Hewlett Packard Enterprise

Technical white paper

Check if the document is available in the language of your choice.

HPE MSA 1060/2060/2062 STORAGE ARRAYS

Best practices

CONTENTS

EXECUTIVE SUMMARY

This paper provides guidance on configuring HPE MSA storage arrays to meet recommended best practices from Hewlett Packard Enterprise. These recommendations help maximize application availability and performance, as well as improve system security. The paper is not a user guide but complements other official documentation that explains MSA storage technology and how to configure array settings. This best practices document focuses on providing clear recommendations and does not provide detailed information on the technologies it references. Technology details in the best practices documents for previous generation arrays have migrated between the HPE MSA Gen6 Virtual Storage Technical Reference Guide and technology-specific documents found in the HPE Support Center.

Intended audience

This paper is for those tasked with the installation, configuration, and ongoing maintenance of HPE MSA storage systems. Additionally, this paper will also assist technical sales staff in designing optimal solutions.

CONNECTIVITY BEST PRACTICES

Connectivity is the fundamental connection between devices. This section covers topics regarding the management network, data network, and path management software.

Naming hosts

Best practices for naming hosts include:

- **Best practice:** Group initiators (IDs) as hosts and define friendly names for them.
  - **System default**: None
  - **Detail:** The default HPE MSA Storage Management Utility (SMU) (web-based user interface) behavior is to not allow the mapping of a volume to a host without first creating a host of one or more initiators. Initiator names such as the World Wide Port Name (WWPN), which is applicable to Fibre Channel and SAS, and the iSCSI Qualified Name (IQN), which is applicable to iSCSI, are composed of long alphanumeric strings that are difficult to remember or recognize. Because it is possible to map volumes to initiators directly from within the CLI, HPE recommends providing port-based naming, which follows meaningful device inventory naming within an organization.
  - **Example**: Host name: dl380_gen10_1

    ID #1 Nickname: 10009cdc7172690a → dl380_gen10_1_port_0

    ID #2 Nickname: 10009cdc71726909 → dl380_gen10_1_port_1

iSCSI

Best practices for configuring an iSCSI connection include:

- **Best practice:** Use three network ports per host.
  - **System default**: None

    Detail: To ensure isolation between management and application traffic, HPE recommends using separate networks for management and iSCSI traffic. Additionally, at least two physical connections should provide connectivity to the data networks. Actual network topography may vary depending on the use of VLANs and network switch virtualization.

NOTE
A host does not require connectivity to either management port to access data.

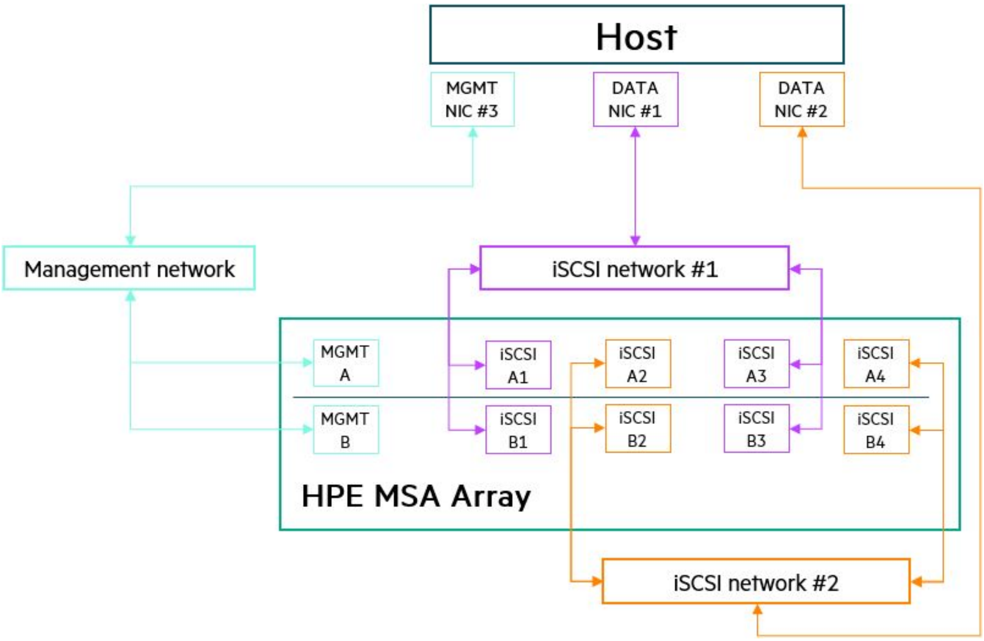

FIGURE 1. A sample iSCSI network topography

- **Best practice:** Use at least two isolated data networks when more than four array host ports are in use.
  - **System default**: None
  - **Detail:** To both maximize performance and minimize application disruption, HPE recommends not configuring more than eight paths to a volume. The more paths configured, the longer the failover time can be when active paths are lost. Using two networks limits an initiator to a maximum of eight paths to a volume (four active/optimized, four active/unoptimized).
  - **Example**:

Controller A, Port 1 (A1):	10.10.10.10/24
Controller B, Port 1 (B1):	10.10.10.11/24
Controller A, Port 2 (A2):	10.10.20.10/24
Controller B, Port 2 (B2):	10.10.20.11/24
Controller A, Port 2 (A3):	10.10.20.12/24
Controller B, Port 2 (B3):	10.10.20.13/24
Controller A, Port 2 (A4):	10.10.20.12/24
Controller B, Port 2 (B4):	10.10.20.13/24
Host data NIC1 #1:	10.10.10.100/24
Host data NIC #2:	10.10.20.100/24

[1] Network Interface Card

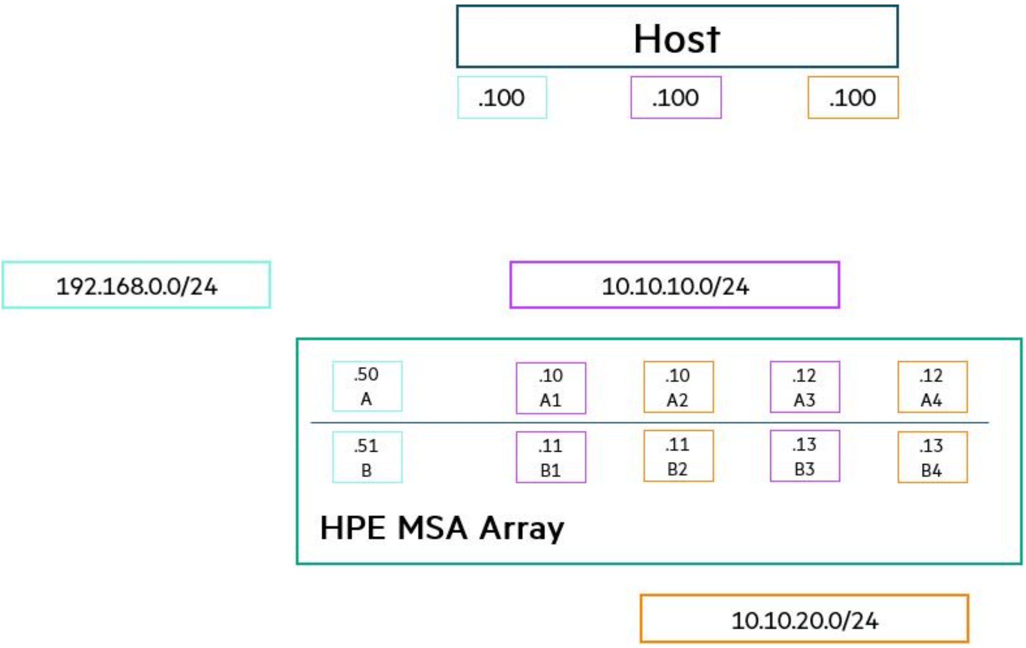

FIGURE 2. Example of iSCSI network subnet assignments

- **Best practice:** Set all devices on the same network to use jumbo frames when configured for an HPE MSA array.
  - **System default**: Disabled (MTU = 1,400)
  - **Detail:** Jumbo frames increase the maximum payload per Ethernet frame and can improve end-to-end performance. However, all devices within the data path must also use jumbo frames to avoid packet fragmentation or loss. HPE MSA arrays advertise a jumbo frame payload of 8,900 bytes. Sending devices usually agree with the MTU advertised by the receiver. If a device is unable to adjust its MTU automatically, do so manually.
  - **Example:** Issuing this command in the MSA CLI enables jumbo frames:

```
set iscsi-parameters jumbo-frames enabled
```

MPIO

Best practices for configuring an MPIO connection include:

- **Best practice:** Install and configure multipath software on connected hosts.
  - **System default**: Not installed
  - **Detail**: Multipath software provides load balancing and tolerance to link failure between a host and a storage array. Without multipath software, a volume mapped to a host appears as multiple physical disks, each with a single path. With multipath software, a volume can appear as a single physical disk that has multiple paths. A path is a connection between an initiator (host bus adapter [HBA] port or iSCSI software initiator) and a target (MSA array host port). When there are multiple active paths to the owning controller/pool for a given volume, multipath software can improve performance by distributing traffic evenly among those paths.
  - **Example**: Issuing the following commands in Microsoft Windows Server® 2016 PowerShell to enable MPIO with the HPE MSA 2060 storage array:

```
Install-WindowsFeature -name MultiPath-io
mpclaim -n -I -d “HPE     MSA 2060 FC”
```

NOTE
There are five spaces between “HPE” and “MSA” in the sample command.

NOTE
HPE MSA storage arrays do not support link aggregation.

- **Best practice**: Modify MPIO timers on Windows Server hosts when connecting to large numbers of logical unit numbers (LUNs).
  - **System default**: 20 seconds
  - **Detail**: Windows Server has a default period of 20 seconds that it retains a multipath pseudo-LUN in memory, even after losing all paths to the device. When this time has passed, pending I/O operations fail and the failure is exposed to applications, rather than continue to recover active paths. When a Windows® host has a large number of volumes (LUNs) mapped, 20 seconds might be too brief a time to wait and can cause long failover times and adversely affect applications.
  - HPE recommends modifying the PDORemovePeriod value within the system registry depending on the protocol used:
    - Fibre Channel: 90 seconds
    - iSCSI: 300 seconds
  - **Example**: Issue the following command at the Windows Server command prompt:

```
reg add HKLM\SYSTEM\CurrentControlSet\Services\mpio\Parameters /t REG_DWORD /v PDORemovePeriod
/d 300 /f
```

Maintaining supported configurations

Best practices for maintaining supported configurations include:

- **Best practice:** Cross-reference support for a configuration via the Hewlett Packard Enterprise Single Point of Connectivity Knowledge for HPE Storage Products (SPOCK) before deploying an HPE storage array or infrastructure upgrades.
  - **System default**: N/A
  - **Detail:** HPE only provides support for configurations listed within SPOCK. SPOCK is a database of tested interoperability between all active components in the data path, such as the operating system, HBAs, SAN switching infrastructure, and arrays, including all firmware and related software. As the storage vendor, the responsibility to issue support statements belongs to HPE, not the operating system vendor. Where possible, HPE works with operating system vendors to ensure support statements match, but this is not always possible. In the event of differing statements, SPOCK is authoritative.
  - **Example**: A SPOCK end-to-end support statement:

    Microsoft Windows Server 2019 → SN1100Q FC HBA → Brocade SN3000B 16Gb FC switch → MSA 2060 FC

BEST PRACTICES FOR MAINTAINING SYSTEM HEALTH

The availability of application data is paramount to any organization. This section provides recommendations to help ensure that the highest levels of availability are maintained.

Users

Best practices for maintaining system health for users includes:

- **Best practice**: Disable unsecure protocols.
  - **System default**: Disabled
  - **Detail**: To minimize the possibility of both unauthorized access and vulnerability to attack, HPE recommends that unsecure protocols remain disabled. These include:
    - Telnet
    - FTP
    - SNMP (unless creating an SNMPv3 user)
    - HTTP
    - Unsecure SMI-S
    - Debug

  - **Example:** Issuing this command in the HPE MSA CLI disables telnet:

    ```
    set protocols telnet disabled
    ```

- **Best practice:** Assign LDAP user groups to the standard role.
  - **System default**: N/A
  - **Detail**: For added security, users affiliated with the Standard role can perform most array management tasks but cannot manage users or clear logs.
  - **Example**: N/A

Firmware

Array firmware provides software that enables features and functionality and includes internal software components that enhance array stability and performance. Maintaining current firmware is critical not just to ensure access to the latest features, but to maintain application availability. Best practices include:

- **Best practice:** Have a tested backup strategy before applying firmware.
  - **System default**: N/A
  - **Detail:** Applying a firmware update to any component introduces a risk of failure and subsequent data unavailability or loss, although such failures are highly unlikely. HPE recommends that before the application of a firmware update, a tested backup solution is in place. Strategies can vary greatly depending on use case ranging from array-based snapshot replication (Remote Snap) to agent-based backup of applications to tape.
  - **Example:** HPE MSA Remote Snap Replication
- **Best practice:** Maintain current versions of all component firmware throughout the data path.
  - **System default**: N/A
  - **Detail**: SPOCK support matrices are continuously updated to reflect changing conditions as new firmware is released. What is supported when first installed might not be supported months later. Firmware releases sometimes bring new features, but often also close security holes and fix bugs. Check SPOCK periodically for new support streams and apply new firmware accordingly. HBAs, switches, storage arrays, and their connected disk enclosures and drives are all examples of firmware that is regularly maintained. Depending on the configuration, other components may also require updating.
  - **Example**: As of August 2020

    Operating system: Windows Server 2019

HBA:	SN1100Q Fibre Channel HBA firmware version 8.08.204
Switch:	Brocade SN3000B 16Gb Fibre Channel switch FOS version 8.2.2c
Array:	MSA 2060 Fibre Channel firmware version IN100R003
Disk:	HPE MSA 960GB SAS RI SFF SSD firmware version 0003

- **Best practice:** Update the HPE MSA and connected components via HPE Smart Component software.
  - **System default**: N/A
  - **Detail**: HPE Smart Component software enforces best practices when updating firmware and provides monitoring of progress that is unavailable elsewhere. Smart Component software increases the probability of a successful firmware update compared to direct updates via the SMU or SFTP.
  - **Example**: N/A
- **Best practice:** Consult the HPE MSA 1060/2060/2062 Storage Management Guide (SMG) for further best practices.
  - **System default**: N/A
  - **Detail**: The HPE MSA sixth-generation SMG contains a list of additional best practices to follow before and during a firmware update. To avoid duplication or misalignment, refer to the SMG in addition to this document. Visit the HPE Support Center to access the SMG.
  - **Example**: N/A

Controller firmware

Best practices for configuring controller firmware include:

- **Best practice:** Keep the Partner Firmware Update (PFU) setting **enabled** to ensure both controllers run the same firmware version.
  - **System default**: Enabled
  - **Detail**: If a controller is replaced or updated outside of Smart Component software, there is a possibility of controller firmware mismatch. The PFU setting ensures that both controllers are running the same firmware version, which is essential to maintaining system stability and functionality.
  - **Example**: Issuing this command in the HPE MSA CLI enables the PFU setting.

    ```
    set advanced-settings partner-firmware-upgrade enabled
    ```
- **Best practice:** Schedule controller firmware updates.
  - **System default**: N/A
  - **Detail:** Updating controller firmware causes each controller to reboot, although at different times. While a controller is rebooting, it is unable to respond to I/O, and the remaining controller temporarily serves I/O for both pools. In dual-pool configurations, there could be a measurable drop in performance during this time. To minimize the impact on applications, schedule controller firmware updates to occur during quiet periods. Although performance should not degrade in single pool configurations, HPE nevertheless recommends scheduling firmware updates to minimize application impact if unexpected problems arise.

    The average time to update firmware is five minutes per controller and no more than 30.
  - **Example:** 19:00, Saturday

Disk drive firmware

Best practices for configuring disk drive firmware include:

- **Best practice:** Stop all application I/O before updating disk drive firmware.
  - **System default**: N/A
  - **Detail:** Disk drive firmware cannot be updated online (while serving I/O) because there are inconsistent practices among drive vendors regarding recovery times. As a result, there is a small chance of I/O timeout. Applications must cease accessing volumes hosted in whole or in part on pools containing target disk drives. It is possible to update drives without stopping I/O so long as neither pool contains the target disk drive type. The Smart Component software assumes all target drives are idle and upgrades their firmware even if I/O is still present.
  - **Example**: N/A
- **Best practice:** Wait for all background tasks to complete before updating disk drive firmware.
  - **System default**: N/A
  - **Detail:** Some background tasks read and write data to disks. If disk drives are targets for firmware updates and are participating in these tasks, there is a possibility of unwanted interruption. Before proceeding with drive firmware updates, check via the activity monitor that the none of the following disk group tasks are active:
    - Initialization
    - Expansion (MSA-DP+ disk groups only)
    - Reconstruction
    - Verification
    - Scrubbing
  - **Example**: N/A
- **Best practice:** Disable disk group background scrub during drive firmware updates if not using the Smart Component.
  - **System default**: Enabled
  - **Detail:** As a scheduled task, background disk scrub has the potential to interfere during a disk drive update. HPE recommends that disk group background scrub be disabled before updating drive firmware and re-enabled when finished. The Smart Component disables background scrubbing automatically but remains at its current setting if updated directly within the array.

- **Example:** Issuing this command in the HPE MSA CLI disables disk group background scrubbing.

```
set advanced-settings background-scrub disabled
```

System monitoring

The monitoring of array health is essential because, without notifications, an administrator would be unable to take the appropriate corrective action promptly. This section contains recommendations that will help administrators receive appropriate and necessary array information at the right time.

- **Best practice:** Configure email notifications of system events.
  - **System default**: Not configured
  - **Detail:** Email notifications are essential so that administrators are made aware of pending issues with the array. When configured, the HPE MSA array sends emails to up to three addresses with information on alerts as they happen. If more than three recipients are required, configure a distribution list of relevant email addresses in the email server software.
  - **Example:** Issuing this sample command in the HPE MSA CLI configures email notification parameters.

```
set email-parameters server smtp.org.net domain org.net email-list admin@org.net sender
msa2060_1 sender-password Password123 port 25 security-protocol TLS alert-notification-level
all
```

- **Best practice:** Enable the managed logs feature and include logs.
  - **System default**: Not configured
  - **Detail:** HPE MSA arrays have a finite amount of storage for log data. The managed logs feature helps reduce the likelihood of losing log data due to wrapping by notifying defined email addresses that the log is nearly full. After a notification is received, an administrator should access the MSA array and download the logs (pull).

    However, HPE recommends also enabling the **include logs** option, which automatically attaches the logs to the notification email (push). Doing so eliminates the possibility of losing historical log data.
  - **Example:** Issuing these sample commands in the HPE MSA CLI enables the managed logs feature.

```
set email-parameters include-logs enabled email-list admin@org.net,,,logcollector@org.net
```

```
set advanced-settings managed-logs enable
```

- **Best practice:** Configure SNMPv3.
  - **System default**: Not configured
  - **Detail:** SNMP is used to monitor events from managed devices centrally and minimizes downtime caused by unacknowledged system events such as component failure. HPE recommends creating an SNMPv3 user for added security.
  - **Example**: N/A
- **Best practice:** Sign up for proactive product advisory notifications.
  - **System default**: N/A
  - **Detail**: To maximize availability, performance, and features, subscribe to HPE Support alerts via the HPE Email Preference Center.
  - **Example**: N/A

Background scrubbing

Best practices for background scrubbing include:

- **Best practice:** Do not disable the background disk group scrubbing feature.
  - **System default**: Enabled
  - **Detail:** The disk group scrubbing feature is important to maintaining array health and maximizing application availability. In addition to both finding and attempting to fix disk errors within a disk group, disk group scrubbing can also reclaim zeroed pages and return capacity to the pool.

- **Example:** Issuing this command in the HPE MSA CLI enables the background disk group scrubbing feature and sets an interval of every 24 hours:

  ```
  set advanced-settings background-scrub enabled background-scrub-interval 24
  ```

IMPORTANT

The only time background scrubbing of disks and disk group should be disabled is before updating disk drive firmware.

- **Best practice:** Do not disable the background disk scrubbing feature.
  - **System default**: Enabled
  - **Detail:** The disk scrubbing feature is similar to the disk group scrubbing feature, except that it checks disks that are not associated with a disk group and has a fixed interval timer of 72 hours.
  - **Example:** Issuing this command in the HPE MSA CLI enables the background disk scrubbing feature.

    ```
    set advanced-settings background-disk-scrub enabled
    ```

Data protection

HPE MSA arrays provide various technologies to aid in the protection of application data and the retention of earlier copies. In addition to low-level mechanisms used to distribute data on disk, snapshots and replication provide further peace of mind. Best practices include:

- **Best practice:** Schedule volume snapshots and configure retention policies.
  - **System default:** Not configured
  - **Detail:** HPE MSA arrays provide redirect-on-write (ROW) volume snapshots that enable the immediate recovery of data from a given point in time. HPE recommends that all volumes have at least one schedule in place for the automatic taking of snapshots.

    Configure retention policies to ensure that snapshots for a given interval do not exceed a defined number. Multiple snapshots schedules allow for finer control over snapshot retention.
  - **Example:** Volume_a001

    Schedule #1: Once per day, retention count = 7

    Schedule #2: Once per week, retention count = 4

    Schedule #3: Once per month, retention count = 12
- **Best practice:** Implement Remote Snap replication.
  - **System default**: Not configured
  - **Detail:** HPE MSA arrays enable replication of volumes to remote arrays over switched Fibre Channel and iSCSI networks. In the event of a disaster at one site, Remote Snap provides an array-based solution to recover access to data as it was when last successfully replicated.
  - **Example**: N/A
- **Best practice:** Maintain a tested backup solution.
  - **System default**: N/A
  - **Detail:** Snapshots are convenient and provide near-instantaneous recovery to a given point in time, but they are not an ideal solution to provide long-term retention of data and might not offer application consistency without the coordination of host-based software. Backup solutions take many forms, including the use of agents within virtual machines that back up data over the network to central locations and to various media. HPE does not recommend a specific backup solution but does recommend that a tested backup solution is configured together with array-based snapshots and replication.
  - **Example**: N/A
- **Best practice:** Consider defining a fixed percentage of pool capacity for use by snapshots.
  - **System default**: 10%
  - **Detail:** To ensure snapshots do not consume more available capacity than desired or to allow them to consume more, consider defining the percentage of a pool's capacity reserved for snapshot data.

- **Example:** Issue the following sample command in the HPE MSA CLI to define 15% of pool capacity to be used by snapshots.

  ```
  set snapshot-space pool A limit 15% middle-threshold 85% limit-policy delete
  ```

- **Best practice:** Enable both the controller-failure and partner-notify settings via the CLI.
  - **System default**: Enabled as of Firmware I110, disabled in previous firmware versions
  - **Detail:** When a single controller is unavailable, these settings instruct the remaining controller to change the cache mode to write-through, which ensures written data is committed to disk at the cost of reduce write performance. After both controllers are operational, the cache mode is reverted to its previous setting, which is usually write-back mode.
  - **Example:** Issue the following commands in the HPE MSA CLI to enable both settings:

    ```
    set advanced-settings controller-failure enable
    set advanced-settings partner-notify enable
    ```

Periodic health checks

Best practices when performing periodic health checks include:

- **Best practice:** Perform regular checks of array health via the HPE Health Check tool.
  - **System default**: N/A
  - **Detail:** The HPE MSA Health Check tool provides a free and easy way to check array health and conformity to best practices that maximize array availability. MSA Health Check is especially helpful in environments where an array does not have access to the internet and where a local manifest is not maintained because it provides a list of available firmware updates. HPE recommends putting a regular schedule in place to upload logs securely to MSA Health Check and to review the subsequent report.
  - **Example**: N/A

STORAGE BEST PRACTICES

"Storage" in this section refers to the underpinning technologies used to distribute and protect data across multiple drives and the data services that provide the resulting capacity, features, and performance. This section targets best practices regarding the configuration of disk drives, disk groups, and pools.

Disk drives

Best practices when configuring disk drives include:

- **Best practice:** Choose the correct drive types for the workload.
  - **System default**: N/A
  - **Detail:** Different drives provide differing ratios of performance, capacity, and cost. Consider workloads before building a solution and ensure that the solution does not use inappropriate drive types.
    - SSD: Suited to workloads that require high random performance (IOPS) and low latency
    - Enterprise-SAS: Optimized for around-the-clock low-end random performance (1-2K IOPS) and high-throughput and low-latency sequential I/O
    - Midline-SAS: Optimized for archival data; not recommended for constant high-workload applications
  - **Example**: N/A
- **Best practice:** Install SSDs in the array enclosure first.
  - **System default**: N/A
  - **Detail:** For optimal performance, install SSDs in the main array enclosure first, and then the closest enclosures as required.
  - **Example**: N/A
- **Best practice:** Replace SSDs when their remaining life reaches 5%.
  - **System default**: N/A

- **Detail:** SSD failure or wear-out is extremely rare. However, if an SSD reaches 0% life left, data loss occurs. To avoid this, the HPE MSA array logs a warning-level event at 5%, 1%, and 0%. HPE recommends replacing SSDs no later than when 5% wear is remaining.
- **Example**: N/A

Choosing disk group types

A disk group is a collection of disks logically teamed together to provide capacity and performance to a pool. Best practices when selecting a disk type group include:

- **Best practice:** Choose an appropriate RAID protection level.
  - **System default**: MSA-DP+
  - **Detail:** RAID provides both resiliency to drive failure and performance benefits. With the introduction of MSA-DP+, HPE recommends that all mechanical hard disk drives (HDDs) take advantage of its benefits and SSDs when in all-flash pool configurations.

NOTE
HPE no longer recommends the use of RAID 5 with mechanical HDDs.

TABLE 1. HPE recommended RAID levels

Disk type	Tier	RAID protection level	Pool type	Detail
SSD	Performance	RAID 1/10	Hybrid or All-flash	• Performance-optimized solutions that require the best random write performance • Low-capacity solutions where RAID types requiring more drives (RAID 5/MSA-DP+) would not be cost-effective
		RAID 5	Hybrid	• Typical configurations that require very high random read performance and medium-high random write performance
		MSA-DP+	All-flash	• Recommended for all-flash configurations
HDDs (all)	Standard or Archive	MSA-DP+	Any	• Recommended for all configurations • Highly resilient to failure • Very fast rebuild times • More performant than RAID 6 and idle spares
		RAID 6		• Recommended when the capacity goal is too low for MSA-DP+ • Medium resiliency to failure • Medium performance • Long rebuild times

  - **Example**: N/A
- **Best practice:** Follow the Power of 2 rule.
  - **System default**: N/A
  - **Detail:** The Power of 2 rule ensures an optimal data layout with parity-based disk groups (RAID 5 and 6). Not following this rule while using HDDs can reduce sequential write performance by 50% or more. See the HPE MSA Gen6 Virtual Storage Technical Reference Guide for detailed information.

TABLE 2. Optimal drive configurations

RAID level	# drives in a disk group	# data chunks in a stripe	# parity chunks in a stripe
RAID 5	3	2	1
RAID 5	5	4	
RAID 5	9	8	
RAID 6	4	2	2
RAID 6	6	4	
RAID 6	10	8	

NOTE
MSA-DP+ disk groups have the Power of 2 rule embedded into their design regardless of drive count.

NOTE
It is not necessary to configure SSDs to follow the Power of 2 rule.

- **Example**: N/A

Sparing

Sparing enables an array to automatically assign idle disks or capacity to rebuild a degraded disk group or stripe zone, thus preventing data loss. Best practices and array behavior vary depending on drive technology, RAID type, and drive form factor.

NOTE
To minimize complexity, examples are given in relation to enclosures. Because an SFF enclosure holds 24 drives and an LFF enclosure holds 12, it is acceptable to interpret the term enclosure to mean the first 12 or 24 drives of that type. For example, if six 10K HDDs are installed within Enclosure 1 and another six are installed within Enclosure 2, only two spares would be required because their sum of twelve is not more than the total drives slots available in a single enclosure.

- **Best practice:** Assign global spares to HDDs tiers that do not use MSA-DP+.
  - **System default:** Dynamic sparing
  - **Detail:** By default, the HPE MSA array uses dynamic sparing, which uses unassigned drives (AVAIL) as needed. However, to ensure that disks are always available for use as spares, HPE recommends allocating drives as global spares as shown in Table 3.

NOTE
To improve reliability, SSD disk groups have their data drained to an HDD tier upon drive failure. Because dedicated SSD spares are costly and could reduce availability should the disk group degrade due to wear-out, HPE recommends when using hybrid pools, to allocate global spares for HDDs only. Additionally, global spares are neither required nor useable by MSA-DP+ disk groups, because sparing is an integral feature of this RAID group type.

TABLE 3. Recommended sparing strategy for non-MSA-DP+ disk groups

Drive type	Tier	Form factor	Spares per enclosure	Spare per additional enclosure
SSD	Performance or SSD Read Cache	SFF / LFF	0 / 0	0 / 0
SSD	All-Flash pool	SFF / LFF	2 / 1	1 / 1
Enterprise SAS (10K/15K)	Standard	SFF	2	1
Midline SAS (7.2K)	Archive	LFF	1	1

  - **Example**: N/A
- **Best practice:** Define an adequate target spare capacity when using MSA-DP+ disk groups.
  - **System default**: Equal to the sum of the two largest drives in the disk group
  - **Detail:** Instead of using unassigned idle spares as needed, MSA-DP+ RAID includes spare capacity within the disk group. By default, spare capacity equals the summed capacity of the two largest drives within the disk group. Because it is possible to expand MSA-DP+ disk groups, HPE recommends setting increasing amounts of target spare capacity before expanding a disk group.

IMPORTANT
The only fixed rule for spare capacity and MSA-DP+ disk groups is that spare capacity must equal the two largest drives' summed capacity within the group. All other rules are best practice guidance that is open for adoption per the individual scenario. For example, guidance provided on a per disk enclosure basis might need adaptation when a disk group spans multiple enclosures in a nonuniform manner, with only a small number of participating drives in each.

There are three scenarios to consider regarding spare capacity for an MSA-DP+ disk group:

- **During the initial creation of the disk group.** It is not possible to define the target spare capacity via the SMU. Therefore, if the quantity of disks is greater than one enclosure, add it to the pool via the CLI. Refer to the HPE MSA 1060/2060/2062 CLI Reference Guide for more information on the spare-capacity switch of the `add disk-group` command. Follow Table 4 for guidance on target spare space capacity.
- **The expansion of a disk group using drives of equal size.** Before expansion, modify the target spare capacity as per Table 4, then add drives and expand the disk group.
- **The expansion of a disk group using larger drives.** Before expansion, modify the target spare capacity as per Table 5, then add drives and expand the disk group.

TABLE 4. Recommended sparing strategy for MSA-DP+ disk groups using drives of the same size

Drive type	Tier	Form factor	Spare capacity for the first disk enclosure	Spare capacity increase per additional disk enclosure
SSD	Performance or SSD Read Cache	SFF / LFF	0 / 0	0 / 0
SSD	All-flash pool	SFF / LFF	2 / 2	1 / 1
Enterprise SAS (10K/15K)	Standard	SFF	= 2x largest drive capacity (default drive count when disk group is created)	+ 1x largest drive capacity
Midline SAS (7.2K)	Archive	LFF		

TABLE 5. Recommended sparing strategy for MSA-DP+ disk groups using drives of different sizes

Drive type	Spare capacity for the first disk enclosure (+1x larger drive)	Spare capacity for the first disk enclosure (+2x larger drives)	Spare capacity increase per additional disk enclosure
SSD (tiering)	0	0	0
SSD (all-flash pool)	= 1x largest drive + 1x next largest capacity	= 2x largest drive capacity (not a default)	+ 1x largest drive capacity
Enterprise SAS (10K/15K)			
Midline SAS (7.2K)			

– **Example**: Table 6 provides examples of spare capacity scenarios using SFF drives.

TABLE 6. Examples of spare capacity scenarios using SFF drives

Detail	Scenario #1 Same size drives (one disk enclosure)	Scenario #2 Same size drives (two disk enclosures)	Scenario #3 +1x larger drive (one disk enclosure)	Scenario #4 +2x larger drive (two disk enclosures)	Scenario #5 +3x larger drive (three disk enclosures)
Starting drive type	1.2 TB SFF	1.2 TB SFF	1.2 TB SFF	1.2 TB SFF	1.2 TB SFF
Starting # disk drives	12	12	12	23	46
Starting # enclosures	1	1	1	1	2
Starting spare capacity	2.4 TB	2.4 TB	2.4 TB	2.4 TB	3.6 TB
Additional drive type	1.2 TB SFF	1.2 TB SFF	1.8 TB SFF	1.8 TB SFF	1.8 TB SFF
# of new drives	12	36	1	2	3

Required # of additional disk enclosures	0	1	0	1	1
Additional spare capacity	0	1.2 TB	600 GB	2.4 TB	3.0 TB
New total spare capacity	2.4 TB (2 * 1.2)	3.6 TB (3 * 1.2)	3 TB (1.8 + 1.2)	4.8 TB (2 * 1.8 + 1.2)	6.6 TB (2 * 1.8 + 1.8 + 1.2)

WARNING

It is only possible to define the target spare capacity via the CLI. If the target spare capacity is to be set higher than the default, it must be defined during disk group creation or before adding new drives. Additionally, a combination of drive capacity and quantities sufficient to reach the target spare capacity must be included or added to the disk group. Not doing so results in a target spare capacity greater than what is available, thus placing the disk group into a degraded state.

Details to consider regarding the scenarios in Table 6 are:

- **Scenario #1**: Added drives are of the same capacity and do not consume a second enclosure. The default rule of spare capacity equaling twice the capacity of the largest drive within the group applies and no changes to spare capacity are required.
- **Scenario #2**: Added drives require the addition of a second enclosure. Because the drive capacities are the same as those already in the disk group, only additional capacity equal to that of one more drive is required (2 + 1).
- **Scenario #3**: Although the addition of a single larger drive does not require an additional enclosure, it does require an increase to the target spare capacity. Because there is only one drive of the larger capacity in the disk group, spare capacity must equal the sum of the larger capacity drive (1.8 TB) and one of the next largest drives in the disk group (1.2 TB).
- **Scenario #4**: With two larger drives added to the disk group, the minimum target spare capacity is equal to two of the larger drives (2 X 1.8 TB). However, both enclosures contain one larger drive each, and the first enclosure contains a majority of smaller drives. In this example, when the target spare capacity minimum of two largest drives is satisfied, the remaining spare capacity can be met with the smaller drives because they dominate the first enclosure. Note that the enclosure order does not always matter concerning spare capacity.
- **Scenario #5**: Like with Scenario #4, the larger drives do not fully populate an enclosure. Because there are two 1.8 TB drives in the second enclosure, this is already enough to meet the basic spare capacity minimums. Although the third enclosure contains only one drive, it is in the majority and should count towards spare capacity. Again like Scenario#4, the first enclosure contains a majority of smaller 1.2 TB drives, so their capacity counts towards the final target spare capacity.

Tiering

Tiering is the process of granularly distributing a volume across different drive types to provide the best balance of cost and performance. Tiering best practices include:

- **Best practice:** Maintain proportions across tiers.
  - **System default**: N/A
  - **Detail:** The HPE MSA tiering engine works in near real-time to deliver the best balance of performance to cost. If a deep combined knowledge of array tiering behavior and workload knowledge is present, it might be possible to successfully work outside of the recommended tier ratios. However, in most cases, following HPE recommendations allows the tiering engine to deliver its value reliably and effectively and with minimal administrator intervention.

NOTE

Although it is supported, HPE does not recommend combining SSDs and Midline SAS drives within the same pool without an intermediary tier of Enterprise SAS drives. Doing so can reduce the effectiveness of the tiering engine and reduce both performance and value. For example, even low capacity configurations using MSA-DP+ would require more SSD read cache than can be configured. If SSDs were used in a performance tier instead, the solution would be neither as cost-effective nor as performant as when using three tiers. However, if known in advance that the working set is and will remain sufficiently small to fit into the available SSD region, then it might be a practical solution. HPE recommends re-evaluating performance and ratios over time.

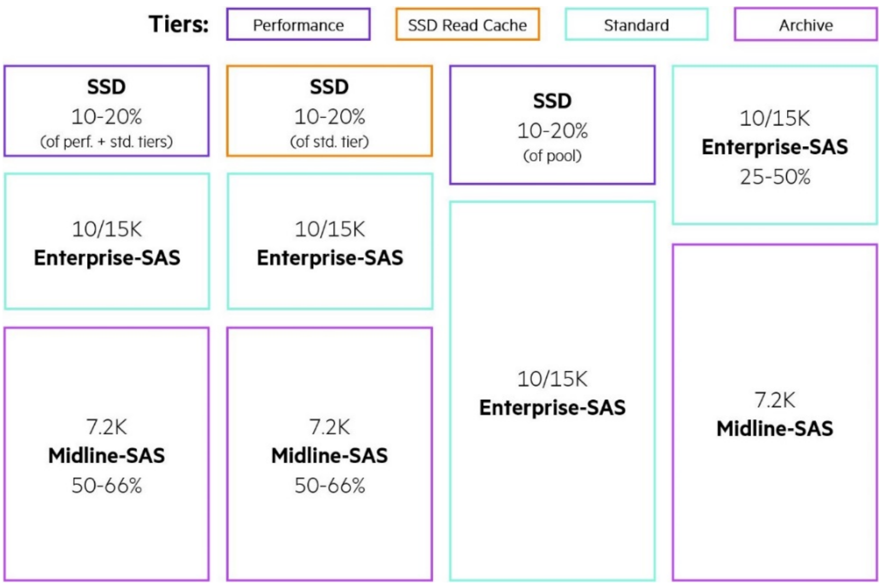

FIGURE 3. Recommended tier ratios

There are two possibilities regarding the recommended quantity of SSDs when combined in a pool with HDDs. In both cases, HPE recommends proportioning SSDs against the standard tier.

TABLE 7. Recommended SSD proportion rules

SSD use case	Flash ratio to maintain	Proportion
Performance tier	10-20%	Percentage of the summed capacity of both the performance tier and the fastest HDD tier
SSD read cache		Percentage of the fastest HDD tier

– **Example**: Table 8 provides examples of disk tier proportions in a three-tier hybrid pool.

TABLE 8. Examples of disk tier proportions in a three-tier hybrid pool

Tier	RAID	# drives	Drive capacity	# disk groups	Capacity	Capacity ratio	Flash ratio (rounded)
Performance	RAID 10	4	1.92 TB	1	3.84 TB	5%	15%
Standard	MSA-DP+	13	2.4 TB	1	21.60 TB	29%	(3.84 / (3.84 + 21.60)) * 100
Archive	MSA-DP+	12	6 TB	1	48.0 TB	65%	N/A

- **Best practice:** Configure SSDs as capacity within the performance tier for generic workloads.
  - **System default**: N/A
  - **Detail:** Mixing SSDs and Enterprise SAS drives as capacity tiers automatically enables tiering and delivers the best balance of cost and performance for most workloads. Refer to the HPE MSA Gen6 Virtual Storage Technical Reference Guide for information on how the HPE MSA tiering engine works and why it is the most effective choice.
  - **Example**:

    Performance tier: 1x RAID 10 disk group

    Standard tier: 1x MSA-DP+ disk group
- **Best practice:** Choose read cache only when a workload is known to have a very low percentage of random writes.
  - **System default**: N/A
  - **Detail:** SSD read cache accelerates random reads, and does not accelerate random writes or sequential I/O. If random writes are frequent, use performance tiering.

- **Example:** Customer requirement

  Random writes: <2K IOPS

  Random reads: >2K IOPS

- **Best practice:** Choose single-tier HDD configurations for throughput-heavy workloads or data archiving.
  - **System default**: N/A
  - **Detail:** HDDs provide an ideal solution for workloads that are not dominantly random. Introducing performance tiering or SSD read cache for such workloads is unlikely to yield tangible benefits but does increase the cost.
  - **Example:** Customer requirement

    Random I/O: <2K IOPS

    Throughput: <8.7 GB/s read

    <5.5 GB/s write

- **Best practice:** Choose single-tier SSD configurations for extremely high throughput-heavy workloads or for large capacities of low-latency storage and datasets that are not candidates for data-reduction techniques.
  - **System default**: N/A
  - **Detail:** SSDs provide low latencies for random I/O and up to 13 GB/s of sequential throughput. Because an HPE MSA array does not offer deduplication or compression, it might be an economical solution for large datasets that are not eligible for compaction. For datasets that are eligible, HPE recommends considering other solutions within the portfolio such as HPE Nimble Storage arrays.
  - **Example:** Customer requirement
    - Random I/O: >5K IOPS
    - Throughput: >8.7 GB/s
    - Random latencies: <10ms
    - Capacity: Many tens of terabytes

- **Best practice:** Periodically review and maintain sufficient SSDs capacity needed to absorb 80% of daily random I/O.
  - **System default**: N/A
  - **Detail:** Recommended tier proportions in this guide are for new installations, and a configuration applied to an HPE MSA array when brought into service might not continue to be effective over the full duration of an array's useful life. For optimal performance, SSDs must have enough capacity to store 80% or more of daily random I/O, which typically accounts for a small fraction of a pool's total size.

    Located in the Capacity area of the SMU, the I/O Workload tool gives administrators a graph of historic daily I/O and a visual representation of the relationship to SSD capacity. HPE recommends using the I/O Workload tool at defined intervals throughout the life of the array to drive changes in the configuration of a pool as needed.
  - **Example:** Customer requirement
    - Random I/O: >5K IOPS
    - Throughput: >8.7 GB/s
    - Random latencies: <10ms
    - Capacity: Many tens of terabytes

- **Best practice:** Do not change the volume tier affinity setting for most workloads.
  - **System default**: No Affinity
  - **Detail:** HPE recommends that volumes use the default tier affinity setting, which is No Affinity. Modifying the affinity setting of a volume could result in the unnecessary degradation of performance for other volumes within the pool and might not yield the expected results.

    There are, however, workloads and scenarios where changing a volume's affinity is warranted. For example, the Archive Affinity setting is particularly useful for infrequently accessed data because it frees capacity in the upper tiers for performance-sensitive applications. Refer to the HPE MSA Gen6 Virtual Storage Technical Reference Guide for further guidance on how and when to use the tier affinity setting.

– **Example**:

Backups:	Archive
HDD images:	Archive
Boot volumes:	Performance
SQL:	No Affinity
Video streaming:	No Affinity
General VM storage:	No Affinity

Single vs. dual pools

HPE MSA arrays ship with two array controllers, each supporting a single pool of storage. A pool contains disk groups that provide capacity and performance independently of the other pool. This section provides best practices regarding the use of a single (asymmetrical) or dual pools (typically, symmetrical).

IMPORTANT

As of Firmware I110, the array defaults for controller-failure and partner-notify settings are enabled although in previous firmware, they were disabled. If a single controller becoming unavailable, these settings cause the remaining controller cache policy to switch to write-through, which will cause a degradation in write performance in return for the assurance that written data is committed to disk. Although the full performance of the remaining controller cannot be realized during single-controller operation, it is still important to consider controller headroom so the overall impact to application performance can be minimized. Refer to the HPE MSA 1060/2060/2062 Storage Arrays Best Practices for further guidance.

- **Best practice:** Use a single pool unless either the performance or capacity goal explicitly requires a second pool.
  - **System default**: N/A
  - **Detail:** Single-controller HPE MSA performance and addressable capacity scales beyond the requirements of most customers. Additionally, single-controller configurations ensure the minimum impact to performance in the unlikely event of controller unavailability during a peak in I/O demand.

    The official HPE tool for sizing is HPE Ninja Online for MSA, which automatically suggests single and dual pool configurations that match the intended performance and capacity goals.

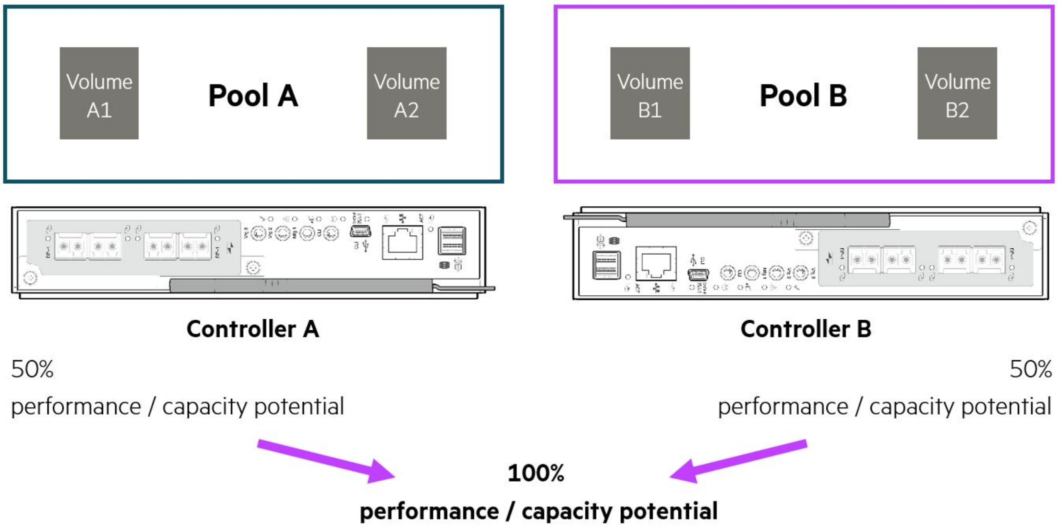

FIGURE 4. Representation of pool and array performance and capacity potential

  - **Example**:

Single pool: 100K IOPS 4K random read performance
4 GB/s sequential read performance
100 TB capacity

Dual pool: 300K IOPS 4K random read performance

13GB/s sequential read performance
1.5 PB capacity

- **Best practice:** Use the Volume Copy feature to rebalance an underperforming array.
  - **System default**: N/A
  - **Detail:** Sixth-generation HPE MSA arrays can address up to 1 PB of storage via a single pool and provide more than adequate performance such that two pools are unlikely necessary. However, for extremely demanding workloads, the array can provide, in most cases, double the capacity and performance by using its second pool.

    If a single-pool array configuration no longer meets requirements, use the Volume Copy feature to assist in the migration of a volume to the other pool. The Volume Copy feature might also help in rebalancing a dual pool configuration where one pool is underutilized.

IMPORTANT
Volume Copy requires the volume to be unmapped from a host and application traffic for that volume halted. After the copy is complete, map the newly copied volume to the host, resume any applications, and consider deleting the source volume.

  - **Example:** Issuing the following command in the HPE MSA CLI copies the volume SourceVol from Pool A to Pool B with a new name "DestVol":

```
copy volume SourceVol destination-pool B name DestVol
```

Thin provisioning

Thin provisioning is the practice of defining volume capacities that, when totaled, exceed the physical capacity of a pool. The principal goal of thin provisioning is to reduce the initial costs of owning an array by both reducing the number of drives that must be initially purchased and the power and cooling costs required to operate them.

- **Best practice:** Monitor pool usage and set appropriate thresholds and notifications.
  - **System default**:
    - Low threshold: 50%
    - Middle threshold: 75%
    - High threshold: Calculated (available pool capacity minus 200 GB)
  - **Detail:** When a pool is overcommitted and has no capacity remaining, incoming writes to previously unwritten areas of any volume are rejected. Additionally, when a pool reaches its high threshold, performance is reduced.

    Configure notifications and set thresholds that allow sufficient time to procure new physical capacity or remove unwanted data from the pool.
  - **Example**: N/A
- **Best practice:** If using VMware ESXi™ 6.5 or later, periodically issue the UNMAP command.
  - **System default**: N/A
  - **Detail:** ESXI 6.5 introduced the automatic release of allocated block storage after the removal of files from a datastore. However, VMware® issues the UNMAP command at a 1 MB granularity but the HPE MSA operates with a 4 MB page size. As a result, pages do not become free automatically but can be released when the UNMAP command is invoked manually.
  - **Example:** ESXi CLI

```
esxcli storage vmfs unmap -l MyDatastore
```

- **Best practice:** Leave the overcommit setting enabled for pools with a large number of snapshots.
  - **System default**: Enabled
  - **Detail:** Overcommit allows snapshots to consume a minimal amount of allocated capacity within the pool. When overcommit is disabled, snapshots consume available space equal to the size of the original volume. In most circumstances, disabling overcommit has an undesired impact on the available capacity of a pool and should be left enabled.
  - **Example**: N/A

Full disk encryption

HPE MSA arrays support data-at-rest encryption via the use of self-encrypting drives (SEDs).

NOTE

Both HPE MSA 1060 and 2060 arrays support full disk encryption. The HPE MSA 2062 array is encryption-capable, but the included 1.92 TB SSDs are not. Because full disk encryption requires every drive in the system to be an SED, it is neither economical nor recommended to use the HPE MSA 2062 array for this purpose.

IMPORTANT

It is not possible to recover a lost passphrase, and it is not possible to access data on a locked system without it. Therefore, HPE strongly recommends keeping a copy of all passphrases in a secure location.

- **Best practice:** Use a different passphrase for each FDE-secured system.
  - **System default**: N/A
  - **Detail:** All HPE MSA arrays that support encryption generate the same lock key and the same passphrase. To increase security, HPE recommends locking each array by using a unique passphrase, which also generates a unique lock key.
  - **Example**: N/A
- **Best practice:** Clear FDE keys before shutdown when moving an entire HPE MSA storage system.
  - **System default**: Not cleared
  - **Detail:** To ensure that data is inaccessible after an array has been relocated, HPE recommends clearing the FDE keys. When the array is powered on again, the disks will be in a secure, locked state, and the original passphrase must be entered to re-access data.
  - **Example**: Issue the following commands in the HPE MSA CLI to clear the FDE keys.

    Before powering down: `clear fde-keys current-passphrase myPassphrase`

    After power is reapplied: `set fde-lock-key myPassphrase`

Capacity expansion

Best practices regarding capacity expansion include:

- **Best practice:** Expand tiers comprising RAID 1/10, 5, or 6 with disk groups of equal proportions and rotational speeds.
  - **System default**: N/A
  - **Detail:** For performance to be consistent across a tier and pool, all disk groups within a tier should exhibit the same performance characteristic and provide equal capacity.
  - **Example**: Before expansion: 2x RAID 6 disk groups, each with ten 2.4 TB Enterprise-SAS drives (Standard tier)

    After expansion: 3x RAID 6 disk groups, each with ten 2.4 TB Enterprise-SAS drives (Standard tier)
- **Best practice:** Expand tiers comprising MSA-DP+ disk groups with drive capacities no greater than a factor of two than the smallest.
  - **System default**: N/A
  - **Detail:** Using drives of capacities that are greater than a factor of two can lead to the inefficient use of the capacity provided by the larger drive and reduces their value.
  - **Example**: 12x 1.2 TB HDD + 1 x 2.4 TB HDD = OK

    12x 600 GB HDD + 1 x 2.4 TB HDD = Not recommended

Volume mapping

Best practices to follow when mapping volumes include:

- **Best practice:** Ensure volumes are mapped via sufficient host ports to meet performance requirements.
  - **System default**: N/A

- **Detail:** If volumes are not attached to hosts via a sufficient number of host ports, array performance may be limited. Consider how many active paths are required to meet the performance potential of the configured system.
- **Example**: Table 9 provides examples of data rates per host port and protocol.

TABLE 9. Data rates per host port and protocol

Link speed	Throughput (MB/s)	IOPS
1Gb iSCSI	110	14,200
10G iSCSI	1,100	142,000
12Gb SAS[2]	5,280	681,600
8Gb FC	880	113,600
16Gb FC	1,760	227,200

- **Best practice:** Do not configure more than eight paths from a host to the array.
  - **System default**: N/A
  - **Detail:** The time for MPIO to recover from multiple path failures might increase unacceptably if there are too many paths to a volume. Because performance cannot benefit from more than eight paths, use Fibre Channel zoning or subnetting to limit available paths.
  - **Example**: N/A
- **Best practice:** Ensure there are redundant paths from a host to the array.
  - **System default**: N/A
  - **Detail:** For application data to remain available during a controller or path outage, ensure that hosts have a minimum of two paths to a volume; one via Controller A and one via Controller B.

    Connect fabric-attached hosts to at least two independent Fibre Channel or Ethernet networks or SAS switches. Connect directly connected hosts with at least one path to each controller. Refer to SPOCK for compatibility information.

    Example: Figure 5 illustrates a host with redundant paths to a volume.

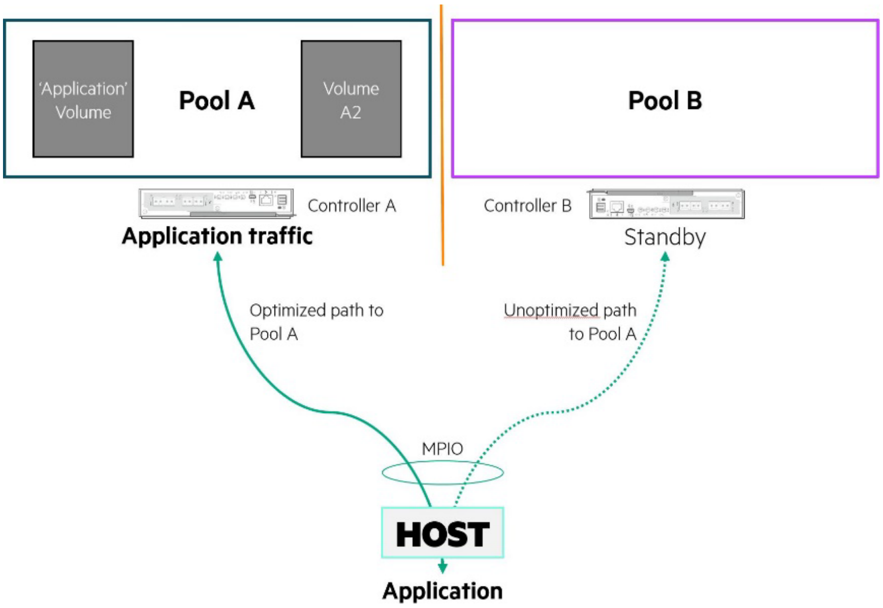

FIGURE 5. Example of a host with redundant paths to a volume

[2] Each SAS port contains four 12Gb SAS lanes.

- **Best practice:** Do not attach a volume to more than one host unless all hosts support the same cluster-aware file system or are cooperatively managed such that non-shared file systems can be used.
  - **System default**: N/A
  - **Detail:** Sharing volumes between multiple hosts can lead to data loss and corruption if the file system or operating system cannot cooperate and temporarily lock disk regions at a granular level. Always refer to operating system documentation for guidance on its file system capabilities and requirements.

IMPORTANT
Some file systems such as Microsoft Cluster Shared Volumes (CSV) require that hosts be in a cluster, whereas others such as VMware vSphere® VMFS do not. Ensure that you place those file systems that do require a cluster configuration in one before mapping volumes.

  - **Example:** The following are examples of file systems that can be shared under the correct circumstances:
    - VMFS is a cluster-aware file system that can be shared without placing hosts in a cluster
    - CSV is a cluster-aware file system that requires hosts participate in a failover cluster

SUMMARY

This best practices guide will help administrators ensure maximal performance and availability of their HPE MSA arrays. Use this guide and the documentation listed to deliver the best configuration for your application needs.

Resources, contacts, or additional links

HPE MSA Gen6 Virtual Storage Technical Reference Guide
https://www.hpe.com/h20195/v2/Getdocument.aspx?docname=a00103247enw

HPE MSA Health Check
www.hpe.com/storage/MSAHealthCheck

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

LEARN MORE AT HPE MSA STORAGE

hpe.com/storage/msa

a00105260ENW, June 2021, Rev. 1