



TECHNICAL WHITE PAPER

BUILDING A LENOVO SAP HANA TDI STACK

WITH PURE STORAGE FLASHARRAY

TABLE OF CONTENTS

- EXECUTIVE SUMMARY** 3
- PURPOSE** 3
- AUDIENCE** 3
- LENOVO SERVER ELEMENTS FOR SAP HANA** 4
 - Lenovo ThinkSystem SR950 4
 - Lenovo ThinkSystem SR650 5
 - Lenovo System x3850/x3950 X6 5
- PURE STORAGE FLASHARRAY FOR SAP HANA TDI ARCHITECTURES** 6
 - Using Pure Storage FlashArray for SAP HANA TDI 6
 - Running SAP HANA and Non-HANA Mixed Workloads 9
 - Data Reduction on Compressed Databases 9
 - Data Copy Management Using Pure Storage FlashRecover Storage Snapshots 9
 - Disaster Tolerance Using Purity ActiveCluster Synchronous Replication 10
- SETUP AND INSTALLATION** 10
 - Compute Node Details for SAP HANA 10
 - FibreChannel Connectivity 11
 - Linux Storage Multi-Pathing Configuration 12
 - Server File System Preparation for SAP HANA 16
 - Implementing the SAP HANA Storage Connector API 17
 - Installing the SAP HANA Database 18
- SIZING PURE STORAGE FLASH ARRAY FOR SAP HANA TDI** 19
 - Sizing the SAP HANA TDI Workload 20
 - Sizing the Non-HANA TDI Workload 21
- CONCLUSION** 22
- RESOURCES** 22

EXECUTIVE SUMMARY

This white paper describes the advantages of using Lenovo servers and Pure Storage FlashArray enterprise storage in an SAP HANA Tailored Datacenter Integration (TDI) environment. Included are best practices and configuration guidelines for setting up and using Pure Storage FlashArray and Lenovo Server with XFS file systems. Our guidelines meet or exceed all SAP performance requirements and ensure the highest availability for database persistence on flash media. This solution is fully supported by all the vendors that deliver components to it. Throughout this document, we highlight key aspects of the simplicity, ease-of-use, efficiency, and reliability of the solution's compute and storage platforms.

PURPOSE

With this document, we aim to detail how we have implemented a performing SAP HANA TDI-based compute and storage solution leveraging Lenovo Server and Pure Storage FlashArray.

We also explain why SAP HANA TDI certified solution components are a great way to determine which combination of hardware and storage delivers the most value for your business.

AUDIENCE

This whitepaper is intended for Lenovo and Pure Storage system integrators, systems and storage administrators, partners, customers, members of Lenovo lab services, and Pure Storage professional services involved in sizing or implementing the infrastructure for an SAP HANA TDI solution.

LENOVO SERVER ELEMENTS FOR SAP HANA

To meet customer needs for SAP HANA TDI, Lenovo offers scalable server solutions that support ever-growing business requirements.

Large memory configurations are available for SAP Business Suite, powered by SAP HANA, and for SAP Business Warehouse and data mart implementations. Users can scale up a single server very easily just by installing additional components into the server.

Users can also create multi-node, scale-out configurations by networking multiple single nodes together, enabling support for larger SAP HANA memory sizes. Lenovo has certified configurations of up to 94 server nodes in a single scale-out cluster.

This modular approach enables a customer to invest in an optimized solution for SAP HANA and grow the infrastructure as the SAP HANA environment grows.

Lenovo offers the following servers for use with SAP HANA:

- ThinkSystem SR950 (two, four, and eight-socket systems based on the latest Intel Skylake technology)
- ThinkSystem SR650 (two-socket system based on the latest Intel Skylake technology)
- System x3850/x3950 X6 (two, four, and eight-socket systems based on Intel Broadwell technology)

For SAP BusinessOne and other SAP HANA entry-level environments, there are additional Lenovo systems available. The following sections briefly introduce each of the aforementioned servers.

Lenovo ThinkSystem SR950

The Lenovo ThinkSystem SR950 is designed for demanding, mission-critical workloads. It is a 4U system that can grow from two to eight Intel Xeon Scalable Family processors, and with 96 DIMM sockets, it supports up to 12 TB of main memory using 128 GB DIMM modules.

Scaling up from a two-socket server works simply by adding a compute tray and system boards within the existing 4U enclosure. After initial installation into a rack, there is no need to pull the server out again. The SR950 is based on a modular service model where all components and options can be removed from the front or rear of the system, even parts that are located in the center of the machine such as fans, memory DIMMs, and processors.



FIGURE 1. Lenovo ThinkSystem SR950

The SR950 packs numerous fault-tolerant and high-availability features into a high-density, 4U rack-optimized design. Detailed technical information is available online in the [Lenovo Press SR950 Product Guide](#).

Lenovo ThinkSystem SR650

The Lenovo ThinkSystem SR650 is a 2-socket 2U rack server for smaller SAP environments such as test, development, or small production environments. It can be equipped with up to two of the latest generation Intel Xeon Scalable family processors and up to 3 TB of main memory in 24 DIMM slots.



FIGURE 2. Lenovo ThinkSystem SR650

Detailed technical information is available online in the [Lenovo Press SR650 Product Guide](#).

Lenovo System x3850/x3950 X6

The Lenovo System x3850 X6 server is a 4U rack-mounted server that can hold two or four Intel Xeon E7 v4 processors (codenamed Intel Broadwell). The Lenovo System x3950 X6 server is an 8U server that can hold up to eight processors from the Intel Xeon E7 v4 family. Both servers, the x3850 X6 and x3950 X6, implement the sixth generation of the Enterprise X-Architecture and can be equipped with up to 24 TB of main memory in the maximum eight-socket configuration.



FIGURE 3. System x3850 X6

With Intel Xeon E7 v4 technology, the x3850 X6 delivers fast application performance, an agile system design, and a resilient platform for mission-critical databases, enterprise applications, and virtualized environments.

Detailed technical information is available online in the [Lenovo Press System x3850 X6 Product Guide](#).

PURE STORAGE FLASHARRAY FOR SAP HANA TDI ARCHITECTURES

Pure Storage FlashArray enables you to transform your data center, cloud, or entire business with an affordable all-flash array capable of consolidating and accelerating all your key SAP applications.

FlashArray performance can make your business smarter by unleashing the power of real-time analytics, driving customer loyalty, and creating new, innovative customer experiences that simply weren't possible with disk.

With Pure Storage FlashArray, organizations can dramatically reduce the complexity of storage to make IT more agile and efficient, accelerating your digital transformation to SAP HANA TDI.

This section provides information on Pure Storage enterprise storage configurations.

Using Pure Storage FlashArray for SAP HANA TDI

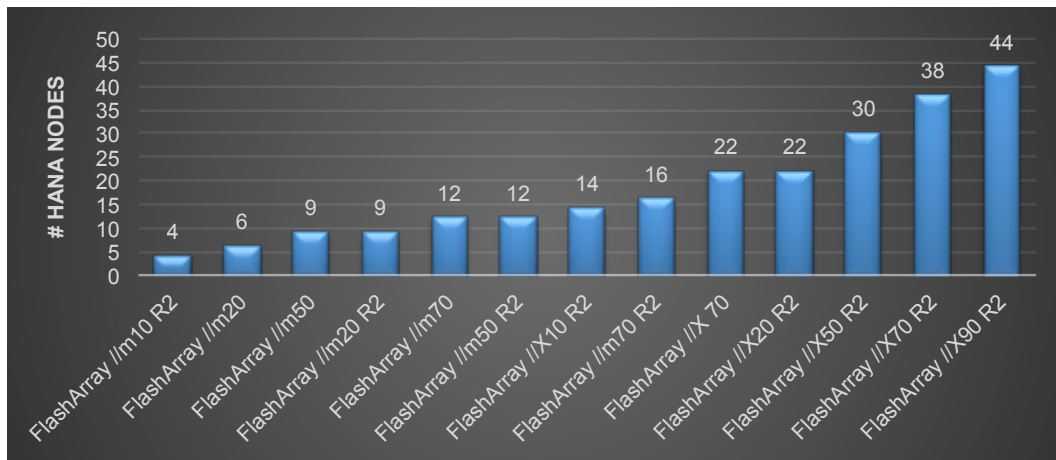
Pure Storage offers various options to set up an SAP HANA TDI landscape. Choose a FlashArray model based on the expected requirements of the custom Tailored Datacenter Integration (TDI) solution.

Effectively protect your data with integrated high availability and synchronous or asynchronous disaster recovery capabilities using Pure Storage ActiveCluster storage replication options.

Additionally, VMware virtualization technology can be used for optimizing deployment and management of your SAP HANA TDI application environment using virtual machines via a single SAP HANA system.

The following table shows FlashArray models and the estimated maximum number of HANA worker nodes that can be connected according to each model:

LICENSING LEVEL	MANAGEABILITY
FLASHARRAY //M20	SERVES UP TO 6 SAP HANA NODES ON A //M20
FLASHARRAY //M50	SERVES UP TO 9 NODES ON A //M50
FLASHARRAY //M70	SERVES UP TO 12 NODES ON A //M70
FLASHARRAY//M10 R2	SERVES UP TO 4 NODES PER //M10 R2
FLASHARRAY//M20 R2	SERVES UP TO 9 HANA NODES PER //M20 R2
FLASHARRAY//M50 R2	SERVES UP TO 12 NODES PER //M50 R2
FLASHARRAY//M70 R2	SERVES UP TO 16 HANA NODES PER //M70 R2
FLASHARRAY //X 70	SERVES UP TO 22 NODES PER //X 70
FLASHARRAY //X10 R2	SERVES UP TO 14 NODES PER //X10 R2
FLASHARRAY //X20 R2	SERVES UP TO 22 NODES PER //X20 R2
FLASHARRAY //X50 R2	SERVES UP TO 30 NODES PER //X50 R2
FLASHARRAY //X70 R2	SERVES UP TO 38 NODES PER //X70 R2
FLASHARRAY //X90 R2	SERVES UP TO 44 NODES PER //X90 R2



TABLES 1 & 2. FlashArray models and estimated max HANA worker nodes

USING PURE STORAGE FLASHARRAY//M

FlashArray//M is tested and validated with a wide range of data center infrastructure and applications to enable seamless deployment and multi-workload consolidation in your data center. The //M chassis leverages internal PCIe networking for controller HA, as well as both PCIe/NVMe and 16Gb SAS for flash and NV-RAM module connectivity. Dual-drive flash modules pack two independent SSDs in each slot, doubling potential performance and density. Drive modules are redundantly connected to both controllers, and can be replaced without performance impact.

USING PURE STORAGE FLASHARRAY//M R2

The FlashArray//M R2 expands upon FlashArray's modular, stateless architecture, which is designed to enable expandability and upgradability for generations. The FlashArray//M R2 leverages a chassis-based design with customizable modules, enabling both capacity and performance to be independently improved over time with advances in compute and flash to meet your business needs today and tomorrow.

USING PURE STORAGE FLASHARRAY//X

SAS provides a scalable enterprise storage solution, but it's a protocol that was developed for spinning disk media. NVMe was developed to take full advantage of all-flash media, providing much greater bandwidth with less IO queuing. Database workloads require high IO bandwidth and low latency, and will therefore typically perform better on NVMe media, which doesn't show performance degradation when overloaded with I/O requests. For SAP HANA database environments that use high I/O log throughput, FlashArray//X is the best choice, as the SAS bus may otherwise become a bottleneck.

USING PURE STORAGE FLASHARRAY//X R2

Pure's new FlashArray//X R2 family and the DirectFlash™ Shelf are based on a next-generation architecture with DirectFlash, NVMe, and NVMe-oF. Media technologies are constantly changing and allowing Pure to optimize for economics and performance (MLC, TLC, QLC, and emerging low-latency media). The new performance and scale configurations (//X50, //X70, and //X90) support existing SAS shelves and our traditional Flash Modules, along with

the new DirectFlash Shelf. The DirectFlash Shelf enables systems to achieve full capacity with 100% DirectFlash NVMe modules, for optimal system performance.

Running SAP HANA and Non-HANA Mixed Workloads

To prevent an SAP HANA workload from being impacted by a non-HANA workload, Purity Operating Environment (the software-defined heart of FlashArray) will allow prioritization of workloads, based on policies.

Storage QoS is a critical feature for both multi-tier consolidation of workloads and multi-tenancy. This feature autonomically protects system workloads against a single workload (usually called a noisy neighbor) consuming all system resources and impacting all tenants. This industry-first solution addresses noisy neighbor problems without having the user define any policies or rules.

Purity constantly monitors incoming I/O and is able to determine when the array is 'too busy' to deliver predictable, sub-ms performance; it automatically detects which volumes are consuming more than their fair share of available resources and throttles that workload to avoid performance degradation on all other workloads.

Purity QoS raises the bar when it comes to one of the most sophisticated features in storage:

- **Always-On QoS** enabled by default. For users who are consolidating workloads that need the same SLA, this functionality delivers fairness of service to all workloads.
- **Performance Classes** allows one-click functionality to assure protection of performance for mission critical apps
- **Performance Limits** enforces performance expectations for targeted workloads

Data Reduction on Compressed Databases

Thanks to the unmatched data reduction capabilities of FlashArray, SAP HANA database persistence is reduced by a factor of 1.6:1 to 2.2:1.

Note that this level of data reduction is achieved on top of the reduction delivered by SAP HANA database compressed tables.

With Purity's strong security features, data-at-rest is automatically encrypted without compromising performance by the storage array. However, when SAP HANA database encryption is used, storage compression benefits may be impacted.

Data Copy Management Using Pure Storage FlashRecover Storage Snapshots

Pure Storage FlashRecover snapshots are an intrinsic part of the Purity Operating Environment. They deliver superior space efficiency, high scalability, and simplicity of volume snapshot management.

Purity Operating Environment does not differentiate between volumes and snapshots. Snapshots are just volumes with static data and are programmatically made non-writable. Unlike snapshots in other storage arrays, FlashRecover snapshots are not owned by a volume, and hence can be used to restore any volume or to create new volumes.

FlashRecover snapshots are thin provisioned, with no dedicated space allocated. Purity data structures allow snapshots to preserve the granular data reduction efficiencies of volumes through global deduplication and compression, and thus volume snapshots require minimal physical capacity on flash drives.

Pure Storage leverages the robust and flexible design of FlashRecover snapshots to incorporate data replication and protection solutions such as fast cloning, application-offloaded workloads, and volume data protection.

Disaster Tolerance Using Purity ActiveCluster Synchronous Replication

With built-in Purity ActiveCluster, SAP HANA can now benefit from the simple-to-deploy, always-on business continuity of active/active storage replication, which provides transparent failover.

ActiveCluster uses a preferred array setting to ensure that hosts have the best possible performance by exposing optimized paths to hosts that prefer that array and non-optimized paths to hosts that do not prefer that array.

Deployments using a uniform storage access configuration have failover-less maintenance of operations: In the event of an array failure, or a replication link failure causing one array to stop I/O service, the hosts experience only the loss of some storage paths but continue using other paths to the available array. There is no storage failover process for the storage administrator to execute, and the workload simply keeps on running. When a loss of communication occurs, both arrays will briefly pause I/O and reach out to the Pure1® Cloud Mediator to determine which array can stay active for each sync-replicated consistency group.

SETUP AND INSTALLATION

This chapter describes the setup of a sample SAP HANA environment running on Lenovo servers with a Pure Storage array for production data and interconnected by a FibreChannel fabric using QLogic HBAs. This chapter can be used as a reference to implement other environments that use Lenovo servers on Pure Storage arrays.

Compute Node Details for SAP HANA

In the example below, four x3850 X6 four-socket systems are used and prepared for an SAP HANA scale-out installation. The servers have the following details:

- Intel Xeon CPU E7- 8880 v4
- QLogic Corp. ISP8324-based 16Gb Fibre Channel to PCI Express HBA
- Intel Corporation I350 Gigabit Network Connection

- SUSE Linux Enterprise Server 12 SP1 for SAP applications, installed onto local storage (128GB LUN for root and 32GB LUN for swap file system)
- Using integrated DM-Multipathing

To achieve maximum performance and reliability for use with SAP HANA, several areas of the server need to be adjusted. Information like minimum firmware levels, required uEFI parameters, or OS-specific settings are documented in Lenovo information and SAP notes. To ensure you get the latest information, please contact Lenovo or Lenovo Professional Services for assistance.

Operating system disk setup has to follow official SAP guidelines. Details are in the [SAP storage requirements whitepaper](#).

In our case, the OS installation looks as follows:

```
se-x3850x6a:~ # lsblk -l
```

NAME	MAJ:MIN	RM	SIZE	RO	TYPE	MOUNTPOINT
sda	8:0	0	1.1T	0	disk	
sda1	8:1	0	148M	0	part	/boot/efi
sda2	8:2	0	128G	0	part	
sda3	8:3	0	32G	0	part	[SWAP]
sda4	8:4	0	150M	0	part	
sda5	8:5	0	128G	0	part	/boot/grub2/x86_64-efi
sda6	8:6	0	828.4G	0	part	/home

In a distributed SAP HANA scale-out environment, we must make sure all systems run the same patch level, are configured using the same time zone, with the same HWCLOCK settings and the same local system date and time settings on all scale-out nodes.

FibreChannel Connectivity

Each lab server has a dual-port QLogic 16 Gbit FibreChannel HBA installed that is based on the QLogic QLE2662 adapters. Each port is connected in a fully redundant fashion to two Lenovo B6505 FC switches.

```
se-x3850x6a:~ # lspci | grep "Fibre Channel"
```

51:00.0 **Fibre Channel**: QLogic Corp. ISP8324-based 16Gb Fibre Channel to PCI Express Adapter (rev 02)

51:00.1 **Fibre Channel**: QLogic Corp. ISP8324-based 16Gb Fibre Channel to PCI Express Adapter (rev 02)

```
se-x3850x6a:~ # lspci -nn | grep -i "fibre"
```

51:00.0 **Fibre Channel** [0c04]: QLogic Corp. ISP8324-based 16Gb **Fibre Channel** to PCI Express Adapter [1077:2031] (rev 02)

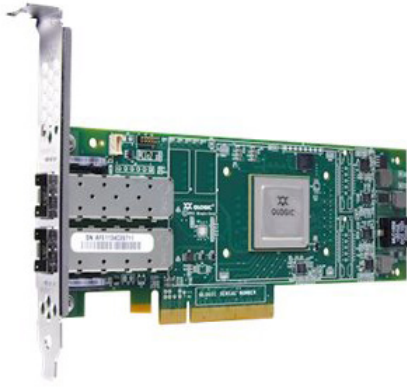


FIGURE 4. QLogic 16 Gbit FC HBA

Each QLogic HBA is running the following code levels with default port settings.

- BIOS version: 3.43
- EFI version: 6.24
- FW version: 8.03.00 7917781

Linux Storage Multi-Pathing Configuration

The redundant paths to the storage array need to be combined by Linux using multi-pathing tools. By default, Linux can talk to each LUN across every path, as can be seen in the example below for the two HBA ports (SCSI address starting with 1 and 4 at the beginning of each line).

```
se-x3850x6a:~ # lsscsi
[0:2:0:0]   disk    IBM      ServerAID M5210  4.29  /dev/sda
[1:0:0:1]   disk    PURE     FlashArray        8888  /dev/sdd
[1:0:0:2]   disk    PURE     FlashArray        8888  /dev/sde
[1:0:0:3]   disk    PURE     FlashArray        8888  /dev/sdc
[1:0:1:1]   disk    PURE     FlashArray        8888  /dev/sdg
[1:0:1:2]   disk    PURE     FlashArray        8888  /dev/sdh
[1:0:1:3]   disk    PURE     FlashArray        8888  /dev/sdf
[4:0:0:1]   disk    PURE     FlashArray        8888  /dev/sdj
[4:0:0:2]   disk    PURE     FlashArray        8888  /dev/sdk
[4:0:0:3]   disk    PURE     FlashArray        8888  /dev/sdi
[4:0:1:1]   disk    PURE     FlashArray        8888  /dev/sdm
[4:0:1:2]   disk    PURE     FlashArray        8888  /dev/sdn
[4:0:1:3]   disk    PURE     FlashArray        8888  /dev/sdl
```

In an SAP HANA scale-out configuration with a local HA server (that is Host Auto-Failover), make sure the **multipath.conf** settings are adjusted accordingly. The SAP HANA Storage Connector API 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

Use **no_path_retry=0** to force a failover, but this may also mean that if the standby server is unavailable, failover may be prevented with a ping-pong effect.

Create a **multipath.conf** file, which, if needed, contains the required configuration settings for the storage API connector:

```
se-x3850x6a:~ # more /etc/multipath.conf

defaults {
    polling_interval      10
    user_friendly_names  no
    fast_io_fail_tmo     10
    dev_loss_tmo         60
}

devices {

    device {
        vendor            "PURE"
        path_selector     "queue-length 0"
        hardware_handler  "1 alua"
        path_grouping_policy group_by_prio
        prio              alua
        failback          immediate
        path_checker      tur
        fast_io_fail_tmo  10
        user_friendly_names no
        no_path_retry    0
        rr_min_io_rq      1
        features        0
        dev_loss_tmo      60
    }
}
}
```

Start the multi-pathing daemon, and enable it so that it is automatically started upon system boot:

```
se-x3850x6a:~ # systemctl enable multipathd
se-x3850x6a:~ # systemctl start multipathd
```

Now the status of the service should be active and running:

```
se-x3850x6a:~ # systemctl status multipathd
multipathd.service - Device-Mapper Multipath Device Controller
   Loaded: loaded (/usr/lib/systemd/system/multipathd.service; enabled)
   Active: active (running)
 Main PID: 8723 (multipathd)
    Status: "running"
   CGroup: /system.slice/multipathd.service
           └─8723 /sbin/multipathd -d -s
```

The output of the multipath daemon is show below. We can see that a fully redundant FC fabric with two independent FC switches is connected on both HBA ports of the server.

```
se-x3850x6a:~ # multipath -ll
3624a937073a0f55f361648b300132027 dm-3 PURE,FlashArray
size=500G features='0' hwhandler='0' wp=rw
|+- policy='service-time 0' prio=1 status=active
| `- 4:0:0:2 sdk 8:160 active ready running
|+- policy='service-time 0' prio=1 status=enabled
| `- 1:0:0:2 sde 8:64 active ready running
|+- policy='service-time 0' prio=1 status=enabled
| `- 1:0:1:2 sdh 8:112 active ready running
`+- policy='service-time 0' prio=1 status=enabled
  `- 4:0:1:2 sdn 8:208 active ready running
3624a937073a0f55f361648b300123023 dm-2 PURE,FlashArray
size=500G features='0' hwhandler='0' wp=rw
|+- policy='service-time 0' prio=1 status=active
| `- 4:0:0:1 sdj 8:144 active ready running
|+- policy='service-time 0' prio=1 status=enabled
| `- 1:0:0:1 sdd 8:48 active ready running
|+- policy='service-time 0' prio=1 status=enabled
| `- 1:0:1:1 sdg 8:96 active ready running
`+- policy='service-time 0' prio=1 status=enabled
  `- 4:0:1:1 sdm 8:192 active ready running
```

To set the correct Linux I/O subsystem parameters for the LUNs, create a file called `/etc/udev/rules.d/10-pure.rules` with the following entry:

```
se-x3850x6a:~ # more /etc/udev/rules.d/10-pure.rules
ACTION=="add|change", KERNEL=="dm-*", \ ATTR{queue/nr_requests}="4096", ATTR{queue/scheduler}="noop"
```

A system reboot is required to activate the changes.

For reference, the following lists the values of essential Linux IO settings. **DEV** is the name of the SAP HANA data and log volumes:

```
# find /sys/block/DEV/queue -name "max_sectors_kb" -exec grep -H . '{}' \;
# find /sys/block/DEV/queue -name "read_ahead_kb" -exec grep -H . '{}' \;
# find /sys/block/DEV/queue -name "max_hw_sectors_kb" -exec grep -H . '{}' \;
# find /sys/block/DEV/queue -name "nr_requests" -exec grep -H . '{}' \;
```

Example using **sda** as the block device to query:

```
se-x3850x6a:~ # find /sys/block/sda/queue -name "max_sectors_kb" -exec grep -H . '{}' \;
/sys/block/sda/queue/max_sectors_kb:256
```

```
se-x3850x6a:~ # find /sys/block/sda/queue -name "read_ahead_kb" -exec grep -H . '{}' \;
/sys/block/sda/queue/read_ahead_kb:512
```

```
se-x3850x6a:~ # find /sys/block/sda/queue -name "max_hw_sectors_kb" -exec grep -H . '{}' \;
/sys/block/sda/queue/max_hw_sectors_kb:256
```

```
se-x3850x6a:~ # find /sys/block/sda/queue -name "nr_requests" -exec grep -H . '{}' \;
/sys/block/sda/queue/nr_requests:4096
```

To find the queue depth setting, run the following command:

```
se-x3850x6a:~ # more /sys/module/qla2xxx/parameters/ql2xmaxqdepth
32
```

Parameter tuning settings are hardware configuration dependent. Please contact your Pure Storage support person to discuss the best possible configuration settings apart from defaults.

Server File System Preparation for SAP HANA

The storage partitions mounted to the host consist of plain LUNs, not LVM volumes or volume groups.

Create the HANA data and log directories on each host (**hananode01**, **hananode02**, **hananode03**, **hananode04**), where the volumes will be mounted:

```
se-x3850x6a:~ # mkdir /hana/data
se-x3850x6a:~ # mkdir /hana/log
se-x3850x6a:~ # mkdir /hana/shared
```

Allow full permissions for HANA data and log directories with their subdirectories:

```
se-x3850x6a:~ # chmod 777 /hana/data
se-x3850x6a:~ # chmod 777 /hana/log
se-x3850x6a:~ # chmod 777 /hana/shared
```

In a scale-out setup, all data and log volumes are presented to all active worker nodes except the standby node. The volumes are automatically mounted by the Storage Connector API, not by the operating system. **In a scale-out setup, do not add an entry in the /etc/fstab file.**

After presenting the LUNs to the hosts, on each host, rescan for new devices:

```
se-x3850x6a:~ # /usr/bin/rescan-scsi-bus.sh
```

Create file systems of type **XFS** on each data and log volume on each of the hosts. Use default XFS settings.

```
se-x3850x6a:~ # ll -l /dev/mapper
total 0
lrwxrwxrwx 1 root root      7 Mar 13 16:42 3624a937073a0f55f361648b300123023 -> ../dm-2
lrwxrwxrwx 1 root root      7 Mar 13 16:42 3624a937073a0f55f361648b300132027 -> ../dm-3
se-x3850x6b:~ # mkfs -t xfs -f /dev/mapper/3624a937073a0f55f361648b300123023
se-x3850x6a:~ # mkfs -t xfs -f /dev/mapper/3624a937073a0f55f361648b300132027
```

Mount the SAP HANA shared NFS volume as **NFSv3** or **NFSv4** on each of the cluster nodes. Verify mounting the HANA shared NFS volume using the following command:

```
se-x3850x6a:~ # mount | grep shared
192.168.27.111:/home/nfsforpurehana on /hana/shared type nfs4 (rw,relatime,vers=4.0,rsize=1048576,
wsize=1048576,namlen=255,hard,proto=tcp,port=0,timeo=600,retrans=2,sec=sys,clientaddr=192.168.27.102,
local_lock=none,addr=192.168.27.111)
```


Modify `/etc/fstab` to include the **NFSv3** shared volume mount point. Verify with the following command:

```
se-x3850x6a:~ # grep nfs /etc/fstab
192.168.27.111:/home/nfsforpurehana /hana/shared nfs defaults 0 0
```

Implementing the SAP HANA Storage Connector API

STONITH¹ is an abbreviation for a high-availability cluster concept that ensures data integrity during node failover situations. It is implemented by SAP HANA and enables the use of a HANA standby node in a scale-out (N:1, 1:1) deployment.

For HANA host auto-failover scenarios, the connector is needed as a fencing mechanism in the following situations:

- When HANA triggers host auto-failover, the connector triggers a power reset of failed HANA nodes
- The standby HANA node that will take on persistence becomes active and will access the storage devices

To use block storage together with SAP HANA HA failover, appropriate re-mounting and IO fencing mechanisms must be set up. SAP HANA offers a ready-to-use storage connector, where if an SAP HANA host fails in a distributed system, the standby host takes over the persistence of the failing host by re-mounting the associated LUNs, together with proper fencing.

There are two fundamentally different storage configurations in a multiple-host system: shared storage devices versus separate storage devices with failover reassignment. Using a local HA node requires setup of shared storage. The LUNs are managed by the SAP HANA Storage API connector, therefore no LUNs are mounted onto the FS by the OS.

The SAP HANA nameserver provides the host auto-failover and system replication takeover process. One of the most important uses of the failover hooks is moving around a virtual IP address in conjunction with STONITH.

The standard **fcClient** comes with the installation package and can be activated by using the parameter `--storage_cfg=/some/path` with `/some/path` pointing to the directory, which contains the **global.ini**.

In an SAP HANA scale-out configuration, if the path to the storage LUNs should be shared, the path needs to be mounted before adding host(s) into the scale-out cluster. If the path is not shared, **set global.ini/[persistence]/basepath_shared=no**, then host failover and standby hosts are not possible.

To employ the **fcClient** during the SAP HANA database installation, the **multipath.conf** and **global.ini** must be prepared beforehand. Create the **global.ini** file as follows:

```
se-x3850x6a:~ # more global.ini
[persistence]
basepath_datavolumes=/hana/data
```

¹ Shoot The Other Node In The Head

```

basepath_logvolumes=/hana/log
use_mountpoints=yes
[storage]
ha_provider=hdb_ha.fcClient
partition_*_*_prtype=5
partition_1_data__wwid= WWN // placeholder for your actual WWN
partition_1_log__wwid= WWN
partition_2_data__wwid= WWN
partition_2_log__wwid= WWN
partition_3_data__wwid= WWN
partition_3_log__wwid= WWN

```

This **global.ini** file will be used during the installation as follows:

```

se-x3850x6a:~ # ./hdblcm --ignore=check_signature_file --component_dirs=/hana/shared/SAP_HANA_
DATABASE/server,/hana/shared/SAP_HANA_CLIENT/client --storage_cfg=/hana/shared/SAP_HANA_DATABASE
--addhosts=hananode01:role=worker,hananode02:role=worker,hananode04:role=standby

```

Installing the SAP HANA Database

In the next step, we install SAP HANA server software, following the official [SAP HANA Server and Installation Guide](#).

We perform a quick CLI-based installation. A single node SAP HANA server installation is started using the following command:

```

se-x3850x6a:~ # ./hdblcm --ignore=check_signature_file --component_dirs=/tmp/SAP_HANA_DATABASE/server,/tmp/
SAP_HANA_CLIENT/client --storage_cfg=/tmp/SAP_HANA_DATABASE

```

A scale-out SAP HANA cluster installation is started using the following command:

```

se-x3850x6a:~ # ./hdblcm --ignore=check_signature_file --component_dirs=/hana/shared/SAP_HANA_
DATABASE/server,/hana/shared/SAP_HANA_CLIENT/client --storage_cfg=/hana/shared/SAP_HANA_DATABASE
--addhosts=hananode01:role=worker,hananode02:role=worker,hananode04:role=standby

```

Note that to specify the scale-out node roles during initial installation, we added the following:

```

addhosts=hananode01:role=worker,hananode02:role=worker,hananode03:role=worker,hananode04:role=standby

```

Alternatively, during the installation wizard, when asked to add additional nodes, say yes and define their roles:

Select roles for host 'hananode04':

Index	Host Role	Description
1	worker	Database Worker
2	standby	Database Standby
3	extended_storage_worker	Dynamic Tiering Worker
4	extended_storage_standby	Dynamic Tiering Standby
5	streaming	Smart Data Streaming
6	rdsync	Remote Data Sync
7	ets_worker	Accelerator for SAP ASE Worker
8	ets_standby	Accelerator for SAP ASE Standby
9	xs_worker	XS Advanced Runtime Worker
10	xs_standby	XS Advanced Runtime Standby

SIZING PURE STORAGE FLASH ARRAY FOR SAP HANA TDI

In general, the potential application workload for the SAP HANA-based landscape has to be estimated either by using the SAP Quicksizer (Greenfield) or the available SAP Sizing reports (Brownfield) for analytical or transactional use cases.

In this chapter we describe in detail only the storage layer-related sizing methodology.

To size an SAP HANA environment, the KPIs of the SAP Enterprise Storage certification apply, which are available as part of the official SAP HWCCT documentation.

The sizing of the persistence layer for the SAP HANA database is dependent on the sizing of the main memory and the number of hosts that belong to the system, as well as on the influence of I/O performance on typical SAP HANA database scenarios: OLAP scenarios are usually optimized for fast loading times, and OLTP scenarios are optimized for low latency log writing.

To size a combined SAP HANA and non-HANA workload on a single Pure Storage FlashArray, a separate SAP HANA sizing effort is completed first, and remaining IOPS are used in sizing the non-SAP HANA workload.

Sizing the SAP HANA TDI Workload

Storage sizing for an SAP HANA workload is done in 2 steps:

- Capacity sizing, using the SAP Storage IO Requirements white paper
- Performance sizing, using the number of SSDs per HANA instance

In the following paragraphs, we will discuss applying the SAP HANA sizing requirements for a Pure Storage array.

The storage sizing is based on the following configuration parameters:

- Select the storage array model
- Select the drive size, number of drives, and expansion shelves required to meet the total capacity for the array
- Select the number of drives as per performance requirements for the array
- Select the number of host port connections for the array

CAPACITY SIZING

The formulas listed for sizing the SAP HANA persistence layer are based on the memory sizing for the target SAP HANA database. Memory sizing has to be done before persistence can be sized.

After database memory has been sized using the SAP HANA Quick Sizer, required storage capacity is sized as follows:

- Size data = 1.2x anticipated net data size on disk
- Size log = [systems ≤ 512GB] Size redo log = 1/2 x RAM
- Size log = [systems > 512GB] Size redo log (min) = 512GB
- Size installation/shared (scale-up) = MIN (1 x RAM; 1 TB)
- Size installation/shared (scale-out) = 1 x RAM_of_worker per 4 worker nodes
- Size backups ≥ Size data + Size redo log

Generally speaking, if all hosts in a system have a certain size main memory, it is recommended that the storage subsystem provide at least three times this amount of persistent space. Depending on the backup strategy, a minimum of half of this amount or more must be added. If a disaster recovery solution (system or storage replication) is added to the overall system, the storage requirements generally double. When storage snapshots are used, a rule of thumb is to add an additional 20% capacity to the online data volume capacity.

Capacity sizing for Pure Storage arrays is done using the FlashArray Capacity and Performance Sizer tool. An SAP HANA workload is a 100% sustained write workload. The storage array sizing is based on volume capacity, number of volumes, annual growth rate, and daily change rate. The number of volumes is calculated based on one data and one log volume per SAP HANA instance; the scale-out SAP HANA shared binaries will be configured on a shared NFS file system outside of the array. If the SAP HANA shared binaries for a scale-up instance are configured using an XFS FS on a block volume, one additional storage block volume is used per SAP HANA instance.

Pure Storage arrays use a default data reduction ratio of 1.9:1 for an SAP HANA workload.

PERFORMANCE SIZING

Performance sizing FlashArray for an SAP HANA workload employs a 100% write workload, or "sustained writes" profile, to define the I/O type. Depending on the type of data (OLTP/OLAP), the SAP HANA block sizes vary between 4 KB and 16 MB. A maximum storage write bandwidth of 2-3 GB/s is used for sustained writes, depending on the array media type and interconnect (SSD/NVMe). Read workload requirements are added for backup purposes, and we reload the database from persistence in the event of a power failure on the server side.

The performance of an SAP HANA workload on an SSD-based FlashArray//M is impacted by the number of SSD drives: System bandwidth increases for //M arrays when adding more SSD drives or shelves, so in some cases using multiple smaller drives may prove more efficient than using fewer larger drives. This is not the case for the NVMe-based FlashArray//X, where the system bandwidth is not impacted or determined by the number of NVMe flash modules. The //M20, //M50 and //M70 arrays can add up to four external shelves, to a maximum capacity of 100 net usable SSDs – a fully populated system containing 120 SSDs when including 2 spare drives per data pack of 12 SSDs.

FlashArray is a symmetric, stateless active/active dual controller array architecture, which provides non-disruptive storage operations that mean consistent and predictable performance for planned and unplanned outages.

FlashArray uses a default RAID level protection of RAID-HA, which provides a minimum of dual parity protection, self-heals when an SSD has failed or is removed, automatically increases the parity protection on highly reduced data sets, and performs priority-based RAID reconstructions that begin with the least protected data residing in the RAID set.

The stateless controller architecture and ability to ensure CPU and memory resources, alongside zero-impact benefits from RAID-HA, produce a non-disruptive storage architecture in case of controller upgrades or replacements, or capacity and performance expansions.

We recommend not exceeding 70% of total array performance. If in doubt, eliminate the bottleneck of SAS bus limitations with FlashArray//M and use FlashArray//X for an SAP HANA database workload.

Sizing the Non-HANA TDI Workload

When SAP HANA workload sizing is completed, it is possible to use remaining IOPS on the system to size a non-HANA workload without impacting the SAP HANA KPI requirements.

Typical workloads that impact SAP HANA KPIs are:

- Workloads that cause 100% 64k random writes, reaching max throughput
- Workloads that cause 256k sequential read, reaching max throughput
- Workloads that cause the front-end ports IOPS max to be reached

Pure Storage arrays use a default data reduction ratio of 3.2:1 for a traditional SAP workload.

CONCLUSION

The test results achieved with the setup in this document clearly demonstrate the high performance, resiliency, and scalability of Lenovo Server and Pure Storage FlashArray for SAP HANA-based application landscapes.

We've demonstrated how customers can achieve these benefits with minimal server tuning and storage configuration setup. These benefits will be even more pronounced on the higher-level FlashArray//X R2 generation of arrays.

RESOURCES

The following is a list of all external resources linked in this paper.

- SAP HANA in the Data Center, Dr. Ralf Czekalla:
<https://apps.support.sap.com/sap/support/knowledge/preview/en/2407186>
- Storage IO Requirements WP V2.10 February 2017
<https://www.sap.com/documents/2015/03/74cdb554-5a7c-0010-82c7-eda71af511fa.html>
- SAP HANA Troubleshooting and Performance Analysis Guide
https://help.sap.com/doc/e344ef1295b6433e88fe084c0768e1cd/1.0.12/en-US/SAP_HANA_Troubleshooting_and_Performance_Analysis_Guide_en.pdf
- SAP HANA Server Installation and Update Guide
http://help.sap.com/hana/SAP_HANA_Server_Installation_Guide_en.pdf
- Lenovo ThinkSystem SR950 Product Guide
<https://lenovopress.com/lp0647>
- Lenovo ThinkSystem SR650 Product Guide
<https://lenovopress.com/lp0644>
- QLogic 16 Gbit FC HBAs Product Guide
<https://lenovopress.com/tips-0954-qlogic-16gb-fc-hba>
- Lenovo System x3850 X6 / x3950 X6 Product Guide
<https://lenovopress.com/tips1250-system-x3850-x6-6241>
<https://lenovopress.com/tips1251-system-x3950-x6-6241>
- Lenovo System x3850 X6 and x3950 X6 Planning and Implementation Guide
<https://lenovopress.com/sg248208.pdf>

© 2018 Pure Storage, Inc. All rights reserved.

Pure Storage, Pure1, and the “P” logo are trademarks or registered trademarks of Pure Storage, Inc. in the U.S. and other countries. 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. Lenovo, the Lenovo logo, and For Those Who Do are trademarks or registered trademarks of Lenovo in the United States, other countries, or both. All other product and service names mentioned are the trademarks of their respective companies.

The Pure Storage product described in this documentation is distributed under a license agreement and may be used only in accordance with the terms of the agreement. The license agreement restricts its use, copying, distribution, decompilation, and reverse engineering. No part of this documentation may be reproduced in any form by any means without prior written authorization from Pure Storage, Inc. and its licensors, if any.

THE DOCUMENTATION IS PROVIDED “AS IS” AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NON-INFRINGEMENT ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID. PURE STORAGE SHALL NOT BE LIABLE FOR INCIDENTAL OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH THE FURNISHING, PERFORMANCE, OR USE OF THIS DOCUMENTATION. THE INFORMATION CONTAINED IN THIS DOCUMENTATION IS SUBJECT TO CHANGE WITHOUT NOTICE.

ps_wp23p_flasharray-and-lenovo-for-sap-hana_01