Dell EMC Solutions for Microsoft Azure Stack HCI

Scalable Hyperconverged Infrastructure with PowerEdge R440, R740xd, R740xd2, and R640 Storage Spaces Direct Ready Nodes

DELL EMC

Notes, cautions, and warnings

NOTE: A NOTE indicates important information that helps you make better use of your product.

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.

© 2018 2019 Dell Inc. or its subsidiaries. All rights reserved. Dell, EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners.

August 2019

Rev. H17517.6

Contents

1

Introduction

This chapter presents the following topics:

Topics:

Introduction

This Dell EMC Solutions for Azure Stack HCI deployment guide focuses on deploying a scalable hyperconverged infrastructure. The guide includes an overview of the solution infrastructure, guidance on how to integrate the solution components, and instructions for preparing and deploying the solution infrastructure. This guide is applicable only to infrastructure that is built by using the validated and certified Dell EMC Storage Spaces Direct Ready Nodes and Microsoft Windows Server 2016 and Windows Server 2019.

Audience and scope

The audience for this document includes systems engineers, field consultants, partner engineering team members, and customers with a fair amount of knowledge in deploying hyperconverged infrastructures with Microsoft Windows Server 2016 or Windows Server 2019 Hyper-V and Storage Spaces Direct.

Customers who have Volume License agreements with Microsoft can order Dell EMC Microsoft Storage Spaces Direct Ready Nodes with the operating system preinstalled at the factory with an OEM license or bare metal.

The Storage Spaces Direct cluster deployment can be done in either of the following ways:

- Dell EMC Services led: Certified deployment engineers ensure accuracy, speed, reduced risk and downtime.
- Customer led: Customers follow the instructions in this guide, provided that they have the qualified level of technical expertise.

NOTE: Instructions in this deployment guide are applicable only to the generally available OS build of Windows Server 2016 with the latest applicable updates and Windows Server 2019 GA build with latest OS updates. These instructions are not validated with Windows Server version 1709. Storage Spaces Direct Ready Nodes do not support the Windows Server Semi-Annual Channel release. Dell EMC recommends that you update the host OS with the latest cumulative updates from Microsoft before starting the Azure Stack HCI cluster creation and configuration tasks.

Assumptions

This deployment guide makes certain assumptions about the necessary prerequisite knowledge of the deployment personnel. This includes the prerequisite knowledge of:

- Dell EMC Microsoft Storage Spaces Direct Ready Nodes and deploying and configuring BIOS and iDRAC settings
- Dell EMC Networking switches and concepts such as Data Center Bridging (DCB) and Virtual Link Trunking (VLT)
- Deploying and configuring Windows Server 2016 or Windows Server 2019 Hyper-V infrastructure

Known issues

Before starting the cluster deployment, review the known issues and workarounds. For a list of known issues, see https://www.dell.com/support/article/sln313305.

2

Virtualization Infrastructure with Dell EMC Solutions for Azure Stack HCI

This chapter presents the following topics:

Topics:

Overview

Dell EMC Solutions for Azure Stack HCI encompasses various configurations of R740xd, R740xd2, and R640 Storage Spaces Direct Ready Nodes and PowerEdge R440 servers, which power the primary compute cluster that is deployed as a hyperconverged infrastructure (HCI). This HCI uses a flexible solution architecture rather than a fixed component design. The following figure illustrates one of the flexible solution architectures, consisting of a compute cluster alongside the redundant top-of-rack (ToR) switches, a separate out-of-band network, and an existing management infrastructure in the data center.

This solution is available in both hybrid and all-flash configurations. For more information on available configurations, see the Dell EMC Ready Nodes for Microsoft Storage Spaces Direct with Hyper-V Solution Overview.

NOTE: For the two-node cluster deployment, a cluster witness must be configured. For information about available options and other references to deployment instructions, see Configuring a cluster witness.

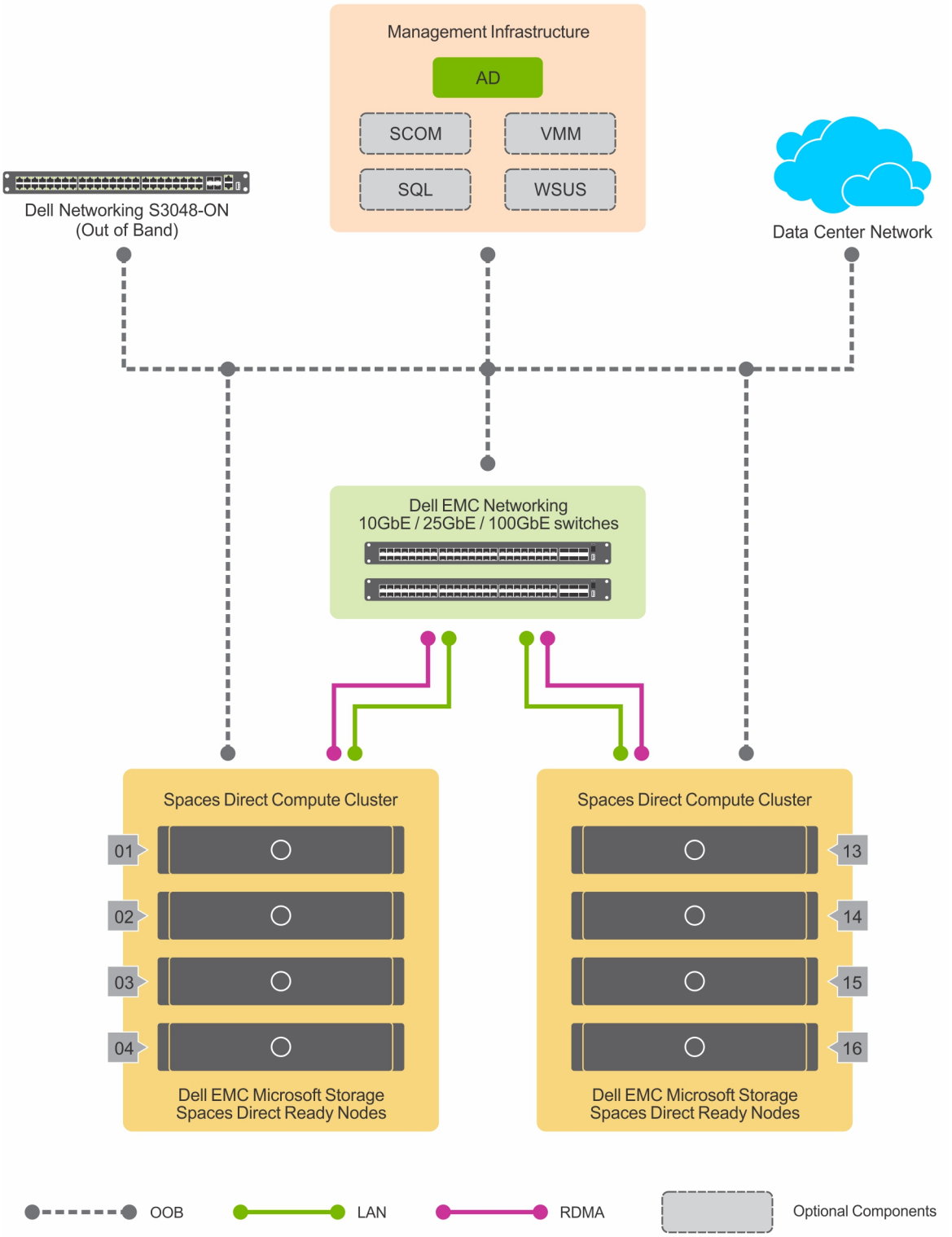

Figure 1. Hyperconverged virtualized solution using Dell EMC Ready Nodes

Dell EMC Solutions for Azure Stack HCI do not include management infrastructure components, such as a cluster for hosting management VMs, services such as Active Directory (AD), Domain Name Service (DNS), and Windows Server Update Services (WSUS), and System Center components such as Virtual Machine Manager (VMM) and Operations Manager (OM). Therefore, the instructions in this guide do not include deployment of these services and components. The instructions in this guide assume that at least an Active Directory domain controller and a DNS server are available in the existing management infrastructure.

The redundant switch configuration that is shown in the preceding figure provides high availability. If you are connecting the nodes to two separate network switches and implementing Switch Embedded Teaming (SET), both switches require access to all subnets so that failover can occur. When you are using RoCE for RDMA, DCB must be configured. DCB provides enhancements to the Ethernet protocol,

which improves the functionality of data center networks. To take advantage of the Mellanox RDMA over Converged Ethernet (RoCE) network adapters, Priority Flow Control (PFC) and Enhanced Transmission Selection (ETS) are required. PFC and ETS are configured on all nodes and all network switches interconnecting the nodes.

This deployment guide provides instructions and PowerShell commands for manually deploying an Azure Stack HCI cluster. For information about configuring host networking and creating an Azure Stack HCI cluster using System Center Virtual Machine Manager (VMM), see https://community.emc.com/docs/DOC-75346.

The subsequent sections provide an overview of the hardware and software components in the virtualized solution based on Dell EMC Ready Nodes.

R740xd Storage Spaces Direct Ready Node

The R740xd Storage Spaces Direct Ready Node, based on the Dell EMC PowerEdge R740xd server, is optimized for software-defined storage implementations that enable converged infrastructure (CI) and hyperconverged infrastructure (HCI) deployments such as the architecture recommended in this deployment guide. With 2 CPU sockets and a wide range of CPU options, this Ready Node provides capabilities to match your computational needs.

This 2U rack mounted server provides high storage density with different drive options for a hybrid configuration. The following table lists the configuration options that are available.

Table 1. System configuration: R740xd Storage Spaces Direct Ready Node

Configuration	Chassis	Drive layout
Hybrid	R740xd Storage Spaces Direct Ready Node (18 drives), 3.5 in. drive form factor	• Up to 6 SSDs in the front bay • Up to 12 HDDs (6 in the front bay, 4 in the internal bay, and 2 in the rear flex bay)
	R740xd Storage Spaces Direct Ready Node (12 drives), 3.5 in. drive form factor	• 2–4 SSDs in the front bay (cache) • 4–8 HDDs in the front bay (capacity)
All-flash	R740xd Storage Spaces Direct Ready Node (24-drive), 2.5 in. drive form factor	4–24 mixed-use or read-intensive SSDs
All-NVMe	R740xd Storage Spaces Direct Ready Node (24 drives with 12 NVMe capable)	4-12 NVMe SSDs
NVMe add-in card (AIC) + HDD	R740xd Storage Spaces Direct Ready Node (12 drives), 3.5 in. drive form factor, with 2 NVMe AICs	• 2 NVMe AICs in PCIe slot 1, 7 (cache) • 12 HDDs in the front bay (capacity)
NVMe small form factor (SFF) + HDD	R740xd Storage Spaces Direct Ready Node (24 drives), 2.5 in. drive form factor, with up to 4 NVMe SFF drives	• 2 or 4 x 2.5 in. NVMe SFF drives in the front bay (cache) • Up to 20 x 2.5 in. HDDs in the front bay (capacity)
NVMe SFF + SSD	R740xd Storage Spaces Direct Ready Node (24 drives), 2.5 in. drive form factor, with up to 4 NVMe SFF drives	• 2 or 4 x 2.5 in. NVMe SFF drives in the front bay (cache) • Up to 20 x 2.5 in. SSDs in the front bay (capacity)

NOTE: The NVMe AICs are not preinstalled in the Ready Node. These cards must be installed before proceeding to the component integration section. See the slot numbers that are noted in the preceding table for slot priority of the NVMe AICs.

The following table provides an overview of the 740xd Storage Spaces Direct Ready Nodes in the solution.

Table 2. System components: R740xd Storage Spaces Direct Ready Node

Component	Specifications

NICs	• Add-in adapter card options:
	  • 1, 2, or 4 QLogic FastLinQ 41262 Dual Port 25 GbE SFP28
	  • Up to 2 Mellanox ConnectX-4 Lx Dual Port 25 GbE SFP28
	  • Up to 2 Mellanox ConnectX-5 Ex Dual Port 100 GbE QSFP28
	• rNDC options:
	  • Dual Port Intel X710 (10 GbE) + Dual Port Intel i350 (1 GbE)
	  • Dual Port Broadcom 57412 (10 GbE) + Dual Port Broadcom 5720 (1 GbE)
	  • Dual Port Qlogic 57800 (10 GbE) + Dual Port Qlogic (1 GbE)
	  • QLogic FastLinQ 41162 Dual Port 10GbE BASE-T & Dual Port 1 GbE BASE-T

Storage adapter	HBA330 Internal SAS HBA
Boot device	BOSS S.1 with 2 x BOSS M.2 devices in RAID 1
Drives	See Dell EMC Ready Nodes for Microsoft Storage Spaces Direct with Hyper-V Solution Overview
LAN switches	2 x 10 GbE, 25 GbE, or 100 GbE Dell EMC Networking switches
OOB switch	1 x S3048

NOTE: Mellanox ConnectX-5 EX 100 GbE NIC is an RDMA card option for select configurations only.

R740xd2 Storage Spaces Direct Ready Node

The R740xd2 Storage Spaces Direct Ready Node provides dense storage capacity configurations. The configurations are optimized for software-defined storage implementations that enable converged infrastructure (CI) and hyperconverged infrastructure (HCI) deployments such as the architecture recommended in this deployment guidance. The server offers a wide range of drive capacity configurations to enable up to 364 TB (using 24 x 3.5 in. SAS/SATA HDDs) of raw storage capacity. With two processor sockets and a wide range of processor options, the R740xd2 Storage Spaces Direct Ready Node offers the capabilities that match not just the storage needs but also the computing needs.

This 2U rack-mounted server is available in hybrid drive configurations. The following table lists the available configuration options.

Table 3. System configuration: R740xd2 Storage Spaces Direct Ready Node

Item	Configuration
Chassis	12 x 3.5 in. front bay 1 + 12 x 3.5 in. front bay 2 with butterfly riser Config 4
Storage controller	HBA330 Internal SAS HBA
Storage – OS Boot	BOSS S.1 with 2 BOSS M.2 devices in RAID 1
Storage – SSD (cache)	Up to 8 SAS (WI/MU)/SATA MU SSDs (Front Bay 2 only)
Storage – HDD (capacity)	Up to 20 x 3.5 in. NL-SAS/SATA HDD (Front Bay 1 and 2)
Network cards	• Add-in: 1 or 2 Mellanox ConnectX-4 Lx Dual Port 25 GbE SFP28 or 1 or 2 QLogic FastLinQ 41262 Dual Port 25 GbE SFP28 • On-board: Broadcom 5720 Dual Port 1 GbE • Mezz: Broadcom 57416 Dual Port 10 GbE SFP+
iDRAC	iDRAC 9

R640 Storage Spaces Direct Ready Node

R640 Storage Spaces Direct Ready Node, a 1U rack server based on PowerEdge R640 server, is optimized for software-defined storage implementations that enable converged infrastructure (CI) and hyperconverged infrastructure (HCI) implementations. This Ready Node

supports up to 2 CPU sockets with a wide range of options in number of cores per CPU socket and 1.5 TB of memory when using DDR4 DIMMs.

The R640 Storage Spaces Direct Ready Node is available in different chassis configurations that offer several internal storage choices.

The following table lists the system configuration options for this Ready Node.

Table 4. System configuration: R640 Storage Spaces Direct Ready Node

Component	Chassis	Drive layout
Hybrid	12 drives, 2.5 in. drive form factor	• 2–4 SSDs (cache) • 4–10 HDDs (capacity)
	10 drives, 2.5 in. drive form factor	• 2–4 x SSDs (cache) • 4–8 x HDDs (capacity)
All-Flash	12 drives, 2.5 in. drive form factor	12 SSDs
	10 drives, 2.5 in. drive form factor	10 SSDs
All-NVMe	10 drives, 2.5 in. drive form factor	4–10 NVMe
NVMe + SSD	10 drives, 2.5 in. drive form factor	• 2 NVMe (cache) • 4–8 SSDs (capacity)
Intel Optane + SSD	10 drives, 2.5 in. drive form factor	• 2 Intel Optane (cache) • 4–8 SSDs (capacity)
NVMe + HDD	10 drives, 2.5 in. drive form factor	• 2 NVMe (cache) • 4–8 HDDs (capacity)

NOTE: The 12-drive chassis configuration in R640 Storage Spaces Direct Ready Node is available only with one PCIe slot and that is used by the BOSS M.2 device for OS RAID. In this chassis configuration, network connectivity is provided by using a Mellanox ConnectX 4 LX or QLogic FastLinQ 41262 rNDC.

The following table provides an overview of the R640 Storage Spaces Ready Node in the solution.

Table 5. System components: R640 Storage Spaces Direct Ready Node

Component	Chassis type	Specifications
NIC	12-drive chassis, Hybrid and All-Flash	Mellanox ConnectX 4 LX or QLogic FastLinQ 41262 rNDC

	10-drive chassis, Hybrid and All Flash	• Add-in adapter card options:
		  • Up to 2 Mellanox ConnectX-4 Lx Dual Port 25 GbE SFP28
		  • Up to 2 QLogic FastLinQ 41262 Dual Port 25 GbE SFP28
		  • Mellanox ConnectX-5 Ex Dual Port 100 GbE QSFP2
		• rNDC card options:
		  • Dual Port Intel X710 (10 GbE) + Dual Port Intel i350 (1 GbE)
		  • Dual Port Broadcom 57412 (10 GbE) + Dual Port Broadcom 5720 (1 GbE)
		  • Dual Port Qlogic 57800 (10 GbE) + Dual Port Qlogic (1 GbE)
		  • QLogic FastLinQ 41162 Dual Port 10 GbE BASE-T + Dual Port 1 GbE BASE-T

	10-drive chassis, NVMe and Intel Optane	• Add-in adapter card options:
		  • Mellanox Connectx-4 LX 25 GbE SFP
		  • QLogic FastLinQ 41262 25 GbE SFP28
		  • Mellanox Connect EX5 (Dual Port 100 GbE)
		• rNDC card options:
		  • Dual Port Intel X710 (10 GbE) + Dual Port Intel i350 (1 GbE)
		  • Dual Port Broadcom 57412 (10 GbE) + Dual Port Broadcom 5720 (1 GbE)
		  • Dual Port Qlogic 57800 (10 GbE) + Dual Port Qlogic (1 GbE)
		  • QLogic FastLinQ 41162 Dual Port 10 GbE BASE-T + Dual Port 1 GbE BASE-T

Storage adapter		HBA330 Internal SAS HBA
Boot device		BOSS S.1 with 2 BOSS M.2 devices in RAID 1
Drives		See Dell EMC Ready Nodes for Microsoft Storage Spaces Direct with Hyper-V Solution Overview
LAN switch		2 x 10 GbE, 25 GbE, or 100 GbE Dell EMC Networking switches
OOB switch		1 x S3048

PowerEdge R440

Dell EMC PowerEdge R440 is a 1U, 2-socket rack server that supports up to 1 TB of physical memory. The PowerEdge R440 server with the 10-drive chassis configuration offers multiple hybrid and all-flash storage options for Microsoft Storage Spaces Direct based hyperconverged infrastructure.

The following table lists the system configuration options for the PowerEdge R440.

Table 6. System configuration: PowerEdge R440

Component	Chassis	Drive layout
Hybrid	PowerEdge R440 (10 drives), 2.5 in. form factor	• 2 SSDs (cache) • 4–8 HDDs (capacity)

Component	Chassis	Drive layout
All-Flash	PowerEdge R440 (10 drives), 2.5 in. form factor	4-10 x SSD (capacity)

The following table provides an overview of the PowerEdge R440 components in the solution.

Table 7. System components: PowerEdge R440

Component	Specifications
NIC	• On-Board Broadcom 5720 Dual Port 1 GB LOM and Broadcom 57416 Dual Port 10 GbE SFP+ Network LOM Mezz Card • QLogic FastLinQ 41262 Dual Port 25 GbE SFP28
Storage adapter	HBA330 Internal SAS HBA
Boot device	BOSS S.1 with 2 x BOSS M.2 devices in RAID 1
Drives	See Dell EMC Ready Nodes for Microsoft Storage Spaces Direct with Hyper-V Solution Overview
LAN switch	2 x 10 GbE or 25 GbE Dell EMC Networking switches
OOB switch	1 x S3048

NOTE: PowerEdge R440 as a cluster node in Storage Spaces Direct infrastructure is validated with QLogic 41262 DP 25 GbE adapter only for converged or nonconverged traffic.

3

Integrating Solution Components

This chapter presents the following topics:

Topics:

Overview

This chapter provides recommendations for server and network switch placement in the racks and port mapping on the top-of-rack (ToR) and out-of-band (OOB) switches. It also provides details about configuring the ToR and OOB switches.

The following figure shows the rack elevation diagrams for the 2-node and 16-node hyperconverged infrastructure (HCI) configurations.

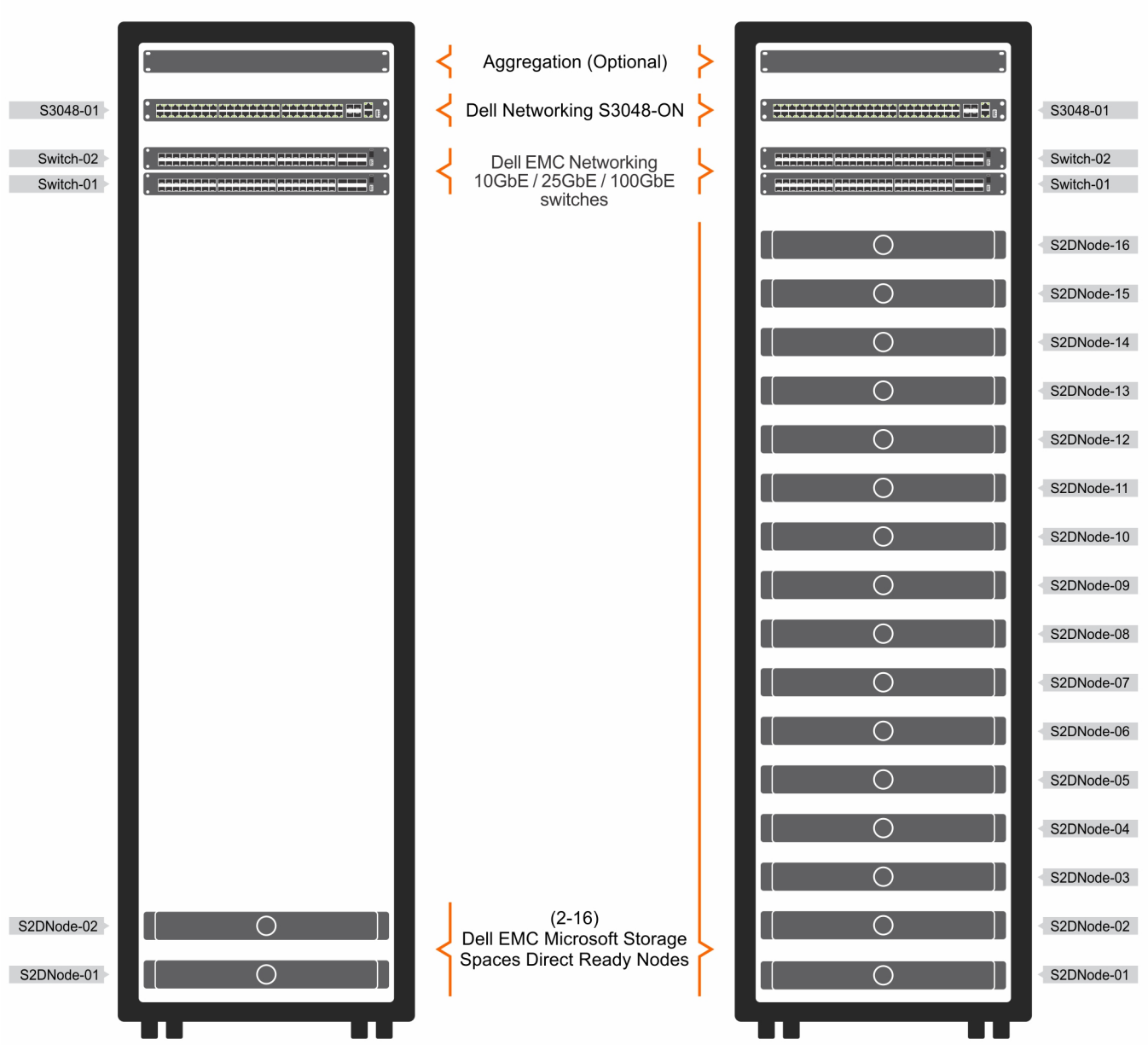

Figure 2. Rack elevation for the 2-node and 16-node HCI configurations

NOTE: Dell EMC does not support expansion to a larger cluster size from a 2-node cluster. A 3-node cluster provides fault-tolerance only for simultaneous failure of a single node and a single drive. If the deployment requires future expansion and better fault-tolerance, consider starting with a 4-node cluster at a minimum.

NOTE: The 2-node configuration built on a Storage Spaces Direct Ready Node has been validated with back-to-back connections for storage connectivity. For more information about the back-to-back connected deployment model, see https://www.dell.com/azurestackhcimanuals.

Network connectivity in Dell EMC Solutions for Azure Stack HCI

There are two physical network topologies used with the Azure Stack HCI Solution—fully converged and nonconverged. Both topologies provide:

- Two top-of-rack (ToR) switches—10 GbE/40 GbE, 25 GbE/100 GbE, or 100 GbE

- A single 1 GbE out-of-band (OOB) management switch

Fully converged network

In this configuration option, each Storage Spaces Direct Ready Node (with the exception of R640 Storage Spaces Direct Ready Node with 12-drive chassis and NVMe configuration, which supports only one network adapter) uses one of the following network adapter options for both host management and storage traffic:

- Up to two Mellanox ConnectX-4 Lx Dual Port 25 GbE SFP28
- Up to two Mellanox ConnectX-5 Ex Dual Port 100 GbE QSFP28
- One, two, or four QLogic FastLinQ 41262 Dual Port 25 GbE SFP28

Each port from each network adapter on each server is connected to a different switch in the redundant ToR network fabric.

The following figure illustrates the connection.

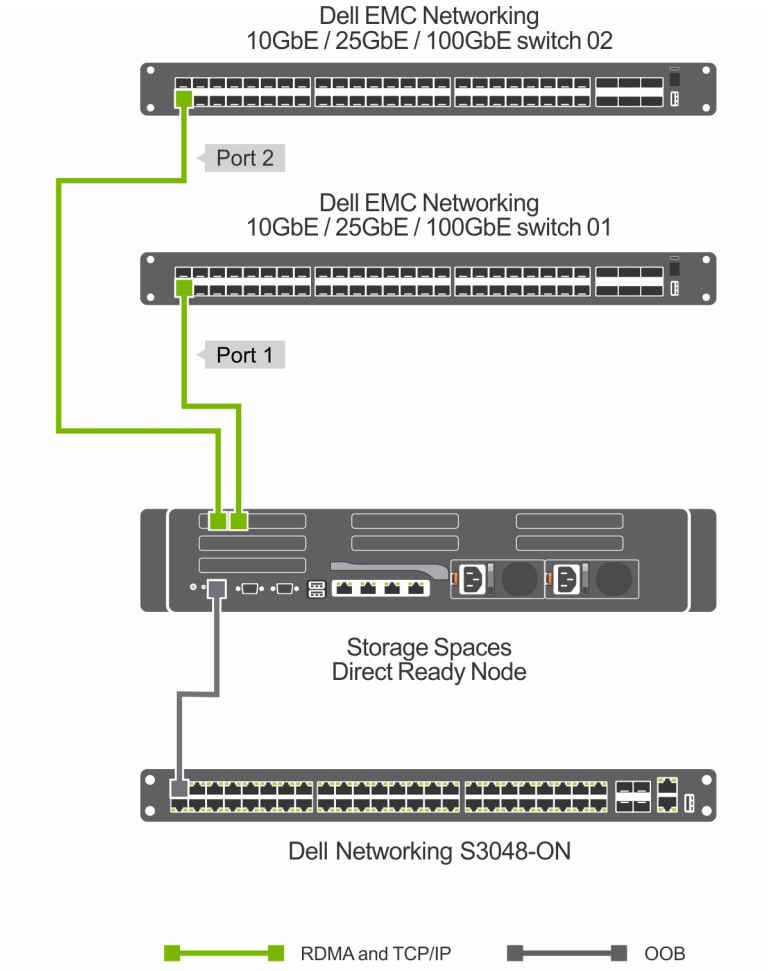

Figure 3. Network and iDRAC connection in a fully converged network configuration

Nonconverged network

In this second network connectivity option, each Storage Spaces Direct Ready Node (with the exception of R640 Storage Spaces Direct Ready Node with 12-drive chassis and NVMe configuration, which supports only one network adapter) uses the following network adapters:

- For storage traffic—One of the following options:
  - Up to two Mellanox ConnectX-4 Lx Dual Port 25 GbE SFP28
  - Up to two Mellanox ConnectX-5 Ex Dual Port 100 GbE QSFP28
  - One, two, or four QLogic FastLinQ 41262 Dual Port 25 GbE SFP28
- For host management and VM traffic—A 10 GbE SFP+ or 1 GbE BASE-T port on one of the following adapters:
  - On-board Intel Ethernet 10G 4P X710/I350
  - Broadcom 57412 DP 10 Gb SFP+ + 5720 DP 1 Gb

- QLogic 57800 DP 10 Gb SFP+ DP 1 Gb BASE-T rNDC

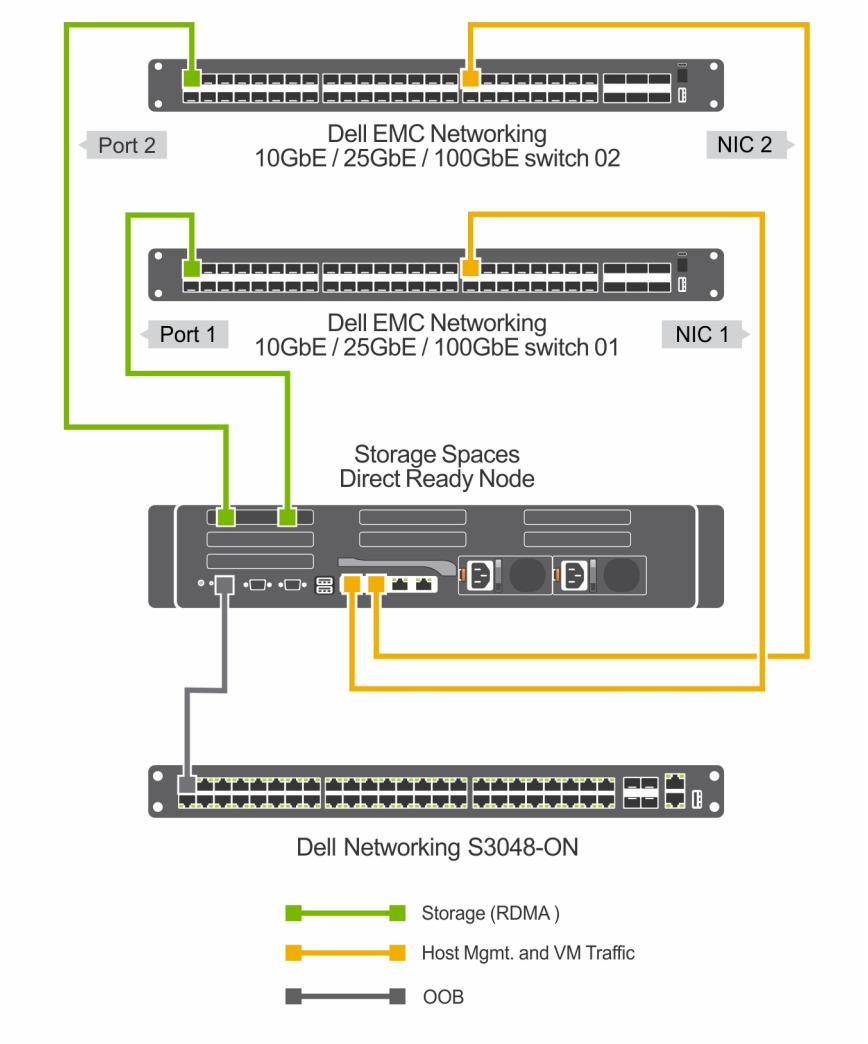

Figure 4. Network and iDRAC connection in a nonconverged network configuration

Networks that use Mellanox ConnectX-4 LX, Mellanox ConnectX-5 EX, or QLogic FastLinQ 41262 network adapters in the server provide Remote Direct Memory Access (RDMA) for storage traffic. RDMA enables significantly increased throughput and lower latency by performing direct memory transfers between servers. Storage Spaces Direct uses SMB for all intra-node communication and uses RDMA with SMB Direct to enhance performance of the overall infrastructure. RDMA with RoCE configuration on Mellanox ConnectX-4 LX and Mellanox ConnectX-5 EX network adapters requires configuration of Data Center Bridging (DCB) and Quality of Service (QoS) in the host operating system as well as the ToR network fabric switch in a fully converged topology.

For all-NVMe configurations that implement iWARP for RDMA using the QLogic FastLinQ 41262 network adapters, implementing DCB in the host OS and ToR network switches is recommended in the fully converged network topology.

The redundant switch configuration shown in Figure 1 provides high availability (HA). If you are connecting the nodes to two separate network switches and implementing Switch Embedded Teaming (SET), both switches require access to all subnets so that failover can occur. When using RoCE for RDMA, DCB must be configured as it provides enhancements to the Ethernet protocol, which improves the functionality of data center networks. To take advantage of the Mellanox RDMA over Converged Ethernet (RoCE) network adapters, Priority Flow Control (PFC), and Enhanced Transmission Selection (ETS) are required. PFC and ETS are configured on all nodes and all network switches interconnecting the nodes.

When deploying the fully converged topology, the Node switch ports must be in Trunk Mode and configured for all Management and Storage Networks VLANs throughout both VLT switches.

In the nonconverged topology, the management network can be connected to the same ToR switches as the storage traffic or it can be connected to a completely different network fabric.

In the dual-NIC configuration with either Mellanox or QLogic adapters, for increased throughput, all four network ports can be connected for storage traffic. This might be necessary for all-flash or all-NVMe configurations.

Each Storage Spaces Direct Ready Node has a dedicated network connection from the Integrated Dell Remote Access Controller (iDRAC) to the Dell Networking S3048-ON switch configured for OOB management.

NOTE: The R640 Storage Spaces Direct Ready Node in a 12-drive chassis configuration can be integrated using the fully converged network topology only.

When using the nonconverged network topology, you can configure the host network in different ways. For detailed information about the host network configuration options, see https://community.emc.com/docs/DOC-73779. For Storage Spaces Direct, Dell EMC recommends a nonconverged network topology with storage traffic on the physical adapters and the host management/virtual traffic implemented as virtual adapters connected to a switch embedded team created using the onboard 10 GbE or 1 GbE ports.

NOTE: In the nonconverged network topology, it is not necessary to configure DCB/PFC in the ToR switches or QoS in the host OS.

NOTE: Both the topologies support tagged and untagged VLANs for storage traffic.

NOTE: For a list of all supported network topology configurations, see https://community.emc.com/docs/DOC-73779.

NOTE: For sample configurations of the ToR switches with DCB and other configurations, see https://community.emc.com/docs/DOC-70310

NOTE: For a list of supported switches for Dell EMC Solutions for Azure Stack HCI, see the support matrix at https://www.dell.com/azurestackhcimanuals.

Port mapping

This section provides recommendations to switch port mapping on a Storage Spaces Direct Ready Node server. The illustrations in the section show how the server NIC ports for both storage and out-of-band (OOB) management traffic are connected to the top-of-rack (ToR) switches.

To make network configuration consistent and easy to understand, follow a consistent mapping across the ToR and OOB switches. The following figures illustrate this mapping.

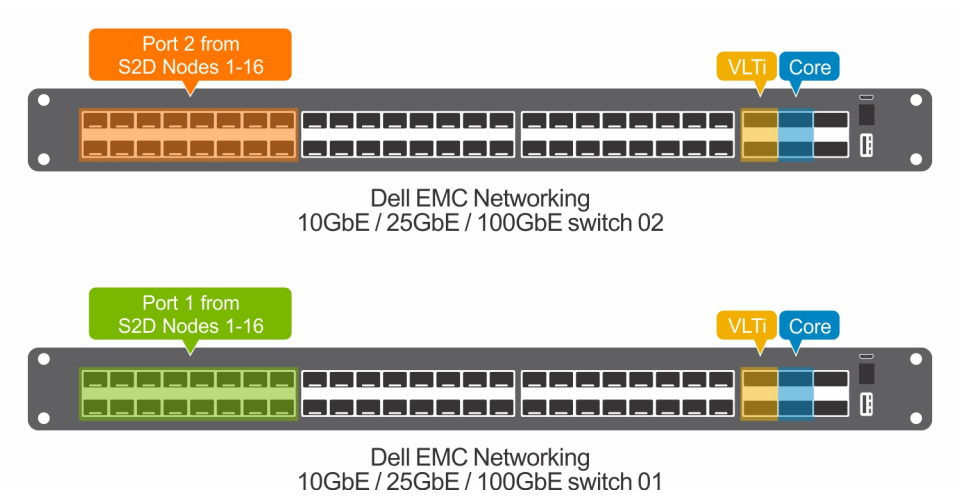

Figure 5. Fully converged port mapping for Ready Nodes

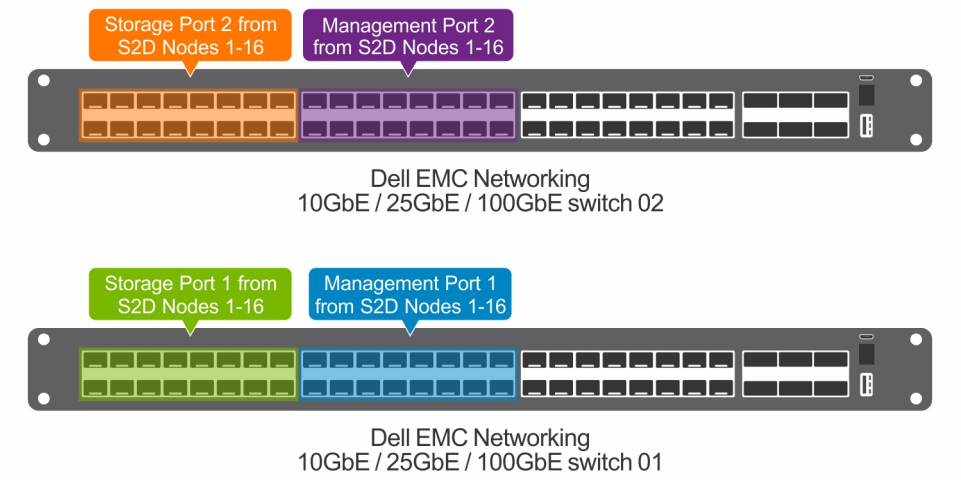

Figure 6. Nonconverged ToR port mapping for Ready Nodes

As shown in the figure, ports from the Mellanox and QLogic adapters from each server are mapped to ports 1-16 on TOR1 and TOR2 respectively. The 40 GbE ports, port 49 and 50 from each ToR, are connected in a Virtual Link Trunking Interconnect (VLT) for inter-switch connectivity.

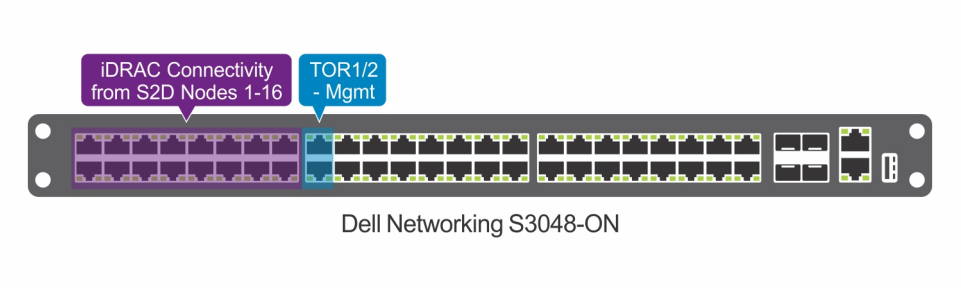

Figure 7. OOB port mapping for Ready Nodes

In another optional deployment scenario for the OOB and ToR network fabric, the OOB switch can be uplinked to the ToR switches by using the 10 GbE ports. Doing so enables connectivity to the OOB interfaces of the cluster components through the ToR switch connections. The following figure illustrates this.

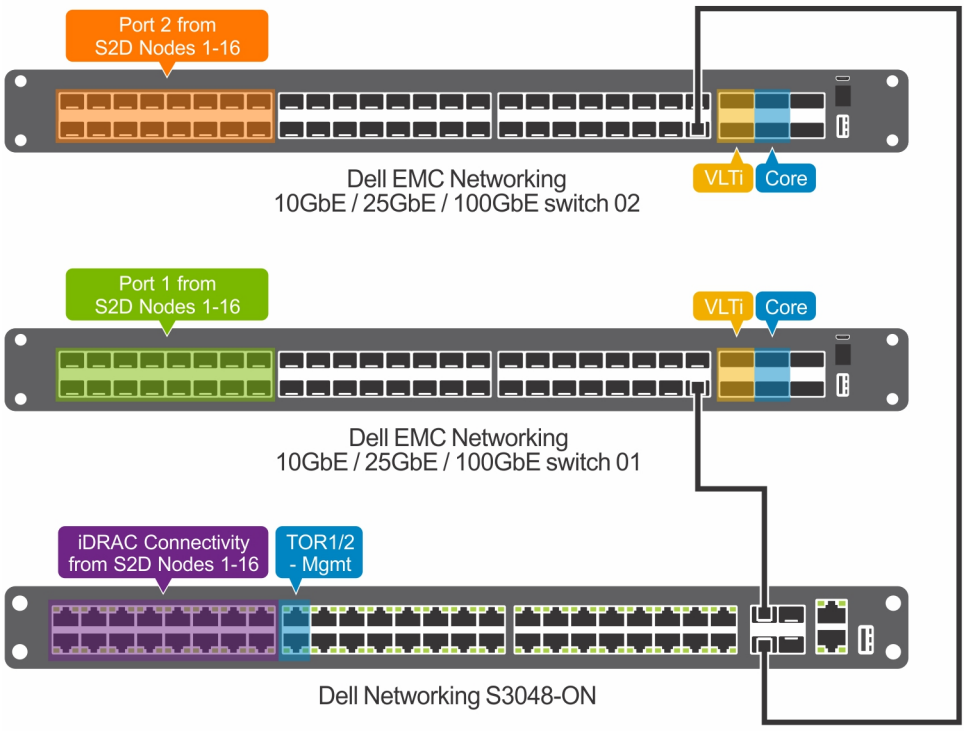

Figure 8. Fully converged network connectivity with Dell EMC Networking S3048 (OOB switch) connecting to Dell EMC Networking S4048 switches

4

Virtualized Infrastructure Deployment Using Dell EMC Ready Nodes

This chapter presents the following topics:

Topics:

Introduction

The Dell EMC Microsoft Storage Spaces Direct Ready Nodes that are described in this guide can be deployed in either of the following ways:

- Manual OS deployment—A manual method of installation starting from OS deployment to cluster creation
- Factory OS deployment—Factory preinstallation of Wiindows Server 2016 or Windows Server 2019 on R740xd, R740xd2, and R640 Storage Spaces Direct Ready Nodes and PowerEdge R440 servers

Each method has deployment prerequisites and required predeployment configuration, including the network switch configuration.

The subsequent sections of this guide describe the deployment prerequisites for each method and provide details about the supported software and firmware versions.

NOTE: Instructions in this deployment guide are applicable only to the generally available OS build of Windows Server 2016 with the latest applicable updates and the Windows Server 2019 GA build. These instructions are not validated with Windows Server version 1709. Storage Spaces Direct Ready Nodes do not support the Windows Server Semi-Annual Channel release. Dell EMC recommends that you update the host OS with the latest cumulative updates from Microsoft before starting the cluster creation and configuration tasks.

NOTE: Each task in this deployment guide requires running one or more PowerShell commands. Dell EMC recommends using these commands to complete the deployment tasks instead of using the UI because there might be scenarios where the UI does not work as expected. For example, the cluster validation UI wizard within the Failover Cluster Manager fails intermittently due to a known issue in the Microsoft code.

Deployment prerequisites

This hyperconverged virtualized solution based on Dell EMC Ready Nodes deployment assumes that the management services required for the OS deployment and cluster configuration are present in the existing infrastructure where the Storage Spaces Direct cluster deployment is being done.

The following table describes different management services, their purpose, and applicability in each of the deployment methods.

Table 8. Management services

Management service	Purpose	Deployment required/optional
Active Directory	User authentication	Required
Domain Name System	Name resolution	Required
Windows Software Update Service (WSUS)	Local source for Windows updates	Optional

Management service	Purpose	Deployment required/optional
MS SQL Server	Database back end for VMM and SCOM	Optional

Software versions

The following table lists the software versions required for the Dell EMC Ready Nodes deployment. As with the management services, the required software depends on the deployment method.

Table 9. Software versions

Component	Version
Operating system	Windows Server 2016 Datacenter or Windows Server 2019 Datacenter
Active Directory forest/domain functional level	Windows Server 2008 R2 or later

Dell EMC validated firmware matrix

This Dell EMC Ready Nodes for Storage Spaces Direct solution is validated and certified with certain firmware versions related to the components in the solution infrastructure. This matrix identifies the Dell EMC validated versions of software and firmware and should be followed for ensuring that the solution infrastructure remains supported and delivers optimal performance.

This support matrix is available at https://www.dell.com/azurestackhcimanuals and gets updated when new revisions of the software and firmware are validated.

Deployment checklist

This section provides a checklist for configuration settings that need to be applied during predeployment configuration and deployment of operating system and other software configuration post OS deployment. For example, network switch configuration requires VLAN ID configuration and IP address space used with each VLAN.

Fill in the checklists provided in the subsequent sections before proceeding to the predeployment configuration.

The section Sample deployment checklists provides completed examples of these checklists for reference.

Management environment checklist

This Dell EMC Ready Nodes deployment is a brownfield deployment and therefore requires information such as active directory domain FQDN, DNS server addresses, and so on.

The following table captures the necessary inputs as a checklist.

Table 10. Management environment checklist

Item	Value
AD domain FQDN	
Domain administrator or equivalent credentials	
DNS server addresses	
WSUS server FQDN (optional)	

Network configuration checklist

Before starting the deployment, identify the IP scope and VLAN information for various traffic classes in the solution infrastructure. The Minimum IP addresses needed column in the following table can be used to identify the correct scope. The value shown in this column is based on the number of components that require the specified traffic class used in this solution. Ensure that the IP scope selected for the traffic class meets the minimum IP addresses requirement.

The IP address requirements and VLAN ID information provided in the following table are only examples. Choose values based on your existing data center architecture.

Consult with the customer network engineering team for VLAN ID and VLAN IP addresses applicable to your solution.

Table 11. Network configuration

Traffic class	Purpose	Minimum number of IP addresses needed	VLAN ID	Tagged/ untagged	Recommended subnet mask	VLAN IP addresses
Out of band	Required for OOB management of server nodes and ToR switches	19		Untagged	/24	
Host Management	Management of cluster and cluster nodes	17		Tagged/ untagged	/25	TOR1: TOR2:
Storage 1	SMB traffic	16		Tagged/ untagged	/27	TOR1: TOR2:
Storage 2	SMB traffic	16		Tagged/ untagged	/27	TOR1: TOR2:

Top-of-rack (ToR) and OOB switch configuration might also require configuring settings such as hostnames, IP routes, DCB priority settings, and so on. The following table captures these items as a checklist.

Table 12. Network configuration checklist

Item	Value
OOB switch hostname	
TOR1 switch hostname	
TOR2 switch hostname	
Enable password	
Additional user/password	
IP route on OOB (if needed)	
IP route on TOR1/TOR2 (if needed)	
DCB bandwidth for SMB traffic	

Host OS network checklist

Dell EMC recommends having consistent host naming and IP addressing across multiple nodes in the virtualized cluster deployment. The host OS network configuration includes naming for the virtual switches and adapters, and assigning hostnames and IP addresses.

The following table provides the checklist for capturing the host OS network switch and adapter details.

Table 13. Host OS network switch and adapter details

Item	Value
Management virtual switch (for nonconverged)	
Storage virtual switch (for nonconverged)	
Virtual switch (for fully converged)	
Management adapter	
Storage 1 adapter	
Storage 2 adapter	

For the host OS configuration in any deployment method, static IP address assignment is recommended for all networks. The following table provides the checklist for capturing the details of the host OS network switch and adapter.

Table 14. Host OS network checklist

	Hostname	Management IP	Storage1 IP	Storage2 IP	OOB IP	OOB hostname
Node 1						
Node 2						
Node 3						
Node 4						
Node 5						
Node 6						
Node 7						
Node 8						
Node 9						
Node 10						
Node 11						
Node 12						
Node 13						
Node 14						
Node 15						
Node 16						

Predeployment configuration

This section describes the predeployment configuration that must be performed before deploying the hyperconverged virtualized solution based on Dell EMC Microsoft Storage Spaces Direct Ready Nodes.

Network switch configuration

When considering the fully converged network topology of the Storage Spaces Direct solution, network resiliency is a critical option that is achieved from both a physical and logical standpoint. The Illustrations below are examples of the Dell EMC/Microsoft network design.

NOTE: Management network redundancy is a combination of either iDRAC or OS DNS/IP resolution.

Dell EMC recommends that you deploy a network topology that supports a dual control plane while sharing a single data plane. The Dell EMC proprietary technology is referred to as Virtual Link Trunk (VLT). This technology provides network resiliency for data I/O.

When configuring switch VLT redundant technology, VRRP provides a virtual floating IP address that any node can reference as a gateway. If a switch fails, the virtual IP is transferred to a peer switch.

Virtual Router Redundancy Protocol (VRRP) is an active/standby First Hop Redundancy Protocol (FHRP). When used with VLT peers, VRRP becomes an active/active protocol. The VRRP virtual MAC address is present as local destination address in the FIB table of both the VLT peers. Using this method, the backup VRRP router forwards intercepted frames whose destination MAC address matches the VRRP virtual MAC address.

Three basic types of networks are required for a standard Storage Spaces Direct deployment—Out-of-band (OOB) management, host management, and storage.

The number of network ports (two or four) used within the storage configuration determines whether you have two or four fault domains. This guidance provides all network topology examples with two storage fault domains.

For sample switch configurations see https://community.emc.com/docs/DOC-70310.

Table 15. Solution network VLANs

VLAN type	Subnet mask	Number of host IP addresses	VLAN ID tag/untag
OOB management		19	Untagged
Host management		17	Tagged or untagged
Storage 1 (Fault Domain 1)		16	Tagged or untagged
Storage 2 (Fault Domain 2)		16	Tagged or untagged

iDRAC and BIOS configuration

The R740xd, R740xd2, and R640 Storage Spaces Direct Ready Nodes are preconfigured at the factory for optimized system BIOS and iDRAC settings. This eliminates the need for customers to manually configure these settings to a recommended baseline.

NOTE: PowerEdge R440 servers are not factory configured with the Microsoft Storage Spaces Direct optimized BIOS and iDRAC configuration settings. For these servers, you must update the BIOS and iDRAC configuration settings before deployment. The list of all optimized configuration settings is available at https://www.dell.com/support/article/us/en/19/sln313842

The integrated Dell Remote Access Controller (iDRAC) in Dell EMC Microsoft Storage Spaces Direct Ready Nodes can be configured to obtain an IP address from DHCP or can be assigned a static IP address. When the OOB network in the environment does not provide DHCP IP addresses, an IPv4 address must be statically configured on each iDRAC network interface manually. This can be done by accessing the physical server console by using KVM or other means.

Perform the following steps to configure the IPv4 address for iDRAC. This method can be used for configuring any additional BIOS settings.

1. Press F12 during the system boot.

Figure 9. Enter iDRAC

NOTE: PowerEdge R440 servers do not have the customized BIOS splash screen.

2. Select **iDRAC Settings**.

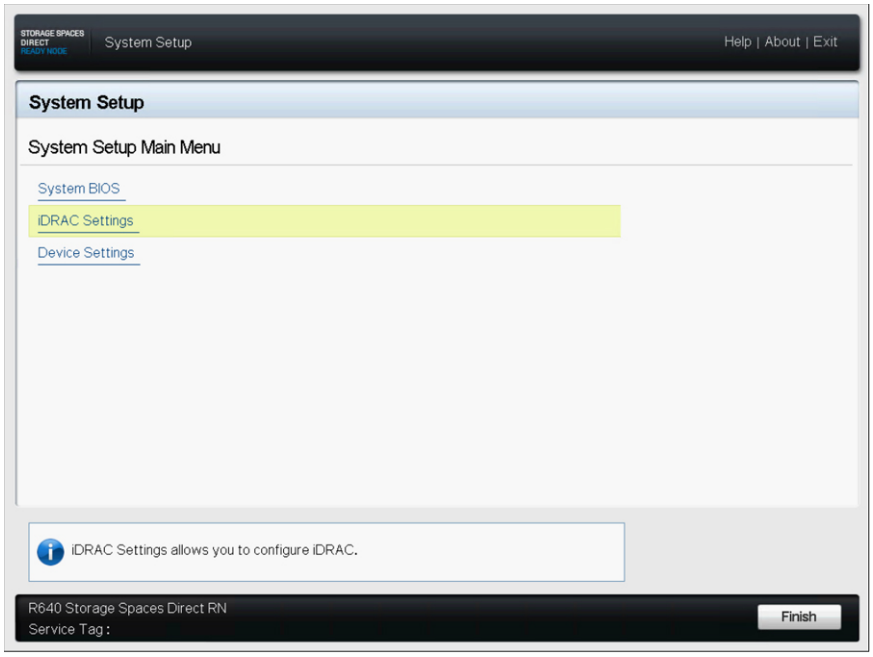

Figure 10. System Setup Main Menu

3. Select **Network**.

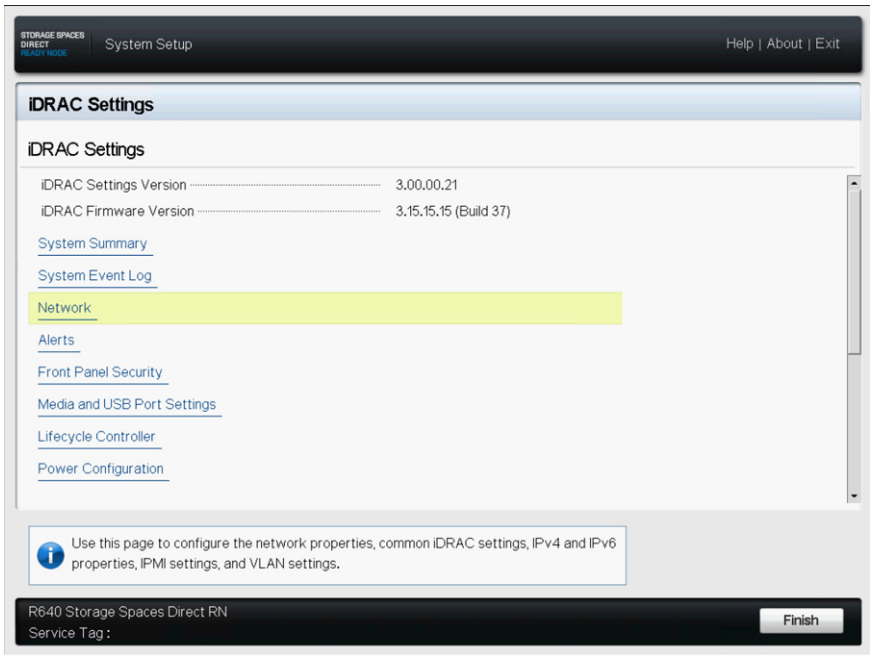

Figure 11. iDRAC Settings

4. In **IPV4 Settings**, for **Enable IPv4**, select **Enabled**.

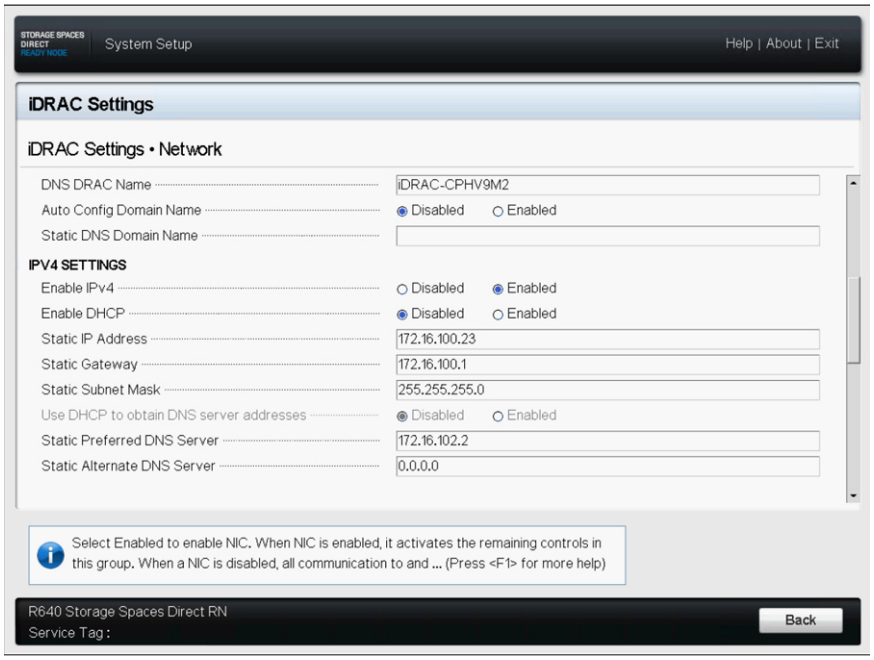

Figure 12. Network Settings

5. Enter the static IPv4 address details.
6. Click **Back** and then click **Finish** to return to the System Setup page.

QLogic NIC configuration

QLogic FastLinQ 41262 network adapter supports both iWARP and RoCE for RDMA. The Storage Spaces Direct Ready Nodes are validated with only iWARP for RDMA when using the QLogic network adapters and therefore based on the network configuration chosen, you must configure the adapter manually to enable iWARP for RDMA.

About this task

Perform the following steps for each port to configure the QLogic network adapters.

Steps

1. Press F2 during system boot to enter System Setup.
2. Click **System BIOS**and select **Device Settings**.

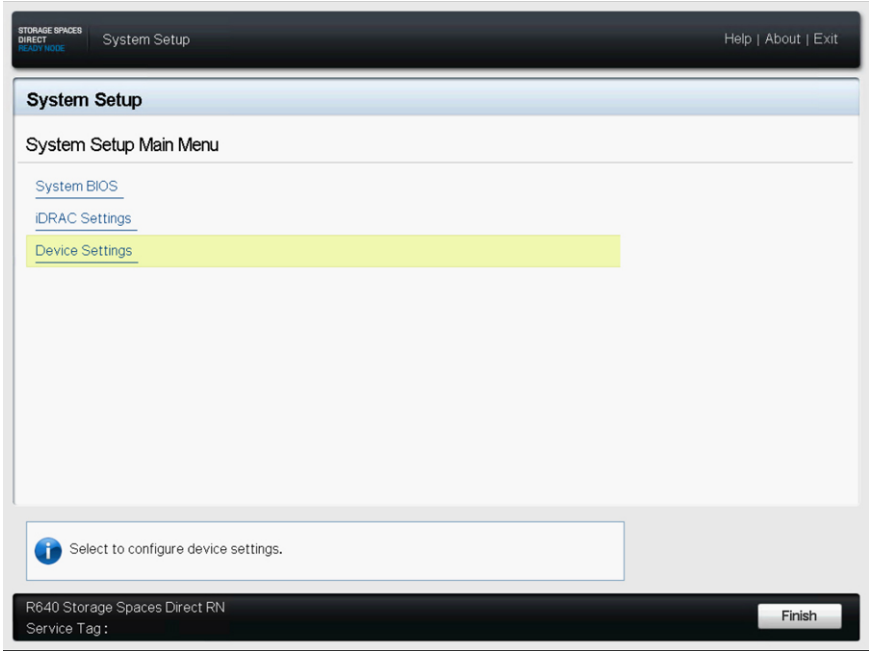

Figure 13. Device settings

3. Select the QLogic network adapter from the list of adapters.

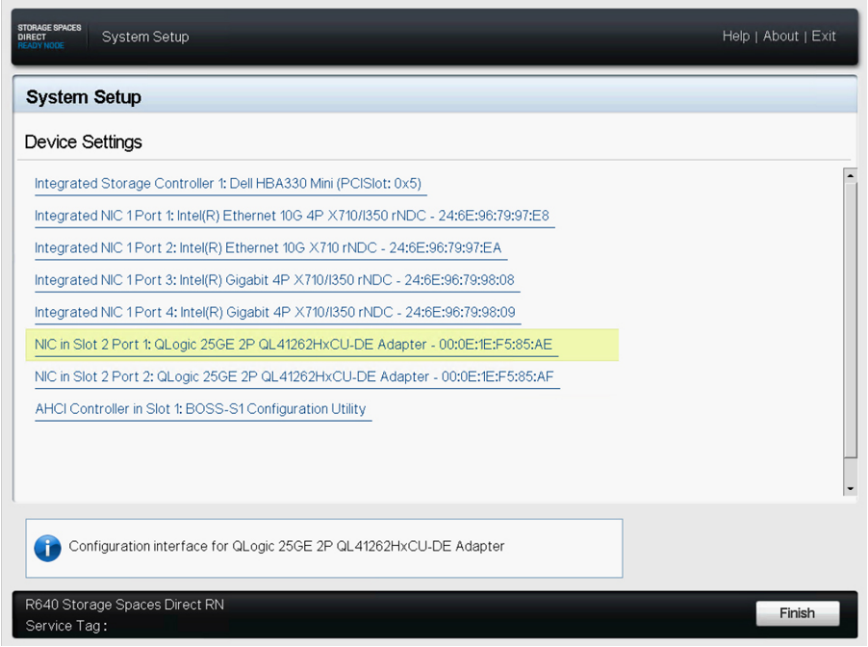

Figure 14. System Setup

4. Click **Device Level Configuration** and ensure that **Virtualization Mode** is set to **None**.

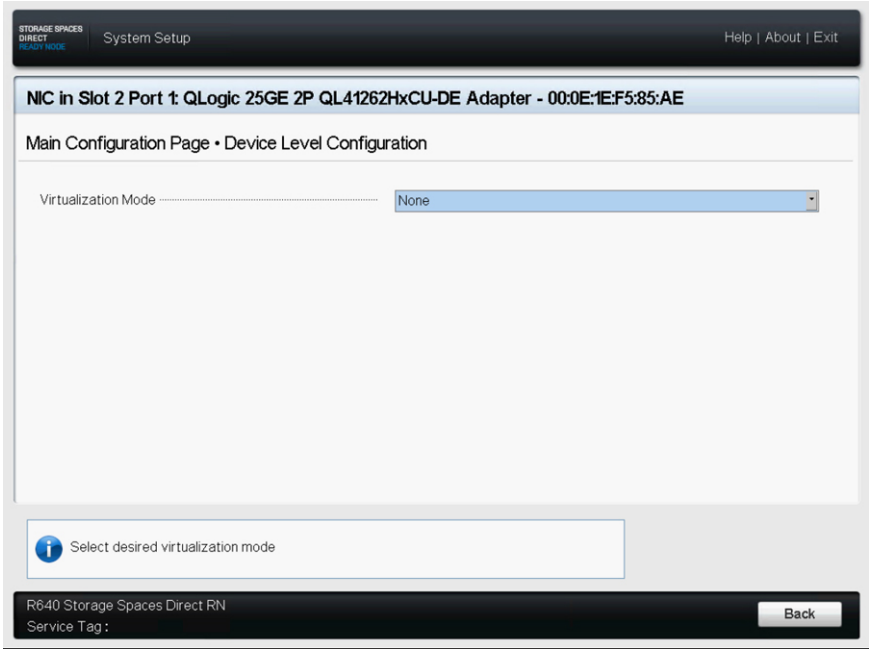

Figure 15. Virtualization Mode

5. Click **Back**, and then click **NIC Configuration**.
6. Select the following options in the **NIC Configuration** page:
   - Link Speed—SmartAN
   - NIC + RDMA Mode—Enabled
   - RDMA Operational Mode—iWARP
   - Boot Protocol—None
   - Virtual LAN Mode—Disabled

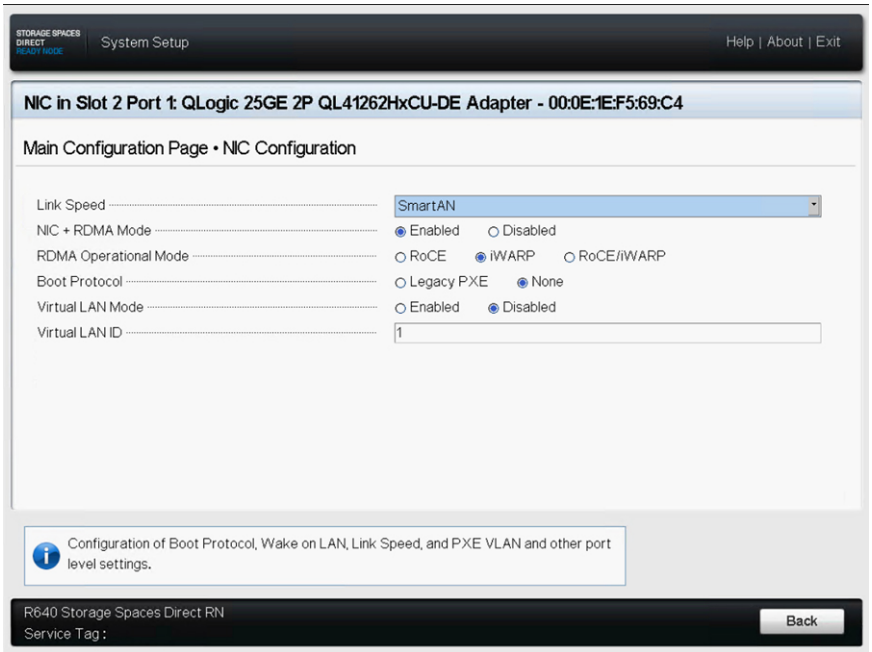

Figure 16. NIC configuration options

7. Click **Back** and click **Data Center Bridging (DCB) Settings**.
8. In the **Data Center Bridging (DCB) Settings** page, set **DCBX Protocol** to **Disabled**.

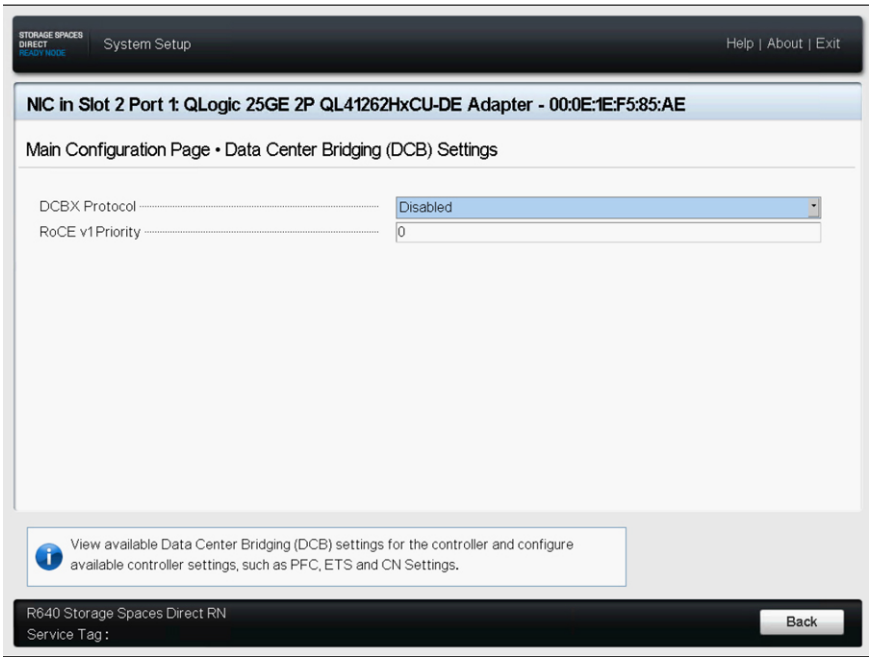

Figure 17. DCB settings

9. Click **Back**, click **Finish**, and then click **Yes** to save the settings.
10. Click **Yes** to return to the Device Settings page.
11. Select the second port of the QLogic adapter and repeat the preceding steps.
12. Click **Finish** to return to the System Setup page.
13. Click **Finish** to reboot the system.

Firmware baselining

The Dell EMC Ready Node has a supported firmware matrix and the nodes in the cluster must comply with the firmware matrix. It is important to ensure that each server has the right firmware revisions for components used within the server.

This can be verified by using the iDRAC system inventory feature or by using a scripted method such as the RACADM command line interface.

Hyperconverged infrastructure deployment

This section describes the steps involved in installing OS on the bare metal servers and deploying the hyperconverged infrastructure (HCI) with Storage Spaces Direct. PowerShell commands are provided to configure cluster deployment from the command line.

Unless mentioned otherwise, perform the following steps on each physical node in the infrastructure that will be a part of Storage Spaces Direct HCI.

OS deployment

There are two methods to deploy the operating system:

- Manual OS deployment—Manual installation from OS deployment to cluster creation
- Factory OS deployment—Factory preinstallation of Windows Server 2016 or Windows 2019 on Dell EMC Ready Nodes

NOTE: The steps in the subsequent sections are applicable to either full OS or Server Core.

NOTE: The command output shown in the subsequent sections might show only Mellanox ConnectX-4 LX adapters as physical adapters. This output is shown only as an example.

NOTE: For the PowerShell commands in this section and subsequent sections that require a network adapter name, run the `Get-NetAdapter` cmdlet to retrieve the correct value for the associated physical network port. The network

adapter names used in the commands in this guide are shown only as examples and might not represent the correct naming convention for what is installed in the system.

Manual OS deployment

Dell Lifecycle Controller and Integrated Dell Remote Access Controller (iDRAC) provide various options for operating systems deployment. Options include manual or unattended installation by using the virtual media and OS deployment feature in Lifecycle Controller.

The step-by-step procedure for deploying the operating system is not within the scope of this guide.

The subsequent steps in this guide assume that Windows Server 2016 or Windows Server 2019 Datacenter Server Core installation on the physical server is complete and that you have access to the iDRAC virtual console of the physical server.

Factory installed OS

If the cluster nodes are shipped from Dell EMC factory with pre-installed Windows Server 2016 or Windows Server 2019 Data Center edition OEM license, the Out of box experience (OOBE) needs to be completed. This includes the following steps:

- Selecting Language and Locale Settings
- Accepting Microsoft and OEM EULA
- Setting up password for the Local Administrator account
- Updating the OS partition size and shrinking it as needed

The factory installed OEM OS is pre-activated and the Hyper-V role is pre-deployed. Therefore, after the OOBE steps are complete, the post OS deployment steps described in section Install Roles and Features should be performed to complete the cluster deployment and Storage Spaces Direct configuration.

Install roles and features

Deployment and configuration of a Windows Server 2016 or Windows Server 2019 Storage Spaces Direct hyperconverged Infrastructure (HCI) cluster requires enabling the following operating system roles and features:

- Hyper-V service (not required if the OS is factory-installed)
- Failover clustering
- Data center bridging (required only when implementing fully converged network topology with RoCE for RDMA, and when implementing DCB for the fully converged topology with iWarp for RDMA)

You can enable these features by running the `Install-WindowsFeature` PowerShell cmdlet:

```
Install-WindowsFeature -Name Hyper-V, Failover-Clustering, Data-Center-Bridging -
IncludeAllSubFeature -IncludeManagementTools –Verbose
```

NOTE: Hyper-V role installation requires a reboot of the system. Because the subsequent steps also require a reboot, the required reboots are combined into one.

Update out-of-box drivers

For certain system components, you might have to update the driver to the latest Dell EMC supported version, as listed in the Supported Firmware and Software Matrix.

Run the following PowerShell command to retrieve a list of all driver versions that are currently installed on the local system:

```
Get-PnpDevice | Select-Object Name, @{l='DriverVersion';e={(Get-PnpDeviceProperty -InstanceId
$_.InstanceId -KeyName 'DEVPKEY_Device_DriverVersion').Data}} -Unique
```

Before configuring host networking, ensure that the out-of-box (OOB) drivers are updated. After identifying the required driver version, download the driver installers from https://www.dell.com/support or by using the Dell EMC Solutions for Azure Stack HCI Update Catalog.

NOTE: The QLogic FastLinQ adapter does not have an in-box driver in Windows Server 2016. Install the driver before attempting host network configuration.

After you download the drivers, attach a folder containing the driver DUP files to the system as a virtual media image. Attach the virtual media image folder as follows:

1. Click **Virtual Media** in the iDRAC virtual console menu.
2. Click **Create Image**.

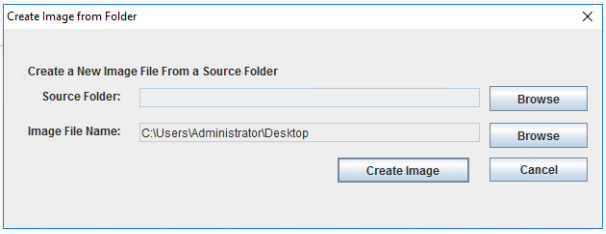

Figure 18. Create new image

3. Click **Browse**, select the folder where the driver DUP files are stored, and, if required, change the name of the image.

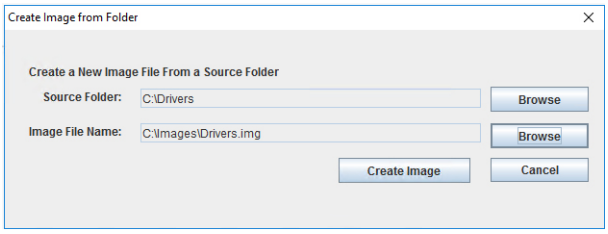

Figure 19. Virtual media image folder

4. Click **Create Image**.
5. Click **Finish**.
6. From the **Virtual Media** menu, click **Connect Virtual Media**.
7. From the **Virtual Media** menu, click **Map Removable Disk**, click **Browse**, and select the image that you created.

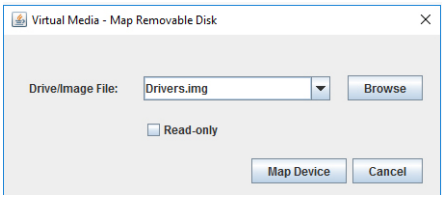

Figure 20. Map device

After the image is mapped, it appears as a drive in the host OS. You can then run the driver DUP files to install the OOB drivers.

Change hostname

By default, the OS deployment assigns a random name as the host computer name. For easier identification and uniform configuration, Dell EMC recommends that you change the hostname to something that is easily identifiable and relevant. You can change the hostname by using the `Rename-Computer` cmdlet:

```
Rename-Computer -NewName S2DNode01 -Restart
```

NOTE: This command induces an automatic restart at the end of rename operation.

Configure firewall

For the cluster operations post-deployment and optional monitoring configuration, certain firewall rules have to be enabled on the cluster nodes. For a complete list of ports or firewall rules that need to be enabled, see Appendix A.

For configuring firewall rules at the command prompt, see technet.microsoft.com/EN-US/library/jj554906(v=wps.630).aspx.

Cluster deployment and configuration

The following sections provide instructions on cluster deployment and Storage Spaces Direct configuration.

Configure host networking

This section focuses on configuring the host networking such as VM switches, VM network adapters, and other QoS and RDMA configurations.

NOTE: All PowerShell commands in this section must be run at the local console to ensure that there are no failures due to network disconnections during configuration.

VM switch and adapter configuration

Dell EMC Solutions for Azure Stack HCI can be configured to implement a fully converged or nonconverged network for storage and management connectivity. Dell EMC recommends a nonconverged network implementation and using physical network adapters for the storage traffic rather than implementing Switch Embedded Teaming (SET). However, in a nonconverged configuration option, if VM adapters need to have RDMA capability, SET configuration is necessary for storage adapters.

NOTE: For specifics of configuration such as VM switch name, adapter names and VLAN IDs, see the Deployment checklist.

You must configure the host OS network based on the physical connectivity of the network (fully converged or nonconverged).

The following sections describe the two different host networking configurations and provide the commands for configuration.

NOTE: The following configuration instructions assume that only two physical ports are used for storage in a nonconverged configuration or for the VM switch in a fully converged configuration. For configuration instructions that use four physical ports for storage or fully converged SET, see https://community.emc.com/docs/DOC-73779.

Fully converged network

The following figure illustrates the fully converged configuration in a Storage Spaces Direct cluster node.

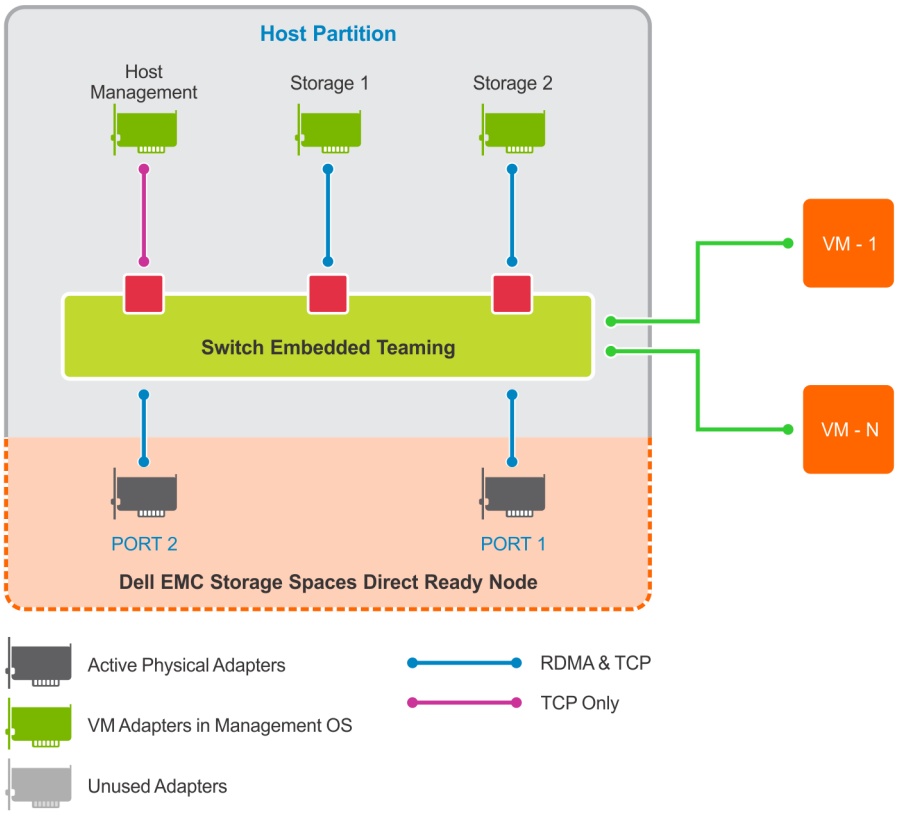

Figure 21. Fully converged VM switch and adapter configuration

The following PowerShell commands can be used to configure the VM switch as a switch embedded team and configure the VM network adapters in the host OS for management, storage 1, and storage 2 traffic.

NOTE: The default load balancing algorithm for Windows Server 2016 SET is Dynamic and in Windows Server 2019 it is Hyper-V port. Dell EMC recommends that you always use the default load balancing algorithm.

Perform the following steps to configure the OS network:

1. Run the following command to create a VM switch in the SET configuration by using the physical network ports from the Mellanox or QLogic network adapters in the system:

```
New-VMSwitch -Name S2DSwitch -AllowManagementOS 0 -NetAdapterName 'SLOT 1 PORT 1','SLOT 1
PORT 2' -MinimumBandwidthMode Weight -Verbose
```

In this example, the arguments to the NetAdapterName parameter represents the physical NIC ports that must be a part of the SET configuration. These interface names can be retrieved by running the `Get-NetAdapter` cmdlet.

NOTE: The minimum Bandwidth Mode set to Weight can be used to shape the VM Network traffic, and it is not used for host OS network adapters. Setting the minimum Bandwidth Mode is optional.

```
Name                      InterfaceDescription                    ifIndex Status
MacAddress             LinkSpeed
----                      --------------------                    ------- ------
----------             ---------
NIC2                      Intel(R) Ethernet 10G X710 rNDC               7 Disconnected
24-6E-96-52-CC-A4        10 Gbps
NIC4                      Intel(R) I350 Gigabit Network Conn...#2       6 Disconnected
24-6E-96-52-CC-C3         0 bps
NIC3                      Intel(R) I350 Gigabit Network Connec...       2 Disconnected
24-6E-96-52-CC-C2         0 bps
NIC1                      Intel(R) Ethernet 10G 4P X710/I350 rNDC       4 Disconnected
24-6E-96-52-CC-A2        10 Gbps
SLOT 1 Port 2             Mellanox ConnectX-4 Lx Ethernet Ad...#2      10 Up
24-8A-07-59-4C-69        10 Gbps
SLOT 1 Port 1             Mellanox ConnectX-4 Lx Ethernet Adapter       8 Up
24-8A-07-59-4C-68        10 Gbps
```

The argument “0” to the AllowManagementOS parameter prevents creation of a VM network adapter in the host operating system.

This command creates a switch embedded team with Switch Independent Teaming mode and Dynamic Load Balancing algorithm settings.

NOTE: For the host management network, do not run the `Set-VMNetworkAdapterVlan` command if DHCP is used (with untagged VLANs) for management.

2. Run the following command to create and configure the host management network adapter:

```
Add-VMNetworkAdapter -ManagementOS -Name 'Management' -SwitchName S2DSwitch -Passthru |
Set-VMNetworkAdapterVlan -Access -VlanId 102 -Verbose
```

3. Run the following command to add the host OS VM network adapters for Storage 1 and Storage 2 traffic and configure the VLAN IDs:

```
Add-VMNetworkAdapter -ManagementOS -Name 'Storage1' -SwitchName S2DSwitch -Passthru | Set-
VMNetworkAdapterVlan -Access -VlanId 103 -Verbose

Add-VMNetworkAdapter -ManagementOS -Name 'Storage2' -SwitchName S2DSwitch -Passthru | Set-
VMNetworkAdapterVlan -Access -VlanId 104 -Verbose
```

NOTE: The number of storage VLANs can be equal to the number of storage adapters or 1.

4. Run the following command to configure IP addresses:

```
#Host Management Adapter
New-NetIPAddress -InterfaceAlias ‘vEthernet (Management)’ -IPAddress 172.16.102.51
DefaultGateway 172.16.102.1 -PrefixLength 25 -AddressFamily IPv4 -Verbose
#Storage 1 Adapter
New-NetIPAddress -InterfaceAlias ‘vEthernet (Storage1)’ -IPAddress 172.16.103.51 -
PrefixLength
27 -AddressFamily IPv4 -Verbose
#Storage 2 Adapter
New-NetIPAddress -InterfaceAlias ‘vEthernet (Storage2)’ -IPAddress 172.16.104.51 -
PrefixLength
27 -AddressFamily IPv4 -Verbose
```

Nonconverged network

The following figure provides a high-level overview of VM switch and management adapter configuration in the Storage Spaces Direct Ready Node in a nonconverged network configuration. This approach includes physical links for the storage traffic and a switch embedded team for host management and VM traffic.

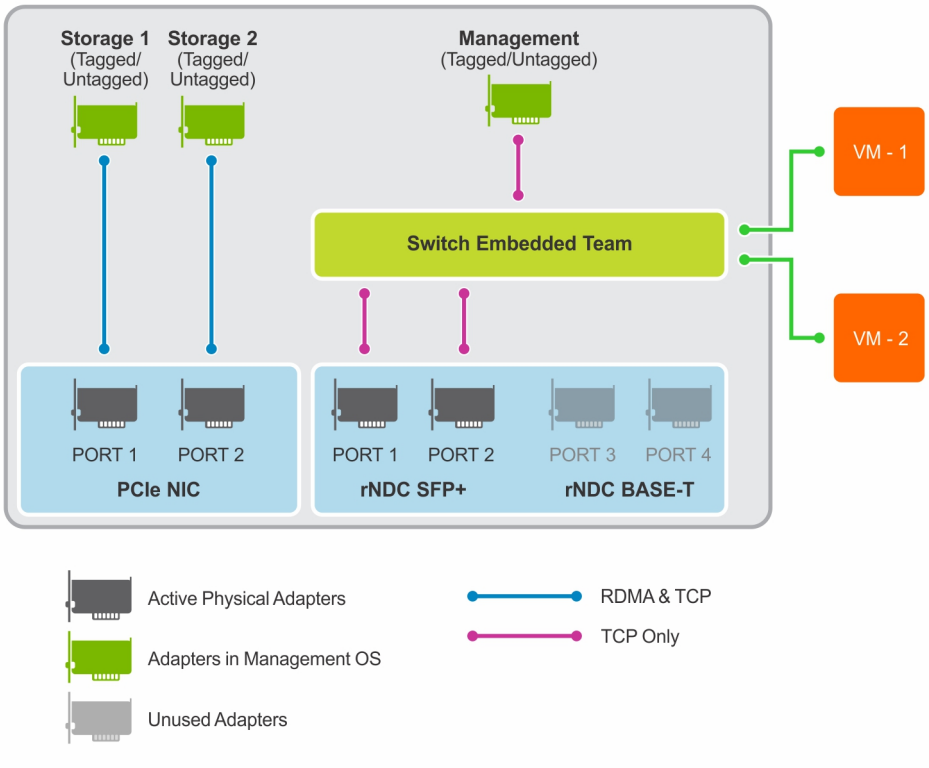

Figure 22. Nonconverged network configuration

NOTE: Nonconverged network configuration can implement a SET VM switch for storage as well. To configure SET for storage, see https://community.emc.com/docs/DOC-73779

NOTE: The default load balancing algorithm for Windows Server 2016 SET is Dynamic and in Windows Server 2019 it is Hyper-V port.

The physical NIC ports from the Mellanox or QLogic network adapters are used for storage traffic, and the SFP+ ports or the 1 Gb BASE-T ports from the rNDC are used to create a VM switch in a SET configuration. The host management and VM adapters are attached to this VM switch.

1. Run the following command to create a VM switch in the SET configuration by using the 10 GbE SFP+ network ports from rNDC:

```
    New-VMSwitch -Name Management -AllowManagementOS 0 -NetAdapterName 'NIC1,'NIC2' -
MinimumBandwidthMode Weight -Verbose
```

2. Run the following command to create and configure the host management network adapter.

   NOTE: For the host management network, do not run the `Set-VMNetworkAdapterVlan` command if DHCP is used (with untagged VLANs) for management.

```
Add-VMNetworkAdapter -ManagementOS -Name 'Management' -SwitchName Management -Passthru |
Set-VMNetworkAdapterVlan -Access -VlanId 102 -Verbose
```

3. Run the following command to configure VLAN IDs, if required, on the physical adapters used for storage traffic:

```
Set-NetAdapter -Name 'SLOT 1 PORT 1' -VlanID 103
Set-NetAdapter -Name 'SLOT 1 PORT 2' -VlanID 104
```

   NOTE: The number of storage VLANs can be equal to the number of storage adapters or 1.

4. Run the following command to configure IP addresses:

```
#Host Management Adapter New-NetIPAddress -InterfaceAlias 'vEthernet (Management)' -
IPAddress 172.16.102.51 DefaultGateway 172.16.102.1 -PrefixLength 25 -AddressFamily IPv4 -
Verbose
#DNS server address
Set-DnsClientServerAddress -InterfaceAlias 'vEthernet (Management)' -ServerAddresses
172.16.102.202
#Storage 1 Adapter
New-NetIPAddress -InterfaceAlias 'SLOT 1 PORT 1' -IPAddress 172.16.103.51 -PrefixLength
27 -AddressFamily IPv4 -Verbose
#Storage 2 Adapter
New-NetIPAddress -InterfaceAlias 'SLOT 1 PORT 2' -IPAddress 172.16.104.51 -PrefixLength
27 -AddressFamily IPv4 -Verbose
```

NOTE: Static IP assignment on the management adapter is not required if a DHCP server is being used for management network.

In this configuration, default gateway and DNS configuration are required only for the host management network.

AD domain join

The cluster nodes must be a part of an Active Directory domain before you can create a cluster. You can perform the domain join task by running the `Add-Computer` cmdlet.

See the Deployment checklist for the domain administrator or equivalent credentials that are required for the domain join.

NOTE: Connecting to AD directory services by using the host management network might require routing to the AD network. Ensure that this routing is in place before proceeding to domain join.

```
$credential = Get-Credential
Add-Computer -DomainName S2dlab.local -Credential $credential -Restart
```

NOTE: This command induces an automatic restart at the end of the domain join operation. You must run this command on each host that will be a part of the Storage Spaces Direct cluster.

NOTE: Optionally, you can add all newly created computer objects from the cluster deployment to a different Organizational Unit (OU) in the AD directory services. In this case, you can use the -OUPath parameter along with the `Add-Computer` cmdlet.

Create host cluster

Before creating a host cluster, ensure that the nodes that will be a part of the cluster are configured as needed and are ready for the cluster creation. This can be done using the `Test-Cluster` cmdlet.

NOTE: The commands in this section need to be executed on only one node in the infrastructure.

NOTE: Before creating the host cluster, execute the `Get-PhysicalDisk` command on all cluster nodes and verify the output to ensure that all disks are in healthy state and there are equal number of disks per node.

NOTE: Validate that the nodes have homogeneous hardware configuration.

```
Test-Cluster -Node S2Dnode01, S2DNode02, S2dNode03, S2dNode04 -Include 'Storage Spaces
Direct', 'Inventory', 'Network', 'System Configuration'
```

NOTE: The `Test-Cluster` cmdlet generates an HTML report of all validations performed and includes a summary of the validation. Review this report before creating a cluster.

Run the `New-Cluster` cmdlet to create the host cluster as follows:

```
New-Cluster -Name S2DSystem -Node S2Dnode01, S2DNode02, S2dNode03, S2dNode04 -StaticAddress
172.16.102.55 -NoStorage -IgnoreNetwork 172.16.103.0/27, 172.16.104.0/27 -Verbose
```

In this command, the StaticAddress parameter is used to specify an IP address for the cluster in the same IP subnet as the host management network. The NoStorage switch parameter specifies that the cluster is to be created without any shared storage.

NOTE: The `New-Cluster` cmdlet generates an HTML report of all configurations performed and includes a summary of the validation. Review the report before enabling Storage Spaces Direct.

Configuring Storage Spaces Direct

After you create the cluster, you can use the `Enable-ClusterS2D` cmdlet to create the Storage Spaces Direct configuration on the cluster.

Do not run the `Enable-ClusterS2D` cmdlet in a remote session. Instead, use the local console session.

```
Enable-ClusterS2D -Verbose
```

The `Enable-ClusterS2D` cmdlet generates an HTML report of all configurations performed and includes a summary of the validation. Review this report. The report is typically stored in the local temporary folder on the node where the `Enable-ClusterS2D` cmdlet was run. The verbose output of the command shows the path to the cluster report.

At the end of the operation, this cmdlet discovers and claims all the available disks into an auto-created storage pool.

You can verify the cluster creation by running any of the following commands:

```
Get-ClusterS2D
Get-StoragePool
Get-StorageSubSystem -FriendlyName *Cluster* | Get-StorageHealthReport
```

Change RDMA mode on QLogic NICs—iWARP only

In the predeployment configuration, the QLogic 41262 NICs are configured to use iWARP for RDMA. However, the driver in the OS defaults to RoCE v2 for RDMA. Change the setting by running the `Set-NetAdapterAdvancedProperty` cmdlet.

NOTE: This change is required only for Ready Nodes with QLogic 41262 adapters that are used for storage traffic.

NOTE: For QLogic 41262 driver versions earlier than 8.37.37.0, the display name of the attribute is 'RDMA Mode'.

```
Set-NetAdapterAdvancedProperty -Name 'SLOT 1 PORT 1' -DisplayName 'NetworkDirect Technology' -
DisplayValue 'iWarp'

Set-NetAdapterAdvancedProperty -Name 'SLOT 1 PORT 2' -DisplayName 'NetworkDirect Technology' -
DisplayValue 'iWarp'
```

RDMA configuration

Storage Spaces Direct hyperconverged Infrastructures with Mellanox ConnectX-4 LX, Mellanox ConnectX-5 EX, and QLogic FastLinQ 41262 adapters use converged network design and RDMA for storage traffic. Hence, it is important to ensure that the physical and virtual adapters used for storage traffic have RDMA enabled.

Perform the following steps to enable RDMA for live migration and storage traffic:

1. Run the following command to enable RDMA on the storage virtual adapters and physical network ports.

   NOTE: You can retrieve the argument to the Name parameter by running the `Get-NetAdapter` cmdlet.

   ```
   Enable-NetAdapterRDMA -Name 'vEthernet (Storage1)', 'vEthernet (Storage2)'
   ```

   ```
       Enable-NetAdapterRDMA -Name 'SLOT 1 PORT 1', 'SLOT 1 PORT 2'
   ```

2. Run the following command to map the storage virtual adapters in the host OS to the physical NIC ports to ensure that the storage traffic uses these affinity rules, and traffic spreads across evenly. This is not required if the storage traffic is using the physical network ports.

   ```
   Set-VMNetworkAdapterTeamMapping -VMNetworkAdapterName "Storage1" -ManagementOS -
   PhysicalNetAdapterName "SLOT 1 PORT 1"
   Set-VMNetworkAdapterTeamMapping -VMNetworkAdapterName "Storage2" -ManagementOS -
   PhysicalNetAdapterName "SLOT 1 PORT 2"
   ```

   For each storage network adapter, the `RDMA Capable` value must be `True`.

NOTE: With the inbox Mellanox drivers, the value appears as `False`. You can fix this issue by installing the updated drivers from Dell EMC. See the Support Matrix for details about the supported versions of firmware and drivers.

3. Verify the RDMA configuration for the virtual storage adapters (in fully converged topology) by running the following command:

```
Get-SmbClientNetworkInterface | Where-Object { $_.FriendlyName -Like "*Storage*}
```

4. Run the following command to enable RDMA for live migration traffic:

```
Set-VMHost –VirtualMachineMigrationPerformanceOption SMB
```

QoS policy configuration

NOTE: For configurations that use QLogic 41262 network adapters with iWARP for RDMA, and NVMe drives for storage, configuring DCB in the OS and top-of-rack (ToR) switches is recommended to ensure optimal storage performance. For other disk configurations such as the hybrid and/or all-flash SSD configurations, configuring DCB for iWARP is not necessary. For all Ready Node configurations using Mellanox adapters in a fully converged network topology, implement the steps that are provided in this section.

The Dell EMC Microsoft Storage Spaces Direct Ready Nodes that are configured with Mellanox ConnectX-4 LX and Mellanox ConnectX-5 EX adapters require DCB configuration on the ToR switches and Quality of Service (QoS) configuration in the host OS. Ensure that the QoS policies are configured to prioritize the SMB traffic related to the storage adapters. Ensure that the QoS configuration in the host OS matches the QoS configuration that is performed in Network switch configuration.

The following table provides an overview of the QoS priorities and the required state for the set of priorities.

NOTE: QoS configuration is needed only for a fully converged network topology. For a nonconverged network, the storage traffic uses dedicated NIC ports; therefore, DCB/PFC and QoS configuration is not needed.

Table 16. QoS priorities

QoS priority	State
0,1,2,4,6,7	Disabled
3	Enabled
5	Enabled

Perform the following steps to configure QoS in the host OS:

1. Run the following command to create a new QoS policy with a match condition set to 445; this displays the TCP port dedicated for Server Message Block traffic:

```
New-NetQosPolicy –Name 'SMB' –NetDirectPortMatchCondition 445 –PriorityValue8021Action 3
New-NetQosPolicy 'Cluster' -Cluster –PriorityValue8021Action 5
```

   The arguments `3` and `5` to the PriorityValue8021Action parameter indicate the IEEE 802.1p values for SMB and cluster traffic. The values must be in the Enabled state, as shown in the preceding table.

2. Run the following commands to map the IEEE 802.1p priority enabled in the system to a traffic class:

```
New-NetQosTrafficClass –Name 'SMB' –Priority 3 –BandwidthPercentage 50 –Algorithm ETS
New-NetQosTrafficClass –Name 'Cluster' –Priority 5 –BandwidthPercentage 1 –Algorithm ETS
```

   The preceding commands specify that the transmission algorithm is Enhanced Transmission Selection (ETS) and the traffic classes SMB and Cluster get bandwidth percentages of 50 and 1 respectively.

   NOTE: In the nonconverged network topology, configuring DCB/PFC in the ToR switches or QoS in the host OS is not necessary.

   NOTE: The argument to `–BandwidthPercentage` shown in this step is an example only. This can be and should be modified based on the infrastructure requirements and type of network configuration.

3. Run the following command to configure flow control for the priorities that are shown in the preceding table:

```
Enable-NetQosFlowControl –Priority 3,5
Disable-NetQosFlowControl –Priority 0,1,2,4,6,7
```

4. Run the following command to enable QoS for the Mellanox network adapter ports. The argument for the InterfaceAlias can be retrieved by using the `Get-NetAdapter` cmdlet.

```
Enable-NetAdapterQos -InterfaceAlias 'SLOT 1 PORT 1','SLOT 1 PORT 2'
```

5. Run the following command to disable DCBX Willing mode in the operating system:

```
Set-NetQosDcbxSetting -Willing $false
```

6. Run the following commands to set Mellanox NIC DCBX mode to `Host in charge`:

```
Set-NetAdapterAdvancedProperty -Name 'SLOT 1 PORT 1' -DisplayName 'Dcbxmode' -DisplayValue
'Host in charge'

Set-NetAdapterAdvancedProperty -Name 'SLOT 1 PORT 2' -DisplayName 'Dcbxmode' -DisplayValue
'Host in charge'
```

Remove host management network from live migration

After you create the cluster, live migration is configured by default to use all available networks. Disable live migration on the host management network by excluding the host management network from the live migration settings.

Run the following PowerShell commands to exclude the host management network:

```
$clusterResourceType = Get-ClusterResourceType -Name 'Virtual Machine'

$hostNetworkID = Get-ClusterNetwork | Where-Object { $_.Address -eq '172.16.102.0' } | Select
-ExpandProperty ID

Set-ClusterParameter -InputObject $clusterResourceType -Name MigrationExcludeNetworks -Value
$hostNetworkID
```

In the preceding command, `172.16.102.0` represents the host management subnet.

Update hardware timeout for Spaces port

NOTE: For performance optimization and reliability, Dell EMC recommends that you update the hardware timeout configuration for the Spaces port.

Run the following PowerShell commands on every node in the cluster to update the configuration in the Windows registry:

```
Set-ItemProperty -Path HKLM:\SYSTEM\CurrentControlSet\Services\spaceport\Parameters -Name
HwTimeout -Value 0x00002710 -Verbose

Restart-Computer -Force
```

This command induces a reboot of the node at the end of the registry update. Perform this update on all Storage Spaces Direct nodes being deployed immediately after initial deployment. Update one node at a time and wait until each node rejoins the cluster.

Update page file settings

To be able to capture the active memory dump when a fatal system error occurs, allocate sufficient size for the page file. Dell EMC recommends that the size be at least 40 GB plus the size of CSV block cache.

Determine the cluster CSV block cache size value by running the following command:

```
$blockCacheMB = (Get-Cluster).BlockCacheSize
```

NOTE: On Windows Server 2016, the default block cache size is 0. On Windows Server 2019, the block cache is set to a value 1024 (1 GB).

To update the page file settings, run the following command:

```
$pageFilePath = "C:\pagefile.sys"
$initialSize = [Math]::Round(40960 + $blockCacheMB)
$maximumSize = [Math]::Round(40960 + $blockCacheMB)

$system = Get-WmiObject -Class Win32_ComputerSystem -EnableAllPrivileges
if ($system.AutomaticManagedPagefile) {
    $system.AutomaticManagedPagefile = $false
    $system.Put()
}

$currentPageFile = Get-WmiObject -Class Win32_PageFileSetting
if ($currentPageFile.Name -eq $pageFilePath)
{
    $currentPageFile.InitialSize = $InitialSize
    $currentPageFile.MaximumSize = $MaximumSize
    $currentPageFile.Put()
}
else
{
    $currentPageFile.Delete()
    Set-WmiInstance -Class Win32_PageFileSetting -Arguments @{Name=$pageFilePath; InitialSize
= $initialSize; MaximumSize = $maximumSize}
}
```

Enabling jumbo frames

Enabling jumbo frames specifically on the interfaces supporting the storage network might help improve the overall read/write performance of the Storage Spaces Direct cluster. It is important to note that an end-to-end configuration of jumbo frames is required to take advantage of the technology. In addition, considerations need to be made when configuring the technology because support for jumbo frame sizes varies between software, NIC, and switch vendors. The lowest value within the data path determines the maximum frame size used for that path.

For the storage network adapters in the host OS, enable jumbo frames by running the `Set-NetworkAdapterAdvancedProperty` cmdlet.

NOTE: Network adapters from different vendors support different jumbo packet sizes. The configured value must be consistent across the host OS and network switch configuration.

For information about configuring jumbo frames at the switch port level, see Sample Switch Configurations.

Configuring a cluster witness

Microsoft recommends configuring a cluster witness for a 4-node Storage Spaces Direct cluster. A cluster witness must be configured for a 2-node cluster.

Cluster witness configuration helps maintain a cluster or storage quorum when there is a node or network communication failure where nodes continue to operate but can no longer communicate with one another.

A cluster witness can be either a file share or a cloud-based witness.

NOTE: If you choose to configure a file share witness, it should exist outside the 2-node cluster.

For information about configuring a file share witness, see https://techcommunity.microsoft.com/t5/Failover-Clustering/New-File-Share-Witness-Feature-in-Windows-Server-2019/ba-p/372149.

For information about configuring a cloud-based witness, see https://docs.microsoft.com/en-us/windows-server/failover-clustering/deploy-cloud-witness.

Recommended next steps

Before proceeding to the next steps and operational management of the cluster, generate a cluster validation report to ensure that all configuration is in order. You can generate the report by running the `test-Cluster` cmdlet:

```
Test-Cluster -Node S2DNode01, S2DNode02, S2DNode03, S2DNode04 -Include 'System
Configuration', 'Inventory', 'Network', 'Storage Spaces Direct'
```

This command generates an HTML report with a list of all the tests that were performed and completed without errors.

Dell EMC recommends that you perform the following steps after you create the host cluster and enable Storage Spaces Direct:

1. Post-deployment verification—This step is recommended to ensure that the infrastructure is functional and ready for operations.
2. OS license activation—By default, the OS is installed in the evaluation mode. Activate the license immediately after OS installation.

NOTE: The OS license activation step is not required if the OS is factory-installed.

For more information about these steps, see the operations guide at https://www.dell.com/azurestackhcimanuals.

5

Deployment Services

Support for installation and configuration issues

Issues that arise during installation and configuration are not covered even if you have purchased Dell ProSupport or ProSupport Plus. Support for installation and configuration issues is provided under a separate paid services package. When you call Dell EMC with an installation and configuration issue, Dell Tech Support will route you to your Account Manager in Dell EMC Sales. The Account Manager will then help you in purchasing the onsite deployment services package.

6

Additional Resources

- iDRAC documentation
- Supported firmware and software matrix
- Storage Spaces Direct overview

A

Firewall Port Requirements

Table 17. Firewall port requirements

Source	Target	Protocol	Port	Comment
Any	Domain Controllers	TCP/UDP	53	DNS
		TCP/UDP	88	Kerberos
		UDP	123	NTP
		TCP	135	RPC, EMP
		UDP	137	NetLogon, NetBIOS Name Resolution
		UDP	138	DFSN, NetLogon, NetBIOS, Datagram Service
		TCP	139	DSFN, NetBIOS Session Service, NetLogon
		TCP/UDP	389	LDAP
		TCP/UDP	445	SMB, CIFS, SMB2, DFSN, LSARPC, NbtSS, NetLogonR, SAMR, SrvSvc
		TCP/UDP	464	Kerberos change/set password
		TCP	636	LDAP (SSL)
		TCP	3268	Global Catalog
		TCP	3269	Global Catalog (SSL)
		TCP	5722	RPC, DFSR (SYSVOL)
		TCP	9389	SOAP
		TCP	1025:5000	RPC, DCOM, EPM, DRSUAPI, NetLogon, SamR, FRS (2003)
		UDP	1025:5000	DCOM, RPC, EPM (2003)
		TCP	49152:65535	RPC, DCOM, EPM, DRSUAPI, NetLogonR, SamR, FRS (2008)
		UDP	49152:65535	DCOM, RPC, EPM (2008)
Local Subnet	All Hosts and VMs	UDP	137:138	Allow Name/Share Resolution
		TCP	139	Allow Name/Share Resolution
Any	Console VM	TCP	3389	Remote Desktop

Source	Target	Protocol	Port	Comment
WSUS (on VMM VM)	Any	TCP	80	SWUS Updates (HTTP)
		TCP	443	SWUS Updates (HTTPS)

B

Sample Deployment Checklists

Table 18. Sample checklist

Item	Value
AD domain FQDN	hci.lab
Domain administrator or equivalent credentials	Username: hci\administrator Password: <DO NOT WRITE IT DOWN>
DNS server addresses	dns.s2dlab.local
WSUS server FQDN (if needed)	wsus.s2dlab.local

Table 19. Sample checklist

Traffic class	Purpose	Minimum number of IP addresses needed	VLAN ID	Tagged/ untagged	Subnet mask	VLAN IP addresses
Out of band	Required for OOB management of server nodes and ToR switches	18	100	Untagged	/24	OOB: 172.16.100.1
Host management	Management of cluster and cluster nodes	17	102	Tagged	/25	TOR1: NA TOR2: NA
Storage 1	SMB traffic	16	103	Tagged	/27	TOR1: NA TOR2: NA
Storage 2	SMB traffic	16	104	Tagged	/27	TOR1: NA TOR2: NA

Table 20. Sample checklist

Item	Value
OOB switch hostname	S2D-OOB
TOR1 switch hostname	S2D-TOR1
TOR2 switch hostname	S2D-TOR2
Enable password	<DO NOT WRITE IT DOWN>
Additional user/password	NA
IP route on OOB (if needed)	NA
IP route on TOR1/TOR2 (if needed)	NA
DCB bandwidth for SMB traffic	50%

Table 21. Sample checklist

Item	Value
Virtual switch	S2DSwitch
Management adapter	Management
Storage 1 adapter	Storage1

Item	Value
Storage 2 adapter	Storage2

Table 22. Sample checklist

Node	Hostname	Management IP	Storage1 IP	Storage2 IP	OOB IP	OOB hostname
Node 1	S2DNode01	172.16.102.51	172.16.103.51	172.16.104.51	172.16.100.51	S2D-DRAC-1
Node 2	S2DNode02	172.16.102.52	172.16.103.52	172.16.104.52	172.16.100.52	S2D-DRAC-2
Node 3	S2DNode03	172.16.102.53	172.16.103.53	172.16.104.53	172.16.100.53	S2D-DRAC-3
Node 4	S2DNode04	172.16.102.54	172.16.103.54	172.16.104.54	172.16.100.54	S2D-DRAC-4
Node 5	S2DNode05	172.16.102.55	172.16.103.55	172.16.104.55	172.16.100.55	S2D-DRAC-5
Node 6	S2DNode06	172.16.102.56	172.16.103.56	172.16.104.56	172.16.100.56	S2D-DRAC-6
Node 7	S2DNode07	172.16.102.57	172.16.103.57	172.16.104.57	172.16.100.57	S2D-DRAC-7
Node 8	S2DNode08	172.16.102.58	172.16.103.58	172.16.104.58	172.16.100.58	S2D-DRAC-8
Node 9	S2DNode09	172.16.102.59	172.16.103.59	172.16.104.59	172.16.100.59	S2D-DRAC-9
Node 10	S2DNode10	172.16.102.60	172.16.103.60	172.16.104.60	172.16.100.60	S2D-DRAC-10
Node 11	S2DNode11	172.16.102.61	172.16.103.61	172.16.104.61	172.16.100.61	S2D-DRAC-11
Node 12	S2DNode12	172.16.102.62	172.16.103.62	172.16.104.62	172.16.100.62	S2D-DRAC-12
Node 13	S2DNode13	172.16.102.63	172.16.103.63	172.16.104.63	172.16.100.63	S2D-DRAC-13
Node 14	S2DNode14	172.16.102.64	172.16.103.64	172.16.104.64	172.16.100.64	S2D-DRAC-14
Node 15	S2DNode15	172.16.102.65	172.16.103.65	172.16.104.65	172.16.100.65	S2D-DRAC-15
Node 16	S2DNode16	172.16.102.66	172.16.103.66	172.16.104.66	172.16.100.66	S2D-DRAC-16