



Reference Architecture for Lenovo Storage V Series Solution for SAP HANA

Last update: **09 July 2018**

Version: 1.0

**Describes SAP HANA TDI
setup on on Lenovo Storage V
Series SAN systems**

**Includes operating system setup and
tuning**

**Provides details about sizing,
configuration and storage
layout**

**Lists the recommended
hardware and software
components**

Levente Szente

Trick Hartman

Gereon Vey



Table of Contents

1	Introduction	1
1.1	Terminology	1
2	Business problem and business value	2
2.1	Business problem	2
2.2	Business value	2
3	Requirements	3
3.1	Functional requirements	3
3.2	Quality Attributes	3
4	Architectural overview	4
5	Component model	6
5.1	Component relationship diagram	8
6	Operational model	9
6.1	Hardware Components	9
7	Deployment considerations	11
7.1	Lenovo Storage sizing	11
7.1.1	Drives	11
7.1.2	Easy Tier	11
7.1.3	Storage performance sizing for SAP HANA	12
7.1.4	Definitions of storage pools and volumes	12
7.1.5	Layout of storage pools and storage volumes	12
7.1.6	Competing storage utilization	13
7.1.7	The use of Lenovo storage compression (RtC)	13
7.2	Systems management	13
7.3	Networking	14
7.4	Operating system setup & tuning	14
7.5	Linux setup & tuning	14
7.6	Multipath changes	15

7.6.1	Using aliases for UUIDs in <code>multipath.conf</code>	17
7.7	IO device tuning	17
7.8	IO scheduler.....	17
7.9	Setup of shared access for the <code>/hana/shared/</code> directory.....	18
7.10	SAN zoning and volume mapping.....	18
7.11	Setup of file systems for DATA and LOG	18
7.12	VMware Virtual Volumes	20
7.13	Setup of <code>/hana/shared/global.ini</code>	20
7.14	<code>hdbparam fileio</code> parameter	20
7.15	Other considerations.....	21
8	Appendix: Lenovo Bill of materials.....	22
	Resources	23
	Document history	24

1 Introduction

This document describes the reference architecture for the SAP HANA Enterprise Storage certification based on the SAP Tailored Data Center Integration (TDI) rules from SAP SE where were used for the Lenovo V Series family of SAN storage systems. The Lenovo and SAP SE teams worked together on the architectural vision and joint engineering effort to create this Reference Architecture. This paper in intended to provide planning, design considerations, and best practices for implementing SAP HANA installation using Lenovo V Series storage products through a storage attached network (SAN) and Lenovo servers.

The architecture described herein has been validated by both Lenovo and SAP SE. SAP SE has certified this configuration on their SAP HANA hardware directory. For a list of all storage systems certified for SAP HANA TDI, please visit: <https://www.sap.com/dmc/exp/2014-09-02-hana-hardware/enEN/enterprise-storage.html>.

The intended audience of this document is IT professionals, technical architects, sales engineers, and consultants to assist in planning, designing and implementing SAP HANA TDI on Lenovo V Series storage. The SAP HANA TDI allows the SAP customer to use external storage to attach to the SAP HANA server.

This document provides the necessary information to select, verify and connect Lenovo V Series Storage to the SAP HANA server through a SAN (fibre channel) attached storage network.

1.1 Terminology

System	The term “system” will be used for a single instance of the SAP HANA platform edition software. For example, if four (4) SAP HANA virtual instances are deployed on a single Lenovo server this reflects four (4) SAP HANA systems in this paper.
Compute Node	Describes the certified server attached to the Enterprise Storage subsystem used for SAP HANA
Storage Node	In this paper, the Lenovo Storage V series family of systems which can be used to host SAP HANA persistent data include: <ul style="list-style-type: none">• Lenovo V3700 V2• Lenovo V3700 V2 XP• Lenovo V5030
Network Node	Describes the networking switches and their configurations necessary to connect Compute and Storage nodes together

2 Business problem and business value

This section provides a summary of the business problems that this reference architecture is intended to help address, and the additional value that this solution provides for storing large amounts of data.

2.1 Business problem

Transitioning to SAP HANA is a big step. One needs not only the computing resources necessary for running an in-memory database, but also the back-end storage that fulfills all of the SAP HANA key performance indicators (KPI). Without an assurance that these KPIs are met, it is not 100% certain that your SAP HANA project will succeed.

Although Lenovo provides a solution based on local storage, it may be your company's policy to use external SAN storage for your mission critical data. This data should be stored on a storage subsystem that meets those SAP HANA KPIs while providing the SAP HANA Storage API.

2.2 Business value

The Lenovo Reference Architecture for Lenovo Storage V Series Solution for SAP HANA provides a tested and certified SAN storage solution for up to 12 SAP HANA worker nodes. This can be used as the basis for a successful SAP HANA installation whether used for analytical workloads like SAP Business Warehouse in scale-out situation, or using this for transactional applications like S/4HANA. Naturally, virtualized SAP HANA setups can easily be implemented as well using VMware ESX to virtualize your entire landscape on top of the Lenovo Reference Architecture for Lenovo Storage V Series Solution for SAP HANA.



Figure 1: Lenovo V Series V3700 V2

3 Requirements

There are several requirements that should be taken into account before starting a deployment of an SAP HANA Enterprise Storage setup. What follows are some functional and additional requirements you should become familiar with before starting.

3.1 Functional requirements

In order to implement Lenovo Reference Architecture for Lenovo Storage V Series Solution for SAP HANA, you need to set up your compute, storage and networking nodes as described in the next chapters. SAP recommends a storage validation for SAP HANA Enterprise Storage setups according to their Key Performance Indicators (KPI) for TDI installations. These can be found in SAP Note 1943937:

<https://launchpad.support.sap.com/#/notes/1943937>

The Lenovo team has affirmed with SAP ICC that the functional requirements listed here are necessary to fulfill the SAP HANA TDI KPI values.

The SAP HANA compute nodes must be sized and configured according to SAPs specification. This can be obtained by the Lenovo Enterprise Technical Services (LETS) team. In addition each SAP HANA compute node needs to have a minimum of two (2) fibre channel ports, we recommend four (4) fibre channel ports per compute node.

A SAN environment is required to attach the Lenovo Storage through fibre channel with the SAP HANA compute nodes. While 8Gb/s fibre channel adapters will suffice, we recommended to use 16Gb/s adapters in all newer configurations.

3.2 Quality Attributes

In addition to the obvious features and functions, there are other requirements that may pertain to the installation when the topic arises. These quality attributes are organized by a number of common themes or subcategories. These include the areas of performance and capacity, availability, usability, security and privacy, maintainability, manageability and flexibility.

Table 1: List of Quality Attributes for SAP HANA Installations

Requirement	Description
Scalability	Solution components scale for growth
Load balancing	Workload is distributed evenly across servers
Fault tolerance	High availability and Disaster Recovery options
Data Recovery	Backup and Restore of SAP HANA
Ease of installation	Reduced complexity for solution deployment
Ease of management/operations	Reduced complexity for solution management
Security	Solution provides means to secure customer infrastructure
High performance	Performance Optimizations necessary

4 Architectural overview

SAP HANA Platform edition is an in-memory database which saves a copy of its persistent data on disk. This disk can be either local, for systems which run only on one server, or it can be remote using some form of attached storage. The second method is useful when implementing SAP HANA in a scale-out or clustered variant. If virtualization is to be used, and the customer wants to use the benefits of their virtualization software for topics like high-availability and disaster recovery, then an externally attached storage system is also very helpful.

The Lenovo Reference Architecture for Lenovo Storage V Series Solution for SAP HANA can run a Single-Host or a Multiple-Host SAP HANA System in a SAN environment. Multiple ThinkSystem servers can be connected to the SAN and share the same storage array.

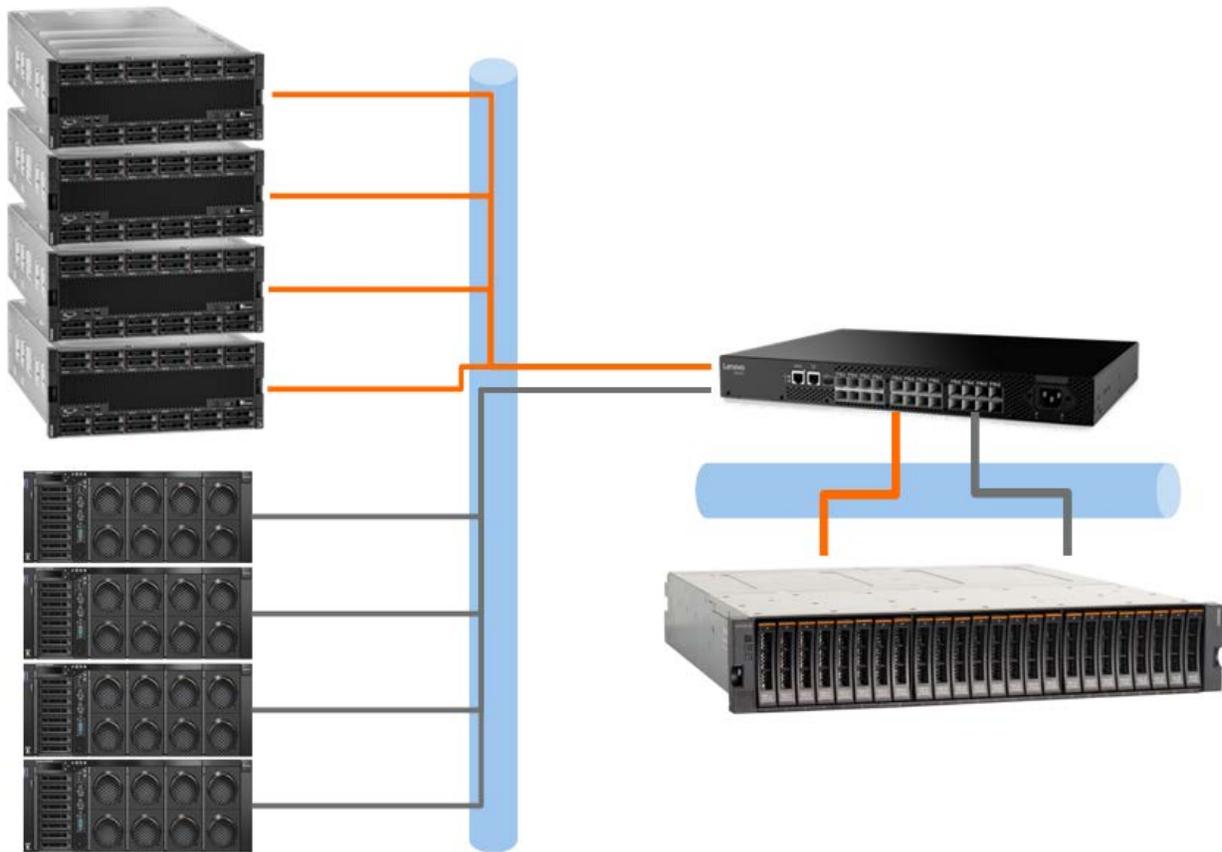


Figure 2: Lenovo Reference Architecture for Lenovo Storage V Series Solution for SAP HANA Architecture

An SAP HANA Multiple-Host System can run on a shared file system, such as NFS or GPFS, or it can run on a non-shared storage which must be used in combination with the SAP Fibre Channel Storage Connector. The Fibre Channel Storage Connector is a ready to use implementation of SAP HANA's Storage Connector API provided by SAP.

5 Component model

The SAP HANA Storage Connector API basically manages the automatic mounting/unmounting of devices on the HANA nodes employing an I/O fencing mechanism. A detailed description can be found in the SAP HANA Fibre Channel Storage Connector Admin Guide available under the link provided in the resources section, below is just an overview.

The Storage Connector API provides hooks for database startup and for failing-over nodes in setups with native multipathing of fibre channel attached devices. The provided hooks are Python calls within the SAP HANA nameserver.

The Storage Connector API will act as an agent for managing the shared volumes from the Lenovo V Series Storage Solution to ThinkSystem servers. The agent then will manage the automatic mounting/unmounting of devices on the HANA nodes.

SCSI-3 Persistent Reservations are written to the devices to prevent any data corruption due to multiple hosts accessing a device at the same time. SCSI-3 Persistent reservations are a standard for a wide range of storage subsystems with multiple ways to ensure I/O fencing. Basically a key must be registered on a device, which is used to make a reservation. A reservation can only be made by the host, which previously registered this key. Finally, only if a device is reserved, the I/O fencing is active.

The following figure shows a common, highly available, setup of storage connection via Fibre Channel:

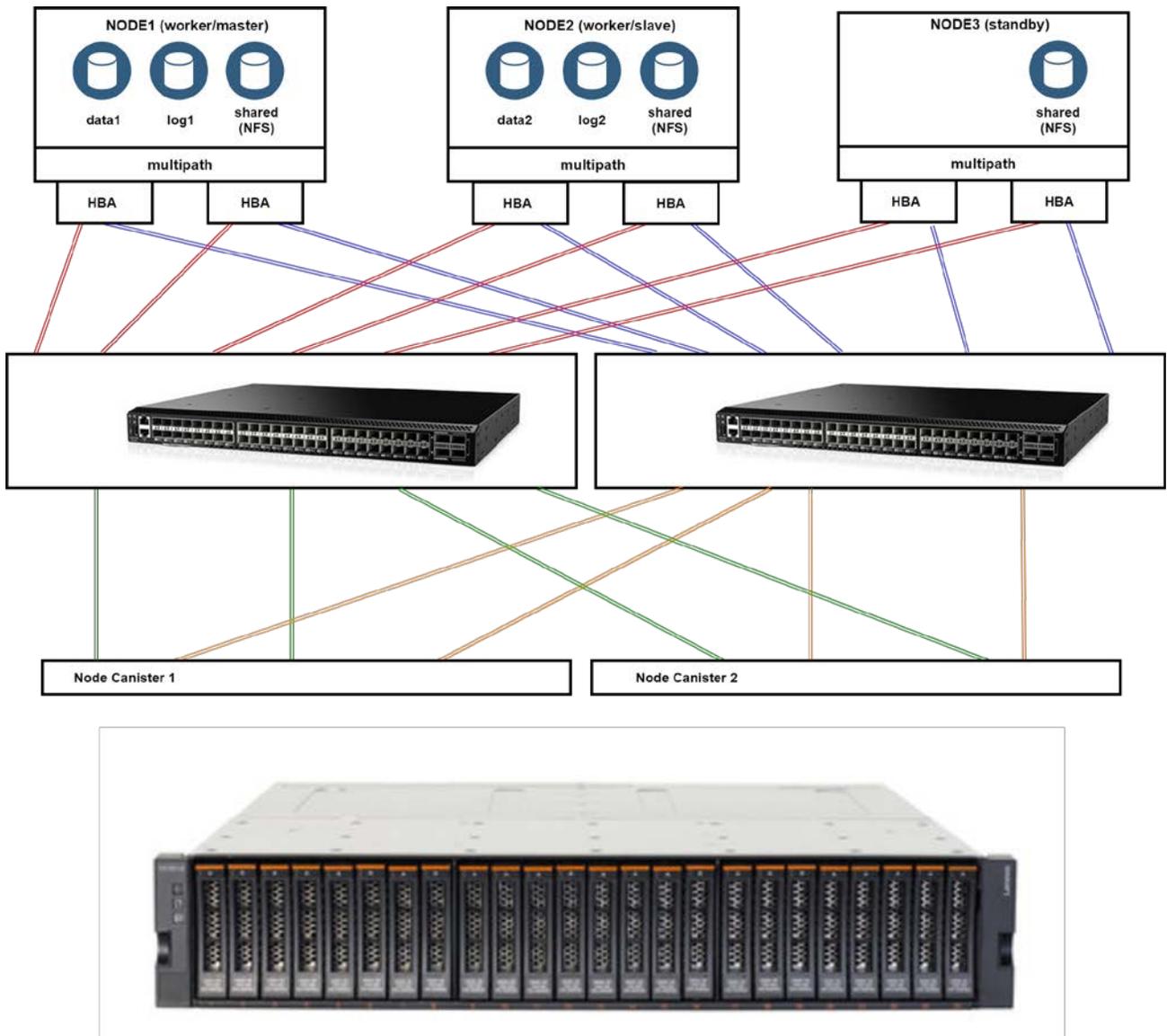


Figure 3: Highly available storage connection via Fibre Channel

All three hosts have each two host bus adapters installed. The first HBA is connected to the fibre channel switch 1 and second HBA is connected to the fibre channel switch 2. Both switches have four connections to the storage subsystem in place, two from each controller. From the host's perspective this means that each LUN (represented as `/dev/mapper/<wwid>` device) is available through sixteen paths (represented as `/dev/sd*devices`), which are managed by the multipath daemon within the Linux kernel. The SCSI-3 persistent reservation tool `sg_persist` directly works on the single paths.

5.1 Component relationship diagram

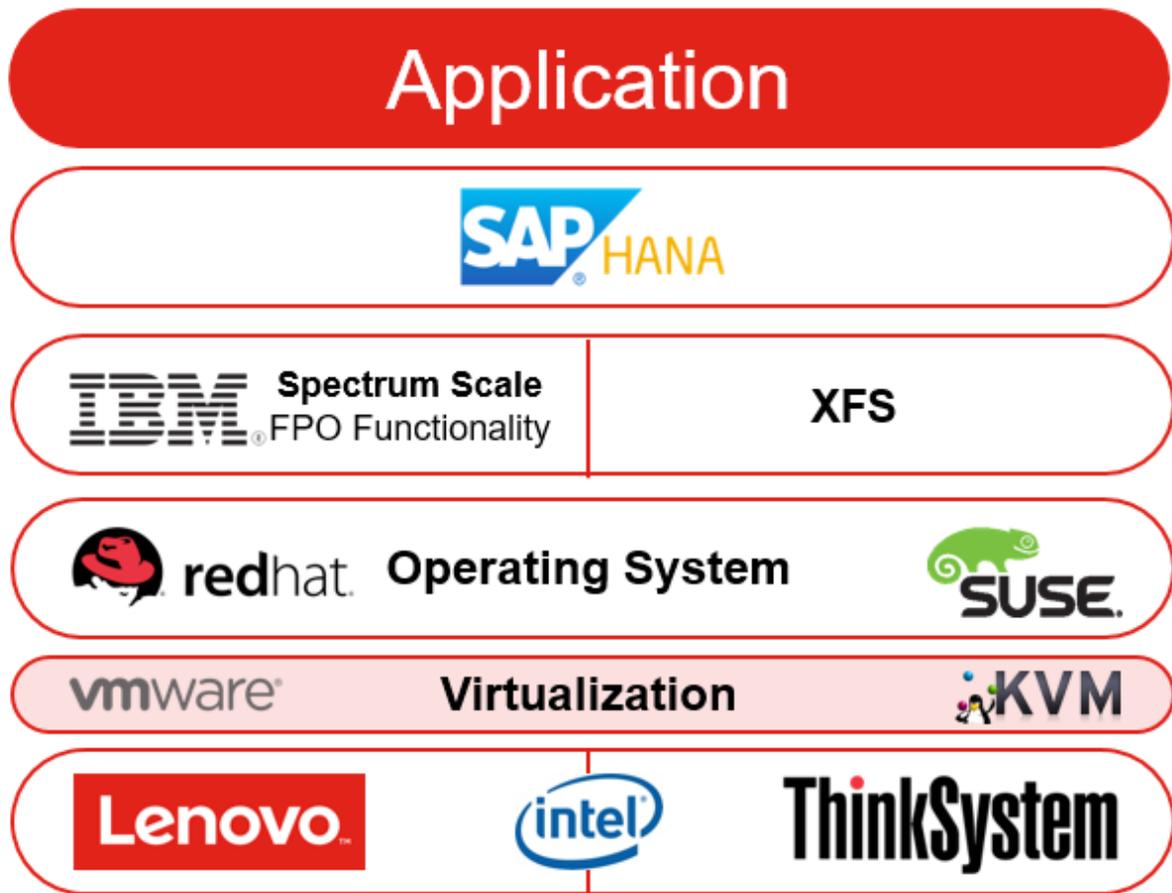


Figure 3: SAP HANA Component Diagram

6 Operational model

This section describes the validated Lenovo V Series Storage. It concludes with an example deployment model.

6.1 Hardware Components

Lenovo Storage V3700 V2, V3700 V2 XP and V5030 are members of the Lenovo Storage V Series family of disk systems. By using IBM Spectrum Virtualize software functions, interoperability, and management tools, they deliver innovation and new levels of storage efficiency with ease of use in an entry-level disk system to enable organizations to overcome their storage challenges.

Lenovo Storage V3700 V2 and V3700 V2 XP storage systems feature two node canisters, with 8 GB cache per canister, upgradeable to 16 GB per canister (V3700 XP only), in a compact, 2U, 19-inch rack mount enclosure. 12 Gb SAS (V3700 V2 XP only) and 1 Gb iSCSI connectivity is standard, with an option for additional 1 Gb iSCSI, 12 Gb SAS, 16 Gb Fibre Channel (FC), or 10 Gb iSCSI or Fibre Channel over Ethernet (FCoE) connectivity.

Lenovo Storage V5030 control enclosure models include two node canisters, each with 16 GB cache upgradeable to 32 GB per canister for a system total of 64 GB cache. 1/10 Gb iSCSI connectivity is standard, with an option for additional 1 Gb iSCSI, 12 Gb SAS, 16 Gb Fibre Channel (FC), or 10 Gb iSCSI or Fibre Channel over Ethernet (FCoE) connectivity.

Lenovo Storage V3700 V2 and V3700 V2 XP support up to 264 drives with up to ten expansion enclosures. They also offer flexible drive configurations with the choice of 2.5-in. and 3.5-in. form factors, 10 K or 15 K rpm SAS and 7.2 K rpm NL SAS hard disk drives (HDDs), and SAS solid-state drives (SSDs).

A Lenovo Storage V5030 system scales up to 504 drives with the attachment of up to 20 Lenovo Storage V5030 expansion enclosures. It also offers flexible drive configurations with the choice of 2.5-inch and 3.5-inch form factors, 10 K or 15 K rpm SAS and 7.2 K rpm NL SAS hard disk drives (HDDs), and SAS flash drives (solid-state drives [SSDs]).

The Lenovo Storage V Series Family of storage systems are designed to offer high levels of system and data availability with the following technologies:

- Dual-active, intelligent node canisters with mirrored cache
- Dual-port drives with automatic drive failure detection and RAID rebuild with global hot spares
- Redundant hardware, including power supplies and fans
- Hot-swappable and customer replaceable components
- Automated path failover support for the data path between the server and the drives

A switch attach solution - or SAN - places a switch between the servers and the controller enclosures within the storage system. Using switches, a SAN shares a storage system among multiple servers, reducing the number of storage systems required for a particular environment. Using switches increases the number of servers that can be connected to the storage system.

DB620S

The Lenovo ThinkSystem DB620S FC SAN Switch provides exceptional price/performance value by delivering market-leading 32 Gb Gen 6 Fibre Channel technology and combining flexibility, simplicity, and enterprise-class functionality to meet the demands of growing flash-based storage environments.

Designed to enable maximum flexibility and reliability, the ThinkSystem DB620S is a compact, 1U rack-mount FC switch that offers low-cost access to industry-leading Storage Area Network (SAN) technology while providing “pay-as-you-grow” scalability to meet the needs of an evolving storage environment.

The DB620S FC SAN Switch offers 24x SFP+ ports that support 4/8/16/32 Gbps speeds. The DB620S FC SAN switch provides easy integration into the existing SAN environments while realizing the benefits of Gen 6 Fibre Channel connectivity, and the switch offers a rich set of standard software features with the options to expand its capabilities as needed.

The DB620S FC SAN Switch features the EZSwitch Setup wizard and can be configured in Access Gateway Mode to simplify deployment. The switch provides full non-blocking performance with Ports On Demand scalability to support SAN expansion and enable long-term investment protection.

7 Deployment considerations

Before starting to deploy a SAP HANA TDI, ensure that you have a SAN environment to attach the Lenovo V Series Storage through fibre channel with the SAP HANA TDI nodes. We recommend using the Lenovo ThinkSystem DB610S or DB620S FC SAN Switch for the SAN.

SAP requires a storage validation for SAP HANA TDI according to their KPI (Key Performance Indicators) – please contact SAP for further details.

Also the SAP HANA nodes must be sized and configured according to SAPs specification (Bill of Material for SAP HANA node).

7.1 Lenovo Storage sizing

All components of the SAN infrastructure must be configured to use a minimum 8GB link speed with 4 links. While 8Gb/s fibre channel adapters will suffice, we recommended to use 16Gb/s adapters in all newer configurations. The Storage FC ports support use of qualified 8-Gbit/s or 16-Gbit/s SFPs.

All listed storage configurations apply to the active SAP HANA TDI nodes only; no additional storage configuration is required for any stand-by node.

Table 2: Lenovo Storage sizing for SAP HANA

Lenovo System Storage	Nr. of SAP HANA Productive Instances
V3700 V2	8
V3700 V2 XP	10
V5030	12

7.1.1 Drives

For your workload you can create disk groups comprised of only SSDs, only spinning disks, or a combination of both. You may choose either any of the Lenovo offered Enterprise HDD or SSD drives. These drives differ in the number of Drive Writes Per Day (DWPD) meaning one full overwrite of the entire drive capacity per day at maximum can be supported. Options of either 1 (Enterprise Capacity), 3 (Enterprise Mainstream) or 10 (Enterprise Performance) DWPD are supported for the Lenovo V Series Storage, for typical HANA workload SSDs with 3 DWPD will suffice.

During lifetime of the any SSD, the capacity utilization should not exceed 80%. If the utilization approaches this limit, additional capacity of the same kind and device capacity need to be added to the storage system.

7.1.2 Easy Tier

Easy Tier is an optional licensed function of the Lenovo V Series Storage that brings enterprise storage enhancements to the entry and midrange segment. It enables automated subvolume data placement throughout different storage tiers to intelligently align the system with its current workload requirements and to optimize storage usage. This function includes the ability to automatically and non disruptively relocate data (at the extent level) from one tier to another tier in either direction to achieve the best available storage performance workload for an specific environment.

Easy Tier reduces the I/O latency for hot spots, but it does not replace storage cache. Easy Tier and storage cache solve a similar access latency workload problem, but these methods weigh differently in the algorithmic construction based on “locality of reference,” recency and frequency. Because Easy Tier monitors I/O performance from the extent end (after cache), it is able to pick up the performance issues that cache cannot solve and complement the overall storage system performance.

In general, the storage environment I/O is monitored on volumes and the entire volume is always placed inside one appropriate storage tier. Determining the amount of I/O on single extents is too complex for monitoring I/O statistics, to move them manually to an appropriate storage tier and to react to workload changes.

7.1.3 Storage performance sizing for SAP HANA

In order to achieve the necessary performance which meets the SAP HANA TDI KPI values, we recommend to use Distributed RAID5 or RAID6 arrays in the Lenovo Reference Architecture for Lenovo Storage V Series Solution for SAP HANA. This is regardless of the model or type of drive.

HANA storage requirements:

- Log volume: ½ of the RAM size for systems ≤ 512GB RAM and min 512 GB for all other systems
- Data volume: 1 x RAM
- HANA shared for scale-up systems: 1 x RAM
- HANA shared for scale-out systems: 1 x RAM of a worker node for each 4 nodes in the cluster

For the latest sizing requirements refer to the latest version of the SAP HANA Storage Requirements whitepaper:

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

7.1.4 Definitions of storage pools and volumes

The SAP HANA data and log volumes have different I/O characteristics, so it is recommended to create the data volumes and the the log volumes in different pools.

For all storage systems create at least two (2) volumes for SAP HANA data and at least two (2) volumes for SAP HANA log for each node.

7.1.5 Layout of storage pools and storage volumes

Each SAP HANA worker node is assigned a separate storage and the Storage Connector is in charge of the failover mechanism.

The shared area is mounted on all nodes (on the standby as well).

The following diagram shows a 2+1 node example and the relationship between the LUNs assigned from the storage and the Linux LVM components: physical volumes (PV), volume groups (VG) and logical volumes (LV) alongside with the shared NFS export.

In this example the /hana/shared storage is assumed to be hosted in the V Series Storage and exported to the hosts via an HA-NFS server, because the Lenovo V Series Storage does not have NFS capability.

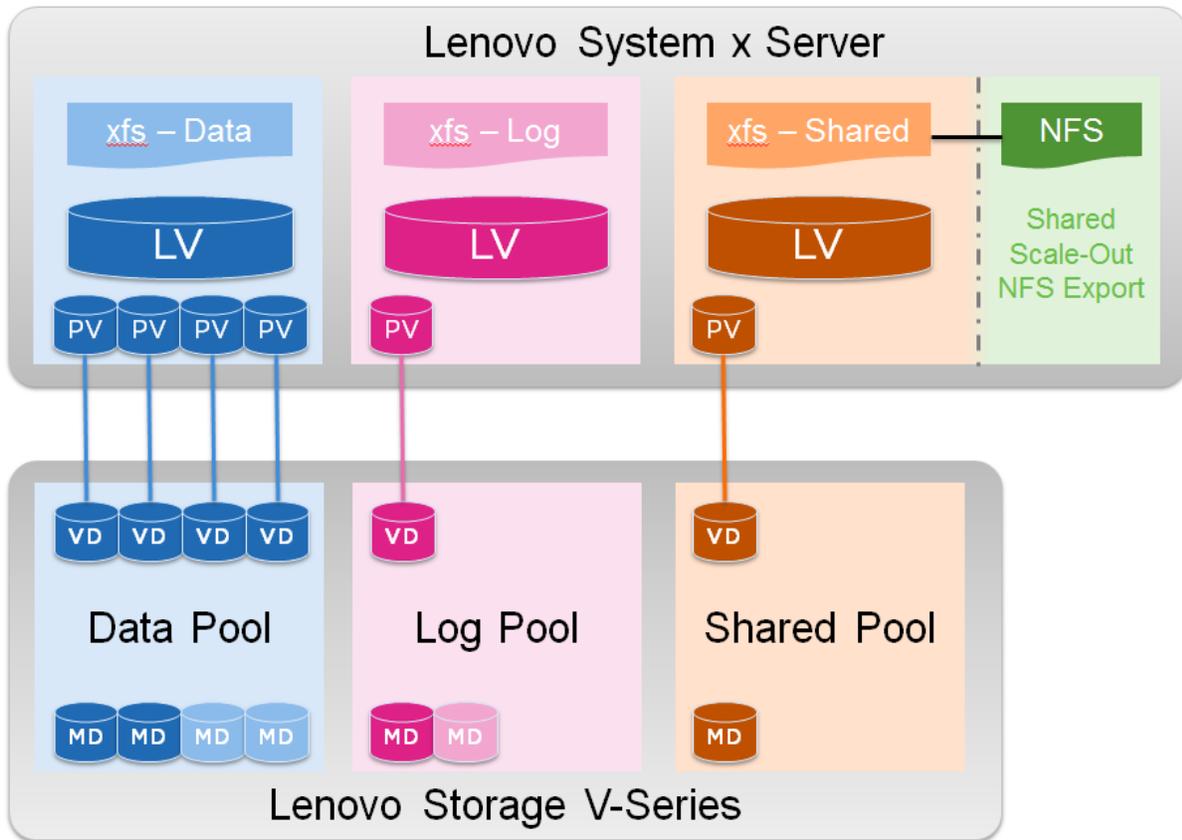


Figure 4: Layout of storage pools and volumes

7.1.6 Competing storage utilization

Every Lenovo Storage system can be shared between SAP HANA (production) and any other SAP or non-SAP (production) workload. To do proper sizing, the entire storage infrastructure needs to be evaluated, including the configurations of RAID controller, number of SSD, number and type of HBA and FC ports/links. Lenovo and Lenovo Business Partners are pleased to provide support for this sizing task.

7.1.7 The use of Lenovo storage compression (RtC)

The Lenovo storage systems provide the option to compress data on VDISK / Volume level. Because the data of the SAP HANA database are “just” integers, these data can not be compressed much. RtC can be used, but the saving might be not high.

7.2 Systems management

To manage the hardware systems, you may use the Lenovo XClarity™ Administrator, which is a centralized resource management solution that reduces complexity, speeds up response, and enhances the availability of Lenovo® server systems and solutions.

The Lenovo XClarity Administrator provides agent-free hardware management for Lenovo’s System x® rack servers. The figure below shows the Lenovo XClarity administrator interface, in which rack servers are managed and are seen on the dashboard. Lenovo XClarity Administrator is a virtual appliance that is quickly

imported into a virtualized environment server configuration.

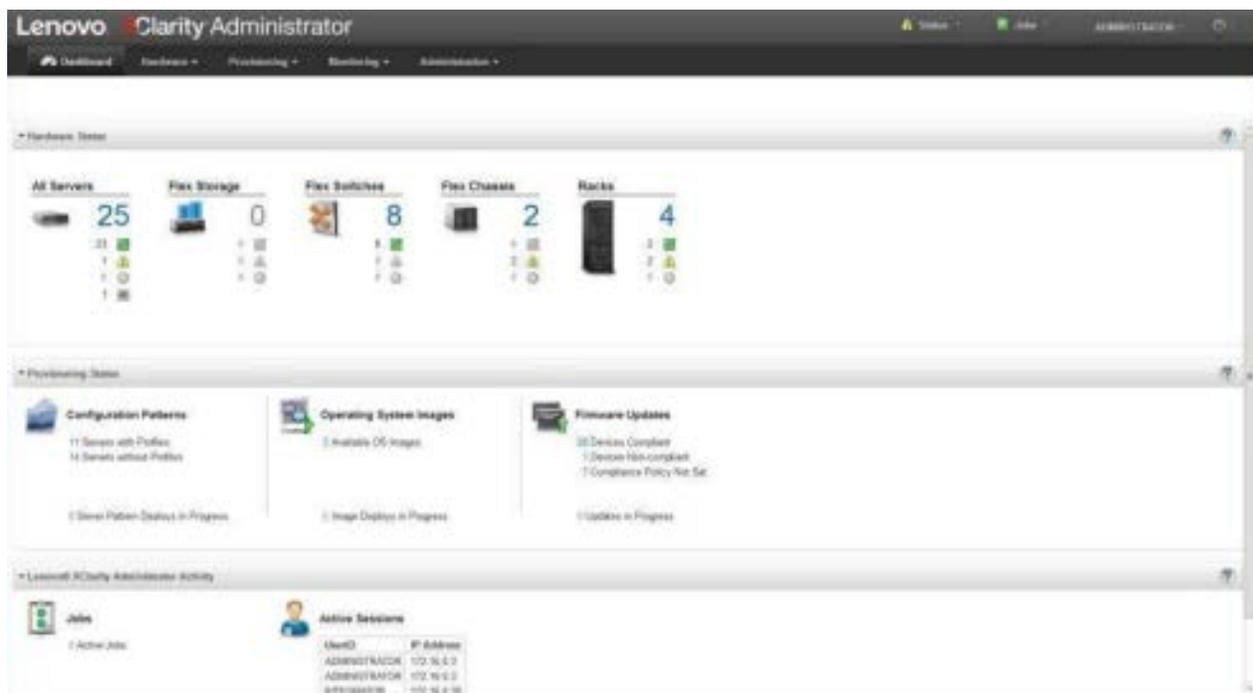


Figure 5: XClarity Administrator interface

7.3 Networking

This optional section is used to describe the network topology integration into the operational model.

Two 10G ports are used for SAP HANA internal communication, one 1G port is used for management communication to other servers.

Bonding setup is used for redundancy, improved performance and high availability.

7.4 Operating system setup & tuning

The operating system should be configured according to the SAP HANA Reference Guide for each supported Operating System and can be found under:

<https://wiki.scn.sap.com/wiki/display/SAPHANA/Linux+Operating+System+with+SAP+HANA+Reference+Guide>

Additional information can be found in the SAP Notes listed there.

7.5 Linux setup & tuning

This section provides guides about the required setup and tuning, additional information can be found in guide “SUSE Linux Enterprise Server 11.x (12x) for SAP Applications Configuration Guide for SAP HANA”:

<https://launchpad.support.sap.com/#/notes/1944799>

7.6 Multipath changes

The Lenovo V Series Storage systems comply with the SCSI-3 standard for Asymmetrical Logical Unit Access (ALUA). ALUA compliant storage systems will provide optimal and non-optimal path information to the host during device discovery, but the operating system must be directed to use ALUA by creating a separate device entry for the Lenovo V Series Storage system in the `/etc/multipath.conf`.

An example for the Lenovo V Series Storage stanza used for multipath configuration is shown below.
`/etc/multipath.conf`:

```
[...]  
    device {  
        vendor "IBM"  
        product "^2145"  
        features "1 queue_if_no_path"  
        path_grouping_policy "group_by_prio"  
        prio "alua"  
        path_checker "tur"  
        path_selector "service-time 0"  
        failback "immediate"  
        rr_weight "uniform"  
        no_path_retry "fail"  
        rr_min_io_rq 10  
        dev_loss_tmo 600  
        fast_io_fail_tmo 5  
    }  
[...]
```

Note: For All Flash configurations set the `rr_weight` parameter to "priorities".

An example of a `multipath -ll` output on a server with Lenovo Storage V3700 V2 Storage:

```
3600507638080808a700000000000012 dm-1 IBM,2145  
size=256G features='1 retain_attached_hw_handler' hwhandler='1 alua' wp=rw  
|+- policy='service-time 0' prio=50 status=active  
| |- 1:0:0:1 sdb 8:16 active ready running  
| |- 1:0:3:1 sdac 65:192 active ready running  
| |- 2:0:2:1 sdbd 67:112 active ready running  
| `-- 2:0:3:1 sdbm 68:0 active ready running  
`+- policy='service-time 0' prio=10 status=enabled  
  |- 1:0:1:1 sdk 8:160 active ready running  
  |- 1:0:2:1 sdt 65:48 active ready running  
  |- 2:0:0:1 sdal 66:80 active ready running  
  `-- 2:0:1:1 sdau 66:224 active ready running
```

The Storage Connector supports two different persistent reservation types:

- Exclusive Access (--prout-type=6): blocks any external read or write access on a device
- Write Exclusive (--prout-type=5): allows reading from, but not writing to a device

When using fcClientLVM, reservation type 5 is mandatory, because the LUNs must be visible to the LVM in order to read metadata from it. The fcClientLVM does not work if the metadata is not accessible.

Additional information is available in SAP Note 1941776: <https://launchpad.support.sap.com/#/notes/1941776>

For using the SAP HANA fcClientLVM SAP recommends the following settings to avoid hangs in the takeover:

- no_path_retry = fail
We recommend a retry setting of "fail" (or 0). This causes the resources to fail over when the connection to the storage is lost. Otherwise, the messages queue and the resource failover cannot occur.
- fast_io_fail_tmo = 5
The fast_io_fail_tmo parameter sets the length of time to wait before failing I/O when a link problem is detected. I/O fails that want to reach the driver.
- path_selector: service-time 0
A service-time oriented load balancer that balances I/O on paths according to the latency. This optimizes the SAP HANA Log I/O which is latency sensitive (as stated by Novell).

The following `lvm.conf` modifications according to the SAP HANA Fibre Channel Storage Connector Admin Guide:

- use_lvmetad: Activation state of the LVM Metadata Daemon. If available in the utilized Linux OS it has to be disabled as it doesn't support the Type 3 locking type.
- locking_type: Type of locking to use. Type 3 uses built-in clustered locking which is mandatory for HANA in combination with LVM

`/etc/lvm/lvm.conf`

The `multipath.conf` must be identical on all nodes belonging to the SAP HANA instance.

```
use_lvmetad = 0
locking_type = 3
```

7.6.1 Using aliases for UUIDs in `multipath.conf`

The `alias` setting overrides the `user_friendly_names` setting. Using alias makes managing and identifying volumes easier. If you want to use alias, please make sure that in the case of a multi-node environment, you fully understand the requirement and handling of consistent multipath device names across all nodes: The `multipath.conf` must be identical on all nodes belonging to the SAP HANA instance.

```
multipaths {
  multipath {
    wwid 3600507638080808a700000000000012
    alias HANA_LSZ_DATA_1
  }
  multipath {
    wwid 3600507638080808a700000000000013
    alias HANA_LSZ_LOG_1
  }
  rr_weight "priorities"
  no_path_retry "fail"
  rr_min_io_rq 10
  dev_loss_tmo 600
  fast_io_fail_tmo 5
}
```

7.7 IO device tuning

If I/O performance issues have been detected, the following parameters should be considered to tune. Do not change these in advance. Consult a performance specialist to find a vital combination:

- Increase `/sys/block/<device>/queue/nr_requests` if the default (128) results in blocked I/O submission. This will indirectly help to optimize the blocking inside SAP HANA.
- Increase `'rr_min_io_rq'` to `'32'` in `multipath.conf`
- Increase the queue depth of devices, e.g.

```
echo 64 > cat /sys/bus/scsi/devices/<device>/queue_depth
```

7.8 IO scheduler

The use of the NOOP scheduler for SAP HANA is recommended and it was changed accordingly.

The default scheduler `cfq` delivers low IO performance for multiple (parallel) streaming read processes. The NOOP scheduler is recommended for setups with devices that do I/O scheduling themselves, such as intelligent storage or in multipathing environments. The DEADLINE scheduler can provide a superior throughput over the CFQ I/O scheduler in cases where several threads read and write and fairness is not an issue. For example, for several parallel readers from a SAN.

The NOOP scheduler performs only minimal merging functions on your data. There is no sorting, and

therefore, this scheduler has minimal overhead. This scheduler was developed for non-disk-based block devices, such as memory devices, SSD. It also does well on storage media that have extensive caching. In some cases, it can be helpful for devices that do I/O scheduling themselves, as intelligent storage, or devices that do not depend on mechanical movement, like SSDs as NOOP scheduler has less overhead it may produce better performance for certain workloads.

More information can be found in the SLES12 System Analysis and Tuning Guide:

https://www.suse.com/documentation/sles-12/singlehtml/book_sle_tuning/book_sle_tuning.html#sec.tuning.io.schedulers.noop

7.9 Setup of shared access for the `/hana/shared/` directory

There are two fundamentally different storage configurations for a Multiple-Host System: shared storage devices or separate storage devices with failover reassignment. Do not confuse "shared storage" with the installation directory `/hana/shared` that must be shared across all hosts. In this document we will only cover the separate storage devices with failover reassignment scenario where a SAN storage must be used in combination with the SAP Fibre Channel Storage Connector, which SAP HANA offers storage technology vendors. During failover, SAP HANA uses the storage connector API to tell the storage device driver to re-mount the required data and logs volumes to the standby host and fence off the same volumes from the failed host.

This scenario requires that the directory `/hana/shared/` is accessible from all nodes.

The shared access can be achieved most easily through a highly available NFS server like the Highly Available NFS Storage with DRBD and Pacemaker: https://www.suse.com/documentation/sle-ha-12/singlehtml/book_sleha_techguides/book_sleha_techguides.html

Setup these systems according to the appropriate implementation guides, and make the NFS file system available on all SAP HANA nodes, including the standby nodes.

Basic setup of NFS Server, add this line to `/etc/exports`:

```
/nfs node1(rw,sync,no_root_squash,no_subtree_check)
node2(rw,sync,no_root_squash,no_subtree_check) node3(...)
```

Basic setup of the SAP HANA nodes (NFS clients), add this line to `/etc/fstab`:

```
nfs-server:/nfs /hana/shared nfs rw,soft,intr,rsize=8192,wsiz=8192 0 0
```

7.10 SAN zoning and volume mapping

Apply standard, best practices zoning within the SAN.

If HANA High Availability node-fail over or HANA scale-out is use, then, and only then, it is required to map all Data and Log volumes to all SAP HANA Nodes.

7.11 Setup of file systems for DATA and LOG

Create the DATA and LOG file systems needed by SAP HANA according to the SAP HANA Storage Requirements documentation from SAP.

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

SAP does neither force to use any specific file system nor has requirements for its configuration. Practical experience showed that XFS is used in general.

- We recommend the use of the **xfs** file system type with a 4KB block size for the DATA and LOG file systems
- If you are using LVM 2:
vgcreate → lvcreate → mkfs.xfs
- Use useful names for volume groups and logical volumes
 - Do not use the hyphen (minus) sign '-' as part of the name, because the Linux device mapper uses this symbol as a separator between the Volume Groups and Logical Volumes names.
- Create Volume Groups and Logical Volumes for DATA and LOG, for each SAP HANA mount point (as listed in global.ini) with exactly one Logical Volume.

DATA:

1. Use at least 2 physical volumes for the Volume Group
2. Create a volume group:
`vgcreate hanadata1vg /dev/dm-1 /dev/dm-2`
3. Create a logical volume:
`lvcreate -l131070 -n hanadata1lv hanadata1vg`
4. Create the file system
`mkfs.xfs /dev/mapper/hanadata1vg-hanadata1lv`
5. Do not mount the file system and do not add any entry to the `/etc/fstab`, mounting will be done by SAP HANA **in case if scale-out or HA node fail-over is used.**

LOG:

1. Use at least 2 physical volumes for the Volume Group
2. Create a volume group:
`vgcreate hanalog1vg /dev/dm-3 /dev/dm-4`
3. Create a logical volume:
`lvcreate -l131070 -n hanalog1lv hanalog1vg`
4. Create the file system
`mkfs.xfs /dev/mapper/hanalog1vg-hanalog1lv`
5. Do not mount the file system and do not add any entry to the `/etc/fstab`, mounting will be done by SAP HANA in case of scale-out or HA node fail-over is used.

SHARED:

In case of scale-out: NFS, non-failover

1. Just add an entry to `/etc/fstab` for the NFS file system to be automatically mounted

7.12 VMware Virtual Volumes

Using VMware Virtual Volumes (VVols) requires the same setup like for a native installation on physical drives. The only deviation is that VMware™ (ESX® server) controls the SAN multipath, not the Linux operating system. The volumes are typically accessible under the path: `/dev/sd...` instead of `/dev/mapper/...`

VMware vSphere Virtual Volumes (VVol) integration brings advanced storage management functions to the virtual machine level for finer VM storage service level control and policy-driven automation. The VMware VVol integration feature requires Spectrum Control Base Edition for Lenovo that can be downloaded from <http://datacentersupport.lenovo.com/us/en/products/storage/lenovo-storage/v3700v2/6535/downloads>

7.13 Setup of `/hana/shared/global.ini`

A dedicated `global.ini` file (located in `/hana/shared`) is only needed if you want to use the node fail-over functionality provided by SAP HANA software and for SAP HANA scale-out; else the SAP HANA installer will create a `global.ini` file. Enter the Logical Volume names as shown in directory `/dev/mapper` into `/hana/shared/global.ini`:

```
[persistence]
basepath_datavolumes = /hana/data/LSZ
basepath_logvolumes = /hana/log/LSZ
use_mountpoints = yes
[storage]
ha_provider = hdb_ha.fcClientLVM
partition_1_data__lvmname = hanadata1vg-hanadata1lv
partition_1_log__lvmname = hanalog1vg-hanalog1lv
partition_2_data__lvmname = hanadata2vg-hanadata2lv
partition_2_log__lvmname = hanalog2vg-hanalog2lv
partition_*_*__prtype = 5
partition_*_*__mountoptions = -t xfs
```

7.14 `hdbparam fileio` parameter

The SAP HANA database parameters used for the Lenovo V Series Family Storage Systems are now default in SAP HANA version 2:

```
async_write_submit_active=on
async_write_submit_blocks=all
async_read_submit=on
max_parallel_io_requests=64
size_kernel_io_queue=512
```

For SAP HANA version 1, follow the instructions in SAP Note 1930979 to set these parameters via the

hdbparam tool.

If you need to change these parameters in SAP HANA version 2, follow the instructions in SAP Note 2399079
– Elimination of hdbparam in HANA 2 to set these parameters via SAP HANA Studio.

7.15 Other considerations

Backup, High Availability and Disaster Recovery is not discussed in this document.

8 Appendix: Lenovo Bill of materials

This appendix contains the bill of materials (BOMs) for different configurations of hardware for [SAP HANA TDI](#) deployments. There are sections for user servers, management servers, storage, and networking.

The BOM lists in this appendix are not meant to be exhaustive and must always be double-checked with the configuration tools. Any discussion of pricing, support, and maintenance options is outside the scope of this document.

Within a specific BOM section, optional items are numbered with alternatives shown as lower-case letters. For example, a Fibre Channel adapter for a compute server is only needed for shared storage connected through a SAN.

Following is the BOM used in our testing, you can choose between the V3700 V2, V3700 V2 XP and V5030 systems and depending on your SAN infrastructure, you can add one or two fibre channel switches and one or two Fibre Channel adapters per server.

Storage System:

Part #	Description	Quantity
6535-EC2	Lenovo Storage V3700 V2 SFF Dual Controller Enclosure	1
01DE347	Lenovo Storage V3700 V2 300GB 2.5" 15K HDD	24

SAN:

Part #	Description	Quantity
6559-HC2	Lenovo ThinkSystem DB610S 32Gb FC SAN Switch	2

Fibre Channel adapters:

Part #	Description	Quantity
ATZC	QLogic 16Gb Enhanced Gen5 FC Dual-port HBA	2
or		
ATZV	Emulex 16Gb Gen6 FC Dual-port HBA	2

Resources

SAP SE Resources

SAP Service Marketplace: <https://service.sap.com/hana>

SAP HANA Administration Guide: http://help.sap.com/hana/SAP_HANA_Administration_Guide_en.pdf

SAP HANA Hardware Directory:

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

SAP HANA Storage Requirements:

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

SAP HANA Fibre Channel Storage Connector Admin Guide:

<https://www.sap.com/documents/2016/06/84ea994f-767c-0010-82c7-eda71af511fa.html>

SAP HANA TDI FAQ:

<https://www.sap.com/documents/2016/05/e8705aae-717c-0010-82c7-eda71af511fa.html>

Lenovo Resources

Quick Installation Guide

https://download.lenovo.com/pccbbs/thinkservers/v3700v2_v5030_quickinstallguide_eng.pdf

Lenovo Storage V3700 V2/V5030 Series Information Center

http://systemx.lenovofiles.com/help/index.jsp?topic=%2Fcom.lenovo.storage.common.nav.doc%2Foverview_storage_vseries.html

Implementing the Lenovo Storage V3700 V2 and V5030 Systems

<https://lenovopress.com/lp0836-implementing-the-lenovo-storage-v3700-v2-and-v5030-systems>

SAP Notes

[1943937](#) - Hardware Configuration Check Tool - Central Note

[171356](#) - SAP Software on Linux: General information

[2205917](#) - SAP HANA DB: Recommended OS Settings for SLES 12 / SLES for Applications 12

[2292690](#) - SAP HANA DB Recommended OS Settings for RHEL 7.2

[2031375](#) - SAP HANA: Transparent HugePages (THP) setting on Linux

[1900823](#) - SAP HANA Storage Connector API

[1941776](#) - Nameserver in distributed environment hangs at startup

Document history

Version 1.0 July 2018

Initial version includes Lenovo Storage V3700 V2, V3700 V2 XP and V5030 Series SAN products.

Trademarks and special notices

© Copyright Lenovo 2018.

References in this document to Lenovo products or services do not imply that Lenovo intends to make them available in every country.

Lenovo, the Lenovo logo, ThinkCentre, ThinkVision, ThinkVantage, ThinkPlus and Rescue and Recovery are trademarks of Lenovo.

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both.

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

Intel, Intel Inside (logos), MMX, and Pentium are trademarks of Intel Corporation in the United States, other countries, or both.

SAP and other SAP products and services mentioned herein as well as their respective logos 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.

Information is provided "AS IS" without warranty of any kind.

All customer examples described are presented as illustrations of how those customers have used Lenovo products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.

Information concerning non-Lenovo products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by Lenovo. Sources for non-Lenovo list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. Lenovo has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-Lenovo products. Questions on the capability of non-Lenovo products should be addressed to the supplier of those products.

All statements regarding Lenovo future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Contact your local Lenovo office or Lenovo authorized reseller for the full text of the specific Statement of Direction.

Some information addresses anticipated future capabilities. Such information is not intended as a definitive statement of a commitment to specific levels of performance, function or delivery schedules with respect to any future products. Such commitments are only made in Lenovo product announcements. The information is presented here to communicate Lenovo's current investment and development activities as a good faith effort to help with our customers' future planning.

Performance is based on measurements and projections using standard Lenovo benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

Photographs shown are of engineering prototypes. Changes may be incorporated in production models.

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