Technical white paper

HPE MSA 1050/2050/2052

Best practices

Contents

Executive summary

This white paper highlights best practices for optimizing and deploying HPE MSA 1050/2050/2052 arrays and should be used together with other HPE MSA manuals. The MSA 1050/2050/2052 array is the fifth-generation MSA storage area network (SAN). MSA technical user documents are available from the HPE MSA Storage webpage. This paper is also designed to convey best practices in the deployment of the MSA 1050/2050/2052 array.

Note
Images shown in this document reflect current software functionality as of the latest firmware available at the time of publication. Features and functionality might vary with different storage system firmware levels.

Intended audience

This white paper is intended for MSA 1050/2050/2052 administrators with previous SAN knowledge. It offers best practices that can contribute to an MSA best customer experience.

Prerequisites

Prerequisites for using this product include knowledge of:

- Networking
- Storage system configuration
- SAN management
- Connectivity methods such as direct attached storage (DAS), Fibre Channel, and serial attached SCSI (SAS)
- iSCSI and Ethernet protocols

Related documentation

In addition to this guide, other documents or materials for this product include:

- HPE MSA System Racking Instructions
- HPE MSA 1050 User Guide
- HPE MSA 2050 User Guide
- HPE MSA 1050/2050 Storage Management Utility (SMU) Reference Guide
- HPE MSA 1050/2050 CLI Reference Guide
- HPE MSA 1050 Quick Start Instructions
- HPE MSA 2050/2052 Quick Start Instructions
- HPE MSA 1050 Cable Configuration Guide
- HPE MSA 2050 Cable Configuration Guide

You can find HPE MSA 1050/2050/2052 documents from the Hewlett Packard Enterprise Information Library for the HPE MSA 1050 and HPE MSA 2050/2052.

Introduction

The MSA models referenced in this paper include the MSA 1050, MSA 2050, and MSA 2052.

MSA 1050

The MSA 1050 is designed for entry-level markets. It features 8 Gb Fibre Channel, 6 Gb/12 Gb SAS, and 1 GbE and 10 GbE iSCSI protocols. The MSA 1050 also features:

- 6 GB cache per controller (4 GB read/write plus 2 GB system memory)
- Support for small form factor (SFF) and large form factor (LFF) solid-state drives (SSDs)
- Two host ports per controller
- 4 Gb/8 Gb Fibre Channel connectivity
- 1 GbE/10 GbE iSCSI connectivity
- 6 Gb/12 Gb SAS connectivity
- Support for MSA fan-out SAS cables
- Support for up to four disk enclosures including the array enclosure
- Support for up to 96 SFF drives and 48 LFF drives
- Support for Performance Tiering software[1]
- Automated Tiering
- Thin Provisioning
- Support for read cache
- Wide striping, which allows more hard disk drives (HDDs) behind a single volume to improve performance (for example, more than 16 drives per volume)
- Support for replication snapshot history and queueing for Remote Snap[2]
- Remote Snap for both Fibre Channel and iSCSI[2]

MSA 2050

The MSA 2050 is a high-performance storage system designed for Hewlett Packard Enterprise (HPE) customers who want 8 Gb or 16 Gb Fibre Channel, 6 Gb or 12 Gb SAS, and 1 GbE or 10 GbE iSCSI connectivity with four host ports per controller. The MSA 2050 is the industry's fastest entry-level array. It provides excellent value for customers who need performance balanced with price to support initiatives such as consolidation and virtualization.

The MSA 2050 delivers this performance by offering:

- Twice the I/O performance compared to the previous generation
- 8 GB cache per controller (4 GB read/write plus 4 GB system memory)
- Support for SFF and LFF SSDs
- Four host ports per controller
- 4 Gb/8 Gb/16 Gb Fibre Channel connectivity
- 1 GbE/10 GbE iSCSI connectivity
- 6 Gb/12 Gb SAS connectivity

[1] A license is required for the Performance Tier for mixed SSD and HDD systems. For a system with only SSDs, a Performance Tier license is not required.
[2] Remote Snap software requires a license.

- Support for both Fibre Channel and iSCSI in a single controller
- Support for up to eight disk enclosures including the array enclosure
- Support for up to 192 SFF drives and 96 LFF drives
- Support for read cache
- Thin Provisioning
- Automated Tiering
- Support for Performance Tiering[3]
- Wide striping, which allows more hard drives behind a single volume for improved performance (for example, more than 16 drives per volume)
- Support for replication snapshot history and queueing for Remote Snap[4]
- Remote Snap for both Fibre Channel and iSCSI[4]
- Volume Copy across pools
- Support for full disk encryption (FDE) using self-encrypting drives (SEDs)[5]

MSA 2052

The MSA 2052 SAN offers an entry-level platform with built-in hybrid flash for application acceleration and high performance. It is ideal for performance-hungry applications and includes 1.6 TB of SSD capacity.

An MSA 2052 storage system supports all the features of the MSA 2050, plus 1.6 TB of SSD capacity, standard. The MSA Advanced Data Services (ADS) Suite is included as a standard feature on the MSA 2052 at no extra charge. The MSA ADS Suite includes the following functionality:

- Performance Tiering and Archive Tiering licenses
- A 512-snapshot license
- Remote Snap license

Note
The ADS Suite license key ships standard with the MSA 2052 array and must be redeemed and installed to enable the services.

The MSA 1050/2050 storage systems ship standard with a license for 64 snapshots and Volume Copy for increased data protection. There is an optional license for 512 snapshots.

The MSA 1050/2050/2052 can replicate data between MSA 1050/2050/2052 arrays using Fibre Channel or iSCSI. The user can also replicate from an MSA 1050/2050/2052 to an MSA 1040/2040/2042 array using virtual volumes only if the MSA 1040/2040/2042 has GL200 or newer firmware. The optional Remote Snap feature is needed for the replication.

Terminology

The following terms are used throughout this white paper. These definitions are key to understanding MSA arrays.

- **Virtual storage:** Data is virtualized not only across a single disk group, but also across multiple disk groups with different performance capabilities and use cases.

[3] A license is required for the Performance Tier for mixed SSD and HDD systems. For a system with only SSDs, a Performance Tier license is not required.
[4] Remote Snap software requires a license.
[5] SEDs are supported only in the MSA 2050.

- **Page:** A page is an individual block of data residing on a physical disk. A page is the smallest unit of data that can be allocated, deallocated, or moved between virtual disk groups in a tier or between tiers. For virtual storage, the page size is 4 MB.
- **Disk group:** A disk group is a collection of disks in a given RAID level.
- **Storage pools:** Storage pools are composed of one or more virtual disk groups. A volume's data on a logical unit number can span all disk drives in a pool. When capacity is added to a system, users benefit from the performance of all spindles in that pool. When leveraging storage pools, an MSA 1050/2050/2052 array supports large, flexible volumes with sizes up to 128 TiB and facilitates seamless capacity expansion. As volumes are expanded, data automatically reflows to balance capacity utilization on all drives.
- **Logical unit number (LUN):** MSA 1050/2050/2052 arrays support 512 volumes and up to 512 snapshots in a system. All these volumes can be mapped to LUNs. The maximum LUN size is 128 TiB.
- **Thin Provisioning:** Thin Provisioning software allows storage allocation of physical storage resources only when they are consumed by an application. Thin Provisioning also allows you to overprovision physical storage pool resources, enabling volumes to grow without having to predict storage capacity up front.
- **Tiers:** Disk tiers consist of an aggregation of one or more disk groups of similar physical disks. MSA 1050/2050/2052 arrays support three distinct tiers:
  - A performance tier with SSDs
  - A standard SAS tier with enterprise SAS HDDs
  - An archive tier using midline SAS HDDs

  LUN-level tiering requires careful planning. You should place applications that require the best performance on disk groups that use high-performance SSDs. You can place applications with lower performance requirements on disk groups consisting of enterprise SAS or midline SAS HDDs.

  The MSA 1050/2050/2052 Automated Tiering engine moves data between available tiers based on the access characteristics of that data. Frequently accessed data contained in pages is migrated to the highest available tier, delivering maximum I/Os to the application. Similarly, cold or infrequently accessed data is moved to lower performance tiers. Data is migrated between tiers automatically so that I/O is optimized in real-time.

  Archive and standard tiers are provided at no charge on the MSA 1050 and MSA 2050 platforms. A performance tier using a fault-tolerant SSD disk group is an add-on feature that requires a license for the MSA 1050/2050. Without a Performance Tier license installed, you can still use SSDs as read cache with the sub-LUN tiering feature. Sub-LUN tiering from SAS midline (MDL) (archive tier) to enterprise SAS (standard tier) drives is provided at no additional charge for the MSA 1050/2050/2052 arrays.
- **Read cache:** Read cache is an extension of the controller cache. Read cache provides a lower-cost way to get performance improvements from SSDs.
- **Automated Tiering:** Automated Tiering is a technology that enables the automatic movement of data between storage tiers based on access trends. In MSA 1050/2050/2052 arrays, Automated Tiering places data in a LUN that is accessed frequently in better-performing media; data that is infrequently accessed is placed in slower media.
- **Array enclosure:** This is the array head or chassis of the MSA that includes the MSA controllers.
- **Disk enclosure:** This is the expansion shelf that is connected to the array enclosure.
- **Storage system:** This is the whole MSA system that includes the array enclosure and disk enclosures.

General best practices

This section outlines some general best practices when administering an MSA storage system.

Become familiar with the array by reading the manuals

The first recommended best practice is to read the corresponding guides for the MSA 1050 and MSA 2050/2052. These documents include the User Guide, the SMU Reference Guide, and the CLI Reference Guide. The appropriate guide depends on the interface that you will use to configure the storage array. Always operate the array in accordance with the user manual. In particular, never exceed the environmental operation requirements. The MSA Remote Snap Software technical white paper is also helpful to review.

Stay current on firmware

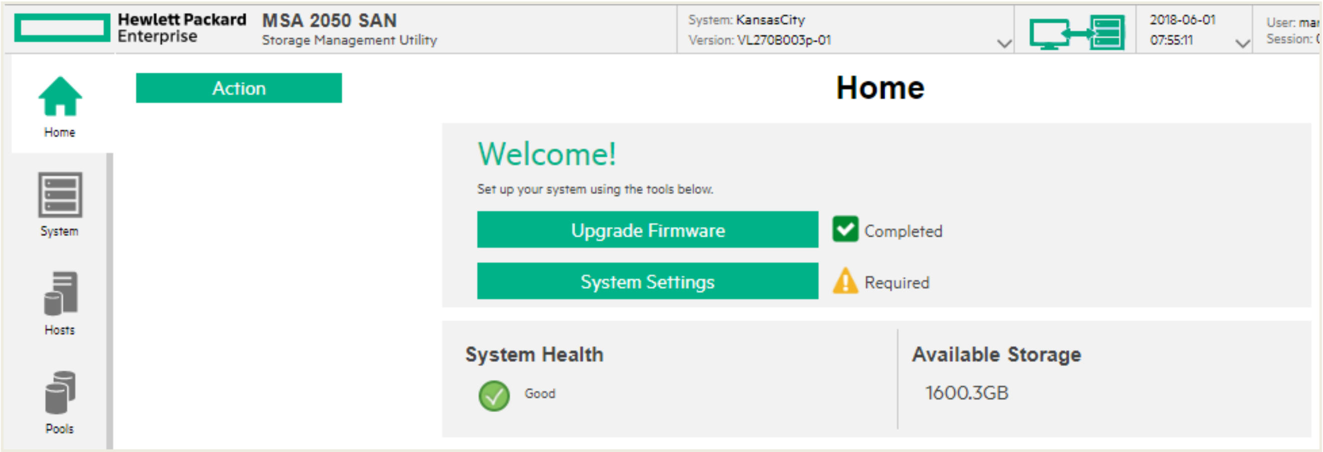
Use the latest controller, disk, and disk enclosure firmware to benefit from the continual improvements in the performance, reliability, and functionality of MSA 1050/2050/2052 arrays. For additional information, see the release notes and release advisories for the respective MSA products. You can find the latest firmware for all components in MSA arrays at https://www.hpe.com/storage/msafirmware. Additional information is available from the HPE Support Center.

Review settings in the Welcome panel of the SMU

The Welcome panel in the SMU provides options for you to quickly and easily set up the MSA system. It also guides you through the firmware update and configuration process. To use guided setup, you must first access the Upgrade Firmware panel where you can review the controller module firmware version and perform recommended updates. When finished, you must configure your system settings by accessing the System Settings panel and completing all required options. After these settings are complete, you can provision the system.

The standard Home panel is not visible until all required actions are completed or acknowledged.

Note
A user with the Manage role must complete the guided setup process.



System: KansasCity
Version: VL270B003p-01
2018-06-01
07:55:11
User: ma
Session: (

Home

Welcome!

Set up your system using the tools below.

Upgrade Firmware ✓ Completed

System Settings ⚠ Required

System Health
Good

Available Storage
1600.3GB

Figure 1. Welcome panel

Use tested and supported configurations

Deploy the MSA array only in supported configurations. Do not risk the availability of critical applications with unsupported configurations. HPE does not recommend or provide support for unsupported MSA configurations.

The HPE primary portal used to obtain detailed information about supported HPE storage product configurations is the HPE Single Point of Connectivity Knowledge (SPOCK). An HPE Passport account is required to enter the SPOCK website.

Understand what a host is from the array perspective

An initiator is like an external port on a host bus adapter (HBA). An initiator port does not equate to a physical server; rather, it is a unique connection on that server. For example, a dual-port Fibre Channel HBA has two ports and therefore two unique initiators. The array shows two separate initiators for that HBA.

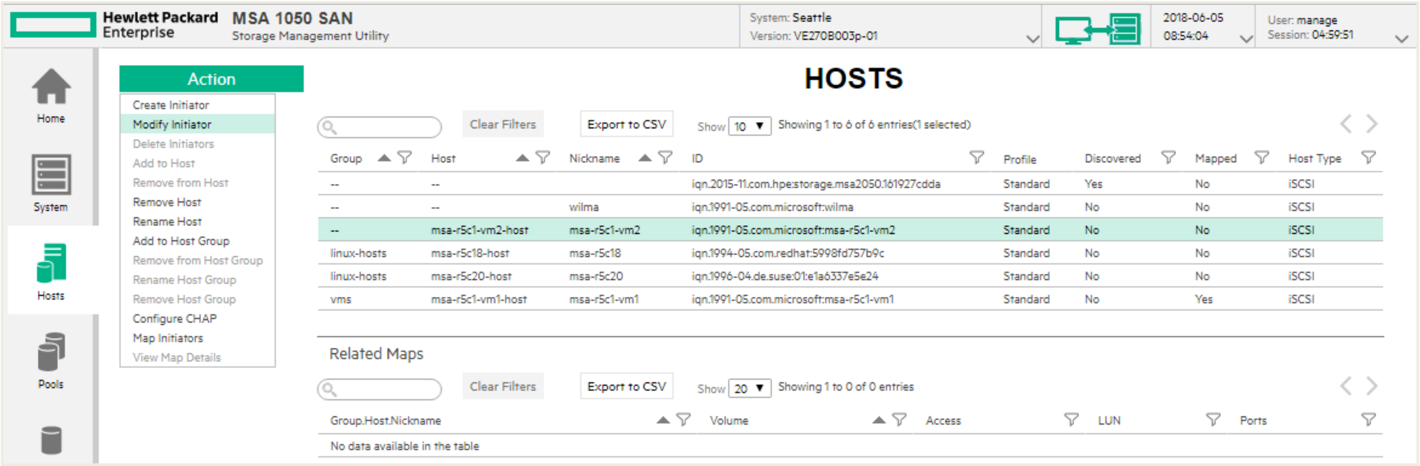
A host is a collection of one or more initiators. MSA firmware can support 512 hosts with multiple initiators per host. An MSA storage system can manage 1024 initiators.

The array supports the grouping of initiators under a single host and grouping hosts into a host group. Grouping the initiators and hosts simplifies mapping operations.

Rename hosts to a user-friendly name

Applying user-friendly names to the hosts enables you to easily identify which hosts are associated with servers and operating systems. A recommended method for acquiring and renaming World Wide Names (WWNs) is to connect one cable at a time and then rename the WWN to an identifiable name. The following procedure outlines the steps needed to rename hosts using the SMU:

1. Log in to the SMU and from the left frame, click **Hosts**.
1. Locate and highlight the WWN (ID) you want to name.
2. From the **Action** drop-down list, click **Modify Initiator**.
3. Enter the initiator nickname and click **OK**.
4. Repeat these steps for additional initiator connections.



HOSTS

Showing 1 to 6 of 6 entries(1 selected)

Group	Host	Nickname	ID	Profile	Discovered	Mapped	Host Type
--	--		iqn.2015-11.com.hpe:storage.msa2050.161927cdda	Standard	Yes	No	iSCSI
--	--	wilma	iqn.1991-05.com.microsoft:wilma	Standard	No	No	iSCSI
--	msa-r5c1-vm2-host	msa-r5c1-vm2	iqn.1991-05.com.microsoft:msa-r5c1-vm2	Standard	No	No	iSCSI
linux-hosts	msa-r5c18-host	msa-r5c18	iqn.1994-05.com.redhat:5998fd757b9c	Standard	No	No	iSCSI
linux-hosts	msa-r5c20-host	msa-r5c20	iqn.1996-04.de.suse:01:e1a6337e5e24	Standard	No	No	iSCSI
vms	msa-r5c1-vm1-host	msa-r5c1-vm1	iqn.1991-05.com.microsoft:msa-r5c1-vm1	Standard	No	Yes	iSCSI

Related Maps

Showing 1 to 0 of 0 entries

Group.Host.Nickname	Volume	Access	LUN	Ports
No data available in the table				

Figure 2. Renaming hosts

The recommended practice is to use initiator nicknaming as outlined in Figure 2. You can use the SMU to aggregate host initiators and group hosts.

Best practice for monitoring array health

Setting up the array to send notifications is important for troubleshooting and log retention.

Configure email and SNMP notifications

The SMU is the recommended method for setting up email and SNMP notifications. You can set up these services easily by using a web browser. To connect, enter the IP address of the MSA 1050/2050/2052 management port.

Email notifications can be sent to as many as three different email addresses. In addition to the normal email notification, HPE recommends enabling managed logs with the **Include logs as an email attachment** option enabled. When this feature is enabled, the system automatically attaches the system log files to the managed log's email notifications that are sent. The managed log's email notification contains the logs for future diagnostic investigation.

The MSA 1050/2050/2052 storage system has a limited amount of space to retain logs. When this log space is exhausted, the oldest entries in the log are overwritten. For most systems, this space is adequate to allow for diagnosing issues seen on the system. The managed logs feature notifies the administrator that the logs are nearing a full state and that older information will soon get overwritten. The administrator can then choose to manually save the logs. If the **Include logs as an email attachment** check box is selected, the segment of logs that is nearing a full state is attached to the email notification. Managed logs attachments can be multiple megabytes in size.

Enabling the managed logs feature allows log files to be transferred from the storage system to a log-collection system to avoid losing diagnostic data. The option is disabled by default.

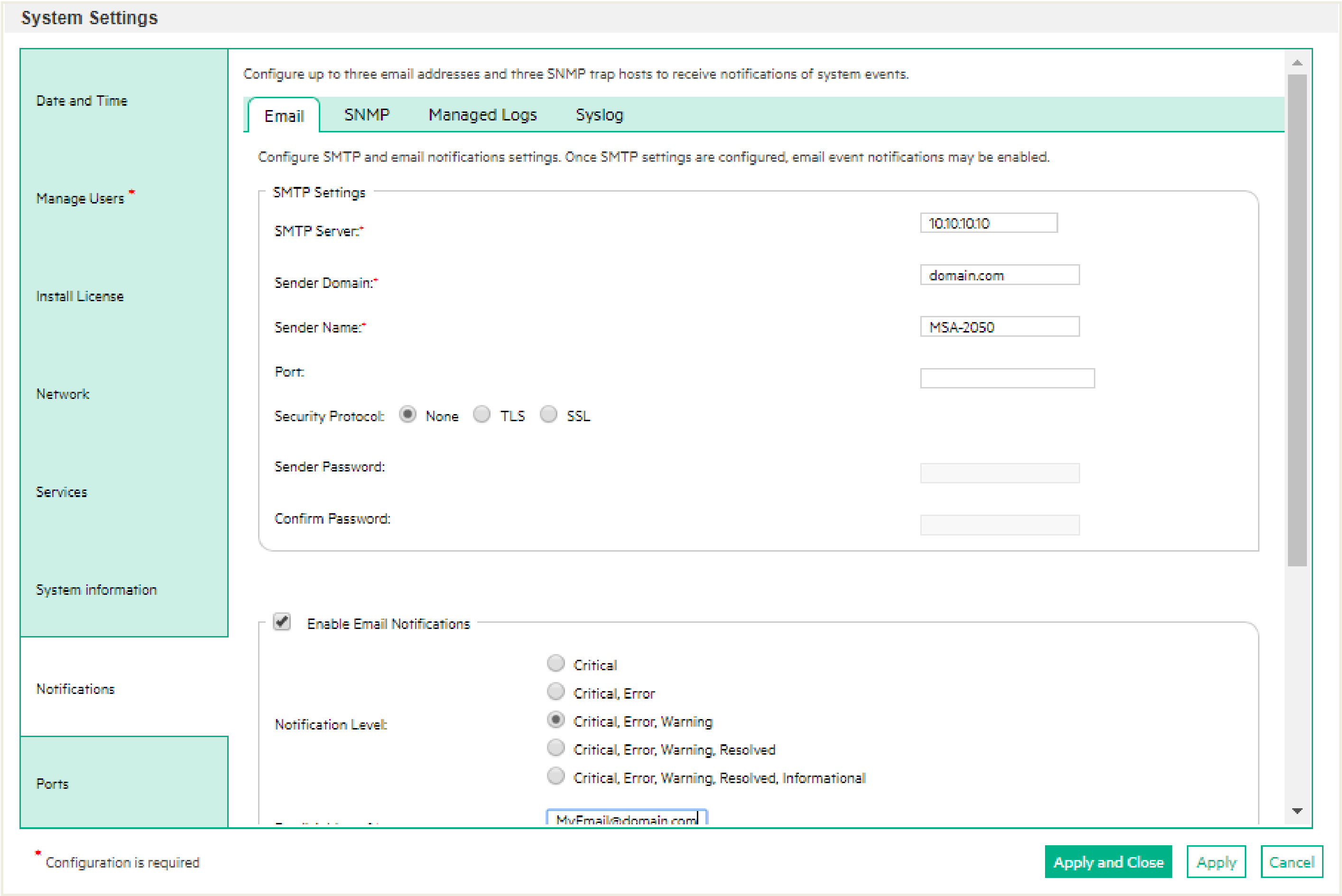
The MSA storage system sends an event for a degraded component, such as a faulty hard drive, only once. If that event alert is missed or overlooked, the system could remain in a degraded state for an extended time. To avoid missing important events, you can enable the MSA health alerts feature, which sends a weekly email stating the health state of the storage system. In most cases, this is a healthy system alert, but it could also be a reminder that attention is required. You should also consider the lack of a health alert as a notification that attention is required.

Important
HPE recommends that at least one method of notification is configured and that you regularly monitor for unhealthy components.

Note
HPE recommends careful monitoring for compact flash events. A list of relevant event codes can be found in the HPE MSA Event Descriptions Reference Guide.

HPE recommends enabling SNMP traps. SNMPv1 traps can be sent to up to three host trap addresses (that is, SNMP servers). To send SNMPv3 traps, create an SNMPv3 user with the trap target account type. Use SNMPv3 traps rather than SNMPv1 traps for greater security. SNMP traps can be useful in troubleshooting issues with MSA 1050/2050/2052 arrays.

For details on accessing the Notifications setup panel to configure alerts and accessing the User Management panel to configure SNMPv3 users and trap targets, refer to the MSA 1050/2050 SMU Reference Guide.



Configure up to three email addresses and three SNMP trap hosts to receive notifications of system events.

Configure SMTP and email notifications settings. Once SMTP settings are configured, email event notifications may be enabled.

SMTP Settings

SMTP Server:* 10.10.10.10

Sender Domain:* domain.com

Sender Name:* MSA-2050

Port:

Security Protocol: ● None ○ TLS ○ SSL

Sender Password:

Confirm Password:

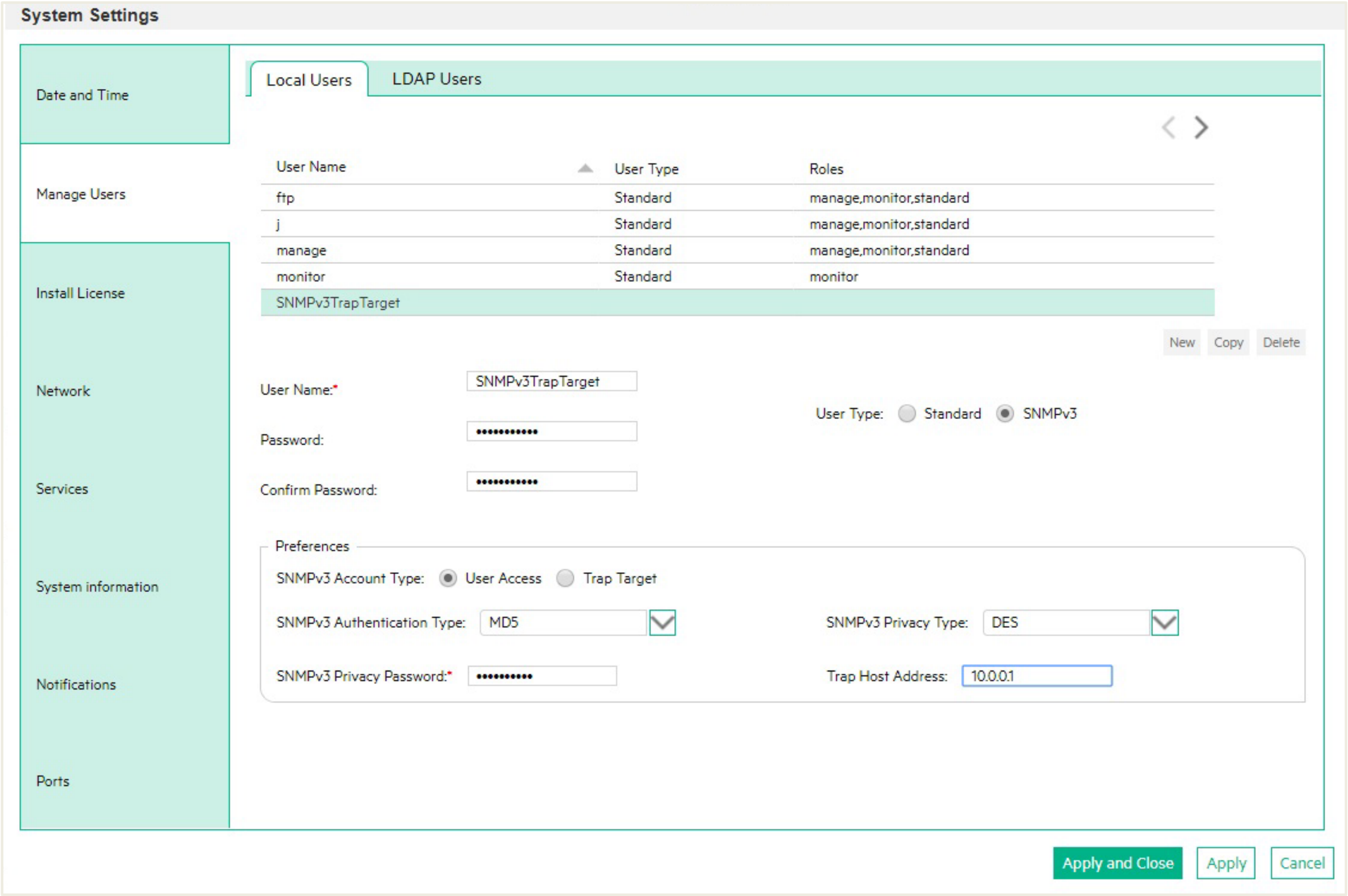
☑ Enable Email Notifications

Notification Level:
- ○ Critical
- ○ Critical, Error
- ● Critical, Error, Warning
- ○ Critical, Error, Warning, Resolved
- ○ Critical, Error, Warning, Resolved, Informational

MyEmail@domain.com

* Configuration is required

Figure 3. Setting up management services



User Name	User Type	Roles
ftp	Standard	manage,monitor,standard
j	Standard	manage,monitor,standard
manage	Standard	manage,monitor,standard
monitor	Standard	monitor
SNMPv3TrapTarget		

User Name:* SNMPv3TrapTarget

Password: •••••••••••

Confirm Password: •••••••••••

User Type: ○ Standard ● SNMPv3

Preferences

SNMPv3 Account Type: ● User Access ○ Trap Target

SNMPv3 Authentication Type: MD5

SNMPv3 Privacy Type: DES

SNMPv3 Privacy Password:* ••••••••••

Trap Host Address: 10.0.0.1

Figure 4. Creating an SNMPv3 user

Set the notification level for email and SNMP

Setting the notification level to **Critical**, **Error**, or **Warning** for email configurations and setting **Warning** for SNMP configurations ensures that events of that level or greater are sent to the destinations (that is, SNMP server or SMTP server) for that notification. HPE recommends setting the notification level to **Warning**.

MSA 1050/2050/2052 notification levels are:

- Warning sends notifications for all Critical, Error, or Warning events.
- Error sends notifications for Critical and Error events.
- Critical only sends notifications for Critical events.

Sign up for proactive product advisory notifications

Sign up for proactive notifications to receive MSA product advisories. Applying the suggested resolutions can enhance the availability of the product. Sign up for the notifications on the HPE subscription page.

Best practices for provisioning storage on MSA 1050/2050/2052 arrays

This section outlines the best methods for optimizing virtual storage features such as Thin Provisioning, wide striping, and automated tiering for MSA 1050/2050/2052 arrays.

Thin Provisioning

Thin Provisioning is a storage allocation scheme that automatically allocates storage as applications need it.

Thin Provisioning dramatically increases storage utilization by removing the need to match purchased capacity to allocated capacity. Traditionally, application administrators purchased storage based on the current capacity and future growth needs. This resulted in surplus capacity and unused space.

With Thin Provisioning, applications can be provided with the capacity required for growth but can begin operating on a smaller amount of physical storage. As the applications fill their storage space, new storage can be purchased as needed and added to the array's storage pools. This results in more efficient storage utilization and reduced power and cooling requirements.

Overcommit is enabled by default. The overcommit setting lets you oversubscribe the physical storage (that is, provision volumes in excess of physical capacity). If you disable overcommit, you can provision virtual volumes only up to the available physical capacity. Overcommit is performed on a per-pool basis by using the **Change Pool Settings** option. Consult the HPE MSA 1050/2050 SMU Reference Guide for more details.

Thresholds and notifications

If you use Thin Provisioning, monitor space consumption and set notification thresholds appropriately for the rate of storage consumption. Users with a Manage role can view and change settings that affect the thresholds and corresponding notifications for each storage pool. The following thresholds and notifications can help determine when more storage needs to be added:

- **Low Threshold:** When this percentage of virtual pool capacity has been used, informational event 462 is generated to notify the administrator. This value must be less than the Mid Threshold value. The default is 50%.
- **Mid Threshold:** When this percentage of virtual pool capacity has been used, event 462 is generated to notify the administrator to add capacity to the pool. This value must be between the Low Threshold and High Threshold values. The default is 75%. If the pool is not overcommitted, the event has an Informational severity. If the pool is overcommitted, the event has a Warning severity.
- **High Threshold:** When this percentage of virtual pool capacity has been used, event 462 is generated to alert the administrator to add capacity to the pool. This value is automatically calculated based on the available capacity of the pool minus 200 GB of reserved space. If the pool is not overcommitted, the event has an Informational severity. If the pool is overcommitted, the event has a Warning severity and the system will use write-through cache mode until virtual pool usage drops back below this threshold.

T10 UNMAP for thin reclaim

UNMAP is the ability to reclaim thinly provisioned storage after the storage is no longer needed. There are procedures to reclaim UNMAP space when using Thin Provisioning and VMware® ESXi.

The user should run the UNMAP command with ESXi to avoid performance issues. In ESXi, the UNMAP command was decoupled from auto reclaim; therefore, use the VMware vSphere® CLI command to run the UNMAP command. Refer to the VMware knowledge base for more details on the UNMAP command and reclaiming space.

Background scrubbing

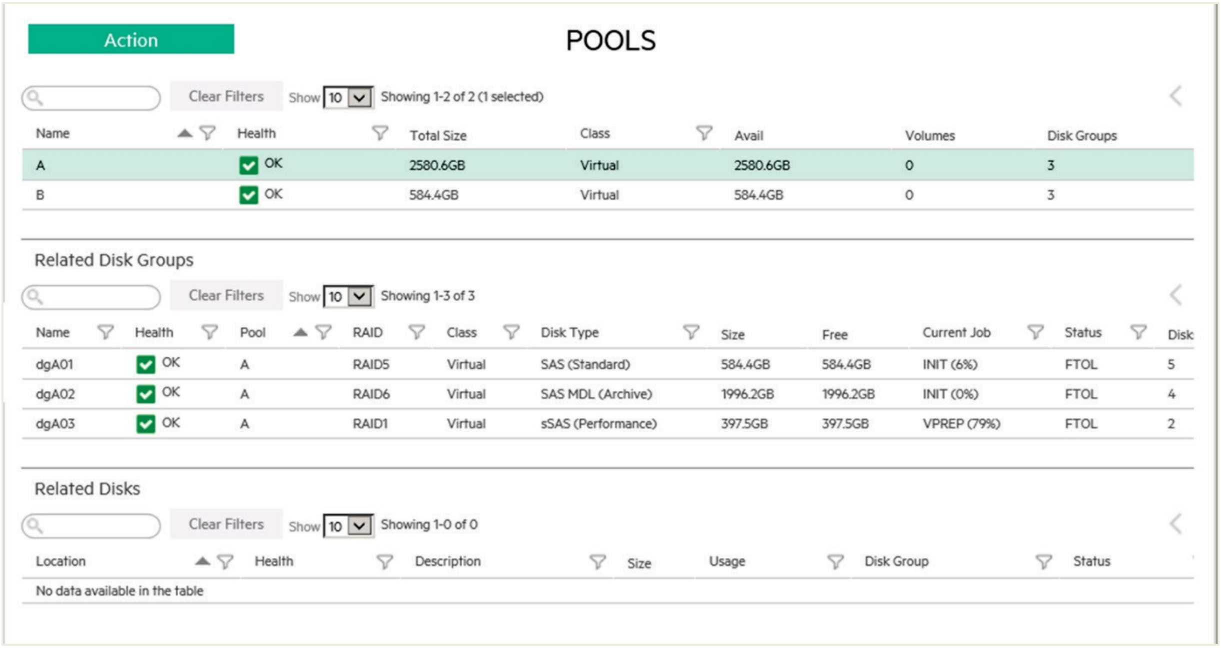
HPE recommends that background scrubbing is enabled for disks that are and are not assigned to disk groups. This process will periodically analyze and fix any parity RAID errors, as well as alert to any media errors. Additionally, zeroed pages will be returned to the pool to which a disk group is associated as unallocated capacity. Refer to the HPE MSA 1050/2050 SMU Reference Guide for more information.

Pool balancing

Creating and balancing storage pools properly can help with performance of the MSA array. HPE recommends keeping pools balanced from a capacity utilization and performance perspective. Pool balancing leverages both controllers and balances the workload across the two pools.

Assuming symmetrical composition of storage pools, you should create and provision storage volumes by the workload that will be used. For example, an archive volume is best placed in a pool with the most available Archive Tier space. For a high-performance volume, create the volume on the pool that is getting the least amount of I/O on the standard and performance tiers.

Determining the pool space can easily be viewed in the SMU. Simply navigate to **Pools** and click the name of the pool.



POOLS

Showing 1-2 of 2 (1 selected)

Name	Health	Total Size	Class	Avail	Volumes	Disk Groups
A	✓ OK	2580.6GB	Virtual	2580.6GB	0	3
B	✓ OK	584.4GB	Virtual	584.4GB	0	3

Related Disk Groups

Showing 1-3 of 3

Name	Health	Pool	RAID	Class	Disk Type	Size	Free	Current Job	Status	Disk
dgA01	✓ OK	A	RAID5	Virtual	SAS (Standard)	584.4GB	584.4GB	INIT (6%)	FTOL	5
dgA02	✓ OK	A	RAID6	Virtual	SAS MDL (Archive)	1996.2GB	1996.2GB	INIT (0%)	FTOL	4
dgA03	✓ OK	A	RAID1	Virtual	sSAS (Performance)	397.5GB	397.5GB	VPREP (79%)	FTOL	2

Related Disks

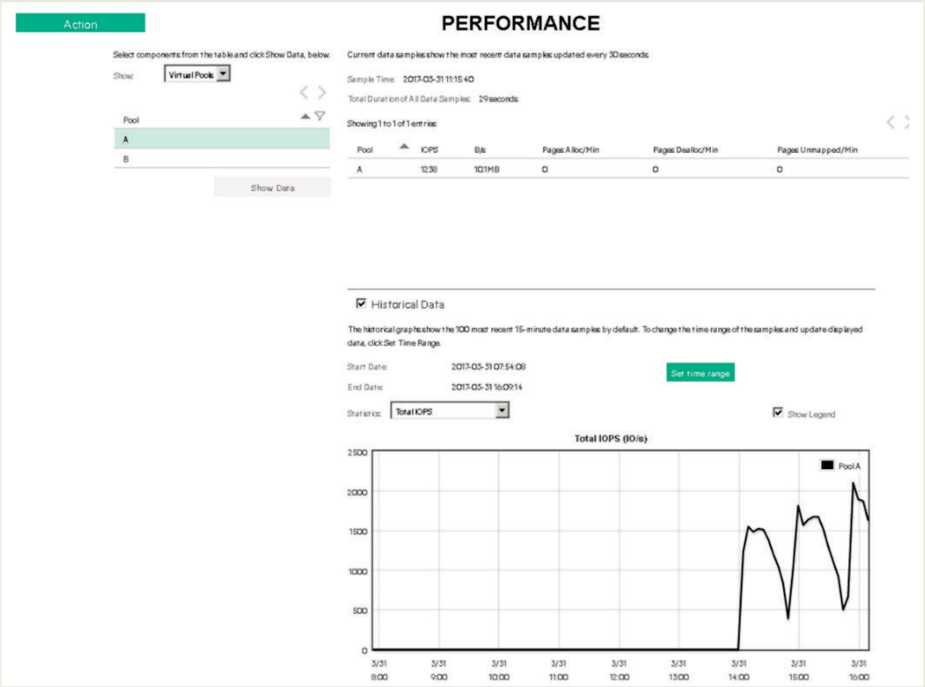
Showing 1-0 of 0

Location	Health	Description	Size	Usage	Disk Group	Status
No data available in the table						

Figure 5. MSA Pool A screen

Viewing the performance of pools or virtual disk groups can also help to determine where to place the archive tier space.

To view specific data, from the SMU, navigate to **Performance**. From the Show drop-down menu, click **Virtual Pools**. Next, click the pool. For real-time data, click **Show Data**. For historical data, select the **Historical Data** check box and click **Set time range**.



Select components from the table and click Show Data, below.

Show: Virtual Pools

Pool
A
B

Current data samples show the most recent data samples updated every 30 seconds.

Sample Time: 2017-03-31 11:15:40

Total Duration of All Data Samples: 29 seconds

Showing 1 to 1 of 1 entries

Pool	IOPS	B/s	Pages Alloc/Min	Pages Dealloc/Min	Pages Unmapped/Min
A	1238	10.1MB	0	0	0

☑ Historical Data

The historical graphs show the 100 most recent 15-minute data samples by default. To change the time range of the samples and update displayed data, click Set Time Range.

Start Date: 2017-03-31 07:54:08

End Date: 2017-03-31 16:09:14

Statistics: Total IOPS

☑ Show Legend

Figure 6. MSA Virtual Pools Performance screen

MSA 1050/2050/2052 arrays also can copy volumes owned by one pool to the other pool. Use this volume copy feature if the storage system becomes unbalanced.

Wide striping

MSA 1050/2050/2052 arrays support the wide striping concept for virtual storage. Wide striping means that when virtual disk groups are in a storage pool, the MSA algorithm evenly distributes the allocated pages of a volume across all disk groups in the storage pool.

Wide striping also allows for rapid expansion of volumes if all pool capacity is consumed. When one or more virtual disk groups are added to the storage pool, the volumes within the pool immediately benefit from the additional capacity. The leveling process begins automatically and redistributes data evenly across all disk groups in the storage pool. Essentially, you can increase storage space and improve storage performance by adding disk groups.

Note

The rebalancing happens automatically; no user interaction is needed.

Expanding virtual volumes

A virtual disk group in a pool might start to fill up. To add more space easily, the MSA uses wide striping to increase the size of the virtual volumes. The recommended method to increase the volume size is to add a new virtual disk group with the same amount of drives and RAID type as the existing virtual disk group.

For example, imagine that a virtual disk group in Pool A is filling up. This virtual disk group has six 900 GB, 10K rpm disk drives in a RAID 6 configuration. The recommended procedure is to create a new virtual disk group on Pool A that also has six 900 GB, 10K rpm disk drives in a RAID 6 configuration.

Using MSA 2052 embedded SSDs

The MSA 2052 includes solid-state drives in SFF and LFF configurations to allow users to accelerate application performance. In addition to including two SSDs in the base configuration as a standard feature for flash acceleration, the MSA 2052 also features a rich set of standard software that includes a license for 512 snapshots as well as Remote Snap replication and Performance Tiering capabilities. Each MSA 2052 model ships standard with 1.6 TB (2 x 800 GB) mixed-use SSDs that can be used as read cache or as a start to building a fully tiered configuration. With the inclusion of the Advanced Data Services software license, customers can choose how to best use these SSDs to provide the optimal flash acceleration for their environment. The MSA 2052 storage system uses a real-time tiering algorithm, which works without user intervention to place the hottest pages of an array on the fastest medium at the right time. The tiering engine moves hot pages up to SSD for flash acceleration and moves cooler, less-used pages back down to spinning disks as workloads change in real-time. All these tasks are performed without any intervention from the IT manager.

Using the MSA 2052 SSDs for read cache

This section details the best methods for using MSA 2052 arrays.

With two SSDs standard on the MSA 2052, each storage pool can have one read cache SSD. The method to construct this setup is:

1. Create a virtual disk group on Pool A using enterprise SAS or midline SAS drives.

Important
Make sure to choose **Pool A** from the drop-down menu.

2. Create the read cache for Pool A by using SSD Number 1.
3. Create a virtual disk group on Pool B by using enterprise SAS or midline SAS drives.

Important
Make sure to choose **Pool B** from the drop-down menu.

4. Create the read cache for Pool B by using SSD Number 2.

Now you have used each SSD of the MSA 2052 as read cache for both pools.

Using the MSA 2052 SSDs for Performance Tiering

With two SSDs standard on the MSA 2052, only one storage pool can have a performance tier disk group without the need to purchase additional SSDs. This is because you need a minimum of two SSDs to construct a RAID 1 disk group to use for auto-tiering onto the performance tier. If you purchase additional SSDs, then you may create RAID 1, 5, or 6 disk groups.

The method to construct this setup is:

1. Create a virtual disk group for Pool A using the SSDs.

Important
Select the correct RAID level and make sure to choose **Pool A** from the drop-down menu.

Now you have provisioned SSDs installed in the MSA 2052 as a disk group to be used as a performance tier.

HPE recommends balancing disk groups across both pools as described in the Creating disk groups section of this white paper. For performance tiering, balancing disk groups requires installing a minimum of two more 800 GB SSDs in addition to the two 800 GB SSDs that ship standard with the MSA 2052. With additional SSDs, you can select different RAID levels and create SSD disk groups for each pool.

Configuring the two SSDs into a single pool is supported, but this configuration results in an unbalanced system because only one pool uses SSDs. Customers do not get the full performance benefits of the MSA 2052 in an unbalanced system; only the pool with the SSDs would receive the full performance benefit. For example, if you configure two SSDs in Pool A but no SSDs in Pool B, then any applications with data in Pool B will not have the performance benefit of the SSDs in Pool A.

Automated Tiering

Automated Tiering is a virtual storage feature that automatically moves data residing in one class of disks to a more appropriate class of disks based on data access patterns:

- Frequently accessed hot data can move to disks with better performance, typically lower capacity, and typically higher costs.
- Infrequently accessed cool data can move to disks with lower performance, greater capacity, and typically lower costs per GB.

Each virtual disk group, depending on the type of disks it uses, is automatically assigned to one of the following tiers:

- **Performance:** This highest tier uses SAS SSDs, which provide the best performance but also the highest cost per GB.
- **Standard:** This middle tier uses enterprise-class 10K rpm or 15K rpm spinning SAS disks, which provide good performance with midlevel cost per GB and capacity.
- **Archive:** This lowest tier uses midline 7.2K rpm spinning SAS disks, which provide the lowest performance with the lowest cost per GB and highest capacity.

How Automated Tiering works

Automated Tiering uses virtualized storage and is accomplished by paging. MSA virtual volumes allocate data into small, 4 MB chunks (known as *pages*) from within the virtual storage pool. These pages are ranked based on a sophisticated algorithm. The page rank is used to efficiently select appropriate pages to move between tiers. The result is that pages can be migrated between tiers automatically such that I/Os are optimized in real-time.

In contrast to data movement at the LUN level, Automated Tiering at the sub-LUN level provides highly efficient data movement. Only a minimum amount of CPU and memory resources are needed to support the data movement between tiers; therefore, movement can happen in real-time rather than in offline batch movements.

Automated Tiering concepts

The MSA tiering algorithm runs every five seconds. Pages are ranked, scanned, and migrated during this period. The MSA algorithm performs the following steps:

1. Pages are ranked by access patterns.
2. A scan looks for highly ranked pages.
3. These highly ranked pages are then migrated up the tier levels.
    a. Pages are only migrated down a tier if space is needed for a highly ranked page.
    b. Only 80 MB of data is migrated every five seconds to avoid degrading system throughput.
4. The MSA tiering algorithm is tuned to avoid thrashing or moving pages back and forth between tiers in brief amounts of time.
5. Infrequently accessed or cold data is only moved to lower tiers as capacity is required for more frequently accessed data. This keeps as much of the data in a better-performing tier as possible.
6. To optimize workloads, sequential writes are initially written to the fastest spinning drive tier with free capacity. A random write workload is initially written to the fastest tier with free capacity. In a three-tiered MSA, the random write workload is written to the performance tier.

Note
The information in Steps 5 and 6 explain the No Affinity volume setting. Consult the Volume Tier Affinity section of this white paper to review various virtual volume settings and how they affect Quality of Service (QoS).

Automated Tiering components

You need a license on the arrays to get the benefits of Automated Tiering. The Automated Tiering function is enabled after you install the proper licenses.

Refer to the following table to determine which license is needed.

Table 1. MSA 1050/2050/2052 Automated Tiering components

Automated Tiering component	MSA 1050	MSA 2050	MSA 2052
Archive Tiering (enterprise SAS to midline SAS)	Standard	Standard	Standard
Performance Tiering (SSD to enterprise SAS)	Optional[6]	Optional[6]	Standard[7]

Note
The ADS Suite license ships pre-installed as standard with the MSA 2052. However, the entitlement still must be redeemed.

Advantages of using Automated Tiering

Benefits of Automated Tiering include:

- Because a virtual pool can have multiple virtual disk groups, each belonging to a different tier, a virtual pool can provide multiple tiers of storage, which lowers total cost of ownership by moving less accessed data to the lower-cost-per-GB media.
- The I/O load is automatically balanced between components in a tier, which improves performance and allows for easy expansion.
- Virtual disk groups can be added or removed without disrupting I/O. Data in virtual disk groups that is being removed is automatically migrated to other disk groups in a storage pool if the other disk groups in the storage pool have enough storage space. If there is not enough space, the system will not delete the disk groups until enough free space is available.

Disk group considerations

Allocated pages are evenly distributed among disk groups in a tier; therefore, create all disk groups in a tier with the same RAID type and number of drives to ensure uniform performance in the tier.

Consider an example where the first disk group in the standard tier consists of five 15K enterprise SAS drives in a RAID 5 configuration. To ensure consistent performance in the tier, any additional disk groups for the standard tier should also be a RAID 5 configuration. Adding a new disk group configured with four 10K enterprise SAS drives in a RAID 6 configuration produces inconsistent performance within the tier because of the different characteristics of the disk groups.

For optimal write sequential performance, parity-based disk groups (RAID 5 and RAID 6) should be created with “the power of 2” method. This method means that the number of data drives (nonparity) contained in a disk group should be a power of 2. See Table 2 for details.

[6] The optional license requires the ADS Suite.
[7] All MSA 2052 models ship standard with the ADS Suite. Software titles included in the ADS Suite include Performance Tiering and Archive Tiering software, Volume Copy and 512 snapshots software, and Remote Snap software.

Table 2. The power of 2 method

RAID type	Total drives per disk group	Data drives	Parity drives
RAID 5	3	2	1
RAID 5	5	4	1
RAID 5	9	8	1
RAID 6	4	2	2
RAID 6	6	4	2
RAID 6	10	8	2

Because of the limitation of 16 disk groups per a pool, RAID type should be considered when creating new disk groups. For example, instead of creating multiple RAID 1 disk groups, consider using a larger RAID 10 disk group.

Drive type and capacity considerations when using tiering

All (HDDs) in a tier should be the same type. For example, do not mix 10K rpm and 15K rpm drives in the same standard tier. The SSD performance tier or read cache should be sized correctly to fit the active data size to gain the best performance boost. Refer to the SSD read cache section of this white paper for more information on sizing read cache data.

The MSA 1050/2050/2052 supports all SSDs that shipped with previous MSA 1040/2040/2042 products, so upgrades are supported. SSDs can be used in the read cache or performance tier. All SSDs are classified into the same tier and interoperate. Different SSD saleable parts have differences in both performance and endurance; mixing different parts should be done with care.

The following table lists which disk group configurations are supported in MSA 1050/2050/2052 arrays.

Table 3. Supported disk group types

Different drive types in the same disk group	Supported in MSA 1050/2050/2052
LFF 15K and SFF 15K	Yes
SSD mainstream endurance[8] (ME), mixed-use (MU), and read-intensive (RI)	Yes
SSD and 15K	No
SSD and 10K	No
10K and 15K	Yes, but not recommended

Disk group RAID type considerations

RAID 6 is recommended when using large capacity MDL SAS drives in the archive tier. The added redundancy of RAID 6 protects against data loss in the event of a second disk failure with large MDL SAS drives.

RAID 5 is commonly used for the standard tier where the disks are smaller and faster, resulting in shorter rebuild times. RAID 5 is used in workloads that typically are both random and sequential in nature.

See the Best practices for SSDs section of this white paper for RAID types used in the performance tier and read cache.

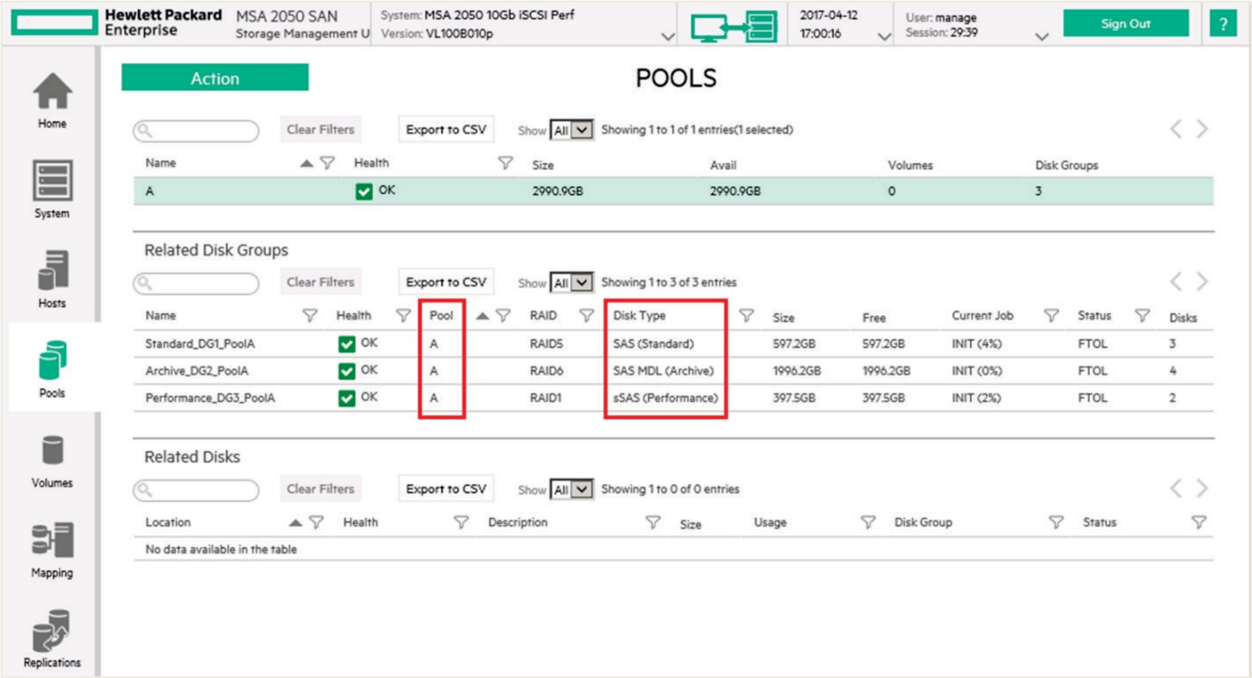
[8] SSD Mainstream Endurance drives are only supported in the MSA 1050/2050/2052 after an upgrade from an MSA 1040/2040/2042.

Creating multiple tiers on one storage pool

When you are creating multiple virtual disk groups, the SMU attempts to alternate disk group ownership between Pools A and B by default. Make sure to select one pool only if you need multiple tiers on one pool.

Important

As shown in Figure 7, make sure to select **Pool A** from the drop-down menu when creating virtual disk groups.



System: MSA 2050 10Gb iSCSI Perf
Version: VL100B010p

2017-04-12
17:00:16

User: manage
Session: 29:39

POOLS

Showing 1 to 1 of 1 entries(1 selected)

Name	Health	Size	Avail	Volumes	Disk Groups
A	✓ OK	2990.9GB	2990.9GB	0	3

Related Disk Groups

Showing 1 to 3 of 3 entries

Name	Health	Pool	RAID	Disk Type	Size	Free	Current Job	Status	Disks
Standard_DG1_PoolA	✓ OK	A	RAID5	SAS (Standard)	597.2GB	597.2GB	INIT (4%)	FTOL	3
Archive_DG2_PoolA	✓ OK	A	RAID6	SAS MDL (Archive)	1996.2GB	1996.2GB	INIT (0%)	FTOL	4
Performance_DG3_PoolA	✓ OK	A	RAID1	sSAS (Performance)	397.5GB	397.5GB	INIT (2%)	FTOL	2

Related Disks

Showing 1 to 0 of 0 entries

Location	Health	Description	Size	Usage	Disk Group	Status
No data available in the table						

Figure 7. MSA virtual pools screen

Volume Tier Affinity

Volume Tier Affinity is an attribute that enables you to define QoS preferences for virtual volumes in a tiered environment. There are three Tier Affinity options:

- Archive—Prefers the lowest tier of service
- Performance—Prefers the higher tiers of service
- No Affinity—Uses the Standard Tiering strategy

Tier Affinity is not the same as tier pinning and does not restrict data to a given tier and capacity. Data on a volume with Archive affinity can still be promoted to a performance tier if that data becomes in demand to the host application.

Note

The Performance affinity does not require an SSD tier and uses the highest performance tier available.

Mechanics of Volume Tier Affinity

Volume Tier Affinity guides the system regarding where to place data from a given volume in the available tiers.

The standard strategy is to prefer the highest spinning disk (non-SSD) tiers for new sequential writes and the highest tier available (including SSD tiers) for new random writes. When data is accessed later by the host application, the data is moved to the most appropriate tier based on demand. Hot data is promoted toward the highest performance tier and cold data is demoted to the lower spinning disk-based tiers. This standard strategy is followed for data on volumes set to **No Affinity**.

For data on volumes set to **Performance** affinity, the standard strategy is followed for all new writes. However, subsequent access to that data will have a lower threshold for promotion, making it more likely for that data to be available on the higher performance tiers.

Preferential treatment is provided to hot data that has the Performance affinity at the SSD tier, making it more likely for Archive or No Affinity data to be demoted out of the SSD tier to make room. This is useful when you know the data will be in demand and want to ensure that it has priority treatment for promotion to and retention in the highest performance tier.

For volumes that are set to Archive affinity, all new writes are initially placed in the archive tier if space is available. If no space is available, new writes are placed on the next highest tier available. Subsequent access to that data will allow for its promotion to the performance tiers as it becomes hot. However, the data will have a lower threshold for demotion and will be moved out of the highest performance SSD tier if there is a need to promote hot data from a lower tier.

Volume Tier Affinity impact in existing environments

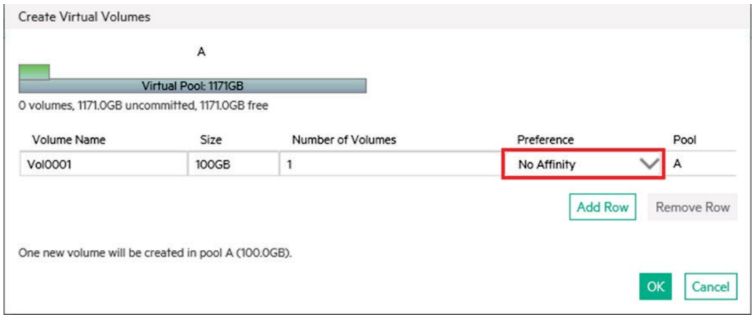
New virtual volumes have a default setting of No Affinity and continue to use the standard tiering strategy.

If the affinity of an existing volume is changed to **Performance**, there is no immediate change made. The current data on the volume is promoted using the performance strategy based on host application needs.

If the affinity of an existing volume is changed to **Archive**, a background operation begins moving data for the affected volume down to the archive tier. This process is performed at a low priority to have minimal effect on host I/O performance. As data in a volume with an Archive affinity becomes hot, it is promoted to the higher tiers automatically. New data written to the volume is targeted to the archive tiers based on the appropriate strategy.

Configuring Volume Tier Affinity

For new virtual volumes, at volume creation time, set the affinity by using the Preference drop-down menu. Note that the default is No Affinity.



Create Virtual Volumes

A

Virtual Pool: 1171GB

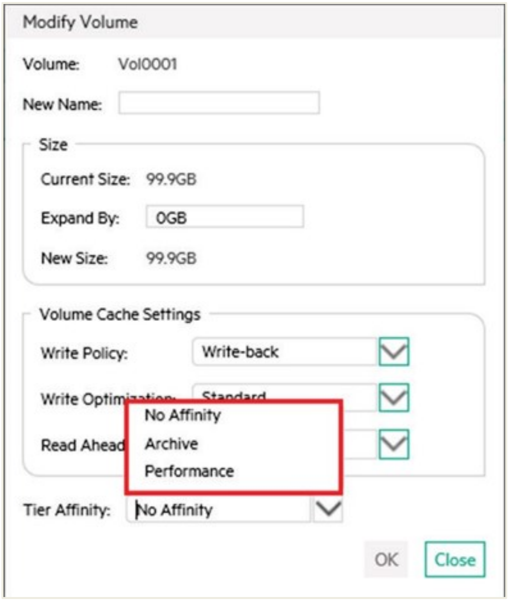
0 volumes, 1171.0GB uncommitted, 1171.0GB free

Volume Name	Size	Number of Volumes	Preference	Pool
Vol0001	100GB	1	No Affinity	A

One new volume will be created in pool A (100.0GB).

Figure 8. Setting Volume Tier Affinity for virtual volumes

For existing virtual volumes, from the SMU, select the **Volumes** tab. Select the volume you want to set an affinity on and from the Action menu, select **Modify Volume**. Then from the Preference drop-down option, set the affinity.



Modify Volume

Volume: Vol0001

New Name:

Size

Current Size: 99.9GB

Expand By: 0GB

New Size: 99.9GB

Volume Cache Settings

Write Policy: Write-back

Write Optimization: Standard

Read Ahead

Tier Affinity: No Affinity

Figure 9. Modifying the Volume Tier Affinity

HPE recommends the default **No Affinity** option for most configurations. This setting attempts to balance the frequency of data access, disk cost, and disk availability by moving the volume's data to the appropriate tier.

If the virtual volume uses mostly random or bursty low-latency workloads such as online transaction processing (OLTP), virtual desktop infrastructure (VDI), or virtualization environments, HPE recommends setting the preference to **Performance**. This setting keeps as much of the volume's data in the performance tier for as long as possible.

If the virtual volume contains infrequently accessed workloads such as backup data or email archiving, HPE recommends setting the preference to **Archive**. This option keeps as much of the volume's data in the archive tier for as long as possible.

Best practices when choosing drives for MSA 1050/2050/2052 storage

The characteristics of applications and workloads are important when selecting drive types for the MSA 1050/2050/2052 array.

Drive types

MSA 1050 and MSA 2052 arrays support SSDs, SAS enterprise drives, and SAS MDL drives. The MSA 2050 and MSA 2052 array supports SSDs, SAS enterprise drives, SAS MDL drives, and SEDs. Selecting the correct drive type is important; you should select the drive types based on the workload and performance requirements of the volumes that will be serviced by the storage system. For sequential workloads, SAS enterprise drives or SAS MDL drives provide a good price-for-performance trade-off over SSDs. If more capacity is needed in a sequential environment, SAS MDL drives are recommended. SAS enterprise drives offer better performance than SAS MDL drives and should also be considered for random workloads when performance is a premium. For high-performance random workloads, SSDs are appropriate.

SAS MDL drives are not recommended for constant high-workload applications. SAS MDL drives are intended for archival purposes.

SSD endurance

MSA fifth-generation arrays support mainstream endurance (ME)[9], mixed-use (MU), and read-intensive (RI) solid-state drives. Drive endurance is not a contributing factor to an array's overall reliability or performance, and HPE has qualified these drives for use with the fifth-generation MSA under all workloads.

Endurance is a measure of how many times per day over a five-year period a drive's full capacity can be re-written. As SSDs grow larger, the ability to write their full capacity within a day diminishes as does drive wear. MSA virtual storage technology further reduces this wear by providing an abstraction layer between application I/O, and the data written to an individual SSD. As a result of these modern technologies, SSD endurance is no longer a concern when drives are installed within fifth-generation MSA arrays.

Using the I/O Workload functionality to determine SSD capacity

A feature named **I/O Workload** has been added to the SMU in VE270/VL270 and later firmware. MSA array controllers keep track of a substantial amount of data pertaining to the I/O dynamics at the page level (4 MB chunks). From this data, the MSA provides visibility into the capacity accessed by a percentage of I/O over a seven-day period. Although some workloads have "transient" data access patterns, many workloads have steady access patterns on small portions of the array's capacity. This behavior produces hot pages in the array that remain hot for a large amount of the array's uptime. You would see substantial benefits if these pages were served from the fastest media in the array, which is ideally an SSD. The MSA's tiering engine works to position the hottest pages on the fastest media at any given time.

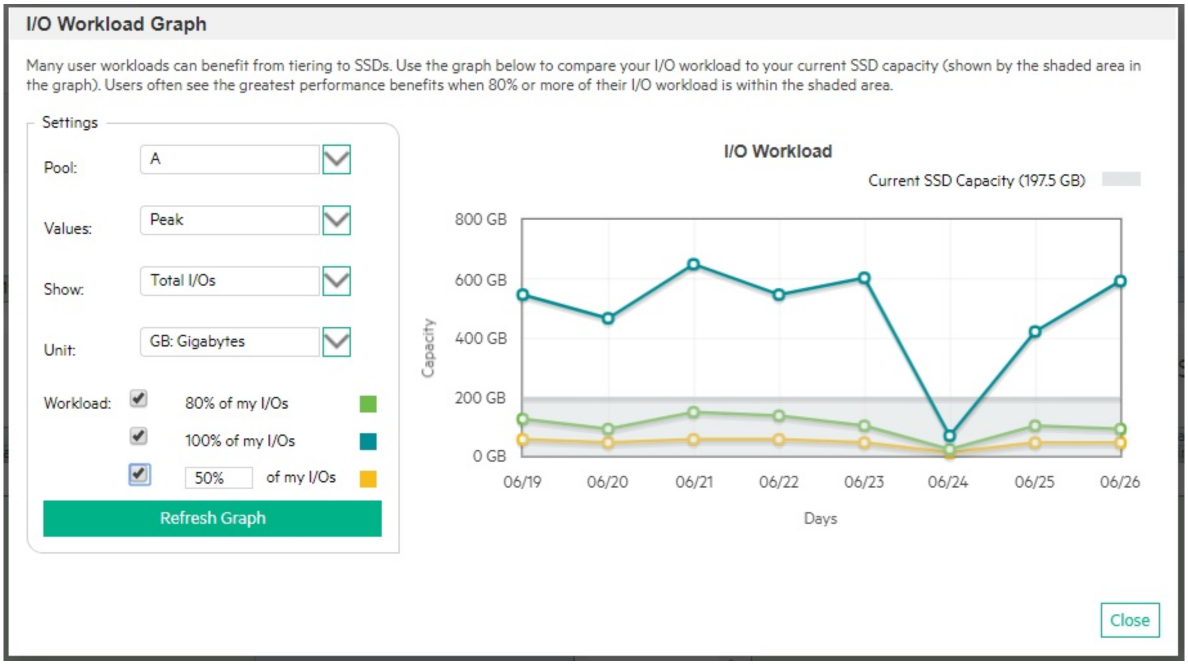
The I/O Workload graph shows the capacity of each selected workload percentage (100%, 80%, or other value). Here are two examples of user scenarios where the I/O Workload graph might be useful, including how to use the graph data:

- **New installation or new SSD installation:** After the MSA array is installed and has had workloads running against it for a week's time, the I/O Workload data provides a representation of what capacity is servicing 100% and 80% of I/O. You can select a custom percentage if needed. In a new installation or in an installation with no SSD tier installed, the 80% line is a reasonable starting point for an SSD tier. Based on SSD RAID settings, you can calculate a starting point for SSD tier sizing based on that week's workload. Although 80% is not a rule, this percentage is a good starting point. Compare these values to the best practice "rule of thumb," which suggests that 5% to 15% of the array's capacity should be SSDs for tiered solutions.
- **Existing SSD tiering or read caching installed and running:** For arrays running with SSDs installed (performance tier or read cache), the I/O Workload graph shows a dotted line that represents the installed SSD capacity. The I/O Workload graphs can be checked periodically to see where the 80% I/O line is compared to the SSD capacity line. Although there are no rules that indicate good or bad situations, you can use the graph with other system performance tools to better understand specific dynamics of the installation and the normal dynamics of a system in the daily activities for a specific environment.

Interpreting the I/O Workload graphs enables customers to strike a balance between the SSD costs compared to performance benefits. For example, some customers might have days where peak usage is far greater than the SSD capacity line because it might be acceptable to have slower performance when the system uses HDDs for a larger percentage of the workload I/O. Their systems might be sized to optimize the cost of storage because of budget constraints. Other customers might want to optimize the system such that a large percentage of daily I/O can reside on SSD media (sized to 80% or 90%).

When combined with other performance monitoring tools, the I/O Workload functionality gives you insight into how workloads and the MSA are working together in a real-world environment. Figure 10 shows a one-week sample sized to 50%, 80%, and 100% I/O using the I/O Workload graph for Pool A. In this example, the 80% graph falls under the SSD capacity, which indicates that the current SSD capacity for Pool A is enough to service the I/O Workload.

[9] Mainstream endurance SSDs are only supported within an MSA 1050/2050/2052 after an upgrade from an MSA 1040/2040/2042.



I/O Workload Graph

Many user workloads can benefit from tiering to SSDs. Use the graph below to compare your I/O workload to your current SSD capacity (shown by the shaded area in the graph). Users often see the greatest performance benefits when 80% or more of their I/O workload is within the shaded area.

Settings

Pool: A

Values: Peak

Show: Total I/Os

Unit: GB: Gigabytes

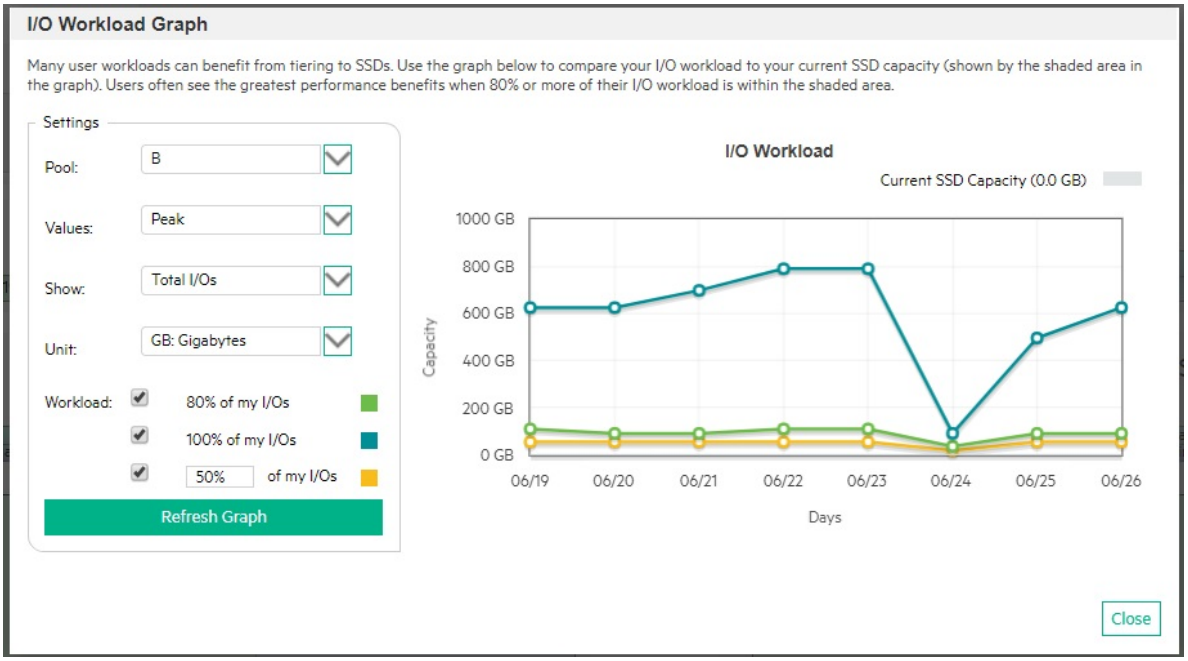
Workload: ☑ 80% of my I/Os 🟩

☑ 100% of my I/Os 🟦

☑ 50% of my I/Os 🟨

Figure 10. I/O Workload graph for Pool A

Figure 11 shows a one-week sample sized to 50%, 80%, and 100% I/O using the I/O Workload graph for Pool B. These graphs show the I/O trend for Pool B.



I/O Workload Graph

Many user workloads can benefit from tiering to SSDs. Use the graph below to compare your I/O workload to your current SSD capacity (shown by the shaded area in the graph). Users often see the greatest performance benefits when 80% or more of their I/O workload is within the shaded area.

Settings

Pool: B

Values: Peak

Show: Total I/Os

Unit: GB: Gigabytes

Workload: ☑ 80% of my I/Os 🟩

☑ 100% of my I/Os 🟦

☑ 50% of my I/Os 🟨

Figure 11. I/O Workload graph for Pool B

To view I/O Workload information, access the footer panel in the SMU. Refer to the MSA 1050/2050 SMU Reference Guide for more details.

Best practices to improve availability

There are many methods to improve availability when using an MSA 1050/2050/2052 array. High availability is always advisable to protect assets in the event of a device failure. This section outlines some options to help you in the event of a failure.

Volume mapping

Using volume mapping correctly can provide high availability from the hosts to the array. For high availability during a controller failover, a volume must be mapped to at least one port accessible by the host on both controllers. HPE recommends that you map a volume via ports on both controllers, which ensures that at least one of the paths is available in the event of a controller failover, thus providing a preferred or optimal path to the volume.

Note
HPE recommends the use of explicit mapping to hosts or initiators, and that default mapping not be used unless unavoidable.

If a controller fails over, the surviving controller reports that it is now the preferred path for all disk groups. When the failed controller is back online, the disk groups and preferred paths switch back to the original owning controller.

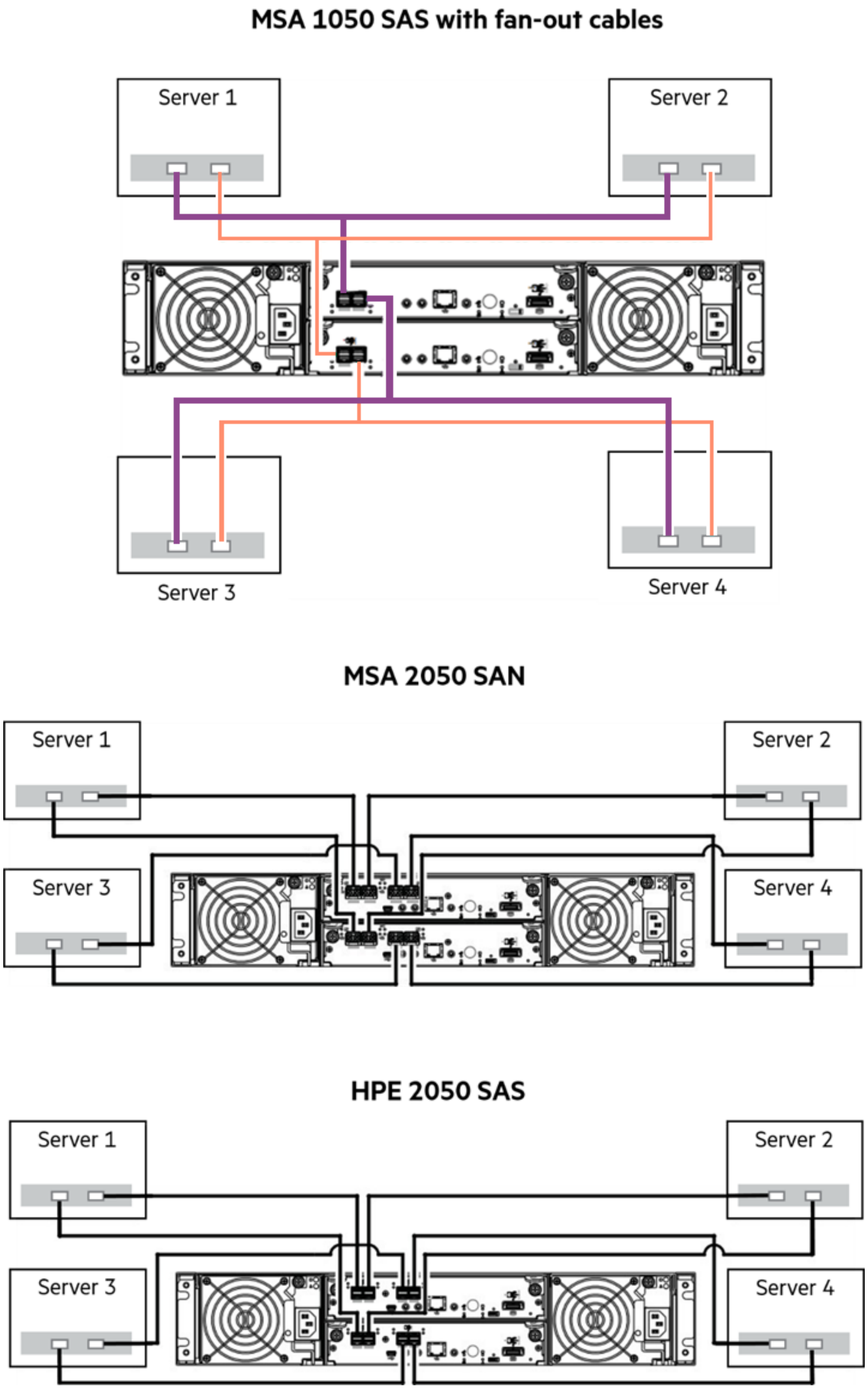
The best practice is to map volumes to two ports on each controller to take advantage of load balancing and redundancy.

Mapping a port creates a mapping to each controller. For example, mapping Port 1 maps host Ports A1 and B1 as well. Mapping Port 2 maps host Ports A2 and B2.

With this in mind, make sure that physical connections are set up correctly on the MSA so that a server has a connection to both controllers on the same port number. For example, on a direct-attached MSA 1050/2050/2052 configuration with multiple servers, make sure that Ports A1 and B1 are connected to Server 1, Ports A2 and B2 are connected to Server 2, and so on.

HPE does not recommend enabling more than eight paths to a single host—that is, two HBA ports on a physical server connected to two ports on the A controller and two ports on the B controller. Enabling more paths from a host to a volume puts additional stress on the operating system's multipath software, which can lead to delayed path recovery in very large configurations.

Note
Volumes should not be mapped to multiple servers at the same time unless the operating systems on the servers are cluster-aware. However, because a server can contain multiple unique initiators, mapping a volume to multiple unique initiators (that are contained in the same server) is supported and recommended. The recommended practice is to put multiple initiators for the same host into a host and map the host to the LUNs, rather than individual maps to initiators.



MSA 1050 SAS with fan-out cables

MSA 2050 SAN

HPE 2050 SAS

Figure 12. Direct attach cabling

Redundant paths

To increase the availability of the array to the hosts, use multiple redundant paths together with multipath software. Redundant paths can also help increase performance from the array to the hosts. Redundant paths can be accomplished in multiple ways. In the case of a SAN-attach configuration, the best practice is to have multiple redundant SAN switches with the hosts having at least one connection into each SAN switch and the array having one or more connections from each controller into each switch.

In the case of a direct attach configuration, the best practice is to have at least two connections to the array for each server. In the case of a direct attach configuration with dual controllers, the best practice is to have at least one connection to each controller.

Multipath software

To fully utilize redundant paths, multipath software should be installed on the hosts. Multipath software allows the host operating system to use all available paths to volumes presented to the host; redundant paths allow hosts to survive SAN component failures. Multipath software can increase performance from the hosts to the array. Table 4 lists supported multipath software by operating system.

Note

More paths are not always better. Enabling more than eight paths to a single volume is not recommended.

Table 4. Multipath and operating systems

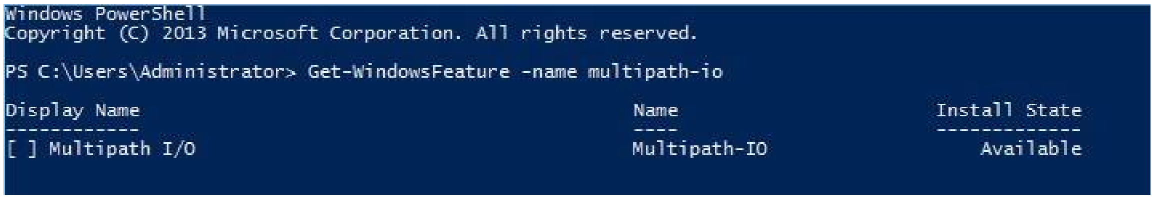
Operating system	Multipath name	Vendor	Product ID
Windows Server® 2012/2016	Microsoft® Multipath I/O (MPIO)	HPE	MSA 1050 SAN MSA 1050 SAS MSA 2050 SAN MSA 2050 SAS
Linux®	DM-Multipath	HPE	MSA 1050 SAN MSA 1050 SAS MSA 2050 SAN MSA 2050 SAS
VMware	Native Multipathing Plug-in (NMP)	HPE	MSA 1050 SAN MSA 1050 SAS MSA 2050 SAN MSA 2050 SAS

Installing MPIO on Windows Server 2012 and 2016

You can use Power Shell commands in Windows Server 2012 and 2016 to install MPIO. Open PowerShell and perform the following steps:

1. Check current installed status of MPIO:

   Get-WindowsFeature –name MultiPath-io

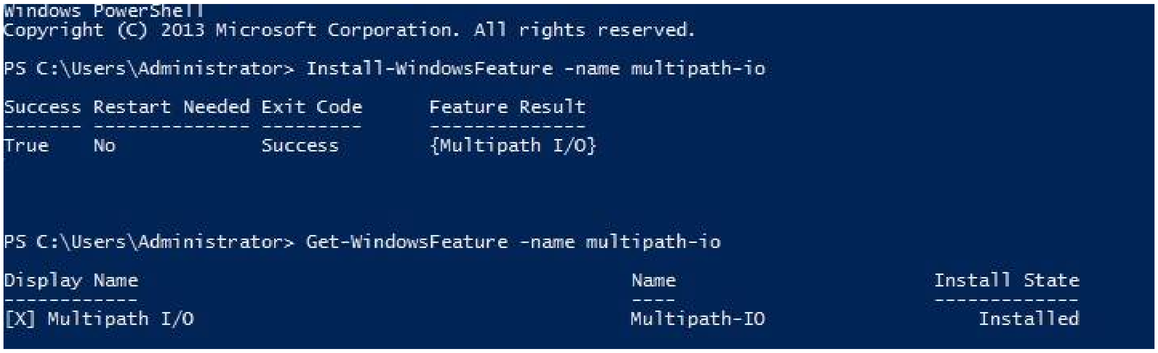


```
Windows PowerShell
Copyright (C) 2013 Microsoft Corporation. All rights reserved.

PS C:\Users\Administrator> Get-WindowsFeature -name multipath-io

Display Name                                            Name                       Install State
------------                                            ----                       -------------
[ ] Multipath I/O                                       Multipath-IO                   Available
```

Figure 13. MSA 1050/2050/2052 MPIO verification command

2. Install MPIO:

   Install-WindowsFeature –name MultiPath-io

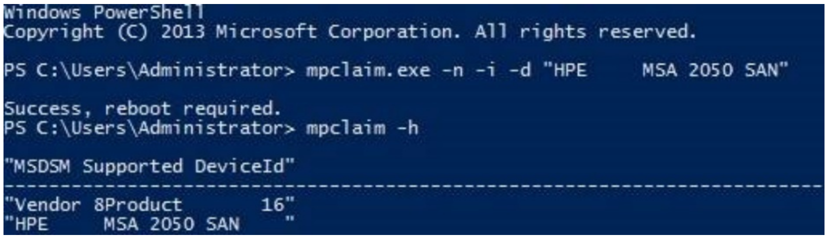


```
Windows PowerShell
Copyright (C) 2013 Microsoft Corporation. All rights reserved.

PS C:\Users\Administrator> Install-WindowsFeature -name multipath-io

Success Restart Needed Exit Code      Feature Result
------- -------------- ---------      --------------
True    No             Success        {Multipath I/O}


PS C:\Users\Administrator> Get-WindowsFeature -name multipath-io

Display Name                                            Name                       Install State
------------                                            ----                       -------------
[X] Multipath I/O                                       Multipath-IO                   Installed
```

Figure 14. MSA 1050/2050/2052 MPIO install and verification commands

3. Claim the devices:

   mpclaim –n –I –d “HPE   MSA 2050 SAN”



```
Windows PowerShell
Copyright (C) 2013 Microsoft Corporation. All rights reserved.

PS C:\Users\Administrator> mpclaim.exe -n -i -d "HPE     MSA 2050 SAN"

Success, reboot required.
PS C:\Users\Administrator> mpclaim -h

"MSDSM Supported DeviceId"
-------------------------------------------------------------------------
"Vendor 8Product       16"
"HPE     MSA 2050 SAN    "
```

Figure 15. MSA 1050/2050/2052 **mpclaim** command

Notes

- A reboot is required after the claim of the devices.
- There are five spaces between HPE and MSA in the **mpclaim** command.
- When running the **mpclaim** command, enter the correct product ID for the MSA product. See Table 4 for more details.

When the MPIO Device Specific Module (DSM) is installed, no further configuration is required. However, after initial installation, you should use Windows Server Device Manager to ensure that MPIO DSM was installed correctly as described in Managing MPIO LUNs.

Long failover times when using MPIO with large numbers of LUNs

Windows servers running MPIO use a default Windows registry PDORemovePeriod setting of 20 seconds. When MPIO is used with a large number of LUNs, this setting can be too brief, causing long failover times that can adversely affect applications.

The Microsoft step-by-step guide Configuring MPIO Timers describes the PDORemovePeriod setting:

“This setting controls the amount of time (in seconds) that the multipath pseudo-LUN will continue to remain in system memory, even after losing all paths to the device. When this timer value is exceeded, pending I/O operations will be failed, and the failure is exposed to the application rather than attempting to continue to recover active paths. This timer is specified in seconds. The default is 20 seconds. The max allowed is MAXULONG.”

If you are using MPIO with a large number of LUNs, one work-around is to edit your registry settings so that HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\PDORemovePeriod is set to a higher value.

Use the following values for MPIO with a large number of LUNs:

- If you are using a Fibre Channel connection to a Windows Server running MPIO, use a value of 90 seconds.
- If you are using an iSCSI connection to a Windows Server running MPIO, use a value of 300 seconds.

Managing MPIO LUNs

The Windows Server Device Manager enables you to display or change devices, paths, and load balance policies. It also allows you to diagnose and troubleshoot the DSM. After initial installation of MPIO DSM, use Device Manager to verify that it has installed correctly.

If MPIO DSM was installed correctly, each MSA 1050/2050/2052 storage volume visible to the host is listed as a multipath disk drive, as shown in Figure 14.

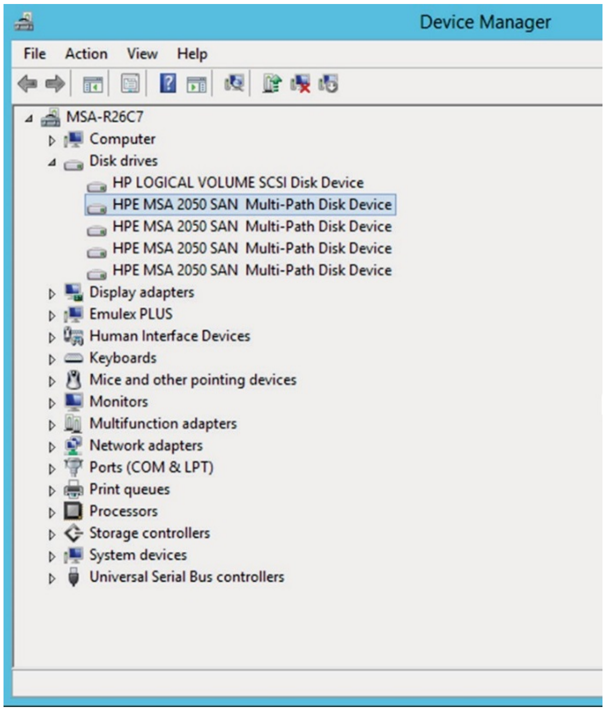
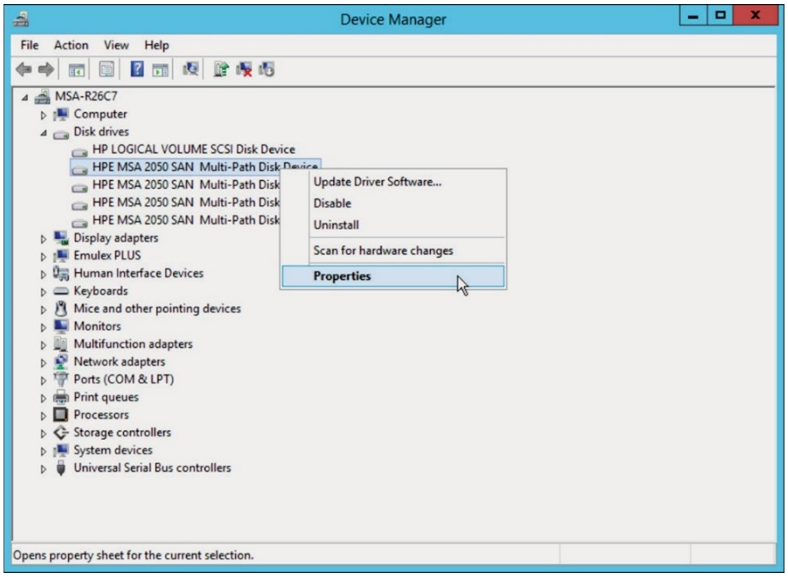

Device Manager
File Action View Help
MSA-R26C7
Computer
Disk drives
HP LOGICAL VOLUME SCSI Disk Device
HPE MSA 2050 SAN Multi-Path Disk Device
HPE MSA 2050 SAN Multi-Path Disk Device
HPE MSA 2050 SAN Multi-Path Disk Device
HPE MSA 2050 SAN Multi-Path Disk Device
Display adapters
Emulex PLUS
Human Interface Devices
Keyboards
Mice and other pointing devices
Monitors
Multifunction adapters
Network adapters
Ports (COM & LPT)
Print queues
Processors
Storage controllers
System devices
Universal Serial Bus controllers

Figure 16. MSA multipath disk devices

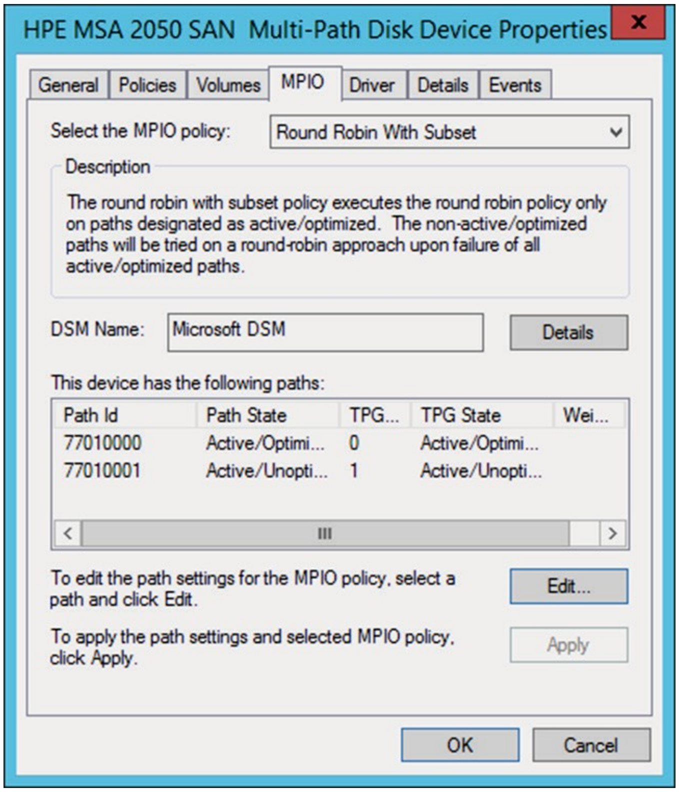
To verify that there are multiple redundant paths to a volume, right-click the multipath disk device and select **Properties**.


Device Manager
File Action View Help
MSA-R26C7
Computer
Disk drives
HP LOGICAL VOLUME SCSI Disk Device
HPE MSA 2050 SAN Multi-Path Disk Device
HPE MSA 2050 SAN Multi-Path Disk
HPE MSA 2050 SAN Multi-Path Disk
HPE MSA 2050 SAN Multi-Path Disk
Update Driver Software...
Disable
Uninstall
Scan for hardware changes
Properties
Display adapters
Emulex PLUS
Human Interface Devices
Keyboards
Mice and other pointing devices
Monitors
Multifunction adapters
Network adapters
Ports (COM & LPT)
Print queues
Processors
Storage controllers
System devices
Universal Serial Bus controllers

Opens property sheet for the current selection.

Figure 17. Selecting Properties for the MSA multipath disk device

Click the **MPIO** tab, which enables you to view or change the load balance policy and view the number of paths and their status.



Select the MPIO policy: Round Robin With Subset

Description

The round robin with subset policy executes the round robin policy only on paths designated as active/optimized. The non-active/optimized paths will be tried on a round-robin approach upon failure of all active/optimized paths.

DSM Name: Microsoft DSM

This device has the following paths:

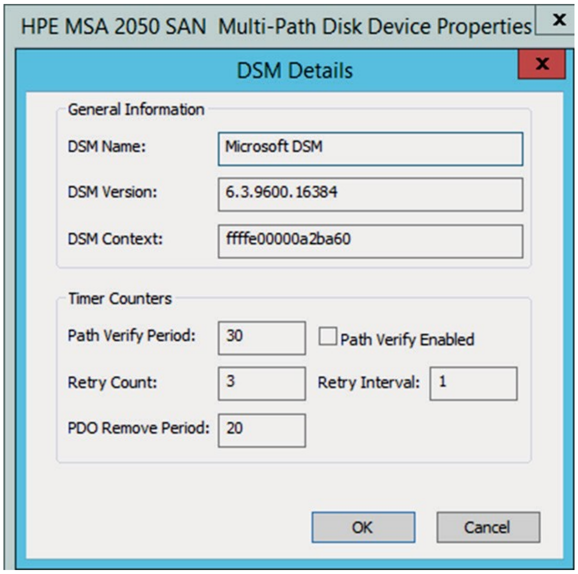
Path Id	Path State	TPG...	TPG State	Wei...
77010000	Active/Optimi...	0	Active/Optimi...	
77010001	Active/Unopti...	1	Active/Unopti...	

To edit the path settings for the MPIO policy, select a path and click Edit.

To apply the path settings and selected MPIO policy, click Apply.

Figure 18. MSA multipath disk device properties—MPIO tab

The Details tab shows additional parameters.



General Information

DSM Name: Microsoft DSM

DSM Version: 6.3.9600.16384

DSM Context: ffffe00000a2ba60

Timer Counters

Path Verify Period: 30 ☐ Path Verify Enabled

Retry Count: 3 Retry Interval: 1

PDO Remove Period: 20

Figure 19. MSA multipath disk device properties—DSM Details

Snapshots

MSA snapshot services enable increased data protection by creating recovery points for data by taking a picture of the data at a specific point in time. Snapshots are then maintained even as data continues to change. In the event of a failure, you can recover to any previous snapshot. Snapshots are a great complement to tape or disk backup strategy.

MSA snapshot functionality is controller based, so host resources are not used. MSA 1050/2050/2052 snapshot services use redirect-on-write capabilities.

HPE recommends using the snapshot functionality for data protection. Review the snapshot space management guidelines in the following sections when using MSA snapshots.

Snapshots on virtual volumes

You need to determine how you manage snapshots on virtual volumes on pools that have overcommit enabled. The MSA offers two options for setting the frequency of snapshot management:

- **Regularly:** If you already maintain snapshots, then you probably do not need to change anything. The system automatically sets the limit at 10% of the pool and only notifies you if a threshold is crossed.
- **Rarely:** If you want the system to actively manage snapshots, including removing old snapshots, consider changing the limit policy to delete. Use the `set snapshot-space` CLI command to manage snapshot space options—the SMU does not provide this ability. Only unmapped snapshots that are leaves of a snapshot tree are considered for deletion. The oldest and lowest priority snapshots are deleted first. Use the `set volume` CLI command to set the retention priority on snapshots—the SMU does not provide this ability.

On pools with overcommit disabled, snapshots are fully provisioned, allocating storage equal to the size of the volume.

Verify the rate of change of your data

Based on the rate of change in the data and the necessary snapshot retention, adjust the snapshot space limit and thresholds accordingly.

Note

Retention policies apply to individual volumes and are inherited. Setting the retention level of a base volume does not affect existing snapshots of that volume. Rather, it affects the snapshots created after setting the retention level.

Consult the HPE MSA 1050/2050 CLI Reference Guide to change snapshot space and snapshot retention policies.

Dual power supplies

The MSA 1050/2050/2052 array enclosure and supported disk enclosures ship with dual power supplies. At a minimum, connect both power supplies in all enclosures. For the highest level of availability, connect the power supplies to separate power sources.

Reverse cabling of disk enclosures

MSA 1050/2050/2052 firmware supports both fault-tolerant (reverse) cabling and straight-through SAS cabling of disk enclosures.

Fault-tolerant cabling allows any disk enclosure to fail or be removed without losing access to other disk enclosures in the chain. For the highest level of fault tolerance, use fault-tolerant (reverse) cabling when connecting disk enclosures.

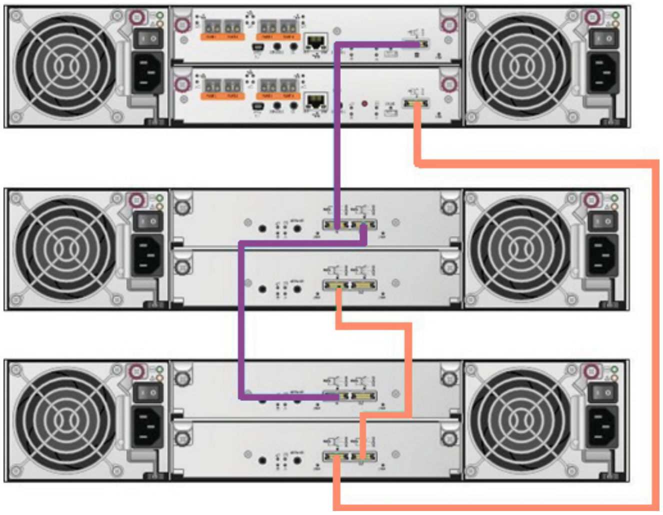


Figure 20. Reverse cabling example using the MSA 2050 system

See the MSA Cable Configuration Guide for more details on cabling the MSA 1050/2050/2052. You can find the HPE MSA 1050/2050 Cable Configuration Guides on the support pages for the MSA 1050 and MSA 2050.

Create disk groups across disk enclosures

HPE recommends striping disk groups across shelf enclosures to enable data integrity in the event of an enclosure failure. A disk group created with RAID 1, 10, 5, or 6 can sustain one or more disk enclosure failures without loss of data depending on RAID type. Disk group configuration should consider MSA drive sparing methods such as global and dynamic sparing.

Drive sparing

HPE strongly recommends drive sparing to help protect data in the event of a disk failure in a fault-tolerant disk group (RAID 1, 5, 6, or 10) configuration. In the event of a disk failure, the array automatically attempts to reconstruct the data from the failed drive to a compatible spare. A compatible spare is defined as a drive that has sufficient capacity to replace the failed disk and is the same media type (that is, SAS SSD, enterprise SAS, midline SAS, or self-encrypting drives). The MSA 1050/2050/2052 supports global and dynamic sparing. The MSA 1050/2050/2052 reconstructs a critical or degraded disk group.

Important

An offline or quarantined disk group is not protected by sparing.

Supported spare types are:

- **Global spare:** This type is reserved for use by any fault-tolerant disk group to replace a failed disk. The array supports up to 64 global spares per system. At least one disk group must exist before you can add a global spare.
- **Dynamic spare:** All available drives are available for sparing. If the storage system has available drives and a disk group becomes degraded, any available drive can be used for disk group reconstruction.

Sparing process

When a disk fails in a redundant disk group, the system first looks for a compatible global spare. If the system does not find a compatible global spare and the dynamic spare option is enabled, the system uses any available compatible disk for the spare. If no compatible disk is available, reconstruction cannot start.

During data reconstruction, the affected disk group is in either a degraded or critical status until the parity data is completely written to the spare, at which time the disk group returns to fault-tolerant status. In the case of global spares, after the failed drive is replaced, the replacement drive needs to be added back as a global spare.

The best practice for sparing is to configure at least one spare for every fault-tolerant disk group in the system. Dynamic sparing is enabled by default.

Note
Warnings and alerts are sent when the last global spare is used in a system.

Drive replacement

In the event of a drive failure, replace the failed drive with a compatible drive as soon as possible. If global sparing is in use, mark the new drive as a global spare so it can be used in the future for any other drive failures.

Implement Remote Snap replication

MSA 1050/2050/2052 Remote Snap software is a form of asynchronous replication that replicates block-level data from a volume on a local system to a volume on a second independent system. The second system can be at the same location as the first or located at a remote site.

The MSA 2050/2052 supports up to four peer connections; the MSA 1050 supports only one. The best practice is to implement Remote Snap replication for disaster recovery.

Use the secured web access (HTTPS) when using Remote Snap replication on the MSA.

Note
You must purchase a license to implement Remote Snap. The ADS Suite includes Remote Snap software but requires a license key from HPE and must be installed on the MSA 2052 to enable Remote Snap.

To obtain a Remote Snap license, go to My License Portal.

For more information, consult the MSA Remote Snap Software technical white paper.

Best practices to enhance performance

This section outlines configuration options for enhancing array performance.

Cache settings

One method to tune the storage system is by choosing the correct cache settings for the volumes. You can set controller cache options for individual volumes to improve I/O performance.

Caution
Only disable write-back caching if you fully understand how the host operating system, application, and adapter move data. If this feature is used incorrectly, you might degrade system performance.

Using write-back or write-through caching

By default, volume write-back cache is enabled. Because controller cache is backed by super-capacitor technology, if the system loses power, data is not lost. For most applications, enabling write-back cache is the best practice.

You can change a volume's write-back cache setting. Write-back is a cache-writing strategy in which the controller receives the data to be written to disks, stores it in the memory buffer, and immediately sends the host operating system a signal that the write operation is complete, without waiting until the data is actually written to the disk. The storage system mirrors all the data from one controller module cache to the other so data can be written to disk even after a controller failover. Write-back cache improves the performance of write operations and the throughput of the controller. This is especially true in the case of random I/O, where write-back caching allows the array to coalesce the I/O to the volumes.

When write-back cache is disabled, write-through becomes the cache-writing strategy. Using write-through cache, the controller writes the data to the disks before signaling to the host operating system that the process is complete. Write-through cache has lower write operation and throughput performance than write-back, but all data is written to nonvolatile storage before confirmation to the host. You can set conditions that cause the controller to change from write-back caching to write-through caching. See the MSA 1050/2050 SMU Reference Guide for ways to set the auto write-through conditions correctly. In most situations, the default settings are acceptable.

In both caching strategies, active/active failover of the controllers is enabled.

Optimizing read-ahead caching

You can optimize a volume for sequential reads or streaming data by changing its read-ahead cache settings. Read ahead is triggered by sequential accesses to consecutive logical block addressing (LBA) ranges. Read ahead can be forward (that is, increasing LBAs) or reverse (that is, decreasing LBAs). Increasing the read-ahead cache size can greatly improve performance for multiple sequential read streams. However, increasing read-ahead size will likely decrease random read performance.

Use the following read-ahead caching options for volumes:

- **Adaptive:** This option works well for most applications. It enables adaptive read ahead, which allows the controller to dynamically calculate the optimum read-ahead size for the current workload. This is the default.
- **Stripe:** This option sets the read-ahead size to one stripe. The controllers treat non-RAID and RAID 1 disk groups internally as if they have a stripe size of 512 KB, even though they are not striped.
- **Specific size options:** These options let you select an amount of data for all accesses.
- **Disabled:** This option turns off read-ahead cache. This is useful if the host is triggering read ahead for what are random accesses. This can happen if the host breaks up the random I/O into two smaller reads, triggering read ahead.

Caution

Only change read-ahead cache settings if you fully understand how the host operating system, application, and adapter move data.

Optimizing cache modes

You can also change the optimization mode for each volume. The two modes are:

- **Standard:** This mode works well for typical applications where accesses are a combination of sequential and random. This method is the default. For example, use this mode for transaction-based and database update applications that write small files in random order.
- **No-mirror:** In this mode, each controller stops mirroring its cache metadata to the partner controller. This improves write I/O response time but at the risk of losing data during a failover. Unified LUN Presentation (ULP) behavior is not affected with the exception that during failover any write data in cache will be lost. In most cases, no-mirror is not recommended and should only be used after careful consideration.

Parameter settings for performance optimization

You can configure storage systems to optimize performance for specific applications by setting the parameters listed in Table 5. This section provides a basic starting point for fine-tuning the system, which should be done during performance baseline modeling.

Table 5. Optimizing performance for various applications

Application	RAID level	Read-ahead cache size	Cache write optimization
Default	5 or 6	Adaptive	Standard
High-Performance Computing	5 or 6	Adaptive	Standard
Mail spooling	1	Adaptive	Standard
NFS_Mirror	1	Adaptive	Standard
Oracle_DSS	5 or 6	Adaptive	Standard
Oracle_OLTP	5 or 6	Adaptive	Standard
Oracle_OLTP_HA	10	Adaptive	Standard
Random 1	1	Stripe	Standard
Random 5	5 or 6	Stripe	Standard
Sequential	5 or 6	Adaptive	Standard
Sybase_DSS	5 or 6	Adaptive	Standard
Sybase_OLTP	5 or 6	Adaptive	Standard
Sybase_OLTP_HA	10	Adaptive	Standard
Video streaming	1, 5, or 6	Adaptive	Standard
Exchange database	5 for data; 10 for logs	Adaptive	Standard
SAP®	10	Adaptive	Standard
SQL	5 for data; 10 for logs	Adaptive	Standard

Host port utilization for maximum performance

Maximum array performance might require using multiple host ports per controller, as well as using multiple ports on the hosts accessing the array. For example, a single 16 Gb Fibre Channel port can produce only about 1.6 GB/s throughput. Likewise, a 1 Gb iSCSI port can produce only about 110 MB/s, which not only limits sequential throughput, but limits IOPS per port as well.

A 1 Gb host port running at 110 MB/s unidirectionally with an 8 KB I/O size only achieves about 14,200 IOPS on that single port, and a maximum of only 113,600 IOPS with all host ports being used.

Other methods to enhance array performance

There are other methods to enhance performance of the MSA 1050/2050/2052. In addition to optimizing cache settings, the performance of the array can be maximized by using the following techniques.

Place higher-performance SSD and SAS drives in the array enclosure

The MSA 1050/2050/2052 controller is designed to have a single SAS link per drive in the array enclosure and only four SAS links to disk enclosures. Placing better-performance drives (that is, SSD and enterprise SAS drives) in the array enclosure allows the controller to utilize the performance of those drives more effectively than if they were placed in disk enclosures. This process helps generate better overall performance.

Fastest throughput optimization

The following guidelines list the general best practices to follow when configuring the storage system for fastest throughput:

- Configure host ports to match the fastest speed that the infrastructure supports.
- Balance disk groups between the two controllers.
- Balance disk drives between the two controllers.
- Set cache settings to match Table 5 for the application.
- Distribute the load across as many drives as possible.
- Distribute the load across multiple array controller host ports.

Creating disk groups

When you create disk groups, the best practice is to add them evenly across both pools. With at least one disk group assigned to each controller, both controllers are active. This active/active controller configuration allows maximum use of a dual-controller configuration's resources.

Choosing the appropriate RAID levels

Choosing the correct RAID level when creating disk groups can be important for performance. However, there are some trade-offs with cost when using the more fault-tolerant RAID levels.

See Table 6 for the strengths and weaknesses of the supported MSA 1050/2050/2052 RAID types.

Table 6. MSA 1050/2050/2052 RAID levels

RAID level	Minimum disks	Allowable disks	Description	Strengths	Weaknesses
1	2	2	Disk mirroring	• Very high performance and data protection • Minimal penalty on write performance • Protection against single disk failure	High redundancy cost overhead; requires twice the storage capacity because all data is duplicated
5	3	16	Block-level data striping with distributed parity	• Best cost/performance for transaction-oriented networks • Very high performance and data protection • Support for multiple simultaneous reads and writes • Optional optimization for large, sequential requests • Protection against single disk failure	Slower write performance than RAID 0 or RAID 1
6	4	16	Block-level data striping with double distributed parity	• Best suited for large sequential workloads • Similar nonsequential read and sequential read/write performance as RAID 5 • Protection against dual disk failure	• Higher redundancy cost than RAID 5 because the parity overhead is twice that of RAID 5 • Not well-suited for transaction-oriented network applications • Slower nonsequential write performance than RAID 5
10 (1+0)	4	16	Stripes data across multiple RAID 1 sub-disk groups	Highest performance and data protection (protection against multiple disk failures)	• High redundancy cost overhead; requires twice the storage capacity because all data is duplicated • Requires a minimum of four disks

Note

You can only create non-RAID (NRAID) and RAID 0 when setting read cache. NRAID and RAID 0 are used with read cache when the data in the read cache SSDs is duplicated on either the standard or archive tier.

Volume mapping

For increased performance, access the volumes from the ports on the controller that owns the disk group, which is the preferred path. Accessing the volumes on the nonpreferred path results in a slight performance degradation.

Optimum performance with MPIO can be achieved with volumes mapped to multiple paths on both controllers. When the appropriate MPIO drivers are installed on the host, only the preferred (optimized) paths are used. The nonoptimized paths are reserved for failover.

Best practices for SSDs

The performance capabilities of SSDs offer an excellent alternative to traditional spinning HDDs in highly random workloads. SSDs cost more in terms of dollars per GB throughput than spinning HDDs; however, SSDs cost much less in terms of dollars per IOPS. Keep this in mind when choosing the numbers of SSDs per MSA 1050/2050/2052 array.

Use SSDs for randomly accessed data

The use of SSDs can greatly enhance array performance. Because there are no moving parts in the drives, data that is random in nature can be accessed much faster.

If you have the Performance Tier license or if the storage system includes only SSDs, you can use SSDs for virtual disk groups. When SSDs are combined with virtual disk groups that consist of other disk classes, improved read and write performance is possible through automated tiered storage. Alternatively, you can use one or two SSDs in read-cache disk groups to increase read performance for pools without a performance tier. The application workload of a system determines the percentage of total disk capacity that should be SSDs for best performance.

Data such as database indexes and TempDB files benefit from SSDs because this type of data is accessed randomly. Another good example of a workload that benefits from the use of SSDs is desktop virtualization. A specific example is VDI, where boot storms require high performance with low latency.

SSD and performance

Some performance characteristics can be met with linear scaling of SSDs. There are also bandwidth limits in MSA 1050/2050/2052 controllers. There is a point where these two curves intersect. At the intersecting point, additional SSDs do not increase performance. See Figure 19.

The MSA 1050/2050/2052 reaches this bandwidth at a low number of SSDs. For the best performance using SSDs on MSA 1050/2050/2052 arrays, use a minimum of four SSDs with one mirrored pair of drives (RAID 1) per controller. RAID 5 and RAID 6 are also good choices for SSDs but require more drives using the best practice of having one disk group owned by each controller. This would require six SSDs for RAID 5 and eight SSDs for RAID 6.

Base the number of SSDs used on the amount of space needed for the highly random, high-performance data set. For example, if the amount of data that must reside in the SSD volumes exceeds a RAID 1 configuration, use a RAID 5 configuration.

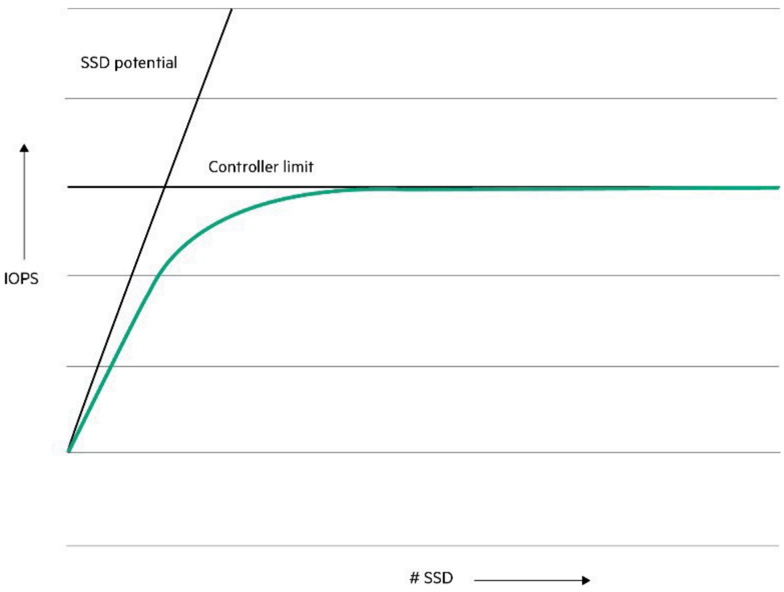


Figure 21. SSD performance potential and the MSA 1050/2050/2052 controller limit

Note
There is no limit to the number of SSDs that can be used in the MSA 1050/2050/2052 array system.

SSD read cache

SSD read cache is a feature that extends the MSA 1050/2050/2052 controller cache.

Read cache is most effective for workloads that are high in random reads. You should size the read cache capacity based on the size of the hot data being randomly read. A maximum of two SSDs per pool can be added for read cache.

HPE recommends beginning with one SSD assigned per storage pool for read cache. Monitor the performance of the read cache and add more SSDs as needed.

There is a 4 TB maximum limit per pool for read cache.

Note
You can have SSDs in a fault-tolerant disk group as a performance tier or as a non-fault-tolerant (up to two disks) disk group as read cache. However, neither pool can have both a performance tier and a read cache. For example, Pool A can have a performance tier and Pool B can have a read cache.

SSD wear gauge

SSDs can be written and erased a limited number of times because of the memory cells on the drives. The SSDs in the MSA 1050/2050/2052 include a wear gauge as well as appropriate events that are generated to help detect the failure. When the wear gauge reaches 0%, data integrity is not guaranteed.

The best practice to prevent data integrity issues is to replace the SSD when the events and gauge indicate less than 5% life remains.

Full disk encryption

Full disk encryption is a data security feature used to protect data on disks that are removed from a storage array. FDE uses special SEDs to secure user data. FDE functionality is available only on the MSA 2050, which supports all MSA features with FDE enabled.

Note
If you plan to use FDE and Performance Tiering, then you must install an ADS license.

The SED is a drive with a circuit built into the drive's controller chipset, which encrypts and decrypts all data to and from the media automatically. The encryption is part of a hash code, which is stored internally on the drive's physical medium. In the event of a drive failure or improper drive removal, you must enter a proper key sequence to gain access to the data stored within the drive.

Full disk encryption on the MSA 2050

The MSA 2050 uses a passphrase to generate a lock key to enable securing the entire storage system. All drives in an FDE-secured system must be SEDs (FDE capable). By default, a system and SED are not secured and all data on the disk can be read or written by any controller. The encryption on the SED conforms to Federal Information Processing Standard (FIPS) 140-2.

To secure an MSA 2050, you must set a passphrase to generate a lock key and then FDE secures the system. Simply setting the passphrase does not secure the system. After an MSA 2050 system has been secured, all subsequently installed disks will be secured automatically using the system lock key. Non-FDE-capable drives are unusable in a secured MSA 2050 system. Unusable drives illuminate amber to indicate an issue.

All MSA 2050 storage systems generate the same lock key with the same passphrase. It is recommended that you use a different passphrase on each FDE-secured system.

Important
Full disk encryption is not supported on the MSA 1050 or 2052.

Caution
The system passphrase should be saved in a secure location. Loss of the passphrase could result in loss of all data on the MSA 2050.

If you are moving the entire storage system, HPE recommends clearing the FDE keys before system shutdown. This locks all data on the disks in case of loss during shipment. Only clear the keys after a backup is available and the passphrase is known. When the system is in the new location, enter the passphrase and the SEDs will be unlocked with all data available.

SEDs that fail in an FDE-secured system can be removed and replaced. Data on the drive is encrypted and cannot be read without the correct passphrase.

Best practices for firmware updates

This section details common firmware update best practices for MSA 1050/2050/2052 arrays. You can download firmware from HPE MSA Storage Firmware and Release Notes.

General MSA 1050/2050/2052 device firmware updates

Follow these best practices when performing device firmware updates:

- As with any other firmware upgrade, ensure that you have a full backup before upgrading.
- Before upgrading the firmware, make sure that the storage system configuration is stable and is not being reconfigured or changed in any way. If any changes are in progress, monitor them using the SMU or CLI. Wait until they are complete before proceeding with the upgrade.
- Do not power cycle or restart devices during a firmware update. If the update is interrupted or there is a power failure, the module could become inoperative. Should this happen, contact HPE customer support.

- After the device firmware update process is completed, confirm that the new firmware version is displayed correctly by using one of the MSA management interfaces—for example, SMU or CLI.

MSA 1050/2050/2052 array controller or I/O module firmware updates

Follow these best practices when performing array controller or I/O module firmware updates:

- HPE recommends that both array controllers run the same firmware version and that the Partner Firmware Update (PFU) setting is enabled.
- HPE recommends that both I/O modules in all enclosures run the same firmware.
- You can update the array controller (or I/O module) firmware in redundant controller systems in online mode only. You can check the firmware upgrade readiness on the system with the CLI command check firmware-upgrade-health.
- When planning for a firmware upgrade, schedule an appropriate time to perform an online upgrade. Because the online firmware upgrade is performed while host I/O is being processed, I/O load can impact the upgrade process. Select a period of low I/O activity to ensure that the upgrade completes as quickly as possible and to avoid disruptions to hosts and applications because of time-outs.
- When planning for a firmware upgrade, allow sufficient time for the update.
  - Plan for a firmware upgrade to take at least 30 minutes; large configurations or heavy I/O workload during upgrades can extend this time.
  - When upgrading or downgrading the firmware, ensure that the Ethernet connection of the management port on each storage controller is available and accessible before starting the process.
  - When using an HPE Smart Component firmware package, the Smart Component process automatically uses best practices to upgrade the controllers in the system. HPE recommends using a Smart Component.

MSA 1050/2050/2052 disk drive firmware updates

Follow these best practices when performing disk drive updates:

- Upgrading disk drives on the MSA 1050/2050/2052 storage system is an offline process. All host and array I/O must be stopped before the upgrade.
- If the drive is in a disk group, verify that it is not being initialized, expanded, reconstructed, verified, or scrubbed. If any of these tasks are in progress, wait for the task to complete or terminate it before performing the update. Also, verify that background scrub is disabled so that it does not start. You can determine this using SMU or CLI interfaces. If you are using a firmware Smart Component, it would fail and report if any of these prerequisites are not met.
- Disk drives of the same model in the storage system must have the same firmware revision. If using a firmware Smart Component, the installer ensures all the drives are updated.

Miscellaneous best practices

This section details various best practices for the MSA 1050/2050/2052.

Boot from storage considerations

When booting from storage on the MSA 1050/2050/2052, set the Volume Tier Affinity to **Performance** even if an SSD performance tier is not available. Selecting this option minimizes operating system boot latencies and avoids delays on the operating system.

Even though the Volume Tier Affinity is set to **Performance**, the data on the boot drive can move to other tiers if it is not accessed.

8 Gb/16 Gb switches and small form-factor pluggable transceivers

The MSA 2050/2052 storage system uses specific small form-factor pluggable (SFP) transceivers that do not operate in HPE 8 Gb or 16 Gb switches. Likewise, HPE Fibre Channel switches use SFPs, which do not operate in the MSA 2050/2052.

The MSA 2050/2052 controllers do not include SFPs. Qualified SFPs for the MSA 2050/2052 are available for separate purchase in four packs. Both 8 Gb and 16 Gb SFPs are available to meet customer needs and budget constraints. All SFPs in an MSA 2050/2052 array should conform to the installation guidelines given in the product QuickSpecs. SFP speeds and protocols can be mixed, but only in the configurations specified in the QuickSpecs for the MSA 1050, MSA 2050, and MSA 2052.

In the unlikely event of an MSA 2050/2052 controller or SFP failure, a field-replaceable unit (FRU) is available. SFPs must be moved from the failed controller to the replacement controller. Refer to the HPE Transceiver Replacement Instructions document for more details.

MSA 1050/2050/2052 iSCSI considerations

When you are using an MSA 1050/2050/2052 SAN controller in an iSCSI configuration, it is a best practice to use at least three network ports per server: two for the storage (private) LAN, and one or more for the public LANs. This ensures that the storage network is isolated from the other networks.

The private LAN is the network that goes from the server to the MSA 1050/2050/2052 SAN controller. This private LAN is the storage network. The public LAN is used for management of the MSA 1050/2050/2052. Isolating the storage and public networks from each other improves performance and increases security.

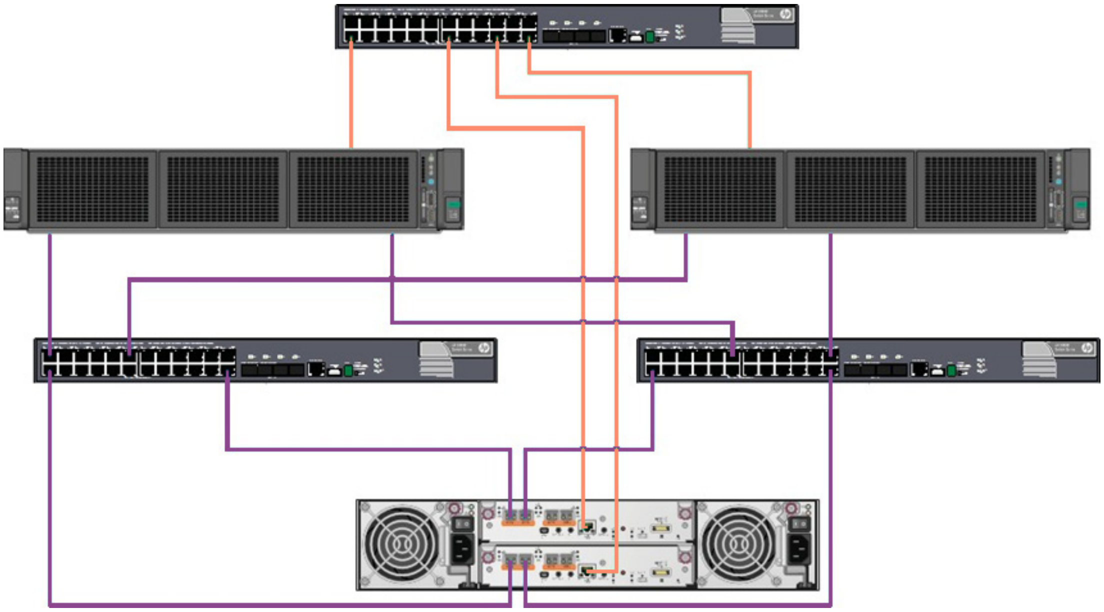


Figure 22. MSA 1050/2050/2052 SAN iSCSI network

IP address scheme for the controller pair

The MSA 2050 SAN controller in iSCSI configurations should have ports on each controller in the same subnets to enable preferred path failover. The suggested means of doing this is to vertically combine ports into subnets. For example, with a subnet mask of 255.255.255.0, the MSA 2050 SAN would be configured as follows:

- Controller A Port 1: 10.10.10.100
- Controller A Port 2: 10.11.10.110
- Controller A Port 3: 10.10.10.120
- Controller A Port 4: 10.11.10.130
- Controller B Port 1: 10.10.10.140

- Controller B Port 2: 10.11.10.150
- Controller B Port 3: 10.10.10.160
- Controller B Port 4: 10.11.10.170

Jumbo frames

A normal Ethernet frame can contain 1500 bytes, whereas a jumbo frame can contain a maximum of 9000 bytes for larger data transfers. The MSA 1050/2050/2052 storage system reserves some of this frame size. The current maximum frame size is 1400 for a normal frame and 8900 for a jumbo frame. This frame maximum can change without notification. If you are using jumbo frames, make sure to enable jumbo frames on all network components in the data path.

Hot-adding disk enclosures

This section details the steps required for upgrading both fault-tolerant and straight-through configurations while the MSA 1050/2050/2052 storage system is online and operational. You can accomplish this without removing the storage from use. Adding disk enclosures while the system is online provides a transparent upgrade path with no disruption of service and ensures data integrity throughout the process.

Adding storage to the existing MSA 1050/2050/2052 storage systems does not significantly impact current data center operations. The process should be performed during a maintenance window where I/O to the affected storage systems will be minimized.

Differences in disk enclosures

There are two types of disk enclosures that you can choose for an MSA 1050/2050/2052 storage system, depending on whether you are using SFF or LFF drives. The two disk enclosure examples are shown in Figure 23 and Figure 24. Note that the backs of the SFF and LFF disk enclosures look the same.

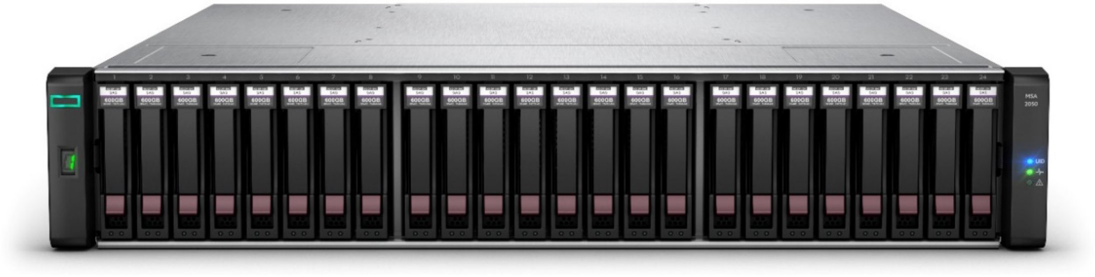


Figure 23. MSA 2050 SFF disk enclosure (without the bezel)

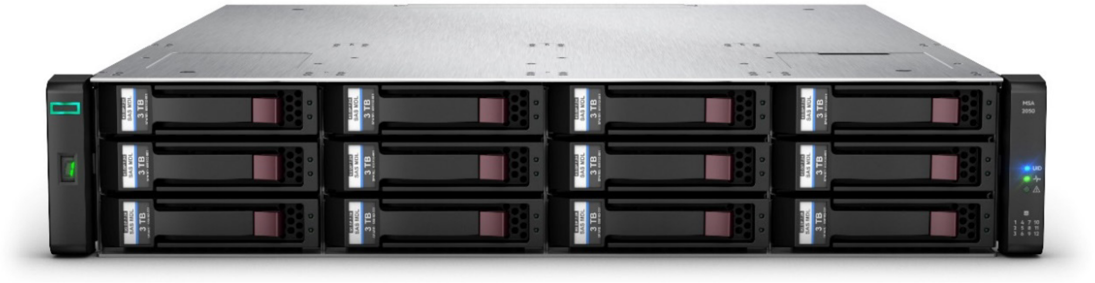


Figure 24. MSA 2050 LFF disk enclosure (without the bezel)

The common feature to note in these examples is that both disk enclosures have two SAS ports per I/O module. The left port on each module is the In port and the right port is the Out port.

Determining the current cabling configuration

If the storage system to be upgraded is an MSA 1050/2050/2052 single-shelf system, adding the second shelf needs no special considerations. This section addresses disk enclosure upgrades where:

- You are adding disk enclosures to an MSA 1050/2050/2052 with at least one disk enclosure.
- You are changing from a straight-through cabling configuration to a fault-tolerant cabling design. To determine the current cabling configuration, refer to Figure 25 and Figure 26.

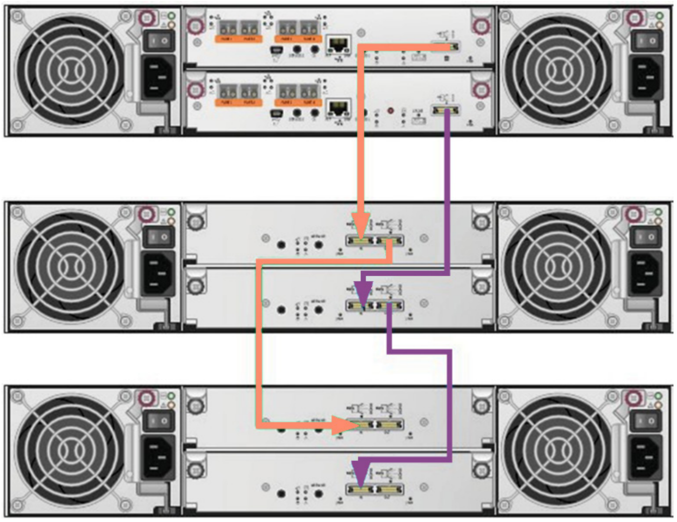


Figure 25. Straight-through cabling configuration

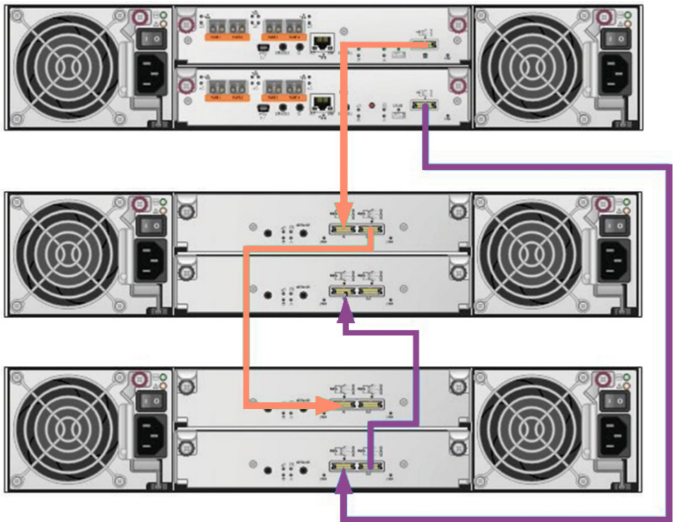


Figure 26. Fault-tolerant cabling configuration

Implementing the upgrade

To ensure data integrity, it is important to determine the type of configuration that is currently implemented and what type of end configuration will be needed going forward. The following steps refer you to the correct section of this white paper to ensure that during the upgrade, there is no risk to data integrity.

If you are currently using:

- An MSA 1050/2050/2052 with a straight-through cabling configuration (Figure 25) and want to add one or more disk enclosures but continue to use the same cabling configuration, refer to the section titled Extending the straight-through cabling configuration.
- An MSA 1050/2050/2052 with a straight-through cabling configuration (Figure 25) and want to reconfigure the system with a fault-tolerant cabling configuration or add a disk enclosure by implementing a fault-tolerant configuration, read the section titled Upgrading from straight-through cabling to a fault-tolerant cabling configuration.
- An MSA 1050/2050/2052 currently configured for fault-tolerance (Figure 26) and need additional disk enclosures, read the section titled Adding storage to an existing fault-tolerant environment.

Extending the straight-through cabling configuration

Perform the following steps to complete the upgrade:

1. Install and mount the new disk enclosure and then install all drives.
2. Power on the new disk enclosure.
3. Use the SMU to ensure that all currently connected disk enclosures are powered on and functioning properly.
4. Connect the new SAS cable (Figure 27, orange cable) from the A port (Out) of the current MSA storage system to the A port (In) of the new storage enclosure.
5. Give the system a few moments to initialize and recognize the new disk enclosure.
6. Restart or reload the SMU for the MSA 1050/2050/2052 and verify that the system has recognized the new storage.
7. Connect the new SAS cable (Figure 27, purple cable) from the B port (Out) of the current MSA storage system to the B port (In) of the new disk enclosure.
8. Give the system a few moments to initialize and recognize the new ports.
9. Ensure that all firmware is at the latest levels for all disk enclosures.
10. Use the SMU to verify proper system operation by validating all drives in the new disk enclosure.

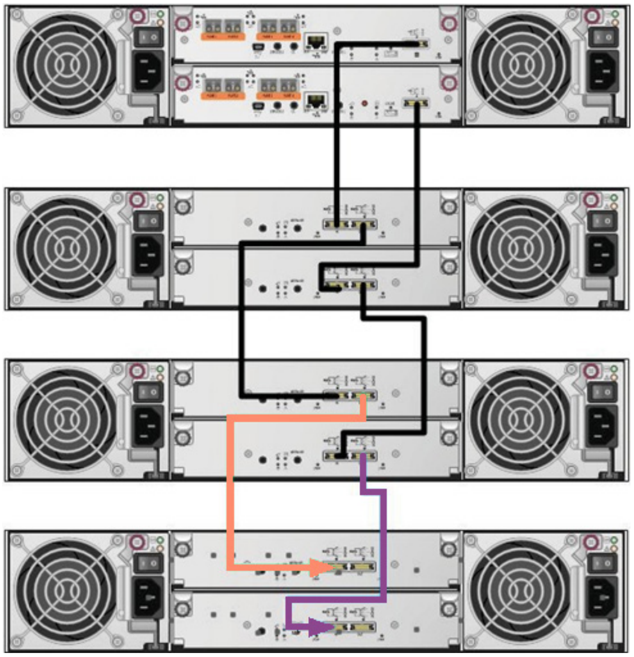
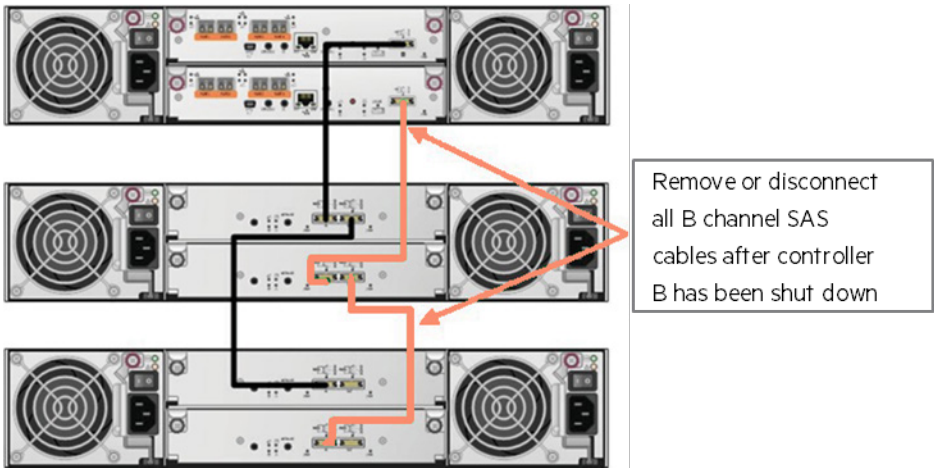


Figure 27. Extended straight-through cabling configuration

Upgrading from straight-through cabling to a fault-tolerant cabling configuration

Perform the following steps to complete the upgrade:

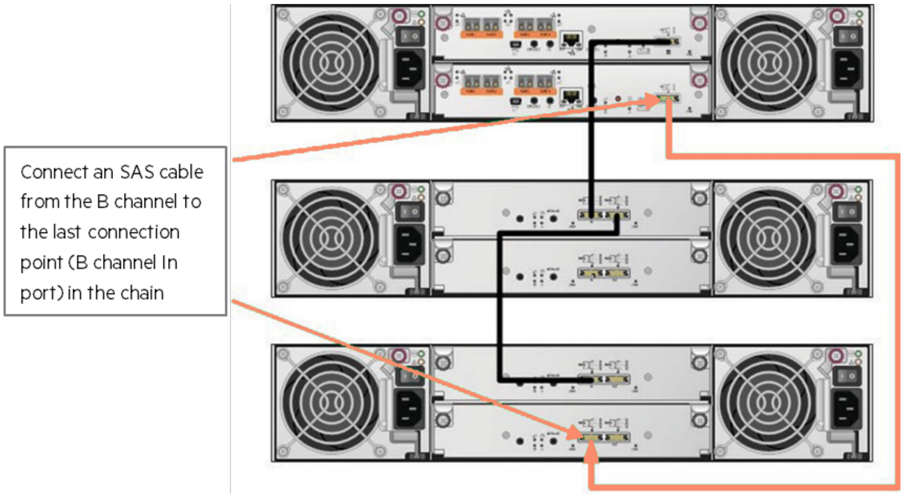
1. Install and mount any new disk enclosures, if needed, and then install all drives.
2. Power on the new disk enclosures.
3. Use the SMU to ensure that all disk enclosures are powered on and functioning properly.
4. Using the SMU or the CLI, shut down the B storage controller to ensure that any pending write or read requests are flushed from the controller B cache.
5. Remove existing cabling from the B channel chain as indicated in Figure 28.



Remove or disconnect all B channel SAS cables after controller B has been shut down

Figure 28. Straight-through cabling configuration

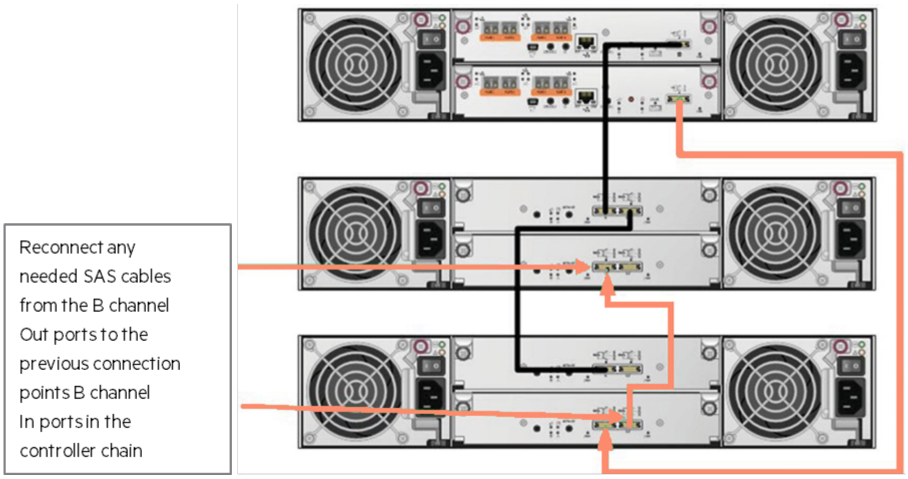
6. Connect an SAS interconnect cable from controller B on the MSA 1050/2050/2052 array to the last B channel In port on the disk enclosures as indicated in Figure 29.



Connect an SAS cable from the B channel to the last connection point (B channel In port) in the chain

Figure 29. Initial B channel cabling configuration

7. Connect or reconnect a SAS interconnect cable from the last B channel Out port to the previous disk enclosure's B channel In port, as indicated in Figure 30.
8. Using the SMU or the CLI, restart the B storage controller.
9. Give the storage system a few moments to initialize and recognize the new ports.
10. Ensure that all firmware is at the latest levels for all enclosures.
11. Verify proper system operation by validating that all drives in the new disk enclosure are recognized.



Reconnect any needed SAS cables from the B channel Out ports to the previous connection points B channel In ports in the controller chain

Figure 30. Completed B channel cabling configuration

Adding additional storage to an existing fault-tolerant environment

Use the following steps to complete the upgrade:

1. Install and mount the new disk enclosure and then install all drives.
2. Power on the new disk enclosures.
3. Use the SMU to ensure that all disk enclosures are powered on and functioning properly.
4. Connect the new SAS cable (Figure 31, orange cable) from the A port (Out) of the current MSA storage system to the A port (In) of the new disk enclosure.
5. Give the system a few moments to initialize and recognize the new disk enclosure.
6. Restart or reload the SMU for the MSA 1050/2050/2052 and verify that the system has recognized the new storage.
7. Using the SMU or the CLI, shut down the B storage controller to ensure that any pending write or read requests are flushed from the controller B cache.

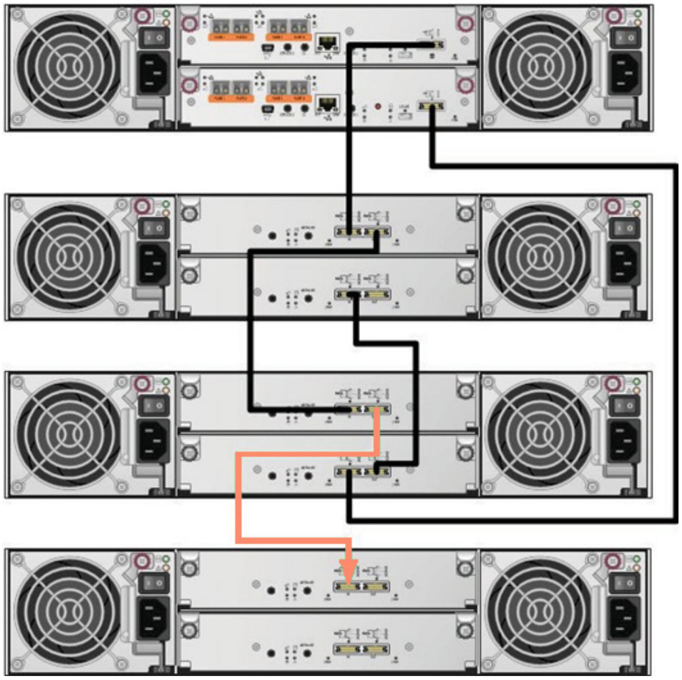


Figure 31. Adding new SAS cable to extend a channel to the additional shelf

8. Disconnect the existing SAS cable (Figure 32, dashed purple cable) from controller B of the current MSA 1050/2050/2052 storage system to the B port (In) of the existing disk enclosure, and reconnect the SAS cable to the last disk enclosure in the chain (Figure 33, purple cable).

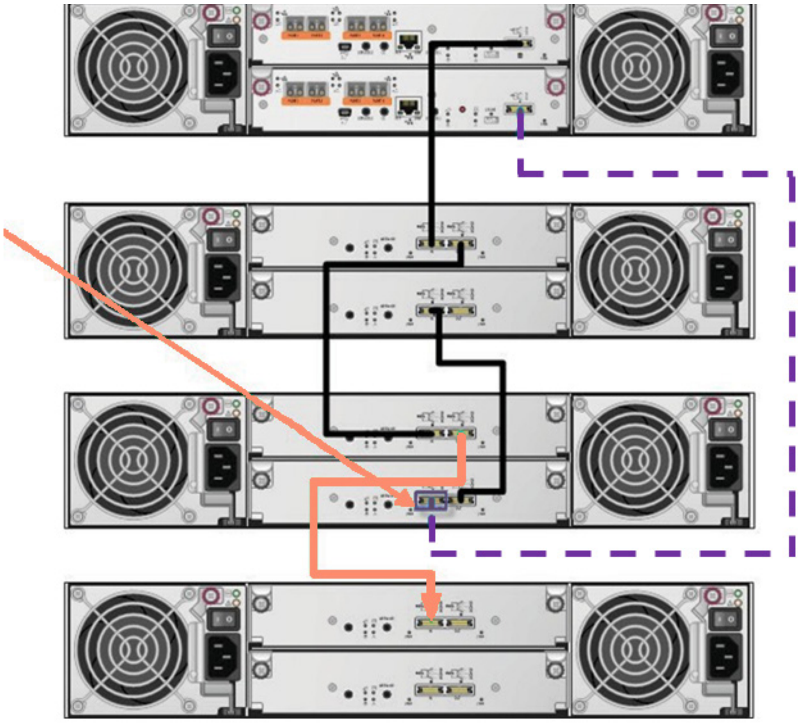
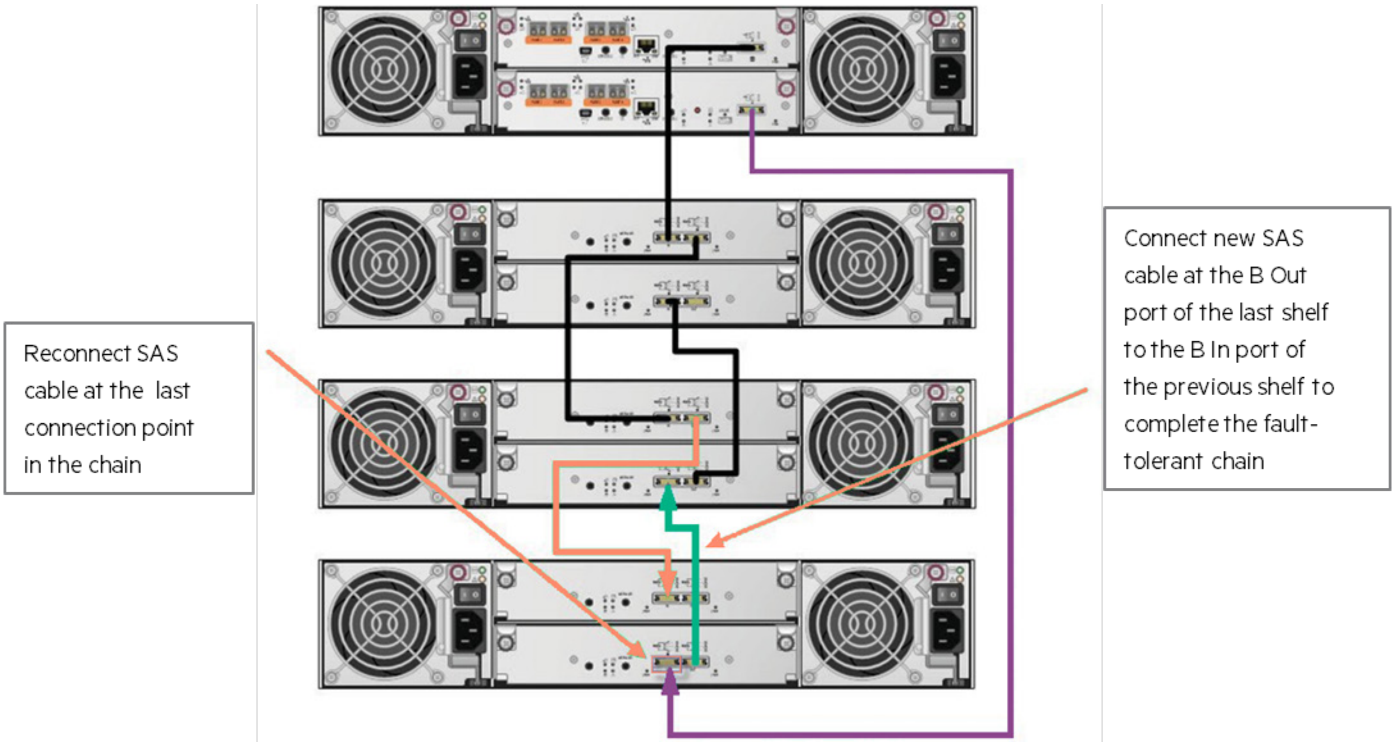


Figure 32. Breaking the B channel chain to extend B channel to an additional shelf

9. Connect the new SAS cable (Figure 33, green cable) from the B port (Out) of the new the MSA 1050/2050/2052 disk enclosure to the B port (In) of the previous disk enclosure.
10. Using the SMU or the CLI, restart the B controller.
11. Give the system a few moments to initialize and recognize the new ports.
12. Ensure that all firmware is at the latest levels for all disk enclosures.
13. Verify proper system operation by validating that the system has recognized all drives in the new disk enclosure.



Reconnect SAS cable at the last connection point in the chain

Connect new SAS cable at the B Out port of the last shelf to the B In port of the previous shelf to complete the fault-tolerant chain

Figure 33. Completing the fault-tolerant chain on the B channel

User management with the MSA

The guided MSA setup requires that you change the passwords for the default users or removing the users. Note that at least one user with the Manage role must exist.

Important

HPE recommends that at least one additional user with administrative rights is created for redundancy.

The MSA provides two different ways to access its management features. The first method is by using the legacy local user logins. The second method is to validate user credentials by using Lightweight Directory Access Protocol (LDAP) against a server running Windows Server 2012 R2 or Windows Server 2016 with Active Directory.

When a user attempts to log in to the array, the storage system checks the input credentials against the local user database first. To avoid confusion, do not duplicate Active Directory user names in the storage system local user database.

To enable the LDAP feature on the MSA storage system, you must disable all unsecure management protocols such as Telnet, HTTP, FTP, and unsecure Storage Management Initiative Specification (SMI-S). Only secure management protocols are supported to prevent sensitive user credentials from being passed through unsecure protocols.

Note

HPE recommends that all unsecure protocols be disabled. These include HTTP, Telnet, unencrypted SMI-S, FTP, Service Debug, and Activity Progress Reporting.

The LDAP feature allows storage administrators to log in to the MSA storage system using their common login, eliminating the need to remember the local user name and password.

There is a user management role known as **standard**. The standard role does not allow user management or firmware updates but allows all other array management features. HPE recommends limiting LDAP user groups to the standard role. Consult the MSA 1050/2050 SMU Reference Guide for information on configuring LDAP.

The array records all management interaction in an audit log, which is limited to 2 MB and wraps when it exceeds that size.

Summary

MSA 1050/2050/2052 administrators should determine the appropriate levels of fault tolerance and performance that best suits their needs. Understanding the workloads and environment for the MSA SAN is also important. Following the configuration options listed in this paper can help optimize the MSA 1050/2050/2052 array.

Resources

HPE MSA 1050 Storage QuickSpecs
hpe.com/support/MSA1050QuickSpecs

HPE MSA 2050 Storage QuickSpecs
hpe.com/support/MSA2050QuickSpecs

HPE MSA 2052 Storage QuickSpecs
hpe.com/support/MSA2052QuickSpecs

MSA 1050/2050 CLI Reference Guide
support.hpe.com/hpsc/doc/public/display?docId=emr_na-a00017709en_us

Sign up for HPE updates
h41360.www4.hpe.com/alerts-signup.php

Learn more at HPE MSA Storage

hpe.com/storage/msa

a00015961ENW, March 2020, Rev.5