

Dell Storage with Microsoft Storage Spaces Best Practices Guide



Notes, cautions, and warnings

-  **NOTE:** A NOTE indicates important information that helps you make better use of your computer.
-  **CAUTION:** A CAUTION indicates either potential damage to hardware or loss of data and tells you how to avoid the problem.
-  **WARNING:** A WARNING indicates a potential for property damage, personal injury, or death.

© 2016 Dell Inc. All rights reserved. This product is protected by U.S. and international copyright and intellectual property laws. Dell and the Dell logo are trademarks of Dell Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.

2016 - 05

Rev. A04

Contents

1 Introduction.....	5
2 Terminology.....	6
3 Hardware best practices.....	7
Recommended switch configuration.....	7
Storage.....	7
Rack requirements.....	7
4 Networking best practices.....	10
IP address requirements	10
Subnet requirements for SOFS configurations.....	10
Subnet requirements for Converged configurations.....	11
Network interface cards.....	11
Remote Direct Memory Access.....	12
5 Windows Server 2012 R2 best practices.....	13
Windows Server 2012 R2 versions.....	13
MPIO configuration.....	13
SMB considerations for SOFS configurations.....	14
UNMAP command.....	15
Configuring Windows Server updates.....	15
6 Microsoft Storage Spaces best practices.....	16
Storage pools.....	16
Configuring storage pools for rebuilds.....	16
Virtual disks (Microsoft Storage Spaces).....	18
Column count.....	19
Interleave size.....	22
Logical sector size.....	22
Storage tiers.....	23
Enclosure awareness.....	26
Write-back cache.....	27
Journal disks.....	28
Adding disk space to a storage pool.....	28
7 Cluster best practices.....	31
Cluster shared volumes.....	31

Cluster shared volumes cache.....	31
Cluster quorum.....	31
8 Scale-Out File Server best practices.....	33
Scale-Out File Share usage.....	33
Creating file shares.....	34
Deduplication.....	34
9 Converged best practices.....	35
Converged use cases.....	35
Logical processor sizing for Hyper-V virtual machines.....	35
Determining the number of available logical processors	35
Calculating logical processor requirements.....	35
Memory Sizing for Hyper-V virtual machines.....	36
Calculating memory requirements.....	36
Dynamic Memory.....	36
Allocating and configuring storage.....	37
Considerations in assigning CSV to specific VMs.....	37
Virtual Hard Disks (VHDX).....	37
Deduplication.....	38
Parity Spaces.....	38
Allocating and configuring network	38
Virtual switches (Hyper-V)	38
NIC Teaming.....	38
10 Getting help.....	39
Contacting Dell.....	39
Quick Resource Locator	39

Introduction

This document describes the best practices for the Dell Storage with Microsoft Storage Spaces (DSMS) solution. The DSMS is a software-defined storage (SDS) platform that uses Microsoft Storage Spaces and Dell servers, storage, and networking. You can deploy and configure the DSMS solution as a Scale-Out File Server (SOFS) or as a Converged solution.

When you configure the DSMS solution as a SOFS, the compute workloads are disaggregated from the storage. The compute workloads use Server Message Block (SMB) to access SOFS file shares on the storage nodes, which are direct SAS attached to the storage enclosures and provisioned and protected by clustered Storage Spaces.

Alternatively, you can configure the DSMS solution as a Converged solution. In this implementation, SOFS is not used and the compute workloads now run directly on the servers, which are direct SAS-attached to the storage enclosures. Clustered Storage Spaces are used to provision and protect the shared storage, which is used to store all the compute workload data.

This document is intended to support DSMS configurations, which have unique solution SKUs called Solution IDs. These DSMS solution IDs are required when ordering a DSMS configuration, and to access the benefits such as performance and sizing, optimized server and storage components, single-payload updates, and solution-level technical support. If you are in a geographical area outside of the US, the solution IDs are not available at this time. However, you may reference the configuration name instead. For the complete list of configuration solution IDs, see [Dell Storage with Microsoft Storage Spaces Configurations and Solution IDs](#) available at Dell Tech Center.

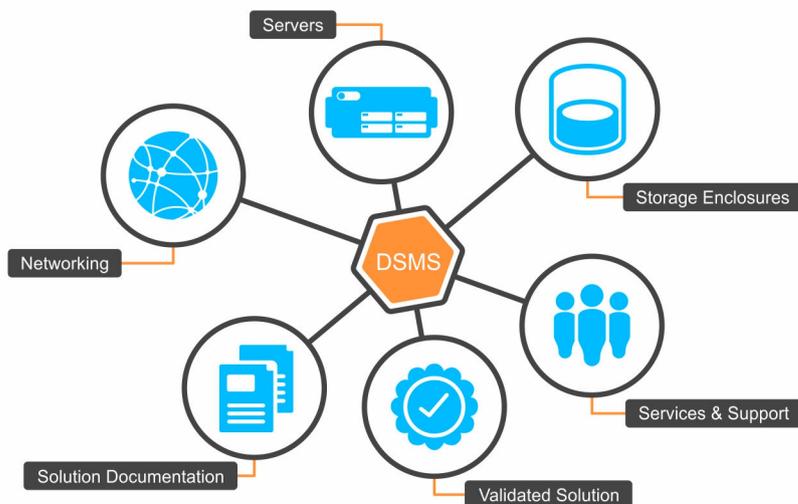


Figure 1. The correct solution ID connects you to all the solution has to offer, including Dell Services and Support.

Terminology

- Resiliency —Refers to the method by which a virtual disk (VD) is protected against physical disk failures.

There are three types of resiliency options for Microsoft Storage Spaces—Simple, Mirrored, and Parity.

- Simple VD —Data is striped across the physical disks and there is no resiliency provided. There is no protection against physical disk failures.
 - Mirrored VD—Data is striped across physical disks and an extra one or two copies of the data is also written to a second set of physical disk to protect against physical disk failures. A two-way mirror can tolerate one physical disk failure, while a three-way mirror can tolerate two physical disk failures.
 - Parity VD—Data is striped across the physical disk along with parity information. There are two types of parity options—single parity and dual parity. Single parity writes one copy of parity information protecting against one hard drive failure, while dual parity writes two copies and protects against two hard disk drive (HDD) failures.
- Scale-Out File Server (SOFS) —A role that enables a cluster to act as an active-active file server, which increases performance and provides highly available SMB file shares.
 - Hyper-V —A role that enables a server to host multiple guest virtual machines (VMs) that share the underlying hardware resources by virtualizing the processors, memory, and storage.
 - Storage node—A physical server, which is a member of a SOFS failover cluster.
 - Compute node—A physical server, which is a member of a failover cluster running a compute workload (for example, Hyper-V and SQL)
 - Storage pools —A group of physical disks, which allows you to manage and use the combined disk space of all the physical disks in the pool.
 - Storage Spaces —Virtual disks (VDs) created from free disk space in a storage pool. You can configure each Storage Space with a specific resiliency and storage tiers.
 - Storage tiers—Enables increased performance by creating and managing two separate data tiers — one SSD tier and one HDD tier. Frequently accessed data is stored on the high-speed SSD tier and less frequently accessed data is stored on the HDD tier.

Hardware best practices

Recommended switch configuration

For SOFS configurations, when integrating the solution in to your switch infrastructure, ensure that redundant network connections exist between the SMB clients and the SOFS storage nodes. SMB client communication to the storage nodes should use 10 GbE-capable switches. Dell recommends deploying two top-of-rack 10 GbE-capable switches to ensure physical switch and path redundancy for this solution. You must configure the switches for jumbo frames with the maximum transmission unit (MTU) frame size set to the highest MTU that the switch supports, typically 9 KB. Also, you must enable flow control on the switch for receiving but not for data transmission.

For Converged configurations, when deploying the smaller configurations with only two server nodes, a switch may not be required for networks used only for inter-node communication, such as heartbeat, CSV redirection, and Hyper-V live migration. For inter-node communication on two node configurations, you can use a direct-attached networking cable to connect the two server nodes together. Refer to the Support Matrix to verify that the adapters support direct connections between two nodes. However, for the larger configurations where a switch is required, Dell recommends two top-of rack switches to ensure physical switch and path redundancy. You must configure the switches for jumbo frames with the MTU frame size set to the highest MTU that the switch supports, typically 9 KB. Also, you must enable flow control on the switch for receiving but not for data transmission.

For more information about a list of the latest validated Dell switches for this solution, see the *Dell Storage with Microsoft Storage Spaces Support Matrix* available at [Dell.com/dsmsmanuals](https://www.dell.com/dsmsmanuals).

Storage

For optimal performance and redundancy, distribute SSDs evenly between each storage enclosure.

For example, for a 4x4 configuration with four DSMS 3060e enclosures and 48 total SSDs, each storage enclosure would contain 12 SSDs.

Rack requirements

When planning to deploy this solution, use Table 1 and Table 2 to estimate the physical space requirements required in the rack. Dell recommends that you to use a 48U rack with a depth of at least 1000 mm when deploying this solution to ensure ease of cable management. However, if the solution includes a DSMS 3060e Storage Enclosure, then a rack with 1200 mm depth is recommended.

The four major components of this solution are the client servers, switches, servers, and the storage enclosures. While the quantity of servers and storage enclosures are defined on the basis of solution that

was ordered, you can customize the client servers and switches to fulfill the application requirements. Each solution's rack space requirements are listed in the following tables.

Table 1. Rack space requirements for SOFS configurations

Server or storage solution	2x2	2x3	3x3	2x4	4x4
DSMS 630 with DSMS 3060e	10U	14U	15U	18U	20U
DSMS 730 with DSMS 1400 series	8U	10U	12U	12U	N/A
DSMS 730 with DSMS 3060e	12U	16U	18U	20U	N/A

Table 2. Rack space requirements for Converged configurations

Server or storage solution	2x1	2x2	2x3	3x3	4x3
DSMS 730 with DSMS 1400 series	6U	8U	10U	12U	14U

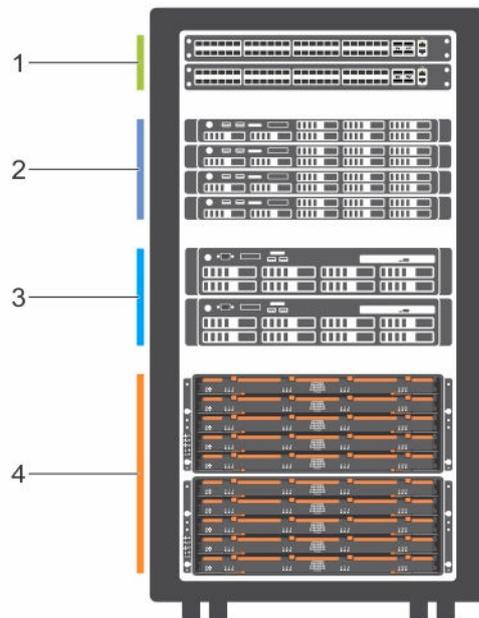


Figure 2. Example rack space usage for SOFS configuration

- | | |
|---|--|
| 1. Two Switches—Dell Networking S4810 (1U each) | 2. Four SMB Client Servers—DSMS 630 (1U each) |
| 3. Two Storage Nodes—DSMS 730 | 4. Two Storage Enclosures—DSMS 3060e (4U each) |

(2U each)

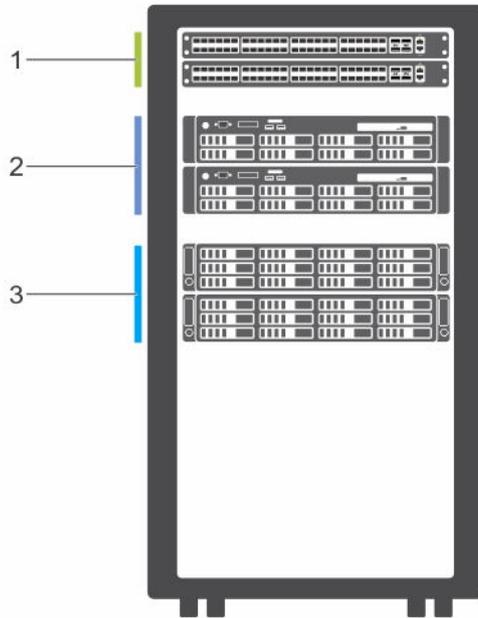


Figure 3. Example rack space usage for Converged configuration

1. Two Switches—Dell Networking S4810 (1U each)
2. Two Compute Servers—DSMS 730 (2U each)
3. Two Storage Enclosures—DSMS 1400 (2U each)

Networking best practices

IP address requirements

Dell recommends that you use static IP addresses for all network ports. Configurations may require up to seven IP addresses for each node for ease of management.

Subnet requirements for SOFS configurations

Dell recommends that you use five distinct subnets for Scale-Out File Server configurations. Two high-speed storage networks are used for external SMB client communication (L3-routable). Two additional subnets are used for cluster communication (do not require L3 routing). The fifth subnet is used for cluster management and Active Directory integration. This subnet can leverage an existing network infrastructure. If the solution requires high-availability, you must have each redundant pair of subnets routed through two physically separate switches.

 **NOTE:** Inside Failover Cluster Manager, ensure that only the SMB subnets (Subnet 2 and 3 in following table) are set to Cluster and Client. The server management and internal communication subnets must be set to Cluster Only. This is done to avoid using 1G network connections for high speed traffic.

Table 3. SOFS configurations -Subnet overview by port

	Static or DHCP address	Subnet mask	Gateway	DNS	Network
iDRAC server management	Either	X	X	X	Subnet 1 or existing network (L3-routable)
Cluster or server management	Either	X	X	X	Subnet 1 or existing network (L3-routable)
	Either	X	X	X	Subnet 1 or existing customer network (L3-routable)
External communication to SMB clients	Static	X	X	X	Subnet 2 (Layer 3-routable)
	Static	X	X	X	Subnet 3 (Layer 3-routable)

	Static or DHCP address	Subnet mask	Gateway	DNS	Network
					routable)
Internal communication between server cluster nodes	Static	X			Subnet 4
	Static	X			Subnet 5

Subnet requirements for Converged configurations

Converged network configurations change on the basis of workload that the solution handles. There is no single solution that can be commonly used for different networking infrastructures.

The following is a list of the different types of network traffic to consider when planning the deployment. Based on the network speed and infrastructure of your environment, you can consolidate many of these networks over one or more physical network interface cards (NICs).

Table 4. Type of network traffic for Converged

Network	Description
iDRAC server management	Network used to manage the physical server remotely using the Integrated Dell Remote Access Controller (iDRAC) utility, Web Interface.
Management	Network used to manage physical server such as Active Directory and Domain Name System integration.
External communication	Network used for communication to applications running on the solution, such as Hyper-V replica, external access to file shares, or VMs.
Internal Communication between server cluster nodes	Network used for the failover cluster, which provides communication for the heartbeat in addition to inter-node communication, such as CSV redirection and Hyper-V live Migration.

Network interface cards

Configure NICs on all clients and servers for jumbo frames with the MTU frame size set to 9 KB. Also, ensure that flow control is enabled on every NIC. Typically, flow control is enabled by default for NICs that support this feature.

To verify that flow control is enabled on every NIC:

1. By using Server Manager, open the **Network Connections** page. Else, enter the PowerShell command `<nepa.cpl>` at the Shell command line interface.
2. Select the NIC, and then click **Properties**.
3. In **Properties**, click **Configure**. In the dialog box, click the **Advanced** tab. Under **Property**, click **Flow Control**, and then select appropriate property from the **Value** drop-down menu.

Also, it is a best practice to rename all networks, in Windows Server Failover Cluster Manager, based on their function by using names such as Manage, External, or Cluster.

Remote Direct Memory Access

Remote Direct Memory Access (RDMA) allows for significantly increased throughput and low latency by performing direct memory transfers between servers.

For Scale-Out File Server configurations, SMB Direct is a feature of Windows Server 2012 R2 that enables the use of RDMA between SMB clients and the server. SMB Direct requires a network adapter, which supports the RDMA protocol. Also, both the storage nodes and the SMB clients must have RDMA adapters installed to use SMB Direct. For more information about a list of supported RDMA adapters, see the *Dell Storage with Microsoft Storage Spaces Support Matrix* available at Dell.com/dsmsmanuals.

Configure RDMA adapters on all SMB clients and servers for jumbo frames with the MTU frame size set to 9 KB. Also, verify that flow control is enabled on every NIC.

Also, when adding supported RDMA adapters to the solution there are several steps required in Windows Server 2012 R2 to ensure SMB Direct is enabled, existing SMB connections are refreshed, and SMB routing is configured (if supported by the adapter). For more information about configuring the RDMA adapters, see the *Dell Storage with Microsoft Storage Spaces Deployment Guide* available at Dell.com/dsmsmanuals.

For Converged configurations, RDMA is typically used for internal communication between server cluster nodes with heavy I/O profiles such as CSV redirection and Hyper-V live Migration. RDMA adapters should not be teamed or added to a Hyper-V virtual switch, because they will lose the RDMA functionality. Configure RDMA adapters on all servers for jumbo frames with the MTU frame size set to 9 KB. Also, ensure that flow control is enabled on every NIC.

Windows Server 2012 R2 best practices

Windows Server 2012 R2 versions

There are two different versions of Windows Server supported with DSMS solutions—Windows Server 2012 R2 Standard Edition and Windows Server 2012 R2 Datacenter Edition. The primary difference between the two editions, as it applies to the DSMS solution, is the licensing for the guest operating systems in the VMs. Standard Edition enables you to install 2012 R2 on the host and up to two VMs. Datacenter Edition enables you to install 2012 R2 on the host and unlimited VMs.

For Scale-Out File Server configurations, VMs are not installed on the storage nodes; therefore, it is recommended to use Windows Server 2012 R2 Standard Edition.

For Converged configurations, the operating system edition selected depends on whether the solution will be running VMs. If you intend to run VMs, Datacenter Edition on the servers is often the best value because it enables you install unlimited VMs running 2012 R2. However, if you do not intend to run VMs with the solution and will be running other applications or workloads then Standard Edition may be more cost effective.

MPIO configuration

Configuring Multipath I/O (MPIO) on each server enables Windows Server to use alternate paths to hard disk drive in the event of a failure and provides load balancing. When properly cabled and when using dual-port SAS physical disk drives, each server has two physical paths to each physical disk drive.

 **NOTE:** The following procedure must be completed before adding disks to a pool.

The recommended global MPIO load balancing policy for this solution is Least Blocks (LB).

Run the following PowerShell command on each server to enable LB:

```
Set-MSDSMGlobalDefaultLoadBalancePolicy -Policy LB
```

This sets the global value to LB. Any new drive added to the solution will be automatically configured as LB. Dell recommends using Least Blocks for both SSDs and HDDs for optimal performance in most applications.

To configure a different load balance policy for different groups of drives use the following code. Dell uses the drive model identifier to make quicker changes. If you have a mix of models of HDDs and SSDs you may have to go through and change them manually.

Run the following command to get the Target Hardware Identifier value:

```
mpclaim -e
```

Run the following command to set MPIO policy for the specified Hardware Identifier:

```
mpclaim -l -t "Target H/W Identifier" Policy Number
```

Policy Number options:

0 = Clear the Policy

1 = Fail-Over Only

2 = Round Robin

6 = Least Blocks

You must run the process for HDD and SSDs. There are other policy number options available but these are the only ones that Dell supports.

If a pool is already created, you must manually acquire the drive Hardware Identifier and set a policy for it.

SMB considerations for SOFS configurations

Dell requires running Microsoft Windows Server 2012 R2 on the servers. Dell recommends that you run Microsoft Windows Server 2012 R2 on all SMB clients. While you can use Microsoft Windows Server 2012 for SMB clients, Dell recommends that you use version R2 to ensure that the SMB clients have access to the new SMB features introduced in Microsoft Windows Server 2012 R2.

Table 5. New and updated features of SMB in Windows Server 2012 R2

Feature	Summary
Automatic rebalancing of Scale-Out File Server clients	Changes the functionality on how SMB clients connect to each scale-out file share. SMB clients are redirected to the node in the cluster, which provides the optimal path to the CSV. This reduces redirected I/O between the storage nodes.
Improved SMB event messages	Provides more information in the SMB event logging, which is used to assist in troubleshooting some network failure conditions.
Hyper-V Live Migration over SMB	Enables the use of advanced SMB features for Live Migration, such as SMB Direct and SMB Multi-channel.
Improved SMB bandwidth management	Expands SMB bandwidth management capability to control different types of SMB traffic.
Support for multiple SMB instances on a Scale-Out File Server	Provides capability for each storage node to be able to segregate intra-cluster CSV and SMB traffic in separate instances.
Improved performance of SMB Direct (SMB over RDMA)	Provides various performance improvements for small I/O workloads.

UNMAP command

Windows Server 2012 R2 periodically issues UNMAP commands to SSDs to optimize SSD storage disk space. However, SAS SSDs typically optimize automatically without being prompted by the OS. Dell recommends disabling UNMAP from being issued by the OS as this may result in more latency when processing these commands.

Run the following PowerShell command to disable UNMAP.

```
Fsutil behavior set disabledeletenotify 1
```

Configuring Windows Server updates

Dell recommends that you enable automatic updates in Windows Update or Windows Server Update Services (WSUS) to ensure that you are receiving the latest updates, which may affect many of the features in this solution.

For more information about required updates and related knowledge base articles, see the *Dell Storage with Microsoft Storage Spaces Support Matrix* available at [Dell.com/dsmsmanuals](https://www.dell.com/dsmsmanuals).

Microsoft Storage Spaces best practices

Storage pools

When setting up pools, for multiple enclosure configurations, ensure that the physical disks are spread evenly across all the enclosures by using multiple pools, rather than assigning a pool to one enclosure.

For example, if you have two DSMS 1420 enclosures, each with 20 HDDs and four SSDs, 40 HDDs and eight SSDs are available for configuring in the pool. In this example, you require a minimum of two pools. Pool 1 contains two SSDs and 10 HDDs from the first DSMS 1420, and two SSDs and 10 HDDs from the second DSMS 1420. Pool 2 contains the remaining disks.

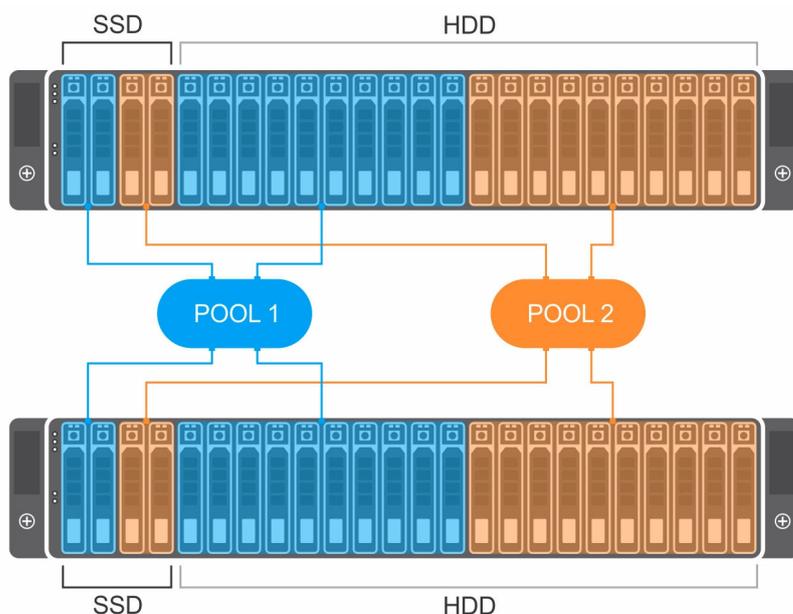


Figure 4. Distribution of physical disks across pools

Configuring storage pools for rebuilds

Microsoft Storage Spaces automatically rebuilds VDs by using free disk space and does not require assigning hot-spares. If a physical disk fails, the failed physical disk is regenerated from free storage disk space without user's intervention. The rebuild begins immediately after a disk failure is detected, limiting the exposure to additional physical disk failure while the VD is running in a degraded state. Dell recommends that you enable automatic rebuilds.

Run the following PowerShell command to enable automatic rebuilds (configured for each pool).

```
Set-StoragePool -FriendlyName <poolName> -RetireMissingPhysicalDisks Always
```

 **NOTE:** Before performing maintenance operations, and when storage pools are configured to automatic rebuild, prior to turning off a storage enclosure, you must change the `RetireMissingPhysicalDisks` attribute of all affected storage pools to `Never`. When the network communication with the drive becomes unavailable because the storage enclosure was turned off for maintenance, changing the attributes prevents the VDs from immediately initiating rebuilds. Ensure that you change `RetireMissingPhysicalDisks` back to the original value after the maintenance operations completes.

When a storage pool is configured for automatic rebuilds, it is important to maintain sufficient free disk space in each pool to enable VD's to automatically rebuild in event of a physical disk failure. The space required is equal to one complete drive with an additional 8 GB of space (for storage spaces overhead) multiplied by the number of disk that fail. This same equation is used for both the HDD and SSD tiers. Here are the equations for sizing the tiers.

Formula to calculate the amount of free space to reserve for automatic rebuilds per storage tier and per pool **without enclosure awareness:**

- Free capacity required in HDD tier per pool = (HDD disk capacity in TiB + 0.0078125 TiB) * (# of disks failures to recover from)
- Free capacity required in SSD tier per pool = (SSD disk capacity in TiB + 0.0078125 TiB) * (# of disks failures to recover from)

Formula to calculate the amount of free space to reserve for automatic rebuilds per storage tier and per pool **with enclosure awareness:**

- Free Space required in HDD tier per pool = (# of disk failures to recover from) * (disk capacity in TiB + 0.0078125) * (# of enclosures) / (# of enclosures - data copies + 1)
- Free Space required in SSD tier per pool = (# of disk failures to recover from) * (disk capacity in TiB + 0.0078125) * (# of enclosures) / (# of enclosures - data copies + 1)

There are two types of rebuild methods — parallel and sequential. During a parallel rebuild process, the data required to rebuild is obtained from multiple physical disks in the pool. Parallel rebuilds are very fast and reduce the time a VD is in a degraded state. However, because of multiple physical disks sourcing the repair data, there is some impact to storage I/O performance during rebuilds. A sequential rebuild only sources repair data from one physical disk in the pool at a time. This means that there is less impact to storage I/O performance during rebuilds. However, the rebuild process takes longer.

Dell recommends parallel rebuilds because it provides the fastest method to ensure all VDs return to their full resiliency. To optimize VD repairs the Microsoft November 2014 KB rollup (KB3000850) is required on all storage and compute nodes. Also, you must perform the tasks listed in the following URL:

<https://technet.microsoft.com/en-us/library/dn858079.aspx>

Run one of the following PowerShell commands for configuring the rebuild method (configured for each pool).

```
Set-StoragePool -FriendlyName <poolName> -RepairPolicy Parallel
```

```
Set-StoragePool -FriendlyName <poolName> -RepairPolicy Sequential
```

Virtual disks (Microsoft Storage Spaces)

When you configure a VD, there is an option to select from various resiliency types—simple, two-way mirror, three-way mirror, single parity, or dual parity.

Dell recommends you to use two-way or three-way mirror VDs. Mirrored VDs are optimized and provide the best performance and resiliency for Hyper-V workloads.

Parity VDs are intended for workloads, which are sequential, such as backup and archiving. Do not use parity VDs for random workloads, because this affects performance. Simple VDs are not recommended, because they provide no resiliency.

Table 6. Pool fault-tolerance and efficiency

Resiliency type	Disk space efficiency	Fault-tolerance for each storage pool	Example enclosures with 60 x 4 TB HDD (3.64 TiB)
Simple	100%	0 Disks	218.4 TiB raw disk space 218.4 TiB usable disk space
two-way mirror	50%	1 Disk	218.4 TiB raw disk space 109.2 TiB usable disk space
three-way mirror	33%	2 Disks	218.4 TiB raw disk space 72.07 TiB usable disk space
Parity	$\frac{\text{Column Count} - 1}{\text{Column Count}}$	1 Disk	Example with column count of 7: 218.4 TiB raw disk space 187.2 TiB usable disk space
Dual parity	$\frac{\text{Column Count} - 3}{\text{Column Count}}$	2 Disks	Example with column count of 7: 218.4 TiB raw disk space

Resiliency type	Disk space efficiency	Fault-tolerance for each storage pool	Example enclosures with 60 x 4 TB HDD (3.64 TiB)
			124.8 TiB usable disk space

 **NOTE:** Creating VDs with thin provisioning is not supported.

Fixed provisioning enables usage of storage tiers and failover clustering, neither of which are supported with thin provisioning.

Storage Space size is based on the number of disks and pools you have. Spread the VDs evenly across all pools, and assign them evenly to the nodes in the cluster to allow for load balancing and resiliency.

You can run the `New-VirtualDisk` Powershell command for more advanced VD creation options.

For example, a new VD called `exampleVD1` was created in storage pool `MyPool1`. This VD was a three-way mirror so the `ResiliencySettingName` attribute was set to `Mirror` and the `PhysicalDiskRedundancy` attribute was set to `2`. Because storage tiers are used, the disk space sizes of the tiers are specified as part of the VD.

```
New-VirtualDisk -FriendlyName "exampleVD1" -StoragePoolFriendlyName "MyPool1"
-ProvisioningType Fixed -ResiliencySettingName Mirror -PhysicalDiskRedundancy 2
-StorageTiers $ssd_tier, $hdd_tier -StorageTierSizes 400GB, 40TB
```

 **NOTE:** Dell recommends that you create storage spaces no larger than 10 TB. For more information, see the *Dell Storage with Microsoft Storage Spaces Support Matrix* available at Dell.com/dsmsmanuals.

Column count

The column count indicates the number of physical disks that Microsoft Storage Spaces stripes data across. The column count has a direct correlation to performance as increasing the column count enables more physical disks to be striped and accessed in parallel during read-and write operations.

You can only configure column count at VD creation by using PowerShell. You cannot set the column count by using the graphical user interface (GUI). Creating a VD in the GUI selects a default column count that may not be optimized for your solution. After a VD is created, the column count cannot be changed. For a VD that uses storage tiers, the column count of the SSD tier and HDD tier must be identical.

Run the following PowerShell command for creating VDs with a specified column count.

```
New-VirtualDisk -FriendlyName <vdName> -StoragePoolFriendlyName <poolName>
-ProvisioningType Fixed -ResiliencySettingName <Simple| Mirror| Parity>
-PhysicalDiskRedundancy <1|2> -NumberOfColumns <#ofColumns>
-StorageTiers<ssdTierObject, hddTierObject> -StorageTierSizes <ssdTierSize ,
hddTierSize>
```

For example, a new VD is created called exampleVD2 in the storage pool MyPool1. This VD is a two-way mirror. The column count is six for this VD so the NumberOfColumns attribute is set to 6.

```
New-VirtualDisk -FriendlyName "exampleVD2" -StoragePoolFriendlyName "MyPool1"
-ProvisioningType Fixed -ResiliencySettingName Mirror -PhysicalDiskRedundancy 1
-NumberOfColumns 6 -StorageTiers $ssd_tier, $hdd_tier -StorageTierSizes 100GB,
15TB
```

Table 7. Column count requirements

Resiliency type	Minimum number of columns	Column-to-disk correlation	Minimum number of disks
Simple	1	1:1	1
Two-way mirror	1	1:2	2*
Three-way mirror	1	1:3	3*
Dual Parity	7	1:1	7
Single Parity	3	1:1	3

 **NOTE:** * This minimum does not account for any overhead in case of disk failure.

When sizing the column count of a VD, consider the following.

Will the VDs be configured to automatically rebuild?

When a VD fails and an automatic rebuilds are enabled, the VD attempts to repair the degraded VD by using existing free disk space in the pool. However, for an automatic rebuild to occur, the VD must not only have free disk space but also sufficient free disks available to restore the intended resiliency level while maintaining the original column count.

To account for this necessity, we must subtract a number of repair disks from the available number in the pool before calculating the column count. For the equations listed below, Dell recommends that this number be set to 1 or 2, in order to keep column count optimal. If you do not intend to enable automatic rebuilds for the VD, subtracting is not required.

Is enclosure awareness enabled for the VD?

When enclosure awareness is enabled, the data copies are spread across the enclosures to allow for failure of a complete enclosure, when also maintaining access to the data. Sufficient free disks to rebuild the VDs must be available in the remaining enclosures, when also maintaining the original column count.

 **NOTE:** Results from the column count equations are rounded down to the nearest whole number. Also, any result where the column count is zero or less than zero indicates an invalid SSD configuration for the pool. Add more SSDs to the pool to achieve a valid column count.

For mirrored VDs:

Without enclosure awareness

$$\text{Column Count} = \frac{\# \text{ of SSDs in the pool} - \# \text{ of Automatic Repair Disks}}{\# \text{ of Data Copies}}$$

With enclosure awareness

$$\text{Column Count} = \frac{\# \text{ of SSDs in the pool}}{\# \text{ Number of Enclosures}} - \# \text{ of Automatic Repair Disks}$$

 **NOTE:** Number of data copies are 2 for two-way mirrored spaces and 3 for three-way mirrored spaces.

For dual parity VDs:

Without enclosure awareness

$$\text{Column Count} = (\# \text{ of disks} - 2)$$

With enclosure awareness

$$\text{Column Count} = (2 * \text{Number of Enclosures}) - 1$$

For example, if you have three DSMS 3060e storage enclosures each with 12 x SSDs and 48 x HDDs, and decide to create two storage pools—MyPool1 and MyPool2, this is how you would calculate the column counts for the VDs in different situations.

Table 8. two-way mirror without enclosure awareness example

Pool Name	MyPool1
HDDs	72
SSDs	18
VD name	columnExample1
Resiliency	two-way Mirror
Enclosure awareness	No
Automatic rebuild	Yes
Column count	8

Table 9. three-way mirror with enclosure awareness example

Pool Name	MyPool2
HDDs	72
SSDs	18
VD name	columnExample2
Resiliency	three-way mirror
Enclosure awareness	Yes
Automatic rebuild	Yes
Column count	4

Interleave size

Microsoft Storage Spaces optimizes performance by striping data across multiple physical disks. The stripe size (interleave size) is set by default to 256 KB. This means that Microsoft Storage Spaces stores 256 KB of data per stripe on each disk. You can configure the stripe size when creating a new VD but only by using PowerShell.

To maximize performance, ensure that the interleave size is as large as the most common I/O of your workload. I/O data that exceeds the interleave size is split into multiple stripes, turning one write operation into multiple write operations and reducing performance. Interleave size is set in bytes.

Run the following PowerShell command to create a VD with a specific interleave size.

```
New-VirtualDisk -FriendlyName <vdName> -StoragePoolFriendlyName <poolName>
-ProvisioningType Fixed -ResiliencySettingName <Simple| Mirror| Parity>
-PhysicalDiskRedundancy <1|2> -NumberOfColumns <#ofColumns> -Interleave
<#ofBytes> -StorageTiers <ssdTierObject, hddTierObject> -StorageTierSizes
<ssdTierSize , hddTierSize>
```

For example, a new VD is created called exampleVD3 in storage pool MyPool1. This VD is a two-way mirror. The most common I/O size for the workload is 64 KB. For this VD, the `Interleave` attribute is set to 65536 (64 KiB converted to bytes).

```
New-VirtualDisk -FriendlyName "exampleVD3" -StoragePoolFriendlyName "MyPool1"
-ProvisioningType Fixed -ResiliencySettingName Mirror -PhysicalDiskRedundancy 1
-NumberOfColumns 4 -Interleave 65536 -StorageTiers $ssd_tier, $hdd_tier
-StorageTierSizes 50GB, 8TB
```

Logical sector size

Older hard drives had a 512B physical sector size, while newer drives have a 4 KB physical sector size and either a 512B logical sector size (512e Drive) or a 4 KB logical sector size (4Kn Drive). If you add or replace disks in your storage pool with a 4Kn drive in the future, Dell recommends that you set your storage

pool's default logical sector size to 4 KB. You cannot add 4Kn drives to a storage pool using a logical sector size of 512B.

To determine the sector size of your current physical disks, run the following PowerShell command.

```
Get-PhysicalDisk | Sort-Object SlotNumber | Select-Object SlotNumber,  
FriendlyName, Manufacturer, Model, PhysicalSectorSize, LogicalSectorSize |  
Format-Table
```

To create a storage pool with a logical sector size of 4 KB, use the command `New-StoragePool` along with the parameter `"-LogicalSectorSizeDefault 4KB"`

Storage tiers

Storage tiers enable the mixing of HDDs and SSDs in one pool to take advantage of the faster SSDs maximizing IOPS and throughput by using the SSDs for the most frequently accessed data. Microsoft Storage Spaces constantly analyzes data usage patterns and moves the most frequently used data from the HDD tier to the SSD tier to speed up access.

 **NOTE:** Storage tiers are only supported with mirroring or simple spaces.

Microsoft Storage Spaces creates a heat map based on the frequency that the data is used. Once a day, an optimization process automatically runs and the most frequently accessed data (hot data) is moved to the SSD tier and less frequently accessed data (cold data) is moved to the HDD tier.

Because the data on the SSD tier is updated only once a day (by default), if required, you can manually optimize data by running the following PowerShell command.

```
defrag.exe /C /H /K /G
```

You must run this command on all nodes in the cluster, because it only optimizes the VDs owned by the node where the command was run.

When validating the performance of your solution, perform benchmarks over the course of several days to allow the SSD tier to optimize for your workload. Alternatively, you can manually optimize the SSD tier. Manually optimizing storage tier data using the command above generates a Storage Tier Optimization Report. The Storage Tier Optimization report provides data on the storage tier, which can be used to identify methods to optimize performance.

You may want the ability to permanently place a frequently used file on the SSD tier. For example, a frequently accessed VHDX file that requires low latency and high throughput. You can accomplish this by pinning files to the SSD tier.

Consider the following before running the command:

- You must run the command from the node, which owns the CSV where the file is stored.
- You must use the local path of the CSV on the node.

Even after pinning the file, it will not move to the SSD tier until the next optimization (or if you run it manually).

Dell recommends you to use pin sparingly, because the primary goal of storage tiers is to allow the heat map process to optimize the tiers.

Run the following Powershell command to pin files to the SSD tier.

```
Set-FileStorageTier -FilePath <localFilePath>  
-DesiredStorageTierFriendlyName<ssdTierName>
```

Run the following Powershell command to unpin files from the SSD tier.

```
Set-FileStorageTier -FilePath <localFilePath>
```

Run the following PowerShell command to review all files currently pinned.

```
Get-FileStorageTier -VolumePath <csvVolumePath>
```

For example, pinning a VHDX file called myVHDX, which is on exampleShare on a CSV labeled Volume3. The storage tier is called MyPool1_SSD.

```
Set-FileStorageTier -FilePath "C:\ClusterStorage\Volume3\Shares\exampleShare  
\myVHDX.vhdx" -DesiredStorageTierFriendlyName "MyPool1_SSD"
```

Storage tiers are configured for each storage pool and it is recommended to create one SSD tier and one HDD tier for each storage pool.

The following equations are for planning purposes, when you start creating VDs in your storage pools. The values from the equations are used to understand the upper limit for capacity planning so that you are always leaving sufficient HDD and SSD raw disk space free across the pool as you create and size each VD for automatic rebuilds. These are not intended to be equations to guide you about sizing the tiers of the individual VDs. If you are not using automatic rebuilds and plan to only perform manual rebuilds, subtracting the capacity for automatic rebuilds in the equations below is not required..

*Raw SSD Usable Capacity Per Pool=(Number of SSDs in the pool*SSD Capacity)-(SSD Capacity+8 GB for automatic rebuild space)*(Number of disk failures to recover from)*

*Raw HDD Usable Capacity Per Pool=(Number of HDDs in the pool*HDD Capacity)-(HDD Capacity+8 GB for automatic rebuild space)*(Number of disk failures to recover from)*

Tier size is the raw value modified by a resiliency factor depending on the type of Storage Space you are creating. For example, 1 for a Simple Space, 1/2 for a two-way Mirror, and 1/3 for a three-way mirror.

SSD Tier Per Pool= Raw SSD Usable Capacity Per Pool

HDD Tier Per Pool=Raw HDD Usable Capacity Per Pool

$$SSD \text{ Tier per VD} = \frac{SSD \text{ Tier per Pool}}{Number \text{ of Virtual Disks}}$$

$$HDD \text{ Tier per VD} = \frac{HDD \text{ Tier per Pool}}{Number \text{ of Virtual Disks}}$$

For example, a pool is created with 48 x 4 TB and 12 x 800 GB disks called tierPool1. Only two VDs exist in this pool, one with two-way mirroring called 2wayVD1 and one with three-way mirroring called 3wayVD1.

2wayVD1 HDD tier was 42.7 TiB, which uses 85.5 TiB of the HDD tier because of the 50 percent resiliency overhead for two-way mirroring. The SSD tier was 2 TiB in size, which uses 4 TiB of SSD tier because of resiliency overhead.

3wayVD2 HDD tier was 27.9 TiB, which uses 55.8 TiB of the HDD tier because of the 66 percent resiliency overhead from three-way mirroring. The SSD tier was 1.2 TiB, which uses 2.4 TiB of SSD tier because of resiliency overhead.

After both VDs are created, there is 33.42 TiB of disk space remaining in the HDD tier and 2.3 TiB of disk space remaining in the SSD tier.

Table 10. Example of disk space values used for storage tiers for pools and VDs

Pool Name	tierPool1
HDDs in pool	48
HDD disk space	4 TB (3.64 TiB)
HDD tier disk space (after automatic rebuild)	141.3 TiB
SSDs in pool	12
SSD disk space	800 GB (745 GiB)

Table 11. Example of disk space values used for storage tiers for pools and VDs

VD name	2wayVD1
Resiliency	two-way mirror
HDD tier disk space	42.76 TiB
SSD tier disk space	2 TiB
SSD tier disk space (after automatic rebuild)	8.18 TiB

Table 12. Example of disk space values used for storage tiers for pools and VDs

VD name	3wayVD2
Resiliency	three-way mirror
HDD tier disk space	27.9 TiB
SSD tier disk space	1.23 TiB

Run the following PowerShell commands when creating a new SSD and HDD storage tier (configured for each pool).

```
New-StorageTier -StoragePoolFriendlyName <poolName> -FriendlyName <ssdTierName>
-MediaType SSD
```

```
New-StorageTier -StoragePoolFriendlyName <poolName> -FriendlyName <hddTierName>
-MediaType HDD
```

For example,

```
New-StorageTier -StoragePoolFriendlyName "tierPool1"
-FriendlyName"tierPool1_SSD" -MediaType SSDNew-StorageTier
-StoragePoolFriendlyName "tierPool1" -FriendlyName"tierPool1_HDD" -MediaType HDD
```

Storage tiers are sized for each VD when the VD is created. You can size the SSD and HDD tiers in the GUI during VD creation or by using PowerShell.

Run the following PowerShell command for creating a VD by using existing storage tiers.

```
New-VirtualDisk -FriendlyName <vdName> -StoragePoolFriendlyName <poolName>
-ProvisioningType Fixed -ResiliencySettingName <Simple| Mirror| Parity>
-PhysicalDiskRedundancy <1|2> -StorageTiers <ssdTierObject, hddTierObject>
-StorageTierSizes <ssdTierSize , hddTierSize>
```

For example, a new VD is created called 2wayVD1 in the storage pool tierPool1. This VD is configured as a two-way mirror. The SSD and HDD tiers for this pool were created in the earlier example. Because the StorageTiers attribute requires objects as its input, the output of the Get-StorageTier PowerShell command is assigned to the \$ssd_tier and \$hdd_tier variables, and then used when creating the VD. The StorageTierSizes attribute is set to the size of each tier for which the VD uses the tier disk space.

```
$ssd_tier = Get-StorageTier -FriendlyName tierPool1_SSD
```

```
$hdd_tier = Get-StorageTier -FriendlyName tierPool1_HDD
```

```
New-VirtualDisk -FriendlyName "2wayVD1" -StoragePoolFriendlyName "tierPool1"
-ProvisioningType Fixed -ResiliencySettingName Mirror -PhysicalDiskRedundancy 1
-StorageTiers $ssd_tier, $hdd_tier -StorageTierSizes 2TB, 42.7TB
```

Enclosure awareness

Enclosure awareness provides fault-tolerance for an entire enclosure failure, by ensuring that data copies are spread across the available enclosures, such that the loss of an entire enclosure still allows access to the data. Enclosure awareness requires at least three storage enclosures.

Table 13. Enclosure configurations for failure coverage

Resiliency level	Storage enclosure failure coverage		
	Two storage enclosures	Three storage enclosures	Four storage enclosures
two-way mirror	1 disk	1 enclosure	1 enclosure
three-way mirror	2 disks	1 enclosure + 1 disk	1 enclosure + 1 disk

Resiliency level	Storage enclosure failure coverage		
	Two storage enclosures	Three storage enclosures	Four storage enclosures
Parity	1 disk	1 disk	1 disk
Dual parity	2 disks	2 disks	1 enclosure + 1 disk

Enclosure awareness is configured when you create a VD.

Run the following PowerShell command for enabling enclosure awareness.

```
New-VirtualDisk -FriendlyName <vdName> -StoragePoolFriendlyName <poolName>
-IsEnclosureAware <${true}|$false> -ProvisioningType Fixed
-ResiliencySettingName<Simple| Mirror| Parity> -PhysicalDiskRedundancy <1|2>
-StorageTiers<ssdTierObject, hddTierObject> -StorageTierSizes <ssdTierSize ,
hddTierSize>
```

For example, a new VD is created with the name exampleVD3 in storage pool MyPool1. This VD uses enclosure awareness so the `IsEnclosureAware` attribute is set to `$true`.

```
New-VirtualDisk -FriendlyName exampleVD3 -StoragePoolFriendlyName MyPool1
-IsEnclosureAware $true -ProvisioningType Fixed -ResiliencySettingName Mirror
-PhysicalDiskRedundancy 1 -StorageTiers $ssd_tier, $hdd_tier
-StorageTierSizes20GB, 10TB
```

Write-back cache

Configuring write-back cache (WBC) allows Microsoft Storage Spaces to use a portion of available disk space of the SSDs to provide low latency cache for incoming write commands. All incoming write requests are cached on SSDs and host immediately receives a write completion status. Without write-back cache, the physical disks operate in write-through mode where incoming write requests hit slower, traditional HDDs before responding with completion. Using WBC can make a significant performance improvement to almost all workloads.

To use WBC, you must have SSDs in a storage pool. The default disk space of WBC is 1 GB and can be changed when creating a VD. Dell recommends that you use the default WBC disk space.

Whenever you create a VD, as long as the SSD tier disk space is greater than either the default or specified WBC disk space and you have the minimum number of SSDs in the storage pool to enable WBC for the resiliency setting and WBC is automatically enabled for the VD.

Table 14. WBC Resiliency Setting

Resiliency setting	Simple	two-way mirroring	three-way mirroring	Single parity	Dual parity
Minimum number of SSDs for WBC (per pool)	1 SSD	2 SSDs	3 SSDs	2 SSDs	3 SSDs

Journal disks

Assigning dedicated journal disks to SSDs in the pool when using VDs with parity resiliency improves sequential write performance. Without dedicated journal disks, the journal disks reside on the same physical disks as the parity VD and this can cause increased seek times and reduce overall throughput of a parity VD.

When using a dedicated journal disk, ensure that the quantity of journal disks scales with the number of parity spaces in the pool. Dedicated journal disks can only be added by using PowerShell.

Run the following PowerShell command for adding a dedicated journal disk to an existing pool.

```
Add-PhysicalDisk -StoragePoolFriendlyName <poolName> -PhysicalDisks  
<physicalDiskObject> -Usage Journal
```

For example, a dedicated journal disk which is an SSD labeled `PhysicalDisk5` is assigned to variable `$pd`. The disk is then added to an existing pool `MyPool1` with the `Usage` attribute set to `Journal`.

```
$pd = Get-PhysicalDisk -CanPool $true -FriendlyName PhysicalDisk5  
Add-PhysicalDisk -StoragePoolFriendlyName "MyPool1" -PhysicalDisks $pd -Usage  
Journal
```

Adding disk space to a storage pool

When planning to add physical disk space to existing storage pools and VD, there are several considerations to make before adding the additional physical disks or storage enclosures. For more information about the expanding an existing DSMS solution, see the *Dell Storage with Microsoft Storage Spaces Deployment Guide*.

When adding a new storage enclosure to the SOFS cluster or just new physical disks to an existing storage enclosure follow these guidelines:

- For an up-to-date list of validated storage enclosures and physical disks to confirm that the new configuration is supported, see the *Dell Storage with Microsoft Storage Spaces Support Matrix*.
- For ensuring that you are following the proper cabling guidance, if adding a new storage enclosure, see the *Dell Storage with Microsoft Storage Spaces Cabling Guide*.

After verifying all newly added disks are available to the cluster, you can now create new or expand an existing storage pool or VD. Dell recommends when expanding a storage pool to add physical disks in a quantity equal to the column count multiplied by the number of data copies plus any additional disks required for automatic rebuilds. For example, for a two-way mirror VD, if the column count is four, you must add a minimum of eight disks to the pool to expand the VD.

Run the following PowerShell command to find out the number of columns used by a specific VD.

```
Get-VirtualDisk -FriendlyName <vdName> | FL NumberOfColumns
```

The reason for the recommendation is to ensure that you are able to expand VDs that are already very low on usable disk space. For a write operation to a VD to be successful, data is striped across the number of disks indicated by the column count. If you add fewer disks to the storage pool, there may be

new free disk space now in the pool, but you may find you are not able to expand the disk space of the VD, because there will not be enough disks available with free space to allow for a full stripe to be written.

For example, for a 2x3 configuration with three DSMS 1400 storage enclosures, each with eight HDDs and four SSDs, a quantity of 24 HDDs and 12 SSDs are included in a storage pool called MyPool1. The pool has one VD that was created by using storage tiers—2wayVD1, with two-way mirroring and a column count of five. The plan is to add an extra DSMS 1400 enclosure with eight new HDDs and four new SSDs.

For this example, MyPool1 is full with data and the HDD tier, to where no usable disk space remains in 2wayVD1. Eight new HDDs are added to MyPool1. However, 2wayVD1 has a column count of five, which means a total of 10 disks are required for a full stripe, one stripe across five disks for first data copy and one stripe across another five disks for the second data copy. Because only eight HDDs were added to the pool after the original disks in the pool are out of disk space, the HDD tier cannot be expanded to take advantage of the newly added disk.

There is one more factor to consider when adding physical disks or storage enclosures with the intent of expanding existing VDs, which were created with enclosure awareness. Enclosure awareness spreads the data copies of each VD across three or more storage enclosures. However, when adding new physical disks or storage enclosures, VDs created on the newly added storage disk space may not be enclosure aware. This occurs if there is no sufficient free disk space or there are no enough physical disks in existing storage enclosures to spread the new data copies in a method, which satisfies the enclosure awareness requirements.

Run the following PowerShell command for adding a new physical disks to an existing pool.

```
Add-PhysicalDisk -StoragePoolFriendlyName <poolName>  
-PhysicalDisks<physicalDiskObject> -Usage AutoSelect
```

Run the following PowerShell command for extending a VD with storage tiers.

```
Resize-StorageTier -FriendlyName <vdName> -Size <newVDSIZE>
```

Run the following PowerShell command for extending a volume.

```
Resize-Partition -DiskNumber <diskNumber> -Size <newVolumeSize>
```

For example, a physical disk, which is labeled PhysicalDisk13, is assigned to variable \$pd. The disk is then added to an existing pool MyPool1 with the Usage attribute set to AutoSelect. A VD named

2wayVD1, is 30 GB and currently exists in the pool. In this example, the VD is extended to 60 GB, and the volume must be extended to match the new VD size.

```
$pd = Get-PhysicalDisk -CanPool $true -FriendlyName PhysicalDisk13
Add-PhysicalDisk -StoragePoolFriendlyName "MyPool1" -PhysicalDisks $pd
-UsageAutoSelect
Resize-StorageTiers -FriendlyName 2wayVD1 -Size 60GB
$vd = Get-VirtualDisk -FriendlyName 2wayVD1
$diskNum = Get-Disk -VirtualDisk $vd
$partNum = Get-Partition -DiskNumber $diskNum.Number
$size = Get-PartitionSupportedSize -DiskNumber $diskNum.Number
Resize-Partition -DiskNumber $diskNum.Number -PartitionNumber
$partNum.PartionNumber -Size $size.SizeMax
```

Cluster best practices

Cluster shared volumes

Dell recommends that no CSVs be greater than 10 TB, to ensure efficient failover from one node to another.

You must have one cluster-shared volume (CSV) for every VD. You should add CSVs in multiples of the cluster nodes to enable the CSVs to load balance across the cluster. For example, if you have a 2x2 configuration, which has two storage nodes, you should create CSVs in multiples of two (for example, 2, 4, 6, and 8 CSVs). If you have a 4x4 configuration, which has four storage nodes, create CSVs in multiples of 4 (for example 4, 8, 12, and 16 CSVs).

To support the addition of cluster nodes in the future, Dell recommends that you have the quantity of CSV to be at least twice the number of cluster nodes. For example, a 2x2 configuration should have four CSVs.

Cluster shared volumes cache

CSV cache allows you to allocate disk space on the storage node for a write-through cache. If you are using Microsoft Storage Spaces with storage tiers, the CSV cache is not used and is not configured.

For a Scale-Out File Server configuration, if you have VDs that are not using storage tiers or you are using VDs with parity resiliency then Dell recommends that you enable CSV cache. For optimal performance, the recommendation is to allocate 50 percent of the storage node memory for the CSV cache.

Run the following PowerShell command to set the CSV cache disk space.

```
(Get-Cluster).BlockCacheSize = <sizeInMB>
```

For example, using a 4x4 configuration with four DSMS 630 as the storage nodes. Each storage node has 128 GB of RAM. CSV cache should be set to 64 GB.

```
(Get-Cluster).BlockCacheSize = 65536
```

Cluster quorum

A witness disk is used when there is a node or network communication failure where the node continue to operate but can no longer communicate between one another. The witness disk acts as part of the voting process to determine which nodes remains as part of the cluster to maintain quorum. Use witness disks on clusters with even number of storage or compute nodes to ensure that there is a vote majority in case of a split-brain situation.

There are two options for cluster quorum witness, a disk witness and file share witness. For SOFS configurations, Dell recommends that you use file share witness because a continuously available SMB share can be created on the SOFS cluster to host the witness disk. You can create the file-share witness disk on a small 3 GB, two-way mirror VD without storage tiers.

For Converged configurations, Dell recommends that the quorum disk be located on its own CSV within any storage pool that is accessible to all nodes.

There are different types of quorum modes:

- Node majority
 - Recommended for clusters with an odd number of cluster nodes (for example, 3x3)
 - All nodes that are online and can communicate, have a vote and the cluster is operational when votes are over 50 percent.
- Node and disk majority
 - Each node and the witness disk, which is part of cluster, has a vote.
- Node and file share majority
 - Recommended for clusters with an even number of cluster nodes (for example, 4x4).
 - Each node and the file share witness has a vote.

Run the following PowerShell command to create the cluster quorum disk.

```
Set-ClusterQuorum <Mode> <PathToResource>
```

For example, a 3x3 configuration with three DSMS 730 Storage nodes. Setting quorum mode to node majority with no file share or disk witness on cluster named MYCLUSTER.

```
Set-ClusterQuorum -NodeMajority -Cluster MYCLUSTER
```

For example, a 4x4 configuration with four DSMS 630 storage nodes. Set quorum mode to node and file share majority with a 1 GB VD setup as a SOFS file share. The cluster is named MYCLUSTER and file share is named FileShareWitness.

```
Set-ClusterQuorum -NodeAndFileShareMajority "\\MYCLUSTER\FileShareWitness"
```

Scale-Out File Server best practices

Scale-Out File Share usage

Scale-Out File Server (SOFS) shares are intended for storage of Hyper-V virtual hard disks and SQL databases. Using SOFS for information worker workloads including shares for end user file shares, containing unstructured data such as home folders, spreadsheets, or PDFs requires some additional considerations. Information worker workloads require metadata changes (file open, close, rename, delete) typically from hundreds of users at once. Scale-Out File Server shares are continuously available, which requires all nodes to synchronize metadata changes, which in the case of information worker workloads result in a potential performance overhead for these shares because of the large number of metadata changes. Based on the information worker workload, the performance impact may vary and in some cases be negligible, such as information worker workloads with newer versions of Microsoft Office. Finally, many features, which are available on a general use file shares may not work on scale-out file shares – such as DFS-R and quotas.

There are three options for providing file shares to information workers on SOFS:

Option 1: Create a SOFS on the SOFS cluster and host the information worker workload directly on the share. As long as the clients accessing the shares are using Windows 8 or later, they will receive all the benefits of using a SOFS – single namespace, load balancing, and more. Workloads on the share may experience some performance overhead because of metadata changes. Also, the scale-out file shares may not support all the features a general use file server provides like DFS-R and quotas.

Option 2: Create a new VM running Windows Server 2012 R2 and store it on a SOFS file share. Install the File Server role and configure it as a File Server for general use. Allocate the disk space (size) of the VHDX file on the basis of worker data you plan to save. Create all information worker file shares inside the VM. After the information worker workloads are run inside of a VHDX, there will be no performance overhead because of SOFS metadata changes. Also, general use file shares have the full capabilities such as DFS-R and quotas. However, because clients access general use file shares they will not receive all the benefits of accessing a SOFS. When the VM running the general-use file server is highly available on the SOFS cluster, the general use file shares running inside the VM will not be. More tasks such as clustering the guest OS may be require to provide continuously available file shares to clients.

Option 3: Create a new SMB file share on the SOFS, but modify the share after you create it by clearing the **Enable continuous availability** check box. By disabling the 'continuous availability' feature, the file share will not have performance overhead because of SOFS metadata changes. However, after the file share is not continuously available, then if access to a node in the cluster hosting the share is lost, there may be momentary loss of connectivity to the workload during the failover of the file share. In many cases, information worker workloads, such as Microsoft applications, cache data locally and a brief outage during the failover may be transparent to the user. Third-party applications may not offer the same level of data consistency, and should be evaluated on a case by case basis. Also, as long as the

clients accessing the shares are using Windows 8 or later, they receive all the benefits of using a SOFS—single namespace, load balancing, and more. But, after the information worker workload is being hosted on SOFS it will not support all the features a general use file server provides such as Distributed File System Replication (DFS-R) and quotas.

Creating file shares

SMB scale-out file shares are created on CSVs. You can create any number of shares on the CSVs based on your requirements. However, the underlying CSV size dictates the total usable disk space available for the file shares. When creating the file share in the **Failover Cluster Manager**, click **SMB Share-Applications**.

As part of the file share configuration, you have an option to select the **Encrypt data access** check box. Selecting this check box enables you to use SMB Encryption for the specified file share, if you required. Using SMB encryption has a significant performance overhead.

Deduplication

Deduplication is a method of reducing storage needs by eliminating redundant data. Deduplication is supported on SOFS, however, there are some important caveats including the fact that it is not supported on Hyper-V hosts or hosts running SQL server. The only exception is VDI VHDs.

Great candidates for deduplication:

- Folder redirection servers
- Software deployment shares
- SQL Server and Exchange Server backup volumes
- VDI VHDs

Deduplication must be evaluated based on the following content:

- Line-of-business servers
- Static content providers
- Web servers
- High-performance computing (HPC)

Not good candidates for deduplication:

- Hyper-V hosts
- WSUS
- Servers running SQL Server or Exchange Server
- Files approaching or larger than 1 TB in size

Converged best practices

Converged use cases

A Converged solution runs compute workloads directly on the servers, which are direct SAS-attached to the storage enclosures. Typically, the most common workload is Hyper-V; however, any other workload that is supported by Windows Server 2012 R2 is supported on the Converged solution. But, if you decide to use the solution for Hyper-V, it must be the exclusive workload and no other workloads should run on the configuration to prevent potential resource starvation to the VMs.

Logical processor sizing for Hyper-V virtual machines

Determining the number of available logical processors

The number of logical processors available for allocation to VMs is dependent on the number of logical processors available in the server. To determine the number of available logical processors on the Hyper-V host server, start Windows Task Manager, click **More Details**, and then click the **Performance** tab. The number of logical processors listed in this tab is the maximum number of processors available for allocation.

Calculating logical processor requirements

Some processing capability must be reserved for the Hyper-V host server so that Storage Spaces and the remaining roles and features running on the server perform as expected. For each server, Dell recommends at least two logical processors be left unallocated to VMs so that they remain available for the Hyper-V host server.

Also, Dell recommends that you design the solution such that in the event of a potential compute node failure when the VMs fail over to the remaining compute nodes, that the logical processor resources in those nodes are able to handle the additional VM load. Dell recommends that the number of logical processors allocated to all VMs hosted by the compute cluster not exceed the number of available logical processors for each node multiplied by the total number of nodes minus one, as shown in the following equation:

*Total Number of processors available for VMs = (Logical Processors available Per Node - 2) * (Total # of nodes - 1)*

For example, a three-node Converged cluster, with 32 logical processors on each node and where two logical processors are reserved for each node, supports up to $(32-2)*(3-1) = 60$ logical processors

available for allocation to all VMs. If each VM was configured to use only one logical processor, the solution could support up to 60 total VMs, or 20 VMs per node.

Memory Sizing for Hyper-V virtual machines

Calculating memory requirements

The host OS must have sufficient memory to provide services such as I/O virtualization, VM snapshot, and management to support the child partitions. The host reserves a minimal amount of memory, called root reserve, which cannot be allocated to VMs. In general, this minimal amount is often too low. This lack of memory invites the risk of the VMs starving the host during periods of high activity, resulting in the host OS running poorly and impacting the Microsoft Storage Spaces and Hyper-V management functions.

Dell recommends that you reserve at least 8 GB of memory for the host OS on each compute node, by not allocating this memory to VMs.

Plan your Converged solution so that when VMs failover to the remaining nodes during a compute node failure, sufficient memory is available to manage the additional VM load. Dell recommends that the amount of memory specified for all VMs hosted by the compute cluster not exceed the memory available for each node multiplied by the number of nodes minus one. This recommendation maintains the guaranteed optimal performance levels in the event of a single compute node failure. This is shown in the following equation:

$$\text{Total amount of Memory available for VMs} = (\text{Memory available Per Node} - 8 \text{ GB}) * (\text{Total \# of nodes} - 1)$$

For example, in a two-node Converged cluster with 128 GB of memory on each node, and 8 GB of RAM reserved for the host, the total memory available to be allocated to all VMs in the solution should not exceed $(128 - 8) * (2 - 1) = 120$ GB of memory. If each VM was configured to use 2 GB of RAM the solution could support up to 60 total VMs, or 30 VMs per node.

In another example, a three-node Converged cluster with 256 GB of memory on each node allows for 496 GB of memory that can be allocated to all VMs hosted by the compute cluster. In this example, if one node fails, then the 248 GB on each of the two remaining nodes is sufficient to provide the allocated memory to VMs and still have 8 GB in reserve for each of the node's host OS.

Dynamic Memory

When Dynamic Memory is enabled during VM configuration, Dell recommends that the sum of the maximum RAM setting for all VMs not exceed the physical memory capacity of the compute cluster, as calculated in the previous section.

Enabling Dynamic Memory will intermittently add a small latency for each reallocation of memory and add some overhead in the host OS related to memory management in the virtualized environment.

Following is a summary of VM settings related to Dynamic Memory. These settings can be configured using the Set-VM PowerShell cmdlet or using the Hyper-V Manager and must be configured on the node where the VM resides.

 **NOTE:** You can enable or disable Dynamic Memory only when a VM is in a stopped state.

Startup RAM

The startup RAM value can be set when creating a VM using the New-VM and the Set-VM cmdlets or it can be set by using the Hyper-V manager for the VM, under the hardware settings for memory.

The Startup RAM value specifies the amount of memory required when a VM is powered on. This value can exceed the minimal RAM level when the memory required to start a guest OS is more than the minimum amount needed to run the guest OS.

Minimum RAM

Specifies the minimum amount of memory to allocate to the VM after the VM has started. Valid values range from 32 MB to the default value, which is the Startup RAM.

Maximum RAM

Sets Maximum RAM that can be assigned to the VM. It is important to set this to a reasonable value on all VMs to prevent a single VM, in error, starving the other VMs. When Dynamic Memory is enabled, its default maximum is 1 TB, which is too high. Be aware that this value can be increased anytime, but to reduce the value, the VM must be turned off.

Memory Buffer

Percentage to reserve as a buffer over current demand to ensure quick response to changes in memory usage by the VM to minimize repeated reallocations of memory.

Memory weight

Sets priority for VM for memory resources compared to other VMs in the situation where there is not sufficient available memory to give every VM its requested amount of memory.

Allocating and configuring storage

Considerations in assigning CSV to specific VMs

- While multiple VMs can share the same CSV, Dell recommends spreading the VMs across multiple CSVs for optimal performance.
- Depending on your workload, you may consider using multiple storage pools, to allow different workloads to be placed on different physical disks.
- Dell recommends running different types of workloads on separate CSVs.

Virtual Hard Disks (VHDX)

Dell recommends that fixed VHDX files with default parameters are used for Converged solutions involving Hyper-V VMs, as this results in optimal performance. If you decide to use dynamically expanding VHDXs, ensure that you have available storage for the VD to expand. If expansion happens and there is no available storage the VMs can suddenly pause and create outages. Also, file fragmentation can occur as the VHDX file increases in size over time, which can lead to overall reduced speed when compared to a VHDX that has been fully allocated at creation.

Deduplication

Deduplication is not recommended on Converged solutions because of CPU overhead, which will have an impact on Storage Spaces and workload performance.

Parity Spaces

Parity spaces are supported with Converged configurations, however, there will be a greater CPU overhead compared to Mirrored spaces due to the parity calculations required.

Allocating and configuring network

Virtual switches (Hyper-V)

Use caution when integrating virtual switches with RDMA adapters as virtual switches cause the adapters to not be able to use the RDMA functionality. For more information about a list of the latest validated RDMA adapters for this solution, see the *Dell Storage with Microsoft Storage Spaces Support Matrix* available at Dell.com/dsmsmanuals.

NIC Teaming

When using NIC Teaming, Dell recommends that you only team adapters of the same speed and function together and also that you do not team RDMA adapters, if you want to use the RDMA functionality.

Getting help

Contacting Dell

Dell provides several online and telephone-based support and service options. If you do not have an active internet connection, you can find contact information about your purchase invoice, packing slip, bill, or Dell product catalog. Availability varies by country and product, and some services may not be available in your area. To contact Dell for sales, technical assistance, or customer-service issues:

1. Go to **Dell.com/support**.
2. Select your country from the drop-down menu on the lower right corner of the page.
3. For customized support:
 - a. Enter your system Service Tag in the **Enter your Service Tag** field.
 - b. Click **Submit**.
The support page that lists the various support categories is displayed.
4. For general support:
 - a. Select your product category.
 - b. Select your product segment.
 - c. Select your product.
The support page that lists the various support categories is displayed.
5. For contact details of Dell Global Technical Support:
 - a. Click [Global Technical Support](#).
 - b. The **Contact Technical Support** page is displayed with details to call, chat, or e-mail the Dell Global Technical Support team.

Quick Resource Locator

Use the Quick Resource Locator (QRL) to get immediate access to DSMS Solution information. This information provides access to reference documentation, a link to the Dell support page, and links to hardware-specific QRL pages, which include how-to videos. You can access this information by visiting www.dell.com/QRL or by using your smartphone or tablet and scanning the QR code below.

