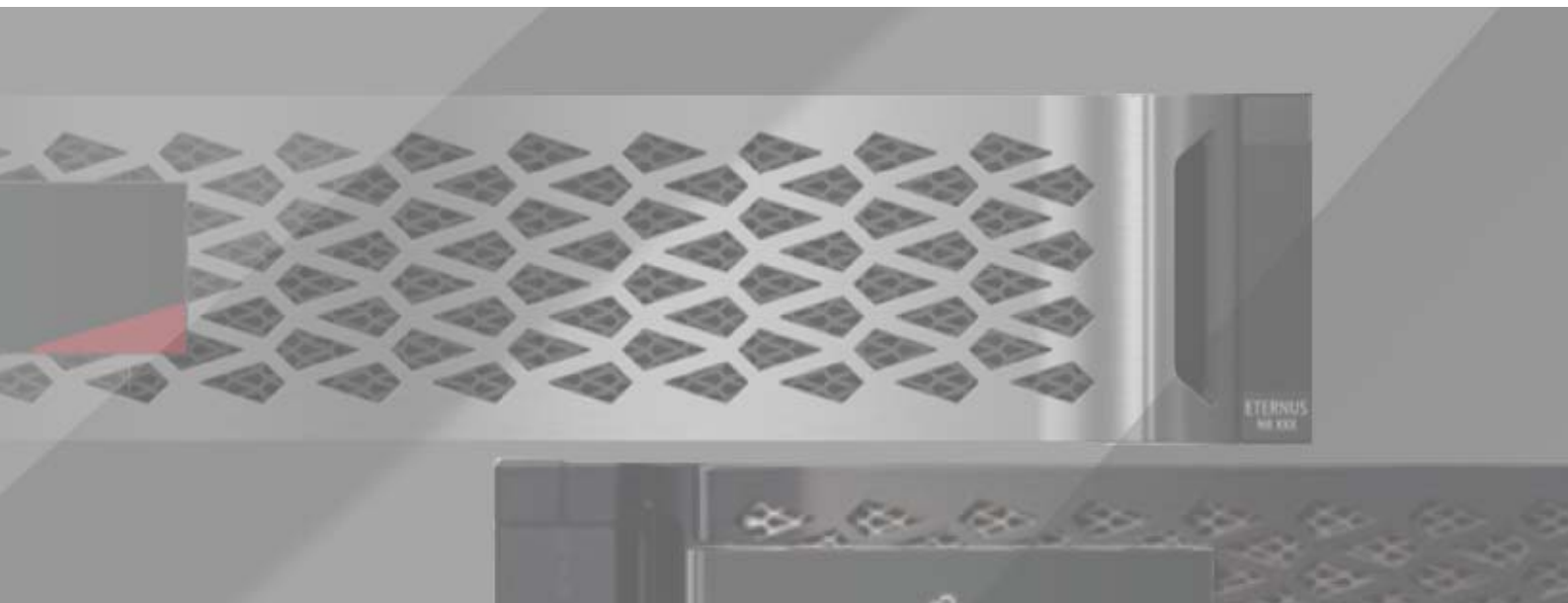


# FUJITSU Storage ETERNUS AB series All-Flash Arrays, ETERNUS HB series Hybrid Arrays

---

## SANtricity OS Dynamic Disk Pools Feature Description and Best Practices



# Table of Contents

<b>1. Introduction .....</b>	<b>7</b>
Overview .....	7
Intended Use .....	7
Units Convention .....	8
<b>2. Technical Overview .....</b>	<b>9</b>
Data Layout .....	9
Operation When a Drive Fails .....	10
Multiple Drive Failures .....	12
Drawer Loss Protection .....	12
Management .....	13
Comparison of DDP and Volume Groups .....	16
Configuration Guidelines .....	17
<b>3. Performance.....</b>	<b>18</b>
<b>4. Best Practices for Configuring Pools .....</b>	<b>20</b>
Choosing Between DDP Technology and Traditional RAID .....	20
How Large Should the Pool Be? .....	21
Analytics Best Practices: Small-Block Random Workloads .....	21
Backup and Video Surveillance Best Practices: Sequential Workloads .....	22
Technical Computing Best Practices: Large-Block Sequential Workloads .....	22
Configuring Pools with Equal Capacity Volumes .....	22
Reconstruction Priority Setting .....	24
<b>5. Summary .....</b>	<b>26</b>
<b>A. Glossary of Terms.....</b>	<b>27</b>
<b>B. Thin Provisioning .....</b>	<b>28</b>

# List of Figures

Figure 1	D-piece and D-stripe example .....	9
Figure 2	24-drive pool .....	10
Figure 3	24-drive pool with one drive that has failed .....	10
Figure 4	Example of time to rebuild an HDD .....	11
Figure 5	Example of time to rebuild an SSD .....	11
Figure 6	4U60 drive shelf .....	12
Figure 7	Create Pool in SANtricity System Manager .....	13
Figure 8	Pool Settings in SANtricity System Manager .....	14
Figure 9	Create Volumes in SANtricity System Manager .....	15
Figure 10	Block Size 16K/64K Performance (Read/Write) .....	18
Figure 11	Block Size 16K Latency (Read:Write=3:1) .....	18
Figure 12	Block Size 16K Latency (Read:Write=1:3) .....	19
Figure 13	Mixing volume groups and pools .....	20
Figure 14	Selecting DDP technology or RAID .....	20
Figure 15	Equal volumes in SANtricity System Manager .....	23
Figure 16	Pool capacity with equal volumes in SANtricity System Manager .....	23
Figure 17	RAID rebuild versus DDP reconstruction .....	24
Figure 18	Drop in IOPS performance during rebuild .....	25

# List of Tables

Table 1	Comparison of DDP and Volume Groups .....	16
Table 2	DDP configuration guidelines .....	17

# Preface

With Dynamic Disk Pools (DDP), storage administrators can group sets of like disks into a pool topology in which all the drives in the pool participate in the I/O workflow. This technology provides faster drive rebuilds than with RAID 5 or RAID 6 and removes the complications of RAID group configurations, so that storage administrators can focus on capacity allocation. This document provides a detailed description of the DDP feature.

Copyright 2020 FUJITSU LIMITED

First Edition  
November 2020

## Trademarks

---

Third-party trademark information related to this product is available at:

<https://www.fujitsu.com/global/products/computing/storage/eternus/trademarks.html>

Trademark symbols such as ™ and ® are omitted in this document.

## About This Manual

---

### Intended Audience

---

This manual is intended for system administrators who configure and manage operations of the ETERNUS AB/HB, or field engineers who perform maintenance. Refer to this manual as required.

### Related Information and Documents

---

The latest information for the ETERNUS AB/HB is available at:

<https://www.fujitsu.com/global/support/products/computing/storage/manuals-list.html>

## Document Conventions

---

### ■ Notice Symbols

The following notice symbols are used in this manual:

#### Caution

Indicates information that you need to observe when using the ETERNUS AB/HB. Make sure to read the information.

#### Note

Indicates information and suggestions that supplement the descriptions included in this manual.

# 1. Introduction

---

## Overview

---

Dynamic Disk Pools (DDP) technology represents a significant advancement in storage system data protection and management. As disk capacities continue to grow without corresponding increases in data transfer rates, traditional RAID rebuild times are getting longer, even up to several days. Slow rebuilds result in much more time with degraded performance and exposure to additional disk failures.

With five issued patents, DDP technology is designed to deliver worry-free storage through effortless management and self-optimization while maintaining predictable performance under any conditions, including recovery from drive failures. With rebuild times that are up to four-times faster than previous methods, DDP technology significantly reduces exposure to multiple cascading disk failures, providing excellent data protection.

The following list identifies the key DDP attributes that enable these benefits:

- Simplified management:
  - Distributed hot spare capacity (known as preservation capacity) eliminates the need for dedicated idle hot spare drives.
  - You can add drives to a pool without reconfiguring RAID.
  - The protection scheme and stripe size are automatic; you do not need to configure them.
- Predictable performance:
  - A deterministic algorithm dynamically distributes data, spare capacity, and protection information across a pool of drives.
  - If a drive fails, segments are recreated elsewhere, which reduces the magnitude and duration of the performance disruption.
  - The large pool of drives reduces hot spots.
- Reduced exposure to multiple disk failures:
  - Through segment relocation, the system returns to an optimal state faster.
  - Through prioritized reconstruction, any stripes that experience multiple drive failures are given the highest priority.

This technical report provides a high-level overview of DDP technology and best-practice guidelines for using pools.

## Intended Use

---

This information is for Fujitsu customers and partners.

## Units Convention

---

In this document, IEC binary units are used when referring to base-2 values, and decimal units are used for base-10 values. Following are examples of binary units:

- **KiB.** Kibibyte, or  $1,024$  bytes
- **MiB.** Mebibyte, or  $1,024^2$  bytes
- **GiB.** Gibibyte, or  $1,024^3$  bytes
- **TiB.** Tebibyte, or  $1,024^4$  bytes
- **PiB.** Petabyte, or  $1,024^5$  bytes

Following are examples of decimal units:

- **KB.** Kilobyte, or  $1,000$  bytes
- **MB.** Megabyte, or  $1,000^2$  bytes
- **GB.** Gigabyte, or  $1,000^3$  bytes
- **TB.** Terabyte, or  $1,000^4$  bytes
- **PB.** Petabyte, or  $1,000^5$  bytes

### Caution

SANtricity System Manager uses binary labels for binary values.



## 2. Technical Overview

With DDP technology, SANtricity OS and management software allows you to create pools in addition to traditional volume groups (generally referred to as RAID groups). A pool can range in size from a minimum of 11 drives to as large as all the drives in a storage system, which is up to 480 disks in the ETERNUS HB5200. Pools can consist of either hard disk drives (HDDs) or solid-state drives (SSDs). In addition, pools and volume groups can coexist in the same system. The following lists some possible combinations with a 24-drive storage system in which all drives have equal capacity. Note that this is an example and other combinations are also available.

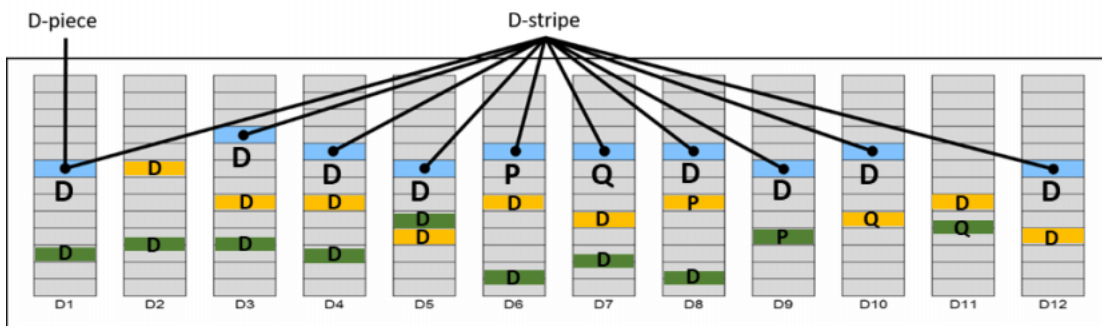
- One 8-drive RAID 10 (4+4) volume group, and one 16-drive pool
- One 24-drive pool
- Two 12-drive pools
- One 5-drive RAID 5 (4+1) volume group, one 4-drive RAID 10 (2+2) volume group, and one 15-drive pool

## Data Layout

Within a pool, volume data is distributed across all drives, regardless of how many drives are assigned to the pool. A volume comprises many virtual stripes, known as D-stripes. Each D-stripe resides on 10 drives that are distributed throughout the pool by an intelligent optimization algorithm. The portion of each D-stripe that resides on a single drive is called a D-piece. Each D-piece is a contiguous section of the physical drive. [Figure 1](#) shows an example of how a D-stripe might be laid out. In this case, the pool consists of 12 drives, but even if it had more, the D-stripe would still only be divided into 10 D-pieces.

Note that the D-pieces do not necessarily reside in the same portion of each drive. More information on D-stripes and D-pieces and their capacities can be found in ["A. Glossary of Terms"](#).

Figure 1 D-piece and D-stripe example



Each D-stripe consists of 8,192 RAID 6 stripes on an ETERNUS AB6100 and 4,096 RAID 6 stripes on all other ETERNUS AB/HB series storage systems. Each of these stripes is composed of ten 128 KiB segments, for a total of 1 GiB on the ETERNUS AB6100 and 512 MiB on all other ETERNUS AB/HB series storage systems. As shown, eight of the segments are data (D), one is parity (P), and one is the RAID 6 Q value.

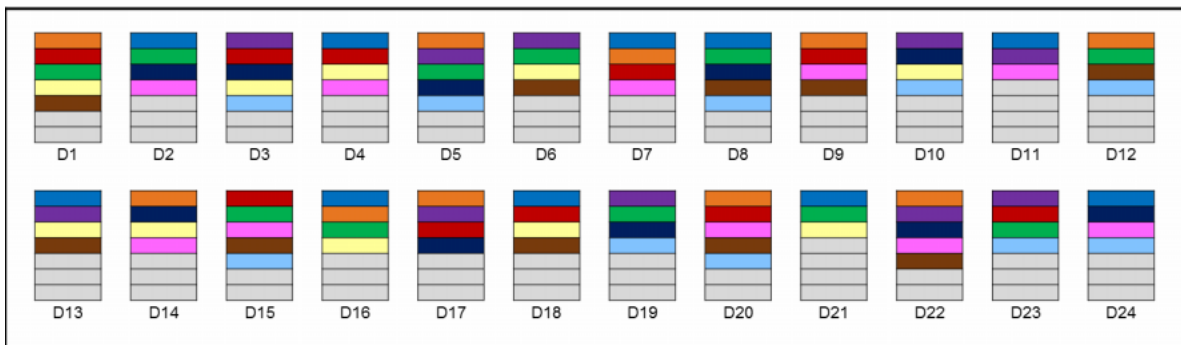
After a pool has been created, a volume can be created within the pool. This volume consists of some number of D-stripes across all the drives within the pool, with the number of data D-stripes equaling the defined volume capacity divided by the D-stripe size. For example, a 2 TiB volume consists of 256 data D-stripes on the ETERNUS AB6100 and 512 data D-stripes on all other ETERNUS AB/HB series storage systems. Allocation of D-stripes for a given volume is performed starting at the lowest available range of logical block addresses (LBAs) for a given D-piece on a given drive.

## Operation When a Drive Fails

A major benefit of DDP technology is that, rather than using dedicated stranded hot spares, the pool itself contains integrated preservation capacity to provide rebuild locations for potential drive failures. This feature simplifies management because you no longer have to plan or manage individual hot spares. It also greatly improves the time of rebuilds and enhances the performance of the volumes themselves during a rebuild.

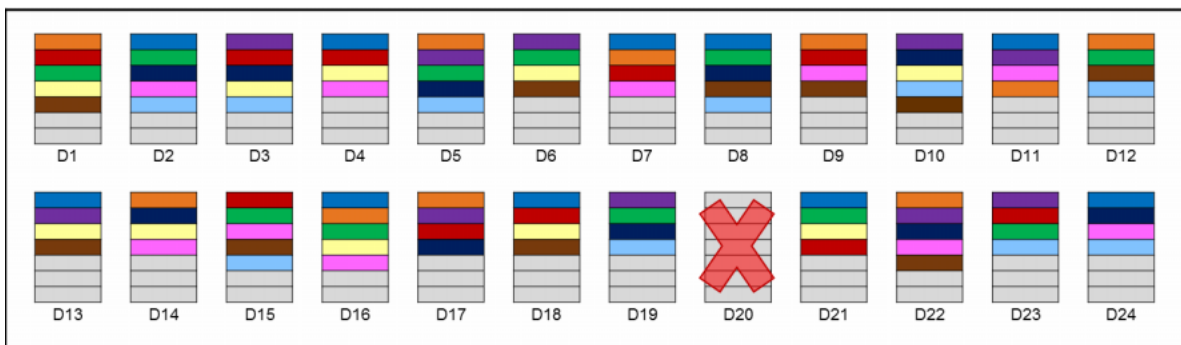
To begin discussion of DDP operation when a drive fails, consider the 24-drive pool that is depicted in [Figure 2](#). Each different color in the diagram represents a D-stripe, each of which contains 10 D-pieces of the same color. The D-pieces are distributed over the pool by the DDP intelligent algorithm, as previously noted.

Figure 2 24-drive pool



Now suppose that one of the drives in the pool fails, as in [Figure 3](#).

Figure 3 24-drive pool with one drive that has failed



When a drive in a pool fails, the D-pieces from the failed drive are reconstructed segment by segment, using the same mechanism that is normally used by RAID 6. The intelligent algorithm chooses other drives in the pool on which to write the rebuilt D-pieces, confirming that no single drive contains two D-pieces from the same drive. The individual D-pieces are reconstructed at the lowest available LBA range on the selected drive.

In [Figure 3](#), drive D20 has failed and its D-pieces have been rebuilt and written to other drives. The rebuild operation runs in parallel across all drives. Because multiple drives participate in the effort, the overall performance effect of this situation is reduced, and the length of time that is needed to complete the operation is also dramatically reduced.

[Figure 4](#) and [Figure 5](#) illustrate the difference in rebuild times between RAID 6 and DDP technology. [Figure 4](#) shows data from an ETERNUS HB1100/HB1200/HB2000 system with various numbers of HDDs, while [Figure 5](#) presents similar data for SSDs. Both were run with the rest of the system in an idle I/O state. These charts show that a pool rebuilds faster than a RAID 6 volume group, and, as the pool spindle count increases, DDP rebuild times go down compared to RAID 5 and RAID 6 rebuild times.

Figure 4 Example of time to rebuild an HDD

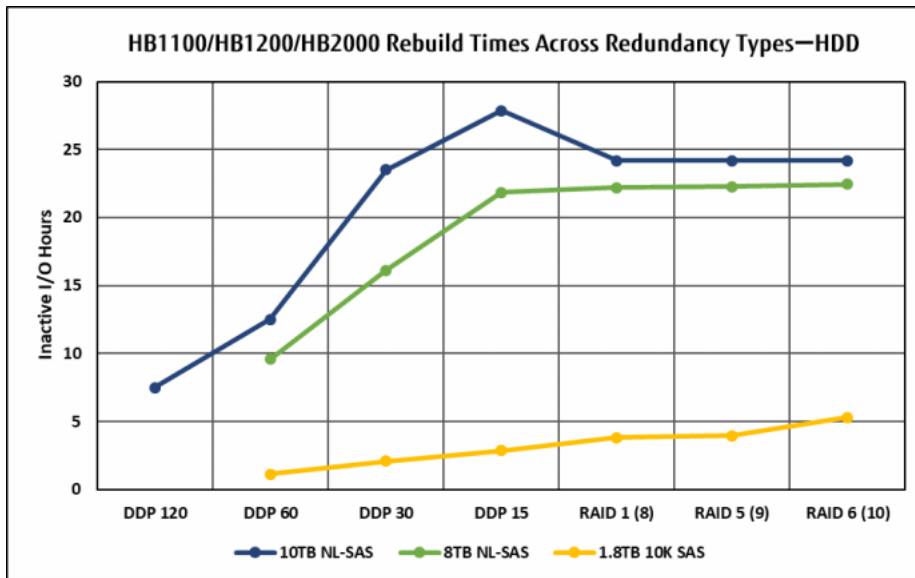
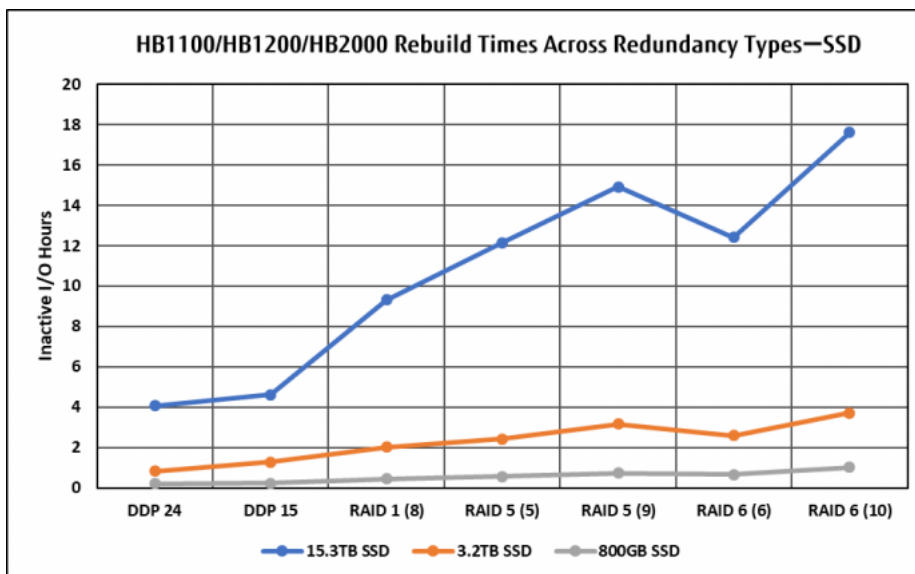


Figure 5 Example of time to rebuild an SSD



## Multiple Drive Failures

To minimize data availability risk, if multiple drives fail within a pool, any D-stripes that are missing two D-pieces are given priority for reconstruction. This approach is called critical reconstruction. After critically affected D-stripes are reconstructed, the rest of the necessary data is then reconstructed.

From a controller resource allocation perspective, there are two user-modifiable reconstruction priorities within the pool:

- Degraded reconstruction priority is assigned for instances in which only a single D-piece must be rebuilt for the affected D-stripes. The default priority for this instance is high.
- Critical reconstruction priority is assigned for instances in which a D-stripe has two missing D-pieces that must be rebuilt. The default priority for this instance is highest.

For very large pools with two simultaneous disk failures, only a relatively small number of D-stripes are likely to encounter the critical situation in which two D-pieces must be reconstructed. As discussed previously, these critical D-pieces are identified and reconstructed initially at the highest priority. This approach returns the pool to a degraded state very quickly so that further drive failures can be tolerated.

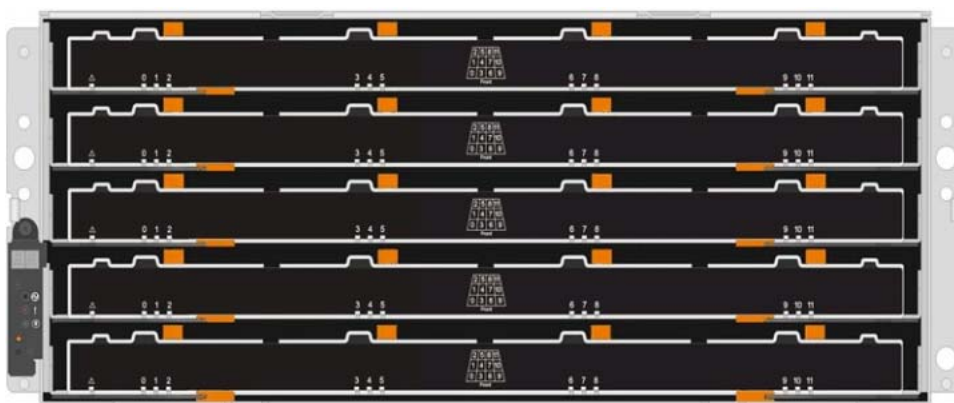
As an example, assume that a pool of 192 drives has been created and has two drive failures. In this case, it is likely that the critical D-pieces would be reconstructed in less than one minute and, after that minute, an additional drive failure could be tolerated. From a mathematical perspective, with the same 192-drive pool, only 5.2% of D-stripes would have a D-piece on one drive in the pool and only 0.25% of the D-stripes would have two D-pieces on those same drives. Therefore, only 48 GiB of data must be reconstructed to exit the critical stage. A very large pool can continue to maintain multiple sequential failures without data loss until there is no additional preservation capacity to continue the rebuilds.

After the reconstruction, the failed drive or drives can be subsequently replaced, although replacement is not specifically required. Fundamentally, this replacement of failed drives is treated in much the same way as an online capacity expansion of the pool. Failed drives can also be replaced before the pool exits from a critical or degraded state.

## Drawer Loss Protection

The 4U60 drive shelf holds up to 60 drives in 4U of rack space. The drives are organized into five drawers, with each drawer containing up to 12 drives. See [Figure 6](#).

Figure 6 4U60 drive shelf



As of SANtricity OS 11.25, it is possible to achieve drawer loss protection (DLP) within a single 4U60 shelf. Drawer loss protection refers to the ability of a pool to withstand the loss of an entire drawer and still maintain I/O operations with no loss of availability.

To enable DLP with a single shelf, the configuration must have at least 15 drives, equally distributed among the drawers. SANtricity System Manager presents DLP candidates during pool creation. The number of drives in these candidates is always a multiple of five. When you add drives to increase pool capacity, you should add them in groups of five, one per drawer to maintain DLP with the new capacity.

DLP can also be enabled for systems with multiple 4U60 shelves. In this case, drives with similar characteristics (drive type, capacity, data assurance, and security) must be distributed equally across all drawers so that proper candidates are presented. All DLP candidates that are presented through the management interfaces consist of an equal number of drives per drawer for all drawers in the pool.

In some complex configurations, it might be necessary to create a smaller DLP pool initially. This situation can occur if similar drives are not equally distributed among the drawers because some of the drawers include other drive types. In that case, you can expand the pool by adding five drives at a time to maintain DLP.

## Management

Configuring pools is simpler than configuring traditional volume groups. You do not have to choose RAID levels, segment sizes, or global hot spares because they are determined by system defaults. The primary decisions for the administrator are what type of drives, how many drives, and whether the pool will have security or DLP.

[Figure 7](#) shows pool creation in SANtricity System Manager. In this case, the administrator has selected a candidate that has 60 HDDs that offer both DLP and full disk encryption (FDE) security.

Figure 7 Create Pool in SANtricity System Manager

**Create Pool**

What is shelf loss protection and drawer loss protection?

Name MyPool

Drive type  
HDD (SAS)

Select a capacity for your pool ...

Free Capacity (GiB)	Total Drives	Secure-Capable	Enable Security?	DA Capable	Shelf Loss Protection	Drawer Loss Protection
420892.00	60	Yes - FDE	Key Required	Yes	No	Yes
337832.00	60	No	N/A	Yes	No	Yes
76692.00	60	Yes - FDE	Key Required	Yes	No	Yes
413632.00	59	Yes - FDE	Key Required	Yes	No	No
332008.00	59	No	N/A	Yes	No	No
75368.00	59	Yes - FDE	Key Required	Yes	No	No
406376.00	58	Yes - FDE	Key Required	Yes	No	No
326184.00	58	No	N/A	Yes	No	No
74048.00	58	Yes - FDE	Key Required	Yes	No	No
399120.00	57	Yes - FDE	Key Required	Yes	No	No

Page 1 of 3 Showing 50 rows per page

Create Cancel

You can change pool settings, but changes are generally not required. [Figure 8](#) shows various pool settings in SANtricity System Manager. Note the default settings for reconstruction and other background operations. Here the administrator can also change the preservation capacity equivalent number of drives. In this case, the default is 3, but the user can select any number from 0 up to a maximum of 10 drives or 20% of the drives, whichever is less.

Figure 8 Pool Settings in SANtricity System Manager

Pool Settings

Properties Settings

Name ?

MyPool

Capacity alerts

☒ Send me a critical alert when...  
- 85 % + of capacity has been allocated.

☐ Send me an early alert when...  
- 50 % + of capacity has been allocated.

Modification priorities

Critical reconstruction priority ?  
Highest (default)

Degraded reconstruction priority ?  
High (default)

Background operation priority ?  
Low (default)

Preservation capacity ?

Why can't I increase my preservation capacity?  
Increase preservation capacity by adding...  
- 3 + drive(s)

Save Cancel

After you create a pool, you can create volumes within the pool in much the same way as with traditional volume groups.

Figure 9 Create Volumes in SANtricity System Manager

The screenshot shows the 'Create Volumes' wizard in SANtricity System Manager, specifically step 3: 'Add/Edit Volumes'. The wizard has four steps: 1. Select Host, 2. Select Workload, 3. Add/Edit Volumes (current), and 4. Review. Below the steps, it says 'I want to create volumes from the following pool(s) and/or volume group(s)...'. A pool named 'MyPool' is selected, with details '(Optimal) (60 drives, 76692.00 GiB capacity)'. A legend indicates 'Allocated' (dark blue), 'Proposed' (light blue), and 'Free' (white). A bar chart shows the pool's capacity, with a 'Proposed' volume of 5000 GiB. Below the bar chart, there are checkboxes for 'Secure-capable Yes (not enabled)' and 'DA Yes'. A table lists the volumes to be created:

Volume Name	Reported Capacity
1	5000 GiB

At the bottom of the table, there is a '+ Add new volume' link. The bottom of the wizard has '< Back', 'Cancel', and 'Next >' buttons.

As with volume groups, pools and pool volumes can also be configured through the CLI or through the REST API.

## Comparison of DDP and Volume Groups

Pools and the volumes within them allow several operations that are similar to operations in traditional volume groups and offer some features that are unique to DDP technology.

[Table 1](#) shows a brief feature comparison between DDP and volume groups.

Table 1 Comparison of DDP and Volume Groups

Feature	DDP and Pool Volumes	Volume Groups and Volume Group Volumes
Snapshot technology	Yes	Yes
Volume copy	Yes	Yes
Synchronous mirroring	Yes except for the ETERNUS AB6100 and the ETERNUS HB1x00	Yes except for the ETERNUS AB6100
Asynchronous mirroring	Yes except for the ETERNUS AB6100	Yes except for the ETERNUS AB6100
Dynamic volume expansion	Yes	Yes
Online capacity expansion and reduction	Yes; add or remove up to 60 drives at a time	Partial; add a maximum of 2 drives, no reduction
Dynamic RAID migration	No; always RAID 6	Yes
Dynamic segment sizing	No; segments are always 128 KiB	Yes
Hot spare	Distributed preservation capacity	Dedicated global hot spare or spares
Dynamic redistribution	Yes; nondisruptive background operation	No; fragmentation and stranded capacity due to deletions
Drive evacuator	Yes	Yes
Shelf and drawer loss protection	Yes	Yes
SSD support	Yes	Yes



## Configuration Guidelines

[Table 2](#) lists several important considerations for when you configure pools.

Table 2 DDP configuration guidelines

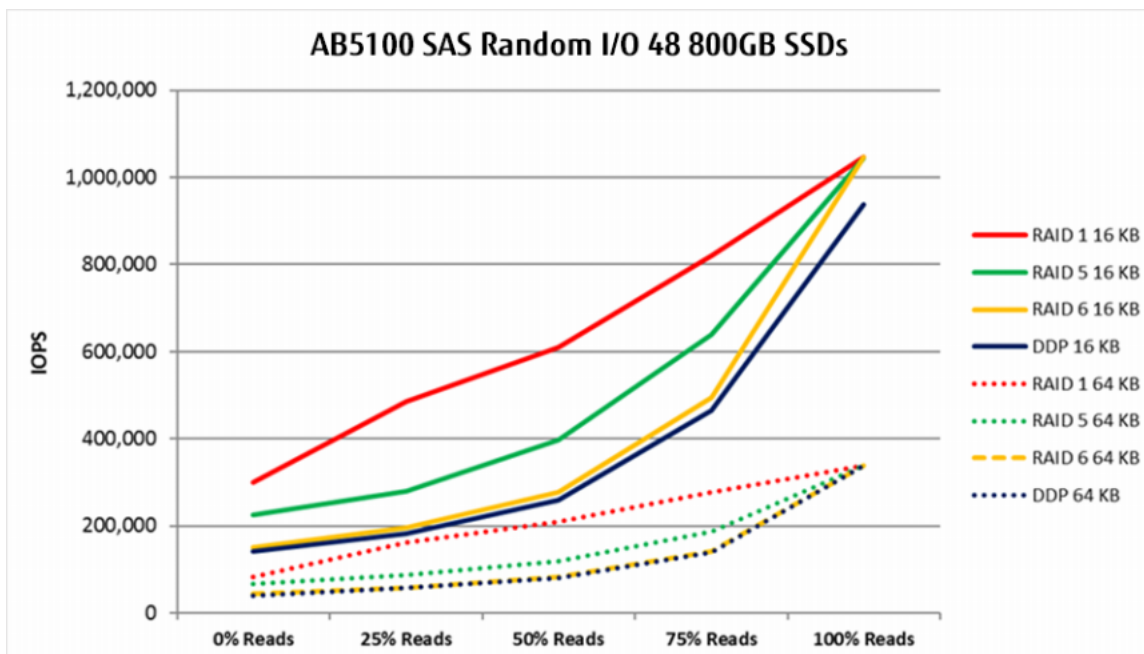
Description	Configuration
Maximum number of pools per system	20
Minimum number of drives per pool	11
Minimum number of drives per pool for drawer loss protection (DLP) with a single 4U60 shelf (*1)	15
Maximum pool capacity (*2) (sum of capacities of all pools in the system)	AB6100-12 PiB
Maximum volume size	All systems—4 PiB
Default preservation capacity by pool size (number of equivalent drives)	11 drives: 1 12–31 drives: 2 32–63 drives: 3 64–127 drives: 4 128–191 drives: 6 192–255 drives: 7 256–384 drives: 8 385–480 drives: 10
Drive types supported	SAS, NL-SAS, SSD, NVMe
<b>Caution</b> All drives in a pool must be of the same type and have the same characteristics (data assurance, security). To avoid losing the capacity of larger drives, all drives should have the same capacity.	
Online addition to or removal from a pool	Up to 60 drives at a time

- \*1: To maintain DLP as capacity is added to a disk pool, drives must be added in groups of five, one drive added per drawer in a 4U shelf.
- \*2: Maximum pool capacity includes RAID protection overhead, pool preservation capacity, usable capacity, and a small DDP-specific reserve based on the size of the pool.

### 3. Performance

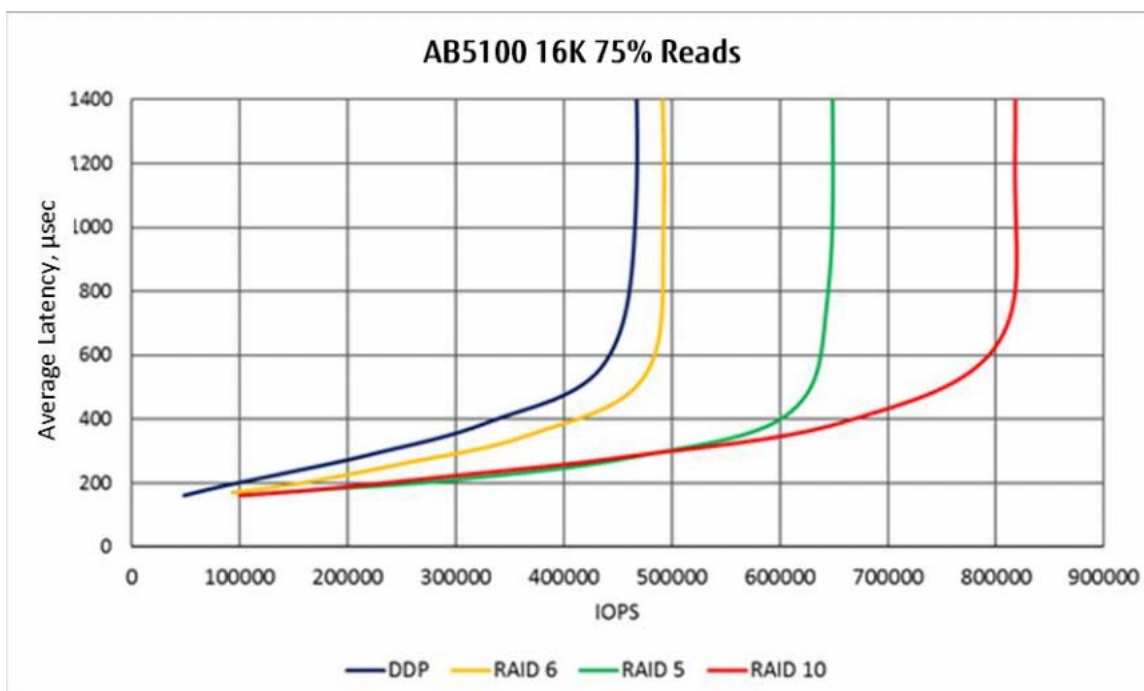
DDP configurations are generally not the highest-performing configuration. However, the delta when compared with RAID 5 and RAID 6 varies from insignificant with 100% read workloads to slightly less performance when running DDP with 100% write workloads. [Figure 10](#) provides a general performance comparison using an ETERNUS AB5100 running 16 KB and 64 KB I/O as the workloads change from 100% write to 100% read.

Figure 10 Block Size 16K/64K Performance (Read/Write)



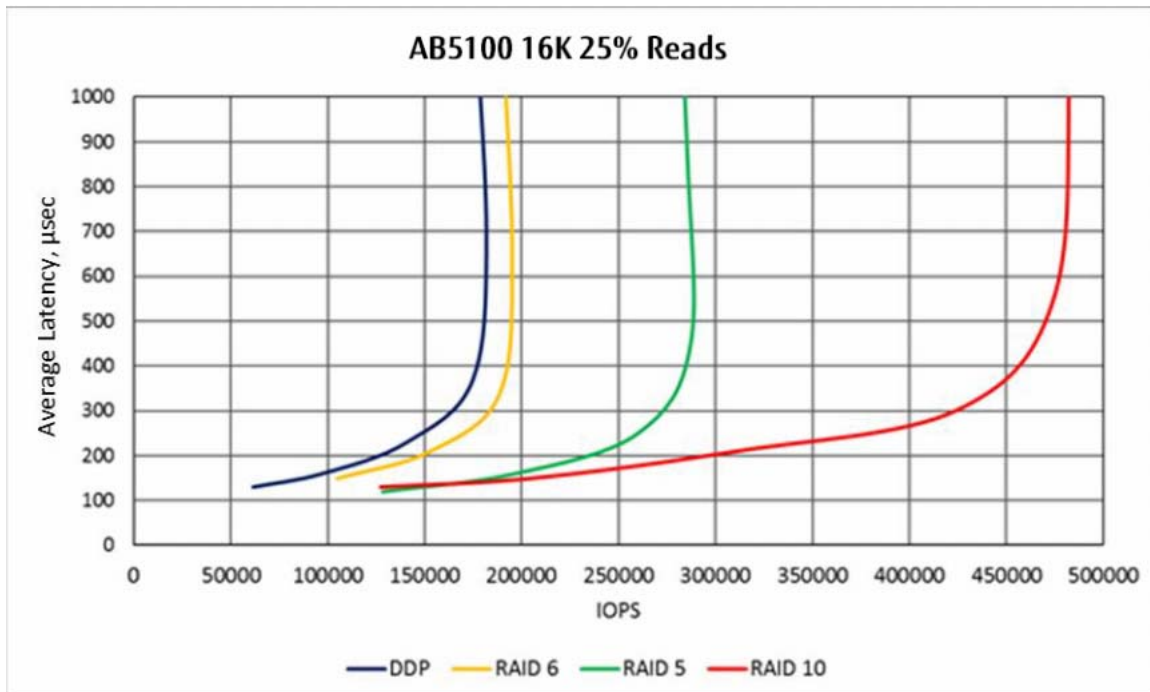
[Figure 11](#) shows a comparison of latency between DDP, RAID 5, RAID 6, and RAID 10 by using a 16 KB I/O size and a 75% read workload.

Figure 11 Block Size 16K Latency (Read:Write=3:1)



As the write component increases in the workload, the IOPS drop as expected, but the latency remains low, as shown in [Figure 12](#).

Figure 12 Block Size 16K Latency (Read:Write=1:3)



These results suggest that, when using DDP technology, IOPS and latency performance, especially at the lower end of the performance range, is very close to other standard RAID choices. At the high end of the performance range, standard RAID offers some performance advantage.

In most cases, performance comes down to a small trade-off of top-end performance for a significant improvement in drive rebuild time that is offered by pools of 30 drives or more.

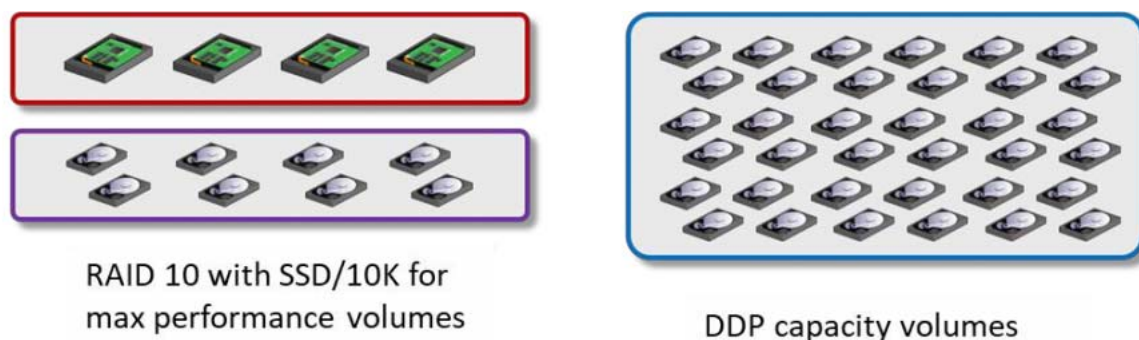
## 4. Best Practices for Configuring Pools

In general, it is a best practice to use DDP for homogeneous environments or for hosting more than one application on a single storage system. DDP technology is designed to perform best in a random workload environment, and several applications on a single storage system produce a random profile to the storage system.

For environments with one or more bandwidth-hungry applications (or applications require the sequential performance) that use the storage system, such as streaming media, configuring one traditional volume group per application is the best practice.

For storage systems with a mix of SSDs, high-performance HDDs, and NL-SAS drives that are used for both high-performance and high-capacity workloads, you might want to configure a mix of volume groups and pools. See [Figure 13](#) for an example.

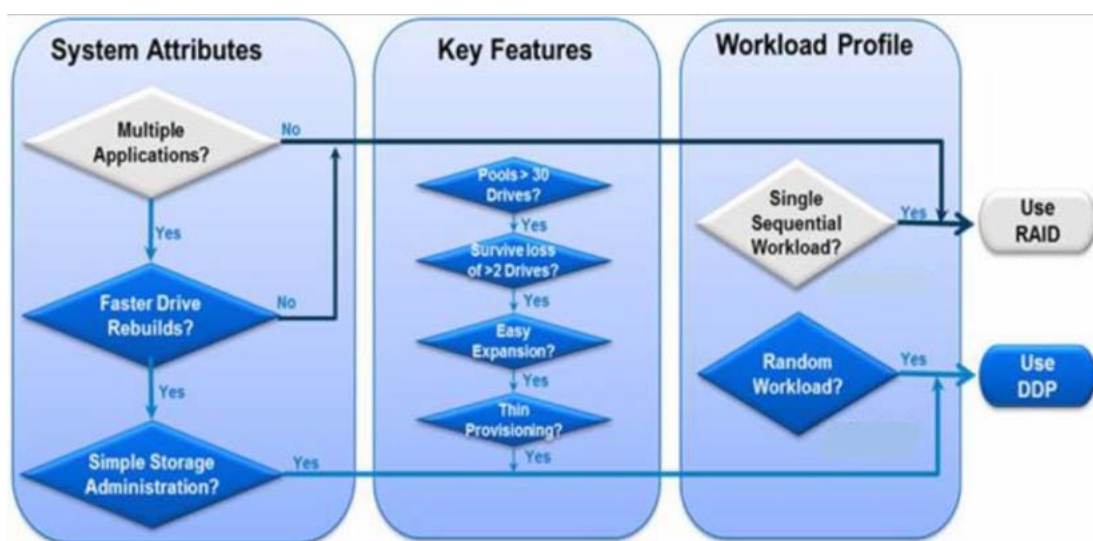
Figure 13 Mixing volume groups and pools



## Choosing Between DDP Technology and Traditional RAID

DDP technology offers several key features that you should consider when you select a storage virtualization technology. [Figure 14](#) compares DDP features with traditional RAID features.

Figure 14 Selecting DDP technology or RAID



#### Caution

If a double drive failure occurs, DDP technology rebuilds critical segments first, allowing the system to survive another drive failure within minutes of the initial double drive failure.

## How Large Should the Pool Be?

From an architectural standpoint, a pool can be expanded up to the total number of disks in the system, assuming that all drives are of the same type. Expansion is limited only by the maximum pool capacity (shown in [Table 2](#)) for the specific storage system model and software version. However, the best practice in multiworkload environments is to create multiple pools of up to 90 drives for midrange system performance requirements and up to 120 drives for low-range system performance requirements.

With all ETERNUS AB/HB series RAID choices, performance varies based on the drive type that is used; for example, SSDs are much faster than NL-SAS drives. This section provides best practice guidelines to help administrators, or anyone who is planning a proof of concept, to provision ETERNUS AB/HB series storage systems to achieve performance goals for IOPS, throughput, and drive rebuilds. Although it is possible to create even larger pools, for NL-SAS HDDs, 60 to 120 drives per pool maintains an optimal balance between performance and reliability.

A single large pool might be a better fit for long-term archive use cases. This approach keeps administrative overhead low while maintaining the data on a very reliable storage platform.

## Analytics Best Practices: Small-Block Random Workloads

Both traditional relational databases and NoSQL databases generate a largely random, small-block read/write workload. This is also the case for OLTP and hypervisor workloads. DDP technology is tuned to perform very well under these types of random workloads. In this type of environment, it is a best practice to put all volumes in a single pool. This approach simplifies system administration and database administration and meets the performance needs of both database log and data storage. Given the higher capacities of disks and the premium on floor space, it's no longer efficient to segment workloads and to isolate database log files to single RAID 1 volumes. Such segmentation creates islands of stranded storage and increases administrative overhead.

Another best practice is to ensure balance between the two RAID controllers, not just the number of volumes but also the function of the volumes. When you use SANtricity management software to create volumes, the system automatically balances the volumes between the controllers. The Automatic Load Balancing feature can also dynamically adjust volume ownership based on controller workload. The automatic load-balancing feature requires appropriate host type selection on Windows, VMware, or Linux (kernel 3.10 or higher).

## Backup and Video Surveillance Best Practices: Sequential Workloads

---

Consider a video surveillance workload that involves multiple cameras recording at once, each in a sequential manner. As multiple streams are written to different parts of the volume, the workload appears random to the storage system. For this scenario, DDP performs well, as discussed in the section "Analytics Best Practices: Small-Block Random Workloads."

For performance-sensitive workloads that are sequential and that typically have larger block transfer sizes, the best practice is to keep these workloads isolated and on their own traditional RAID volume group, rather than configuring pools. This type of workload describes many backup and surveillance applications. However, you should use performance sizing tools to determine whether volume groups are a requirement, given the rebuild benefits of DDP technology. Some backup implementations have shown that a DDP configuration meets the performance requirements while also delivering the rebuild benefits of DDP technology.

If performance is your main goal, the use of a traditional volume group with one volume is the best practice. If shorter rebuild times, better degraded-mode performance, and ease of administration are your goals, the use of pools of 12 to 30 drives can better meet your requirements.

## Technical Computing Best Practices: Large-Block Sequential Workloads

---

Technical computing workloads such as Lustre or IBM Spectrum Scale generate large-block sequential reads and writes, demanding the most from the underlying media. The best practice for this type of environment is to use a traditional RAID volume group with one volume per volume group.

## Configuring Pools with Equal Capacity Volumes

---

Because SANtricity System Manager and the CLI require the administrator to enter the desired capacity for each volume to be created, the user must calculate these capacities before configuring the system.

To begin this calculation, it is useful to know that the usable capacity of a pool is always a multiple of the D-stripe size. The calculation requires you to determine the number of D-stripes per volume and then multiply by the D-stripe size to obtain the number of gibibytes per volume. As an example, suppose that the pool has 161,456 GiB of usable capacity and the user wants to carve it into five equal volumes. The calculation is as follows:

- 1 Find the number of D-stripes in the pool:  $\text{Total capacity in GiB} / \text{D-stripe size} = 161,456 / 4 = 40,364$ .
- 2 Determine number of D-stripes per volume:  $40,364 / 5 = 8,072.8$ .
- 3 Round down to a whole number of D-stripes per volume = 8,072.
- 4 Multiply the number of D-stripes by 4 to obtain the gibibytes per volume:  $8,072 * 4 = 32,288$  GiB.

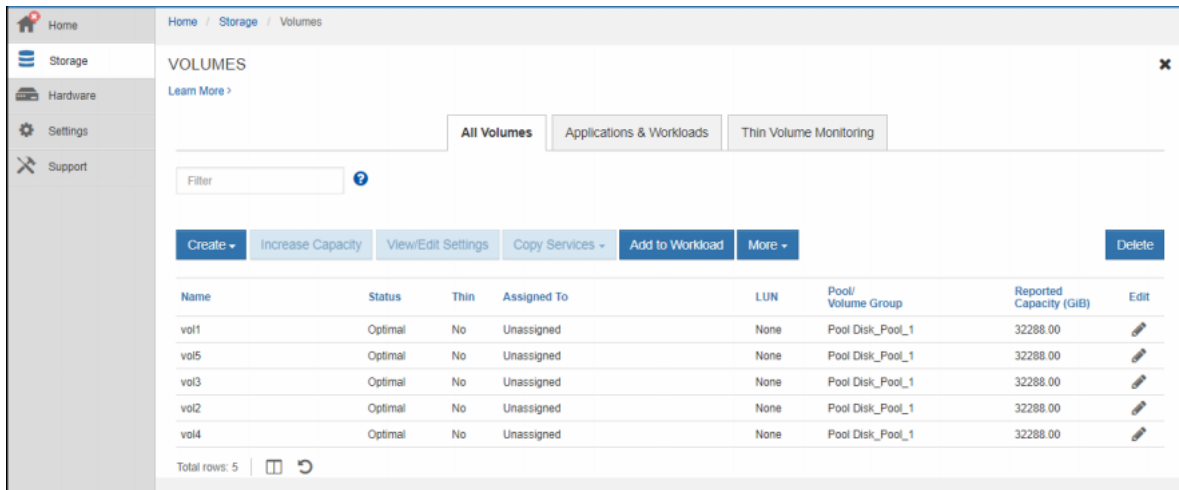
### Caution

In this example, four D-stripes are left over because of the remainder in step 2. If the total number of D-stripes is not divisible by the number of volumes that you want, you can't create equal volumes that use all the pool capacity.

#### 4. Best Practices for Configuring Pools Configuring Pools with Equal Capacity Volumes

[Figure 15](#) demonstrates this example in SANtricity System Manager, showing the five equal volumes in the pool. [Figure 16](#) shows the pool capacity and the allocated capacity. Note the free capacity of 16.00 GiB as described in this example. It is also important to note that for this simple calculation to work properly, preferences must be set to display capacity values in GiB.

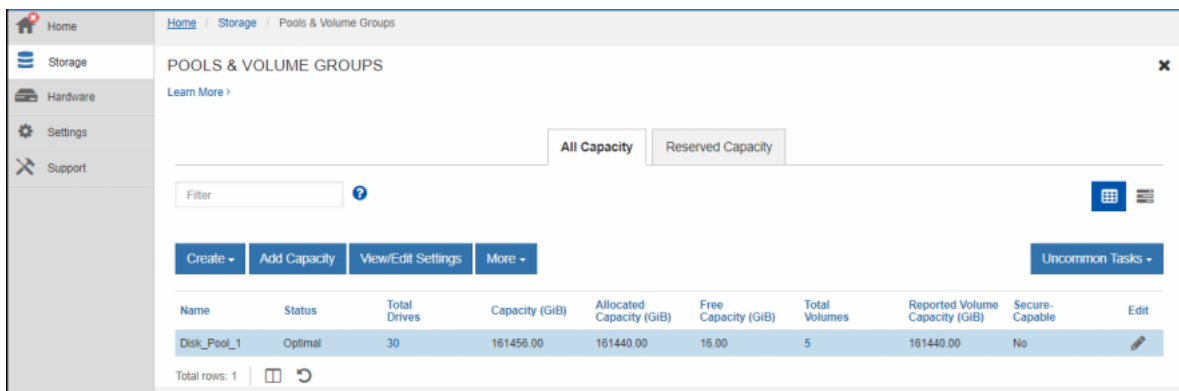
Figure 15 Equal volumes in SANtricity System Manager



The screenshot shows the 'VOLUMES' page in SANtricity System Manager. The left sidebar contains navigation links: Home, Storage, Hardware, Settings, and Support. The main content area has tabs for 'All Volumes', 'Applications & Workloads', and 'Thin Volume Monitoring'. Below the tabs is a filter input field and a row of action buttons: 'Create', 'Increase Capacity', 'View/Edit Settings', 'Copy Services', 'Add to Workload', 'More', and 'Delete'. A table lists five volumes (vol1, vol5, vol3, vol2, vol4) with columns for Name, Status, Thin, Assigned To, LUN, Pool/Volume Group, Reported Capacity (GiB), and Edit. All volumes are 'Optimal', 'No' thin, 'Unassigned', 'None' LUN, and belong to 'Pool Disk\_Pool\_1' with a capacity of 32288.00 GiB.

Name	Status	Thin	Assigned To	LUN	Pool/Volume Group	Reported Capacity (GiB)	Edit
vol1	Optimal	No	Unassigned	None	Pool Disk_Pool_1	32288.00	
vol5	Optimal	No	Unassigned	None	Pool Disk_Pool_1	32288.00	
vol3	Optimal	No	Unassigned	None	Pool Disk_Pool_1	32288.00	
vol2	Optimal	No	Unassigned	None	Pool Disk_Pool_1	32288.00	
vol4	Optimal	No	Unassigned	None	Pool Disk_Pool_1	32288.00	

Figure 16 Pool capacity with equal volumes in SANtricity System Manager



The screenshot shows the 'POOLS & VOLUME GROUPS' page. The left sidebar is the same as Figure 15. The main content area has tabs for 'All Capacity' and 'Reserved Capacity'. Below the tabs is a filter input field and a row of action buttons: 'Create', 'Add Capacity', 'View/Edit Settings', 'More', and 'Uncommon Tasks'. A table lists one pool (Disk\_Pool\_1) with columns for Name, Status, Total Drives, Capacity (GiB), Allocated Capacity (GiB), Free Capacity (GiB), Total Volumes, Reported Volume Capacity (GiB), Secure-Capable, and Edit. The pool is 'Optimal', has 30 drives, a capacity of 161456.00 GiB, allocated capacity of 161440.00 GiB, free capacity of 16.00 GiB, 5 total volumes, and a reported volume capacity of 161440.00 GiB.

Name	Status	Total Drives	Capacity (GiB)	Allocated Capacity (GiB)	Free Capacity (GiB)	Total Volumes	Reported Volume Capacity (GiB)	Secure-Capable	Edit
Disk_Pool_1	Optimal	30	161456.00	161440.00	16.00	5	161440.00	No	

Using the CLI to create equal volumes is similar to using SANtricity System Manager. To illustrate, consider an example of a disk pool with 229.996 TB of usable capacity and nine equal volumes. The following command displays the pool attributes:

```
show diskPool ["Disk_Pool_1"];
```

The portion of the result that shows capacity is as follows:

```
Total capacity:                234.346 TB
Preservation capacity:          4,455.000 GB(3 Drives)
Usable capacity:                229.996 TB
```

To create nine equal volumes, enter the following command nine times with different volume names:

```
create volume diskPool="Disk_Pool_1"
userLabel="vol1"
capacity=26168GB;
```

To see the result, display the pool attributes again:

```
show diskPool ["Disk_Pool_1"];
```

The values that were returned for capacity are as follows:

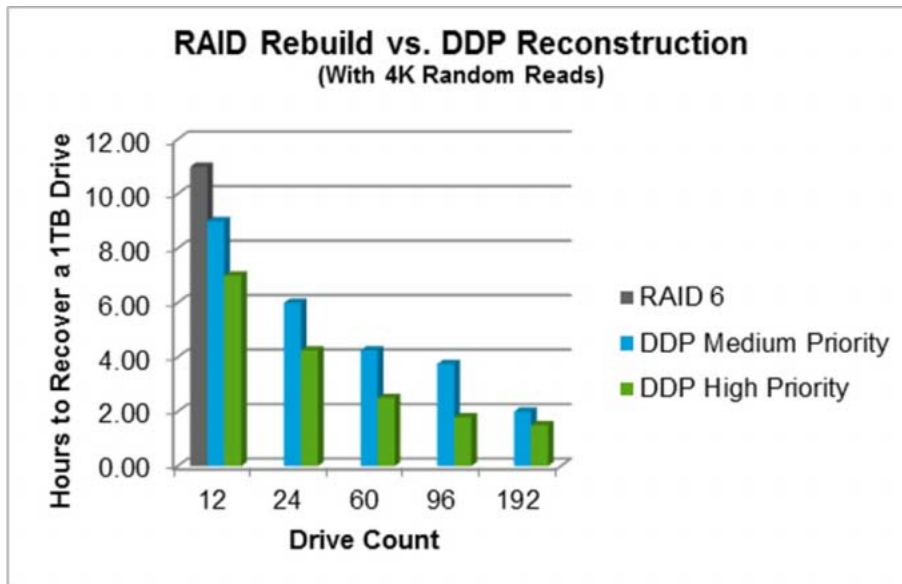
Total capacity:	234.346 TB
Preservation capacity:	4,455.000 GB(3 Drives)
Usable capacity:	229.996 TB
Unusable capacity:	0.000 MB
Used capacity:	229.992 TB
Volumes:	(9), 229.992 TB
Repositories:	(0), 0.000 MB
Free Capacity:	(1), 4.000 GB
Percent full:	99%

## Reconstruction Priority Setting

The reconstruction priority setting can be changed to optimize for the fastest rebuild times (highest-priority setting) or to minimize how storage system performance is affected during a drive rebuild (lowest-priority setting). The default setting and best practice are high priority, balancing a fast rebuild time with maintaining acceptable system performance.

The rebuild of a degraded pool is faster than the rebuild of a traditional volume group, as illustrated in Figure 17 [Figure 17](#), during this degraded state, the I/O performance to the host server or servers is affected very little compared with a traditional volume group.

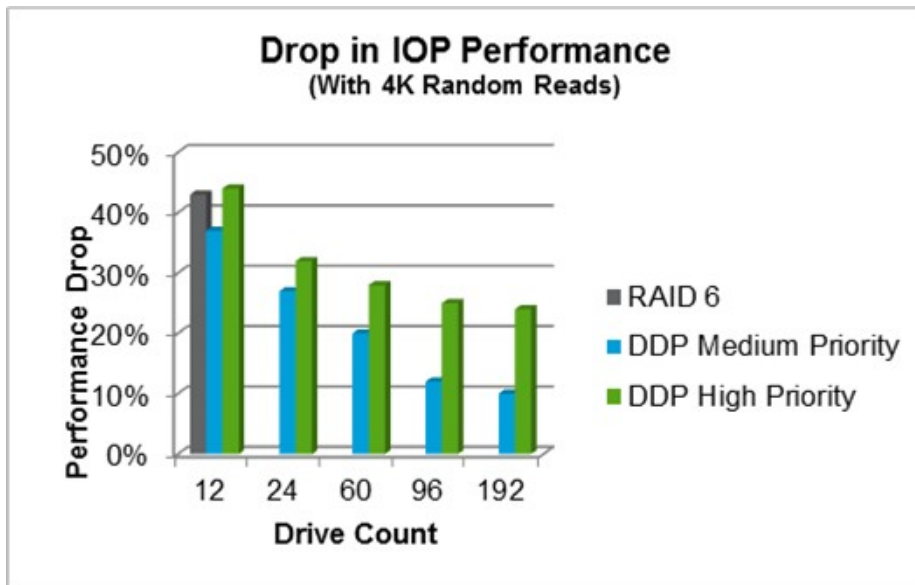
Figure 17 RAID rebuild versus DDP reconstruction



[Figure 18](#) shows an example of performance under various drive-failure scenarios. The chart reflects that, with larger disk pool sizes, there is less effect on the performance with DDP technology. Rebuilds and the transition from a degraded or critical state happen faster than those of a traditional volume group.



Figure 18 Drop in IOPS performance during rebuild



This performance delta occurs because DDP technology decouples drive loss protection from a total drive rebuild. The DDP feature rebalances the data across the remaining drives, affecting only a portion of the drives in the storage system that are required to calculate parity. Traditional RAID affects all the drives in the volume, reducing overall performance. The expected performance degradation and length of time to complete the required rebuild operations can vary based on many factors, including workload pattern and rebuild priority settings.

## 5. Summary

---

Drive capacities continue to increase and add to rebuild times for disk failures, leaving storage systems at the risk of data loss. Prolonged drive rebuilds negatively affect the performance of the system for extended periods of time, making it difficult for administrators to meet SLAs and affecting business application response needs. With the introduction of SANtricity Dynamic Disk Pools technology, the ETERNUS AB/HB series, administrators now have a choice that provides many advantages over traditional volume groups. This technology:

- Substantially improves reconstruction times (by up to 8 times) and limits exposure to additional drive failures by prioritizing critical segment reconstruction
- Reduces the performance impact during rebuild
- Enables online addition or removal of drives, up to 60 at a time
- Eliminates dedicated idle hot spares

The ETERNUS AB/HB series storage systems support both DDP and traditional volume groups. They can be mixed and matched in a single system to provide superior configuration flexibility to meet any workload requirement.

# A. Glossary of Terms

The following table defines terms that are used in this document that might be unfamiliar to the reader.

Term	Description
Critical segment reconstruction	When a D-stripe has two D-pieces that are affected by multiple drive failures, the system must re-create them on other drives in the pool. They are re-created by RAID 6 reconstruction of each of the segments within the affected D-pieces. These segments are called critical segments.
D-piece	The portion of a D-stripe that is contained on one drive.
D-piece size	The D-piece size depends on the platform. For all storage systems except the ETERNUS AB6100, a D-piece is 512 MiB in total capacity. For the ETERNUS AB6100, a D-piece is 1 GiB in total capacity.
Drawer loss protection (DLP)	When a pool has DLP, it can withstand failure or removal of one entire drawer of a drive shelf without losing availability.
D-stripe	The DDP element that comprises volumes. Each D-stripe resides on exactly 10 drives in the pool, regardless of pool size. D-stripes are distributed throughout a pool by an intelligent algorithm to optimize performance and failure tolerance.
D-stripe size	The D-Stripe size depends on the platform. For all storage systems except the ETERNUS AB6100, a D-stripe contains 4 GiB of user data and 1 GiB of RAID 6 parity. For the ETERNUS AB6100, a D-stripe contains 8 GiB of user data and 2 GiB of RAID 6 parity.
GB, MB, TB, PB	These units are used when referring to capacity by base 10 values. I/O throughput values and raw drive capacities are expressed in base 10 values.
GiB, MiB, TiB, PiB	These units are used when referring to capacity by base 2 values. Pool, volume group, and volume capacities are all expressed in base 2 values.
Logical unit number (LUN)	When a volume is assigned to a host, it acquires a reference number by which the host knows the volume.
Pool	The container in DDP technology that houses volumes.
RAID 6 stripe	A subset of a D-stripe that contains eight data segments and two parity segments. The P and Q parity segments are distributed among the drives that are involved in a D-stripe. There are 8,192 RAID 6 stripes in a D-stripe on the ETERNUS AB6100 and 4,096 RAID 6 stripes in a D-stripe on all other storage systems.
RAID group	An industry term that refers to containers that house volumes and that use traditional RAID. The ETERNUS AB/HB refers to them as volume groups.
Segment	The portion of a RAID 6 stripe that resides on a single drive. A segment is 128 KiB in size, and there are 8,192 segments in a D-piece on the ETERNUS AB6100 and 4,096 segments in a D-piece on all other storage systems.
Volume	The unit of capacity that is used to store application data. A volume resides in a volume group or in a pool and is assigned a LUN when it is associated with a host.
Volume group	The ETERNUS AB/HB series name for a RAID group.

## B. Thin Provisioning

---

The ETERNUS AB/HB series storage systems support thin provisioning when they are used with the DDP feature, but for most workloads, a thick volume is a better choice.

### Caution

Thin provisioning is not supported with RAID 10, RAID 5, RAID 6, or on the ETERNUS AB6100.

The thin provisioning feature provides overcommit capacity, but it does not deliver the same level of IOPS or throughput performance that thick volumes do. As a result, thin volumes are generally not recommended for the ETERNUS AB/HB series storage systems or for any transactional workload. The maximum thin-provisioned volume size is 256 TiB.

If you originally configured thin volumes and want to change to thick volumes, you must copy existing data to a new thick volume. Therefore, you should carefully plan the time to copy data, the associated cost of conversion capacity, and application cutover logistics.

---

FUJITSU Storage  
ETERNUS AB series All-Flash Arrays,  
ETERNUS HB series Hybrid Arrays  
SANtricity OS Dynamic Disk Pools  
Feature Description and Best Practices

P3AG-5672-01ENZ0

Date of issuance: November 2020  
Issuance responsibility: FUJITSU LIMITED

---

- The content of this manual is subject to change without notice.
- This manual was prepared with the utmost attention to detail.  
However, Fujitsu shall assume no responsibility for any operational problems as the result of errors, omissions, or the use of information in this manual.
- Fujitsu assumes no liability for damages to third party copyrights or other rights arising from the use of any information in this manual.
- The content of this manual may not be reproduced or distributed in part or in its entirety without prior permission from Fujitsu.

FUJITSU