Microsoft HCI Solutions from Dell Technologies

Deployment Guide

D&LL Technologies

Notes, cautions, and warnings

(i) **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.**

Contents

Introduction

Topics:

- Document overview
- Audience and scope
- Known issues

Document overview

This deployment guide provides an overview of Microsoft HCI Solutions from Dell Technologies, guidance on how to integrate solution components, and instructions for preparing and deploying the solution infrastructure. Microsoft HCI Solutions from Dell Technologies includes:

- Dell EMC Integrated System for Microsoft Azure Stack HCI (based on Azure Stack HCI OS v20H2)
- Dell EMC HCI Solutions for Microsoft Windows Server (based on Windows Server 2016/2019 OS)

For end-to-end deployment steps, use the information in this guide with the information in Network Configuration and Host Configuration Options.

This guide applies only to infrastructure that is built by using the validated and certified AX nodes from Dell Technologies and Windows Server 2016, Windows Server 2019, and the Azure Stack HCI operating system.

Audience and scope

The audience for this document includes systems engineers, field consultants, partner engineering team members, and customers with knowledge about deploying hyperconverged infrastructures (HCIs) with Windows Server 2016, Windows Server 2019, the Azure Stack HCI operating system, Hyper-V, and Storage Spaces Direct. Customers who do not have Volume License agreements with Microsoft can order AX nodes from Dell Technologies with a factory-installed operating system and OEM license or as a bare-metal installation.

An Azure Stack HCI cluster can be deployed in the following ways:

- Dell Technologies Services led—Certified deployment engineers ensure accuracy, speed, and reduced risk and downtime.
- Customer led—Customers who have the qualified level of technical expertise follow the instructions in this guide.
- Using Windows Admin Center—Deployment services engineers or customers perform the Azure Stack HCI solution deployment using Windows Admin Center.

(i) **NOTE:** Instructions in this deployment guide are applicable only to:

- The generally available (GA) operating system build of Windows Server 2016 with the latest applicable updates
- The Windows Server 2019 GA build with the latest operating system updates
- The GA operating system build of the Azure Stack HCI operating system with the latest cumulative update

These instructions are not validated with Windows Server version 1709. AX nodes from Dell Technologies do not support the Windows Server Semi-Annual Channel release. Dell Technologies recommends that you update the host operating system with the latest cumulative updates from Microsoft before starting the Azure Stack HCI cluster creation and configuration tasks.

Assumptions

This guide assumes that deployment personnel understand the following concepts and tasks:

- AX nodes from Dell Technologies
- Deploying and configuring BIOS and integrated Dell Remote Access Controller (iDRAC) settings

- Deploying and configuring Windows Server core operating system Hyper-V infrastructure

Known issues

Before starting the cluster deployment, see Dell EMC Solutions for Microsoft Azure Stack HCI - Known Issues.

Solution Overview

Topics:

- Solution introduction
- Deployment models
- Solution integration and network connectivity

Solution introduction

Microsoft HCI Solutions from Dell Technologies include various configurations of AX nodes. These AX nodes power the primary compute cluster that is deployed as a HCI. The HCI uses a flexible solution architecture rather than a fixed component design.

For information about supported AX nodes and operating system support for each of the AX nodes, see the Support Matrix for Microsoft HCI Solutions.

The solutions are available in both hybrid and all-flash configurations. For more information about available configurations, see the AX nodes specification sheet.

Deployment models

Microsoft HCI Solutions from Dell Technologies offer the following types of cluster infrastructure deployments:
- Switchless storage networking
- Scalable infrastructure
- Stretched cluster infrastructure

(i) **NOTE:** This guide does not provide deployment instructions for stretched cluster infrastructure. For information about this infrastructure, see the Dell EMC Integrated System for Microsoft Azure Stack HCI: Stretched Cluster Deployment Reference Architecture Guide.

Switchless storage networking

This Microsoft HCI Solutions from Dell Technologies infrastructure type offers two to four nodes in a switchless configuration for storage traffic. This infrastructure can be implemented using any of the validated and supported AX nodes. However, the number of nodes in a cluster varies between the AX node models and the number of network adapters that each model supports.

Switchless storage networking offers two full-mesh configurations:
- Single-link
- Dual-link

For more information about these configurations, see Network Integration and Host Configuration Options.

(i) **NOTE:** For the two-node cluster deployment, configure a cluster witness. For details, see Configuring a cluster witness.

(i) **NOTE:** The switchless storage configuration has been validated with back-to-back connections for storage connectivity.

The following figure shows the switchless storage networking infrastructure:
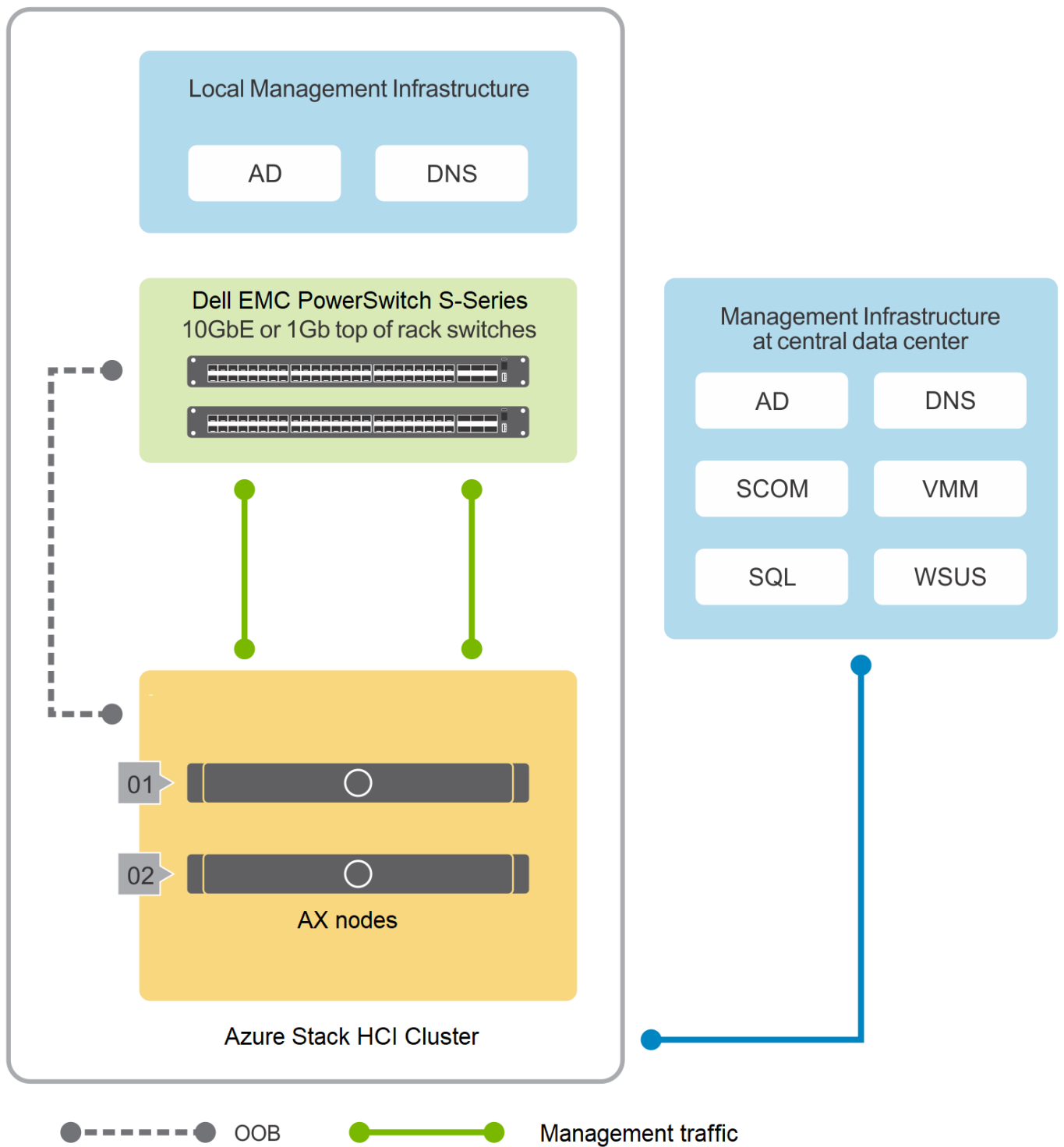
Figure 1. Switchless storage networking

Scalable infrastructure

The scalable offering within Microsoft HCI Solutions from Dell Technologies encompasses various AX node configurations. In this Windows Server HCI solution, as many as 16 AX nodes power the primary compute cluster.

The following figure illustrates one of the flexible solution architectures. It includes the Azure Stack HCI cluster, redundant top-of-rack (ToR) switches, a separate out-of-band (OOB) network, and an existing management infrastructure in the data center.
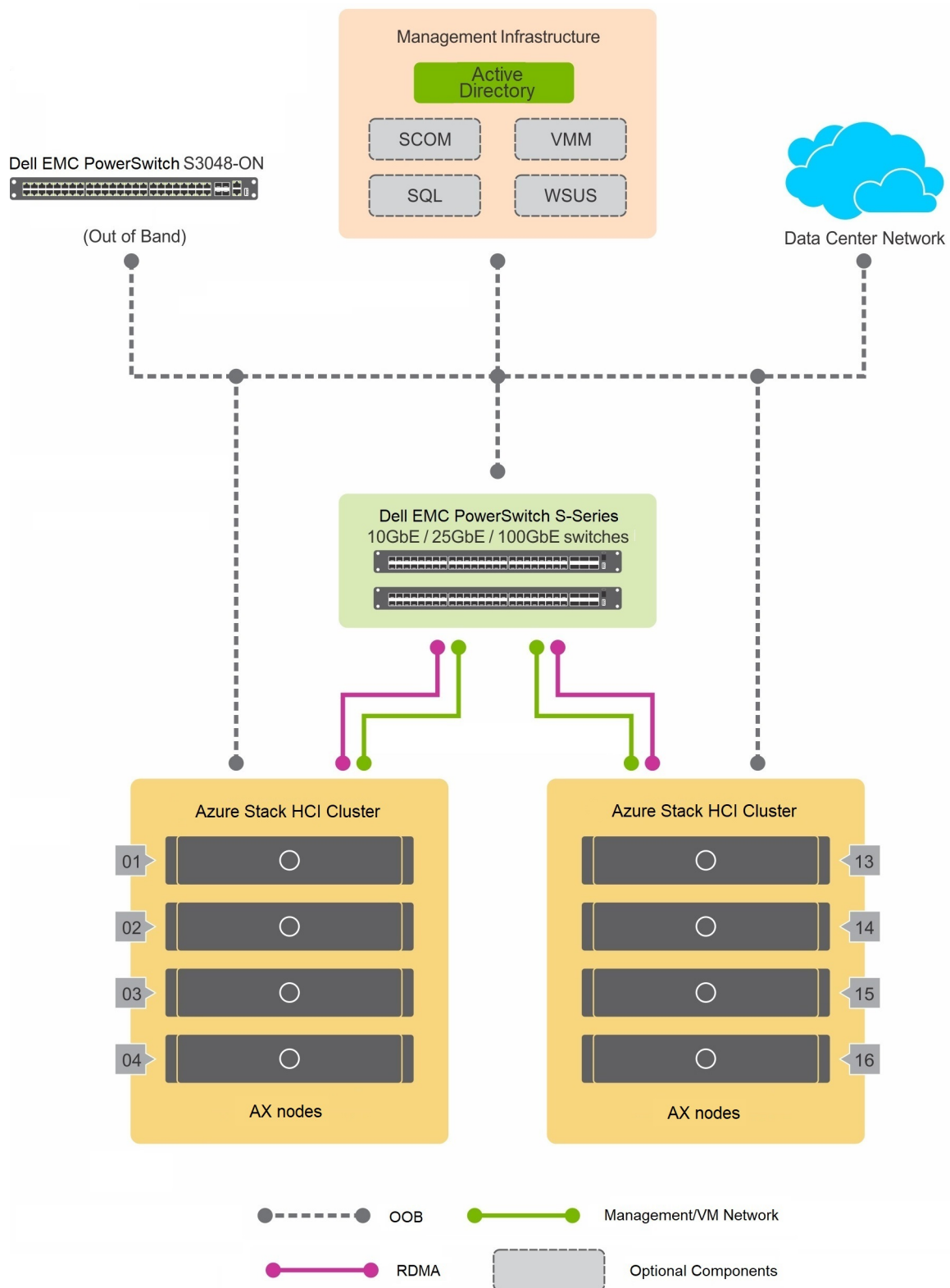
Figure 2. Scalable solution architecture

Microsoft HCI Solutions from Dell Technologies do not include management infrastructure components such as a cluster for hosting management VMs and services such as Microsoft Active Directory, Domain Name System (DNS), Windows Server Update Services (WSUS), and Microsoft System Center components such as Operations Manager (SCOM). The instructions in this guide do not include deployment of any of these services and components, and they assume that at least an Active Directory domain controller is available in the existing management infrastructure. In a remote office scenario, Dell Technologies recommends that you deploy either an Active Directory replica or read-only domain controller (RODC) at the remote office.

If you are using an RODC at the remote site, connectivity to the central management infrastructure with a writeable domain controller is mandatory during deployment of the Azure Stack HCI cluster.

(i) **NOTE:** Dell Technologies does not support expansion of a two-node cluster to a larger cluster size. A three-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 four-node cluster at a minimum.

(i) **NOTE:** For recommended server and network switch placement in the racks, port mapping on the top-of-rack (ToR) and OOB switches, and details about configuring the ToR and OOB switches, see Network Integration and Host Network Configuration Options.

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 by using System Center Virtual Machine Manager (VMM), see Preparing and Using SCVMM for Azure Stack HCI Network and Cluster Configuration.

Stretched cluster infrastructure

The Azure Stack HCI operating system added a new feature to support disaster recovery between two sites using Azure Stack HCI clusters. With Storage Replica as its foundation, stretched clusters support both synchronous and asynchronous replication of data between two sites. The replication direction (uni- or bi-directional) can be configured for either an active/passive or active/active stretched cluster configuration.

(i) **NOTE:** Stretched clustering infrastructure is supported only with the Azure Stack HCI operating system. For more information, see the Dell EMC Integrated System for Microsoft Azure Stack HCI: Stretched Cluster Deployment Reference Architecture Guide.

Solution integration and network connectivity

Each of the variants in Microsoft HCI Solutions from Dell Technologies supports a specific type of network connectivity. The type of network connectivity determines the solution integration requirements.

- For information about all possible topologies within both fully converged and nonconverged solution integration, including with switchless storage networking and host operating system network configuration, see Network Integration and Host Network Configuration Options.
- For switchless storage networking, carry out the server cabling according to the instructions detailed in Cabling Instructions.
- For sample switch configurations for these network connectivity options, see Sample Network Switch Configuration Files.

Fully converged network connectivity

In the fully converged network configuration, both storage and management/VM traffic use the same set of network adapters. The adapters are configured with Switch Embedded Teaming (SET). In this network configuration, when using RoCE, you must configure data center bridging (DCB).

The following table shows when to configure DCB based on the chosen network card and switch topology:

Table 1. DCB configuration based on network card and switch topology

Network card on node	Switch topology	
	Fully converged[a]	Nonconverged
Mellanox (RoCE)	DCB (required)	DCB (required) for storage network adapters only[b]
QLogic (iWARP)	DCB (not required)	No DCB

a. For simplicity, all converged sample switch configurations have DCB configured for both Mellanox and QLogic cards. However, for QLogic (iWARP), except for high-performance configurations such as all-NVMe, DCB is not required.

b. Updated guidance on Mellanox nonconverged topology.

(i) **NOTE:** DCB should be enabled only on storage network adapters.

Nonconverged network connectivity

In the nonconverged network configuration, storage traffic uses a dedicated set of network adapters either in a SET configuration or as physical adapters. A separate set of network adapters is used for management, VM, and other traffic classes. In this connectivity method, DCB configuration is optional for Qlogic (iWARP), but mandatory for Mellanox (RoCE) adapters.

The switchless storage networking deployment model also implements nonconverged network connectivity without the need for network switches for storage traffic.

Solution Deployment

Topics:

Introduction to solution deployment

Microsoft HCI Solutions from Dell Technologies can be deployed in either of the following ways:
- Manual operating system deployment—Begin by manually installing the operating system on AX nodes.
- Factory-installed operating system deployment—Begin with AX nodes that have factory-installed Windows Server 2019 with the Desktop Experience feature or the Azure Stack HCI operating system. The core edition is not available as a factory-installed operating system for Windows Server 2019.

Each deployment method has its own prerequisites, including configuring network switches, as described in this guide.

(i) **NOTE:** Instructions in this deployment guide are applicable only to:
- Windows Server 2016 generally available (GA) build with the latest applicable updates
- Windows Server 2019 GA build
- Azure Stack HCI operating system

AX nodes from Dell Technologies do not support the Windows Server Semi-Annual Channel release. Dell Technologies recommends that you update the host operating system with the latest cumulative updates from Microsoft before starting the cluster creation and configuration tasks.

(i) **NOTE:** Each task in this deployment guide requires running one or more PowerShell commands. Dell Technologies recommends using these commands to complete the deployment tasks because the UI might not work as expected in some scenarios. 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

Dell Technologies assumes that the management services that are required for the operating system deployment and cluster configuration are in the existing infrastructure where the Azure Stack HCI cluster is being deployed.

The following table describes the management services:

Table 2. Management services

Management service	Purpose	Required/optional
Active Directory	User authentication	Required
Domain Name System	Name resolution	Required
Windows Software Update Service (WSUS)	Local source for Windows updates	Optional
SQL Server	Database back-end for System Center VMM and System Center Operations Manager (SCOM)	Optional

Predeployment configuration

Before deploying AX nodes, complete the required predeployment configuration tasks.

Complete the following predeployment configuration before deploying the Azure Stack HCI solution.

ⓘ **NOTE:** If the cluster has persistent memory devices, preoperating-system deployment configuration as well as postoperating-system deployment configuration is required. For more information, see Appendix: Persistent Memory for Azure Stack HCI.

Configuring network switches

Based on the selected network topology, configure the ToR network switches to enable storage and VM/management traffic.

ⓘ **NOTE:** The switchless storage networking deployment model requires configuration of switches that are deployed for host management and OOB traffic only. Storage traffic uses full-mesh connectivity between the nodes.

ⓘ **NOTE:** Management network redundancy is a combination of either iDRAC or operating system DNS/IP resolution. Dell Technologies recommends that you deploy a network topology that supports a dual control plane while sharing a single data plane. Virtual Link Trunking (VLT) is Dell Technologies proprietary technology that provides network resiliency for data I/O.

Configuring switch VLT redundancy with Virtual Router Redundancy Protocol (VRRP) provides a virtual floating IP address that any node can reference as a gateway. If a switch fails, the virtual IP address is transferred to a peer switch. 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 the local destination address in the forwarding information base (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.

A standard Storage Spaces Direct deployment requires three basic types of networks—OOB management, host management, and storage. The number of network ports (two or four) that are used within the storage configuration determines whether you have two or four fault domains.

For sample switch configurations, see Sample Network Switch Configuration Files.

For configuration choices and instructions about different network topologies and host network configuration, see Network Integration and Host Network Configuration Options.

Configuring iDRAC and BIOS

The AX nodes are factory-configured with optimized system BIOS and iDRAC settings. This preconfiguration eliminates the need for you to manually configure the settings to a recommended baseline.

The iDRAC in AX 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, you must manually set a static IPv4 address on each iDRAC network interface. You can access the physical server console to set the addresses by using KVM or other means.

Configuring BIOS settings including the IPv4 address for iDRAC

Perform these steps to configure the IPv4 address for iDRAC. You can also perform these steps to configure any additional BIOS settings.

Steps

1. During the system boot, press F12.
2. At **System Setup Main Menu**, select **iDRAC Settings**.
3. Under **iDRAC Settings**, select **Network**.
4. Under **IPV4 SETTINGS**, at **Enable IPv4**, select **Enabled**.
5. Enter the static IPv4 address details.
6. Click **Back**, and then click **Finish**.

Configuring QLogic NICs

The QLogic FastLinQ 41262 network adapter supports both iWARP and RoCE.

When used with the QLogic network adapters, the AX nodes are validated only with iWARP. Manually configure the adapter to enable iWARP based on the chosen network configuration.

Configuring the QLogic NIC

Configure the QLogic network adapter for each port.

Steps

1. During system startup, press F2 to enter System Setup.
2. Click **System BIOS** and select **Device Settings**.
3. Select the QLogic network adapter from the list of adapters.
4. Click **Device Level Configuration** and ensure that **Virtualization Mode** is set to **None**.
5. Click **Back**, and then click **NIC Configuration**.
6. On the **NIC Configuration** page, select the following options:
   - Link Speed: SmartAN
   - NIC + RDMA Mode: Enabled
   - RDMA Operational Mode: iWARP
   - Boot Protocol: None
   - Virtual LAN Mode: Disabled
7. Click **Back**, and then click **Data Center Bridging (DCB) Settings**.
8. On the **Data Center Bridging (DCB) Settings** page, set **DCBX Protocol** to **Disabled**.
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.

Operating system deployment

These instructions are for manual deployment of the Windows Server 2016, Windows Server 2019, or Azure Stack HCI operating system on AX nodes from Dell Technologies. Unless otherwise specified, perform the steps on each physical node in the infrastructure that will be a part of Azure Stack HCI.

(i) **NOTE:** The steps in the subsequent sections are applicable to either the full operating system or Server Core.

> (i) **NOTE:** The command output that is shown in the subsequent sections might show only Mellanox ConnectX-4 LX adapters as physical adapters. The output is shown only as an example.

> (i) **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 that are used in the commands are shown only as examples; they might not represent the correct naming convention for the adapters in your system.

Manual operating system deployment

Dell Lifecycle Controller and iDRAC provide operating system deployment options. Options include manual installation or unattended installation by using virtual media and the operating system deployment feature in Lifecycle Controller for Windows Server 2016, Windows Server 2019, and the Azure Stack HCI operating system.

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

The remainder of this guide assumes that:
- Windows Server 2016, Windows Server 2019, or the Azure Stack HCI operating system installation on the physical server is complete.
- You have access to the iDRAC virtual console of the physical server.

> (i) **NOTE:** For information about installing the operating system using the iDRAC virtual media feature, see the "Using the Virtual Media function on iDRAC 6, 7, 8 and 9" Knowledge Base article.

> (i) **NOTE:** The Azure Stack HCI operating system is based on Server Core and does not have the full user interface. For more information about using the Server Configuration tool (Sconfig), see Deploy the Azure Stack HCI operating system.

Factory-installed operating system deployment

If the cluster nodes are shipped from Dell Technologies with a preinstalled operating system, complete the out-of-box experience (OOBE):
- Select language and locale settings.
- Accept the Microsoft and OEM EULAs.
- Set up a password for the local administrator account.
- Update the operating system partition size, and shrink it as needed.
  > (i) **NOTE:** Partition size adjustment is not available with the factory-installed Azure Stack HCI operating system.

The OEM operating system is preactivated, and the Hyper-V role is predeployed for Windows Server 2019. After completing the OOBE steps, perform the steps in Installing roles and features to complete the cluster deployment and Storage Spaces Direct configuration. For the Azure Stack HCI operating system, roles and features are preinstalled.

The Azure Stack HCI operating system factory image has multilingual support for these languages: English, German, French, Spanish, Korean, Japanese, Polish, and Italian. To change language:
- Run the following PowerShell commands. `<LANGUAGE>` can be en-US, fr-FR, ja-JP, ko-KR, de-DE, pl-PL, it-IT, or es-ES.
  - `Set-WinUserLanguageList <LANGUAGE>`
  - `Set-WinSystemLocale -systemlocale <LANGUAGE>`
- Reboot after running the commands.

> (i) **NOTE:** For all of the languages except Polish the main screen changes to the appropriate font. For Polish, the English font is used.

If you have purchased a license for a secondary operation system to run your virtual machines, the VHD file is located in the `C:\Dell_OEM\VM` folder. Copy this VHD file to a virtual disk in your Azure Stack HCI cluster and create virtual machines using this VHD.

> (i) **NOTE:** Do not run your virtual machine with the VHD file residing on the BOSS device (for example, c:\). It should always be on `c:\clusterstorage\<VD1>`

Installing roles and features

Deployment and configuration of a Windows Server 2016, Windows Server 2019, or Azure Stack HCI operating system cluster requires enabling specific operating system roles and features.

Enable the following roles and features:

- Hyper-V service (not required if the operating system is factory-installed)
- Failover clustering
- Data center bridging (DCB) (required only when implementing fully converged network topology with RoCE and when implementing DCB for the fully converged topology with iWARP)
- BitLocker (optional)
- File Server (optional)
- FS-Data-Deduplication module (optional)
- RSAT-AD-PowerShell module (optional)

Enable these features by running the `Install-WindowsFeature` PowerShell cmdlet:

```
Install-WindowsFeature -Name Hyper-V, Failover-Clustering, Data-Center-Bridging, BitLocker,
FS-FileServer, RSAT-Clustering-PowerShell, FS-Data-Deduplication -IncludeAllSubFeature
-IncludeManagementTools -verbose
```

ⓘ **NOTE:** Install the `storage-replica` feature if Azure Stack HCI operating system is being deployed for a stretched cluster.

ⓘ **NOTE:** Hyper-V and the optional roles installation require a system restart. Because subsequent procedures also require a restart, the required restarts are combined into one (see the Note in the "Changing the hostname" section).

Verifying firmware and software compliance with the support matrix

Microsoft HCI Solutions from Dell Technologies are validated and certified with certain firmware versions that are related to the solution infrastructure components.

Use the validated firmware and software, as specified in the support matrix, to ensure that the solution infrastructure remains supported and delivers optimal performance.

You can verify compliance and update the nodes with an Azure Stack HCI online or offline catalog by using Dell EMC OpenManage Integration with Windows Admin Center. To verify compliance and to update firmware and drivers on a stand-alone node, see the Dell EMC HCI Solutions for Microsoft Windows Server—Managing and Monitoring the Solution Infrastructure Life Cycle.

Updating out-of-box drivers

For certain system components, you might have to update the driver to the latest Dell Technologies supported version.

About this task

ⓘ **NOTE:** This section is optional if you are using OpenManage Integration with Windows Admin Center to update the nodes.

Steps

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

```
Get-PnpDevice | Select-Object Name, @{l='DriverVersion';e={(Get-PnpDeviceProperty -
InstanceId $_.InstanceId -KeyName 'DEVPKEY_Device_DriverVersion').Data}} -Unique |
Where-Object {($_.Name -like "*HBA*") -or ($_.Name -like "*mellanox*") -or ($_.Name
-like "*Qlogic*") -or ($_.Name -like "*X710*") -or ($_.Name -like "*intel*") -or
```

```
($_.Name -like "*Broadcom*")  -or
($_.Name -like "*marvell*") }
```

2. Update the out-of-box drivers to the required versions, if necessary.

   For the latest Dell Technologies supported versions of system components, see the Support Matrix for Microsoft HCI Solutions.

   Download the driver installers from https://www.dell.com/support or by using the Dell EMC Azure Stack HCI Solution Catalog.

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

3. Attach a folder containing the driver DUP files to the system as a virtual media image:
   a. In the iDRAC virtual console menu, click **Virtual Media**.
   b. In the **Create Image from Folder** window, click **Create Image**.
   c. Click **Browse**, select the folder where the driver DUP files are stored, and, if required, change the name of the image.
   d. Click **Create Image**.
   e. Click **Finish**.
   f. From the **Virtual Media** menu, select **Connect Virtual Media**.
   g. Select **Map Removable Disk**, click **Browse**, and select the image that you created.
   h. Click **Map Device**.

   After the image is mapped, it appears as a drive in the host operating system.

4. Go to the driver DUP files and run them to install the updated out-of-box drivers.

Changing the hostname

By default, the operating system deployment assigns a random name as the host computer name. For easier identification and uniform configuration, Dell Technologies recommends that you change the hostname to something that is relevant and easily identifiable.

Change the hostname by using the `Rename-Computer` cmdlet:

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

(i) **NOTE:** This command induces an automatic restart at the end of rename operation.

Configuring host networking

Configure Microsoft HCI Solutions from Dell Technologies to implement a fully converged or nonconverged network for storage and management connectivity.

Complete the following steps to configure host networking:

1. Create VM switches (based on topology).
2. Create VM adapters, and configure VLANs and IP addresses.
3. Map the VM storage adapters (based on topology).
4. Enable RDMA for the storage adapters.
5. Change RDMA settings on the QLogic NICs.
6. Configure the QoS policy.
7. Disable the DCBX Willing state in the operating system.

(i) **NOTE:** Dell Technologies recommends implementing a nonconverged network and using physical network adapters for the storage traffic rather than using SET. However, in a nonconverged configuration, if virtual machine adapters must have RDMA capability, SET configuration is necessary for storage adapters.

For more information about each of the preceding steps and all possible topologies within both fully converged and nonconverged solution integration (including switchless storage networking solution integration) and host operating system network configuration, see Network Integration and Host Network Configuration Options.

NOTE: The host operating system network configuration must be complete before you join cluster nodes to the Active Directory domain.

Joining cluster nodes to an Active Directory domain

Before you can create a cluster, the cluster nodes must be a part of an Active Directory domain.

NOTE: Connecting to Active Directory Domain Services by using the host management network might require routing to the Active Directory network. Ensure that this routing is in place before joining cluster nodes to the domain.

You can perform the domain join task by running the `Add-Computer` cmdlet on each host that will be a part of the Azure Stack HCI cluster.

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

```
$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.

Deploying and configuring a host cluster

After joining the cluster nodes to an Active Directory domain, you can create a host cluster and configure it for Storage Spaces Direct.

Creating the host cluster

Verify that the nodes are ready for cluster creation, and then create the host cluster.

Steps

1. Run the `Test-Cluster` cmdlet:

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

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

2. Run the `Get-PhysicalDisk` command on all cluster nodes.

   Verify the output to ensure that all disks are in the healthy state and that the nodes have an equal number of disks. Verify that the nodes have homogenous hardware configuration.

3. Run the `New-Cluster` cmdlet to create the host cluster.

   NOTE: For the `-IgnoreNetwork` parameter, specify all storage network subnets as arguments. Switchless configuration requires that all storage network subnets are provided as arguments to the `-IgnoreNetwork` parameter.

   ```
   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.

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

Enabling Storage Spaces Direct

After you create the cluster, run the `Enable-ClusterS2D` cmdlet to configure Storage Spaces Direct on the cluster. Do not run the cmdlet in a remote session; instead, use the local console session.

Run the `Enable-ClusterS2d` cmdlet as follows:

`Enable-ClusterS2D -Verbose`

The `Enable-ClusterS2D` cmdlet generates an HTML report of all configurations and includes a validation summary. Review this report, which is typically stored in the local temporary folder on the node where the cmdlet was run. The verbose output of the command shows the path to the cluster report. At the end of the operation, the cmdlet discovers and claims all the available disks into an auto-created storage pool. Verify the cluster creation by running any of the following commands:

`Get-ClusterS2D`

`Get-StoragePool`

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

Configuring the host management network as a lower-priority network for live migration

After you create the cluster, live migration is configured by default to use all available networks.

During normal operations, using the host management network for live migration traffic might impede the overall cluster role functionality and availability. Rather than disabling live migration traffic on the host management network, configure the host management network as a lower-priority network in the live migration network order:

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

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

`$otherNetworkID = (Get-ClusterNetwork).Where({$_.ID -ne $hostnetworkID}).ID`

`$newMigrationOrder = ($otherNetworkID + $hostNetworkID) -join ';'`

`Set-ClusterParameter -InputObject $clusterResourceType -Name MigrationNetworkOrder -Value $newMigrationOrder`

Updating the page file settings

To help ensure that the active memory dump is captured if a fatal system error occurs, allocate sufficient space for the page file. Dell Technologies recommends allocating at least 50 GB plus the size of the CSV block cache.

About this task

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

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

   (i) **NOTE:** On Windows Server 2016, the default block cache size is 0. On Windows Server 2019 and the Azure Stack HCI operating system, the default block cache size is 1 GB.

2. Run the following command to update the page file settings:

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

$pageFilePath = "C:\pagefile.sys"
$initialSize = [Math]::Round(51200 + $blockCacheMB)
$maximumSize = [Math]::Round(51200 + $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}
}
```

Configuring a cluster witness

A cluster witness must be configured for a two-node cluster. Microsoft recommends configuring a cluster witness for a four-node Azure Stack HCI cluster. Cluster witness configuration helps maintain a cluster or storage quorum when a node or network communication fails and 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.

(i) **NOTE:** If you choose to configure a file share witness, ensure that it is outside the two-node cluster.

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

Azure onboarding for Azure Stack HCI operating system

Clusters deployed using Azure Stack HCI operating system must be onboarded to Microsoft Azure for full functionality and support. For more information, see Connect Azure Stack HCI to Azure.

Best practices and recommendations

Dell Technologies recommends that you follow the guidelines that are described in this section.

Disable SMB Signing

Storage Spaces Direct uses RDMA for SMB (storage) traffic for improved performance. When SMB signing is enabled the network performance of SMB traffic is significantly reduced.

For more information, see Reduced networking performance after you enable SMB Encryption or SMB Signing in Windows Server 2016.

(i) **NOTE:** By default, SMB Signing is disabled. If SMB Signing is enabled in the environment through a Group Policy Object (GPO), you must disable it from the domain controller.

Update the hardware timeout for the Spaces port

For performance optimization and reliability, update the hardware timeout configuration for the Spaces port.

The following PowerShell command updates the configuration in the Windows registry and induces a restart of the node at the end of the registry update. Perform this update on all Storage Spaces Direct nodes immediately after initial deployment. Update one node at a time and wait until each node rejoins the cluster.

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

Enable jumbo frames

Enabling jumbo frames specifically on the interfaces supporting the storage network might help improve the overall read/write performance of the Azure Stack HCI cluster. An end-to-end configuration of jumbo frames is required to take advantage of this technology. However, support for jumbo frame sizes varies among software, NIC, and switch vendors. The lowest value within the data path determines the maximum frame size that is used for that path.

For the storage network adapters in the host operating system, 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 operating system and network switch configuration.

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

Recommended next steps

Before you proceed with operational management of the cluster, Dell Technologies recommends that you validate the cluster deployment, verify that the infrastructure is operational, and, if needed, activate the operating system license.

1. Run the `test-Cluster` cmdlet to generate a cluster validation report:

   ```
   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.

2. If the operating system was not factory-installed, activate the operating system license.

   By default, the operating system is installed in evaluation mode. For information about activating the operating system license as well as the management and operational aspects for the Azure Stack HCI solution, see the Operations Guide—Managing and Monitoring the Solution Infrastructure Life Cycle.

Deployment services

Dell Technologies recommends using the company's deployment services to install Microsoft HCI Solutions from Dell Technologies. Issues that arise during do-it-yourself installations and configuration are not covered even if you have purchased Dell Technologies ProSupport or ProSupport Plus. Support for installation and configuration issues is provided under a separate paid services package.

When you call Dell Technologies with an installation and configuration issue, Dell Tech Support routes you to your Sales Account Manager. The Account Manager will then help you to purchase the onsite deployment services package.

References

Topics:

- Dell Technologies documentation
- Microsoft documentation

Dell Technologies documentation

These links provide more information from Dell Technologies:

- iDRAC documentation
- Support Matrix for Microsoft HCI Solutions
- Operations Guide—Managing and Monitoring the Solution Infrastructure Life Cycle

Microsoft documentation

The following link provides more information about Storage Spaces Direct:

Storage Spaces Direct overview

Appendix A: Persistent Memory for Windows Server HCI

Topics:

Configuring persistent memory for Windows Server HCI

Intel Optane DC persistent memory is designed to improve overall data center system performance and lower storage latencies by placing storage data closer to the processor on nonvolatile media. The persistent memory modules are compatible with DDR4 sockets and can exist on the same platform with conventional DDR4 DRAM DIMMs. In App Direct Mode, the operating system distinguishes between the RAM and persistent storage memory spaces.

Intel Optane DC persistent memory provides an ideal capacity to be used as a cache device (SBL) for Microsoft Storage Spaces Direct. Storage data is interleaved between the Intel Optane DC persistent memory DIMMs within each CPU socket to achieve the best performance. A single region per CPU socket is configured in the system BIOS. Thus, a dual CPU server platform provides Storage Spaces Direct with two persistent memory cache devices. These high-endurance, write-intensive cache devices can be used to enhance the performance of many slower-performing NVMe/SAS/SATA devices that are used for storage capacity, as shown in the following figure:
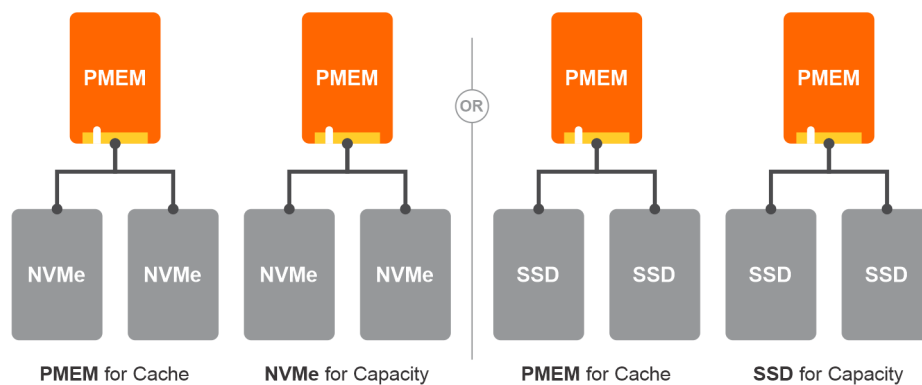

Figure 3. Persistent memory

Persistent memory requirements

Persistent memory requirements for Microsoft HCI Solutions from Dell Technologies:

*   AX-640 nodes
*   2 x Intel Xeon Cascade Lake-SP Gold or Platinum CPUs (models 52xx, 62xx, or 82xx) per server
*   12 x 32 GB RDIMMs in DIMM slots A1-A6 and B1-B6 (white slots) per server, totaling 384 GB of RAM per server
*   12 x 128 GB Intel Optane DC Persistent DIMMs in DIMM slots A7–A12 and B7–B12 (black slots) per server, totaling 2 x 768 GB cache devices per server
*   Windows Server 2019 Datacenter

Configuring persistent memory BIOS settings

Configure the BIOS to enable persistent memory.

Steps

1. During system startup, press F12 to enter **System BIOS**.
2. Select **BIOS Settings** > **Memory Settings** > **Persistent Memory**.
3. Verify that **System Memory** is set to **Non-Volatile DIMM**.
4. Select **Intel Persistent Memory**.
   The **Intel Persistent Memory** page provides an overview of the server's Intel Optane DC persistent memory capacity and configuration.
5. Select **Region Configuration**.
   To be used as a Storage Spaces Direct cache device, Intel Persistent Memory must be configured in App Direct Interleaved mode. App Direct Interleaved Mode creates two regions—one region for each CPU socket, as shown in the following figure:
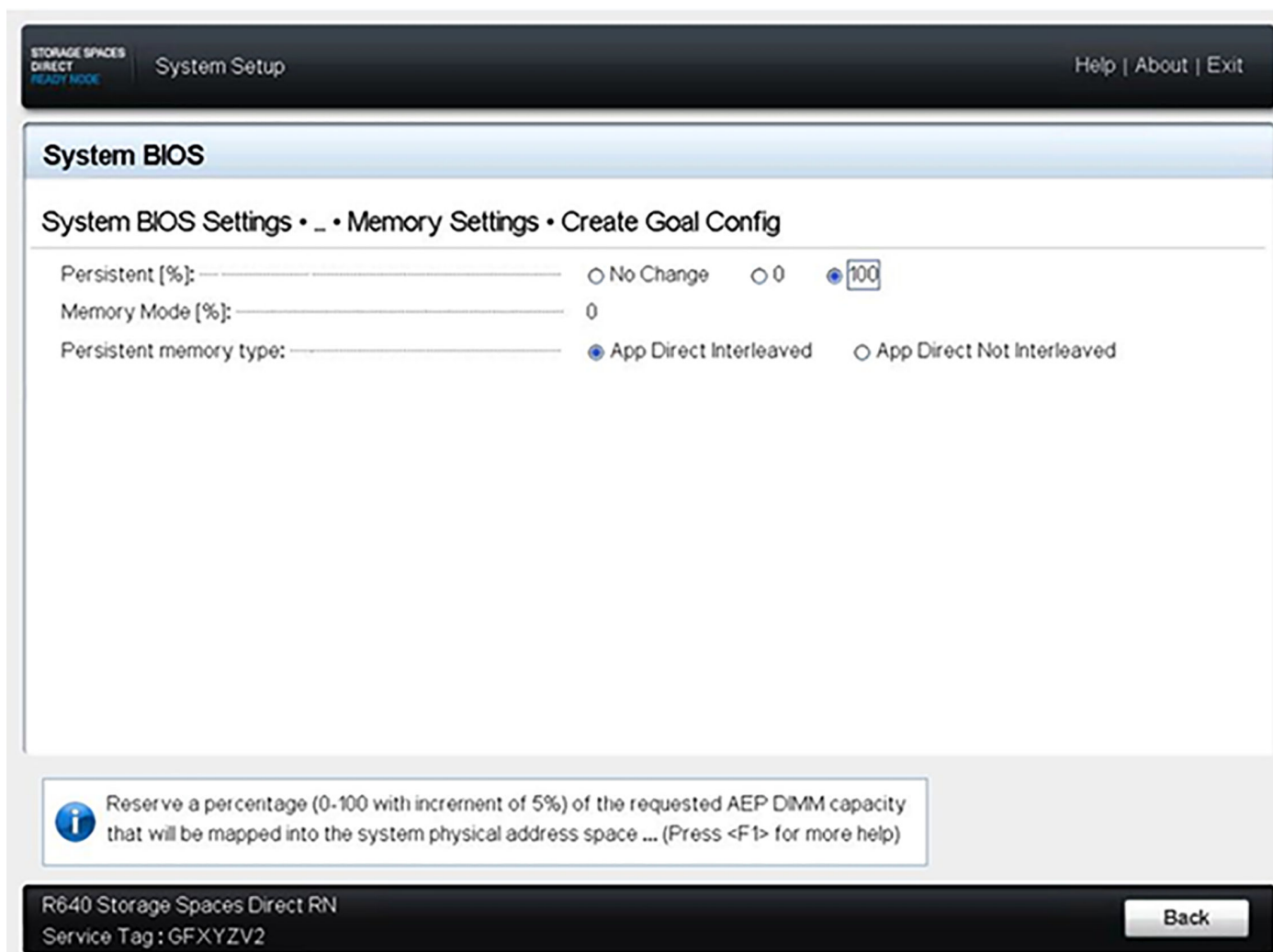

Figure 4. Persistent memory region configuration

6. If the configuration is missing or incorrect, select **Create goal config** to reconfigure the persistent memory regions.

   CAUTION: Performing these steps erases all previous persistent memory regions.

   a. In **Create Goal Config**, for **Persistent (%)**, select **100**.
   b. For **Persistent memory type**, select **App Direct Interleaved**.
      A warning is displayed. All Intel Persistent Memory data is erased when changes are saved to the BIOS configuration.

7. Exit the BIOS and save the configuration.

Configuring Windows Server HCI persistent memory hosts

Three types of device objects are related to persistent memory on Windows Server 2019: the NVDIMM root device, physical INVDIMMs, and logical persistent memory disks. In Device Manager, physical INVDIMMs are displayed under **Memory devices**, while logical persistent disks are under **Persistent memory disks**. The NVDIMM root device is under **System Devices**. The `scmbus.sys` driver controls the NVDIMM root device.

The `nvdimm.sys` driver controls all NVDIMM devices, while the `pmem.sys` driver controls the logical disks. Both the nvdimm.sys and pmem.sys drivers are the same for all types of persistent memory, such as NVDIMM-N and Intel Optane DC Persistent Memory (INVDIMM).

The following figure shows the Device Manager on a system with 12 INVDIMMs across dual CPU sockets and two persistent storage disks:
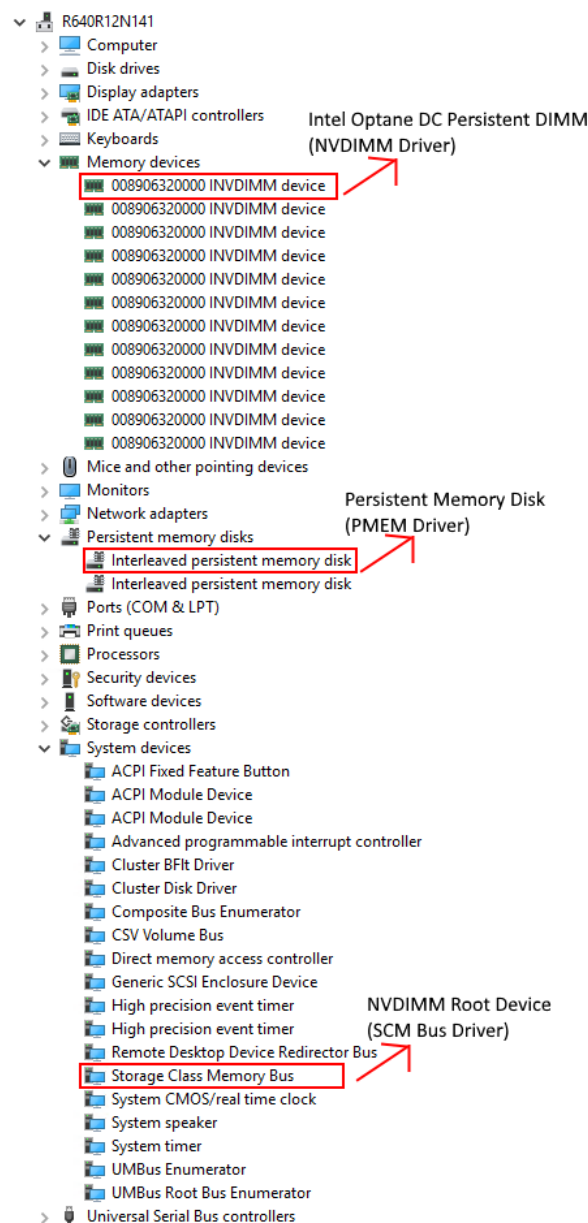

Figure 5. Device Manager example

Managing persistent memory using Windows PowerShell

Windows Server 2019 provides a `PersistentMemory` PowerShell module that enables user management of the persistent storage space.

```
PS C:\> Get-Command -Module PersistentMemory
CommandType     Name                        Version     Source
-----------     ----                        -------     ------
Cmdlet          Get-PmemDisk                1.0.0.0     PersistentMemory
Cmdlet          Get-PmemPhysicalDevice      1.0.0.0     PersistentMemory
Cmdlet          Get-PmemUnusedRegion        1.0.0.0     PersistentMemory
Cmdlet          Initialize-PmemPhysicalDevice   1.0.0.0     PersistentMemory
Cmdlet          New-PmemDisk                1.0.0.0     PersistentMemory
Cmdlet          Remove-PmemDisk             1.0.0.0     PersistentMemory
```

- `Get-PmemDisk`—Returns one or more logical persistent memory disks that were created by `New-PmemDisk`. The returned object includes information about size, health status, and the underlying physical NVDIMM devices.
- `Get-PmemPhysicalDevice`—Returns one or more physical persistent memory NVDIMM devices. The returned object includes information about size, firmware, physical location, and health status. In App Direct Interleaved mode, each INVDIMM device displays its full capacity as `Persistent memory size`.

  (i) **NOTE:** The Intel Optane DC Persistent Memory firmware might have to be periodically updated. For the supported firmware version, see the Support Matrix for Microsoft HCI Solutions. After identifying the required firmware version, download the Intel Optane DC Persistent Memory Firmware Package from Dell Technologies Support or use the Microsoft HCI Solutions from Dell Technologies Update Catalog.

- `Get-PmemUnusedRegion`—Returns aggregate persistent memory (Pmem) regions that are available for provisioning a logical device. The returned object has a unique region ID, total size, and list of physical devices that contribute to the unused region.
- `Initialize-PmemPhysicalDevice`—Writes zeroes to the label storage area, writes new label index blocks, and then rebuilds the storage class memory (SCM) stacks to reflect the changes. This cmdlet is intended as a recovery mechanism and is not recommended for normal use.
- `New-PmemDisk`—Creates a disk out of a given unused region. This cmdlet writes out the labels to create the namespace, and then rebuilds the SCM stacks to expose the new logical device. The new logical persistent disk is added in Device Manager under `Persistent memory disks`. `Get-PhysicalDisk` displays the storage device as MediaType `SCM`.
- `Remove-PmemDisk`—Removes the given persistent memory disk. This cmdlet accepts the output of `Get-PmemDisk`. It deletes the namespace labels and then rebuilds the SCM stacks to remove the logical device.

Configuring persistent memory as SCM logical devices

On each server node, verify unused persistent memory regions and configure them as new SCM logical devices:

1. Run `Get-PmemPhysicalDevice` to verify that 12 INVDIMM physical devices are available and healthy:

```
PS C:\> Get-PmemPhysicalDevice
DeviceId    DeviceType              HealthStatus    OperationalStatus PhysicalLocation
FirmwareRevision Persistent memory size Volatile memory size
--------    ----------              ------------    -----------------
----------------    ----------------    --------------------    --------------------
1    008906320000 INVDIMM device    Healthy    {Ok}    A7    102005395    126 GB
   0 GB
1001    008906320000 INVDIMM device    Healthy    {Ok}    B7    102005395    126 GB
      0 GB
101    008906320000 INVDIMM device    Healthy    {Ok}    A10    102005395    126 GB
      0 GB
1011    008906320000 INVDIMM device    Healthy    {Ok}    B8    102005395    126 GB
      0 GB
1021    008906320000 INVDIMM device    Healthy    {Ok}    B9    102005395    126 GB
      0 GB
11    008906320000 INVDIMM device    Healthy    {Ok}    A8    102005395    126 GB
   0 GB
1101    008906320000 INVDIMM device    Healthy    {Ok}    B10    102005395    126 GB
      0 GB
111    008906320000 INVDIMM device    Healthy    {Ok}    A11    102005395    126 GB
      0 GB
```

```
1111    008906320000 INVDIMM device    Healthy    {Ok}    B11    102005395    126 GB
      0 GB
1121    008906320000 INVDIMM device    Healthy    {Ok}    B12    102005395    126 GB
      0 GB
121    008906320000 INVDIMM device    Healthy    {Ok}    A12    102005395    126 GB
      0 GB
21    008906320000 INVDIMM device    Healthy    {Ok}    A9    102005395    126 GB
    0 GB
```

2. Run `Get-PmemUnusedRegion` to verify that two unused Pmem regions are available, one region for each physical CPU:

```
PS C:\> Get-PmemUnusedRegion

RegionId    TotalSizeInBytes    DeviceId
--------         ---------------    --------
      1         811748818944       {1, 111, 21, 101...}
      3         811748818944       {1001, 1111, 1021, 1101...}
```

3. Run the `Get-PmemUnusedRegion | New-PmemDisk` script to create two Pmem disks, one for each Pmem region:

```
PS C:\> Get-PmemUnusedRegion | New-PmemDisk
Creating new persistent memory disk. This may take a few moments.
Creating new persistent memory disk. This may take a few moments.
```

4. Run the `Get-PhysicalDisk | ?{$_.MediaType -eq "SCM"}` script to verify that both Pmem disks are available as physical disk devices:

```
PS C:\> Get-PmemDisk
DiskNumber    Size    HealthStatus    AtomicityType    CanBeRemoved
PhysicalDeviceIds
----------         ----    -----------    -------------    ------------
----------------
11         756 GB    Healthy    None    True    {1, 111, 21, 101...}
12         756 GB    Healthy    None    True    {1001, 1111, 1021, 1101...}
```

When you run the `Enable-ClusterS2D` command to enable Storage Spaces Direct, the SCM logical devices are automatically detected and used as cache for NVMe and SSD capacity devices.