event, from a scheduled maintenance window to unplanned outages or failures.

The most obvious threat are stolen disks or improperly decommissioned disks. When a disk is replaced, you must ensure that data that is on the disk is properly destroyed. This requirement is difficult to fulfill for failed drives that cannot be accessed.

By using the Lenovo System Solution for SAP HANA, such data can be encrypted on a file system level. It uses the encryption capabilities that are built into Spectrum Scale 4.1 code. Any data that is written out by an SAP HANA database to its persistency layer (that is, to flash or disk storage) is encrypted as it goes through the file system. No unencrypted data is on physical media.

Spectrum Scale encrypts data blocks when it receives them from the application. As a result, Spectrum Scale data that travels over the local network also is encrypted before it is sent out by a node. All Lenovo System Solutions for SAP HANA that are running with the Spectrum Scale data replication feature for high-availability or disaster recover (Spectrum Scale parameter r=2/3, m=2/3) has their data traffic encrypted in addition to the data on their local storage devices. Spectrum Scale metadata is *not* encrypted; it comprises the names and permissions of the files, directories, and other elements, such as Spectrum Scale Extended Attributes (EAs) that are not used in a System solution for SAP HANA.

Figure 6-20 shows what parts of a scale-out System solution are secured. All the orange elements in the Figure 6-20 on page 122 are encrypted. These elements include data that is written to local disks and data that is sent out over the network to other nodes to be stores on their disks (single-node installations do not send out Spectrum Scale traffic to the network).

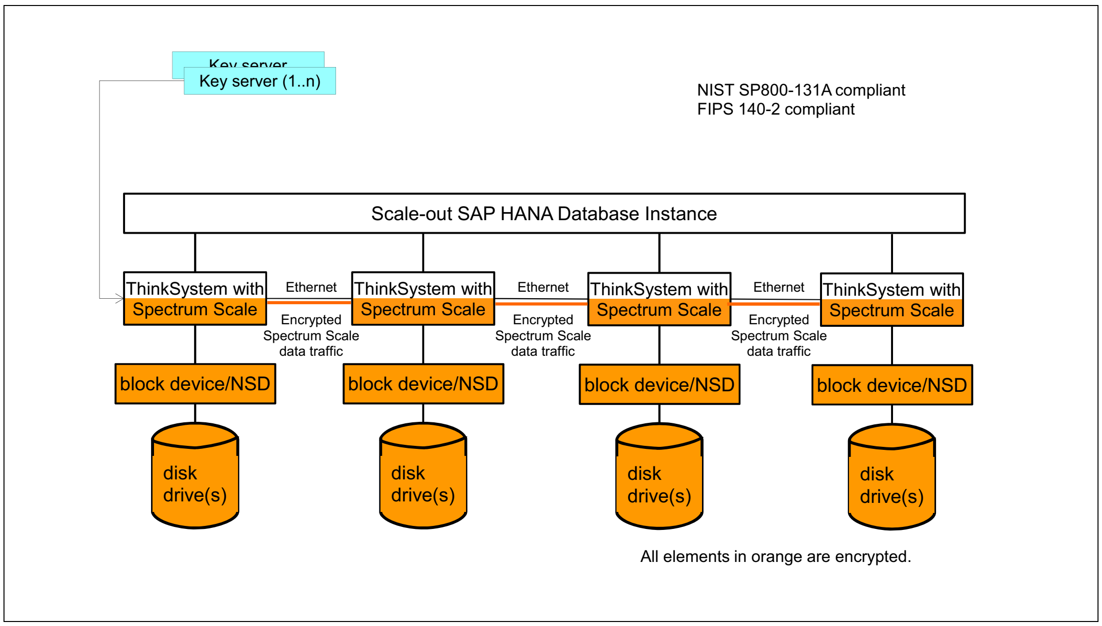


Figure 6-20 Encrypting SAP HANA data by using Spectrum Scale

Spectrum Scale Encryption includes the following features:

- Encryption is built in Spectrum Scale without dependency on third-party products.
- Spectrum Scale encryption feature is compliant with NIST SP800-131A and FIPS 140-2.
- Encryption uses AES algorithm and key lengths of 128, 192, or 256 bits.
- Support for multiple encryption keys in a single file system (perfectly suited for multitenancy or multipurpose SAP HANA environments).
- Fine-grained Spectrum Scale policies enable encrypting only subsets of your file system.

Running an encrypted SAP HANA environment safeguards you from the following threats:

- Prevention of data leaks from stolen or lost disks.
- Prevention of data leaks from improperly discarded disks.
- Prevention of eavesdropping on the network, which is especially important in HA and DR scenarios when data flows over a network that is owned by a third-party ISP.
- Secure delete is a cryptographic operation (as opposed to digital shredding by overwriting), which gives the ability to destroy arbitrarily large subsets of a file system in a predetermined amount of time.

The key server that is holding the encryption keys is external to the environment. Spectrum Scale uses the Key Management Interoperability Protocol (KMIP), which is an open standard by OASIS. At the time of this writing, the only verified and supported key server by Spectrum Scale is the IBM Security Key Lifecycle Management (ISKLM) server (formerly Tivoli Key Lifecycle Manager), which supports KMIP.

ISKLM is available on the following hardware platforms and operating systems:

- x86-64 bit; Windows and Linux (RHEL and SLES)
- Power; AIX
- System z; z/OS and Linux (RHEL and SLES)
- SPARC; Solaris 10

ISKLM is validated to Federal Information Processing Standard (FIPS) 140-2 Level 1 and supports FIPS 140-2 Level 3-validated hardware for enhanced security requirements. To use the Spectrum Scale encryption feature, clients must also purchase the Advanced Edition license of Spectrum Scale (by default, the Lenovo SAP HANA installation includes the Spectrum Scale Standard Edition license).

Customers who have encryption in place for the storage or tape environment often have an ISKLM or Tivoli Key Lifecycle Manager installation that can be used to store the Spectrum Scale encryption keys.

Implementing encryption on the active production system is one step in the overall security policy. Do not forget to also include your backup and archival systems in the overall data security strategy. You can operate an end-to-end encrypted environment with ISKLM and its capability to also manage encrypted backup and tape archives. By using ISKLM, you can transfer your encrypted SAP HANA data between the different systems and storage devices without the need to decrypt it along the way.

For more information about Spectrum Scale Encryption, see this website:

http://www-01.ibm.com/support/knowledgecenter/SSFKCN_4.1.0/com.ibm.cluster.gpfs.v4r1.gpfs200.doc/bl1adv_encryption.htm

For more information about how to use and implement Spectrum Scale Encryption for SAP HANA solutions, contact your account representative, business partner, or the Lenovo SAP Center of Competency at: SAPsolutions@lenovo.com.

6.8.2 Securing SAP HANA communication channels

Encrypting SAP HANA data is only one piece to the overall security puzzle. You also must ensure that data that is in transit is secure and cannot be eavesdropped. This requirement applies to the following communication channels:

- SAP HANA external communication
- SAP HANA internal communication

For more information about securing external SAP HANA communication channels, see 6.8.3, “Securing SAP HANA access” on page 124.

Securing internal communication channels is applicable to all SAP HANA deployments, except for single-node installations without HA or DR. All installations with more than one node contributing to an SAP HANA database instance must think about securing the internal 10 Gbit Ethernet communication network.

SAP supports the Secure Sockets Layer (SSL) protocol for communication between multiple hosts in a distributed SAP HANA environment. SSL protects the following communication channels:

- Cluster-internal communication in a distributed, scale-out SAP HANA deployment
- System Replication traffic between two SAP HANA systems

To implement SSL, you must deploy certificates on every participating host. All certificates must be signed by a central certificate authority (CA). The CA can be an external, trusted company, a CA internal to the customer, or a self-signed environment that is for SAP HANA only.

SSL provides authentication of the communication peer and encrypts data that is going across the network.

6.8.3 Securing SAP HANA access

SAP groups accounts that access SAP HANA data into the following logical categories:

- Database user
- Technical database user

This categorization is purely logical and does not manifest in different parameters or rights. The purpose is to determine if SAP HANA accounts correspond to a real person or an account that is used for communication between different SAP systems.

The following methods of authentication are supported by SAP HANA:

- User name with password
- Kerberos
- Security Assertion Markup Language (SAML)
- Logon and Assertion tickets
- X.509 client certificates

User authentication can be audited.

For more information about security in an SAP HANA system, see the SAP HANA Security Guide, which is available at this website:

https://help.sap.com/viewer/b3ee5778bc2e4a089d3299b82ec762a7/2.0.04/en-US

7

Business continuity and resiliency for SAP HANA

This chapter describes individual SAP HANA high availability (HA) and disaster recovery (DR) deployment options. It explains basic terminology and deployment options for single-node systems and scale-out systems. It also describes the backup options of SAP HANA.

This chapter includes the following topics:

7.1 Overview of business continuity options

Because of the relevance of SAP software environments in today's world, it is critical that these systems do not experience unexpected downtime. Hardware manufacturers, such as Intel or Lenovo, invest much into the reliability and availability features of their products; however, IT systems are still exposed to many different sources of errors.

For that reason, it is crucial to consider business continuity and reliability aspects of IT environments when SAP HANA systems are planned. Business continuity in this context refers to designing IT landscapes with failure in mind. Failure spans from single component errors, such as a hard disk drive (HDD) or a network cable, up to the outage of the whole data center because of an earthquake or a fire. Different levels of contingency planning must be done to cope with these sources of error.

Developing a business continuity plan highly depends on the type of business a company is doing, and it differs (among other factors) by country, regulatory requirements, and employee size.

This section introduces the three main elements of business continuity planning:

- Implementing HA
- Planning for DR
- Taking backups regularly

These elements have different objectives for how long it takes to get a system online again, for the state in which the system is after it is online, and for the end-to-end consistency level of business data when an IT environment comes online again. These values feature the following definitions:

- *Recovery Time Objective* (RTO) defines the maximum tolerated time to get a system online again.
- *Recovery Point Objective* (RPO) defines the maximum tolerated time span to which data must be restored. It also defines the amount of time for which data is tolerated to be lost. An RPO of zero means that the system must be designed to not lose data in any of the considered events.

  The most common approach to achieve an RPO of zero is to implement HA within the primary data center plus an optional synchronous data replication to an offsite location (usually a second data center).
- *Recovery Consistency Objective* (RCO) defines the level of consistency of business processes and data that is spread out over multitier environments.

It is important to understand the difference between HA and DR. HA covers a hardware failure (for example, one node becomes unavailable because of a faulty processor, memory DIMM, storage, or network failure). DR covers the event when multiple nodes in a scale-out configuration fail, or a whole data center goes down because of a fire, flood, or other disaster, and a secondary site must take over the SAP HANA system.

The ability to recover from a disaster, or to "tolerate" a disaster without major effect, is sometimes also referred to as *disaster tolerance* (DT).

All of the single-node and scale-out solutions described in Chapter 6, “SAP HANA IT landscapes with Lenovo solutions” on page 99 can be enhanced with HA and DR capabilities. All scenarios require at least one other copy of the data to be available in the system so that the SAP HANA database can survive the outage of a server, including the data that is on it.

HA in scale-out clusters is implemented by introducing standby nodes. During normal operation, these nodes do not actively participate in processing data, but they do receive data that is replicated from the worker nodes. If a worker node fails, the standby node takes over and continues data processing after the data was loaded into memory. For more information about node takeover, see “Example of a node takeover” on page 93.

> **Note:** When referring to DR, the words *primary site*, *primary data center*, *active site,* and *production site* mean the same thing. Similarly, *secondary site*, *back up site*, and *DR site* are also used interchangeably.
>
> The primary site hosts your production SAP HANA instance during normal operation.

When an SAP HANA side-car scenario is run (for example, SAP CO-PA Accelerator, sales planning, or smart metering), the data still is available in the source SAP Business Suite system. Planning or analytical tasks run slower without the SAP HANA system being available, but no data is lost. More important is the situation in which SAP HANA is the primary database, such as when SAP Business Suite with SAP HANA or Business Warehouse is used with SAP HANA as the database. In those cases, the production data is available solely within the SAP HANA database, and according to the business service level agreements, prevention of a failure is necessary.

HA and DR solutions for SAP HANA can be at the following levels:

- On the infrastructure level

  By replicating data that is written to disk by the SAP HANA persistency layer synchronously or asynchronously, standby nodes can recover lost data from failed nodes. Data replication can happen within a data center (for HA), across data centers (for DR), or both (for any combination of HA and DR). This feature is known as storage replication, which can be realized with Spectrum Scale or a feature of external storage.

  Backups can be used as a business continuity option as well, if the RTO and RPO requirements are soft enough. At this level backups that are replicated or otherwise shipped from the primary site to the secondary site. More detail is described in 7.4.7, “Backup and restore as a DR strategy” on page 145.

- On the application level

  All actions that are performed on an active SAP HANA instance to a passive (or active read enabled) instance that is running in receive-only mode are replicated. Essentially, the passive instance runs the same instructions as the active instance, except for accepting user requests. This feature is known as *SAP HANA System Replication* (HSR).

The available features to implement business continuity are listed in Table 7-1. Also listed are the scenarios to which the features are applicable.

Table 7-1 Business continuity options for SAP HANA

Level	Technology	RTO	RPO	Suitable for HA	Suitable for DR
Infrastructure level (storage)	Storage replication using external storage	Minutes	Zero	Yes	Yes
	Spectrum Scale based synchronous replication	HA: Zero DR: Minutes	Zero	Yes	Yes
	Spectrum Scale based asynchronous replication	Minutes to hours	Above zero (Seconds to hours)	No	Yes
	Backup - Restore	Usually hours	Hours to days	No	Yes
Application level (system)	HSR (synchronous)	Minutes	Zero	Yes[a]	Yes
	HSR (asynchronous)	Minutes	Seconds	No	Yes

a. With SUSE HA Extension only. For more information, see the following Note.

> **Note:** HSR does not support automatic failover to a standby system. As the name implies, it replicates data to a standby system only. Manual intervention is required for the failover to occur. As an alternative, the tasks can be automated using a cluster manager that is, for example, part of SUSE HA Extension.
>
> SAP uses the term *Host Auto-Failover* to describe the capability of an automatic failover to a standby system.

The Spectrum Scale based storage replication and HSR are described next.

The rest of this chapter then describes all supported scenarios for HA and DR of single-node and scale-out SAP HANA environments that are based on Lenovo ThinkSystem solutions. For more information about conventional backup methods, see 7.4, “Backup and restore” on page 139.

7.1.1 Spectrum Scale based storage replication

Lenovo uses Spectrum Scale predominantly in scale-out clusters for SAP HANA. For single node implementations, we recommend to use SAP HANA System Replication. The file system includes built-in replication features with which you can have multiple copies of one file that is stored on disk that is spread out on different servers in a multi-server environment. Those copies are referred to as *replicas* and are all identical to each other. For more information about replicas, see , “Spectrum Scale extensions for shared-nothing architectures” on page 87.

Lenovo uses Spectrum Scale in its SAP HANA solutions in a way that allows any node in a cluster to fail without losing data. For more information about what happens when a node fails and how a standby node takes over the workload without taking down the service, see 5.2.7, “Scaling out SAP HANA using Spectrum Scale” on page 89.

Expanding into a DR setup involves another replica of the data to be stored at a remote location.

The major difference between a single site HA (as described in 5.2.7, "Scaling out SAP HANA using Spectrum Scale" on page 89) and a multi-site DR solution is the placement of the replicas within Spectrum Scale. In a single-site HA configuration, there are two replicas[1] of each data block in one cluster. In contrast, a multi-site DR solution holds a third replica in the remote or secondary site. This configuration ensures that when the primary site fails, a complete copy of the data is available at the second site and operations can be resumed at this site.

A two-site solution implements the concept of a synchronous or asynchronous data replication on a file system level between both sites by using the replication capabilities that are built into Spectrum Scale.

For more information about the wording and concepts that are used by Spectrum Scale to implement HA, see the paper that is available at this website:

https://www.ibm.com/support/knowledgecenter/en/STXKQY/ibmspectrumscale_welcome.html

7.1.2 SAP HANA System Replication

HSR is a feature that was introduced with SAP HANA 1 SPS05 and was improved in subsequent revisions. In an environment that uses HSR, the primary and the secondary system must be configured identically in terms of SAP HANA worker nodes. The number of standby nodes can differ (since SAP HANA 1 SPS06). Every SAP HANA process that is running on the primary system's worker nodes must have a corresponding process on a secondary worker node to which it replicates its activity.

In SAP HANA 1.0 the only difference between the primary and secondary system is the fact that one cannot connect to the secondary HANA installation and run queries on that database. They can also be called active and passive systems. With SAP HANA 2.0 the feature of active-active read enabled was implemented. Here the secondary system can be utilized for analytical purposes (read only). To ensure the right performance in case of a failover, the secondary system should have at least the same capacity as the primary.

Upon start of the secondary HANA system, each process establishes a connection to its primary counterpart and requests the data that is in main memory, which is called a *snapshot*. After the snapshot is transferred, the primary system continuously sends the log information to the secondary system that is running in recovery mode. At the time of this writing, HSR does not support replaying the logs immediately as they are received; therefore, the secondary site system acknowledges and persists the logs only. To avoid having to replay hours or days of transaction logs upon a failure, HSR asynchronously transmits a new incremental data snapshot periodically.

Among other criteria, the following criteria must be met to enable the HSR feature:

- The SAP HANA software revision of the target environment must be the same or higher than the source environment.
- Both systems must use the same SID and instance numbers.
- On both systems, "instance number + 1" must be free because it is used for replication purposes.

[1] In Spectrum Scale terminology, each data copy is referred to as a replica. This term also applies to the primary data, which is called the first replica. This term indicates that all data copies are equal.

Replication modes

HSR can be set to one of the following modes:

- Synchronous mode (in-memory)

  This mode makes the primary site system acknowledge the change after it is committed in main memory on the secondary site system, but not yet persisted on disk.

- Synchronous with Full Sync mode (on disk)

  This mode makes the primary site system wait until the change is committed and persisted on the secondary site system. Log write returns success if the log is written on the secondary site as well. If the link between the sites goes down, database transaction suspends until the link is restored. Although no data is lost, the system is not usable until the link is restored.

- Asynchronous mode

  This mode makes the primary site system commit transaction when the replicated log is sent to the DR site. The system does not wait for an acknowledgment from the remote site.

  Asynchronous replication allows for greater distances because a high latency between the primary and secondary system does not prevent a production workload from running at maximum performance as with synchronous replication.

In all synchronous modes, the effect is defined by the transmission time from the primary to its corresponding secondary system process. When HSR is run in synchronous mode, you must add the time that it takes to persist the change on a disk in addition to the transmission delay.

Live-replication stops if the connection between the two data centers is lost. Then, after a (configurable) time expires on the primary site system, it resumes work without replication.

When the connection is restored, the secondary site system requests a delta snapshot of what changes were done since the connection was lost. Live replication can then continue after this delta is received on the secondary site system.

Operating Modes

There are two operating modes that HSR can be set to:

- Delta data shipping

  This mode instructs the secondary system to receive the data from the primary system only but not replay it into the standby system.

- Continuous log replay (available since SAP HANA 1 SPS11)

  This mode instructs the secondary system to immediately replay all received transactions from the primary system. This reduces the takeover time.

- Active / active (read enabled) (available since SAP HANA 2 SPS00)

  This is similar to the continuous log replay, but additionally the secondary system can be used for analytical purposes because users can access the information and hence utilize the system as depicted in Figure 7-1 on page 131.

  More details are available in the SAP HANA platform documentation in the administration guide and SAP Note 2732012 and 2737255.

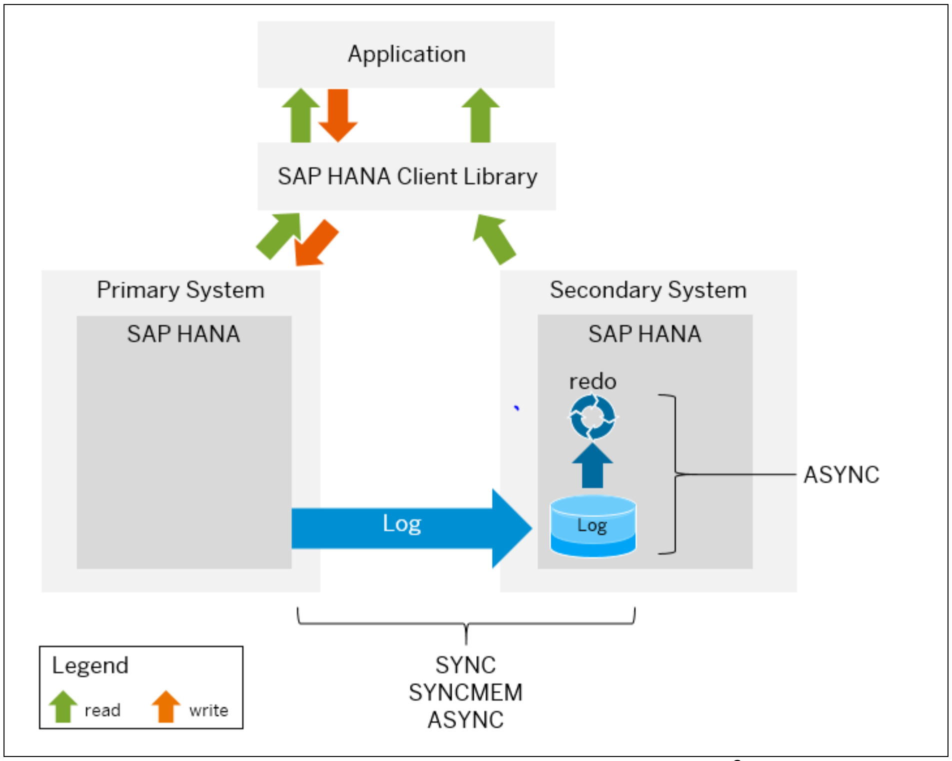


Figure 7-1 Operation mode "Active - active - read enabled" for HSR[2]

System failover

If there is a failover to the secondary system (*system takeover*), manual intervention is required to change the secondary site system from recovery mode to active mode. SAP HANA automatically loads all row-based tables into memory and rebuilds the row store indexes. In the next step, all logs since the last received snapshot are replayed. After this step finishes, the system can accept incoming database connections. A restart of the SAP HANA database instance is not required.

An optional feature (called *SYNCMEM*) enables the primary system to share information about which columns are loaded into main memory. The secondary system can use this information to preload those tables in main memory. Preloading reduces the duration of a system takeover operation.

Hosting a non-productive instance at the secondary system

If the secondary system is intended to host a non-productive instance, preloading must be disabled. In such a scenario, SAP HANA operates with a minimal main memory footprint on the secondary system so that the remaining memory can be used for a non-productive SAP HANA installation.

If a system takeover is triggered, both instances (the one receiving the replication data and the non-productive instance) must be stopped. The secondary system must be reconfigured to use all available main memory, and then a takeover operation is run. Because you must

[2] source: https://help.sap.com/viewer/6b94445c94ae495c83a19646e7c3fd56/2.0.04/en-US/fe5fc53706a34048bf4a3a93a5d7c866.html Copyright SAP SE

restart the SAP HANA processes, the time for a system takeover and a subsequent system performance ramp-up is longer when compared to when no non-productive instance is hosted and tables preload is enabled.

A non-productive instance cannot share storage with production data. For this reason, Lenovo adds extra storage to extend the locally available storage capacity to hold the data and log files of the non-productive system.

Multitier System Replication

Multitier System Replication was introduced with SAP HANA 1 SPS07, with which you can cascade replication over several databases. Multiple data centers, locally and geographically remote, can for a cluster to provide business continuity, dependent on the business requirements for RPO and RTO. An example is depicted in Figure 7-2.

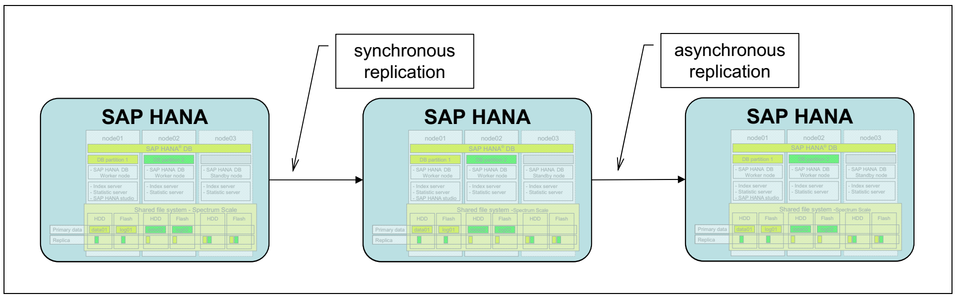


Figure 7-2 SAP HANA Multitier System Replication

More information is available here:

https://help.sap.com/viewer/6b94445c94ae495c83a19646e7c3fd56/2.0.04/en-US/ca6f4c62c45b4c85a109c7faf62881fc.html

Multitenant Database Containers

When multiple databases are consolidated into one single SAP HANA system using the MDC feature, then System Replication has to be implemented at the system level. That means no replication on tenant level is supported.

The primary and secondary systems must be identical at the time System Replication is initially set up. This requirement includes the tenant databases.

For more information about HSR, see the following publications:

- *Introduction to High Availability for SAP HANA*:

  https://help.sap.com/viewer/6b94445c94ae495c83a19646e7c3fd56/2.0.04/en-US/6d252db7cdd044d19ad85b46e6c294a4.html

- *How to Perform System Replication for SAP HANA*:

  https://help.sap.com/viewer/4e9b18c116aa42fc84c7dbfd02111aba/2.0.04/en-US

- *SAP HANA Platform documentation (the version of SAP HANA can be selected on the upper right from a drop down menu)*

  http://help.sap.com/hana_platform

7.1.3 Special considerations for DR and long-distance HA setups

The distance between the data centers that are hosting the SAP HANA servers must be within a certain range to keep network latency to a minimum. This range allows synchronous replication to occur with limited affect on the overall application performance (which is also referred to as *Metro Mirror distance*). It does not matter for what exact purpose the data is replicated to the second data center (for DR or for long-distance HA). In both cases, data must be transferred between the two locations, which affects the process.

Application latency is the key indicator for how well a long-distance HA or DR solution performs. The geographical distance between the data centers can be short. However, the fiber cable between them might follow another route. The Internet service provider (ISP) usually routes through one of its hubs, which leads to a longer physical distance for the signal to travel, and therefore a higher latency. Another factor is the network equipment between the two demarcation points on each site. More routers and protocol conversions along the line introduce a higher latency.

> **Attention:** Regarding latency, make sure to specify the layer at which you are measuring it. Network engineers refer to network latency, but SAP prefers to use application latency.
>
> *Network latency* refers to the low-level latency that network packets experience when traveling over the network from site A to site B. Network latency does not necessarily include the time that it takes for a network packet to be processed on a server.
>
> *Application latency* refers to the delay that an SAP HANA database transaction experiences when it occurs in a DR environment. This value is sometimes also known as *end-to-end latency*. It is the sum of all delays as they occur while the database request is in flight and includes, in addition to network latency, packet extraction in the Linux TCP/IP stack, Spectrum Scale code execution, or processing the SAP HANA I/O code stack.

Data can be replicated from one data center to the other data center synchronously or asynchronously.

Synchronous data replication refers to any write request that is issued by the application is committed to the application only after the request is successfully written on both sides. To maintain the application performance within reasonable limits, the network latency (and therefore the distance) between the sites must be limited to Metro Mirror distances. The maximum achievable distance depends on the performance requirements of the SAP HANA system. In general, an online analytical processing (OLAP) workload can work with higher latencies than an online transaction processing (OLTP) workload. The network latency mainly is dictated by the connection between the two SAP HANA clusters. This inter-site link typically is provided by a third-party ISP.

SAP provides a How-to Guide for the required network for SAP HANA System Replication:

https://www.sap.com/documents/2014/06/babb2b55-5a7c-0010-82c7-eda71af511fa.html

7.2 HA and DR for single-node SAP HANA

Customers that are running their SAP HANA instance on a single node can still implement redundancy to protect against a failure of this node. Best practice is to use SAP HANA System Replication for that.

Table 7-2 lists the different options that can be implemented with Lenovo workload-optimized solutions for SAP HANA.

Table 7-2 Overview of HA and DR options for single-node SAP HANA solutions with HSR

Characteristic	**Single-node SAP HANA installation with**		
	HA	**Stretched HA**	**DR**
Required data centers	1	2 or 3 (metro distance)	2 (metro distance or higher)
RTO	Seconds	Seconds	Minutes
RPO	Zero	Zero	Zero or higher
Replication method	sync	sync	sync or async
Automatic failover	Yes	Yes	Yes, with a cluster manager like SUSE HA extension
Can host non-production	Yes	Yes	Yes
Number of SAP HANA server nodes	2	2	2
Tolerated node failures	1	1	1

All three options are configured on the application level. Two SAP HANA instances are installed on two different physical systems. The tools are either SAP HANA Studio, SAP HANA Cockpit or hdbnsutil.

The following sections describe each of the solutions.

A DR setup that uses HSR is shown in Figure 7-3. *repaint picture!!!*

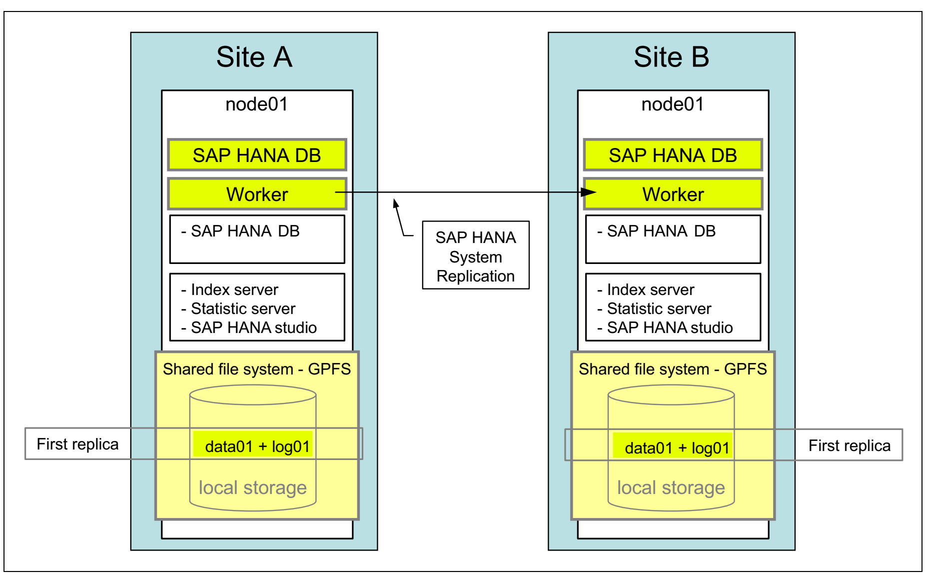


Figure 7-3 Detailed view of single-node DR solution by using SAP HANA System Replication

A failover to site B always requires manual intervention and does not occur automatically.

HSR can be configured to work in synchronous or asynchronous mode. Different physical distances can be realized in each mode. For more information, see “Replication modes” on page 130.

Different architectures are possible for the network that is connecting the two SAP HANA building blocks. Figure 7-4 shows a setup with redundant switches for the SAP HANA replication network.

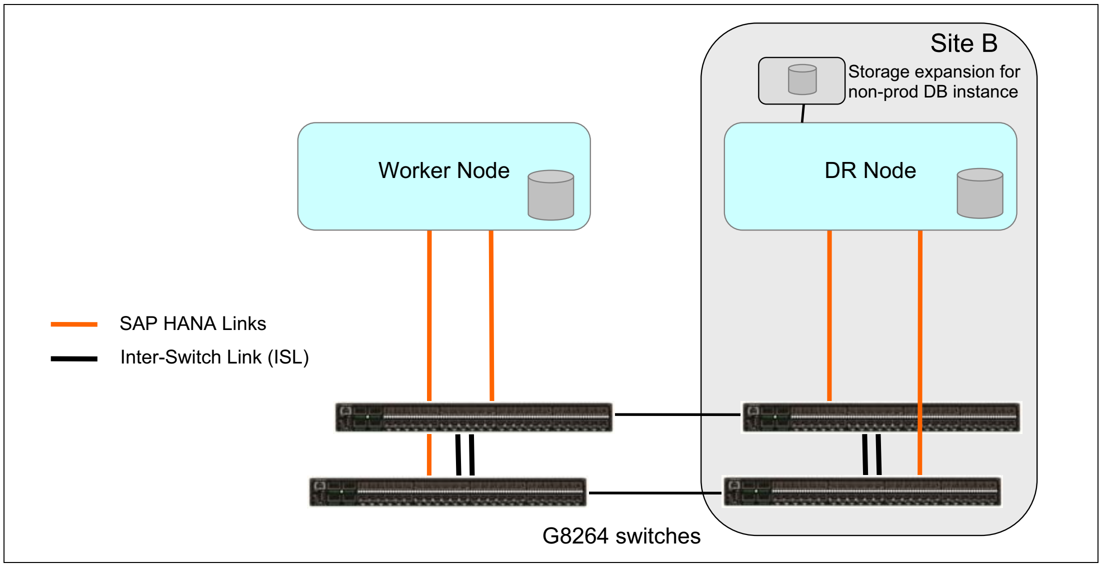


Figure 7-4 Network setup of single node with SAP HANA System Replication (dual inter-site links)

Because SAP HANA processes are not in an active standby mode on the DR node, a non-production instance can be hosted by using extra dedicated storage. Lenovo uses the D1224 for SAP HANA unit to provide this extra space. The expansion units add up to 24 2.5-inch HDDs that are directly connected to the server through an SAS interface. A second file system is created over those drives for the extra SAP HANA instance. This second file system is visible to this node only. Data is not replicated to a second node.

Designs without the network switches are also supported, as shown in Figure 7-5. It is possible to have these inter-site links going across VPNs (or other forms of overlay networks) if the latency and throughput requirements are met.

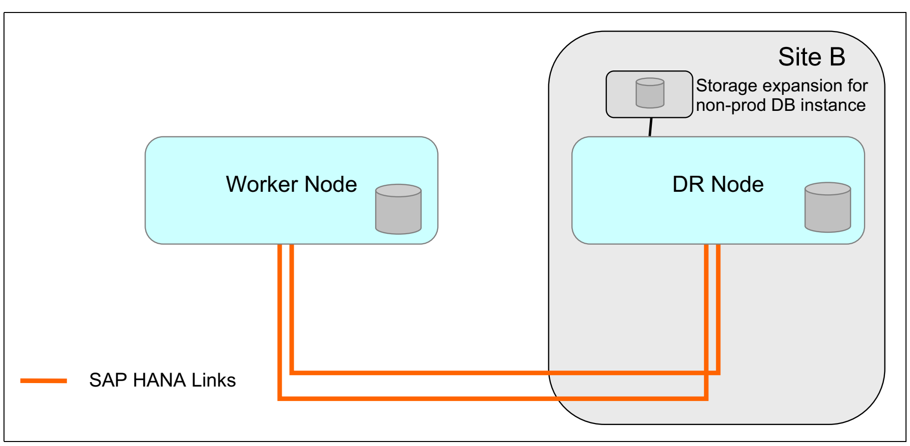


Figure 7-5 Network setup of single node with SAP HANA System Replication (no switches)

Intermediate architectures with redundant switches in the primary data center on site A and only one link to site B are possible (similar to Figure 7-4). Implementations with only one link for HSR can be implemented when no redundancy is required. Details should be discussed with the customer.

7.3 HA and DR for scale-out SAP HANA

Scale-out SAP HANA installations can implement two levels of redundancy to keep their database instance from going offline. The first step is to add a server node to the scale-out cluster that acts as a hot-standby node. When a node fails, the data is being loaded into memory on the stand-by node and the application continues to run. Dependent on the required SLA, this can be a reasonable option. But the data is not preloaded into memory on this stand-by node.

The second step is to set up another scale-out cluster in a distinct data center that takes over operation if there is a disaster at the primary site. This DR capability should be implemented by using HSR. It replicates all of the required data to the DR site nodes.

Table 7-3 lists the available options for protecting scale-out SAP HANA environments when Lenovo workload-optimized solutions for SAP HANA are used.

Table 7-3 HA and DR options for scale-out SAP HANA installations

Characteristic	Scale-out SAP HANA installation with	
	HA (by using Spectrum Scale)	**HA and DR (by using Spectrum Scale and HSR)**
Required data centers	1	2
Geographical distance	Not applicable	Metro distance or higher
RTO	Seconds	Seconds for HA, minutes for DR
RPO	Zero	Zero or higher
Replication method	Spectrum Scale (synchronous)	Spectrum Scale (synchronous) plus HSR (synchronous or asynchronous)
Automatic HA failover	Yes	Yes
Automatic DR failover	Not applicable	Yes, requires a cluster manager like SUSE HA Extension
Can host non-production	No	Yes, at DR site

Although technically it is feasible to implement DR with no HA exclusively in the primary site cluster, it is not a preferred practice to do so and it can cause issues. The first fault that occurs in the primary site triggers a DR event that leads to a failover of the production instance to the DR site. Therefore, always implement HA capabilities alongside DR to avoid these failovers. With HA, the first fault is handled within the primary site and no failover to the DR data center is triggered.

7.3.1 HA by using Spectrum Scale storage replication

SAP HANA uses the concept of standby nodes, which take over operation if a node that is designated as an active worker node fails. Standby nodes can replace any cluster node's role. For this reason, the standby node must have the same memory configuration as the other cluster nodes.

For a takeover to occur, the standby node requires access to the file system of the failed node. This requirement is fulfilled by Spectrum Scale, which is a shared file system. By using the FPO extension of Spectrum Scale, you can keep the shared file system feature in a local storage only environment. For more information about FPO, see , "Spectrum Scale extensions for shared-nothing architectures" on page 87.

To achieve HA in a scale-out SAP HANA installation, Lenovo uses the data replication feature that is built into the file system. Spectrum Scale replication ensures that two valid physical copies of the data are always in the file system. The concept of replicas is transparent to the application, which means that SAP HANA is not affected if a server that is holding one replica goes offline. Access requests to a file that lost one replica are automatically served from the second replica.

Any file system operation is run on multiple nodes (the number depends on different parameters, such as the size of the data to be written). If both data replicas are persisted successfully, Spectrum Scale signals an I/O operation as complete. This measure ensures an RPO of zero.

For more information about HA in a scale-out SAP HANA installation, see "Scale-out solution with high-availability capabilities" on page 91. This section also describes what happens when one cluster member goes offline and its data is no longer accessible.

More than one standby node in a cluster is a supported scenario. Even with multiple standby nodes, there still is a small time window in which only one data replica exists. This duration depends on the speed of the network, load on the system, and amount of data in the database. Multiple standby nodes are not a replacement for a backup strategy.

7.3.2 HA by using Spectrum Scale replication plus DR by using SAP HANA Replication

Frequently, customers choose HSR to implement DR capabilities. In such a scenario, HA within the primary data center still uses Spectrum Scale storage replication, but data replication to the DR site is handled by SAP HANA on application level with the HSR feature.

From a Spectrum Scale perspective, such a setup requires two distinct clusters, one at each site. HSR replicates data from the primary cluster to the DR site cluster. Figure 7-6 shows a four-node scenario with three worker nodes and one standby node (for HA) at each site. HSR replicates data from each node to its corresponding node at the DR site.

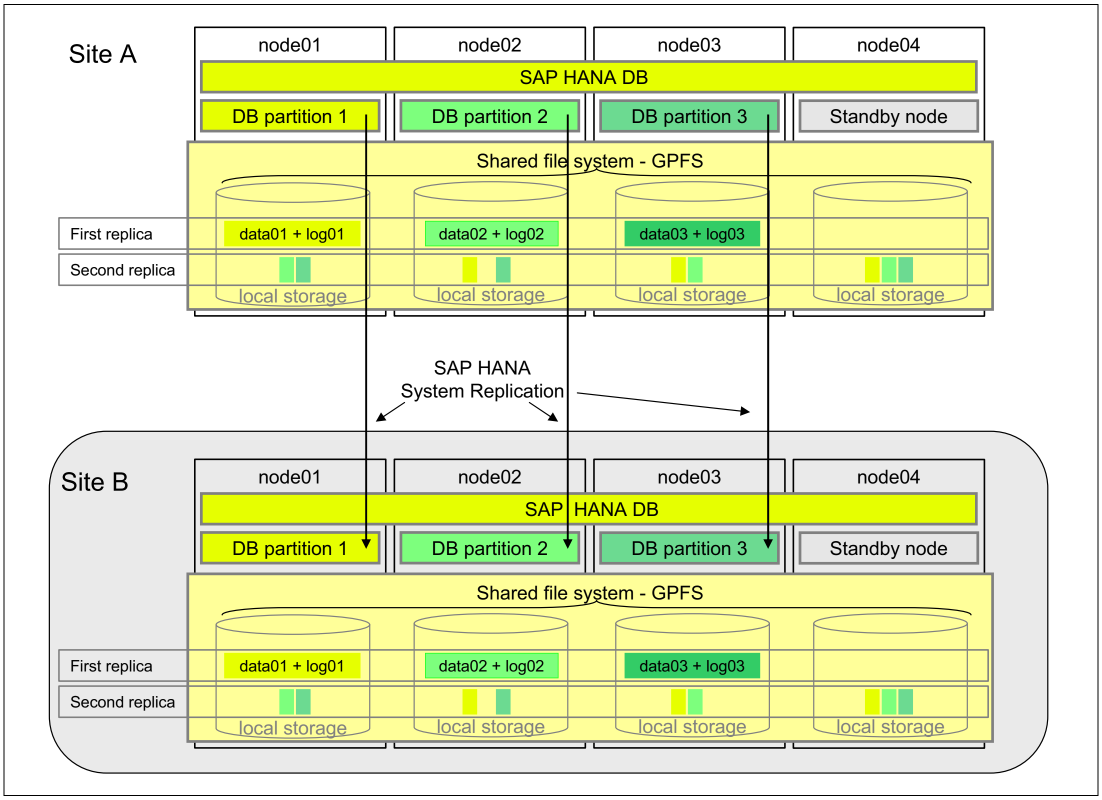


Figure 7-6 DR of a scale-out configuration by using SAP HANA System Replication

This replication can be configured to work in synchronous or asynchronous mode. Synchronous mode ensures an RPO of zero, but limits the maximum distance between the two sites. Asynchronous mode allows for bigger distances, but with the risk of losing data that is still unacknowledged when the disaster happens (RPO > 0).

For more information details about HSR and the two replication modes, see 7.1.2, “SAP HANA System Replication” on page 129.

Single-node failures are handled within the primary data center. The standby node takes over operation of the failed node. This process occurs automatically without administrative intervention.

Multi-node failures, such as a full data center outage or other disasters, require a failover of SAP HANA to the DR site. This failover is a manual activity. SAP HANA administrators must promote DR site nodes from running in recovery mode to active workers (or standby, respectively).

With the release of SAP HANA 1.0 SPS 08, it is possible to use different networks for HSR traffic. You can use the front-end client network, back-end SAP HANA network, or a new dedicated replication network that spans both sites. By having the choice between multiple networks, you can better adapt to different customer situations.

During normal operation, the nodes at the DR site B are running in recovery mode. By using this mode, you can host a non-productive SAP HANA instance (such as development or test) on these idling nodes. If there is a disaster, this non-productive SAP HANA instance must be stopped before the production instance can be started at site B.

This extra SAP HANA instance needs its own storage space for persistency and logs. The D1224 for SAP HANA unit is used to provide this extra space. The expansion units add up to 24 2.5-inch HDDs that are connected directly to the server through a SAS interface. A second file system is created over those drives for the added SAP HANA instance. Depending on the size of this extra database instance, one or more DR site nodes must be used. Every node in scope requires an external storage expansion. This second file system is created with a replication factor of two to implement HA in the non-production instance.

7.4 Backup and restore

Because SAP HANA plays a critical role in the overall SAP IT landscape, it is important to back up the data in the SAP HANA database and restore it. This section described the basic information about backing up and recovery for the operating system, for SAP HANA and the available products that are certified for use with SAP HANA.

7.4.1 Basic operating system backup and recovery

On the ThinkSystem solutions, Lenovo implemented a concept of a backup operating system partition that a known-good state can be fallen back to after something occurs on the primary active operating system installation. This feature was made available with the summer 2014 release of the Lenovo solution. It requires more partitions on the RAID 1 on which the operating system is installed. The feature includes a set of tools with which data can be synced from the active operating system partition onto the backup partition and vice versa. Operating system-level updates or change to certain components can be tested and easily rolled back if not successful.

This feature does not replace a full backup of the operating system installation because the data remains on the server and even on the same disks. For entry-level environments or less critical environments, it provides an easy-to-use tool at no added cost that is included as part of the solution.

7.4.2 Basic database backup and recovery

Saving the savepoints and the database logs technically is impossible in a consistent way, and thus does not constitute a consistent backup from which it can be recovered. Therefore, a simple file-based backup of the persistency layer of SAP HANA is insufficient.

Backing up the database

A backup of the SAP HANA database must be triggered through the SAP HANA Cockpit or through the SAP HANA SQL interface. SAP HANA then creates a consistent backup, which consists of one file per SAP HANA service on each cluster node. SAP HANA always performs a full backup. Incremental backups are not supported by SAP HANA.

SAP HANA internally maintains transaction numbers, which are unique within a database instance, especially in a scale-out configuration. To create a consistent backup across a scale-out configuration, SAP HANA chooses a specific transaction number and all nodes of the database instance write their own backup files, including all transactions up to this transaction number.

The backup files are saved to a defined staging area that might be on the internal disks, an external disk on an NFS share,[3] or a directly attached SAN subsystem. In addition to the data

[3] SAP Note 1820529 lists network file systems that are unsuitable for backup and recovery.

backup files, the SAP HANA configuration files and backup catalog files must be saved to be recovered. For point-in-time recovery, the log area also must be backed up.

With the ThinkSystem solution for SAP HANA, one of the 1 Gbit network interfaces of the server can be used for NFS connectivity, or another 10 Gbit network interface must be installed (if a PCIe slot is available). You can add a Fibre Channel host bus adapter (HBA) for SAN connectivity. For more information about supported hardware additions to provide more connectivity, see the Quick Start Guide for the Lenovo Systems Solution for SAP HANA, is available at this website:

https://support.lenovo.com/us/en/docs/UM103346

Restoring a database backup

It might be necessary to recover the SAP HANA database from a backup in the following situations:

- The data area is damaged

  If the data area is unusable, the SAP HANA database can be recovered up to the last committed transaction if all the data changes after the last complete data backup are still available in the log backups and log area. After the data and log backups are restored, the SAP HANA databases uses the data and log backups and the log entries in the log area to restore the data and replay the logs to recover. It also is possible to recover the database by using an older data backup and log backups if all relevant log backups that are made after the data backup are available.[4]

- The log area is damaged

  If the log area is unusable, the only way to recover is to replay the log backups. Therefore, any transactions that are committed after the most recent log backup are lost, and all transactions that were open during the log backup are rolled back.

  After the data and log backups are restored, the log entries from the log backups automatically are replayed to recover. It is also possible to recover the database to a specific point if it is within the existing log backups.

- The database must be reset to an earlier point because of a logical error

  To reset the database to a specific point, a data backup from before that point to recover to and the subsequent log backups must be restored. During recovery, the log area might be used as well, depending on the point to which the database is reset. All changes that are made after the recovery time are (intentionally) lost.

  Since SAP HANA 2 SPS03 the feature of “secondary time travel” of HSR can mitigate the same challenge. More information is available in the Administration guide of the SAP HANA platform:

  https://help.sap.com/viewer/4e9b18c116aa42fc84c7dbfd02111aba/2.0.04/en-US

- You want to create a copy of the database

  You might want to create a copy of the database for various purposes, such as creating a test system.

A database recovery is started from the SAP HANA cockpit or, starting with SAP HANA 1.0 SPS07, from the command line.

Certain restrictions apply when a backup is restored. Up to and including SAP HANA 1.0 SPS06, the target SAP HANA system was required to be identical to the source regarding the number of nodes and node memory size. Starting with SPS07, it is possible to recover a backup that is taken from an m-node scale-out system and restore it on an n-node scale-out

[4] For help with determining the files that are needed for a recovery, see SAP Note 1705945.

environment. Memory configuration also can be different. You must configure m-index server instances on the n-node target environment to restore the backup, which means that nodes can have more than one index server. Such a configuration does not provide the best performance, but it might be sufficient for test or training environments.

When a backup image is restored from a single-node configuration into a scale-out configuration, SAP HANA does not repartition the data automatically. The correct way to bring a backup of a single-node SAP HANA installation to a scale-out solution is by using the following process:

1. Back up the data from the stand-alone node.
2. Install SAP HANA on the master node.
3. Restore the backup into the master node.
4. Install SAP HANA on the subordinate and standby nodes as appropriate and add these nodes to the SAP HANA cluster.
5. Repartition the data across all worker nodes.

For more information about the backup and recovery processes for the SAP HANA database, see *SAP HANA Backup and Recovery Guide*, which is available at this website:

https://help.sap.com/viewer/product/SAP_HANA_PLATFORM/2.0.04/en-US

7.4.3 File-based backup tool integration

By using the mechanisms that are described in 7.4.1, “Basic operating system backup and recovery” on page 139, virtually any backup tool can be integrated with SAP HANA. Backups can be triggered programmatically by using the SQL interface, and the resulting backup files that are written locally then can be moved into the backup storage by the backup tool. Backup scheduling can be done by using scripts that are triggered by the standard Linux job scheduling capabilities or other external schedulers. Because the Backint backup interface was introduced to SAP HANA with SPS05, a file-based backup tool integration is the only option for pre-SPS05 SAP HANA deployments.

7.4.4 Database backups by using storage snapshots

With Lenovo DM storage solutions that run ONTAP data management software, in combination with SnapCenter data protection software, you can meet all of the following challenges of SAP HANA backup and restore operations:

- Long backup operations with performance degradation on production SAP systems
- Unacceptable system downtime due to long restore and recovery operations
- Shrinking backup windows because of the criticality of the applications
- The need for a flexible solution to mitigate logical corruption

In addition, with the Snapshot technology that is included in ONTAP software, you can create backups or execute restore operations of any size dataset in a matter of seconds. SAP HANA supports the use of storage-based snapshot copies as a valid backup operation with documented interfaces.

Backup Operations

SnapCenter and the plug-in for SAP HANA use ONTAP Snapshot technology and the SAP HANA SQL backup interface to give you an SAP-integrated backup solution. SnapCenter gives you automated workflows for backup operations, including retention management for data backups, for log backups, and for the SAP HANA backup catalog.

And for long-term retention, SnapCenter manages the optional replication of application-consistent backups to an off-site secondary location. Your off-site backup storage can be either a physical storage system on the premises or a Cloud Volumes ONTAP instance that runs in Amazon Web Services (AWS) or in Microsoft Azure.

Figure 7-7 shows an overview of the solution architecture.

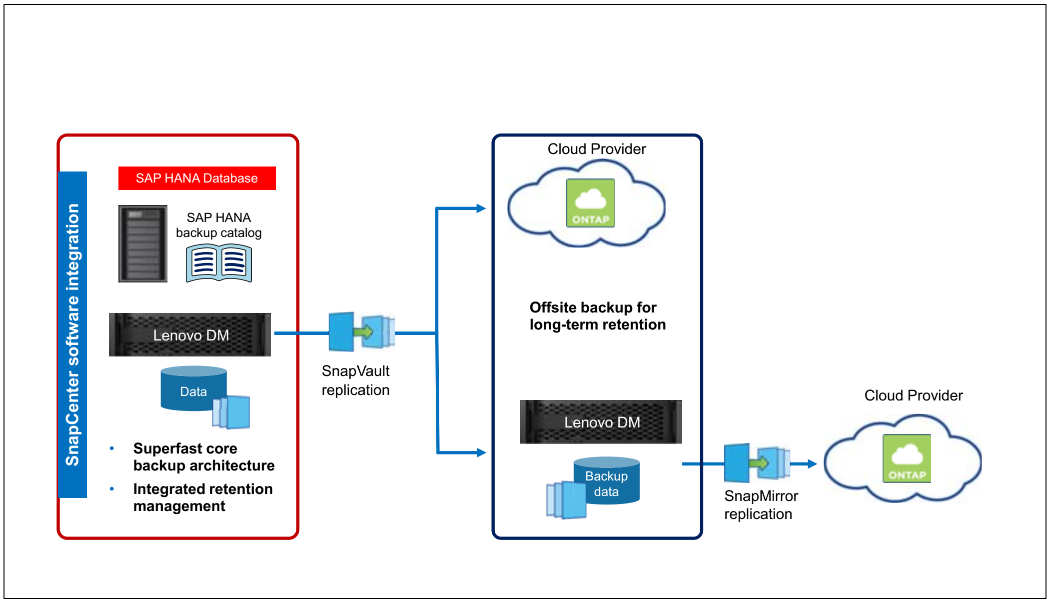


Figure 7-7 Application Integrated backup architecture

Evaluation of customer data has shown that for SAP HANA, the average backup time with Snapshot copies is in the range of a few minutes. In the customer scenario in Figure 7-8 on page 143, a complete backup for a 2.3TB database took 2 minutes and 11 seconds. This is an improvement of 60 – 100 times faster backup operations.

The largest contributor to the overall backup duration is the time that SAP HANA needs to write the synchronized backup savepoint. The amount of time that is required to write the savepoint is a function of the memory of the SAP HANA system and the activity on the system. The storage snapshot operation is performed in a matter of seconds, independent of the size of the database.

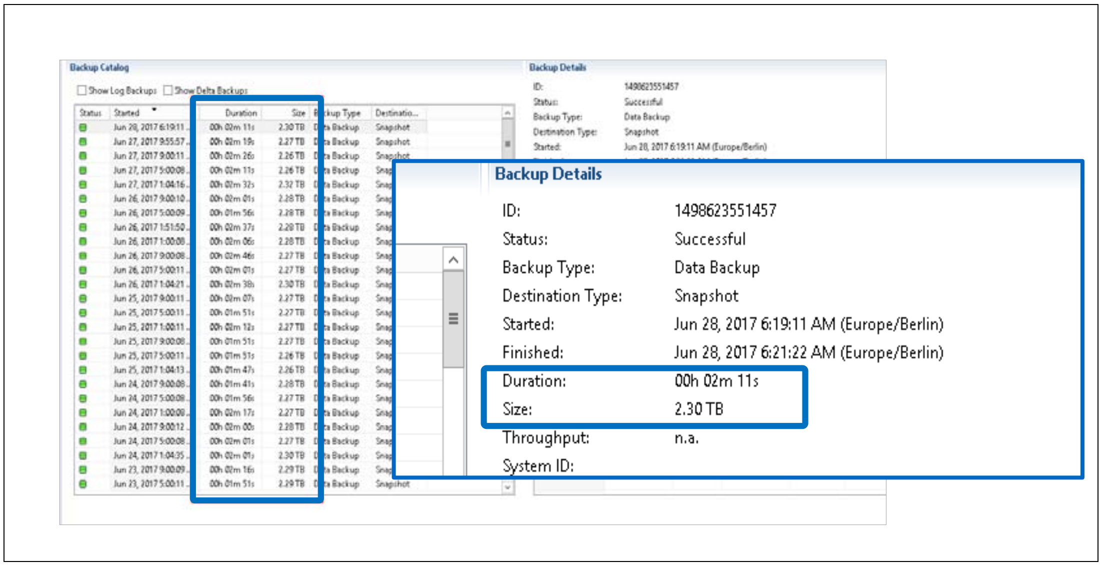


Figure 7-8 Customer example: 2.3TB backed up in 2 min 11 sec

7.4.5 Database backups by using Spectrum Scale snapshots

The scale-out clusters from Lenovo solutions for SAP HANA can use Spectrum Scale as the file system on which SAP HANA runs. Spectrum Scale supports a snapshot feature with which you can take a consistent and stable view of the file system that can then be used to create a backup (which is similar to enterprise storage snapshot features). While the snapshot is active, Spectrum Scale stores any changes to files in a temporary delta area. After the snapshot is released, the delta is merged with the original data and any further changes are applied on this data.

Taking only a Spectrum Scale snapshot does not ensure that you have a consistent backup that you can use to perform a restore. SAP HANA must be instructed to flush out any pending changes to disk to ensure a consistent state of the files in the file system. With the release of SAP HANA 1.0 SPS07, a snapshot feature is introduced that prepares the database to write a consistent state to the data area of the file system (the log area is merged into the data area so only data must be considered for snapshotting). While this snapshot is active, a Spectrum Scale snapshot must be triggered. SAP HANA can then be instructed to release its snapshot. By using Linux copy commands or other more sophisticated backup tools, the data can then be stored in a backup place (NFS share, SAN storage, or other places).

The use of snapshots has much less effect on the performance of the running database than to trigger a file-based backup. Triggering a Spectrum Scale snapshot works in single-node and scale-out environments. The time that it takes to activate a snapshot depends on the amount of data in the file system and the current load on it.

7.4.6 Backup tool integration with Backint for SAP HANA

Since SAP HANA 1.0 SPS05, SAP provides an application programming interface (API) that can be used by manufacturers of third-party backup tools to back up the data and redo logs of an SAP HANA system.[5] By using this “Backint for SAP HANA” API, a full integration with SAP HANA cockpit can be achieved with which backups can be configured and run by using Backint for SAP HANA. With Backint, instead of writing the backup files to local disks,

5 For more information, see SAP Note 1730932 “Using backup tools with Backint”.

dedicated SAN disks, or network shares, SAP HANA creates data stream pipes. Pipes are a way to transfer data between two processes: one is writing data into the pipe, and the other is reading data out of the pipe. This configuration makes a backup by using Backint a one-step backup. No intermediate backup data is written (unlike with a file-based backup tool integration that writes to local disk first), which relieves the local I/O subsystem from the backup workload.

Backing up through Backint

The third-party backup agent runs on the SAP HANA server and communicates with the third-party backup server. SAP HANA communicates with the third-party backup agent through the Backint interface. After the user starts a backup through the SAP HANA Cockpit or by running **`hdbsql`**, SAP HANA writes a set of text files that describe the parameterization for this backup, including version and name information, stream pipe location, and the backup policy to use. Then, SAP HANA creates the stream pipes. Each SAP HANA service (for example, index server, name server, statistics server, and XS engine) has its own stream pipe to which to write its own backup data. The third-party backup agents read the data streams from these pipes, and pass them to the backup server.

Currently, SAP HANA does not offer backup compression; however, third-party backup agents and servers can compress the backup data and further transform it, for example, by applying encryption. Finally, SAP HANA transmits backup catalog information before the third-party backup agent writes a file reporting the result and administrative information, such as backup identifiers. This information is made available in SAP HANA Cockpit.

Restoring through Backint

As described in “Restoring a database backup” on page 140, a database restore might be necessary when the data area or log area is damaged to recover from a logical error or to copy the database. This process can be achieved by using data and log backups that were performed previously.

A restore operation can be started through the SAP HANA Cockpit. For the first step, SAP HANA shuts down the database. SAP HANA then writes a set of text files that describe the parameterization for this restore, including a list of backup identifiers and stream pipe locations. After receiving the backup catalog information from the third-party backup tool, SAP HANA performs a series of checks to ensure that the database can be recovered with the backup data available. Then, SAP HANA establishes the communication with the third-party backup agents by using stream pipes, and requests the backup data from the backup server. The backup agents then stream the backup data that is received from the backup server through the stream pipes to the SAP HANA services. As a final step, the third-party backup agent writes a file reporting the result of the operation for error-handling purposes. This information is made available in SAP HANA Cockpit.

Backint certification

Backup tools that use the Backint for SAP HANA interface are subject to certification by SAP. For more information about the certification process, see the SAP Note: 1730932 and this website:

`http://scn.sap.com/docs/DOC-34483`

To determine which backup tools are certified for Backint for SAP HANA, see the following website and search for enter the search term “HANA-BRINT”

`https://www.sap.com/dmc/exp/2013_09_adpd/enEN/#/d/partners`

As of July 2019 there are multiple solutions certified by SAP. May 2016, the following tools are certified by SAP:

- Apsara HBR 1.6.0
- Arcserve Backup r17.5
- Azure Backup BackInt 1.0
- Bacula Enterprise Edition 10
- Libelle BusinessShadow 6.5
- Cohesity DataProtect 6.0 for SAP HANA
- Commvault V11
- Hitachi Data Protection Suite V11
- DBackup V8.0
- DellEMC Networker 9.2.1
- EISOO AnyBackup 6.0
- Emory 1.0.0 - Cloud Backup for SAP HANA
- Google Cloud Storage
- IBM Spectrum Protect for ERP 7.1[6]
- Micro Focus Data Protector 10
- Veritas NetBackup 8.1.1 and 8.1.2
- NetWorker 9.1
- Rubrik Cloud Data Management v5.0
- SEP sesam 4.4
- HPE StoreOnce Catalyst Plug-in for SAP HANA 2
- Veeam Backup & Replication v9.5 Update 4

7.4.7 Backup and restore as a DR strategy

The use of backup and restore as a DR solution is a basic way of providing DR. Depending on the RPO, it might be a viable way to achieve DR. The basic concept is to back up the data on the primary site regularly (at least daily) to a defined staging area, which might be an external disk on an NFS share or a directly attached SAN subsystem (this subsystem does not need to be dedicated to SAP HANA). After the backup is done, it must be transferred to the secondary site, for example, by a simple file transfer (can be automated) or by using the replication function of the storage system that is used to hold the backup files.

Following a company's DR strategy, SAP HANA must run on the backup site. Therefore, an SAP HANA system must be on the secondary site, which is similar to the system on the primary site at minimum regarding the number of nodes and node memory size.

During normal operations, this system can run other non-productive SAP HANA instances, for example, quality assurance (QA), development (DEV), test, or other second-tier systems. If the primary site goes down, the system must be cleared from these second-tier HANA systems and the backup can be restored. Upon configuring the application systems to use the secondary site instead of the primary one, operation can be resumed. The SAP HANA database recovers from the latest backup if there is a disaster.

[6] IBM Spectrum Protect 7.1 (formerly called TSM for ERP) is only certified for HANA on Power, not on Intel x86.

8

SAP HANA operations

This chapter describes the operational aspects of running an SAP HANA system.

This chapter includes the following topics:

8.1 Installation services

The Lenovo System Solution for SAP HANA features the complete software stack, including the operating system, the filesystem, and the SAP HANA software. Because of the nature of the software stack, and dependencies on how the Lenovo System Solution for SAP HANA is used at the client location, the software stack cannot be preinstalled completely at manufacturing. Therefore, installation services are recommended. The Lenovo System Solution for SAP HANA offers the following installation services:

- Performing an inventory and validating the delivered system configuration
- Verifying and updating the hardware to the latest level of basic input/output system (BIOS), firmware, device drivers, and operating system patches as required
- Verifying and configuring the Redundant Array of Independent Disks (RAID) configuration
- Finishing the software preinstallation according to the client environment
- Configuring and verifying network settings and operation
- Performing system validation
- Providing onsite skills transfer (when required) of the solution and preferred practices, and delivering postinstallation documentation

To ensure the correct operation of the system, installation services for the Lenovo System solution for SAP HANA should be performed by trained personnel who are available from Lenovo Professional Services or Lenovo Business Partners depending on your geography.

8.2 Lenovo SAP HANA Operations Guide

The Operations Guide describes the operations of a Lenovo System Solution for SAP HANA and covers the following topics:

- System Check

  How to obtain Spectrum Scale cluster status and configuration, file system status and configuration, disk status and usage, quotas, SAP HANA application status, and network information from the switches.

- Troubleshooting
  - Actions to take after a server node failure, removing the SAP HANA node from the cluster, and installing a replacement node.
- Single Node Operations
- Cluster operations
- Hard Drive Operations
  - Checking the drive configuration and the health of the drives.
  - Replacing failed hard disk drives (HDDs), solid-state drives (SSDs), and High input/output operations per second (IOPS) devices and reintegrating them into Spectrum Scale.
  - Driver and firmware upgrades for the High IOPS devices, HDDs, and SSDs.
- Software Updates

  Checklists for what to do to update the Linux kernel, the drivers and firmware, including instructions on how to perform a rolling upgrade, where applicable.

- Operating System Upgrades
  - Both for SLES and RHEL
- Drive operations:
  - Checking the drive configuration and the health of the drives.
  - Replacing failed hard disk drives (HDDs), solid-state drives (SSDs), and High input/output operations per second (IOPS) devices and reintegrating them into Spectrum Scale.
  - Driver and firmware upgrades for the High IOPS devices, HDDs, and SSDs.
- References to related documentation, pointing to important documentation from Lenovo, SAP, and SUSE or RedHat.

The *Lenovo SAP HANA Operations Guide* is being optimized and extended continuously based on new developments and client feedback. The latest version of this document can be found in SAP Note 1650046.[1]

8.3 Interoperability with other platforms

To access the SAP HANA database from a system (SAP or non-SAP), the SAP HANA database client must be available for the platform on which the system is running. Platform availability of the SAP HANA database client is documented in the Product Availability Matrix (PAM) for SAP HANA, which is available at the following website (search for “HANA”):

http://support.sap.com/pam

If there is no SAP HANA database client available for a certain platform, SAP HANA can still be used in a scenario with replication by using a dedicated SAP Landscape Transformation server (for SAP Business Suite sources) or an SAP BusinessObjects Data Services server that is running on a platform for which the SAP HANA database client is available. By using this configuration, data can be replicated into SAP HANA, which then can be used for reporting or analytic purpose by using a front end that supports SAP HANA as a data source.

8.4 Monitoring SAP HANA

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

The SAP tool for the administration and monitoring of the SAP HANA appliance is the SAP HANA Studio or the SAP HANA Cockpit with which you can monitor the following aspects of the overall system state:

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

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

[1] SAP Notes can be accessed at http://support.sap.com/notes. An SAP S-user ID is required.

Note: For more information about the administration and monitoring of SAP HANA, see the *SAP HANA Administration Guide*, which is available at this website:

http://help.sap.com/hana_platform

8.5 Installing more agents

Many organizations have processes and supporting software in place to monitor, back up, or otherwise interact with their servers. Because SAP HANA is delivered in a well defined model, there are restrictions regarding extra software (for example, monitoring agents) to be installed on to the appliance. SAP permits the installation and operation of external software if the prerequisites that are described in SAP Note 1730928 are met.

Only the software that is installed by the hardware partner is recommended on the SAP HANA system. For the Lenovo solution for SAP HANA the following categories of agents are defined:

- Supported

  Lenovo provides a solution that covers the respective areas; no validation by SAP is required.

- Tolerated

  Solutions that are provided by a third party can be used on the Workload Optimized Solution for SAP HANA. It is the clients' responsibility to obtain support for such solutions. Such solutions are not validated by Lenovo and SAP. If issues with such solutions occur and cannot be resolved, the use of such solutions might be prohibited in the future.

- Prohibited

  These types of solutions must not be used on the Lenovo solution for SAP HANA. The use of these solutions might compromise the performance, stability, or data integrity of SAP HANA.

Do not install more software on the SAP HANA system that is classified as prohibited for use on the SAP HANA appliance. For example, initial tests show that some agents can decrease performance or possibly corrupt the SAP HANA database (for example, virus scanners).

In general, all added software must be configured to not interfere with the functions or performance of the SAP HANA system. If any issues with the SAP HANA system occur, you might be asked by SAP to remove all added software and to reproduce the issue.

For more information about the agents that are supported, tolerated, or prohibited for use on the SAP HANA system, see the Quick Start Guide. The most recent version is available at this website:

https://support.lenovo.com/us/en/docs/UM103346

8.6 Software and firmware levels

The Lenovo Solution for SAP HANA includes several components that might need to be upgraded (or downgraded) depending on different support organizations' recommendations. These components can be split up into the following general categories:

- Firmware

- Operating system
- Hardware drivers
- Software

The Lenovo SAP HANA solution support team reserves the right to perform basic system tests on these levels when they are deemed to have a direct effect on the SAP HANA appliance. In general, Lenovo does not give specific recommendations about which levels are allowed for the SAP HANA system.

The Lenovo Solution for SAP HANA development team provides new images for the SAP HANA system at regular intervals. Because these images include dependencies regarding the hardware, operating system, and drivers, use the latest image for maintenance and installation of SAP HANA systems. These images can be obtained through Lenovo Support. Part number information is contained in the Quick Start Guide.

If the firmware level recommendations for the hardware components of the SAP HANA system are given through the individual Lenovo support teams that fix known code bugs, it is the client's responsibility to upgrade or downgrade to the recommended levels as instructed by Lenovo Support.

If the operating system recommendations for the SUSE or RedHat Linux components of the SAP HANA system are given through the SAP, SUSE, RedHat, or Lenovo support teams that fix known code bugs, it is the client's responsibility to upgrade or downgrade to the recommended levels, as instructed by SAP through an explicit SAP Note or allowed through an OSS Customer Message. SAP describes their operational concept, including updating of the operating system components, in SAP Note 1599888 - *SAP HANA: Operational Concept*. If the Linux kernel is updated, it is recommended that extra care is taken whenever a software driver recompile is required, as described in the *Lenovo SAP HANA Operations Guide* (see 8.2, "Lenovo SAP HANA Operations Guide" on page 148).

If a driver or file system recommendation to update the software is given through the individual support teams (ThinkSystem, Linux, or file system) to fix code bugs, it is *not* recommend to update these drivers without first asking the Lenovo SAP HANA solution support team through an SAP OSS Customer Message.

If recommendations for upgrades of other hardware or software components of the SAP HANA appliance are given through the individual Support teams that fix code bugs, it is the client's responsibility to upgrade or downgrade to the recommended levels as instructed by Support.

8.7 Lenovo and SAP Integrated support process

The deployment of SAP HANA as an integrated solution, which combines software and hardware from Lenovo and SAP, also is reflected in the support process for the Lenovo Solution for SAP HANA. All SAP HANA models that are offered by Lenovo include SUSE Linux Enterprise Server (SLES) for SAP Applications with SUSE 1-year or 3-year priority support or Red Hat Enterprise Linux for SAP HANA with 1-year or 3-year support. The hardware features a 1-year or 3-year limited warranty, including customer-replaceable unit (CRU) and onsite support.

SAP integrates the support process with SUSE, RedHat, and Lenovo as part of the HANA system solution-level support. If you encounter software problems on your SAP HANA system, see the following SAP Online Service System (SAP OSS) website:

https://launchpad.support.sap.com/

At the website, create a service request ticket by using a subcomponent of BC-HAN or BC-DB-HDB as the problem component. Lenovo support works closely with SAP, SUSE, and RedHat and is dedicated to supporting SAP HANA software and hardware issues.

Send all questions and requests for support to SAP by using their OSS messaging system. A dedicated Lenovo representative is available at SAP to work on this solution. Even if it is clearly a hardware problem, an SAP OSS message should be opened to provide the best direct support for the Lenovo System Solution for SAP HANA.

When an SAP support message is opened, use the text template that is provided in the Quick Start Guide when it is obvious that you have a hardware problem. This procedure expedites all hardware-related problems within the SAP support organization. Otherwise, the SAP support teams helps you with the questions regarding the SAP HANA appliance in general.

Lenovo provides a script to get an overview of the current system status and the configuration of the running system. The **`saphana-support-lenovo.sh`** script is preinstalled in the `/opt/lenovo/saphana/bin` directory. The most recent version can be found in SAP Note 1661146. Previously, the script was called `saphana-support-ibm.sh` and located in `/opt/ibm/saphana/bin` directory. Both scripts can coexist on one system.

Before you contact support, ensure that you take the following steps to try to solve the problem:

- Use the troubleshooting information in your system documentation and the diagnostic tools that is included with your system. Information about diagnostic tools is available in the *Problem Determination and Service Guide* on the Documentation CD that is included with your system.
- See the following support website to check for technical information, hints, tips, and new device drivers
- For SAP HANA software-related issues, you can search the SAP OSS website for problem resolutions. The OSS website includes a knowledge database of known issues and can be accessed at the following website:

  https://launchpad.support.sap.com/

The main SAP HANA information source is available at the following website:

https://help.sap.com/hana_platform

If you have a specific operating system question or issue, contact SUSE regarding SUSE Linux Enterprise Server for SAP Applications. For more information, see the following SUSE website:

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

Media is available for download at the following website:

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

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

If you are running RedHat Enterprise Linux for SAP HANA, contact your RedHat support representative or online at the following website:

https://access.redhat.com/home

A

Additional topics

This appendix includes the following topic:

A.1 Supporting Lenovo organizations

The following section lists the organizations within Lenovo dedicated to work on topics related to SAP.

A.1.1 Lenovo SAP Center of Competence

The Lenovo SAP Center of Competence (Lenovo SAP CoC) is the key support function of the Lenovo and SAP alliance. It serves as a single point of entry for all SAP-related questions for clients that are using System x and ThinkSystem hardware. As a managed question and answer service, it has access to a worldwide network of experts on technology topics about System x and ThinkSystem products in SAP environments. You can contact the Lenovo SAP CoC by using the following email address:

SAPsolutions@lenovo.com

> **Note:** The Lenovo SAP CoC does not provide product support. For more information about product support for the Lenovo System x solution for SAP HANA, see 8.7, “Lenovo and SAP Integrated support process” on page 151.
>
> If you need support for other Lenovo products, consult the product documentation for more information about how to get support.

A.1.2 Lenovo Executive Briefing Center and Innovation Center

Lenovo operates three Executive Briefing Centers (EBC) around the world. The EMEA EBC covers Europe, Middle East, and Africa and is located in Stuttgart, Germany. The EBC covering America is located with the Lenovo DCG head quarters in Morrisville, NC, in the USA. The third EBC covers Asia Pacific and is located in Beijing, China.

All Centers have a showcase area where Lenovo product are featured and facilities to hold technology briefings and solution workshops. The briefing coordinator reaches out to the respective Lenovo subject matter experts that help to demonstrate how technology can solve business problems.

Each Briefing Center has an Innovation Center co-located with it that can be used for different purposes to further evaluate a certain technology. Examples of such engagements are providing customer access to a specific solution to try it out, or to run proof of concept and proof of technology implementations of custom scenarios.

More information can be found on the following web site:

- https://www.lenovo.com/us/en/customer-centers/dcg/raleigh

Related publications

The publications that are listed in this section are considered suitable for a more detailed description of the topics that are covered in this book.

Lenovo Press publications

The following Lenovo Press publications provide more information about the topics in this document. Some publications that are referenced in this list might be available in softcopy only:

- *Lenovo System x3850 X6 and x3950 X6 Planning and Implementation Guide*, SG24-8208

You can search for, view, download, or order these documents and other Lenovo Press publications, papers, drafts, and other materials, at the following website:

http://lenovopress.com

Online resources

The following websites are also relevant as further information sources:

- System x solution for SAP HANA:

  http://shop.lenovo.com/us/en/systems/solutions/alliances/sap/#tab-sap_hana
- SAP In-Memory Computing - SAP Help Portal:

  http://help.sap.com/hana

Help from Lenovo

Lenovo support and downloads:

http://www.lenovo.com/support

Lenovo services:

http://www.lenovo.com/services

Lenovo

In-memory Computing with SAP HANA on Lenovo Systems

Lenovo

In-memory Computing with SAP HANA on Lenovo Systems

Introduces the Lenovo Solution for SAP HANA

Explores the SAP HANA features and use cases

Explains the delivery models

Describes business continuity and operational disciplines

The SAP HANA platform has come a long way. Experience grew over the past nine years.

The sixth edition of this Lenovo Press publication describes in-memory computing systems from Lenovo and SAP that are based on Lenovo ThinkSystem products and SAP HANA. It covers the basic principles of in-memory computing, describes the Lenovo Thinksystem hardware offerings including the possibilities around persistent memory, and explains the corresponding SAP HANA IT landscapes using these offerings.

This book also describes the architecture and components of the Lenovo System solution for SAP HANA. The SAP HANA operational disciplines are explained in detail: Scalability options, high availability and disaster recovery, backup and restore, and virtualization possibilities for SAP HANA systems.

This book is intended for SAP administrators and technical solution architects. It is also for Lenovo Business Partners and Lenovo employees who want to know more about the SAP HANA offering and available Lenovo solutions for SAP clients.

LENOVO PRESS

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

At Lenovo Press, we bring together experts to produce technical publications around topics of importance to you, providing information and best practices for using Lenovo products and solutions to solve IT challenges.

lenovopress.com

SG24-8086-04