Dell EMC Host Connectivity Guide for VMware ESXi Server

Regulatory Model: P/N 300-002-304
July 2020
Rev. 53

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.

Contents

Figures

Tables

PREFACE

As part of an effort to improve its product lines, Dell EMC periodically releases revisions of its software and hardware. Therefore, some functions that are described in this document might not be supported by all versions of the software or hardware currently in use. The product release notes provide the most up-to-date information on product features.

Contact a technical support professional when a product does not function correctly or does not function as described in this document.

NOTE: This document was accurate at publication time. To find the latest version of this document, go to Dell EMC Online Support (https://www.dell.com/support).

Purpose

This guide describes the features and setup procedures for VMware ESXi Server host interfaces to Dell EMC PowerMax, VMAX All Flash, VPLEX, XtremIO, VNX, and CLARiiON storage systems. This document is meant to assist in the installation and configuration of VMware ESXi Server attached to Dell EMC PowerMax, VMAX, VPLEX, XtremIO, VNX, and CLARiiON systems.

Audience

This guide is intended for use by storage administrators, system programmers, or operators who are involved in acquiring, managing, or operating PowerMax, VMAX series, VPLEX, XtremIO, VNX, and CLARiiON, and host devices.

Readers of this guide are expected to be familiar with the following topics:

- Dell EMC PowerMAx, VMAX series, VPLEX, XtremIO, VNX, and CLARiiON system operation
- VMware ESXi Server operating environment

Related documentation

The following Dell EMC publications provide additional information:

- For the most up-to-date information for supported server and HBA combinations, always see the Dell EMC Simple Support Matrix available on Dell EMC E-Lab Navigator.
- For VMware-specific documentation, such as the *VMware ESXi Server Release Notes*, *ESXi Server Administration Guide*, and the *ESXi Server Installation Guide*, go to https://www.vmware.com/support.html.
- For a list of supported guest operating systems, see the *VMware Guest Operating System Installation Guide*, located at http://partnerweb.vmware.com/GOSIG/home.html.

Typographical conventions

Table 1. Typographical conventions

Bold	Used for names of interface elements, such as names of windows, dialog boxes, buttons, fields, tab names, key names, and menu paths (what the user specifically selects or clicks)
Italic	Used for full titles of publications that are referenced in text
`Monospace`	Used for: • System code • System output, such as an error message or script • Pathnames, filenames, prompts, and syntax • Commands and options
Monospace italic	Used for variables
`Monospace bold`	Used for user input
[ ]	Square brackets enclose optional values
|	Vertical bar indicates alternate selections - the bar means "or"

Table 1. Typographical conventions (continued)

{ }	Braces enclose content that the user must specify, such as x or y or z
...	Ellipses indicate nonessential information that is omitted from the example

Where to get help

The Dell EMC Online Support page provides access to licensing information, product documentation, advisories, and downloads, as well as how-to and troubleshooting information. This information may resolve a product issue before contacting Customer Support.

To access the Dell EMC Online Support page:

1. Go to https://www.dell.com/support.
2. Type a product name in the **Enter a Service Tag, Serial Number, Service Request, Model, or Keyword** search box.
3. Select the product from the list that appears. When you select a product, the **Product Support** page loads automatically.
4. (Optional) Add the product to the **My Products** list by clicking **Add to My Saved Products** in the upper right corner of the **Product Support** page.

Product information

For documentation, release notes, software updates, or information about Dell EMC products, go to Dell EMC Online Support (https://www.dell.com/support).

Comments and suggestions

Your comments and suggestions will help to continue to improve the accuracy, organization, and overall quality of the user publications. Send comments and suggestions about this document to techpubcomments@dell.com.

1

Introduction to VMware Infrastructure

Topics:

VMware vSphere

This chapter provides information about the VMware infrastructure vSphere, including:

- vSphere 6.5
- vSphere 6.7
- vSphere 7.0

vSphere 6.5

VMware vSphere 6.5 supports 512 LUNs and 2,048 paths, and improves the storage infrastructure scalability for customers.

VMware vSphere 6.5 also supports the automated UNMAP process by which Virtual Machine File System (VMFS) tracks the deleted blocks and reclaims deleted space from the backend array in the background for minimal storage I/O impact. UNMAP works at a guest operating system level with newer versions of Windows and Linux.

vSphere 6.7

VMware vSphere 6.7 supports up to 1024 LUNs and 4096 paths per host. At the VM level, the number of vDisks has increased from 16 to 64 disks. This equates to 256 disks using PVSCSI adapters. This enables the customer to have larger guest clusters and reduce the overhead of managing multiple smaller clusters.

vSphere 6.7 also supports XCOPY to be configured with specific parameters to optimize the storage array's XCOPY operations. It has been implemented in two stages; support for arrays using the standard SCSI T10 plug-in (VMW_VAAIP_T10) and support for vendor-specific plugins (vmkAPI).

vSphere 6.7 also adds new features to enhance the functionality of Virtual Volumes (vVols) such as end-to-end support for IPv6 and SCSI-3 Persistent Group Reservations support.

vSphere 7.0

In vSphere 7.0, 32-bit userworld support has been deprecated and provides 64-bit userworld support through partner devkits and will retain 32-bit userworld support through this major release. Support for 32-bit userworlds will be permanently removed in the next major ESXi release. To avoid loss of functionality, ensure that any vendor-supplied VIBs such as PowerPath/VE and VAAI-NAS Plugin are migrated to 64-bit version before upgrading beyond the vSphere 7.0 release.

The vSphere 7.0 storage stack added support for FC-NVMe and NVMe-oF over Ethernet using RoCE v2 RDMA, which is the first release to cooperate with NVMe supported external storage system.

If the array supports Write Exclusive-All Registrants (WEAR) type SCSI 3 Persistent Reservation can claim support for Clustered VMDK feature in 7.0, which means customer can use either RDM device or virtual disk as quorum disk for building up VM-based cluster.

VMware ESXi Server

VMware ESXi Server is the main building block of the VMware infrastructure.

The VMware ESXi Server provides a platform for multiple virtual machines sharing hardware resources. These resources include processors, memory, storage, and networking resources. It lets virtual machines perform all the functions of a physical machine. The VMware ESXi server maximizes hardware utilization efficiency while minimizing installation capital and operating cost.

VMware ESXi Server consists of two main components:

- VMkernel
- VMware ESXi Server utilities and functions

The interaction between the two components forms a dynamic and reliable virtualized environment, providing virtual machines with high availability, resource management, operational automation, and security features that improve service quality levels even to the most resource-intensive mission-critical applications.

VMkernel

VMkernel is the ESXi Server virtualized layer that runs on bare metal. It handles CPU and memory directly without going through a third-party operating system. VMkernel uses Scan-Before-Execution (SBE) to handle special or privileged CPU instructions.

To access other hardware, including network and storage devices, VMkernel modules are used. Some of the modules are derived from the Linux kernel modules.

VMkernel can provide services including CPU scheduling, memory management, and virtual switch data processing for virtual machines to access the underlying physical hardware of the server on which they are built.

VMkernel manages all the operating systems on the machine, including both the service console and the guest operating systems running on each virtual machine.

VMkernel interfaces with three major components: hardware, guest operating system, and the service console.

VMware ESXi Server utilities and functions

The following table describes the useful utilities and functions of ESXi 6 and later versions:

Table 2. Utilities and functions of VMware ESXi Server

Utility/Function	Description
vmkfstools	Command to create and manipulate virtual drives, file systems, logical volumes, and physical storage devices on an ESXi host.
vmware-cmd	All vCLI 4.1 commands have been renamed. Significant additions have been made to ESXiCLI. Many tasks that are previously performed with a vicfg- command are now performed with ESXiCLI.
resxtop	Monitors how ESXi hosts use resources in real time. Runs in interactive or batch mode.

Control interface

This section describes the following:

- VMware Web UI
- VMware vSphere Web Client
- VMware vCenter Server

VMware web UI

VMware Web UI is a free option allowing administrators to monitor and manage the server remotely through a web-based graphical interface by typing the host IP in a browser and logging in as administrator.

One of the main disadvantages of VMware web UI, compared to vSphere client and vSphere web client is that you can manage only one ESXi server at a time.

VMware vSphere Web Client

The VMware vSphere Web Client is a web browser-based application that user can use to connect to vCenter Server systems and manage your vSphere infrastructure. It is the primary method for system administrators and users to interact with the virtual data center environment created by vSphere.

VMware vCenter Server

VMware vCenter Server is a centralized management application that enables you manage virtual machines and ESXi hosts centrally. vSphere Client is used to access vCenter Server and ultimately manage ESXi Servers.

Generally, a new version of vCenter Server is compatible with previous ESXi versions, while it is not valid the opposite way. For details about ESXi, vCenter Server, and vSphere Client version compatibility, see the vSphere Compatibility Matrices available at the VMware website.

2

Connectivity

Topics:

Fibre Channel

Fibre Channel (FC) is a Gb speed network technology. A Fibre Channel Storage Area Network (FC SAN) is a collection of FC nodes that communicate with each other, typically through fiber optic media.

Node - A node is defined as a member of the FC network. A node is provided as physical and logical connection to the network by a physical port on an FC switch. Every node requires the use of specific drivers to access the network.

Fabric switches - FC nodes communicate with each other through one or more FC switches, also called fabric switches. The primary function of a fabric switch is to provide a physical connection and logical routing of data frames between the attached devices.

Fabric zoning

Use a single-initiator zoning or a single-initiator-single-target zoning with ESXi hosts. Single-initiator-single-target is the preferred zoning practice.

NOTE: Dell EMC does not support multiinitiator zones in a VMware ESXi Server fabric environment.

Design the zoning on the fabric by creating zone sets that contain the initiator and the targets.

The World Wide Port Names (WWPNs) and World Wide Names (WWN) of the array ports are bolded. See the following individual switch vendor documents for detail zone method:

- Cisco
- Brocade

iSCSI

The Internet Small Computer Systems Interface (iSCSI) protocol enables the transport of SCSI blocks through Transmission Control Protocol/Internet Protocol (TCP/IP) network.

iSCSI works by encapsulating SCSI commands into TCP packets and sending it over an IP network. An example is shown in the following figure:

Figure 1. SCSI commands encapsulated by Ethernet headers

iSCSI is IP-based traffic and can be routed or switched using standard (100 Mb/s, 1G, 10G, 25G, 40G, and 100G) Ethernet equipment.

Traditional Ethernet adapters (NICs) are designed to transfer packetized file level data among PCs, servers, and storage devices, such as NAS appliances.

For NIC to process block level data, the data needs to be placed into a TCP/IP packet before sending over the IP network. This block level data packet creation and TCP/ IP processing is done using iSCSI drivers. This process, which is known as software iSCSI (SW iSCSI), is

extremely CPU intensive and lowers the overall server performance. For more information about SW iSCSI, see the VMware ESXi SW iSCSI section.

The TCP/IP processing performance bottleneck has been the driving force behind the development of TCP/IP Offload Engines (TOE) on adapter cards. A TOE removes the TCP/IP processing from the host CPU and completes TCP/IP processing and packet creation on the HBA. Thus, a TCP/IP offload storage NIC operates like a storage HBA rather than a standard NIC. This process is also known as Hardware iSCSI (HW iSCSI).

NOTE: See the Dell EMC Online Support for supported HW iSCSI initiators.

Always ensure that the hardware iSCSI initiators are successfully installed and recognized by the system.

VMware ESXi SW iSCSI

This section describes SW iSCSI initiator ports configuration.

NOTE: Isolate iSCSI traffic from other network traffic.

ESXi SW iSCSI supports the following:

- Supports both send and static targets discovery
- Requires only VMkernel port to be in the network configuration
- Supports multipathing
- Supports NIC-teaming
- Supports ALUA, failover mode 4 for Unity, VNX series
- Supports Dell EMC PowerPath/VE

NOTE: For more information about Dell EMC PowerPath/VE, see the Dell EMC Website.

Setting the SW iSCSI

See VMware documentation for version-specific configurations. The procedures may vary depending on the ESXi version you are using.

- VMware vSphere 7.0
- VMware vSphere 6.7
- VMware vSphere 6.5

NOTE: Once the iSCSI initiator ports on the ESXi Server are configured, iSCSI storage must be presented to the ESXi Server. See the latest *Dell EMC Simple Support Matrix* on Dell EMC E-Lab Navigator for the up-to-date information about which Dell EMC arrays that are supported by iSCSI attached to VMware ESXi Server.

Network configurations for ESXi

In a two or more NICs environments, you can set up SW iSCSI using a single vSwitch or dual vSwitch network configuration, as shown in the following figure:

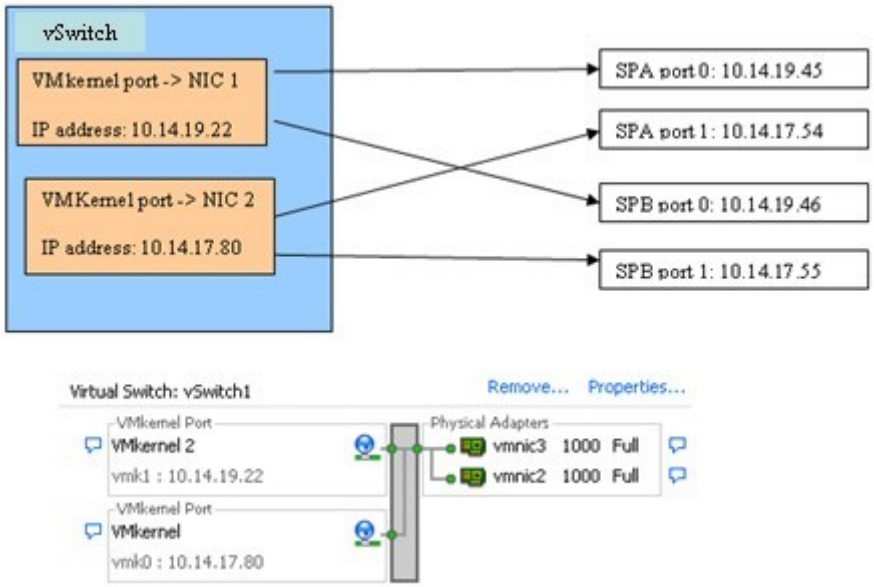


Figure 2. Two NICs on a single vSwitch iSCSI configuration

In ESXi, you can configure a single vSwitch containing two NICs to use NIC teaming or port binding to provide failover capabilities. You can enable port binding by overriding the vSwitch failover order such that each NIC is only bound to one VMkernel port. See Set up 1:1 VMkernel to network adapters mapping section for the procedure to perform port binding.

NOTE: Dell EMC recommends having two NICs/VMkernel ports on different subnets. Ensure the Storage Processor (SP) ports belonging to the same SP are also on different subnets.

Similarly, two vSwitches can be created on ESXi and each vSwitch can be bound to one or more NICs, as shown in the following figure:

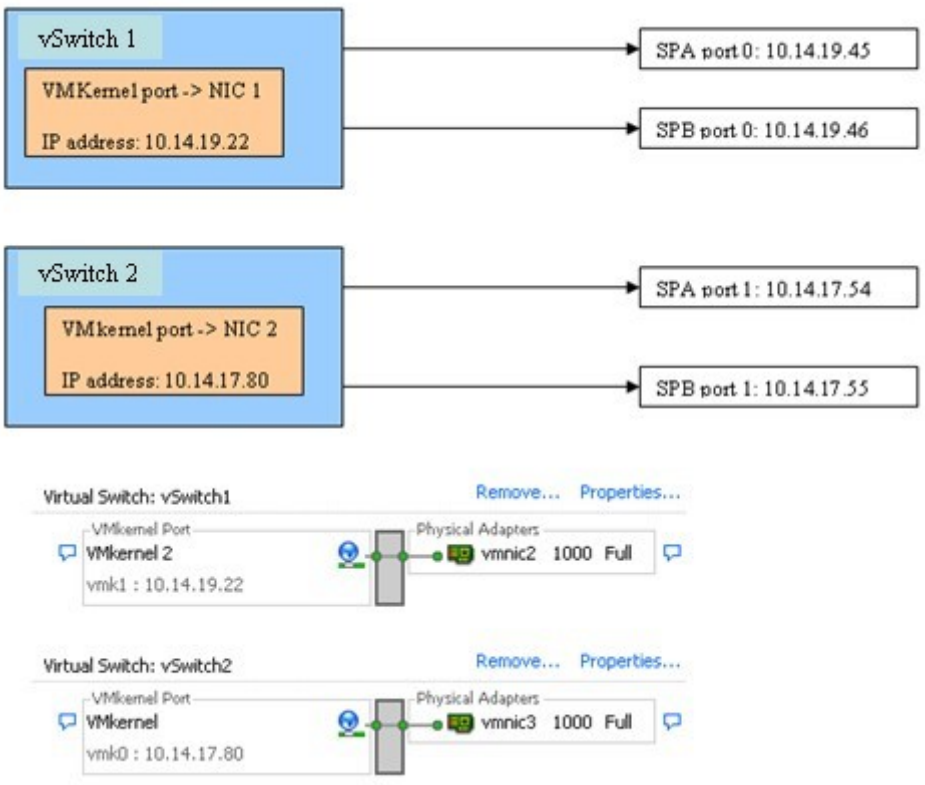


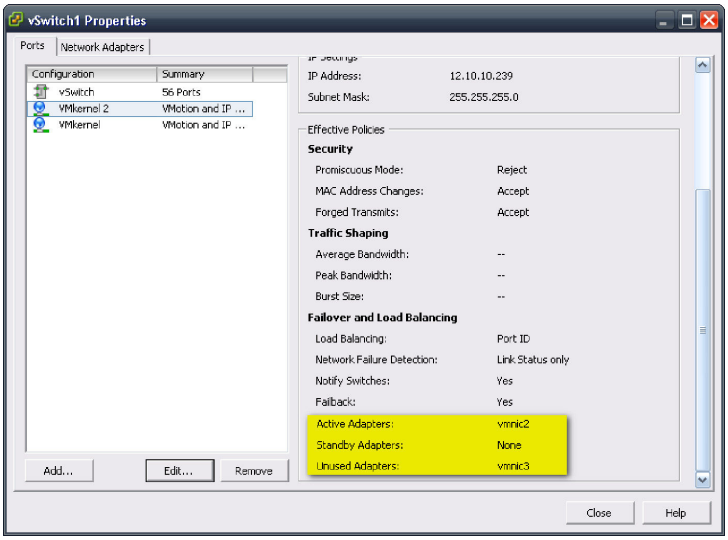
Figure 3. Two NICs in dual vSwitch iSCSI configuration

Set up 1:1 VMkernel to network adapters mapping

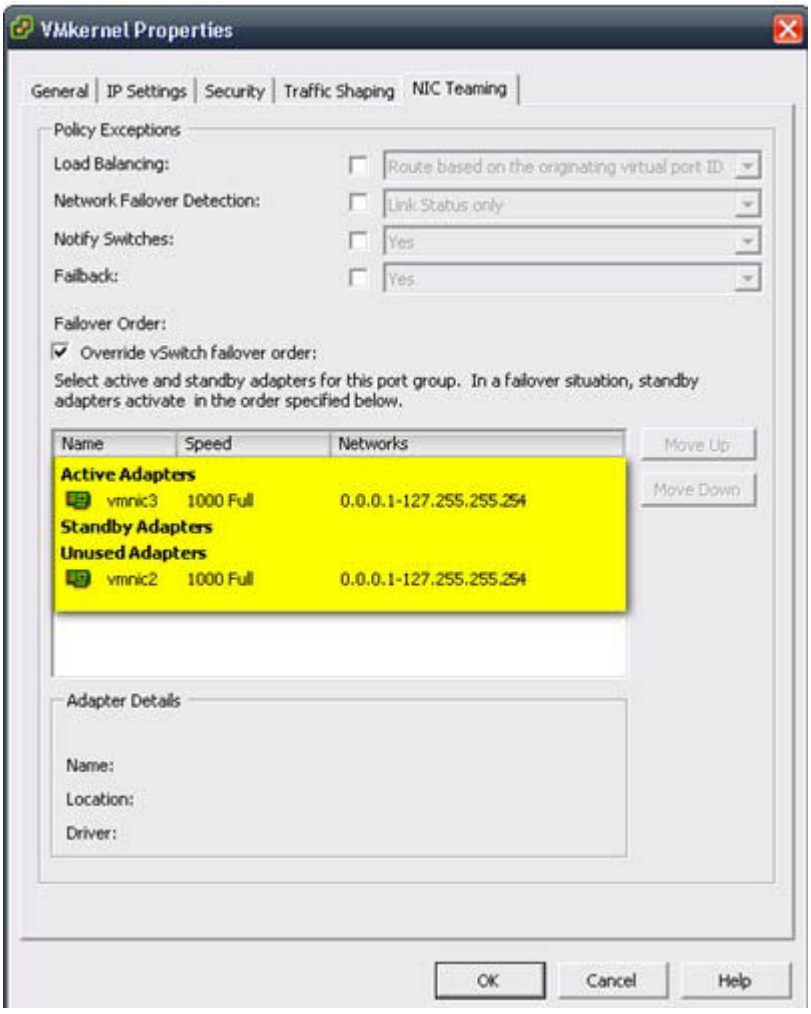
To set up 1:1 VMkernel to network adapters mapping, complete the following steps:

Steps

1. In the **vSwitch Properties** window, select **VMkernel** and click **Edit**.



2. Select the **NIC Teaming** tab and click the **Override vSwitch failover order** check box.



3. Set one active adapter for each VMkernel and move the rest of the adapters to Unused Adapters.
   Repeat this step for every VMkernel on the vSwitch.

NOTE: Make sure the number of VMkernel and Network adapters are same.

4. Activate the host-based multipathing by connecting each VMkernel port (for example: vmk0 and vmk1) to the iSCSI initiator from the service console by running the following commands:

   For vSphere 5:

   ```
   # esxcli iscsi networkportal add -n vmk0 -A vmhba32
   # esxcli iscsi networkportal add -n vmk1 -A vmhba32
   # esxcli iscsi networkportal list -A vmhba32
   ```

   Both vmk0 and vmk1 should be listed and there should be two separate paths to the same target.

5. To remove VMkernel ports from iSCSI initiators, ensure that there are no active sessions between hosts and targets, and run the following command from the service console:

   For vSphere 5:

   ```
   # esxcli iscsi networkportal remove -n vmk0 -A vmhba32
   ```

Network-attached storage

Network-attached storage (NAS) is a file-level computer data storage server that is connected to a computer network providing data access to a heterogeneous group of clients.

NAS removes the responsibility of file serving from other servers on the network. They typically provide access to files using network file sharing protocols such as NFS and SMB/CIFS.

NOTE: The entire network infrastructure must also support Jumbo Frames.

When creating NFS-based storage for ESXi, Dell EMC recommends creating VMware NFS Datastores instead of general-purpose NFS file systems. VMware NFS Datastores are optimized to provide better performance with ESXi. Install the VAAI plug-in before creating NFS Datastores to enable the following optimizations:

- Use 10 Gbps for the best performance.
- Configure Jumbo Frames (MTU of 9000) on all NAS ports.
- Use network trunking and multipathing to provide port failover and greater aggregate bandwidth for NAS connections to a single DM.

Setting up configuration

See the VMware documentation for version-specific configurations. The procedure may vary depending on the ESXi version.

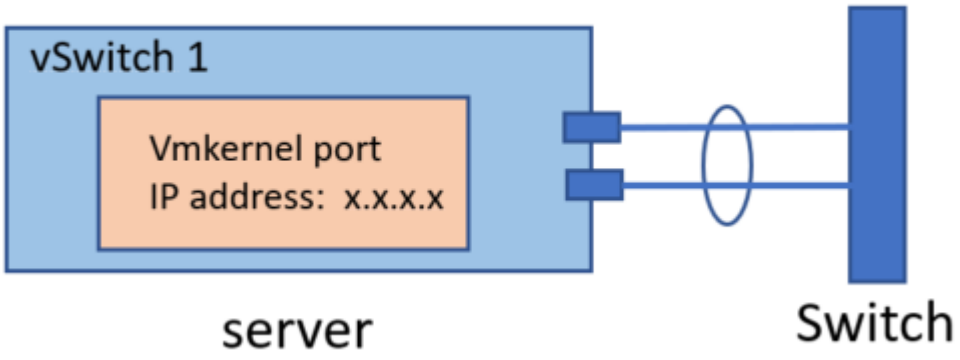
About this task

- VMware vSphere 7.0
- VMware vSphere 6.7
- VMware vSphere 6.5

For high availability, the LAN of NFS connection needs to be designed with availability, downtime-avoidance, isolation, and no single point of failure.

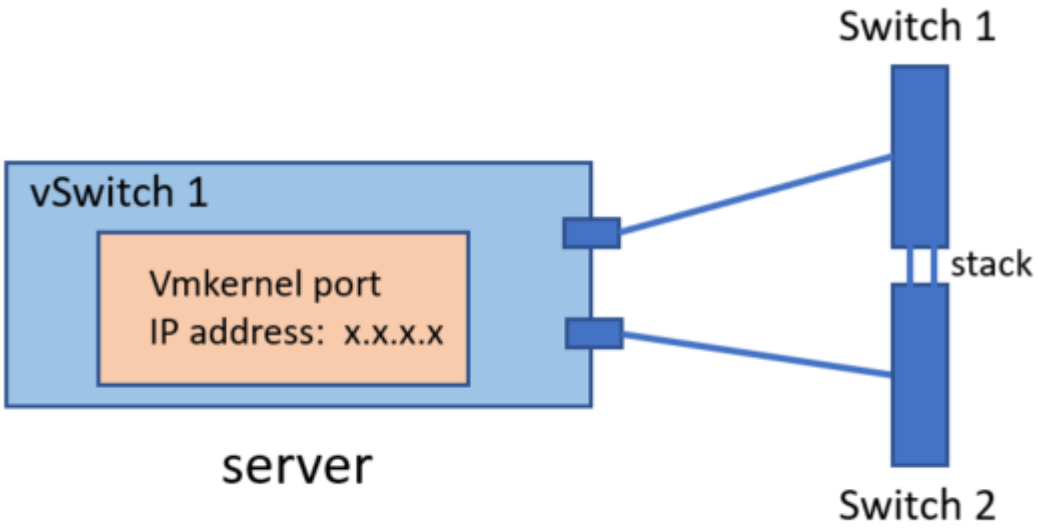
Steps

1. To avoid single point of failure at NIC level, connect two NICs to the same LAN switch. Configure as teamed at the switch and enable IP hash failover at the ESXi server.

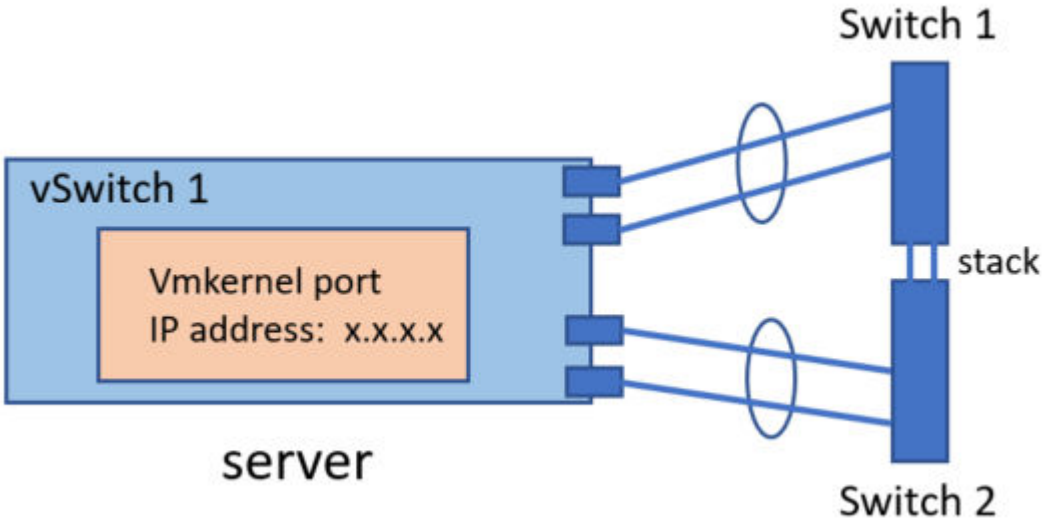


2. To avoid a single point of failure at switch level, configure multiple NICs with IP hash failover and two pairs going to separate LAN switches.

   The following figure shows two NICs with each pair configured as teamed at the respective LAN switches:



   The following figure shows four NICs with each pair configured as teamed at the respective LAN switches:



NVMe

Non-Volatile Memory Express (NVMe) is a standardized protocol that is designed for high-performance multi-queue communication with NVM devices. ESXi supports the NVMe protocol to connect to local and networked storage devices.

Setting up configuration

See the VMware documentation for version-specific configurations. The procedure may vary depending on the ESXi version.

- VMware vSphere 7.0

3

Managing Storage and Disk Paths in VMware ESXi Environments

Topics:

ESXi storage architecture and multipathing overview

The VMkernel has both storage subsystems and support for a limited number of file systems, along with relevant caching. The storage code supports a few of Host Bus Adapters (HBAs) including parallel SCSI, SAS, Fibre Channel, and iSCSI. These HBAs connect a wide variety of active/active and active/passive storage arrays which have been certified for use with the VMkernel.

The primary file system of the VMkernel uses the VMware Virtual Machine File System (VMFS). VMFS is a clustered file system that is designed and optimized to support large files such as virtual drives and swap files. The VMkernel also supports the storage of virtual drives on NFS file systems. The following figure shows the basics of the VMkernel core, with special attention to the storage stack:

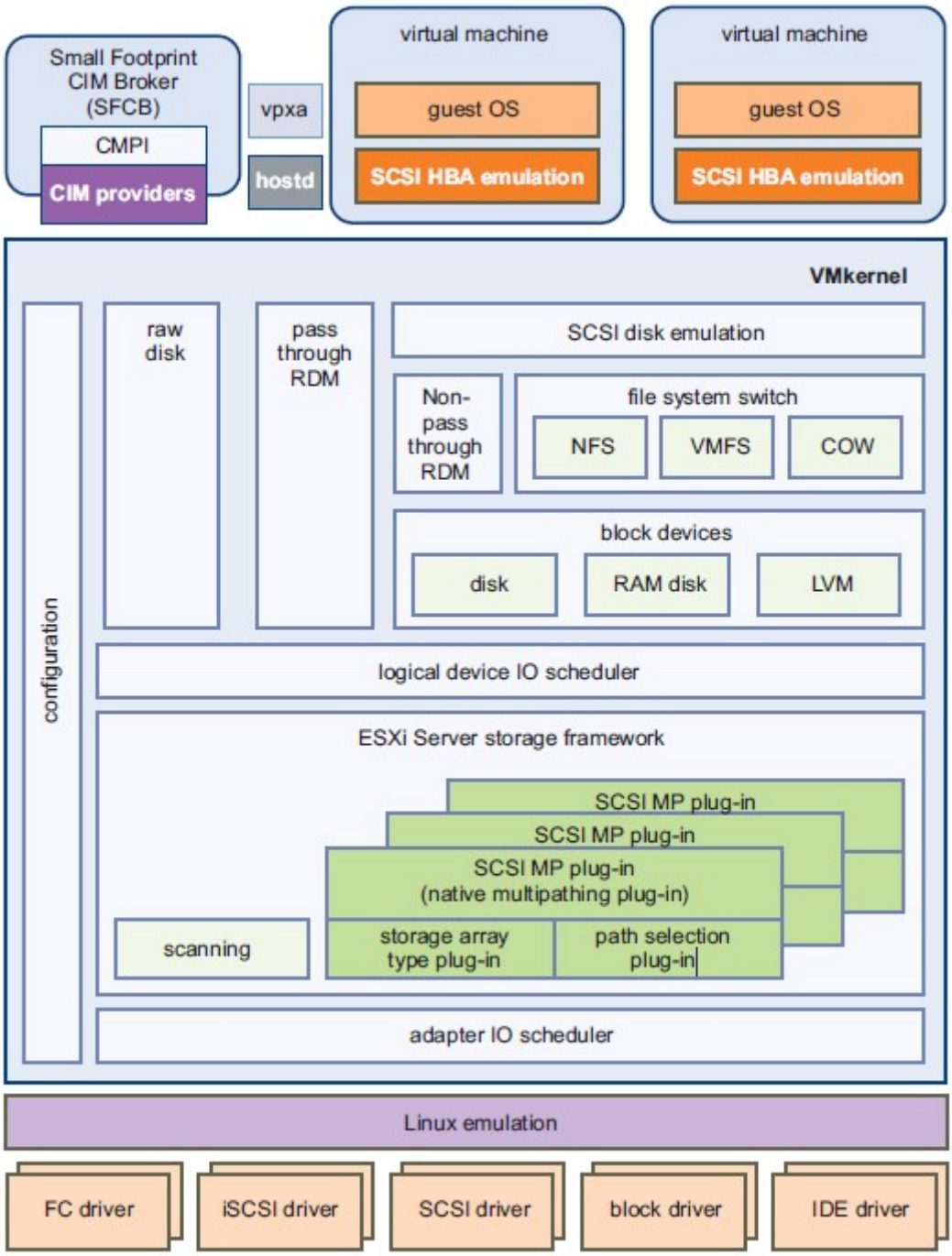


Figure 4. VMkernel and Storage

The storage I/O path provides virtual machines with access to storage devices through device emulation. This device emulation enables a virtual machine to access files on a VMFS or NFS file system as on SCSI devices.

In addition, vSphere has a special VMkernel layer called Pluggable Storage Architecture (PSA) framework. This Architecture provides storage virtualization functions such as fair share scheduling of I/O requests from multiple VMs. VM-transparent multipathing sits between the logical device I/O scheduler and the adapter I/O scheduler layers, as shown in the previous figure.

Native multipathing in VMware ESXi server

A multipathing plug-in is a module that supports the storage virtualization actions of path failover and path selection. The default VMware-supplied VMkernel module that does these operations, is called the Native Multipathing Plug-In (NMP).

Major components

The NMP provides support for multiple different storage arrays using loadable Storage Array Type Plug-ins (SATPs) and support for multiple different path selection algorithms using loadable Path Selection Plug-ins (PSPs).

Claim rules

When ESXi boots, the PSA framework discovers the physical paths to the attached storage devices. The PSA framework issues a SCSI inquiry command on each path to obtain the Vital Product Data (VPD) for the attached device.

The framework then uses this VPD, along with information about the HBA to which the path is attached, and a set of system-wide claim rules to determine which Multipathing Plugin (MPP) should be given management of each path. Each claim rule consists of a path specification and a plug-in name (NMP or a third-party MPP).

The claim rules are numbered. For each physical path, the PSA framework runs the claim rules starting with the lowest numbered rule first. The attributes of the physical path are compared to the path specification in the claim rule. If there is a match, the MPP that is associated with the claim rule can manage the physical path. If there is no match, or if the MPP declines to claim the given physical path, the physical path is compared to the path specification in the next claim rule. This process continues until a match is found and the physical path is claimed by an MPP. The highest numbered claim rule is the default claim rule and will cause the physical path to be claimed by the NMP if no other MPP claims it.

Path policies

When a virtual machine issues an I/O request to the VMkernel on a device that is managed by the NMP, the NMP determines which physical path to use for the I/O request by the PSP. NMP supports the following three basic paths select policies:

- **Fixed**-This is the default policy for Active/Active array and always uses the single preferred path to access the disk. An alternate path is taken only when the preferred path is not available to access the disk. The ESXi host automatically returns to the preferred path when the path becomes available.
- **Most Recently Used (MRU)**-ESXi Server uses the single most recently used path to the disk. If this path becomes unavailable, the host switches to an alternative path and continues to use the new path while it is available. There is no automatic failback in the MRU policy. The ESXi host does not automatically return to the preferred path.
- **Round Robin (RR)**-Round Robin uses an automatic path selection, rotating through all available paths and enabling load balancing across the paths. For Active/Passive arrays, only active paths on the owning SP are used in RR. For Active/Active (but not ALUA) arrays, all paths are used in RR.

Path configuration

When you run the `esxcli storage nmp path list` command to view path details, the result displays a statement `does not support path configuration`, which is not an error message. It indicates that there is no interface to change the path configuration. The following figure shows an example:

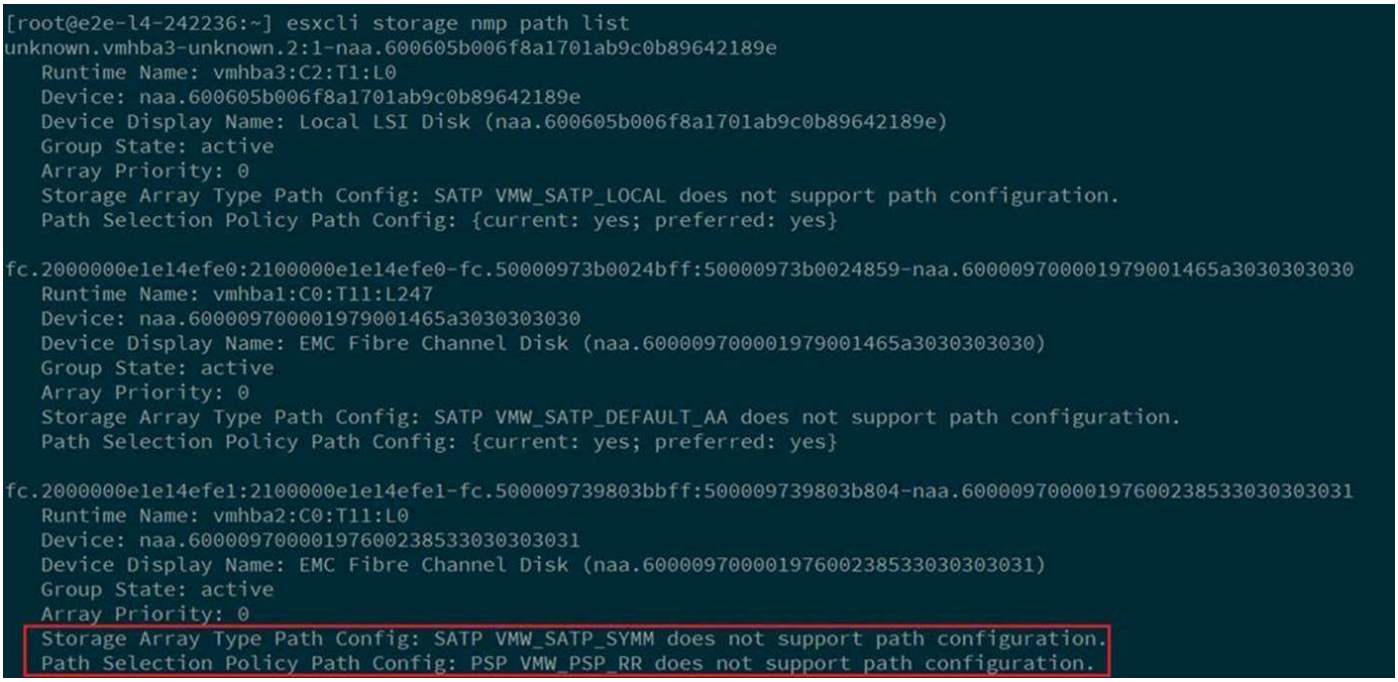


Figure 5. Path configuration

Commands

See vSphere Command-Line Interface Concepts and Examples for details about storage relevant ESXi CLI usage.

Supported storage types

See Storage/SAN Compatibility Guide for the list of storage types that are supported by the NMP.

Third-party multipathing in VMware ESXi server

If your storage array has advanced failover, load-balancing or path management features, and you find that the default NMP does not provide optimal functionality, VMware allows its partners to develop their own plugins (SATP, PSP, MPP) for further control.

PowerPath/VE is a Dell EMC developed MPP. It manages the following functionality:

- Performs physical path claiming and unclaiming.
- Registers and de-registers logical devices, and associates physical paths with logical devices.
- Processes I/O requests to logical devices such as load balancing and handling failures and retry requests.
- Supports management tasks such as *abort* or *reset* of logical devices.

Major components

The PowerPath/VE configuration has two major components that must be deployed:

- PowerPath driver (emcp module)
- Remote PowerPath CLI (rpowermt server)

The PowerPath/VE driver (emcp module) resides in ESXi host within kernels so that multipathing and load-balancing functions are transparent to VMware and to any guest operating system running on VMware.

The rpowermt CIM client enables communication between the VMware ESXi host and the rpowermt host for the management of PowerPath/VE using Remote PowerPath CLI (rpowermt) server where PowerPath remote multipathing CLI and licenses are installed.

Electronic License Management (ELM) is used for PowerPath/VE in VMware ESXi server.

Claim rules

PowerPath/VE and VMware NMP can co-exist on the same vSphere host. PowerPath/VE and NMP can share the same HBAs but they cannot manage the same device simultaneously. Claim rules are used to assign storage devices either to PowerPath/VE or to NMP devices.

During PowerPath/VE installation, PowerPath/VE claims all supported Dell EMC storage types by default. If you do not want all these devices under PowerPath/VE control, but want some Dell EMC devices under NMP control, change the claim rules.

To manage some of the Dell EMC storages by NMP, you need to add a claim rule by defining it with a number between 201 and 250. The PowerPath/VE default numbers are between 250 and 360. The number of rules you must add for each device depends on the number of HBAs in the vSphere host and the array type.

After PowerPath/VE is installed, use the VMware command `esxcli storage core claimrule list` at the SSH or command prompt to list the current set of claim rules as shown in the following figure:

Rule Class	Rule	Class	Type	Plugin	Matches	XCOPY Use Array Reported Values	XCOPY Use Multiple Segments	XCOPY Max Transfer Size
MP	0	runtime	transport	NMP	transport=usb	false	false	0
MP	1	runtime	transport	NMP	transport=sata	false	false	0
MP	2	runtime	transport	NMP	transport=ide	false	false	0
MP	3	runtime	transport	NMP	transport=block	false	false	0
MP	4	runtime	transport	NMP	transport=unknown	false	false	0
MP	101	runtime	vendor	MASK_PATH	vendor=DELL model=Universal Xport	false	false	0
MP	101	file	vendor	MASK_PATH	vendor=DELL model=Universal Xport	false	false	0
MP	250	runtime	vendor	PowerPath	vendor=DGC model=*	false	false	0
MP	250	file	vendor	PowerPath	vendor=DGC model=*	false	false	0
MP	260	runtime	vendor	PowerPath	vendor=EMC model=SYMMETRIX	false	false	0
MP	260	file	vendor	PowerPath	vendor=EMC model=SYMMETRIX	false	false	0
MP	270	runtime	vendor	PowerPath	vendor=EMC model=Invista	false	false	0
MP	270	file	vendor	PowerPath	vendor=EMC model=Invista	false	false	0
MP	280	runtime	vendor	PowerPath	vendor=HITACHI model=*	false	false	0
MP	280	file	vendor	PowerPath	vendor=HITACHI model=*	false	false	0
MP	290	runtime	vendor	PowerPath	vendor=HP model=*	false	false	0
MP	290	file	vendor	PowerPath	vendor=HP model=*	false	false	0
MP	300	runtime	vendor	PowerPath	vendor=COMPAQ model=HSV111 (C)COMPAQ	false	false	0
MP	300	file	vendor	PowerPath	vendor=COMPAQ model=HSV111 (C)COMPAQ	false	false	0
MP	310	runtime	vendor	PowerPath	vendor=EMC model=Celerra	false	false	0
MP	310	file	vendor	PowerPath	vendor=EMC model=Celerra	false	false	0
MP	320	runtime	vendor	PowerPath	vendor=IBM model=2107900	false	false	0
MP	320	file	vendor	PowerPath	vendor=IBM model=2107900	false	false	0
MP	330	runtime	vendor	PowerPath	vendor=IBM model=2810XIV	false	false	0
MP	330	file	vendor	PowerPath	vendor=IBM model=2810XIV	false	false	0
MP	340	runtime	vendor	PowerPath	vendor=XtremIO model=XtremApp	false	false	0
MP	340	file	vendor	PowerPath	vendor=XtremIO model=XtremApp	false	false	0
MP	350	runtime	vendor	PowerPath	vendor=NETAPP model=*	false	false	0
MP	350	file	vendor	PowerPath	vendor=NETAPP model=*	false	false	0
MP	360	runtime	vendor	PowerPath	vendor=COMPELNT model=*	false	false	0
MP	360	file	vendor	PowerPath	vendor=COMPELNT model=*	false	false	0
MP	370	runtime	vendor	PowerPath	vendor=EMC model=VMAXLUNZ	false	false	0
MP	370	file	vendor	PowerPath	vendor=EMC model=VMAXLUNZ	false	false	0
MP	65535	runtime	vendor	NMP	vendor=* model=*	false	false	0

Figure 6. List showing the current set of claim rules

In the preceding figure:

- **Rule** is the claim rule number that corresponds to the device that is managed.
- **Class** is the status of the claim rule. The claim rule has been added persistently.
- **Runtime** means that the claim rule has been fully loaded into vSphere memory.
- **Type** is the type of claim or unclaim operation to perform. Valid values include: transport, vendor.
- **Plugin** is the plug-in that is managing the device. Valid values include: NMP, PowerPath.
- **Matches** is the criteria that are selected to be applied the devices by the claim rule.

For more information about management of PowerPath/VE, see the PowerPath/VE for VMware vSphere Installation and Administration Guide. This guide is updated with every new release.

Path policies

Supported Dell EMC arrays are classified in three types of I/O path control mode: Active/Active, Active/Passive and ALUA.

Each of storage system with different I/O path control mode has complement load balancing and failover policies as follows:

Symmetrix optimization (so)-I/O requests are routed to paths based on an algorithm that considers path load. Load is a function of the number, size, and type of I/O queued on each path.

- This policy is available for the following storage classes: symm
- This policy is the default policy with a PowerPath license for Symmetrix storage systems.

CLARiiON optimization (co)-I/O requests are assigned to paths based on an algorithm that considers path load and logical device priority.

- This policy is available for the following storage classes: CLARiiON, Unity, and VNX

- This policy is the default policy with a PowerPath license for Unity, VNX, and CLARiiON storage systems.

Adaptive (ad)-I/O requests are assigned to paths based on an algorithm that considers path load and logical device priority.

- This policy is available for the following storage classes: SC, VPLEX, and XtremIO
- This policy is the default policy with a PowerPath license for VPLEX, Dell EMC SC, and XtremIO storage systems.

Round Robin (rr)-I/O requests are assigned to each available path in rotation.

- This policy is available for the all storage classes.
- This policy is not a default policy on any storage system.

After PowerPath/VE installed on the VMware vSphere host, the default load balancing and failover policies are **so** for Symmetrix devices, `co` for Unity, VNX, and **ad** for VPLEX, VNXe.

When PowerPath/VE is installed but not licensed, the rpowermt server is unable to display the information for the vSphere host. Upon installation of a valid PowerPath/VE for VMware vSphere license, host display and management capabilities are enabled through the rpowermt server.

Verify information about HBAs and/or devices that are managed by PowerPath/VE by running the following command:

```
rpowermt display dev=all class=<class>|all host=<hostname|host ip>.
```

If required, change the load-balancing policy by running the following command:

```
rpowermt set policy=rr dev=all class=<class>|all host=<hostname|host ip>.
```

Command options are:

- class = {all|symm|vnx|clariion|vplex|invista|netapp|hitachi|hpxp|ess|xtremio|unity|sc|generic}

  This command option limits the command to the specified type of storage system. The option **all** specifies all storage-system types. The default is **all**.
- dev = device|all

  This command option displays information about the specified devices. The option **all** specifies all devices. If dev is not specified, a summary HBA listing is displayed, with one line per HBA. The supported device name formats for rpowermt display dev are:

  - PowerPath/VE pseudo device name
  - Native path identifier - Format of the native device path is, `hwpath:Cx:Ty:Lz value`

PowerPath commands

The PowerPath Command Line Interface (CLI) allows users to use PowerPath to manage storage resources.

For details about PowerPath CLI, see the CLI and System Messages Reference. Unlike other PowerPath platforms, which use a local CLI to manage their environment, PowerPath/VE uses a remote CLI (rpowermt) to configure and manage PowerPath and paths under its control on vSphere hosts in the PowerPath/VE environment. You need to install the PowerPath remote multipathing rpowermt (RTOOLS) CLI package or the PowerPath Management Appliance. For more information, see the PowerPath/VE for VMware vSphere Installation and Administration Guide.

Supported storage types

For the list of storage systems that PowerPath or VE currently supports, see the support matrix available on Dell EMC E-Lab Navigator.

VMware High-Performance Plug-in

The VMware High-Performance Plug-in (HPP) replaces the NMP for high-speed devices, such as NVMe. The HPP improves the performance of ultra-fast flash devices that are installed locally on the ESXi host and it is the default plug-in that claims NVMe-oF targets. To support multipathing, the HPP uses the Path Selection Schemes (PSS). A particular PSS is responsible for selecting physical paths for I/O requests.

4

Host Connectivity with Dell EMC Products

Topics:

PowerMax/VMAX All Flash

This section provides support information about the PowerMax, VMAX All Flash, and VMAX3 series.

Dell EMC PowerMax

Dell EMC PowerMax provides a platform that is ready for current and next generation data center data storage requirements. The PowerMax Family uses NVMe technology for customer application data.

PowerMax is built using a 100% NVMe storage back-end, allowing it to reach optimal IOPS densities by eliminating the flash media issues that are found in traditional SAS and SATA interfaces. NVMe over Fibre Channel 32 Gb front-end connections are supported starting with the PowerMaxOS 5978.444 release running on PowerMax storage systems allows for end-to-end NVMe accessibility, the PowerMaxOS 5978.444 and later releases running on PowerMax storage systems. It also introduces capability for traditional Fibre Channel FCP connections with an operation link speed of 32 Gb. NVMe over Fibre Channel and 32 Gb FCP both require the installation of the 32 Gb Fibre Channel I/O module. The PowerMaxOS 5978.479 and later releases leverages the 32 Gb Fibre Channel I/O module for 32 Git SRDF link connection support. All PowerMaxOS releases maintain downward compatibility with the legacy 16 Gb Fibre Channel I/O module.

There are two PowerMax models as follows:

PowerMax 2000-PowerMax 2000 is designed to provide high efficiency and flexibility to the data center, providing 1.7 million IOPS (8 K RRH) and up to 1 PB of effective capacity in 20U total space. It is the entry NVMe scale out array sold with the Essentials and Pro software packages.

PowerMax 8000-PowerMax 8000 designed to provide optimal scalability, performance, and IOPS density. It can consolidate disparate workloads on a mass scale as 8 Bricks can support 10 million IOPS (8 K RRH) and provide up to 4 PB of effective capacity in just two floor tiles of space. It is the flagship NVMe scale out array sold with the Essentials and Pro software packages.

Dell EMC PowerMax Operating System environment

PowerMaxOS 5978 is the only supported software operating environment for PowerMax storage arrays. It manages the storage and controls communications with the host systems and introduces support of dedupe and other new feature enhancements.

For detailed product specifications including the 32G Frontend I/O and 32G SRDF protocol support information, see Dell EMC PowerMax Family documentation available on https://shop.dellemc.com and *Dell EMC PowerMax Family Product Guide* on Dell EMC Online Support.

Dell EMC VMAX All Flash

Dell EMC VMAX All Flash range of storage arrays use only high-density flash drives. The range contains four models that combine high scale, low latency, and rich data services:

- VMAX 250F with a maximum capacity of 1.16 PBe (Petabytes effective)
- VMAX 450F with a maximum capacity of 2.3 PBe
- VMAX 850F with a maximum capacity of 4.4 PBe
- VMAX 950F with a maximum capacity of 4.42 PBe

Dell EMC VMAX All Flash Operating Environment

Dell EMC VMAX All Flash storage systems are supported with both HYPERMAX OS 5977 and PowerMaxOS 5978, however not all new feature enhancements that are introduced with PowerMax storage systems running PowerMaxOS 5978 would be available with Dell EMC VMAX All Flash storage systems also running PowerMaxOS 5978.

HYPERMAX OS 5977 provides emulations that perform specific data service and control functions in the HYPERMAX environment, and it introduces an open application platform for running data services and provide file system storage with eNAS and embedded management services for Unisphere. The entire feature set available with HYPERMAX OS 5977 running on Dell EMC VMAX All Flash storage systems would be available with PowerMaxOS 5978 running on Dell EMC VMAX All Flash storage systems except FCoE front-end connections.

For detailed product specifications including the Frontend I/O protocol support information, see *VMAX All Flash Product Guide* and *VMAX All Flash: Family Overview* available on https://www.dellemc.com/.

PowerMax/VMAX All Flash/VMAX3 - Device types

There are two classes of devices available on PowerMax/VMAX All Flash/VMAX3 storage arrays:

- Compatibility device - The Compatibility device is the default VMAX device type, advertises T10 SPC-3 compliancy and behaves almost like the legacy VMAX devices.
- Mobility device - The Mobility device is an optionally configurable device type that is supported with HYPERMAX OS 5977.811.784 and later, advertises T10 SPC-4 compliancy, and has numerous INQUIRY and VPD page changes implemented intended to allow for support of future feature enhancements. Implicit Asymmetric Logical Unit Access (ALUA) is advertised by the Mobility device, however the Mobility device is configured into only a single active target port group per PowerMax or VMAX storage system and a second target port group is configured on a remote PowerMax or VMAX system only when the Mobility devices are configured in an SRDF Metro configuration which is restricted only to Windows and Solaris operating system environments for now. For all supported SRDF/Metro Mobility ID/ALUA, see the Support Matrix available on Dell EMC E-Lab Navigator.

  Mobility devices are also not currently supported with NDM migration.

Multipathing in PowerMax/VMAX Mobility device

When using VMware Native Multipathing Plug-In (NMP) with VMAX3/PowerMax mobility device in ALUA mode, you must set the Storage Array Type Plug-In (SATP) as *VMW_SATP_ALUA*.

About this task

NOTE: Default SATP for VMAX3/PowerMax devices is VMW_SATP_SYMM. It is loaded when VMAX3/PowerMax devices are probed, as shown in the following example:

```
# esxcli storage nmp satp rule list
Name        Device        Vendor Model    Driver    Transport    Options
Rule Group    Claim Options    Default PSP    PSP Options Description
------------------    ------------------------------------
-------    ----------------    ------    ---------
--------------------------    ----------
----------------------------------    -----------    -----------
-------------------------------------------------------------
-----------
VMW_SATP_SYMM    EMC    SYMMETRIX system

~ # esxcli storage nmp satp list
VMW_SATP_ALUA    VMW_PSP_MRU    Placeholder (plugin not loaded)
```

Perform the following steps to set the SATP as *VMW_SATP_ALUA*:

Steps

1. Run the following command to add a new rule where the ALUA VMAX devices are claimed by ALUA SATP with Round Robin Path Select Policy:

   ```
   esxcli storage nmp satp rule add -V EMC -M SYMMETRIX -s VMW_SATP_ALUA -c tpgs_on -P
   VMW_PSP_RR
   ```

   Once the rule is added, it appears under *esxcli storage nmp satp rule* list, as shown in the following example:

```
VMW_SATP_ALUA           EMC         SYMMETRIX
user               tpgs_on                              VMW_PSP_RR
```

2. Reboot the host to apply the new rule.

   Once the rule is applied, you can verify it when the VMW_SATP_ALUA Plug-In is loaded. The VMAX3/PowerMax mobility devices are claimed with all paths as Active-Optimized.

3. Run the following commands to verify the device and path status:

```
~ # esxcli storage nmp satp list
VMW_SATP_ALUA    VMW_PSP_MRU    Supports non-specific arrays that use the ALUA protocol

~ # esxcli storage nmp device list naa.600009700bb96bea18bf0083000000c5
Device Display Name: EMC iSCSI Disk (naa.600009700bb96bea18bf0083000000c5)
Storage Array Type: VMW_SATP_ALUA
Storage Array Type Device Config:
{implicit_support=on;explicit_support=off; explicit_allow=on;alua_followover=on;
{TPG_id=1,TPG_state=AO}}
Path Selection Policy: VMW_PSP_RR Path Selection Policy Device Config:
{policy=rr,iops=1000,bytes=10485760,useANO=0; lastPathIndex=2:
NumIOsPending=0,numBytesPending=0}
Path Selection Policy Device Custom Config: Working Paths: vmhba6:C0:T4:L9, vmhba7:C0:T4:L9,
vmhba7:C0:T0:L9, vmhba6:C0:T2:L9 Is Local SAS Device: false  Is USB: false
Is Boot USB Device: false

~ # esxcli storage nmp path list | less
iqn.1990-07.com.emulex:wnh12h15skyhawk-4000006c0000,iqn.1992-04
.com.emc:600009700bb96bea18bf018300000000,t,0-naa.600009700bb96 bea18bf0083000000c5
Runtime Name: vmhba6:C0:T2:L9
Device: naa.600009700bb96bea18bf0083000000c5 Device Display Name: EMC iSCSI Disk
(naa.600009700bb96bea18bf0083000000c5)
Group State: active

Array Priority: 0
Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=115,RTP_health=UP}
Path Selection Policy Path Config: PSP VMW_PSP_RR does not support path configuration.

iqn.1990-07.com.emulex:wnh12h15skyhawk-4000006c0000,iqn.1992-04
.com.emc:600009700bb96bea18bf018300000002,t,0-naa.600009700bb96 bea18bf0083000000c5
Runtime Name: vmhba7:C0:T0:L9
Device: naa.600009700bb96bea18bf0083000000c5 Device Display Name: EMC iSCSI Disk
(naa.600009700bb96bea18bf0083000000c5)
Group State: active Array Priority: 0
Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=1139,RTP_health=UP}
Path Selection Policy Path Config: PSP VMW_PSP_RR does not support path configuration.

iqn.1990-07.com.emulex:wnh12h15skyhawk-4000006c0000,iqn.1992-04
.com.emc:600009700bb96bea18bf018300000003,t,0-naa.600009700bb96 bea18bf0083000000c5
Runtime Name: vmhba7:C0:T4:L9
Device: naa.600009700bb96bea18bf0083000000c5 Device Display Name: EMC iSCSI Disk
(naa.600009700bb96bea18bf0083000000c5)
Group State: active Array Priority: 0
Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=1140,RTP_health=UP}
Path Selection Policy Path Config: PSP VMW_PSP_RR does not support path configuration.

iqn.1990-07.com.emulex:wnh12h15skyhawk-4000006c0000,iqn.1992-04
.com.emc:600009700bb96bea18bf018300000001,t,0-naa.600009700bb96 bea18bf0083000000c5
Runtime Name: vmhba6:C0:T4:L9
Device: naa.600009700bb96bea18bf0083000000c5 Device Display Name: EMC iSCSI Disk
(naa.600009700bb96bea18bf0083000000c5)
Group State: active Array Priority: 0
Storage Array Type Path Config:
{TPG_id=1,TPG_state=AO,RTP_id=116,RTP_health=UP}
Path Selection Policy Path Config: PSP VMW_PSP_RR does not support path configuration.
```

NOTE: The Mobility devices already presented to the host before the new rule are not affected unless they are unmapped from the host, and mapped again. The other option is to change SATP rule of the existing mobility devices manually without the unmap operation.

NOTE: There is no need to change the default system VMW_SATP_SYMM rule. VMware NMP can claim both VMAX3/PowerMax compatibility devices and mobility devices simultaneously by different SATP. VMAX3/PowerMax mobility devices are claimed by VMW_SATP_ALUA while VMAX3/PowerMax compatibility devices are claimed by VMW_SATP_SYMM.

Local replication services

TimeFinder SnapVX

Dell EMC TimeFinder SnapVX creates and manages point-in-time snapshots of critical data that can be used for backups, decision support, and to refresh data warehouse, test, and development environments. SnapVX snapshots do not require target volumes. SnapVX snapshots share back-end allocations with the source volume and other snapshots on the source volume.

TimeFinder SnapVX is supported on VMAX All Flash arrays running HYPERMAX OS 5977 and later, and snapshots are always consistent. Consistency across multiple arrays is achieved when source devices are in a composite group.

SnapVX provides very low impact snapshots and clones for VMAX LUNs. SnapVX supports up to 256 snapshots per source volume, which are tracked as versions with less overhead and simple relationship tracking. Users can assign names to identify their snapshots, and can set automatic expiration dates on each snapshot. SnapVX provides the ability to manage consistent point-in-time copies for storage groups with a single operation. Up to 1024 target volumes can be linked per source volume, providing read/write access as pointers or full-copy clones. TimeFinder in HYPERMAX OS also provides compatibility modes for users who rely on their TimeFinder Mirror, Clone, or VP Snap command scripts. This allows users to leverage their existing scripts while learning how to take advantage of the new features of SnapVX.

For details, see the *TimeFinder SNAPVX Local Replication* document available on PowerMax and VMAX All Flash Technical Documentation page on https://www.dellemc.com/ or https://www.dell.com/support.

Remote Replication Services

SRDF

The Symmetrix Remote Data Facility (SRDF) maintains real-time or near real-time copies of data on a production storage array at one or more remote storage arrays and only one SRDF device mirror can be read/write enabled at any point in time.

SRDF/Metro

HYPERMAX OS 5977.691.684 and Solutions Enabler/Unisphere for VMAX 8.1 first introduced the support for SRDF/Metro for VMAX3 and VMAX All Flash families of storage arrays and all versions of PowerMaxOS on Dell EMC PowerMax arrays.

With SRDF/Metro, the SRDF secondary device is read/write accessible to the host and takes on the external identity of the primary device (geometry, device WWN, and so on). By providing this external identity on the secondary device, both the primary and secondary devices may then appear as a single virtual device across the two SRDF paired arrays for presentation to a single host or host cluster.

With both the local and remote devices being read/write accessible concurrently, the host or hosts (in a cluster) can read and write to both primary and secondary devices with SRDF/Metro ensuring that each copy remains current, consistent, and addressing any write conflicts which may occur between the paired SRDF devices. A single PowerMax/VMAX3/VMAX All Flash array may simultaneously support multiple SRDF groups that are configured for SRDF/Metro operations and multiple SRDF groups that are configured for non-SRDF/Metro operations. PowerMaxOS Q3 2019 release with SRDF/Metro supports online device expansion though which devices could be expanded within an active SRDF/Metro group. These operations can be accomplished using the latest version of Unisphere for PowerMax or Solutions enabler. Both SRDF and SRDF/Metro support FC or GiGE links for connectivity between arrays.

For more detailed information about SRDF/metro configuration, see *Dell EMC SRDF Introduction* and *SRDF/Metro Overview and Best Practices Technical Note* available on https://www.dell.com/support.

Non-Disruptive Migration services

Dell EMC's Non-Disruptive Migration (NDM) allows user to perform online data migrations that are simple and completely non-disruptive to the host and application. NDM is designed to help automate the process of migrating hosts and applications to a new PowerMax array with no downtime. NDM leverages SRDF replication technologies to move the application data to the new array. It also uses auto-provisioning, with PowerPath or a supported host multipathing solution, to manage host access to the data during the migration process.

NDM provides PowerMax user the following benefits:

- Allows non-disruptive migration with hosts and applications online. Potentially supported source storage systems are legacy VMAX v2, VMAX3 series, VMAX All Flash systems, and PowerMax. Potentially supported target systems are VMAX3 series, VMAX All Flash systems, and PowerMax.
- Ease of use with control operations that automate the setup and configuration of the migration environment.
- Managed by familiar and simple user interfaces using Solutions Enabler and Unisphere.
- Migrations can be easily canceled and failed back to the source array for any reason prior to commit.
- Built in and does not require any additional software or licensing costs.

For more detailed features, see the *Dell EMC PowerMax and VMAX: Non-Disruptive Migration Best Practices and Operational Guide* white paper available on https://www.dellemc.com/.

See Dell EMC Simple Support Matrix available on Dell EMC E-Lab Navigator for host interoperability with various operating system platforms and multipathing software that is supported for Non-Disruptive migrations.

Use advanced query option on Dell EMC E-lab Navigator for specific configuration search.

While doing the PowerMax data migration for VMware platform, it is recommended to enable the `consistent_lun` property for initiator group. Use the following command:

```
# symaccess -sid <sid> -type initiator -name <ig_name> set consistent_lun on
```

PowerMax/VMAX All Flash-Storage provisioning

PowerMax and VMAX All flash series initial configuration is performed by a Dell EMC Customer Engineer (CE) through the PowerMax or VMAX Management Module Control Station (MMCS). The CE will configure the storage arrays settings for each Fibre Channel port. The procedures in this document assume that any switches and storage systems to be used in this configuration have been installed, and that the front-end adapter ports have been connected to the appropriate switch ports and switch zoning is completed.

NOTE: It is recommended to use Acess Logic (ACLX) to mask volumes.

Storage provisioning operations can be accomplished by Solutions enabler or Unisphere for PowerMax software.

Dell EMC Solutions Enabler

Dell EMC Solutions Enabler installation provides your host with SYMAPI, CLARAPI, and STORAPI shared libraries for use by Solutions Enabler applications, and the Symmetrix Command Line Interface (SYMCLI) for use by storage administrators and systems engineers. SYMCLI is a specialized library of UNIX-formatted commands that can be invoked one at a time. It supports single command-line entries and scripts to map and perform control operations on devices and data objects toward the management of your storage complex. It also monitors device configuration and status of devices that make up the storage environment. The target storage environments are typically PowerMax and VMAX arrays.

For detailed procedure, see *Dell EMC Solutions Enabler Array Controls and Management Guide* available on https://www.dell.com/support.

Dell EMC Unisphere for PowerMax

Dell EMC Unisphere for PowerMax enables management and monitoring of PowerMax arrays along with legacy VMAX All Flash and VMAX3 arrays. Unisphere is installed on a dedicated Windows or Linux server, or deployed as a Virtual Appliance (vAPP).

For detailed procedure, see *Dell EMC Unisphere for PowerMax Installation Guide* available on Dell EMC online Support under PowerMax and VMAX technical documents and videos section.

Dell EMC PowerMax/VMAX All Flash/VMAX3 - Director bit settings

Dell EMC PowerMax/VMAX All Flash/VMAX3 supports various director bit settings against different front-end protocol such as Fibre Channel FCP, NVMe over Fibre Channel (NVMeOF), iSCSI, and FCoE. They vary between switch-based connectivity and direct attach connectivity with various operating system types.

For detailed director bit settings, see the Simple Support Matrix available on Dell EMC E-Lab Navigator.

Dell EMC PowerMax Fibre Channel FCP connectivity

Dell EMC PowerMax arrays running with PowerMaxOS Q3 2019 introduces 32 Gb/s Fibre Channel modules support. Earlier versions of PowerMaxOS on PowerMax arrays and HYPERMAX OS on VMAX All Flash arrays supports 16 Gb/s front-end module for host connectivity.

For details, see the *Understanding storage provisioning* and *Understanding Host Management* sections in *Dell EMC Unisphere for PowerMax Product Guide* available on https://www.dellemc.com/ or Dell EMC Oline Support.

Dell EMC PowerMax NVMeoFC connectivity

This section describes the procedures for installing a Dell EMC approved Emulex and QLogic adapter into a VMware host environment and configuring the host for connection to a Dell EMC storage array using NVMe over Fibre Channel.

This section contains the following information:

- Prerequisites for first-time installation
- HBA firmware and BIOS verification
- NVMeoFC configuration

Prerequisites for first-time installation

This section lists the requirements before you install the HBA for the first time.

NOTE: Dell EMC does not support mixing different types of Fibre Channel adapter (including different types from the same vendor) in a server.

- See *Dell EMC Simple Support Matrix* available on E-Lab Navigator or contact your Dell EMC representative for the latest information about qualified adapters and drivers.

  Dell EMC supports both in-kernel and out-of-kernel drivers.

  NOTE: The installation of the in-kernel driver occurs when you install ESXi.

  If your installation requires an out-of-kernel driver, download it from Dell EMC or Vendor' website. Follow the links to your adapter for the appropriate operating system and version.
- See vendor's Fibre Channel Host Adapter (HBA) product documentation to properly install an HBA in your server.
- vSphere supports NVMeoFC allowing connectivity to external NVMe arrays using FC starting from vSphere 7.0.
- Deployment of vSphere 7.0 on PowerMax 8000 and PowerMax 2000 arrays that are configured with FC-NVMe connectivity requires support fix 104661 with PowerMaxOS 5978.479.479 code version. Contact Dell EMC support to have proper ePack installed.

HBA firmware and BIOS verification

The following procedure describes how to get HBA's firmware and BIOS version:

Steps

1. Run the following command to list HBAs:

   ```
   esxcfg-scsidevs -a
   ```

   The console displays results similar to the following:

```
[root@localhost:~] esxcfg-scsidevs -a
vmhba0  lsi_mr3          link-n/a  sas.51866da07697ef00                  (0000:02:00.0) Broadcom PERC H730P Mini
vmhba1  vmw_ahci         link-n/a  sata.vmhba1                           (0000:00:11.4) Intel Corporation Wellsburg RAID Controller
vmhba2  vmw_ahci         link-n/a  sata.vmhba2                           (0000:00:1f.2) Intel Corporation Wellsburg RAID Controller
vmhba3  lpfc             link-up   fc.20000000c97179c0:10000000c97179c0  (0000:82:00.0) Emulex Corporation Emulex LPe12000 8Gb PCIe Fibre Channel Adapter
vmhba4  lpfc             link-up   fc.20000000c97179c1:10000000c97179c1  (0000:82:00.1) Emulex Corporation Emulex LPe12000 8Gb PCIe Fibre Channel Adapter
vmhba5  lpfc             link-n/a  fc.200000109b8c4154:100000109b8c4154  (0000:81:00.0) Emulex Corporation Emulex LPe36000 Fibre Channel Adapter
vmhba6  lpfc             link-up   fc.200000109b8c4155:100000109b8c4155  (0000:81:00.1) Emulex Corporation Emulex LPe36000 Fibre Channel Adapter
vmhba7  qlnativefc       link-up   fc.2000f4e9d453ccba:2100f4e9d453ccba  (0000:03:00.0) QLogic Corp QLE2742 Dual Port 32Gb Fibre Channel to PCIe Adapter
vmhba8  qlnativefc       link-up   fc.2000f4e9d453ccbb:2100f4e9d453ccbb  (0000:03:00.1) QLogic Corp QLE2742 Dual Port 32Gb Fibre Channel to PCIe Adapter
vmhba64 brcmnvmefc       link-down fc.200000109b8c4154:100000109b8c4154  (0000:81:00.0) Emulex Corporation Emulex LPe36000 Fibre Channel Adapter
vmhba65 brcmnvmefc       link-up   fc.200000109b8c4155:100000109b8c4155  (0000:81:00.1) Emulex Corporation Emulex LPe36000 Fibre Channel Adapter
vmhba66 qlnativefc       link-up   fc.2000f4e9d453ccba:2100f4e9d453ccba  (0000:03:00.0) QLogic Corp QLE2742 Dual Port 32Gb Fibre Channel to PCIe Adapter
vmhba67 qlnativefc       link-up   fc.2000f4e9d453ccbb:2100f4e9d453ccbb  (0000:03:00.1) QLogic Corp QLE2742 Dual Port 32Gb Fibre Channel to PCIe Adapter
```

2. Run the following command to check the HBA vendor:

   ```
   esxcli hardware pci list | grep vmhba5 -A 25 -B 5
   ```

   This step considers *vmhba5* from the result of previous step as an example.

   The console displays results similar to the following:

```
[root@localhost:~] esxcli hardware pci list | grep vmhba5 -A 25 -B 5
   Address: 0000:81:00.0
   Segment: 0x0000
   Bus: 0x81
   Slot: 0x00
   Function: 0x0
   VMkernel Name: vmhba5
   Vendor Name: Emulex Corporation
   Device Name: Emulex LPe36000 Fibre Channel Adapter
   Configured Owner: VMkernel
   Current Owner: VMkernel
   Vendor ID: 0x10df
   Device ID: 0xf400
   SubVendor ID: 0x10df
   SubDevice ID: 0xf410
   Device Class: 0x0c04
   Device Class Name: Fibre Channel
   Programming Interface: 0x00
   Revision ID: 0x00
   Interrupt Line: 0xff
   IRQ: 255
   Interrupt Vector: 0x00
   PCI Pin: 0x00
   Spawned Bus: 0x00
   Flags: 0x3001
   Module ID: 31
   Module Name: lpfc
   Chassis: 0
   Physical Slot: 2
   Slot Description: PCIe Slot 2
   Device Layer Bus Address: s00000002.00
   Passthru Capable: true
```

Vendor ID **0x10df** indicates it is a Dell EMC branded HBA.

3. Check current firmware and BIOS version as shown in the following figure:

```
[root@localhost:~] /usr/lib/vmware/vmkmgmt_keyval/vmkmgmt_keyval -l -i vmhba5/Emulex |more
Listing keys:
Name:   adapter
Type:   string
value:
lpfc Adapter Page

Emulex LightPulse FC SCSI 12.4.293.2
Emulex LPe35002-M2-D 2-Port 32Gb Fibre Channel Adapter on PCI bus 0000:81 device 00 fn 0 port 1 Link Speed: 0 Gb

BoardNum:       0
FW Version:     12.6.182.8   ##################### firmware version
HW Version:     0000000d
ROM Version:    12.6.182.8
SerialNum:      MY04VDY3FLPB396SM01LA00
PCI ID:         10df f400 10df f410
.....................................
```

See vendor's product documentation to update HBA firmware and boot BIOS to Dell EMC-qualified versions.

NVMeoFC configuration

The following procedure describes how to configure an HBA for NVMeoFC connection:

About this task

NOTE: The Emulex LPe3500x HBA supports NVMeoFC by default, and does not require additional configuration.

Steps

1. Enable NVMeoFC support (optional).
   a. Run the following commands to enable NVMeoFC support for QLogic HBA:

```
[root@localhost:~]  esxcli system module parameters list -m qlnativefc | grep nvme
ql2xnvmesupport             int           Enable NVMeoFC support in driver.  Default is disabled.
[root@localhost:~]
[root@localhost:~]
[root@localhost:~] esxcli system module parameters set -p ql2xnvmesupport=1 -m qlnativefc
[root@localhost:~]
[root@localhost:~]
[root@localhost:~] esxcli system module parameters list -m qlnativefc | grep nvme
ql2xnvmesupport             int   1       Enable NVMeoFC support in driver.  Default is disabled.
[root@localhost:~]
[root@localhost:~]
[root@localhost:~]
[root@localhost:~] reboot
```

2. Run the following command to set hostname and get host NQN:

```
esxcli system hostname set --host=e2e-14-9732
esxcli nvme info get
```

   The console displays results similar to the following:

```
Host NQN: nqn.2014-08.com.vmware:nvme:e2e-14-9732
```

```
[root@localhost:~] esxcli system hostname set --host=e2e-14-9732
[root@e2e-14-9732:~]
[root@e2e-14-9732:~]
[root@e2e-14-9732:~] esxcli nvme info get
   Host NQN: nqn.2014-08.com.vmware:nvme:e2e-14-9732
[root@e2e-14-9732:~]
[root@e2e-14-9732:~]
```

3. Run the following command to list NVMe adapter:

```
esxcli nvme adapter list
```

   The console displays result similar to the following:

```
[root@e2e-14-9732:~] esxcli nvme adapter list
Adapter  Adapter Qualified Name          Transport Type  Driver      Associated Devices
-------  ------------------------------  --------------  ----------  ------------------
vmhba66  aqn:qlnativefc:2100f4e9d453ccba  FC              qlnativefc
vmhba67  aqn:qlnativefc:2100f4e9d453ccbb  FC              qlnativefc
```

4. Run the following command to connect to NVMe controllers and list NVMe devices.

   NOTE: Host reboot initiates NVMe fabric discovery, and controller autoconnect if storage view is configured at PowerMax side.

```
esxcli nvme controller list
```

   The console displays result similar to the following:

```
[root@e2e-14-9732:~] esxcli nvme controller list
Name                                                                                                  Controller Number  Adapter  Transport Type  Is Online
----------------------------------------------------------------------------------------------------  -----------------  -------  --------------  ---------
nqn.1992-04.com.emc:nvme:PowerMax_8000:00:000197600238#vmhba32#500009739803bbff:500009739803b8c8               258  vmhba32  FC                   true
nqn.1992-04.com.emc:nvme:PowerMax_8000:00:000197600238#vmhba33#500009739803bbff:500009739803b88a               259  vmhba33  FC                   true
nqn.1992-04.com.emc:nvme:PowerMax_8000:00:000197600238#vmhba33#500009739803bbff:500009739803b8ca               262  vmhba33  FC                   true
nqn.1992-04.com.emc:nvme:PowerMax_8000:00:000197600238#vmhba32#500009739803bbff:500009739803b888               263  vmhba32  FC                   true
[root@e2e-14-9732:~]
```

```
esxcli nvme namespace list
```

The console displays result similar to the following:

```
[root@e2e-14-9732:~] esxcli nvme namespace list
Name                                  Controller Number  Namespace ID  Block Size  Capacity in MB
------------------------------------  -----------------  ------------  ----------  --------------
eui.6000097000019760023853303036383 9               259          1673         512           46875
eui.6000097000019760023853303036384 1               259          1674         512           46875
eui.6000097000019760023853303036384 3               259          1676         512           46875
eui.6000097000019760023853303036393 2               259          1682         512           46875
eui.6000097000019760023853303036393 4               259          1684         512           46875
eui.6000097000019760023853303036393 5               259          1685         512           46875
eui.6000097000019760023853303036383 9               258          1673         512           46875
eui.6000097000019760023853303036393 6               259          1686         512           46875
eui.6000097000019760023853303036393 9               259          1689         512           46875
eui.6000097000019760023853303036384 1               258          1674         512           46875
eui.6000097000019760023853303036394 1               259          1690         512           46875
eui.6000097000019760023853303036384 3               258          1676         512           46875
eui.6000097000019760023853303036394 2               259          1691         512           46875
eui.6000097000019760023853303036393 2               258          1682         512           46875
eui.6000097000019760023853303036393 4               258          1684         512           46875
eui.6000097000019760023853303036393 5               258          1685         512           46875
eui.6000097000019760023853303036393 6               258          1686         512           46875
eui.6000097000019760023853303036393 9               258          1689         512           46875
eui.6000097000019760023853303036394 1               258          1690         512           46875
eui.6000097000019760023853303036394 2               258          1691         512           46875
eui.6000097000019760023853303036383 9               262          1673         512           46875
eui.6000097000019760023853303036384 1               262          1674         512           46875
eui.6000097000019760023853303036384 3               262          1676         512           46875
eui.6000097000019760023853303036383 9               263          1673         512           46875
eui.6000097000019760023853303036393 2               262          1682         512           46875
eui.6000097000019760023853303036384 1               263          1674         512           46875
eui.6000097000019760023853303036393 4               262          1684         512           46875
eui.6000097000019760023853303036384 3               263          1676         512           46875
eui.6000097000019760023853303036393 5               262          1685         512           46875
eui.6000097000019760023853303036393 2               263          1682         512           46875
eui.6000097000019760023853303036393 6               262          1686         512           46875
eui.6000097000019760023853303036393 9               262          1689         512           46875
eui.6000097000019760023853303036393 4               263          1684         512           46875
eui.6000097000019760023853303036394 1               262          1690         512           46875
eui.6000097000019760023853303036393 5               263          1685         512           46875
eui.6000097000019760023853303036394 2               262          1691         512           46875
eui.6000097000019760023853303036393 6               263          1686         512           46875
eui.6000097000019760023853303036393 9               263          1689         512           46875
eui.6000097000019760023853303036394 1               263          1690         512           46875
eui.6000097000019760023853303036394 2               263          1691         512           46875
```

5. Manually discover fabric and connect to NVMe controllers if host reboot is not allowed (optional).

```
[root@e2e-14-9732:~] esxcli nvme fabrics discover -a vmhba33 -w 50:00:09:73:98:03:b8:8a -W 50:00:09:73:98:03:bb:ff
Transport Type Address Family Subsystem Type Controller ID Admin Queue Max Size Transport Address Transport Service ID Subsystem NQN Connected
-------------- -------------- -------------- ------------- ------------------- ---------------------------------------- ------------------- --------
Fibre Channel Fibre Channel NVM 2438 15 nn-0x500009739803bbff:pn-0x500009739803b88a none nqn.1992-04.com.emc:nvme:PowerMax_8000:00:000197600238 true
[root@e2e-14-9732:~]
[root@e2e-14-9732:~] esxcli nvme fabrics discover -a vmhba33 -w 50:00:09:73:98:03:b8:ca -W 50:00:09:73:98:03:bb:ff
Transport Type Address Family Subsystem Type Controller ID Admin Queue Max Size Transport Address Transport Service ID Subsystem NQN Connected
-------------- -------------- -------------- ------------- ------------------- ---------------------------------------- ------------------- --------
Fibre Channel Fibre Channel NVM 3462 15 nn-0x500009739803bbff:pn-0x500009739803b8ca none nqn.1992-04.com.emc:nvme:PowerMax_8000:00:000197600238 true
[root@e2e-14-9732:~]
[root@e2e-14-9732:~] esxcli nvme fabrics discover -a vmhba32 -w 50:00:09:73:98:03:b8:c8 -W 50:00:09:73:98:03:bb:ff
Transport Type Address Family Subsystem Type Controller ID Admin Queue Max Size Transport Address Transport Service ID Subsystem NQN Connected
-------------- -------------- -------------- ------------- ------------------- ---------------------------------------- ------------------- --------
Fibre Channel Fibre Channel NVM 3331 15 nn-0x500009739803bbff:pn-0x500009739803b8c8 none nqn.1992-04.com.emc:nvme:PowerMax_8000:00:000197600238 true
[root@e2e-14-9732:~]
[root@e2e-14-9732:~] esxcli nvme fabrics discover -a vmhba32 -w 50:00:09:73:98:03:b8:88 -W 50:00:09:73:98:03:bb:ff
Transport Type Address Family Subsystem Type Controller ID Admin Queue Max Size Transport Address Transport Service ID Subsystem NQN Connected
-------------- -------------- -------------- ------------- ------------------- ---------------------------------------- ------------------- --------
Fibre Channel Fibre Channel NVM 2307 15 nn-0x500009739803bbff:pn-0x500009739803b888 none nqn.1992-04.com.emc:nvme:PowerMax_8000:00:000197600238 true
```

Dell EMC PowerMax iSCSI connectivity

Dell EMC PowerMax array supports high-density quad port 10 Gb/s interface module. For detailed iSCSI host connectivity information, see *Dell EMC PowerMax: iSCSI Implementation for Dell EMC Storage Arrays Running PowerMaxOS* available on https://www.dellemc.com/ or Dell EMC Online Support under PowerMax and VMAX All Flash Technical Documentation page.

Midrange Storage

This section describes host connectivity of the Dell EMC Midrange storage arrays, including Unity and PowerStore.

ESXi host in the Unity and VNX series environment

This section gives information about the ESXi host connectivity in the Unity and VNX series environment.

This section describes the following topics:

- Unity and VNX series failover modes
- Unity arrays
- Setting up a host to use Unity VMware vStorage VMFS FC datastores
- Using multipath management software with ESXi hosts
- Configuring the Unity VMware vStorage VMFS datastores for the host
- Manually setting up the host connection to a Unity FC VMFS datastore
- Setting up a host to use Unity vVol datastores

Unity and VNX series failover modes

Unity and VNX series systems with Asymmetric Logical Unit Access (ALUA) mode are supported with ESXi 6.0, ESXi 6.5, and ESXi 6.7. For VNX series systems, the default failover mode for the ESXi host is failover mode 1 with the storage type as Active/Passive. When ESXi host is registered in failover mode 4, ALUA mode is enabled. In such a case, the Unity and VNX series systems behave similarly to an Active or Active array. The ESXi server applies the Round Robin policy to VNX series devices in a VMware native multipathing environment by default.

Unity arrays

For Unity arrays, Dell EMC recommends configuring ESXi with Round Robin Path Selection Plug-in with an IOPS limit of 1. For more details, see the VMware Knowledge Base article 2069356.

iSCSI

When configuring LUNs on ESXi that you access through iSCSI, disable *DelayedACK* on ESXi. For more details, see the VMware Knowledge Base article 1002598.

VMware NFS datastores

When creating NFS-based storage for ESXi, Dell EMC recommends creating VMware NFS Datastores instead of general-purpose NFS file systems. VMware NFS Datastores are optimized to provide better performance with ESXi. Install the VAAI plug-in before creating NFS

Datastores to enable these optimizations. When creating VMware NFS Datastores, Dell EMC recommends using the default 8K Host I/O Size. Choose a different Host I/O Size if all applications that will be hosted in the NFS Datastore primarily use the selected I/O size.

VVols

When configuring VVol NFS datastores, Dell EMC recommends creating at least two vVols-enabled NAS Servers; one on SPA and one on SPB.

Setting up a host to use Unity VMware VMFS FC datastores

Prerequisites for setting up a host to use VMware vStorage VMFS datastores with FC

Before you set up a host to use Unity VMware vStorage VMFS FC datastores, the Unity system host and network requirements that are described in this section must be met. Unity automatically detects VMware ESXi hosts on the SAN. It is possible to add ESXi hosts manually but enabling the storage system to detect ESXi hosts automatically provides the highest level of functionality.

SAN requirements

For a host to connect to FC LUNs and Block VVol datastores on the Unity System:

- The host must be in a SAN environment with the storage system.
- The host must be zoned so that the host and the storage system are visible to each other over the SAN.

For a multipathing environment, each Unity FC LUN for the host must have a minimum of two paths (four paths recommended) associated with it. These paths should be on different switches to ensure high availability.

Path management SAN requirements

When implementing a highly available SAN between a host and the Unity system, ensure that:

- A LUN or VMware vStorage VMFS Datastore is visible to both SPs.
- You can configure multiple paths for a LUN. These paths should be associated with separate physical ports on the same SP.
- Each LUN must present the same LUN ID to all hosts.

NOTE: If the host connects to both SPs and has the required multipath software, directly attaching a host to a storage system is supported.

Using multipath management software with ESXi hosts

This section describes how you can use multipath management software with ESXi host.

Multipath management software manages the paths between the host and the Unity system to provide access to the Unity storage if one of the paths fails. The following types of multipath management software are available for an ESXi host that is connected to a Unity system:

- ESXi Native Multipathing (NMP) on any ESXi host
- Dell EMC PowerPath/VE software on an ESXi 6.0 or later host

For the supported versions of the PowerPath/VE software, see the Unity Support Matrix on the Dell EMC E-Lab Navigator.

Setting up a Unity system for multipath management software

For a Unity system to operate with hosts running multipath management software, each FC LUN on the Unity system should be associated with multiple paths.

Install PowerPath/VE software

This topic describes the procedure to install the PowerPath/VE software.

Steps

1. On the ESXi host, download the latest PowerPath/VE version from the PowerPath/VE software download section on Dell EMC Online Support.
2. Install PowerPath/VE as described in the appropriate *PowerPath/VE Installation and Administration Guide* for the operating system of host available on Dell EMC Online Support.

   If the host is running the most recent version, and a patch exists for this version, install the patch as described in the readme file.
3. After the installation is complete, reboot the host.

4. When the host is restarted, verify that the PowerPath/VE service has started.

Configuring ESXi host native failover

ESXi hosts include native failover for managing the I/O paths between the server and storage system. Native failover provides multiple paths from the server to the storage system. To use the ESXi host native failover with your storage system, you must implement one of the following failover policies:

- Round Robin (default)
- Fixed with failover mode
- Most Recently Used (MRU)

For more information about these policies, see the VMware ESXi configuration information about the VMware Website.

Configure the native failover policy for FC connections

This topic describes the procedure to configure the native failover policy for FC connection.

Steps

1. Log into VMware vSphere client as administrator.
2. From the Inventory panel, select **server**, and click the **Configuration** tab.
3. Under **Hardware**, click **Storage** and select **Datastore (LUN)**.
4. Click **Properties** > **Manage Paths**.
5. In the **Manage Paths** page, under **Policy**, select the policy that you want as follows:
   - **Fixed (VMware)** for fixed native failover policy
   - **Round Robin (VMware)** for Round Robin native failover policy
   - **Most Recently Used (VMware)** for MRU native failover policy
6. If the policy is not set to the policy you want, in the policy selection dialog, select the correct policy.
7. If you selected the **Fixed (VMware)** policy, under **Paths**, select the preferred path.

   You can statically balance the load using the fixed policy by selecting different paths for each datastore. To choose a different path as preferred, right-click the path and click **preferred**.
8. Click **Close**.

Configure the Unity VMware VMFS datastores for the host

Perform the following steps to configure the Unity VMware datastores for the host:

Steps

1. Use Unisphere UI or CLI to perform the following substeps:
   a. Discover VMware ESXi hosts.
   b. Create Unity VMFS datastores.

   For more information about these tasks, see the *Unisphere Online Help* or the *Unisphere CLI User Guide*.

   NOTE: When you create Unity VMFS datastores for discovered VMware ESXi hosts, the system automatically configures the hosts to use the datastores. If you select multiple discovered hosts for the datastores, all the selected hosts are configured automatically. The VMware VMFS datastores are presented directly to the ESXi hosts and not to the hosts virtual machines. If an ESXi host cannot see the VMFS datastore after you create the datastore, manually configure the connection to the VMFS datastore.
2. Search for the VMFS datastores (generic host objects only) and perform the following substeps:
   a. From the **Hosts** and **Clusters panel**, select **server**.
   b. Select **Manage**.
   c. Under **Storage**, select **Storage Adapters**.
   d. Right-click **FC HBA** and then click **Rescan**.

      The **Rescan** dialog box displays.
   e. In the **Rescan** dialog box, select **FC HBA**, and then click **Rescan**.
   f. Under **Storage**, select **Storage Devices**.
3. Click **Rescan**.

Manually setting up the host connection to a Unity FC VMFS datastore

Perform the following steps to manually set up the host connection to a Unity FC VMFS datastore:

Steps

1. Log into VMware vSphere client as an administrator.
2. Rescan for new storage devices as follows:
   a. From the **Inventory** panel, select the server, and click the **Configuration** tab.
   b. Under **Hardware**, click **Storage Adapters**.
      A list of adapters displays.
   c. From the list of adapters, select **FC HBA**, and then click **Rescan**.
      The **Rescan** dialog box displays.
   d. From the **Rescan** dialog box, select **FC HBA**, and then click **Rescan**.
   e. In the **Rescan** dialog box, select both **Scan for New Storage Devices** and **Scan for New VMFS Volumes**, and click **OK**.
3. Add each VMFS datastore to the ESXi host as follows:
   a. From the **Inventory** panel, select the host and click the **Configuration** tab.
   b. Under **Hardware**, click **Storage**, and click **Add Storage**.
   c. On the **Select Disk/LUN** page, select the **Unity VMFS** datastore that you want to use for the datastore, and click **Next**.
   d. On the Current Disk Layout page, review the current virtual drive layout, and click **Next**.
   e. On the **Disk/LUN-Properties** page, enter the exact same name that was used to create the datastore on the storage system. You can find this name using Unisphere.
   f. On the **Disk/LUN-Formatting** page, if needed, adjust the file system values and the capacity for the datastore, and click **Next**.
   g. On the **Ready to Complete** page, review the datastore information, and click **Finish**.

Results

The datastore (VMFS volume) is created on the Unity VMFS datastore for the ESXi host.

Setting up a host to use Unity VVol datastores

The following subsections explain the procedures to set up host Unity datastores:

- Adding the storage system as a VASA provider in vSphere
- Configuring Unity file vVol datastores for the host

Adding the storage system as a VASA provider in vSphere

Perform the following steps to add the storage system as a VASA provider in vSphere:

Steps

1. In vSphere, click **Storage** and click **Manage**.
2. Select **Storage provider**, and then click **Add icon**.
3. Enter the name of the storage system.
   It is recommended that you use the same system name as Unisphere.
4. Enter the VASA provider URL for the Unity system. Use the following format: `https://<management IP address>:8443/vasa/version.xml`.
5. Enter the Unisphere credentials and click **OK**.
   After you click OK, it may take few minutes for the registration to complete.

Configuring Unity file VVol datastores for the host

Use the Unisphere UI or CLI to perform the following steps:

Steps

1. Discover VMware ESXi hosts.

2. Create capability profiles, NAS protocol endpoint servers (VVols-enabled NAS servers), and File VVol datastores.
   You must create a VVol-enabled NAS server on each SP. For more details, see the *Unisphere Online Help* or *Unisphere CLI Guide*.

Setting up the connection to a File vVol datastore

Prerequisites

If the ESXi host cannot detect the storage system File VVol datastore, or if you are using generic host objects, you must manually set up the connection.

Steps

1. Log into vSphere as an administrator.
2. Perform the following steps for each File vVol datastore:
   a. On the **Inventory** panel, select the host and click **Configuration** tab.
   b. On the **Hardware** panel, click **Storage**, and then **Add Storage**.
   c. Select **Network File System** as the file storage type.
   d. Enter the following information:
      - For folder, enter the path to the Unity share.
      - For name, enter the name of the new VMware datastore.
   e. Perform a rescan in vSphere to ensure that the VMware NFS datastore is visible to the ESXi host.

Adding an iSCSI target to Block vVol datastores

Perform the following steps to add an iSCSI target to Block vVol datastores:

Steps

1. Under **Hosts and Clusters**, click **Manage** > **Storage**.
2. Under **Storage Adapters**, select the **iSCSI Software Adapter**.
3. On the **Targets** tab, click **Add**.

Configuring the Unity Block VVol datastores for the host

Perform the following steps to configure the Unity Block VVol datastores for the host:

Prerequisites

For iSCSI only: Make sure that the software iSCSI adapter is already added. If not, add the software iSCSI adapter. For details, see vSphere documentation.

Steps

1. Enter the **iSCSI** details and click **OK**.
2. Use Unisphere UI or CLI and perform the following steps:
   a. Discover VMware ESXi hosts.
   b. Create Unity Block VVol datastores.
      For more details, see the *Unisphere Online Help* or *Unisphere CLI User Guide*.
3. Rescan for the VMFS datastores as follows:
   a. In the **Hosts and Clusters** panel, click **server**.
   b. Click **Manage**.
   c. Under **Storage**, click **Storage Adapters**.
   d. For iSCSI, perform the following steps:
      i. Select the **iSCSI** storage adapter in the list of adapters.
      ii. Click **Rescan icon**.
   e. For FC, perform the following steps:
      i. Right-click **FC HBA** and then click **Rescan**.

ii. In the **Rescan** dialog box, select **FC HBA**, and then click **Rescan**.
iii. Under **Storage**, select **Storage Devices**.
iv. Click **Rescan icon**.

Storage configuration

General recommendations for storage pools

Dell EMC Unity supports two types of storage pools: traditional pools and dynamic pools. The following recommendations are applicable to both types of pool.

Dell EMC recommends using fewer storage pools within Dell EMC Unity to reduce complexity and increase flexibility. However, it may be appropriate to configure multiple storage pools to:

- Separate workloads with different I/O profiles
- Separate pools where FAST Cache is active and not active
- Dedicate resources to meet specific performance goals
- Separate resources for multi-tenancy
- Create smaller failure domains

Storage pool capacity

Storage pool capacity is used for multiple purposes:

- To store all data written into storage objects (LUNs, file systems, data stores, and VVols) in that pool.
- To store data that is needed for Snapshots of storage objects in that pool.
- To track changes to replicated storage objects in that pool
- To perform efficient data relocations for FAST VP

Storage pools must maintain free capacity to operate properly. By default, Dell EMC Unity will raise an alert if a storage pool has less than 30% free capacity, and will begin to automatically invalidate Snapshots and Replication sessions if the storage pool has less than 5% free capacity. Dell EMC recommends that a storage pool always have at least 10% free capacity.

Raid protection

Dell EMC Unity applies RAID protection to the storage pool to protect user data against drive failures. Choose the RAID type that best suits your needs for performance, protection, and cost.

- RAID-1/0 provides the highest level of performance from a given set of drive resources, with the lowest CPU requirements; however, only 50% of the total drive capacity is usable.
- RAID-5 provides the best usable capacity from a set of drive resources, but at lower overall performance and availability than RAID-1/0.
- RAID-6 provides better availability than RAID-5 and better usable capacity than RAID-1/0, but has the lowest performance potential of the three RAID types.

Traditional pools

Traditional Storage Pools apply RAID protection to individual groups of drives within the storage pool. Traditional pools are the only type of pool available on Dell EMC Unity hybrid systems, and are also available on all-Flash systems.

Raid protection

For traditional pools, Dell EMC recommends RAID-5 for drives in Extreme Performance and Performance tiers, and RAID-6 for drives in the Capacity tier.

Assuming that roughly the same number of drives will be configured in a traditional pool, Dell EMC recommends smaller RAID widths as providing the best performance and availability, at the cost of slightly less usable capacity.

Example: When configuring a traditional pool tier with RAID-6, use 4+2 or 6+2 as opposed to 10+2 or 14+2.

When choosing RAID-1/0, 1+1 can provide better performance with the same availability and usable capacity as larger RAID widths (assuming that the same total number of drives are used), and also provides more flexibility.

All-flash pool

All-flash pools provide the highest level of performance in Dell EMC Unity. Use an all-flash pool when the application requires the highest storage performance at the lowest response time.

Snapshots and Replication operate most efficiently in all-flash pools. Data Reduction is only supported in an all-flash pool.

FAST Cache and FAST VP are not applicable to all-flash pools.

Dell EMC recommends using only a single drive size and a single RAID width within an all-flash pool.

Dynamic pools

Dynamic Storage Pools apply RAID protection to groups of drive extents from drives within the pool, and allow for greater flexibility in managing and expanding the pool. Dynamic pools are only available on Dell EMC Unity all-Flash systems, and therefore must be all-Flash pools; dynamic pools cannot be built with HDDs.

RAID protection

At the time of creation, dynamic pools use the largest RAID width possible with the number of drives that are specified, up to the following maximum widths:

- RAID-1/0: 4+4
- RAID-5: 12+1
- RAID-6: 14+2

With dynamic pools, there is no performance or availability advantage to smaller RAID widths. To maximize usable capacity with parity RAID, Dell EMC recommends to initially create the pool with enough drives to guarantee the largest possible RAID width.

- For RAID-5, initially create the pool with at least 14 drives.
- For RAID-6, initially create the pool with at least 17 drives.

Spare capacity

Hot spares are not needed with dynamic pools. A dynamic pool automatically reserves the capacity of one drive, as spare space in the pool, for every 32 drives. If a drive fails, the data that was on the failed drive is rebuilt into the spare capacity on the other drives in the pool. Also, unbound drives of the appropriate type can be used to replenish the spare capacity of a pool, after the pool rebuild has occurred.

Example: For an All-Flash pool, use only 1.6 TB SAS Flash 3 drives, and configure them all with RAID-5 8+1.

Hybrid pool

Hybrid pools can contain HDDs (SAS and NL-SAS drives) and flash drive, and can contain more than one type of drive technology in different tiers. Hybrid pools typically provide greater capacity at a lower cost than all-flash pools, but also typically have lower overall performance and higher response times. Use hybrid pools for applications that do not require consistently low response times, or that have large amounts of mostly inactive data.

Performance of a hybrid pool can be improved by increasing the amount of capacity in the flash drive tier, so that more of the active dataset resides on and is serviced by the flash drives. See the FAST VP section.

Hybrid pools can have up to three tiers (Extreme Performance, Performance, and Capacity). Dell EMC recommends using only a single drive speed, size, and RAID width within each tier of a hybrid pool.

Example:

- For the Extreme Performance tier, use only 800 GB SAS flash 2 drives, and configure them all with RAID-5 8+1.
- For the Performance tier, use only 1.2 TB SAS 10K RPM drives, and configure them with RAID-5 4+1.
- For the Capacity tier, use only 6 TB NL-SAS drives, and configure them all with RAID-6 6+2.

Storage object types

By default, Dell EMC Unity creates thin storage objects. Thin storage objects are virtually provisioned and space efficient. In general, Dell EMC recommends using thin storage objects, as they provide the best capacity utilization, and are required for most features. Thin storage objects are recommended when any of the following features will be used:

- Data Reduction
- Snapshots
- Thin Clones

- Asynchronous Replication

Thick storage objects will reserve capacity from the storage pool, and dedicate it to that particular storage object. Thick storage objects guarantee that all advertised capacity is available for that object. Thick storage objects are not space efficient, and therefore do not support the use of space-efficient features. If it is required to enable a space-efficient feature on a thick storage object, it is recommended to first migrate the thick storage object to a thin storage object, and enable the feature during the migration (for Data Reduction) or after migration has completed (for Snapshots, Thin Clones, and Asynchronous Replication).

In addition to capacity for storing data, storage objects also require pool capacity for metadata overhead. The overhead percentage is greater on smaller storage objects. For better capacity utilization, Dell EMC recommends configuring storage objects that are at least 100GB in size, and preferably at least 1TB in size.

Features

FAST VP

Fully Automated Storage Tiering (FAST) for Virtual Pools (VP) accelerates performance of a specific storage pool by automatically moving data within that pool to the appropriate drive technology, based on data access patterns. FAST VP is applicable to hybrid pools only within a Dell EMC Unity hybrid system.

The default and recommended FAST VP policy for all storage objects is **Start High then Auto-tier**. This policy places initial allocations for the storage object in the highest tier available, and monitors activity to this storage object to determine the correct placement of data as it ages.

FAST VP is most effective if data relocations occur during or immediately after normal daily processing. Dell EMC recommends scheduling FAST VP relocations to occur before backups or nightly batch processing. For applications which are continuously active, consider configuring FAST VP relocations to run constantly.

Dell EMC recommends maintaining at least 10% free capacity in storage pools, so that FAST VP relocations can occur efficiently. FAST VP relocations cannot occur if the storage pool has no free space.

FAST Cache

FAST Cache is a single global resource that can improve performance of one or more hybrid pools within a Dell EMC Unity hybrid system. FAST Cache can only be created with SAS Flash 2 drives, and is only applicable to hybrid pools. Dell EMC recommends to place a Flash tier in the hybrid pool before configuring FAST Cache on the pool. FAST Cache can improve access to data that is resident in the HDD tiers of the pool.

Enable FAST Cache on the hybrid pool if the workload in that pool is highly transactional, and has a high degree of locality that changes rapidly.

For applications that use larger I/O sizes, have lower skew, or do not change locality as quickly, it may be more beneficial to increase the size of the Flash tier rather than enable FAST Cache.

FAST Cache can increase the IOPS achievable from the Dell EMC Unity system, and this will most likely result in higher CPU utilization (to service the additional I/O). Before enabling FAST Cache on additional pools or expanding the size of an existing FAST Cache, monitor the average system CPU utilization to determine if the system can accommodate the additional load. See Table 3 for recommendations.

Data Reduction

Dell EMC Unity Data Reduction by compression is available for Block LUNs and VMFS datastores in an all-flash pool starting with Dell EMC Unity OE 4.1. Data reduction via compression is available for file systems and NFS datastores in an all-flash pool starting with Dell EMC Unity OE 4.2. Beginning with Dell EMC Unity OE 4.3, data reduction includes both compression and deduplication.

Be aware that data reduction increases the overall CPU load on the system when storage objects service reads or writes of reduceable data, and may increase latency. Before enabling data reduction on a storage object, Dell EMC recommends to monitor the system and ensure that the system has available resources to support data reduction (See Table 3 to the Hardware Capability Guidelines). Enable data reduction on a few storage objects at a time, and then monitor the system to be sure it is still within recommended operating ranges, before enabling data reduction on more storage objects.

For new storage objects, or storage objects that are populated by migrating data from another source, Dell EMC recommends to create the storage object with data reduction enabled, before writing any data. This provides maximum space savings with minimal system impact.

Advanced Deduplication

Dell EMC Unity Advanced Deduplication is an optional extension to Data Reduction, that you can enable to increase the capacity efficiency of data reduction enabled storage objects. Beginning with Dell EMC Unity OE 4.5, advanced deduplication is available for storage objects in dynamic pools on Dell EMC Unity 450F, 550F, and 650F All-Flash systems.

As with data reduction, advanced deduplication is only applied to data when it is written to the storage object. LUN Move can be utilized to deduplicate existing data on Block storage objects.

For new storage objects, or storage objects that will be populated by migrating data from another source, it is recommended to create the storage object with advanced deduplication enabled, before writing any data. This provides maximum space savings with minimal system impact.

Snapshots

Dell EMC recommends including a Flash tier in a hybrid pool where snapshots will be active.

Snapshots increase the overall CPU load on the system, and increase the overall drive IOPS in the storage pool. Snapshots also use pool capacity to store the older data being tracked by the snapshot, which increases the amount of capacity used in the pool, until the snapshot is deleted. Consider the overhead of snapshots when planning both performance and capacity requirements for the storage pool.

Before enabling snapshots on a storage object, it is recommended to monitor the system and ensure that existing resources can meet the additional workload requirements (See Table 2 for Hardware Capability Guidelines). Enable snapshots on a few storage objects at a time, and then monitor the system to be sure it is still within recommended operating ranges, before enabling more snapshots.

Dell EMC recommends to stagger snapshot operations (creation, deletion, and so on). This can be accomplished by using different snapshot schedules for different sets of storage objects. It is also recommended to schedule snapshot operations after any FAST VP relocations have completed.

Snapshots are deleted by the system asynchronously; when a snapshot is in the process of being deleted, it will be marked as *Destroying*. If the system is accumulating Destroying snapshots over time, it may be an indication that existing snapshot schedules are too aggressive; taking snapshots less frequently may provide more predictable levels of performance. Dell EMC Unity will throttle snapshot delete operations to reduce the impact to host I/O. Snapshot deletes will occur more quickly during periods of low system utilization.

Thin Clones

Dell EMC recommends including a flash tier in a hybrid pool where thin clones will be active.

Thin clones use snapshot technology to provide space-efficient clones of block objects. Consider the overhead of snapshots when planning performance and capacity requirements for a storage pool which will have thin clones.

Asynchronous replication

Dell EMC recommends including a Flash tier in a hybrid pool where asynchronous replication is active. This is applicable to both the source and the destination pools.

Dell EMC recommends configuring multiple replication interfaces per SP, and distributing replication sessions across them. Link Aggregation Control Protocol (LACP) can also be used to aggregate bandwidth for a replication interface. Configure Jumbo frames (MTU 9000) when possible.

Asynchronous replication takes snapshots on the replicated storage objects to create the point-in-time copy, determine the changed data to transfer, and maintain consistency during the transfer. Consider the overhead of snapshots when planning performance and capacity requirements for a storage pool that has replicated objects.

When possible, fill the source storage object with data before creating the replication session. The data will then be transmitted to the destination storage object during initial synchronization. This is typically the fastest way to populate the destination storage object with asynchronous replication.

Setting smaller RPO values on replication sessions will not make them transfer data more quickly; but smaller RPOs result in more frequent snapshot operations. Choosing larger RPOs, or manually synchronizing during nonproduction hours, may provide more predictable levels of performance.

Synchronous replication/Metrosync for file

Dell EMC recommends including a Flash tier in a hybrid pool where synchronous replication will be active. This is applicable to both the source and the destination pools.

Synchronous replication transfers data to the remote system over the first Fibre Channel port on each SP. When planning to use synchronous replication, it may be appropriate to reduce the number of host connections on this port. When the CNA ports are configured as FC, CNA port 4 is defined as the replication port. If the CNA ports are configured as Ethernet, then port 0 of the lowest numbered FC I/O Module is the replication port.

When possible, create the synchronous replication session before filling the source storage object with data, as this alleviates the need to perform initial synchronization of the replication session. This is typically the fastest way to populate the destination storage object with synchronous replication.

When sizing a disaster recovery solution using synchronous replication, consider provisioning a destination system that has similar performance capabilities as the source system. This can help maintain the same level of application performance after a failover event.

SAN Copy

SAN Copy provides one-time migration of Block resources from a third-party array, using either iSCSI or FC connections. When using FC, note that SAN Copy must use different ports than the FC ports which are designated for Synchronous Replication. This is true even if Synchronous Replication is not actively being used.

To lessen the impact of SAN Copy migrations on other host activity, consider reducing the number of host connections on the FC ports used for SAN Copy.

NDMP

Dell EMC Unity supports 2-way NDMP for file data, which enables the system to send file data directly to a backup device using FC connections. Make sure that NDMP uses different ports than the FC ports which are designated for Synchronous Replication. This is true even if Synchronous Replication is not actively being used.

To lessen the impact of 2-way NDMP backups on other host activity, consider reducing the number of host connections on the FC ports that are used for NDMP.

Data at Rest Encryption

Data at Rest Encryption (D@RE) is Controller Based Encryption that does not impact performance; therefore Dell EMC recommends ordering Dell EMC Unity systems as encryption-enabled, if appropriate for your environment.

NOTE: Encryption can only be enabled at the time of system installation with the appropriate license.

If encryption is enabled, Dell EMC recommends making external backups of the encryption keys after system installation, and immediately following any change in the system's drives (such as, creating or expanding a storage pool, adding new drives, replacing a faulted drive, and so on).

Host I/O limits

Dell EMC recommends setting Host I/O Limits on workloads which might monopolize pool resources and starve other applications of their required performance. Consider some of the following opportunities to utilize Host I/O Limits:

- Limit the bandwidth available to large-block applications, such as backup jobs, which may be increasing the latency on other small-block workloads.
- Limit the IOPS capability of Thin Clones which are used in Test/Dev environments, so that they do not impact the capability of the associated source objects.
- Limit the IOPS / bandwidth that is available to non-essential applications that are hosted on the same Dell EMC Unity system as your critical applications.

Application considerations

Host alignment for block LUNs

Alignment only needs to be done for host operating systems which still use a 63-block disk header. If alignment is required, perform the alignment using a host-based method, and align with a 1MB offset.

See the *Host Connectivity Guide* on Dell EMC Online Support to determine if alignment is required for your operating system, and how to perform the alignment.

Hosts

Under the **ACCESS** category in the main navigation menu, users can configure hosts (Windows or Linux/UNIX) for storage access. Before a network host can access block storage or NFS file systems, the user must define a configuration for the host and associate it with a storage resource. SMB file systems can automatically be accessed by authorized users once provisioned. Users can use the Hosts page, as shown in the following figure to configure host configurations. This can be done on an individual host-by-host basis or through subnet and netgroup configurations that allow access to multiple hosts or network segments. For block resources, before the user starts to configure a host, the user should ensure that initiator interfaces are configured and initiator registration completed. Once a host configuration is completed, users can go to the properties of a storage resource and specify the hosts, subnets, or netgroups from which they want the resource to be accessed.

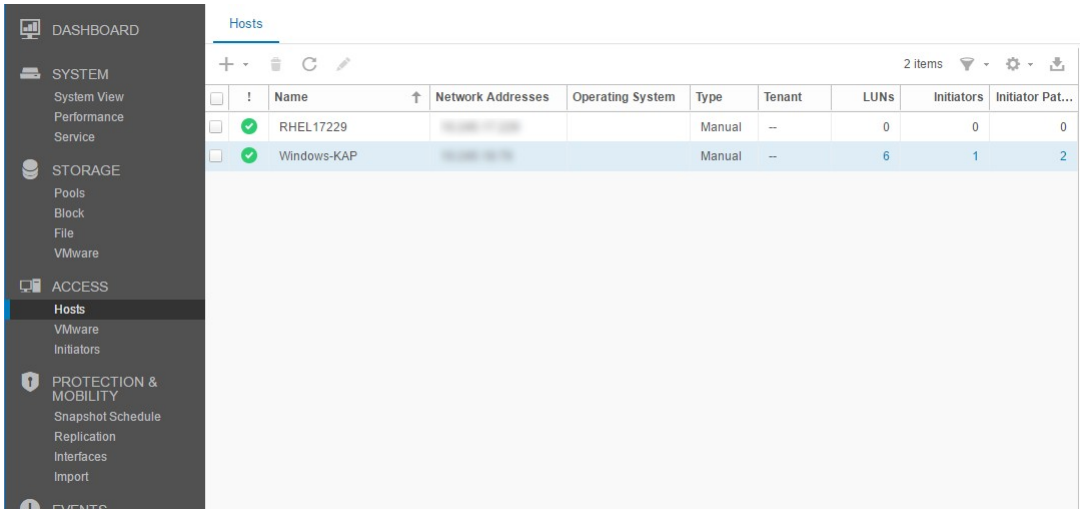


Figure 7. Hosts

VMware (ACCESS)

The VMware host access page is specifically for VMware ESXi hosts and their associated vCenter servers. Unisphere provides VMware discovery capabilities through the VMware page, as shown in the following figure. These discovery capabilities collect virtual machine and datastore storage details from vSphere and display them in the context of the storage system. Imported VMware hosts automatically register their initiators, allowing for ease of management. The vCenters tab allows users to add a vCenter and associated ESXi hosts in a single workflow, while the ESXi hosts tab allows users to add standalone ESXi hosts as needed. The Virtual Machines tab and Virtual Drives tab display imported information about virtual machines and their VMDKs from any added ESXi host.

For more information about VMware access and integration capabilities, see the *Dell EMC Unity: Virtualization Integration white paper* on Dell EMC Online Support.

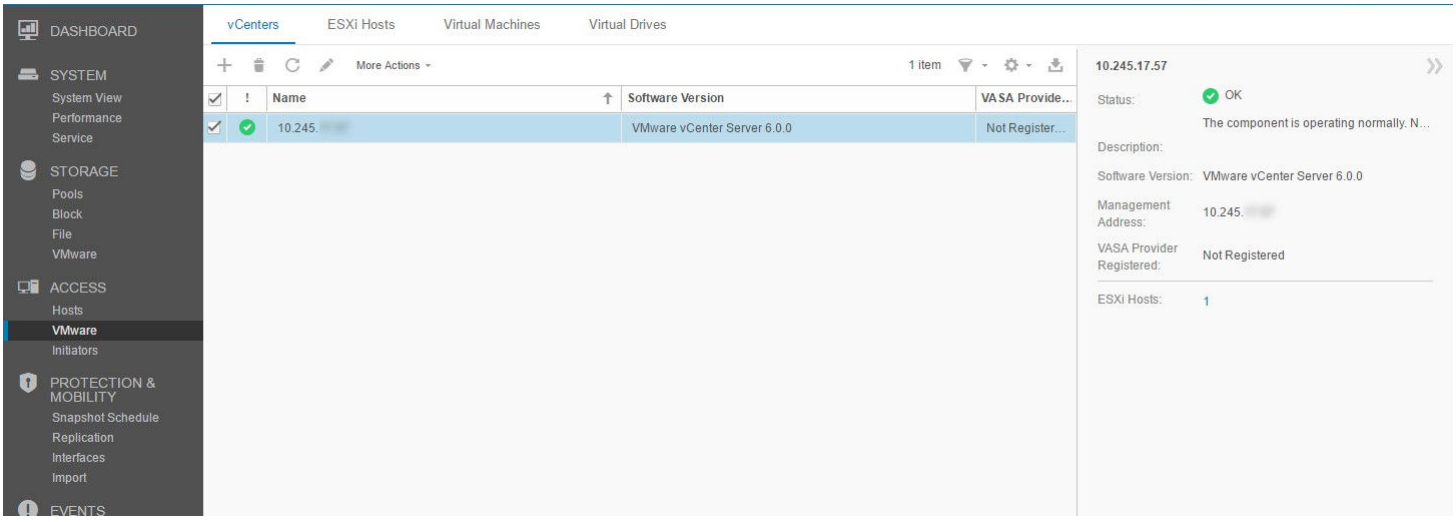


Figure 8. VMware (Access)

Initiators

To ensure that hosts can access block storage resources, the user must register initiators between the storage system and configured hosts. On the Initiators page, as shown in the following figure, users can manually register one or more Fibre Channel or iSCSI initiators. Initiators are endpoints from which Fibre Channel and iSCSI sessions originate, where each initiator is uniquely identified by its World Wide Name (WWN) or iSCSI Qualified Name (IQN). The link between a host initiator and a target port on the storage system is called the initiator path. Each initiator can be associated with multiple initiator paths. The Initiator Paths tab shows all data paths that are currently available to the initiators connected to the system either by FC or iSCSI. For iSCSI paths to show up, iSCSI interfaces must be configured on the Block Page. These initiators can then be discovered and registered by hosts using the iSCSI initiator tool (that is, the Microsoft iSCSI Initiator). For Fibre Channel paths, FC zoning on the appropriate switch is needed for the initiator paths to be seen as available by the system. Once the paths are available, users can configure their connected hosts on the Hosts Page.

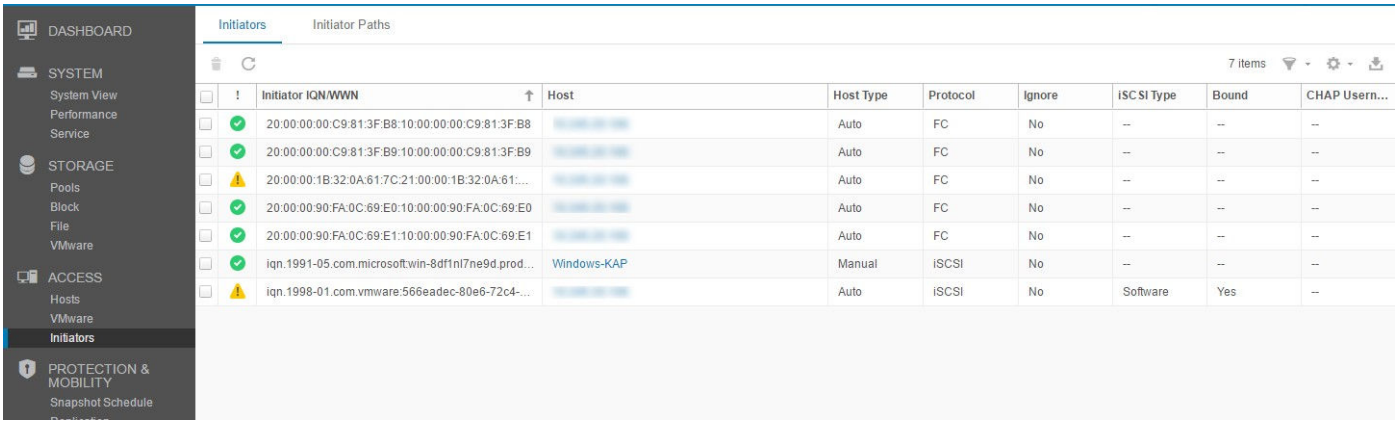


Figure 9. Initiators

With the release of Dell EMC Unity OE version 4.3, Initiators can now have advanced settings customized through Unisphere. In order to access these settings, select an Initiator and then click the pencil icon to bring up the Edit Initiator window. Clicking Advanced at the bottom to reveal the Initiator Source Type, Fail-over Mode, Unit Serial Number, and LunZ Enabled settings, as shown in the following figure. For more information about configuring Host Initiator Parameters, please reference the Online Help through Unisphere.

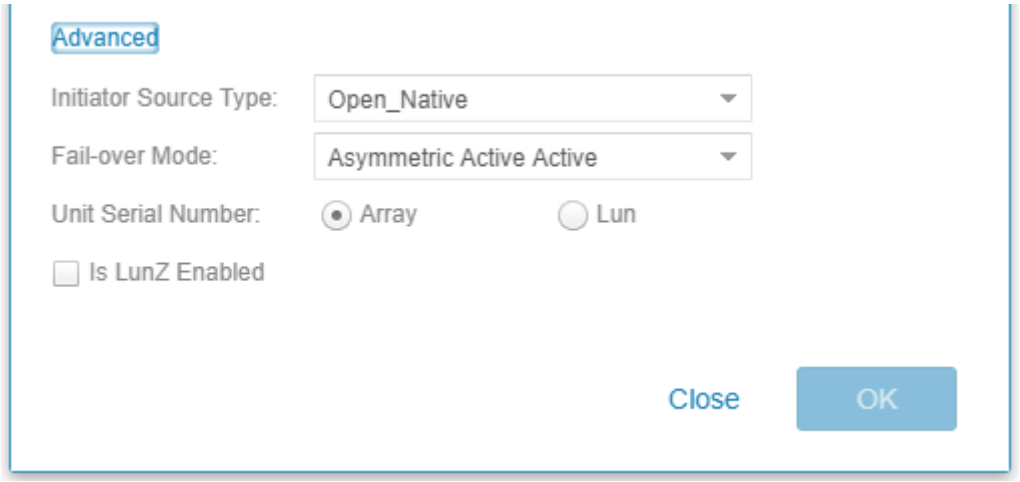


Figure 10. Edit Initiator - Advanced

Snapshot schedule

Dell EMC Unity enables you to take point-in-time snapshots for all storage resources (block or file) to meet protection and recovery requirements in the event of corruption or accidental deletion.

The Snapshot Schedule page, as shown in the figure , enables users to set the system to periodically take snapshots of storage resources automatically. Automating these administrative tasks takes away some of the management aspects of data protection. After enabling a snapshot schedule for a resource, each snapshot that is taken is time-stamped with the date and time of when it was created, and contains a point-in-time image of the data in the storage resource. The default snapshot schedules available on the system are:

- Default protection - A snapshot is taken at 08:00 (UTC) every day, and the snapshot is retained for 2 days.

- Protection with shorter retention - A snapshot is taken at 08:00 (UTC) every day, and the snapshot is retained for 1 day.
- Protection with longer retention - A snapshot is taken at 08:00 (UTC) every day, and the snapshot is retained for 7 days.

NOTE: Times are displayed in a user's local time in a 12-hour format and that default snapshot schedules cannot be modified; but custom snapshot schedules can be configured by selecting the intervals, times, and days for the system to take snapshots regularly.

With the Dell EMC Unity OE version 4.4 or later, user-defined Snapshot Schedules can be replicated using the Synchronous Replication connection that is established between two physical systems. Reference the new **Sync Replicated** column in the Snapshot Schedule page, as shown in the following figure. Applying a replicated Snapshot Schedule is only enabled in synchronously replicated file resources.

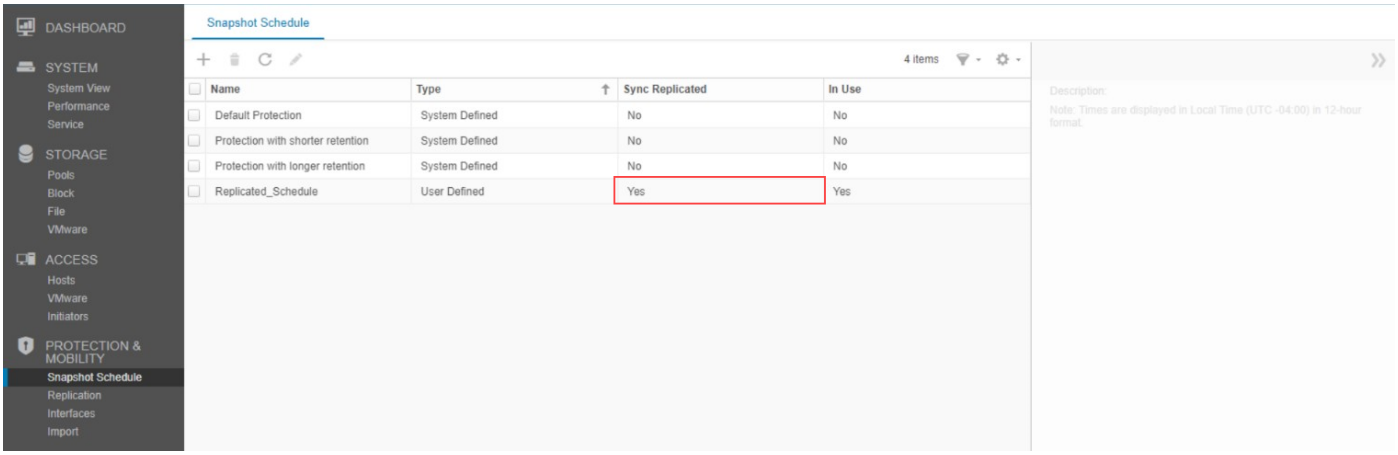


Figure 11. Snapshot schedule

For more information about the snapshot technology available on Dell EMC Unity systems, see the *Dell EMC Unity: Snapshots and Thin Clones* and *Dell EMC Unity: MetroSync for File* white papers on Dell EMC Online Support.

Front-end connectivity

Dell EMC Unity provides multiple options for front-end connectivity, using on-board ports directly on the DPE, and using optional I/O Modules. This section discusses recommendations for the different types of connectivity.

In general, front-end ports need to be connected and configured symmetrically across the two storage processors (SPs), to facilitate high availability and continued connectivity if there is SP failure.

Example - A NAS Server is configured so that NAS clients connect using port 0 of the first I/O Module on SPA; therefore port 0 of the first I/O Module on SPB must be cabled so that it is accessible to the same networks.

For best performance, Dell EMC recommends using all front-end ports that are installed in the system, so that workload is spread across as many resources as possible.

Example - configuring the 4-port Fibre Channel I/O Module, zone different hosts to different ports so that all eight ports across the 2 SPs are used; do not zone all hosts to the first port of each I/O Module.

Fibre Channel

When configured for Fibre Channel (FC), Dell EMC Unity CNA ports and I/O Module ports can be configured with 8 GB or 16 GB SFPs. All FC ports can negotiate to lower speeds. 16 GB FC is recommended for the best performance.

Dell EMC recommends single-initiator zoning when creating zone sets. For high availability purposes, a single host initiator should be zoned to at least one port from SPA and one port from SPB. For load balancing on a single SP, the host initiator can be zoned to two ports from SPA and two ports from SPB. When zoning additional host initiators, zone them to different SP ports when possible, to spread the load across all available SP ports.

Utilize multipathing software on hosts that are connected using FC, such as Dell EMC PowerPath, which coordinates with the Dell EMC Unity system to provide path redundancy and load balancing.

iSCSI

Dell EMC Unity supports iSCSI connections on multiple 1 Gb/s and 10 GB/s port options. 10GBase-T ports can autonegotiate to 1 GB/s speeds. 10 GB/s is recommended for the best performance. If possible, configure Jumbo frames (MTU 9000) on all ports in the end-to-end network, to provide the best performance.

To achieve optimal iSCSI performance, use separate networks and VLANs to segregate iSCSI traffic from normal network traffic. Configure standard 802.3x Flow Control (Pause or Link Pause) on all iSCSI Initiator and Target ports that are connected to the dedicated iSCSI VLAN.

Dell EMC Unity supports 10 GbE and 1GBase-T ports that provide iSCSI offload. Specifically, the CNA ports (when configured as 10GbE or 1GBase-T) and the 2-port 10GbE I/O Module ports provide iSCSI offload. Using these modules with iSCSI can reduce the protocol load on SP CPUs by 10-20%, so that those cycles can be used for other services.

Utilize multipathing software on hosts that are connected using iSCSI, such as Dell EMC PowerPath, which coordinates with the Dell EMC Unity system to provide path redundancy and load balancing.

Network-attached storage (NAS)

Dell EMC Unity supports NAS (NFS, FTP, and/or SMB) connections on multiple 1 GB/s and 10 GB/s port options. 10GBase-T ports can auto-negotiate to 1 GB/s speed. 10 GB/s is recommended for the best performance. If possible, configure Jumbo frames (MTU 9000) on all ports in the end-to-end network, to provide the best performance.

Dell EMC recommends configuring standard 802.3x Flow Control (Pause or Link Pause) on all storage ports, switch ports, and client ports that are used for NAS connectivity.

Dell EMC Unity provides network redundancy for NAS using Link Aggregation Control Protocol (LACP) and Fail-Safe Networking (FSN). Combine FSN and LACP with redundant switches to provide the highest network availability. In addition to redundancy, LACP can also improve performance with multiple 1GBase-T connections, by aggregating bandwidth. LACP can be configured across any Ethernet ports that have the same speed, duplex, and MTU.

NOTE: LACP cannot be enabled on ports that are also used for iSCSI connections.

While LACP creates a link aggregation with multiple active links, FSN provides redundancy by configuring a primary link and a standby link. The standby link is inactive unless the entire primary link fails. If FSN is configured with links of different performance capability (such as a link aggregation of 10 GB/s ports, and a stand-alone 1 GB/s port), Dell EMC recommends that you configure the highest performing link as the primary.

NAS Servers are assigned to a single SP. All file systems that are serviced by that NAS Server have I/O processed by the SP on which the NAS Server is resident. For load-balancing, Dell EMC recommends that you create at least two NAS Servers per Dell EMC Unity system: one on SPA, and one on SPB. Assign file systems to each NAS Server such that front-end workload is approximately the same for each SP.

Connectivity Options

The following tables provide maximum expected IOPS and bandwidth from the different ports that are available in the Dell EMC Unity system. (The capability of a port does not guarantee that the system can reach that level, nor does it guarantee that performance will scale with additional ports. System capabilities are highly dependent on other configuration parameters.)

SAS ports are used by the SPs to move data to and from the back-end drives; all other ports can be used to provide access to hosts.

Table 3 provides maximum expected IOPS and bandwidth from a 12Gb SAS port. The base Dell EMC Unity configuration contains four ports.

The following tables provide maximum expected IOPS from the different ports that are available in the Dell EMC Unity system. (The capability of a port does not guarantee that the system can reach that level, nor does it guarantee that performance will scale with additional ports. System capabilities are highly dependent on other configuration parameters.)

Table 3. Ports and expected IOPS

Port	Maximum IOPS per Port
16 GB FC CNA or 4-port I/O Module	45,000
8 GB FC CNA	45,000
10 GbE iSCSI CNA or 2-port I/O Module	25,000
10 GbE iSCSI 4-port I/O Module	30,000
10GBase-T iSCSI	30,000

Table 3. Ports and expected IOPS (continued)

Port	Maximum IOPS per Port
On-board or 4-port I/O Module	
1GBase-T iSCSI CNA, On-board or 4-port I/OModule	3,000

The following table provides maximum expected IOPS from the front-end ports which provide File protocols (NFS and SMB).

Table 4. Front-end ports and expected IOPS

Port	Maximum IOPS per Port
10 GbE NAS CNA or 2-port I/O Module	60,000
10 GbE NAS 4-port I/O Module	60,000
10GBase-T NAS On-board or 4-port I/O Module	60,000
1GBase-T NAS CNA, On-board or 4-port I/OModule	6,000

PowerStore

PowerStore provides operational simplicity and agility, utilizing a container-based microservices architecture, advanced storage technologies, and built-in machine learning. It is a versatile platform with a performance-centric design that delivers multi-dimensional scale, always on data reduction, and support for next generation media. It brings the simplicity of public cloud to on-premises infrastructure, streamlining operations with a built-in machine learning engine and seamless automation, while offering predictive analytics to easily monitor, analyze, and troubleshoot the environment. PowerStore is also highly adaptable, providing the flexibility to host specialized workloads directly on the appliance and modernize infrastructure without disruption.

For the most up-to-date support information, see the *Dell EMC Simple Support Matrix*, available on Dell EMC E-Lab Navigator.

Fibre Channel HBA configuration

This section describes the recommended configuration that must be applied when attaching hosts to PowerStore array using Fibre Channel.

This section applies only to Fibre Channel. If you are using only iSCSI with vSphere, see iSCSI HBA configuration.

Review Fibre Channel SAN guidelines before you proceed.

Fibre Channel SAN guidelines

This section describes the best practices for attaching hosts to a PowerStore array in a highly available resilient and optimal Fibre Channel SAN.

Recommended configuration values summary

The following table summarizes the recommended configuration values that are related to Fibre Channel SAN:

Table 5. Recommended configuration values

Validation	Impact	Severity
Use two separate fabrics.	Redundancy	Mandatory

Table 5. Recommended configuration values (continued)

Validation	Impact	Severity
Each host should be zoned to both nodes of each appliance.	Redundancy	Mandatory
Balance the hosts between the nodes of the appliance to provide a distributed load across all target ports.	Performance	Recommended
Maximum number of paths per appliance per volume per host: **8**	Performance	Warning
Recommended number of paths per volume per host: **4**	Performance	Warning
Link speed must be consistent across all paths to the PowerStore array.	Performance	Warning
Maximum ISL Hops: **2**	Performance	Recommended

Prerequisites

Before installing HBAs on an ESXi host, ensure that the following prerequisites are met:

- For supported FC HBA models and drivers, see Dell EMC E-Lab Navigator.
- Verify that all HBAs have supported driver, firmware, and BIOS versions.
- Verify that all HBAs BIOS settings are configured according to E-Lab recommendations.

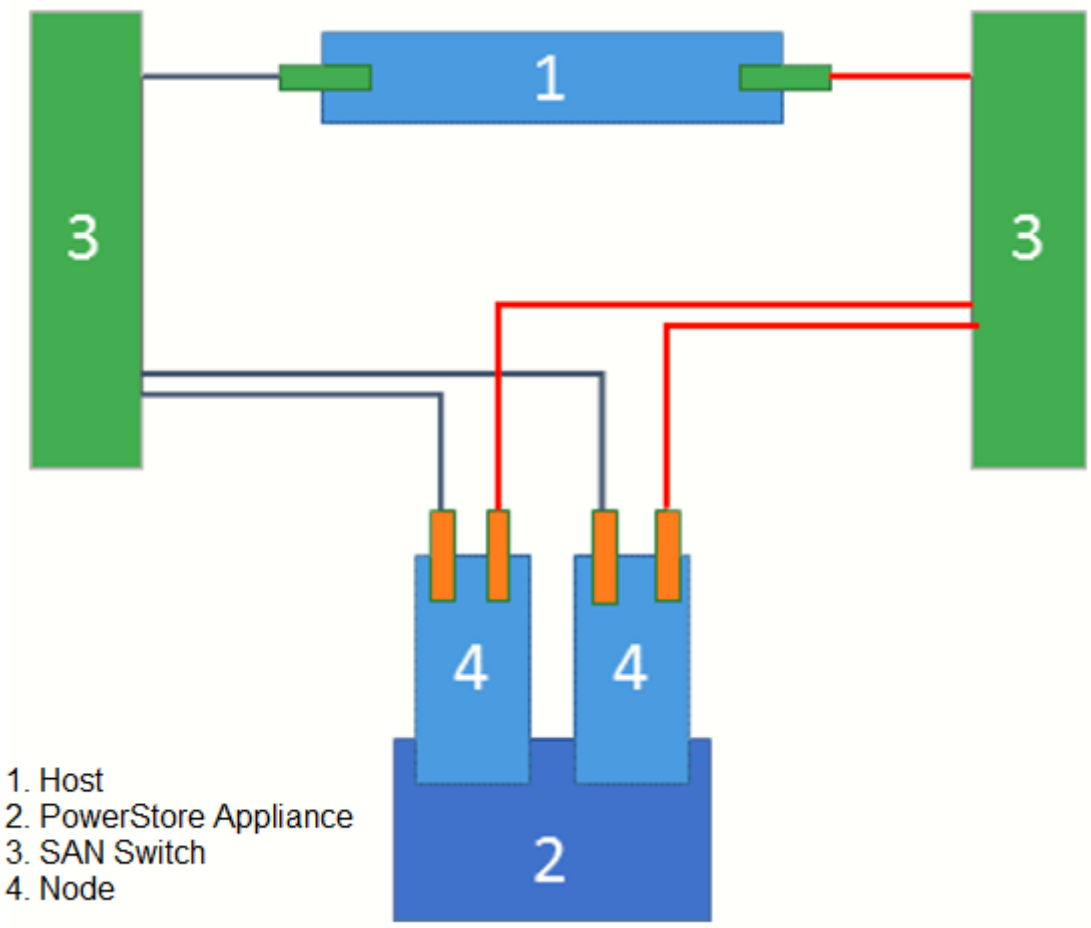
Recommended zoning configuration

When setting up a Fibre Channel SAN infrastructure, follow these guidelines:

- Use two separate fabrics. Each fabric must be on a different physical FC switch for resiliency.
- Balance the hosts between the two nodes of the appliance.
- The PowerStore array can be shipped with various Extension modules for Fibre Channel. If your PowerStore array contains more than one Extension I/O module per node, distribute the zoning among all I/O modules for highest availability and performance.
- Use single initiator zoning scheme: Utilize single-initiator per multiple-target (1: many) zoning scheme when configuring zoning with a PowerStore array.
- Keep a consistent link speed and duplex across all paths between the host and the PowerStore array.
- Host I/O latency can be severely affected by SAN congestion. Minimize the use of ISLs by placing the host and storage ports on the same physical switch. When this is not possible, ensure that there is a sufficient ISL bandwidth, and both the host and PowerStore array interfaces are separated by no more than two ISL hops.
- The optimal number of paths depends on the operating system and server information. To avoid multipathing performance degradation, do not use more than eight paths per device per host. It is recommended to use four paths.
- For more information about zoning best practices, see the *Networked Storage Concepts and Protocols techbook*, available on Dell EMC Online Support.
- PowerStore supports direct attachment of hosts to the appliances.
- With Federation, it is recommended to zone the host to as many appliances as possible, to achieve best load distribution across the cluster. Ensure that you keep the minimum/optimal zoning recommendations for each appliance.

NOTE: A federation is designed to provide a better load balancing, not for better resiliency. To perform volume migration between appliances, a host must be zoned to both appliances.

The following figure describes a simple connectivity.



For implementation instructions, see the Fibre Channel switch *User Manual*.

iSCSI Configuration

This section describes the recommended configuration that must be applied when attaching hosts to PowerStore array using iSCSI.

NOTE:

This section applies only for iSCSI. If you are using only Fibre Channel with vSphere and PowerStore, see Fibre Channel HBA configuration.

Review the iSCSI SAN guidelines before you proceed.

Prerequisites

Before configuring iSCSI with vSphere, ensure that you met the following prerequisites:

Follow the VMware recommendations for installation and setup of the appropriate NIC/iSCSI HBA for your system. It is recommended to install the latest driver version (patch), as described in the VMware support site for each specific NIC/iSCSI HBA.

For supported NIC/ iSCSI HBA models and drivers, see Dell EMC E-Lab Navigator .

NOTE: PowerStore arrays support only one subnet for iSCSI connections.

For example: From the host, NIC1 and Node A-0 and Node B-0 are on one network subnet. NIC2 and Node A-1 and Node B-1 are on a different subnet. This example connects NIC1 to Node A-0 and Node B-0, and NIC2 to Node A-1 and Node B-1.

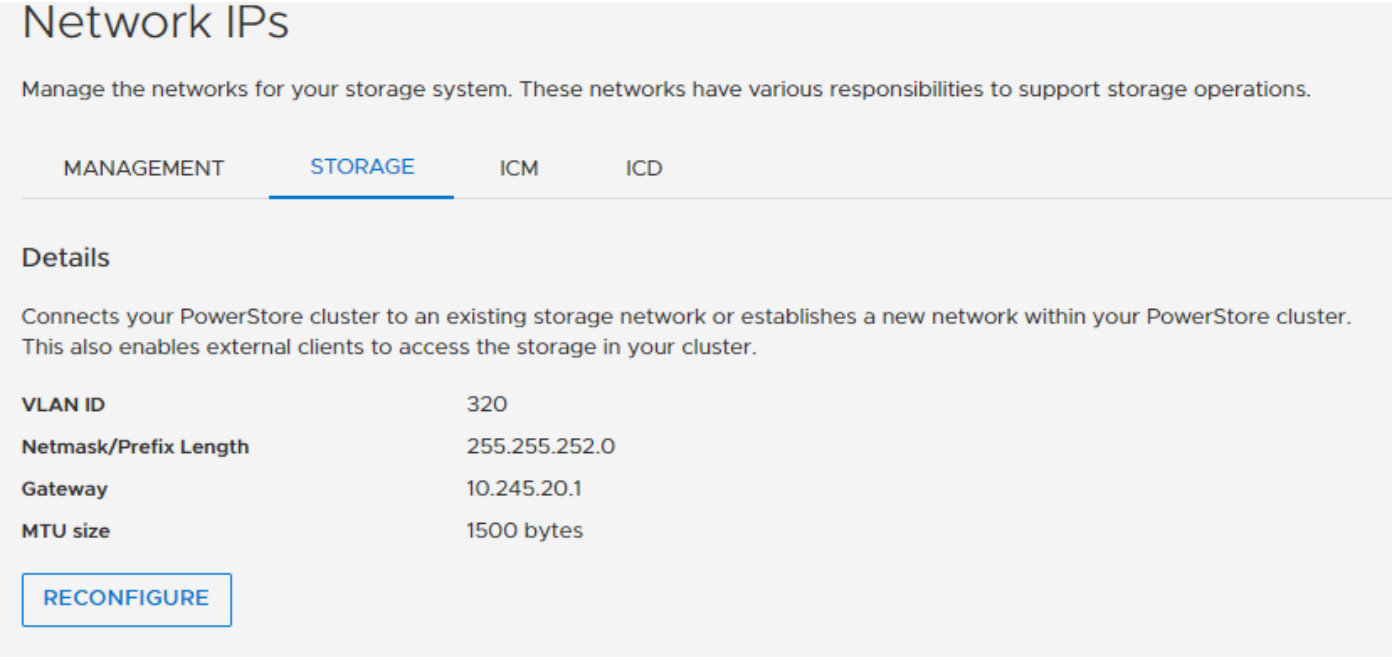
NOTE: NIC1 and NIC2 could also be on the same subnet, but is not recommended it except for PowerStore arrays, that support only one subnet for iSCSI connections.

PowerStore iSCSI reconfiguration

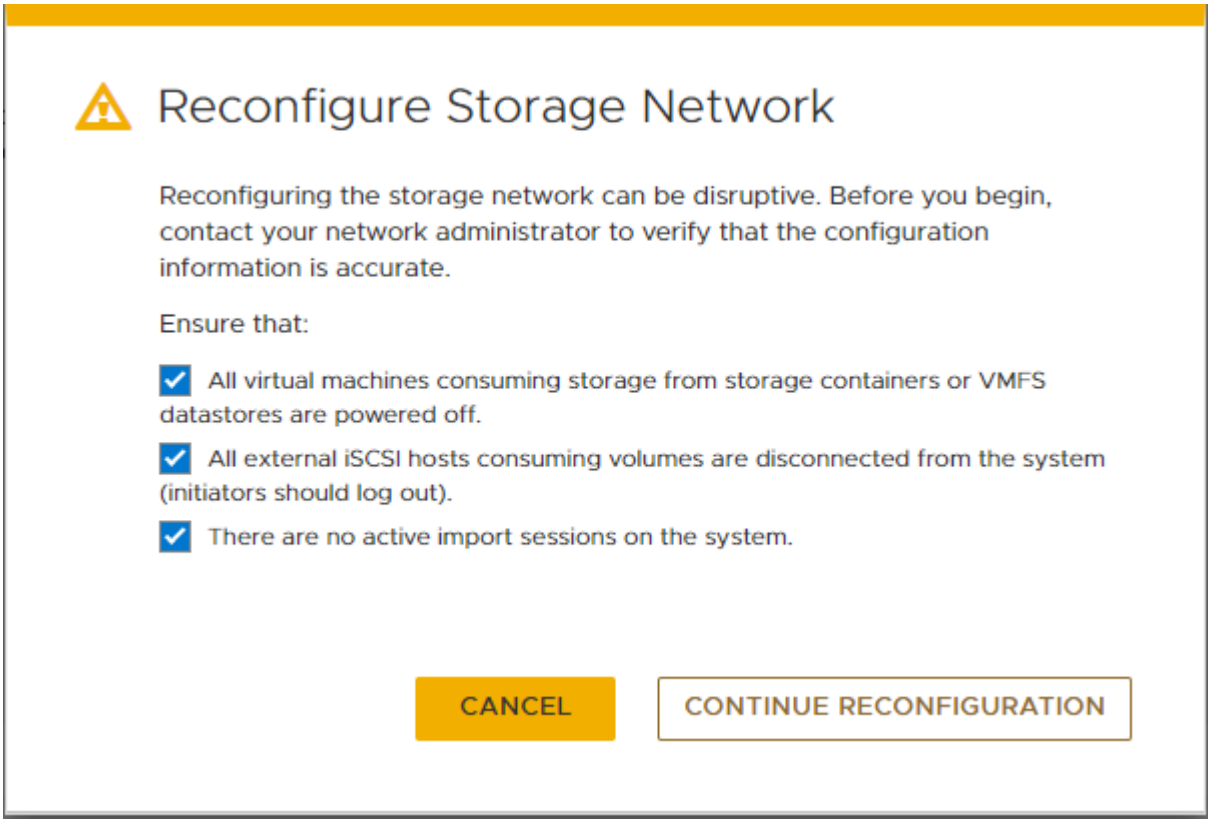
Reconfigure iSCSI IP in PowerStore web UI

Steps

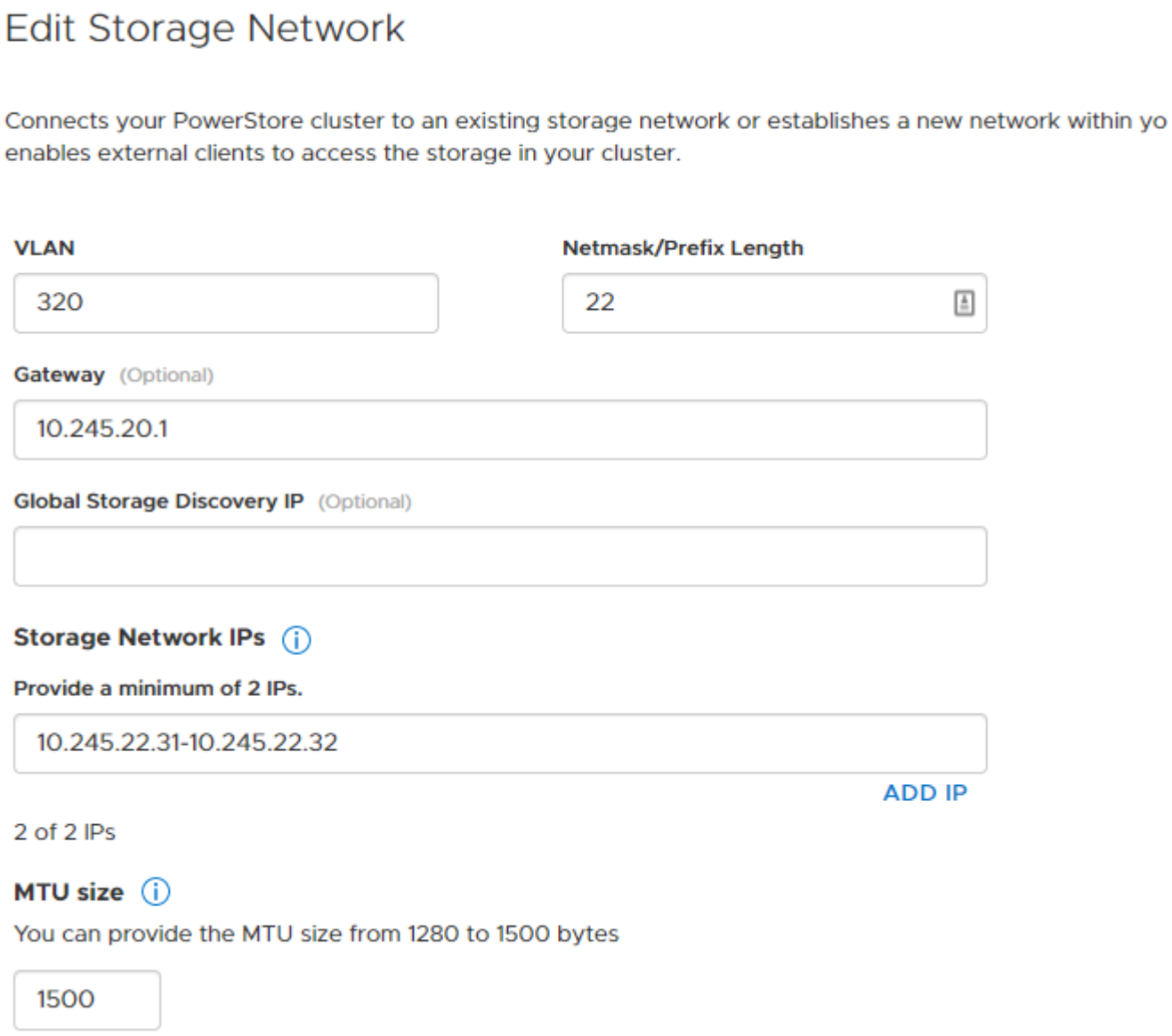
1. Log in to PowerStore web-based UI.
2. Go to **Setting** > **Network IPs** > **Storage.**
   This step assumes that PowerStore is configured during installation.



3. Click **RECONFIGURE**.
   A window similar to the following figure appears:



4. Select all the check boxes, and click **CONTINUE RECONFIGURATION**.
   A window similar to the following figure appears:



5. Enter the IP, VLAN, and Gateway information.

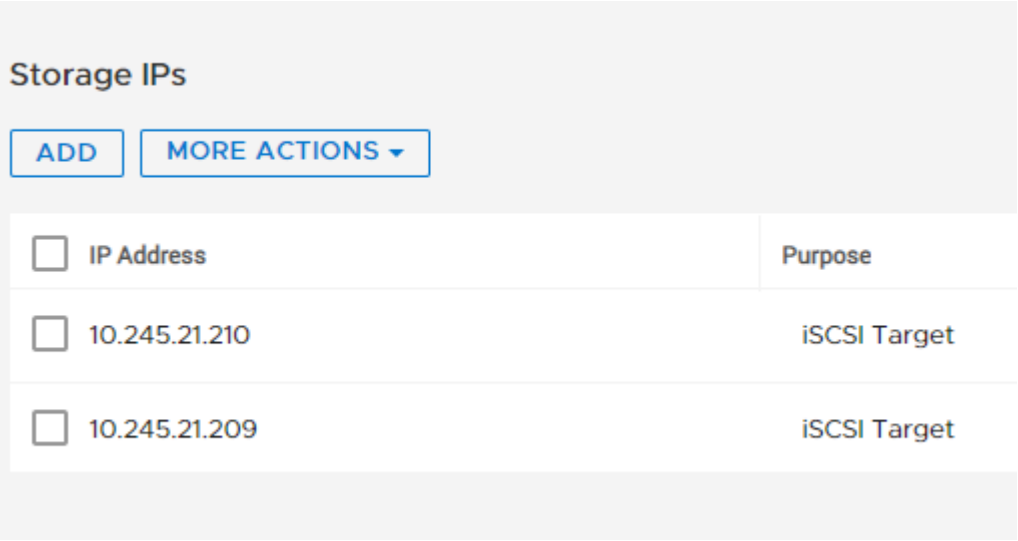
NOTE: The storage IP cannot be in the same subnet of management network.

Routing in storage network is supported, but it is not recommended.

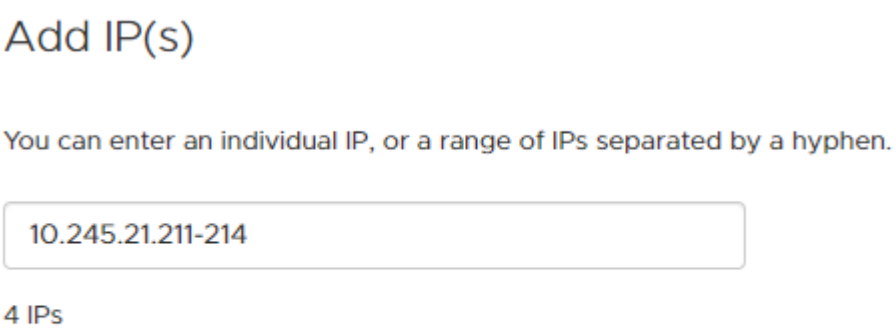
Add new IP for additional interfaces

Steps

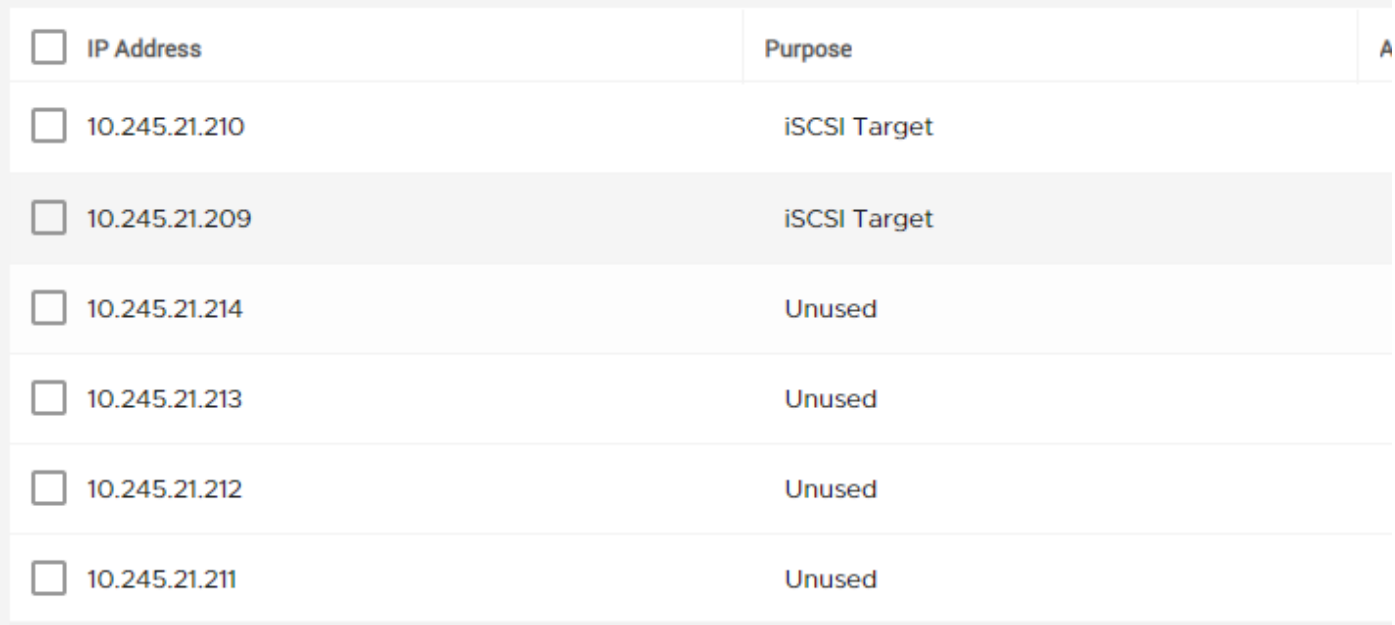
1. In Storage IPs, click **ADD**.



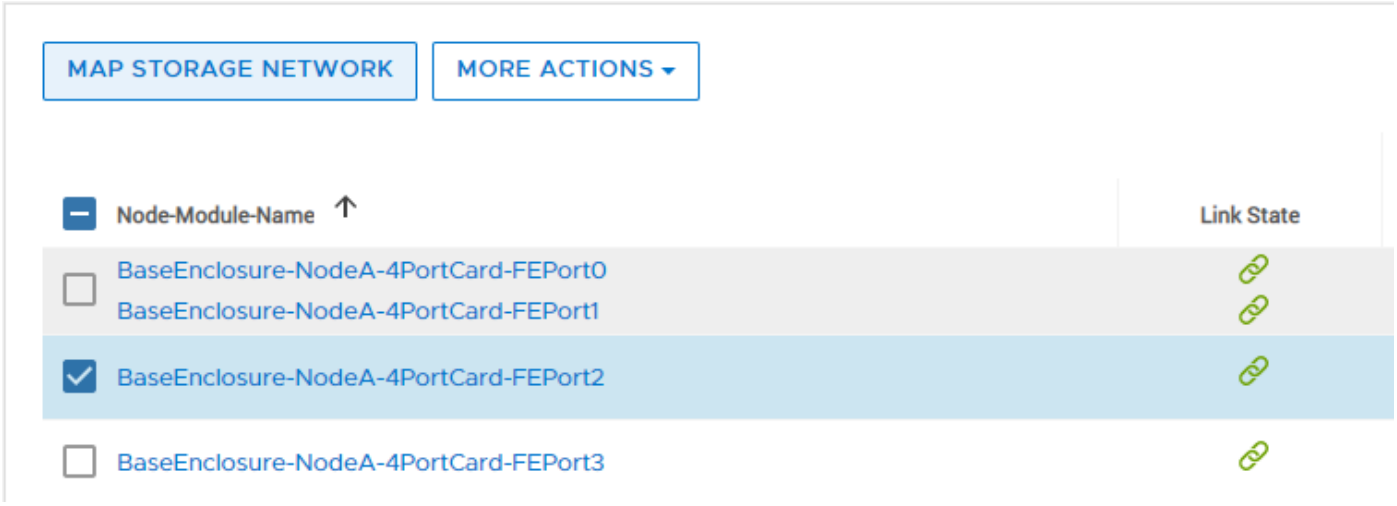
2. Enter the IP details.



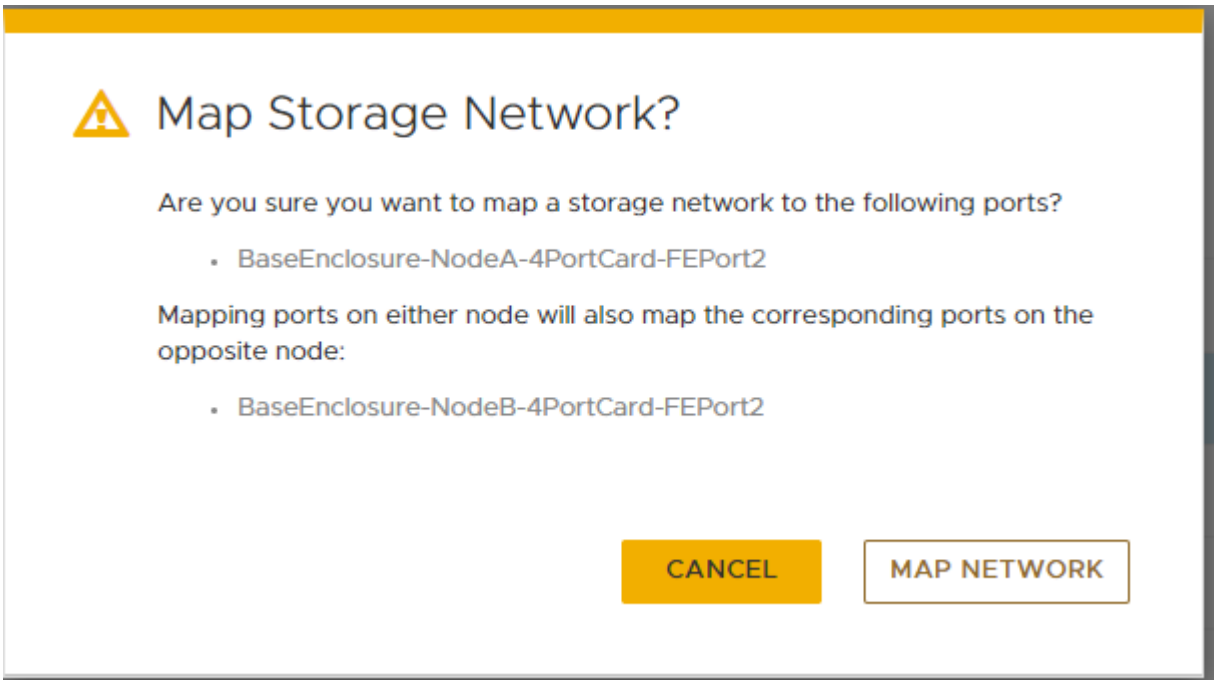
The new IPs show as unused.



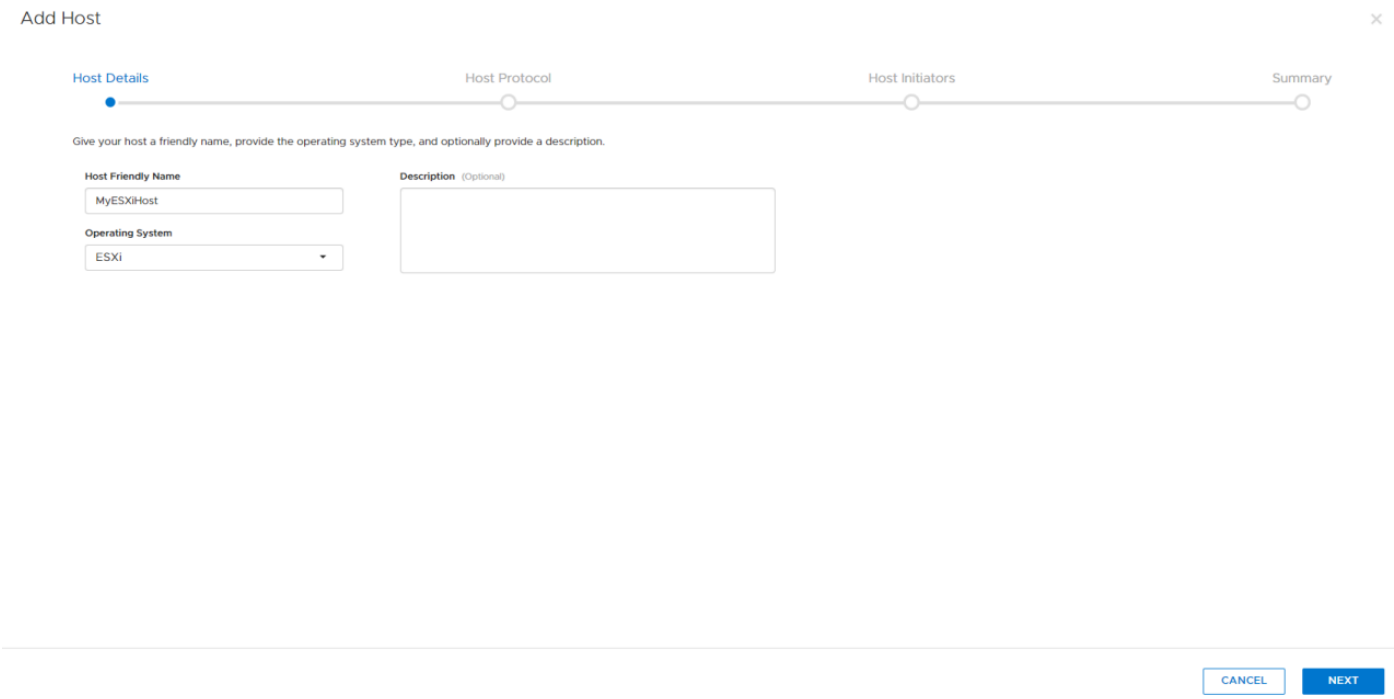
3. Go to **Appliance** > **Ports**. Find the up links and click **MAP STORAGE NETWORK**.



4. Click **MAP NETWORK**.



Once it mapped, the new configured will be assigned to the interface.



Example

Jumbo Frames

Follow these guidelines for configuring Jumbo Frames with ESXi hosts:

When using iSCSI with ESXi hosts and PowerStore, it is recommended to configure end-to-end Jumbo Frames (MTU=9000) for optimal performance. With Jumbo Frames, Ethernet frames are set larger than 1500 bytes (up to 9000 bytes).

NOTE: When using Jumbo Frames, ensure that all ports (ESXi server, switches, and storage) are configured with the correct MTU value. With VMware ESXi, the correct MTU size must be set on the virtual switch as well.

For details about configuring Jumbo Frames with iSCSI on ESXi, see the *VMware KB Article 1007654* on the VMware website.

Delayed ACK

For optimal traffic, it is recommended to disable Delayed ACK on ESXi.

For optimal iSCSI traffic between the ESXi hosts and PowerStore, especially during periods of network congestion, it is recommended to disable Delayed ACK on ESXi. By disabling Delayed ACK, the ESXi host sends an ACK acknowledgment segment for every received data segment (rather than delaying the sending of ACK acknowledgment segments, while receiving a stream of TCP data segments).

For details on the Delayed ACK vSphere parameter and how to disable it using the vSphere Client, see *VMware KB Article 1002598* on the VMware website.

NOTE: The recommended method for configuring the Delayed ACK vSphere setting is per discovered iSCSI target. This allows to disable Delayed ACK only for PowerStore iSCSI targets.

vStorage API for Array Integration settings

This section describes the necessary settings for configuring vStorage API for Array Integration (VAAI) for the PowerStore storage array.

VAAI is a vSphere API that offloads vSphere operations such as virtual machine provisioning, storage cloning, and space reclamation to storage arrays that supports VAAI. The PowerStore storage array fully supports VAAI.

To ensure optimal performance of PowerStore storage, you must enable VAAI on the ESXi host before using PowerStore from vSphere.

Confirming that VAAI is enabled on the ESXi host

Follow these instructions to confirm that VAAI is enabled on the ESXi host:

Before using the PowerStore storage, confirm that following VAAI features are enabled on the ESXi host.

NOTE: VAAI is enabled by default on vSphere version 6.5 and later.

Verify that the following parameters are set to **1**. This represents that the parameters are enabled.

- DataMover.HardwareAcceleratedMove
- DataMover.HardwareAcceleratedInit
- VMFS3.HardwareAcceleratedLocking

If any of the above parameters are not enabled, click the **Edit** icon and click **OK** to adjust them.

Multipathing software configuration

Configuring vSphere Native Multipathing

PowerStore supports the VMware vSphere Native Multipathing (NMP) technology. This section describes the guidelines for configuring native multipathing for PowerStore volumes. Follow these guidelines when configuring vSphere with native multipathing.

NOTE: VMware ESXi 6.7, Path Release ESXi670-201912001 and later version include SATP rules for PowerStore. Round Robin path selection policy is the default policy that is used to claim PowerStore volumes, which requires no manual configuration.

For best performance, it is recommended to:

- Set the native Round Robin path selection policy on PowerStore volumes that are presented to the ESXi host.
- Set the vSphere NMP Round Robin path switching frequency to PowerStore volumes from the default value (1000 I/O packets) to **1**.

These settings ensure optimal distribution and availability of load between I/O paths to the PowerStore storage.

Configuring vSphere NMP Round Robin as the default path policy for all PowerStore volumes

This topic provides the procedure to configure vSphere NMP Round Robin as the default path policy for all PowerStore volumes using the ESXi command line.

About this task

NOTE: Use this method when no PowerStore volume is presented to the host. The PowerStore volumes that are already presented to the host are not affected by this procedure (unless they are unmapped from the host).

Steps

1. Open an SSH session to the host as root.
2. Run the following command to configure the default path policy for newly defined PowerStore volumes as Round Robin with path switching after each I/O packet:

```
esxcli storage nmp satp rule add -c tpgs_on -e "PowerStore" -M
PowerStore -P VMW_PSP_RR -O iops=1 -s VMW_SATP_ALUA -t vendor -V
DellEMC
```

This command also sets the vSphere NMP Round Robin path switching frequency for newly defined PowerStore volumes to 1.

NOTE: Using this method does not impact any non-PowerStore volume that is presented to the ESXi host.

Configuring vSphere NMP Round Robin on an PowerStore volume already presented to the ESXi host

This topic provides the procedure to configure vSphere NMP Round Robin on a PowerStore volume that is already presented to the ESXi host, using ESXi command line:

About this task

NOTE: Use this method only for PowerStore volumes that are already presented to the host. For volumes not yet presented to the host, see to Configuring vSphere NMP Round Robin as the default path policy for all volumes.

Steps

1. Open an SSH session to the host as root user.
2. Run the following command to obtain the NAA of PowerStore LUNs presented to the ESXi host:

   ```
   # esxcli storage nmp path list | grep PowerStore -B1
   ```

3. Run the following command to modify the path selection policy on the PowerStore volume to Round Robin:

   ```
   # esxcli storage nmp device set --device <naa_id> --psp VMW_PSP_RR
   ```

   The following example shows a sample command:

   ```
   # esxcli storage nmp device set -device naa. 68ccf098003f1461569ea4750e9dac50 --psp
   VMW_PSP_RR
   ```

4. Run the following command to set the vSphere NMP Round Robin path switching frequency on PowerStore volumes from the default value (1000 I/O packets) to 1:

   ```
   # esxcli storage nmp psp roundrobin
   deviceconfig set -- device="<naa_id>" --iops=1
   --type=iops
   ```

   The following example shows a sample command:

   ```
   # esxcli storage nmp psp roundrobin deviceconfig set
   --device="naa. 68ccf098003f1461569ea4750e9dac50" --iops=1
   --type=iops
   ```

   To modify all existing LUNs in a batch format, use the following shell command:

   ```
   # for i in `esxcfg-scsidevs -l|grep "PowerStore" -B7|grep "^naa."`;do ESXIicli storage nmp
   device set --device $i --psp VMW_PSP_RR;done
   # for i in `esxcfg-scsidevs -l|grep "PowerStore" -B7|grep "^naa."`;do ESXIicli storage nmp
   psp roundrobin deviceconfig set --type=iops --iops=1 --device=$i; done
   ```

Next steps

NOTE: Using this method does not impact any non-PowerStore volumes that are presented to the ESXi host.

For more details, see the VMWare KB articles 1017760 and 2069356 on the VMware website.

Post configuration tasks

When host configuration is completed, you can use the PowerStore storage from the host. For details about creating, presenting, and managing volumes that can be accessed from the host using either GUI or CLI, see the *PowerStore Storage Array User Guide* that matches the version running on your PowerStore array.

Dell EMC Virtual Storage Integrator (VSI) Unified Storage Management version 6.2 and later can be used to provision from within vSphere Client Virtual Machine File System (VMFS) datastores and Raw Device Mapping volumes on PowerStore. Furthermore, Dell EMC VSI Storage Viewer version 6.2 and later extends the vSphere Client to facilitate the discovery and identification of PowerStore storage devices that are allocated to VMware ESXi hosts and virtual machines.

For more information about using these two vSphere Client plugins, see the *VSI Unified Storage Management Product Guide* and the *VSI Storage Viewer Product Guide*.

Disk formatting

Consider the following guidelines when creating volumes in PowerStore for a vSphere host:

- Disk logical block size - The only logical block (LB) size that is supported by vSphere for presenting to ESXi volumes is 512 bytes.

  NOTE: For details about formatting a newly created volume (using either the web-based UI or the GUI), see the *PowerStore Storage Array User Guide* that matches the version running on your PowerStore array.

- Disk alignment - Unaligned disk partitions may substantially impact I/O to the disk. With vSphere, datastores and virtual disks are aligned by default as they are created. Therefore, no further action is required to align these in ESXi.

  With virtual machine disk partitions within the virtual disk, alignment is determined by the guest operating system. For virtual machines that are not aligned, consider using tools such as UBERalign to realign the disk partitions as required.

Presenting PowerStore volumes to the ESXi host

Follow these guidelines for presenting PowerStore volumes to the ESXi host.

NOTE: Using data reduction and /or encryption software on the host side will affect the PowerStore array data reduction.

NOTE: When using iSCSI software initiator with ESXi and PowerStore storage, it is recommended to use only lower case characters in the IQN to correctly present the PowerStore volumes to ESXi. For more details, see the *VMware KB article 2017582* on the VMware website.

When adding host groups and hosts to allow ESXi hosts to access PowerStore volumes, specify ESXi as the operating system for newly created hosts.

The following figure demonstrates setting the operating system field for a host using the web-based UI:

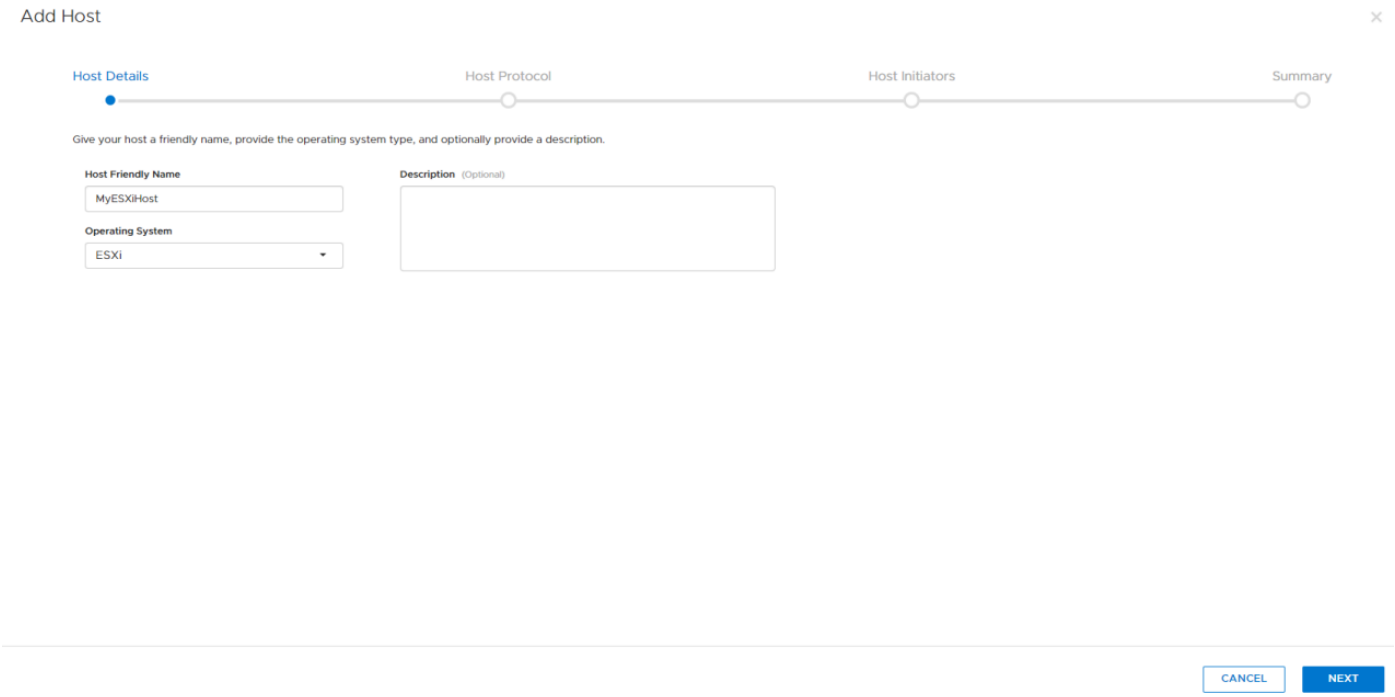


Figure 12. Setting the operating system field for a host

VMware paravirtual SCSI controllers

Follow these guidelines for configuring virtual machines with paravirtual SCSI controllers.

For optimal resource utilization of virtual machines with PowerStore, it is recommended to configure virtual machines with paravirtualized SCSI controllers. VMware paravirtual SCSI controllers are high-performance storage controllers that can provide higher throughput and lower CPU usage. These controllers are best suited for high-performance storage environments.

NOTE: Virtual machines with paravirtualized SCSI controllers that are to be part of an MSCS cluster, are supported with vSphere 6.0 and later.

For more details about configuring virtual machines with paravirtualized SCSI controllers, see the *vSphere Virtual Machine Administration Guide* available on the vSphere documentation.

Virtual Machine formatting

Follow these recommendations for formatting virtual machines on PowerStore storage.

For optimal space utilization with vSphere 6.*x*, it is recommended to format virtual machines on PowerStore, using Thin Provisioning.

Virtual Machine formatting using Thin Provisioning

Follow this procedure for virtual machine formatting using thin provisioning.

About this task

In thin provisioning format, in-guest space reclamation is available, if the following requirements are fulfilled:

- Thin virtual disks
- VM hardware version 11
- ESXi 6.x
- EnableBlockDelete set to **1**
- Guest operating system support of UNMAP

NOTE: Some guest operating systems that support unmapping of blocks, such as Linux-based systems, do not generate UNMAP commands on virtual disks in vSphere 6.0. This occurs because the level of SCSI support for ESXi 6.0 virtual disks is SCSI-2, while Linux expects 5 or higher for SPC-4 standard. This limitation prevents the generation of UNMAP commands until the virtual disks can claim support for at least SPC-4 SCSI commands.

For more details on virtual volumes and UNMAP, see VMware KB# 2112333.

Steps

1. From vSphere Web Client, launch the **Create New Virtual Machine** wizard.
2. Using the wizard up, go to the **2f Customize Hardware** screen.
3. In the **Customize Hardware** screen, click **Virtual Hardware**.
4. Go to the **New Hard Disk** option.
5. Select the **Thin Provision** option to format the virtual disk of the virtual machine.

   For details about migrating a virtual machine from thick provision to thin provision, see VMware KB# 2014832.

Virtual machine formatting with Thick-Provision VMDK

Follow this procedure for virtual machine formatting with Thick Provision VMDK.

About this task

When optimal performance is required or when vSphere version is lower than 6.x, format virtual machines on PowerStore storage, using Thick Provision Eager Zeroed. Using this format, the required space for the virtual machine is allocated in zeroed on creation time. However, with native PowerStore data reduction, thin provisioning and VAAI support, no actual physical capacity allocation occurs.

The advantages of Thick Provision Eager Zeroed format are:

- Logical space is allocated and zeroed on virtual machine provisioning time, rather than scattered with each I/O send by the virtual machine to the disk (when Thick Provision Lazy Zeroed format is used).
- Thick provisioning is managed in the PowerStore Storage Array rather than in the ESXi host (when Thin Provision format is used).

Steps

1. From vSphere Web Client, launch the **Create New Virtual Machine** wizard.
2. Using the wizard, proceed up to the **2f Customize Hardware** screen.
3. In the **Customize Hardware** screen, click **Virtual Hardware**.
4. Go to the **New Hard Disk** option.
5. Select the **Thick Provision Eager Zeroed** option to format the virtual disk of the virtual machine.
6. Proceed using the wizard to complete creating the virtual machine.

PowerStore virtualization

Virtualization architecture and configuration

PowerStore X model configuration

The PowerStore X model configuration provides a hypervisor layer in addition to block storage. VMware ESXi is the base operating system running on the PowerStore hardware, with the PowerStore software running on a VM. The PowerStore X model appliance uses 50% of the system resources and storage space to run the PowerStore software VMs.

This application-centric configuration enables you to run applications in other VMs on the PowerStore hardware.

The PowerStore Resource Balancer automatically manages the placement of VMs and vVols to balance the load across the PowerStore X model cluster. Resources are moved and deployed to locations based on rules to facilitate high availability (HA).

In general, this behavior does not impact the daily operation or management of a PowerStore X model appliance. However, there are some considerations for storage administrators. The PowerStore X model uses federated rather than distributed storage. This configuration may impact where a storage administrator chooses to locate resources within a data center, and how VM loads are set up for satisfactory performance.

PowerStore T model configuration

In the PowerStore T model configuration, the PowerStore software runs directly on the PowerStore hardware.

To add VMs in this configuration, you must add ESXi hosts and a vCenter server connection with a VASA provider. Some of the virtualization features available on PowerStore X model appliances are not available in this configuration.

For PowerStore T model appliances, the PowerStore Resource Balancer can manage the placement of traditional volumes but not VMs or vVols. You can migrate vVol from one appliance to another manually from the Virtual Volumes card of the associated VM.

Hypervisor configuration for PowerStore X model appliances

vSphere Distributed Switch

You cannot use an existing vSphere Distributed Switch (vDS) configuration with your PowerStore X model appliance. PowerStore creates its own vDS and does not use the existing vDS configuration.

Network configuration for PowerStore X model appliances

Plan your network configuration before setting up your PowerStore X model appliance.

The PowerStore X model appliance has some specific requirements and considerations that do not apply to PowerStore T model appliances. For detailed information about network configuration with PowerStore appliances, see the *PowerStore Configure Switches and External Networks Guide*. This document includes:

- Planning and preparation information
- Configuration recommendations and best practices
- Network configuration, VLAN, and IP address requirements
- Switch configuration and cabling procedures
- Network configuration validation instructions

NOTE: If the configuration is modified, there are some consequences and limitations that impact HA.

NTP server requirements

To prevent issues with lag time between vCenter and ESXi hosts when creating your cluster, ensure that they are using the same NTP server.

VPLEX

This section describes host connectivity of the Dell EMC VPLEX in VMware environment.

Overview

For detailed information about VPLEX, See documentation available at Dell EMC Online Support.

Documentation

See the following documents for configuration and administration operations:

- *Dell EMC VPLEX with GeoSynchrony 6.x Product Guide*
- *Dell EMC VPLEX with GeoSynchrony 6.x CLI Guide*
- *Dell EMC VPLEX with GeoSynchrony 6.x Configuration Guide*
- *Dell EMC VPLEX Hardware Installation Guide*
- *Dell EMC VPLEX Release Notes*
- *Implementation and Planning Best Practices for Dell EMC VPLEX Technical Notes*
- *VPLEX online help*, available on the Management Console GUI
- *VPLEX Procedure Generator*, available at Dell EMC Online Support
- *Dell EMC Simple Support Matrix, Dell EMC VPLEX and GeoSynchrony*, available on Dell EMC E-Lab Navigator.

For the up-to-date support information, See Dell EMC E-Lab Navigator.

Prerequisites

Before configuring VPLEX in the Windows environment, complete the following on each host:

- Confirm that all necessary remediation has been completed. This ensures that OS-specific patches and software on all hosts in the VPLEX environment are at supported levels according to the Dell EMC E-Lab Navigator.
- Confirm that each host is running VPLEX-supported failover software and has at least one available path to each VPLEX fabric.

  NOTE: See Dell EMC E-Lab Navigator for the most up-to-date support information and prerequisites.
- If a host is running PowerPath, confirm that the load-balancing and failover policy is set to **Adaptive**.

Provisioning and exporting storage

VPLEX with GeoSynchrony v4.x

To begin using VPLEX, you must provision and export storage so that hosts and applications can use the storage. Storage provisioning and exporting refers to the following tasks required to take a storage volume from a storage array and make it visible to a host:

Steps

1. Discover available storage.
2. Claim and name storage volumes.
3. Create extents from the storage volumes.
4. Create devices from the extents.
5. Create virtual volumes on the devices.
6. Create storage views to allow hosts to view specific virtual volumes.
7. Register initiators with VPLEX.

8. Add initiators (hosts), virtual volumes, and VPLEX ports to the storage view. You can provision storage using the GUI or the CLI. For more information, see *Dell EMC VPLEX Management Console Help* or the *Dell EMC VPLEX CLI Guide*, located on Dell EMC Online Support.

Example

The following figure shows the provisioning and exporting process:

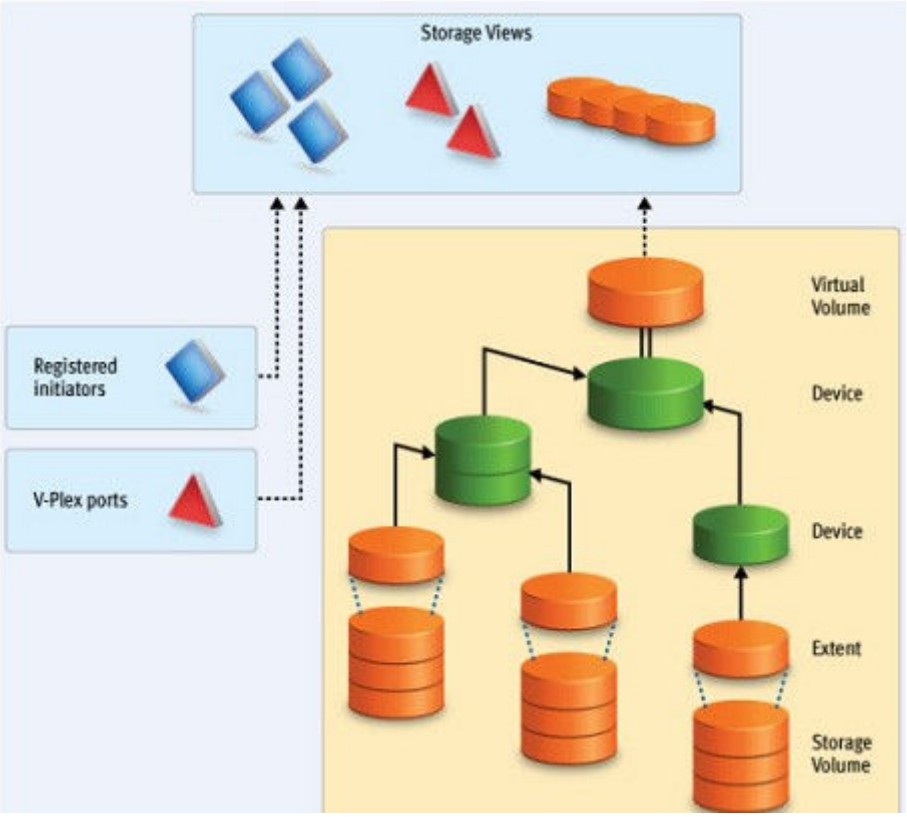


Figure 13. Provisioning and exporting process

VPLEX with GeoSynchrony v5.x

VPLEX allows easy storage provisioning among heterogeneous storage arrays. After a storage array LUN volume is encapsulated within VPLEX, all of its block-level storage is available in a global directory and coherent cache. Any front-end device that is zoned properly can access the storage blocks.

Two methods are available for provisioning: EZ provisioning and Advanced provisioning. For more information, see the *Dell EMC VPLEX with GeoSynchrony 5.5 Product Guide* on Dell EMC Online Support.

VPLEX with GeoSynchrony v6.x

VPLEX provides easy storage provisioning among heterogeneous storage arrays. Use the web-based GUI to simplify everyday provisioning or to create complex devices.

Use the following ways to provision storage in VPLEX:

- Integrated storage provisioning (VIAS-VPLEX Integrated Array Services based provisioning)
- EZ provisioning
- Advanced provisioning

All provisioning features are available in the Unisphere for VPLEX GUI.

For more information, see the *Dell EMC VPLEX with GeoSynchrony 6.1 Product Guide* on Dell EMC Online Support.

Storage volumes

A storage volume is a LUN exported from an array. When an array is discovered, the storage volumes view shows all exported LUNs on that array. You must claim, and optionally name, these storage volumes before you can use them in a VPLEX cluster. Once claimed, you

can divide a storage volume into multiple extents (up to 128), or you can create a single full-size extent using the entire capacity of the storage volume.

NOTE: To claim storage volumes, the GUI supports only the Claim Storage wizard, which assigns a meaningful name to the storage volume. Meaningful names help you associate a storage volume with a specific storage array and LUN on that array and are useful during troubleshooting and performance analysis.

This section contains the following topics:

- Claiming and naming storage volumes
- Extents
- Devices
- Distributed devices
- Rule sets
- Virtual volumes

Claiming and naming storage volumes

You must claim storage volumes before you can use them in the cluster (except the metadata volume, which is created from an unclaimed storage volume). Only after claiming a storage volume you can use it to create extents, devices, and then virtual volumes.

Extents

An extent is a slice (range of blocks) of a storage volume. You can create a full-size extent using the entire capacity of the storage volume, or you can carve the storage volume up into several contiguous slices. Extents are used to create devices and then virtual volumes.

Devices

Devices combine extents or other devices into one large device with specific RAID techniques, such as mirroring or striping. Devices can only be created from extents or other devices. A device's storage capacity is not available until you create a virtual volume on the device and export that virtual volume to a host.

You can create only one virtual volume per device. There are two types of devices:

- Simple device - A simple device is configured using one component, which is an extent.
- Complex device - A complex device has more than one component, combined using a specific RAID type. The components can be extents or other devices (both simple and complex)

Distributed devices

Distributed devices are configured using storage from both clusters and are only used in multi-cluster plexes. A distributed device's components must be other devices and those devices must be created from storage in different clusters in the plex.

Rule sets

Rule sets are predefined rules that determine how a cluster behaves when it loses communication with the other cluster, for example, during an inter-cluster link failure or cluster failure. In these situations, until communication is restored, most I/O workloads require specific sets of virtual volumes to resume on one cluster and remain suspended on the other cluster.

VPLEX provides a Management Console on the management server in each cluster. You can create distributed devices using the GUI or CLI on either management server. The default rule set used by the GUI causes the cluster used to create the distributed device to detach during communication problems, allowing I/O to resume at the cluster. For more information about creating and applying rule sets, see the *Dell EMC VPLEX CLI Guide* on Dell EMC Online Support.

There are cases in which all I/O must be suspended resulting in a data unavailability. VPLEX with functionality of VPLEX Witness.

VPLEX with functionality of VPLEX Witness: When a VPLEX Metro configuration is augmented by VPLEX Witness, the resulting configuration provides the following features:

- High availability for applications in a VPLEX Metro configuration (no single points of storage failure)
- Fully automatic failure handling in a VPLEX Metro configuration
- Improved failure handling in a VPLEX configuration
- Better resource utilization

For information about VPLEX Witness, see the *Dell EMC VPLEX with GeoSynchrony 5.5, 6.1 Product Guide* on Dell EMC Online Support.

Virtual volumes

Virtual volumes are created on devices or distributed devices and presented to a host through a storage view. You can create virtual volumes only on top-level devices and always use full capacity of the device.

System volumes

VPLEX stores configuration and metadata on system volumes that are created from storage devices. There are two types of system volumes:

- Metadata volumes
- Logging volumes

Each of these volumes is briefly discussed in the following sections:

Metadata volumes

VPLEX maintains its configuration state, referred as metadata, on storage volumes provided by storage arrays. Each VPLEX cluster maintains its own metadata, which describes the local configuration information for this cluster and any distributed configuration information that is shared between clusters.

For more information about metadata volumes for VPLEX with GeoSynchrony v4.x, see *Dell EMC VPLEX CLI Guide*, on Dell EMC Online Support.

For more information about metadata volumes for VPLEX with GeoSynchrony v5.x, see *Dell EMC VPLEX with GeoSynchrony 5.0 Product Guide*, on Dell EMC Online Support.

For more information about metadata volumes for VPLEX with GeoSynchrony v6.x, see the *Dell EMC VPLEX with GeoSynchrony 5.0 Product Guide*, on Dell EMC Online Support.

Logging volumes

Logging volumes are created during initial system setup. It is required in each cluster to track any blocks written during a loss of connectivity between clusters. After an inter-cluster link is restored, the logging volume is used to synchronize distributed devices by sending only changed blocks over the inter-cluster link.

For more information about logging volumes for VPLEX with GeoSynchrony v4.x, see *Dell EMC VPLEX CLI Guide*, on Dell EMC Online Support.

For more information about logging volumes for VPLEX with GeoSynchrony v5.x, see *Dell EMC VPLEX with GeoSynchrony 5.0 Product Guide*, on Dell EMC Online Support.

For more information about logging volumes for VPLEX with GeoSynchrony v6.x, see *Dell EMC VPLEX with GeoSynchrony 5.0 Product Guide*, on Dell EMC Online Support.

Required storage system setup

Product documentation and installation procedures for connecting a VMAX, Symmetrix, Unity, VNX series, and CLARiiON storage system to a VPLEX instance are available on Dell EMC Online Support.

Required VMAX series FA bit settings

For VMAX series to VPLEX-connections, configure the VMAX series FC directors (FAs) as shown in following table:

NOTE: Dell EMC recommends that you download the latest information before installing any server.

Table 6. Required Symmetrix FA bit settings for connection to VPLEX

Set	Do not set	Optional
SPC-2 Compliance (SPC2)	Disable Queue Reset on Unit Attention (D)	Linkspeed
SCSI-3 Compliance (SC3)	AS/400 Ports Only (AS4)	Enable Auto-Negotiation (EAN)
Enable Point-to-Point (PP)	Avoid Reset Broadcast (ARB)	

Table 6. Required Symmetrix FA bit settings for connection to VPLEX

Set	Do not set	Optional
Unique Worldwide Name (UWN) Common Serial Number (C)	Environment Reports to Host (E) Soft Reset (S) Open VMS (OVMS) Return Busy (B) Enable Sunapee (SCL) Sequent Bit (SEQ) Non-Participant (N) OS-2007 (OS compliance)	VCM/ACLX[a]

a. You must set VCM/ACLX bit, if VPLEX is sharing VMAX series directors with hosts that require conflicting bit settings. For any other configuration, the VCM/ACLX bit can be either set or not set.

NOTE: When setting up a VPLEX-attach version 4.x or earlier with a VNX series or CLARiiON system, you must set the initiator type to `CLARiiON Open` and Failover Mode to `1`. ALUA is not supported.

When setting up a VPLEX-attach version 5.0 or later with a VNX series or CLARiiON system, the initiator type can be set to CLARiiON Open and the Failover Mode set to 1 or Failover Mode 4 since ALUA is supported.

If you are using the LUN masking, set the VCM/ACLX flag. You must use VCM/ACLX, if sharing array directors with hosts which require conflicting flag settings.

NOTE: The FA bit settings that are listed in Table 4 are for connectivity of VPLEX to Dell EMC VMAX series only. For host to Dell EMC VMAX series FA bit settings, see the Dell EMC E-Lab Navigator.

Initiator settings on back-end arrays

See the *VPLEX Procedure Generator,* on Dell EMC Online Support, to verify the initiator settings for storage arrays when configuring the arrays for use with VPLEX.

Host connectivity

This section describes host connectivity of Dell EMC VPLEX in a VMware environment.

For the most up-to-date information about qualified switches, hosts, host bus adapters (HBAs), and software, see *Dell EMC Simple Support Matrix*, available on Dell EMC E-Lab Navigator or contact your Dell EMC customer representative.

The latest Dell EMC-approved HBA drivers and software are available for download at the following websites:

- http://www.broadcom.com/
- https://cavium.com/
- http://www.brocade.com/

The Dell EMC HBA installation and configurations guides are available at the Dell EMC-specific download pages of these websites.

NOTE: Direct connect from an HBA to a VPLEX engine is not supported.

Exporting virtual volumes to hosts

A virtual volume can be added to more than one storage view. All hosts included in the storage view can access the virtual volume. The virtual volumes that are created on a device or distributed device are not visible to hosts (or initiators) until you add them to a storage view. For failover purposes, you can group two or more front-end VPLEX ports to export the same volumes.

About this task

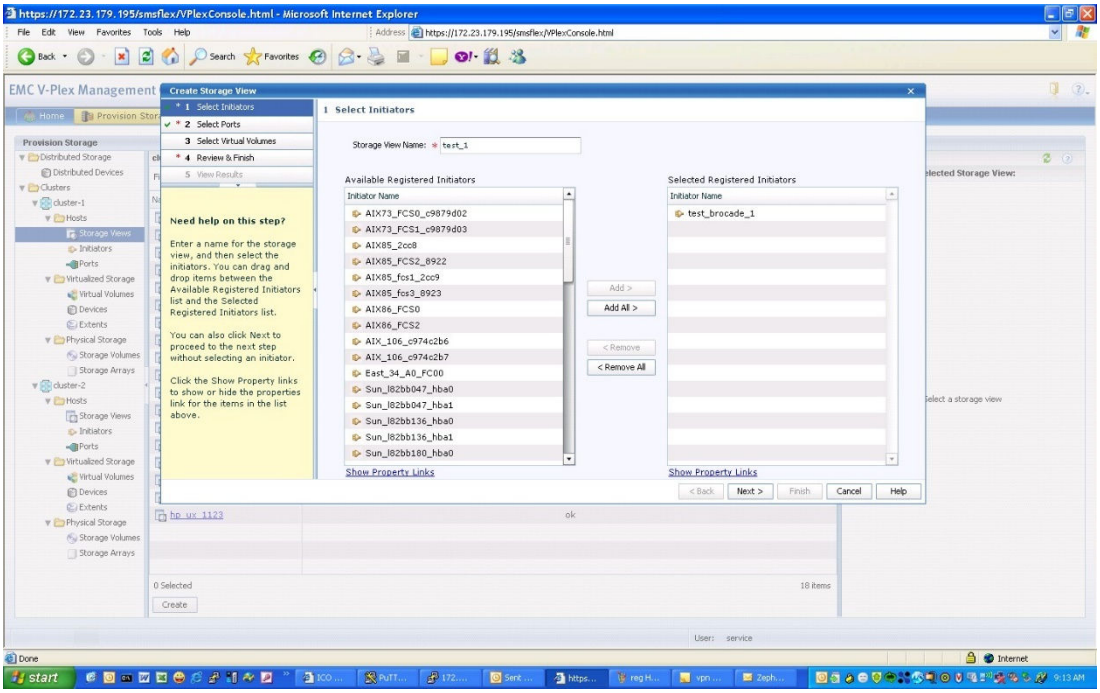
A volume is exported to an initiator as a LUN on one or more front-end port WWNs. Typically, initiators are grouped into initiator groups; all initiators in such a group share the same view on the exported storage. They can see the same volumes by the same LUN numbers on the same WWNs.

You must register an initiator with VPLEX to see any exported storage. The initiator must also be able to communicate with the front-end ports over a Fibre Channel (FC) switch fabric. Direct connect is not supported. Registering an initiator attaches a meaningful name to the WWN, typically the server's DNS name. This enables you to audit the storage view settings to determine which virtual volumes a specific server can access.

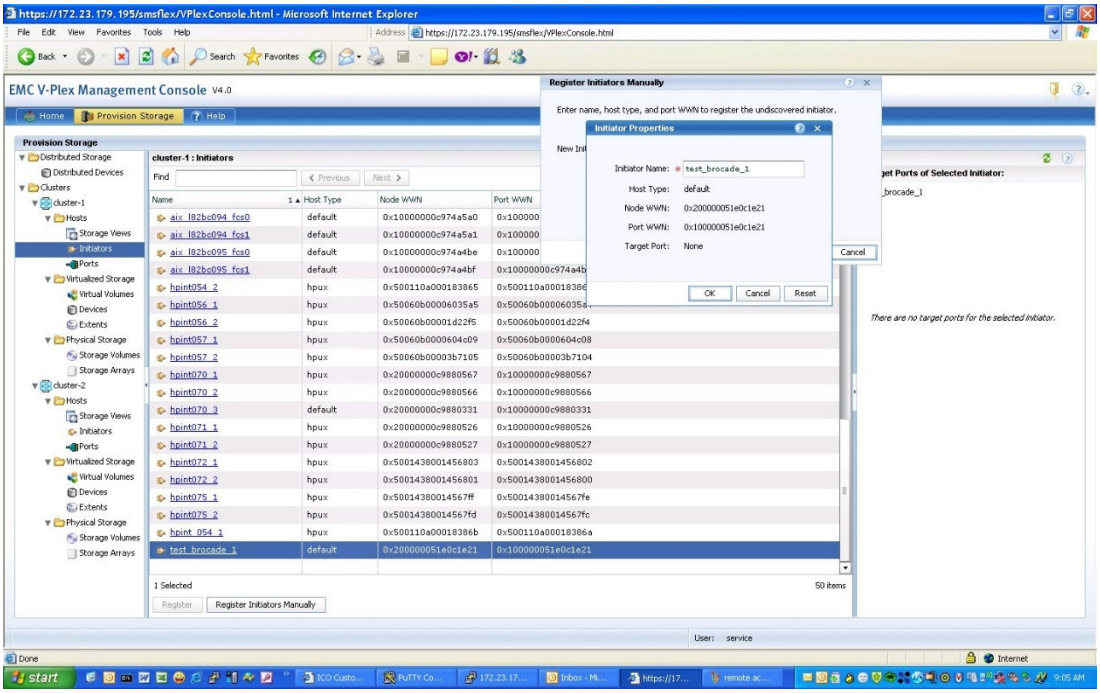
The following steps describe how to export virtual volumes:

Steps

1. Create a storage view.

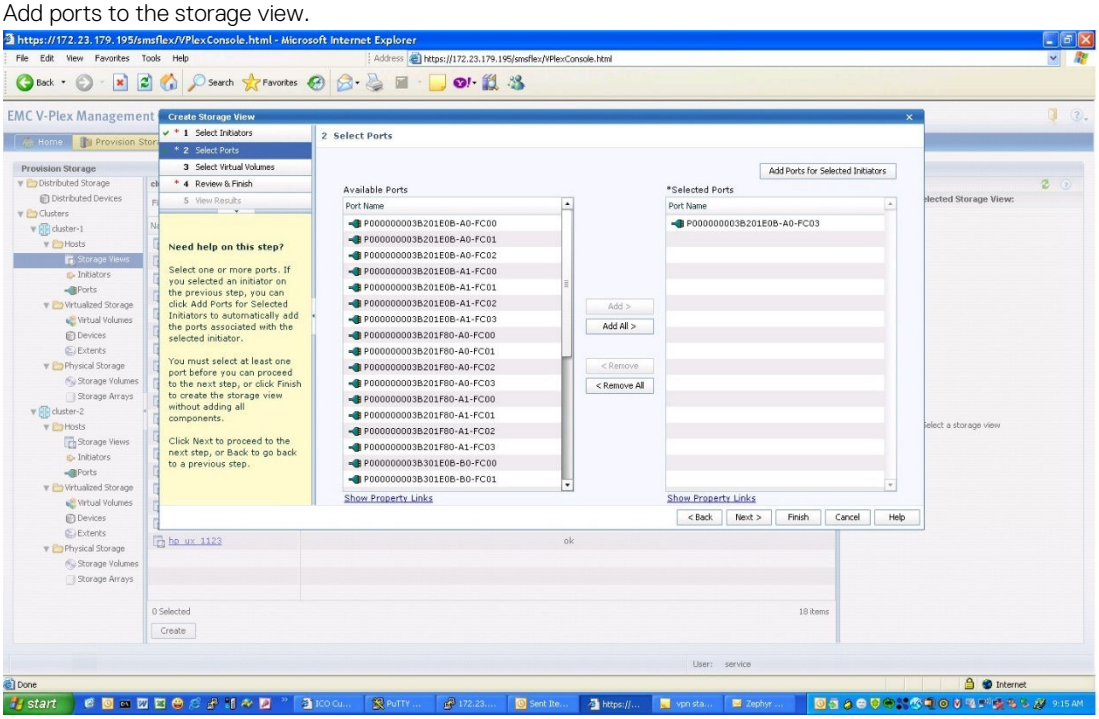


2. Register the initiators.

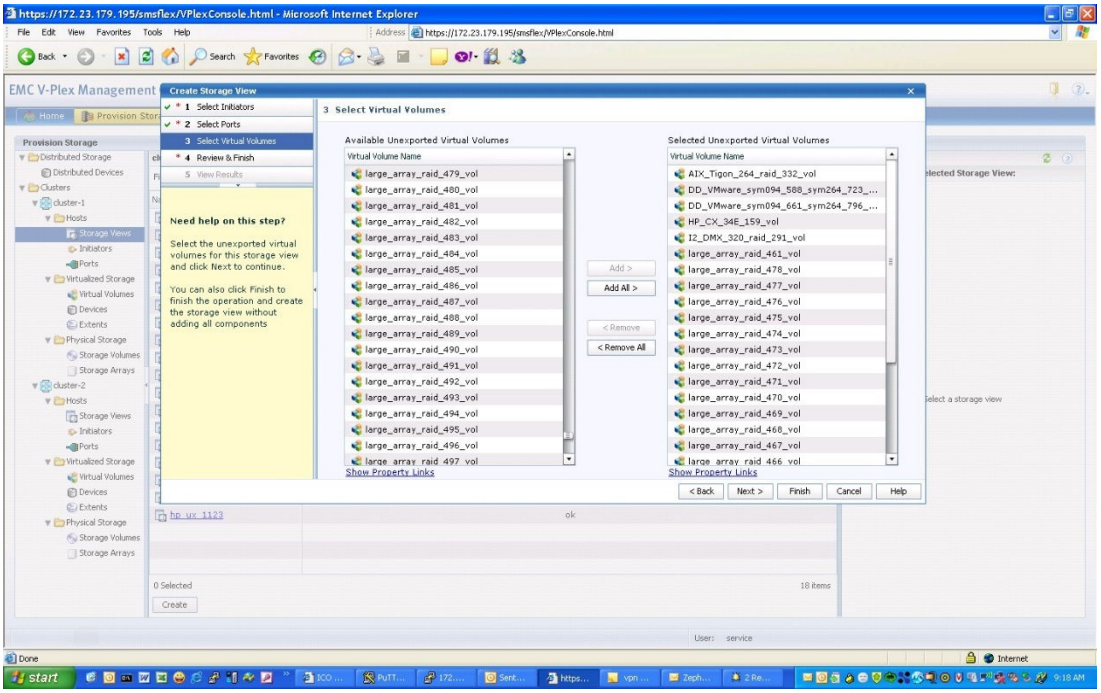


NOTE: When initiators are registered, you can set their type.

3. Add ports to the storage view.



4. Add virtual volumes to the storage view.



Front-end paths

This section describes the following:

- Viewing the WWN for an HBA port
- VPLEX ports
- Initiators

Viewing the WWN for an HBA port

Each HBA port has a World Wide Name (WWN) associated with it. WWNs are unique identifiers that the VPLEX engine uses to identify its ports and Host Initiators.

You can use one of the following ways to view WWNs:

- Switch's name server output
- Dell EMC ControlCenter or solutions enabler
- `syminq` command (VMAX or Symmetrix users)

VPLEX ports

The virtual volumes that are created on a device are not visible to hosts (initiators) until you export them. Virtual volumes are exported to a host through front-end ports on the VPLEX directors and HBA ports on the host/server. For failover purposes, two or more front-end VPLEX ports can be used to export the same volumes. To provide maximum redundancy, a storage view will have two VPLEX ports assigned to it, preferably from two different VPLEX directors. When volumes are added to a view, they are exported on all VPLEX ports in the view, using the same LUN numbers.

Initiators

For an initiator to see the virtual volumes in a storage view, it must be registered and included in the storage view's registered initiator list. The initiator must also be able to communicate with the front-end ports over Fibre Channel connections through a fabric.

A volume is exported to an initiator as a LUN on one or more front-end port WWNs. Initiators are grouped so that all initiators in a group share the same view of the exported storage (they can see the same volumes by the same LUN numbers on the same WWN host types).

Ensure that you specify the correct host type in the `Host Type` column as you cannot change this attribute in the **Initiator Properties** dialog box once the registration is complete. To change the host type after registration, you must unregister the initiator and then register it again using the correct host type.

Configuring VMware ESXi hosts to recognize VPLEX volumes

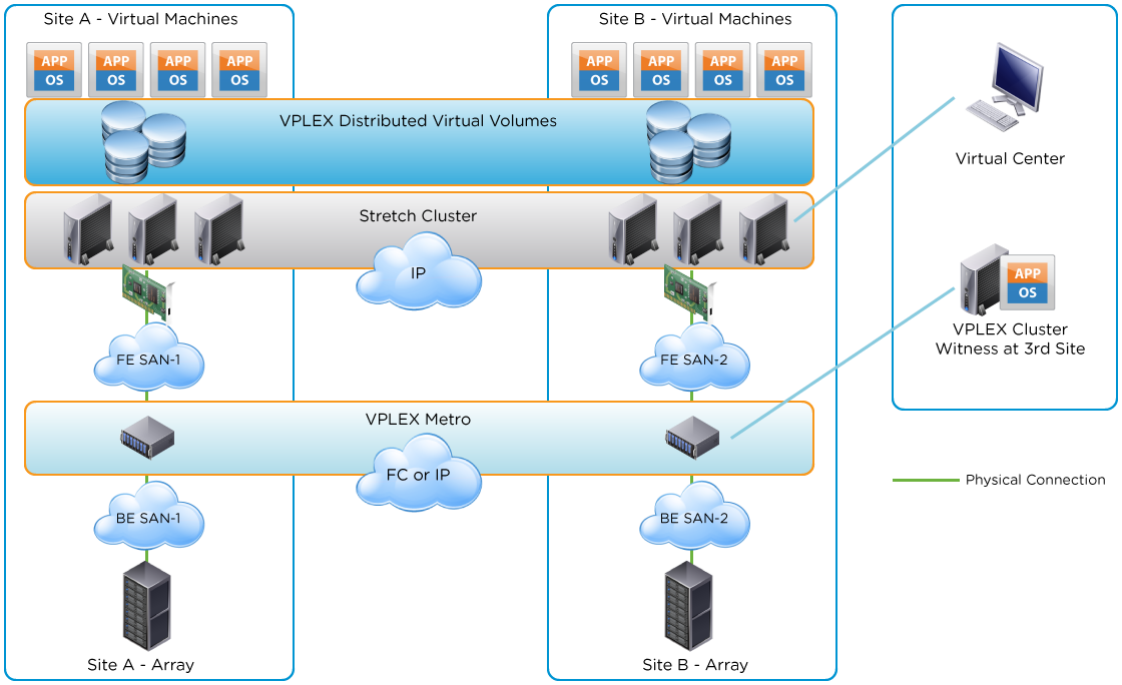
VMware ESXi inbox driver automatically recognizes the volumes after LUN-masking is done properly.

Configuring VMware vSphere cluster to work with VPLEX Metro solution

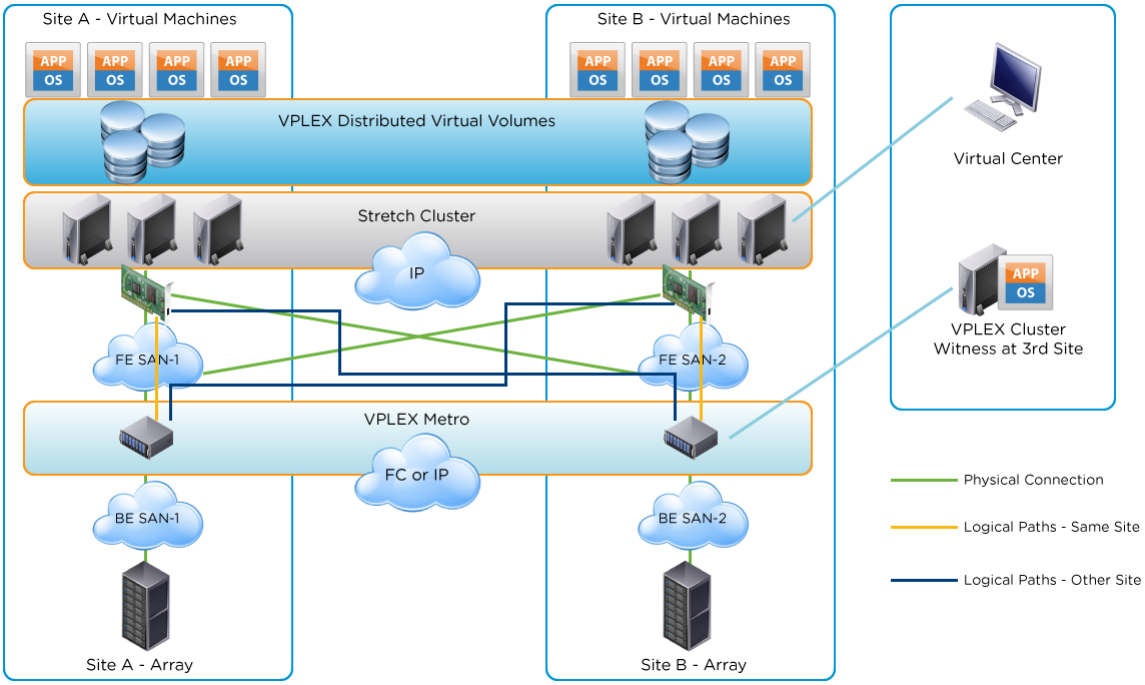
Introduction of the VPLEX witness to a high availability environment, enabling the VPLEX solution to increase the overall availability of the environment by arbitrating a pure communication failure between two primary sites and a true site failure in a multi-site architecture.

A VPLEX Metro solution spread across the two data centers provides the distributed storage to the ESXi hosts. Based on the host SAN connections to the VPLEX storage cluster, there are two different types of deployments as follows:

- **Non-uniform host access**- This type of deployment involves the hosts at either site and see the storage volumes through the same site storage cluster only. The following figure illustrates a typical configuration that is validated by Dell EMC using host clusters with a VPLEX Metro.



- **Uniform host access (Cross-Connect)**-This deployment involves establishing a front-end SAN across the two sites, so that the hosts at one site can see the storage cluster and the other site at the same.



For more information, see the VMware Knowledge Base article 2007545.

Configuring VMware vSphere cluster parameters in non-uniform host access deployment

By default, a VMware HA/DRS cluster is created across the two sites using ESXi 6.0, ESXi 6.5 or ESXi 6.7 hosts and managed by vCenter Server 6.0, 6.5 or 6.7. The ESXi boot disk is located in the internal drives or local volume specific to the hosts and not in the Distributed Virtual Volume itself. The virtual machine runs on the preferred site of the Distributed Virtual Volume. VPLEX consistency group has auto resume set to true.

When VPLEX inter-site link fails, paths between vSphere cluster nodes and non-preferred VPLEX cluster site fail and I/O is suspended. It protects data if VPLEX is not able to sync data between VPLEX cluster nodes at two sites. The paths remain inaccessible even after inter-

site link failure recovers. This is because the virtual machine opens handles on VPLEX devices and does not allow the VPLEX devices to come back online after an unplanned permanent device loss (PDL). In that case, you need to reboot the vSphere cluster nodes connecting to non-preferred site to recover.

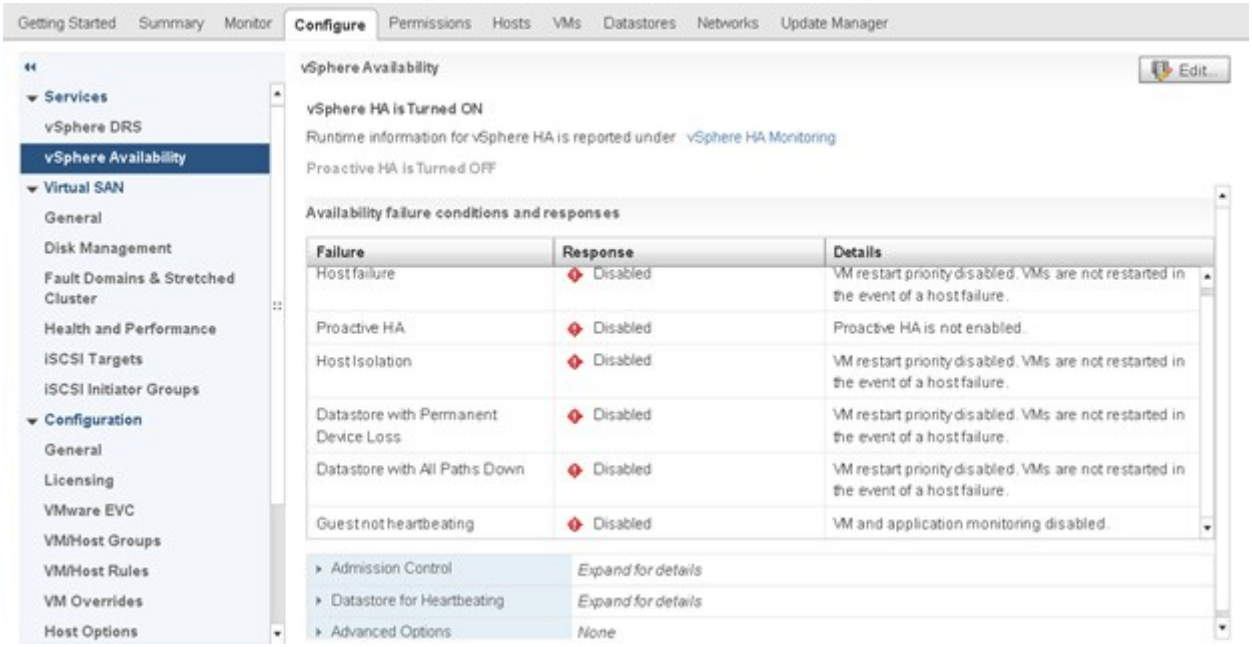
To avoid this scenario, change the default parameters to recover after an unplanned PDL.

Configuring VMware vSphere 6.5 and 6.7

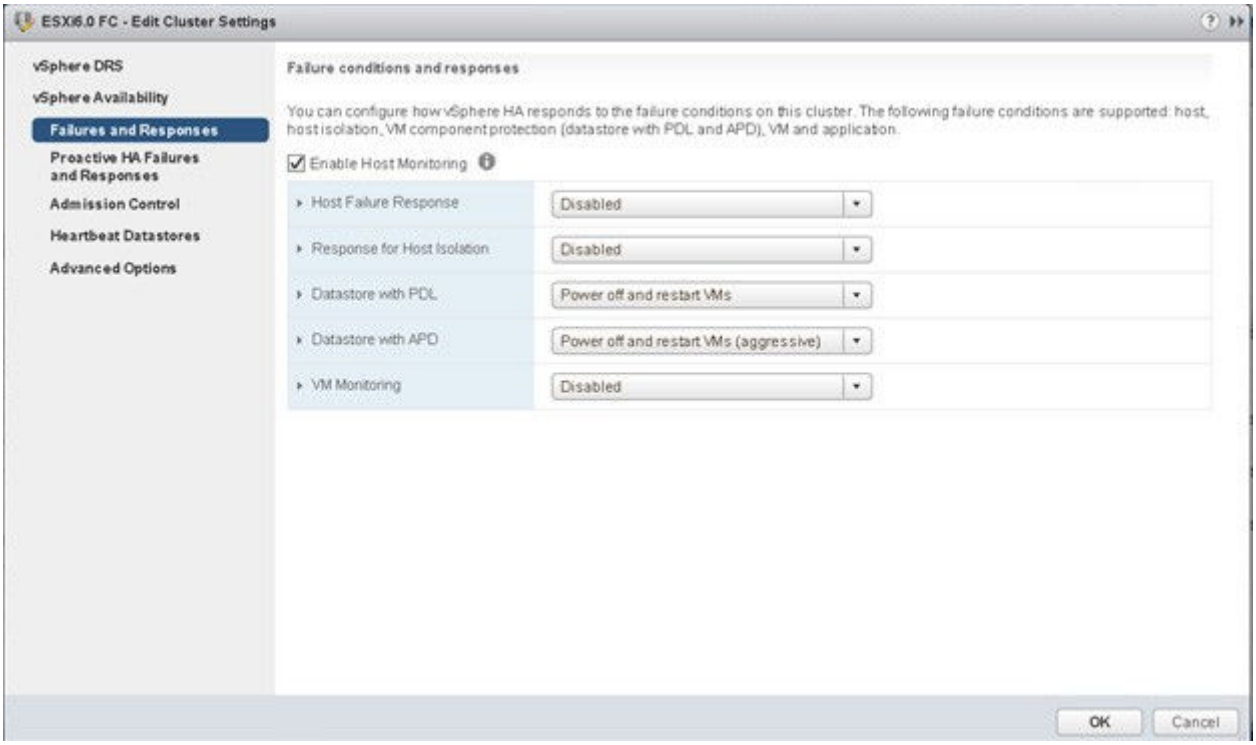
Perform the following steps to configure VMware vSphere cluster parameters in Non-uniform Host Access deployment:

Steps

1. Log in to vCenter Server with vSphere Web Client.
2. Right-click **vSphere cluster** and click **Settings** on the dialog box.
3. Navigate to **vSphere Availability** under **Service** on the left panel as shown in the following figure:



   The **Availability failure conditions and responses** table lists the default parameters.
4. On the upper right corner, click **Edit**.
5. On the left panel, under **vSphere Availability**, click **Failures and Responses**.
   Failure conditions and responses page displays as shown in the following figure:



6. Set the following values:
   - **Datastore with PDL** = Power off and restart VMs
   - **Datastore with APD** = Power off and restart VMs
7. Click **OK**.

Configuring VMware vSphere 6.0

Perform the following steps to configure VMware vSphere cluster parameters:

Steps

1. Log in to vCenter Server with vSphere Web Client.
2. Right click **vSphere cluster** and click **Settings** on the dialog box.
3. Navigate to **vSphere HA** under **Service** on the left panel.
4. On the upper right corner, click **Edit**.

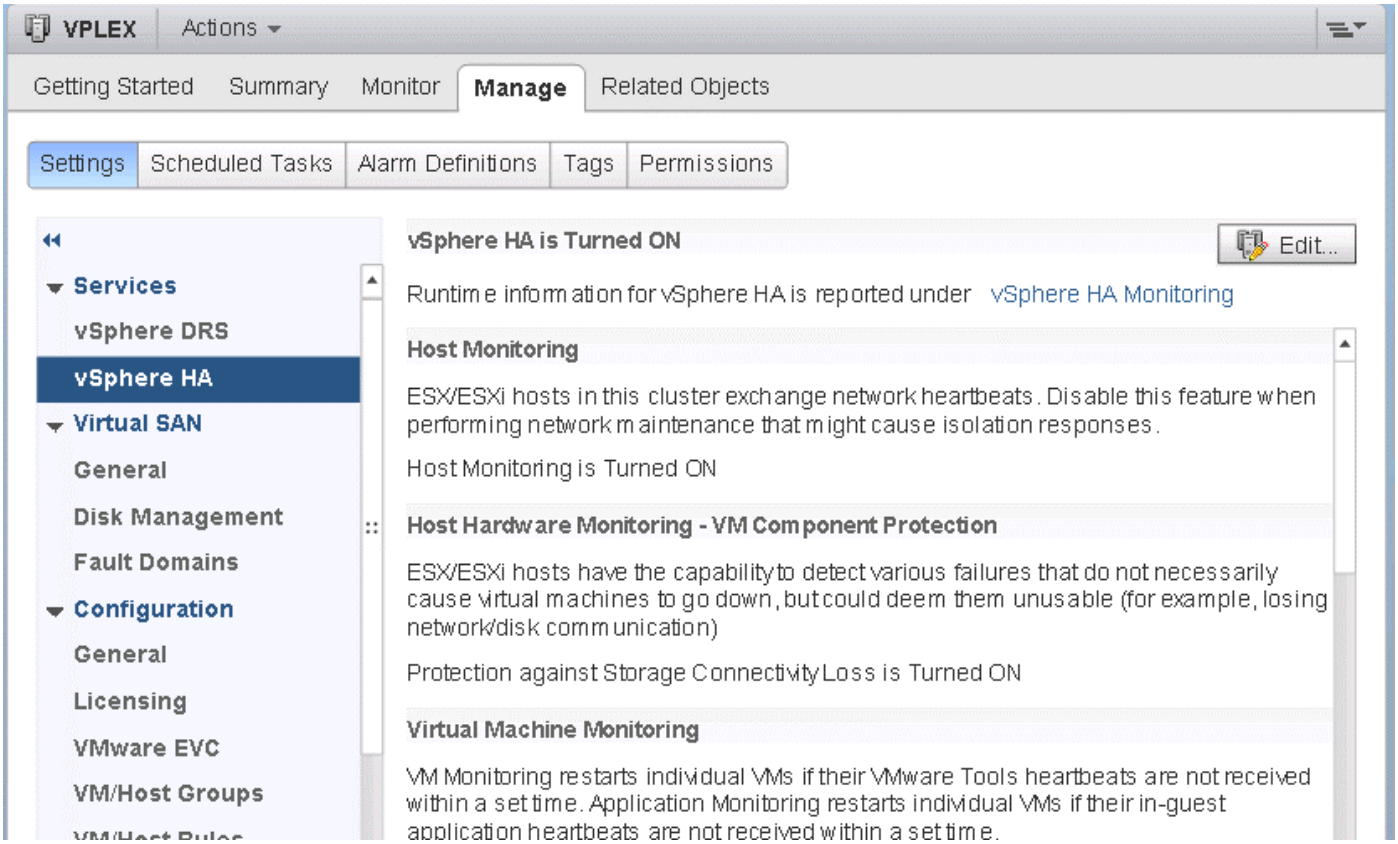


Figure 14. Default parameters for availability failure conditions and responses in vSphere 6.0

5. Under **Virtual Machine Monitoring**, select **VM Monitoring Only**, then expand **Failure conditions** and **VM response**.
6. Set the following values:
   - **Response for Datastore with Permanent Device Loss(PDL)** = Power off and restart VMs
   - **Response for Datastore with All Paths Down(APD)** = Power off and restart VMs (aggressive)

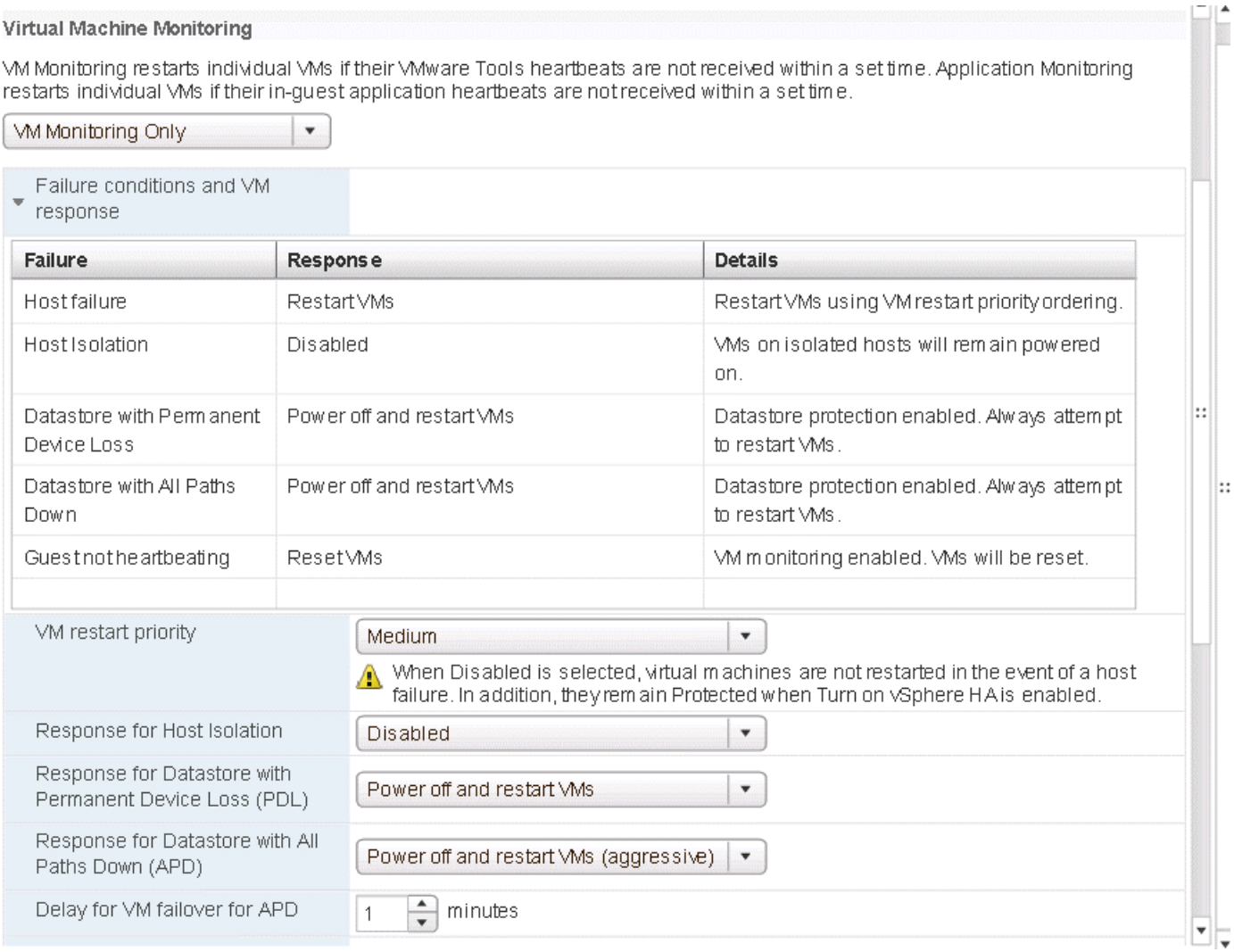


Figure 15. Failure conditions and Responses parameters setting in vSphere 6.0

7. Click **OK**.

XtremIO

This section describes host connectivity of the Dell EMC XtremIO in VMware environment.

NOTE: In hosts running a hypervisor, such as VMware ESXi or Microsoft Hyper-V, it is important to ensure that the logical unit numbers (LUNs) of XtremIO volumes are consistent across all hosts in the hypervisor cluster. Inconsistent LUNs may affect operations such as VM online migration or VM power-up.

NOTE: When using Jumbo Frames with VMware ESXi, the correct MTU size must be set on the virtual switch as well.

Best practices for zoning and subnetting

This section describes the best practices for allowing a host and the XtremIO cluster to communicate using 4, 8 or 16 paths per device.

Recommended configuration values summary

The following table summarizes all used and recommended variables and their values for zoning and subnetting:

Validation	Impact	Severity
Multipathing: Max number of paths shall not exceed 16.	Performance	Warning
Multipathing: Recommended number of paths is 8.	Performance	Warning
Multipathing: Link speed should be consistent across all paths to the XtremIO cluster.	Performance	Warning
Multipathing: Duplex setting should be consistent across all paths to the XtremIO cluster.	Performance	Warning
Balance the hosts between the Storage Controllers to provide a distributed load across all target ports.	Performance	Recommendation
FC-configured zoning: Each host Initiator Group should be at least one path for two Storage Controllers belonging to the same X-Brick.	Stability	Mandatory
iSCSI-configured subnetting: Each host Initiator Group should be at least one path for two Storage Controllers belonging to the same X-Brick.	Stability	Mandatory
iSCSI subnetting: Configuration should not allow traffic between different iSCSI IP subnets.	Performance	Normal
iSCSI MTU value: If jumbo frames are required for iSCSI traffic, all ports (server, switches, and storage) must be configured with the correct MTU value.	Performance	Warning
iSCSI Flow control features: Flow control must be disabled on all ports (server, switches, and storage).	Performance	Warning
iSCSI TCP Offloading: Enable the TCP Offloading Engine (TOE) on the host iSCSI interfaces.	Performance	Warning
iSCSI NIC configuration: Use a dedicated NIC or iSCSI HBA for XtremIO iSCSI and do not partition the iSCSI interface (that is disable NIC Partitioning - NPAR). Dell EMC recommends using interfaces individually rather than using NIC Teaming (Link Aggregation).	Performance	Recommendation

General guidelines

- The optimal number of paths depends on the operating system and server information. To avoid multipathing performance degradation, do not use more than 16 paths per device. Dell EMC recommends using eight paths.

  NOTE: This recommendation is not applicable to Linux hosts connected to XtremIO. On such hosts, more than 16 paths per device can be used (if required).
- Balance the hosts between the Storage Controllers to provide a distributed load across all target ports.
- Host I/O latency can be severely affected by SAN congestion. Minimize the use of ISLs by placing the host and storage ports on the same physical switch. When this is not possible, ensure that there is sufficient ISL bandwidth and that both the Host and XtremIO interfaces are separated by no more than two ISL hops. For more information about proper SAN design, see the Networked Storage Concepts and Protocols Techbook.
- Keep a consistent link speed and duplex across all paths between the host and the XtremIO cluster.
- To ensure continuous access to XtremIO storage during cluster software upgrade, verify that a minimum I/O timeout of 30 seconds is set on the HBAs of all hosts that are connected to the affected XtremIO cluster. Similarly, verify that a minimum timeout of 30 seconds is set for all applications that are using storage from the XtremIO cluster.

NOTE: See the Dell EMC KB article 167514 for references to *Dell EMC Host Connectivity Guides*. These guides provide the procedures that are required for adjusting the HBA minimum I/O timeout.

Minimal zoning/subnetting configuration

To prevent a host path from going down when two Storage Controllers (from separate X-Brick blocks) are down, while the XtremIO cluster remains active (because the failed Storage Controllers are not in the same X-Brick), follow these guidelines:

- When configuring zoning/subnetting from the host to the XtremIO cluster, the minimal zoning/subnetting configuration for each host Initiator Group should be at least one path for two Storage Controllers belonging to the same X-Brick.
- A host port must be zoned to at least two Storage Controllers ports from the same X-Brick. This zone can be expanded to additional Storage Controllers from other X-Brick blocks. Moreover, other host ports can be zoned to Storage Controllers from other X-Brick blocks.

NOTE: The diagrams throughout this chapter illustrate possible implementations of these guidelines. Other possible implementations exist, but are not illustrated.

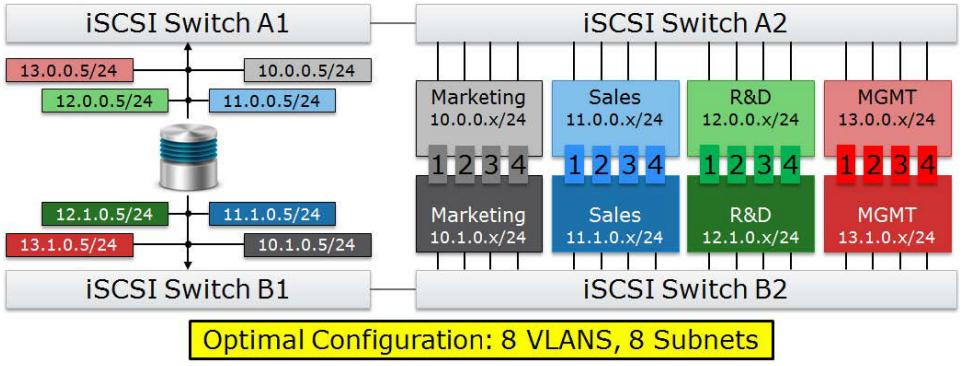
iSCSI SAN guidelines

When setting up an iSCSI SAN infrastructure, follow these guidelines:

- If jumbo frames are required for iSCSI traffic, ensure that all ports (server, switches, and storage) are configured with the correct MTU value.

  NOTE: When using Jumbo Frames with VMware ESXi, the correct MTU size must be set on the virtual switch as well.
- Disable flow control features on the server, switches, and array ports.
- Make sure that the different iSCSI IP subnets cannot transmit traffic between them.
- Use Virtual LAN (VLAN) technology to partition a single iSCSI network link into multiple distinct domains. If possible, use a dedicated VLAN for XtremIO iSCSI traffic and a dedicated IP subnet for each iSCSI fabric. Do not configure iSCSI routes between the different subnets.

The following figure shows optimal VLAN and IP Subnetting configuration:



- Enable the TCP Offloading Engine (TOE) on the host iSCSI interfaces, to offload the TCP packet encapsulation from the CPU of the Host to the NIC or iSCSI HBA, and free up CPU cycles.
- Dell EMC recommends using a dedicated NIC or iSCSI HBA for XtremIO iSCSI and not to partition the iSCSI interface (in other words, disable NIC Partitioning - NPAR).
- When using XtremIO iSCSI, Dell EMC recommends using interfaces individually rather than using NIC Teaming (Link Aggregation), to combine multiple interfaces into a single virtual interface.

NOTE: See the user manual of the FC/iSCSI switch for instructions about real implementations.

Fibre Channel SAN guidelines

When setting up a Fibre Channel (FC) SAN infrastructure, follow these guidelines:

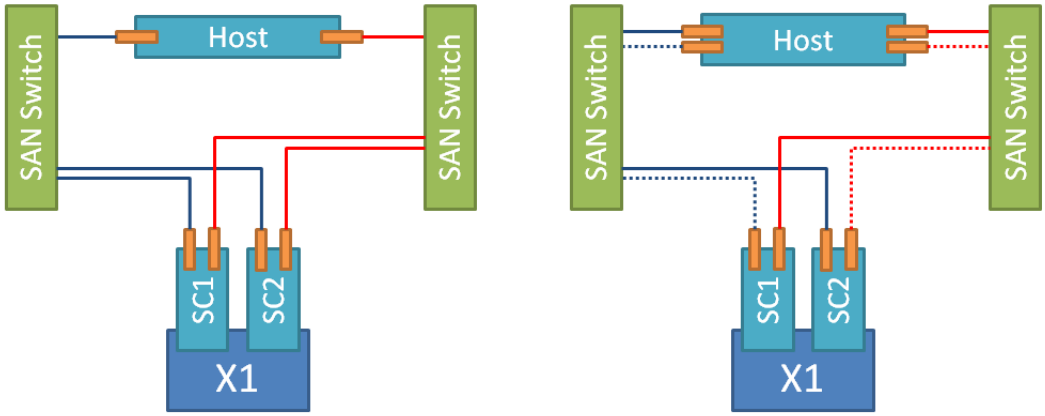
- Use a single-initiator per single-target (1:1) zoning scheme. If the FC switch zone count limitation has been reached, it is also possible to use single-initiator per multiple-target (1: many) zoning scheme.

NOTE: See the user manual of the FC/iSCSI switch for instructions about real implementations.

10 TB starter X-Brick (5 TB) and single X-Brick cluster

In a 10 TB starter X-Brick (5 TB) or a single X1 X-Brick configuration, a host may have up to four paths per device. On an X2 X-Brick configuration with all iSCSI or FC, a host may have up to eight paths per device.

The following figure shows the logical connection topology for four paths. This topology applies to both dual and quad HBA/NIC host architecture:

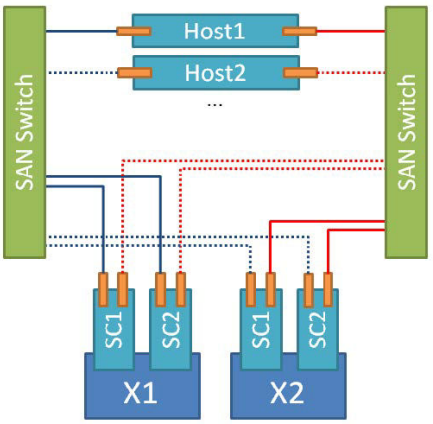


Dual X-Brick clusters

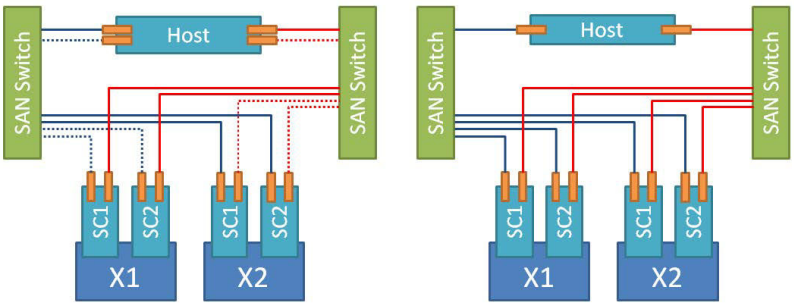
In a dual X-Brick configuration, a host may have up to eight paths per device.

NOTE: When partially zoning/subnetting multiple servers to the XtremIO array, ensure that the I/O load of the server is distributed equally across all X-Brick. For minimal zoning/subnetting configuration guidelines, see Minimal zoning/subnetting configuration.

The following figure shows the logical connection topology for four paths. This topology applies to a dual HBA/NIC host architecture:



The following figure shows the logical connection topology for eight paths. This topology applies to both dual and quad HBA/NIC host architecture:

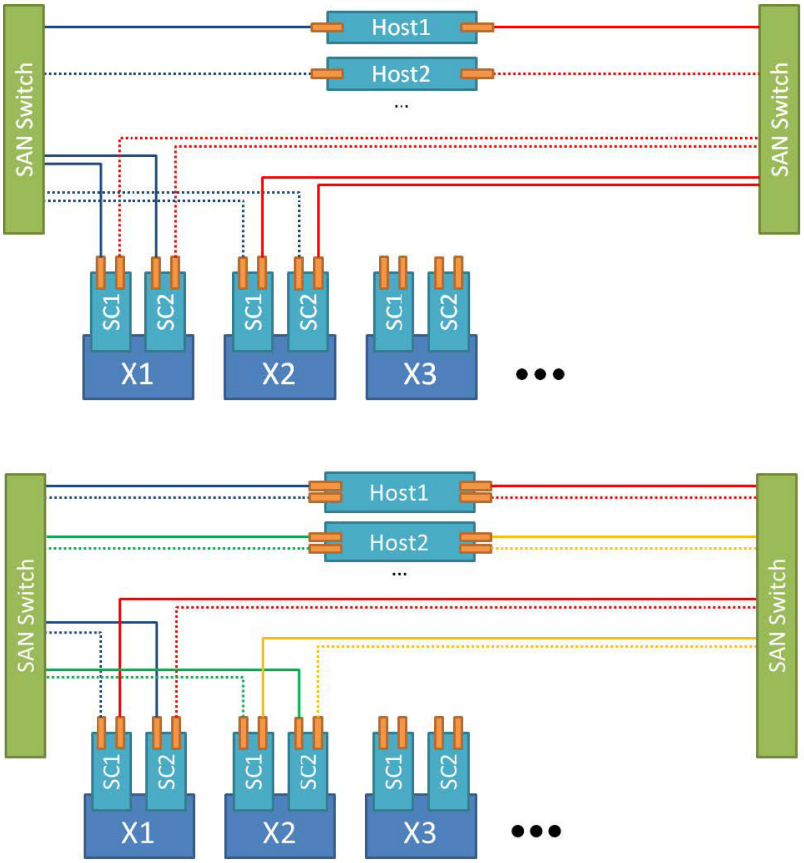


Clusters with multiple X-Brick blocks(three or more)

In a multiple X-Brick configuration (three or more), a host may have up to 16 paths per device.

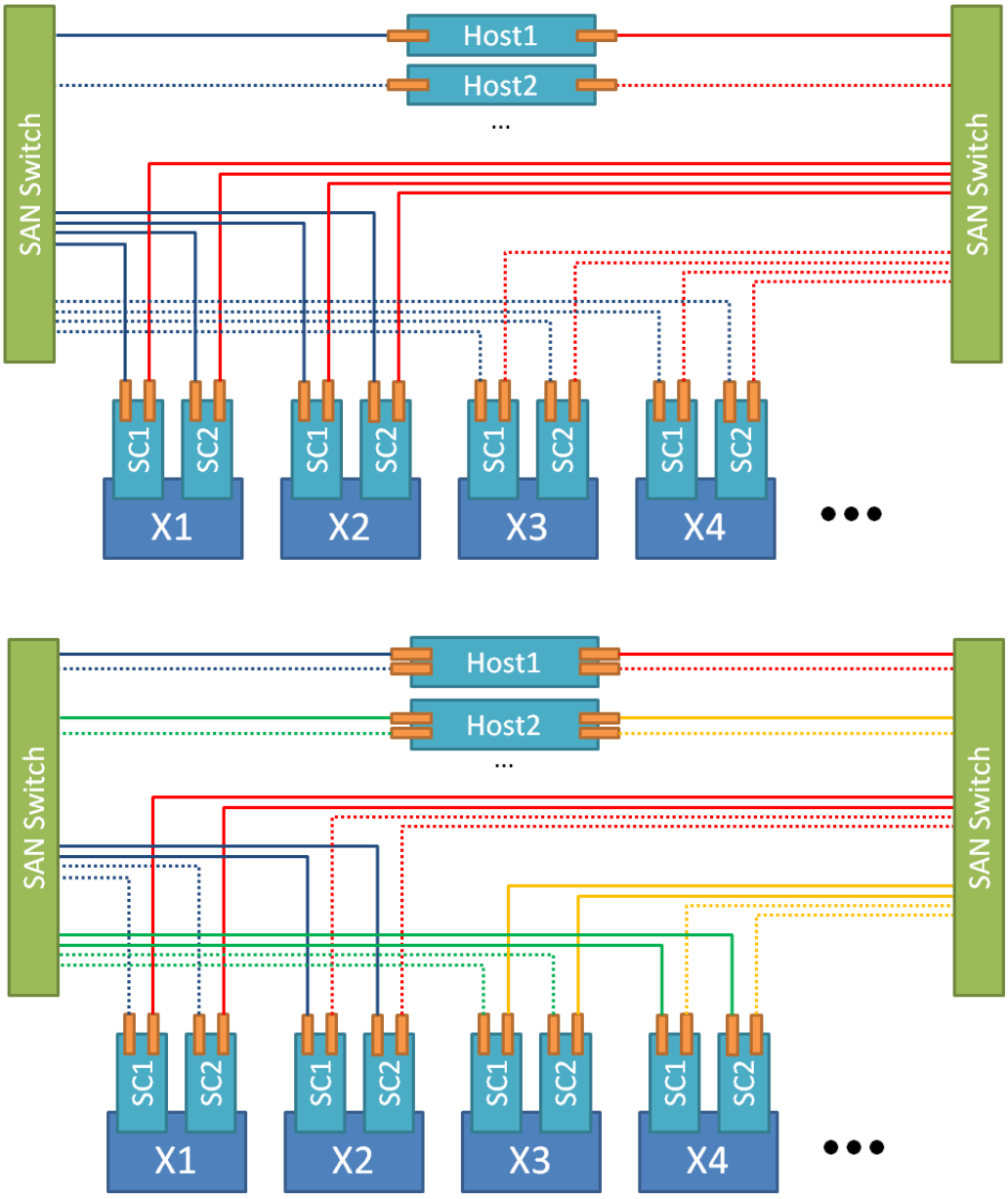
NOTE: When partially zoning/subnetting multiple servers to the XtremIO array, ensure that the I/O load of the server is distributed equally across all X-Brick blocks. For minimal zoning/subnetting configuration guidelines, see Minimal zoning/subnetting configuration.

The following figures show the logical connection topology for four paths. This topology applies to both dual and quad HBA/NIC host architecture:



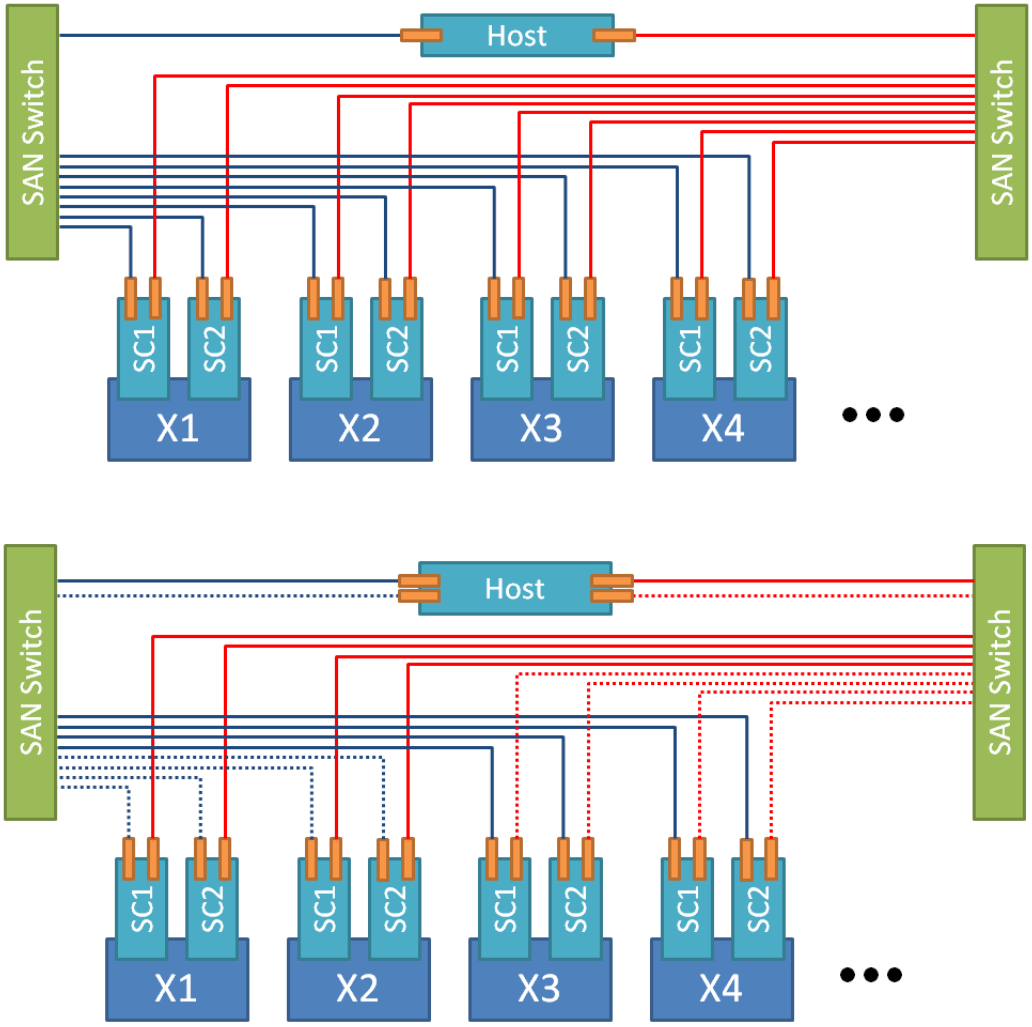
The following figures show the logical connection topology for eight paths. This topology applies to both dual and quad HBA/NIC host architecture.

NOTE: For clusters with an odd number of X-Brick blocks, change these examples to accommodate the cluster configuration and try to balance the host load among the X-Brick blocks of the cluster. The following figures show an eight paths connection topology with four X-Brick blocks (or more):



The following figures show the logical connection topology for 16 paths. This topology applies to both dual and quad HBA/NIC host architecture.

NOTE: For clusters with an odd number of X-Brick blocks, change these examples to accommodate the cluster configuration and try to balance the host load among the X-Brick blocks of the cluster. The following figure shows a 16-path connection topology with four X-Brick blocks (or more):



Recommended configuration values summary

The following table summarizes all used variables and their values when configuring hosts for VMware vSphere:

NOTE: The variable setting recommendations that are detailed in this section can be validated by running the ESXi Host Validation Script on ESXi hosts connected to XtremIO. For details, see Executing the ESXi Host Validation Script on ESXi Hosts.

Validation	Impact	Severity
Timeout: Minimum I/O of 30 s is set on HBAs of all hosts.	Stability	Mandatory
Hypervisor assigned LUNs: The LUN of XtremIO volumes should be consistent across all hosts in the hypervisor cluster.	Stability	Mandatory
iSCSI configuration: Configure end-to-end Jumbo Frames.	Performance	Recommendation
iSCSI configuration: Disable Delayed ACK on ESXi.	Stability	Recommendation
iSCSI configuration: Adjust LoginTimeOut to 30.	Stability	Recommendation
LUN Queue Depth recommended values: • Qlogic: 256 • Emulex: 128 • UCS: 32 for MultiStorage Setup and 128 for XtremIO Only Setup	Performance	Warning
HBA Queue depth recommended values:	Performance	Warning

Validation	Impact	Severity
• Qlogic: N/A • Emulex: 8192 (same as *lpfc_hba_queue_depth*) • UCS: 1024 - I/O throttle		
ESXi Configuration: *config.vpxd.ResourceManager.maxCostPerHost* • 10 TB Starter X-Brick (5 TB) and a Two X-Brick cluster - 16 concurrent full clone operations • Two X-Brick blocks - 16 concurrent full clone operations • Four X-Brick blocks - 32 concurrent full clone operations • Six X-Brick blocks - 48 concurrent full clone operations • MultiStorage Setting - 8	Performance	Recommendation
ESXi Configuration: • DataMover.HardwareAcceleratedMove= 1 • DataMover.HardwareAcceleratedInit= 1 • VMFS3.HardwareAcceleratedLocking=1	Performance	Mandatory
ESXi Configuration: VAAI should be enabled.	Performance	Mandatory
ATS setting: vmkfstools mode=public ATS-only	Stability	Mandatory
FC Adapter Policy IO Throttle Count: • Scope/Granularity - Per vHBA • MultiStorage Setting - 256 • XtremIO Only Setting - 1024	Performance	Recommendation
ESXi Configuration: *Disk.SchedNumReqOutstanding* • Granularity - LUN (global in vSphere 5.5 and earlier) • MultiStorage Setting - 32 • XtremIO Only Setting - Same as LUN queue depth (for vSphere 6.5 and later) or 256 (for vSphere 6.0 and earlier)	Performance	Warning
ESXi Configuration: *Disk.SchedQuantum* • Granularity - Global • MultiStorage Setting - 8 • XtremIO Only Setting - 64	Performance	Recommendation
ESXi Configuration: *Disk.DiskMaxIOSize* • Granularity - Global • MultiStorage Setting - 4096 (4 MB) • XtremIO Only Setting - 4096 (4 MB)	Stability and Performance	Mandatory
ESXi Configuration: *MaxHWTransferSizeDataMover/MaxHWTransferSize* • Scope/Granularity - Global • MultiStorage and XtremIO X2 clusters Setting - 4 MB • XtremIO X1 clusters only Setting - 256 KB	Performance	Performance
Path selection policy: Native round-robin path selection policy on XtremIO volumes that are presented to the ESXi host.	Stability and Performance	Mandatory
Path switching: NMP Round robin path switching frequency to XtremIO volumes set to 1.	Performance	Recommendation
Connectivity: Adjust Connectivity mode to APD (default setting with XIOS version 4.0.25-22 or later).	Stability	Recommendation
Alignment: Guest operating system virtual machines should be aligned.	Storage, efficiency, and performance	Warning

Validation	Impact	Severity
Validation: With XtremIO version 4.0 (or higher), specify ESXi as the operating system for each defined initiator.	Stability and serviceability	Mandatory
Virtual machine configuration: Configure virtual machines with Paravirtualized SCSI controllers.	Stability and Performance	Recommendation
Virtual machines guest operating system: Use the maximum queue depth of the virtual SCSI controller.	Performance	Recommendation
RDM volumes: In Guest operating system - Span RDM volumes used by the virtual machine across SCSI controllers.	Performance	Recommendation
Validation: For space reclamation at ESXi level with ESXi versions 5.5 through 6.0, set reclaim-unit argument to 20000.	Storage efficiency	Mandatory

iSCSI Configuration

This section describes the issues that should be addressed when using iSCSI with XtremIO, for optimal performance (with or without an iSCSI HBA).

NOTE: This section applies only for iSCSI. If you are using only Fibre Channel with vSphere and XtremIO, go to Fibre Channel HBA Configuration.

NOTE: See iSCSI SAN Guidelines, before you proceed.

Prerequisites

Follow the VMware recommendations for installation and setup of the appropriate NIC/iSCSI HBA for your system. Install the latest driver version (patch), as described in the VMware support site for each specific NIC/iSCSI HBA.

See Dell EMC E-Lab Navigator for details about supported NIC/iSCSI HBA models and drivers.

Jumbo Frames

When using iSCSI with ESXi hosts and XtremIO, Dell EMC recommends configuring end-to-end Jumbo Frames (MTU=9000) for optimal performance. With Jumbo Frames, Ethernet frames are set to be larger than 1500 bytes (up to 9000 bytes).

NOTE: When using Jumbo Frames, ensure that all ports (ESXi server, switch, and storage) are configured with the correct MTU value. With VMware ESXi, the correct MTU size must be set on the virtual switch as well.

For further details on configuring Jumbo Frames with iSCSI on ESXi, see the VMware Knowledge Base article 1007654.

Delayed ACK

For optimal iSCSI traffic between the ESXi hosts and XtremIO, especially during periods of network congestion, Dell EMC recommends disabling the Delayed ACK on ESXi. By disabling Delayed ACK, the ESXi host would send an ACK acknowledgment segment for every received data segment rather than delaying the sending of ACK acknowledgment segments, while receiving a stream of TCP data segments.

For further details on the Delayed ACK vSphere parameter and how to disable it, using the vSphere Client, see the VMware Knowledge Base article 1002598.

NOTE: The recommended method for configuring the delayed ACK vSphere setting is per discovered iSCSI target. This allows to disable delayed ACK only for XtremIO iSCSI targets.

Login Timeout

When establishing an iSCSI session between the initiator and target, the Login Timeout vSphere setting controls for how long the ESXi host attempts to login to the iSCSI target before failing the login and retrying.

The default setting for `LoginTimeOut` is **5**, that is, by default an iSCSI session ceases retries after 20 seconds (5 times the `LoginRetryMax` vSphere setting, which is by default set to 4).

To optimize iSCSI session behavior with XtremIO and to better handle periods of network disruptions and NDU, adjust the `LoginTimeOut` to **30**.

Adjust LoginTimeout using vSphere Client on ESXi 5.x or earlier versions

Perform the following steps to adjust `LoginTimeOut` using vSphere Client on ESXi 5.x or earlier versions:

Steps

1. Launch vSphere Client and connect to the ESXi host.
2. Go to **Configure** > **Storage Adapters** > **iSCSI Software Adapter** > **Properties.**
3. Under **Adapter Details**, select **Advanced Options**. Click **Edit** and scroll down to LoginTimeout.
4. Change the value from 5 seconds to 30 seconds.

Adjust LoginTimeOut using command line on ESXi 6.0 or later versions

Perform the following procedures to adjust LoginTimeOut, using command line on ESXi 6.0 or later:

Steps

1. Connect to the ESXi host as *root*.
2. Run the following command:

   ```
   esxcli iscsi adapter param set -A adapter_name -k LoginTimeout -v value_in_sec
   ```

   Example:

   ```
   esxcli iscsi adapter param set -A vmhba64 -k LoginTimeout -v 30
   ```

Fibre Channel HBA Configuration

When using Fibre Channel with XtremIO, the following FC Host Bus Adapters (HBA) issues should be addressed for optimal performance.

NOTE: This section applies only for Fibre Channel. If you are using only iSCSI with vSphere and XtremIO, go to Host Parameters Settings and continue with the rest of this chapter.

Prerequisites

To install one or more Dell EMC-approved HBAs on an ESXi host, follow the procedures in one of these documents, according to the FC HBA type:

- For Qlogic and Emulex HBAs - Typically the driver for these HBAs is preloaded with ESXi. Therefore, no further action is required. For details, see the vSphere and HBA documentation.
- For Cisco UCS fNIC HBAs (vSphere 5.x and later) - see the Cisco UCS Virtual Interface Card Drivers for ESXi Installation Guide for complete driver installation instructions.

NOTE: To avoid latency or connectivity loss issues between ESXi hosts with Cisco UCS fNIC FC HBAs and XtremIO storage, Dell EMC recommends upgrading the Cisco UCS fNICs driver to version 1.6.0.25 (or later). For details, see the Dell EMC Knowledge Base article 494792.

See Dell EMC E-Lab Navigator for details about supported FC HBA models and drivers.

Queue Depth

NOTE: The FC HBA recommendations in this section are applicable to the following FC HBAs:

- **Qlogic - adapters with names that start with *ql***
- **Emulex - adapters with names that start with *lpfc***
- **Cisco - adapters with names that start with *fnic***

NOTE: For Cisco UCS nfNIC adapters that start with nfnic, Cisco UCSnfNIC driver does not allow to change the queue depth setting for adapter. Contact Cisco for more information.

See Dell EMC E-Lab Navigator for all supported FC HBA models and drivers.

NOTE: Changing queue depth settings is designed for advanced users. Increasing queue depth may cause hosts to overstress other arrays that are connected to the ESXi host, resulting in performance degradation while communicating with them. To avoid this, in mixed environments with multiple array types that are connected to the ESXi host, compare the XtremIO recommendations with those of other platforms before applying them.

Queue depth is the number of SCSI commands (including I/O requests) that can be handled by a storage device at a given time. A queue depth can be set on either of the following:

- Initiator level - HBA queue depth
- LUN level - LUN queue depth

The LUN queue depth setting controls the amount of outstanding I/O requests per a single path. On vSphere, the LUN queue depth can be adjusted through the ESXi CLI.

The HBA queue depth (also referred to as execution throttle) settings control the number of outstanding I/O requests per HBA port.

The HBA queue depth should be set to the maximum value. This can be done on the HBA firmware level, using the HBA BIOS or CLI utility provided by the HBA vendor:

- Qlogic - Execution Throttle - This setting is no longer read by vSphere and is not relevant when configuring a vSphere host with Qlogic HBAs.
- Emulex - lpfc_hba_queue_depth - No need to change the default (and maximum) value (8192).
- Cisco UCS fNIC - The I/O Throttle setting determines the total number of outstanding I/O requests per virtual HBA. With Cisco UCS fNIC, Dell EMC recommends setting the I/O throttle to 1024.

For optimal operation with XtremIO storage, it is also recommended to adjust the LUN queue depth of the FC HBA.

NOTE: For further information about adjusting LUN queue depth with ESXi, see the VMware Knowledge Base article 1267.

NOTE: If the HBA queue depth is set to a value lower than the LUN queue depth, it may limit the LUN queue depth to a lower value than set.

NOTE: The setting adjustments for Cisco UCS fNIC HBA detailed in this section, apply to VMware vSphere only. Since these settings are global to the UCS chassis, they may impact other blades in the UCS chassis running a different operating system (for example Windows).

The following table summarizes the default and recommended queue depth settings for VMware vSphere:

Table 7. Queue depth settings for VMware vSphere

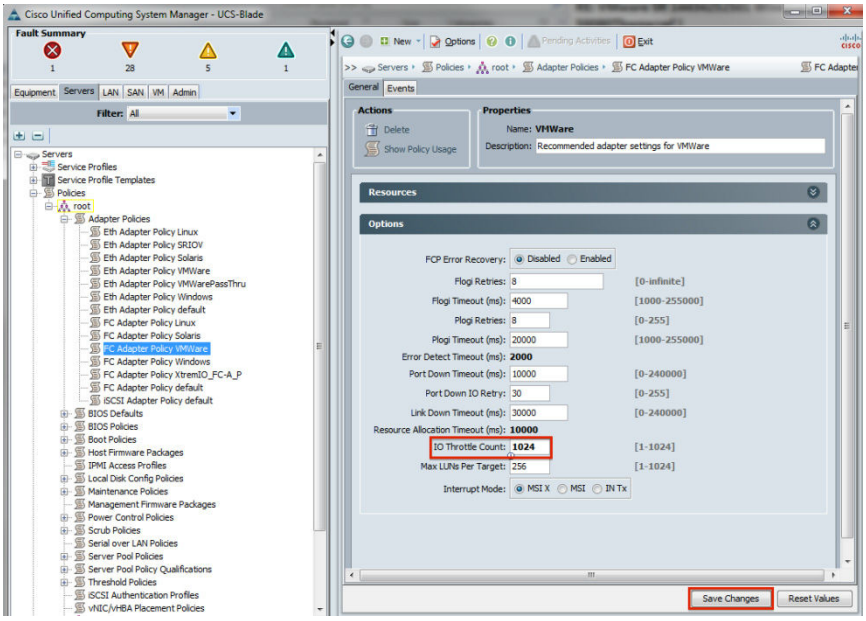
	LUN Queue Depth	HBA Queue Depth
Default Value	Qlogic: 64 Emulex: 30 UCS: 32	Qlogic: N/A Emulex: 8192 UCS: 16
Recommended Value	Qlogic: 256 Emulex: 128 UCS: 128	Qlogic: N/A Emulex: 8192 (maximum) UCS: 1024

Setting the HBA I/O Throttle of the Cisco UCS fNIC HBA

Perform the following procedures to set HBA I/O throttle of the Cisco UCS fNIC HBA:

Steps

1. In the UCSM navigation tree, click the **Servers** tab.
2. In the navigation tree, expand the **Policies and Adapter Policies**.
3. Click the **FC Adapter Policy Linux** or **FC Adapter Policy VMWare**.
4. In the main window, expand the **Options** drop-down.
5. Configure the **I/O Throttle Count** field to `1024`.
6. Click **Save Changes**.



NOTE: For more details about Cisco UCS fNIC FC adapter configuration, see the Cisco UCS fNIC Tunables Guide.

Setting the LUN Queue Depth on a host running vSphere 5.x or later

Perform the following procedures to set the LUN queue depth on a host running vSphere 5.x or later version:

Steps

1. Open an SSH session to the host as *root* .
2. Run one of the following commands to verify which HBA module is loaded:

HBA Vendor	Command
Qlogic	`esxcli system module list | egrep "ql|Loaded"`
Emulex	`esxcli system module list | egrep "lpfc|Loaded"`
Cisco UCS fNIC	`esxcli system module list | egrep "fnic|Loaded"`

Example (for a host with Emulex HBA):

```
# esxcli system module list | egrep "lpfc|Loaded"
Name                              Is Loaded        Is Enabled
lpfc                              true             true
lpfc820                           false            true
```

In this example, the native lpfc module for the Emulex HBA is loaded on ESXi.

3. Run one of the following commands on the currently loaded HBA module to set the LUN queue depth:

NOTE: The commands in the following table refer to the *Qlogic qla2xxx/qlnativefc*, *Emulex lpfc*, and *Cisco UCS fNIC* modules. Use an appropriate module name based on the output of the previous step.

HBA Vendor	Command

Qlogic	vSphere 5.x:
	`esxcli system module parameters set -p ql2xmaxqdepth=256 -mqla2xxx`
	vSphere 5.5 (with Qlogic native drivers):
	`esxcli system module parameters set -p ql2xmaxqdepth=256 -m qlnativefc`
	vSphere 6.x:

HBA Vendor	Command
	`esxcli system module parameters set -p ql2xmaxqdepth=256 -m qlnativefc`
Emulex	`esxcli system module parameters set -plpfc0_lun_queue_depth=128 -m lpfc`
Cisco UCS fNIC	`esxcli system module parameters set -p fnic_max_qdepth=128 -m fnic`

NOTE: The command for Emulex HBA sets the LUN queue depth for the lpfc0 Emulex HBA. If another Emulex HBA is connected to the XtremIO storage, change *lpfc0_lun_queue_depth* accordingly. For example, if lpfc1 Emulex HBA is connected to XtremIO, replace l*pfc0_lun_queue_depth* with *lpfc1_lun_queue_depth.*

NOTE: If all Emulex HBAs on the host are connected to the XtremIO storage, replace *lpfc0_lun_queue_depth* with *lpfc_lun_queue_depth.*

4. Reboot the ESXi host.
5. Open an SSH session to the host as *root*.
6. Run the following command to confirm that queue depth adjustment is applied: `esxcli system module parameters list -m <driver>`

   NOTE: When using the command, replace <driver> with the module name, as received in the output of step 2 (for example, lpfc, qla2xxx and qlnativefc).

Example

Examples:

- For a vSphere 5.x host with Qlogic HBA and LUN queue depth set to 256:

```
# esxcli system module parameters list -m qla2xxx | grep ql2xmaxqdepth
ql2xmaxqdepth int 256 Max queue depth to report for target devices.
```

- For a vSphere 5.5 host with Qlogic HBA (with native drivers) and LUN queue depth set to 256:

```
# esxcli system module parameters list -m qlnativefc | grep ql2xmaxqdepth
ql2xmaxqdepth int 256 Maximum queue depth to report for target devices.
```

- For a vSphere 6.x host with Qlogic HBA and LUN queue depth set to 256:

```
# esxcli system module parameters list -m qlnativefc | grep ql2xmaxqdepth
ql2xmaxqdepth int 256 Maximum queue depth to report for target devices.
```

- For a host with Emulex HBA and LUN queue depth set to 128:

```
# esxcli system module parameters list -m lpfc | grep lpfc0_lun_queue_depth
lpfc0_lun_queue_depth int 128 Max number of FCP commands we can queue to a specific LUN
```

  If LUN queue depth is set for all Emulex HBAs on the host, run the following command instead:

```
# esxcli system module parameters list -m lpfc | grep
lpfc_lun_queue_depth
lpfc_lun_queue_depth int 128 Max number of FCP commands we
can queue to a specific LUN
```

Host parameters settings

This section details the ESXi host parameters settings necessary for optimal configuration when using XtremIO storage.

NOTE: The following settings may cause hosts to overstress other arrays that are connected to the ESXi host, resulting in performance degradation while communicating with them. To avoid the performance degradation in mixed environments with multiple array types that are connected to the ESXi host, compare these XtremIO recommendations with the other platforms before applying them.

When using XtremIO storage with VMware vSphere, Dell EMC recommends setting the following parameters to their maximum values:

- `Disk.SchedNumReqOutstanding` - Determines the maximum number of active storage commands (I/Os) allowed at any given time at the VMkernel. The maximum value is 256.

NOTE: In vSphere 5.5 or later, the `Disk.SchedNumReqOutstanding` parameter can be set on a specific volume rather than on all volumes that are presented to the host. It should be set only after XtremIO volumes are presented to the ESXi host using ESXi command line.

NOTE: When using vSphere 6.5 or later, the `SchedNumReqOutstanding` parameter should be equal to or lower than the LUN queue depth.

- `Disk.SchedQuantum` - Determines the maximum number of consecutive sequential I/Os allowed from one VM before switching to another VM (unless this is the only VM on the LUN). The maximum value is 64.

In addition, the following parameter setting is required:

- `Disk.DiskMaxIOSize` - Determines the maximum I/O request size passed to storage devices. With XtremIO, it is required to change it from 32,767 (default setting of 32 MB) to 4,096 (4 MB). This adjustment allows a Windows VM to EFI boot from XtremIO storage with a supported I/O size of 4 MB.

  NOTE: For details on the possible Windows EFI boot issue with XtremIO storage (in case the above maximum I/O block size setting adjustment cannot be done), see the VMware Knowledge Base article 2137402.

  NOTE: For details on adjusting the maximum I/O block size in ESXi, see the VVMware Knowledge Base article 1003469.

These setting adjustments should be carried out on each ESXi host that is connected to XtremIO cluster using either the vSphere Client or the ESXi command line.

NOTE: When using Cisco UCS Blade servers with vSphere and XtremIO, Dell EMC recommends modifying the `Max LUNs per target` parameter to 1024 on the UCS server profile. This modification is to avoid a possible data unavailability to Cisco UCS Blade servers during an upgrade to VMware vSphere version 6.7 U1. For more information, contact Cisco Support.

Adjusting the ESXi host parameters for XtremIO storage

To adjust the ESXi host parameters for XtremIO storage, perform one of the following procedures that is applicable to your system:

Adjust the ESXi Host Parameters for XtremIO Storage using the vSphere WebUI client

Perform the following procedures to adjust the ESXi host parameters for XtremIO storage:

Steps

1. Launch the vSphere Web client and go to **Home** > **Hosts and Clusters**.
2. In the left menu section, locate the **ESXi host** and click it.
3. In the right pane, click **Manage** > **Settings**.
4. From the **System** section, click **Advanced System Settings**.
5. Locate the `Disk.SchedNumReqOutstanding` parameter. Click the **Edit** icon and set the parameter to its maximum value (256).

   NOTE: Do not perform this step in a host using vSphere 5.5 or later, where the parameter is set on a specific volume using ESXi command line.

6. Locate the `Disk.SchedQuantum` parameter. Click the **Edit** icon and set it to its maximum value (64).
7. Locate the `Disk.DiskMaxIOSize` parameter. Click the **Edit** icon and set it to 4096.
8. Click **OK** to apply the changes.

Adjust the ESXi host parameters for XtremIO storage using ESXi host command line (for vSphere 5.0 and 5.1)

Perform the following procedures to adjust the ESXi host parameters for XtremIO storage:

Steps

1. Open an SSH session to the host as *root*.

2. Run the following commands to set the `SchedQuantum`, `SchedNumReqOutstanding`, and `DiskMaxIOSize` parameters, respectively:

```
esxcli system settings advanced set --int-value 64
--option /Disk/SchedQuantum
```

```
esxcli system settings advanced set --int-value 256
--option /Disk/SchedNumReqOutstanding
```

```
esxcli system settings advanced set --int-value 4096
--option /Disk/DiskMaxIOSize
```

Adjust the ESXi Host parameters for XtremIO storage using the ESXi host command line (for vSphere 5.5 or later)

Perform the following procedures to adjust the ESXi host parameters for XtremIO storage:

Steps

1. Open an SSH session to the host as *root* .
2. Run the following commands to set the `SchedQuantum` and `DiskMaxIOSize` parameters, respectively:

```
esxcli system settings advanced set --int-value 64
--option /Disk/SchedQuantum
```

```
esxcli system settings advanced set --int-value 4096
--option /Disk/DiskMaxIOSize
```

3. Run the following command to obtain the NAA for XtremIO LUNs presented to the ESXi host and locate the NAA of the XtremIO volume:

```
esxcli storage nmp path list | grep XtremIO -B1
```

4. For vSphere versions 6.0 and above, run the following command to verify that the LUN queue depth of your device is set correctly:

   `esxcli storage core device list |grep "XtremIO Fibre" -A 31 |grep "XtremIO Fibre\|Device Max Queue Depth:"`

   Example:

```
# esxcli storage core device list |grep "XtremIO Fibre" -A
31 |grep "XtremIO Fibre\|Device Max Queue Depth:"
     Display Name: XtremIO Fibre Channel Disk
(naa.514f0c5291800001)
     Device Max Queue Depth: 128
     Display Name: XtremIO Fibre Channel RAID Ctlr
(naa.514f0c5000000000)
     Device Max Queue Depth: 128
```

   If the LUN queue depth is incorrect, see Setting the LUN Queue Depth on a Host Running vSphere 5.x or Above.

5. For each LUN, run the following command to set the `SchedNumReqOutstanding` parameter according to the vSphere version:

   - For versions prior to vSphere 6.5, set the `SchedNumReqOutstanding` parameter to its maximum value (256): `esxcli storage core device set -d naa.xxx -O 256`
   - In vSphere 6.5, VMware limits the `SchedNumReqOutstanding` parameter to be equal to or lower than the LUN queue depth. Set the following values according to the HBA:
     - Emulex - 128
     - Cisco UCS FNIC - 128
     - Qlogic - 256
     - Software iSCSI initiator - 128
     - Hardware iSCSI initiator - 256

vCenter Server parameter settings

About this task

The maximum number of concurrent full cloning operations should be adjusted based on the XtremIO cluster size. The vCenter Server parameter `config.vpxd.ResourceManager.maxCostPerHost` determines the maximum number of concurrent full clone operations allowed (the default value is 8). Adjusting the parameter should be based on the XtremIO cluster size as follows:

- 10 TB Starter X-Brick (5 TB) and a single X-Brick - 8 concurrent full clone operations
- Two X-Brick blocks - 16 concurrent full clone operations
- Four X-Brick blocks - 32 concurrent full clone operations
- Six X-Brick blocks - 48 concurrent full clone operations

To adjust the maximum number of concurrent full cloning operations:

Steps

1. Launch vSphere WebUI client to log in to the vCenter Server.
2. From the top menu, select **vCenter Inventory List** .
3. From the left menu, under **Resources**, Click **vCenter Servers**.
4. Select **vCenter** > **Manage Tab** > **Settings** > **Advanced Settings**.
5. Click **Edit.**
6. Locate the `config.vpxd.ResourceManager.maxCostPerHost` parameter and set it according to the XtremIO cluster size.
   If you cannot find the parameter, type its name in the *Key* field and the corresponding value in the *Value* field.
7. Click **Add**.
8. Click **OK** to apply the changes.

vStorage API for Array Integration (VAAI) Settings

VAAI is a vSphere API that offloads vSphere operations such as virtual machine provisioning, storage cloning, and space reclamation to storage arrays that supports VAAI. XtremIO Storage Array fully supports VAAI.

To ensure optimal performance of XtremIO storage from vSphere, VAAI must be enabled on the ESXi host before using XtremIO storage from vSphere. Failing to do so may expose the XtremIO cluster to the risk of datastores becoming inaccessible to the host.

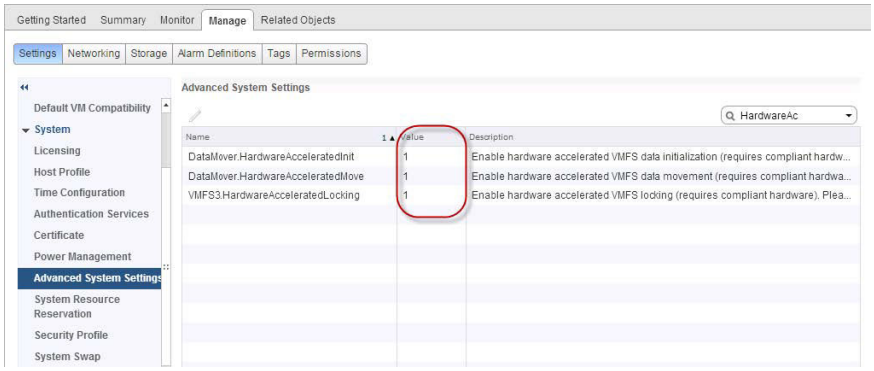
Enabling VAAI features

Confirming that VAAI is enabled on the ESXi host

In vSphere version 5.x or later versions, VAAI is enabled by default. Before using the XtremIO storage, confirm that VAAI features are enabled on the ESXi host. Perform the following procedures to confirm that VAAI is enabled on the ESXi host:

Steps

1. Launch the vSphere Web Client and go to **Home** > **Hosts and Clusters**.
2. In the left menu section, locate the **ESXi host**, and click it.
3. In the right pane, click **Manage** > **Settings**.
4. From the **System** section, click **Advanced System Settings**.
5. Verify that the following parameters are enabled (set to 1):
   - `DataMover.HardwareAcceleratedMove`
   - `DataMover.HardwareAcceleratedInit`
   - `VMFS3.HardwareAcceleratedLocking`



If any of the parameters that are mentioned earlier are not enabled, adjust them by clicking the **Edit** icon and click **OK** .

VMware ATS Heartbeat

Dell EMC recommends keeping VMware ATS Heartbeat setting at its default setting (enabled). This recommendation also applies during an XtremIO software upgrade (NDU).

NOTE: In cases where datastore connectivity or I/O responsiveness is impacted and an `ATS Miscompare detected between test and set HB images` log message is displayed in the vmkernel.log file, see the VMware Knowledge Base article 2113956 for guidance on how to proceed.

Manually setting VAAI on Datastore

About this task

NOTE: If VAAI setting is enabled after a datastore was created on XtremIO storage, the setting does not automatically propagate to the corresponding XtremIO Volumes. The setting must be manually configured to avoid data unavailability to the datastore.

Perform the following procedure on all datastores created on XtremIO storage before VAAI is enabled on the ESXi host. To manually set VAAI setting on a VMFS-5 datastore that is created on XtremIO storage with VAAI disabled on the host:

Steps

1. Confirm that the VAAI Hardware Accelerator Locking is enabled on this host. See Confirming that VAAI is enabled on the ESXi Host for details.
2. Using the following vmkfstools command, confirm that the datastore is configured as *public ATS-only*:

   ```
   # vmkfstools -Ph -v1 <path to datastore> | grep public
   ```

   - In the following example, a datastore volume is configured as *public*:

     ```
     # vmkfstools -Ph -v1 /vmfs/volumes/datastore1 | grep public
     Mode: public
     ```

   - In the following example, a datastore volume is configured as *public ATS-only*:

     ```
     # vmkfstools -Ph -v1 /vmfs/volumes/datastore2 | grep public
     Mode: public ATS-only
     ```

3. If the datastore was found with mode *public*, change it to *public ATS-only* by executing the following steps:
   a. Unmount the datastore from all ESXi hosts on which it is mounted (except one ESXi host).
   b. Access the ESXi host on which the datastore is still mounted.
   c. Run the following *vmkfstools* command to enable ATS on the datastore:

      ```
      # vmkfstools --configATSOnly 1 <path to datastore>
      ```

   d. Click **0** to continue with ATS capability.
   e. Repeat step 2 to confirm that ATS is set on the datastore.
   f. Unmount datastore from the last ESXi host.
   g. Mount datastore on all ESXi hosts.

Tuning VAAI XCOPY with XtremIO

The VAAI XCOPY chunk size is set, using the `MaxHWTransferSize` parameter. By default, vSphere instructs the storage array to copy data in 4 MB chunks. To optimize VAAI XCOPY operation with XtremIO, Dell EMC recommends adjusting the chunk size to 256 KB.

To adjust the VAAI XCOPY chunk size to 256 KB, run the following CLI commands according to the vSphere version running on your ESXi host:

- For vSphere versions earlier than 5.5:

  ```
  esxcli system settings advanced list -o /DataMover/MaxHWTransferSize
  esxcli system settings advanced set --int-value 0256 --option /DataMover/MaxHWTransferSize
  ```

- For vSphere version 5.5 and later:

  ```
  esxcli system settings advanced set --int-value 0256 --option /DataMover/MaxHWTransferSize
  ```

Disabling VAAI in ESXi

About this task

In some cases (mainly for testing purposes), it is necessary to temporarily disable VAAI. As a rule, you should enable VAAI on an ESXi host that is connected to XtremIO. Avoid disabling VAAI if possible. If not, disable it temporarily

NOTE: For further information about disabling VAAI, see the VMware Knowledge Base article 1033665.

NOTE: Disabling the Atomic Test and Set (ATS) parameter can cause data unavailability in ESXi 5.5 for volumes that are created natively as VMFS5 datastore.

To disable VAAI on the ESXi host:

Steps

1. Browse to the host in the vSphere Web Client navigator.
2. Select the **Manage** tab and click **Settings**.
3. In the **System** section, click **Advanced System Settings**.
4. Click **Edit** and modify the following parameters to set to 0 to disable VAAI:
   - `DataMover.HardwareAcceleratedMove`
   - `DataMover.HardwareAcceleratedInit`
   - `VMFS3.HardwareAcceleratedLocking`
5. Click **OK**.

Configuring VMware vSphere with XtremIO Storage in a Multiple Storage Array Configuration

The table in this section shows the recommended vSphere settings when multiple storage arrays are connected to vSphere and to XtremIO.

For reference, this table also includes the corresponding recommendations for settings when vSphere is connected to XtremIO storage only.

Parameter name	Scope/Granularity	MultiStorage setting	XtremIO only setting
FC Adapter Policy IO Throttle Count	Per vHBA	256	1,024
fnic_max_qdepth	Global	32	128
Disk.SchedNumReqOutstanding	LUN[a]	32	256[b]
Disk.SchedQuantum	Global	8	64
Disk.DiskMaxIOSize	Global	4 MB	4 MB

Parameter name	Scope/Granularity	MultiStorage setting	XtremIO only setting
XCOPY (/DataMover/ MaxHWTransferSize)	Global	4 MB	256 KB
config.vpxd.ResourceManager. maxCostPerHost	vCenter	8	8 per X-Brick (up to 48)

a. If you are using vSphere 5.5 or later version, the parameter can be set on a specific volume, as noted here. With earlier vShpere versions, this is an global setting.ESXi
b. If you are using vSphere 6.5, this parameter should be set to a different value. See Adjusting the ESXi Host Parameters for XtremIO Storage for more details.

- FC Adapter Policy IO Throttle Count - The total number of I/O requests that can be outstanding on a per-virtual host bus adapter (vHBA) in UCS. For details, see Setting the HBA I/O Throttle of the Cisco UCS fNIC HBA.
- fnic_max_qdepth - A UCS FC HBA driver level setting that manages the total number of I/O requests that can be outstanding on a per-LUN basis. For details, see the steps relevant to the Cisco UCS fNIC HBA in Setting the LUN Queue Depth on a Host Running vSphere 5.x or Above.
- Disk.SchedNumReqOutstanding - The total number of outstanding commands that are permitted from all virtual machines collectively on the host to a LUN. For details, see Adjusting the ESXi Host Parameters for XtremIO Storage.
- Disk.SchedQuantum - The maximum number of consecutive "sequential" I/Os that are allowed from one VM before forcing a switch to another VM. For details, see Adjusting the ESXi Host Parameters for XtremIO Storage.
- Disk.DiskMaxIOSize - The maximum I/O size that ESXi allows before splitting I/O requests. For details, see Adjusting the ESXi Host Parameters for XtremIO Storage.
- XCOPY (/DataMover/MaxHWTransferSize) - The maximum number of blocks used for XCOPY operations. For details, see Tuning VAAI XCOPY with XtremIO.
- config.vpxd.ResourceManager.maxCostPerHost - The maximum number of concurrent full clone operations allowed (the default value is 8). For details, see vCenter Server Parameter Settings.

Multipathing Software Configuration

NOTE: You can use Dell EMC Virtual Storage Integrator (VSI) Path Management to configure path management across Dell EMC platforms, including XtremIO. For information about using this vSphere Client plug-in, see the *Dell EMC VSI Path Management Product Guide*.

Configuring vSphere Native Multipathing

XtremIO supports the VMware vSphere Native Multipathing (NMP) technology. This section describes the procedure that is required for configuring native vSphere multipathing for XtremIO volumes.

For best performance, Dell EMC recommends doing the following:

- Set the native Round Robin path selection policy on XtremIO volumes that are presented to the ESXi host.

  NOTE: With NMP in vSphere versions below 5.5, clustering is not supported when the path policy is set to Round Robin. For details, see vSphere MSCS Setup Limitations in the Setup for Failover Clustering and Microsoft Cluster Service guide for ESXi 5.0 or ESXi 4.x. In vSphere 5.5, Round Robin PSP (PSP_RR) support is introduced. For details, see the MSCS support enhancements in vSphere 5.5 in the VMware Knowledge Base article 2052238.
- Set the vSphere NMP Round Robin path switching frequency to XtremIO volumes from the default value (1000 I/O packets) to 1.

These settings ensure optimal distribution and availability of load between I/O paths to the XtremIO storage. Starting from ESXi 5.5 P08 (build number 4179633), ESXi 6.0 P03 (build number 4192238) and ESXi 6.5 (build number 4564106), the NMP SATP rule for a newly provisioned XtremIO LUN is *VMW_SATP_DEFAULT_AA* by default, and the path policy for XtremIO LUNs is *VMP_PSP_RR, iops=1*.

With NMP in vSphere 6.7 (or later), a new *latency* suboption is introduced to enable VMW_PSP_RR path congestion awareness. With this new *latency* suboption, the PSP actively monitors all paths and considers the path latency and pending I/Os on each active path.

It is still recommended to use the Round Robin with IOPS=1 with native multipathing for all vSphere versions.

For vSphere version 6.7 or later, you can choose between the recommended option and the *latency* suboption. No differences were recorded using either of these two alternatives with XtremIO.

To configure this setting on an XtremIO Volume that is presented to the ESXi host with the *latency* suboption, run the following ESXi CLI commands:

```
esxcfg-advcfg -s 1 /Misc/EnablePSPLatencyPolicy
```

```
esxcli storage nmp psp roundrobin deviceconfig set -d
<Device_ID> -type=latency
```

If the installed ESXi version is below ESXi 5.5 P08 or ESXi 6.0 P03, and XtremIO LUN is already presented to the host, Dell EMC recommends performing one of the following actions:

- Upgrade the host to ESXi 5.5 P08, ESXi 6.0 P03, or ESXi 6.5 and then reboot the host.
- Change the NMP Round-Robin configuration. See Configure vSphere NMP Round Robin on an XtremIO volume already presented to the ESXi host, using ESXi command line.

To set vSphere NMP Round-Robin configuration for XtremIO volumes that are not yet presented to the host, Dell EMC recommends configuring vSphere NMP Round Robin as the default pathing policy, using the ESXi command line. See Configure vSphere NMP Round Robin as the default pathing policy for all XtremIO volumes, using the ESXi command line for the detailed procedure.

For XtremIO volumes already presented to the host, Dell EMC recommends configuring vSphere NMP Round Robin on an XtremIO volume in an ESXi host, using ESXi command line (per volume, for each host where the volume is presented). See Configure vSphere NMP Round Robin on an XtremIO volume already presented to the ESXi host, using ESXi command line for the detailed procedure.

Configure vSphere NMP Round Robin as the default pathing policy for all XtremIO volumes, using the ESXi command line

About this task

NOTE: Use this method when no XtremIO volume is presented to the host. XtremIO volumes already presented to the host are not affected by this procedure (unless they are unmapped from the host).

Steps

1. Open an SSH session to the host as *root*.
2. Run the following command to configure the default pathing policy for newly defined XtremIO volumes to Round Robin with path switching after each I/O packet:

   ```
   esxicli storage nmp satp rule add -c tpgs_off -e "XtremIO
   Active/Active" -M XtremApp -P VMW_PSP_RR -O iops=1 -s
   VMW_SATP_DEFAULT_AA -t vendor -V XtremIO
   ```

   This command also sets the vSphere NMP Round Robin path switching frequency for newly defined XtremIO volumes to one (1).

   NOTE: Using this method does not impact any non-XtremIO volume that is presented to the ESXi host.

Configure vSphere NMP Round Robin on an XtremIO volume already presented to the ESXi host, using ESXi command line

About this task

NOTE: Use this method only for XtremIO volumes that are already presented to the host. For volumes not yet presented to the host, use the method that is described in Configure vSphere NMP Round Robin as the default pathing policy for all XtremIO volumes, using the ESXi command line.

Steps

1. Open an SSH session to the host as *root*.
2. Run the following command to obtain the NAA of XtremIO LUNs presented to the ESXi host: `#esxcli storage nmp path list | grep XtremIO -B1`.
3. Run the following command to modify the path selection policy on the XtremIO volume to Round Robin: `#esxcli storage nmp device set --device <naa_id> --psp VMW_PSP_RR`.

Example:

```
#esxcli storage nmp device set --device naa.514f0c5e3ca0000e
--psp VMW_PSP_RR
```

4. Run the following command to set the vSphere NMP Round Robin path switching frequency on XtremIO volumes from the default value (1000 I/O packets) to 1: `#esxcli storage nmp psp roundrobin deviceconfig set --device="<naa_id>" --iops=1 --type=iops`

   Example:

```
#esxcli storage nmp psp roundrobin deviceconfig set
--device="naa.514f0c5e3ca0000e" --iops=1 --type=iops
```

NOTE: Using this method does not impact any non-XtremIO volumes that are presented to the ESXi host.

For more details, see the following:

- VMware Knowledge Base article 1017760
- VMware Knowledge Base article 2069356

PowerPath Multipathing with XtremIO

XtremIO supports multipathing using Dell EMC PowerPath/VE on VMware vSphere. PowerPath/VE versions 5.9 SP1 and later provide Loadable Array Module (LAM) for XtremIO Array devices. With this support, XtremIO devices running versions 2.2 and later are managed under the XtremIO class.

NOTE: For the most updated information about PowerPath support with XtremIO storage, see the XtremIO Simple Support Matrix on Dell EMC E-Lab Navigator.

PowerPath/VE provides enhanced path management capabilities for up to 32 paths per logical device and intelligent dynamic I/O load-balancing functionalities that are transparent to VMware vSphere, and to Guest operating systems. Having multiple paths enables the host to access a storage device even if a specific path is unavailable. Multiple paths share the I/O traffic to a storage device, using intelligent load-balancing policies which enhance I/O performance and increase application availability. Dell EMC PowerPath is the recommended multipathing choice.

PowerPath/VE features include:

- Multiple paths - enables higher availability and I/O performance.
- Path management insight capabilities - PowerPath characterizes I/O patterns and aids in diagnosing I/O problems due to flaky paths or unexpected latency values. The following capabilities are provided:
  - Read and write - in MB/seconds per LUN
  - Latency distribution - the high and low watermarks per path
  - Retries - the number of failed I/Os on a specific path
  - Virtual machine performance monitoring support - available from Dell EMC PowerPath 6.0 SP1 and later
  - Support performance monitoring - for a specific VM or for all VMs in the environment
- Autostandby - automatically detects intermittent I/O failures and places paths in autostandby (also known as flaky paths).
- Remote monitoring and management:
  - PowerPath Management Appliance (PPMA)
  - Remote PowerPath CLI (rpowermt)
  - VSI for VMware vSphere Path Management
  - VMware vCenter Update Manager
  - VMware Auto Deploy

Additional PowerPath related information:

- For detail on the PowerPath/VE releases supported for your VMware vSphere host, see the XtremIO Simple Support Matrix on Dell EMC E-Lab Navigator.
- For details on class support with XtremIO for your host, see the *Dell EMC PowerPath/VE release notes* for the PowerPath/VE version you are installing.
- For details on installing and configuring PowerPath/VE with XtremIO class on your host, see the *Dell EMC PowerPath on VMware vSphere Installation and Administration Guide* for the PowerPath/VE version you are installing. This guide provides the required information for placing XtremIO volumes under PowerPath/VE control.

NOTE: The PowerPath with XtremIO class support installation procedure is fully storage-aware. All required PowerPath settings with XtremIO storage are automatically done when PowerPath is installed on your host. This includes settings such as the PowerPath multipathing policy that does not require manual setting.

Post configuration steps - Using the XtremIO storage

When host configuration is completed, you can use the XtremIO storage from the host. For details on creating, presenting, and managing volumes that can be accessed from the host through GUI or CLI, see the *XtremIO Storage Array User Guide* that matches the version running on your XtremIO cluster.

Dell EMC Virtual Storage Integrator (VSI) Unified Storage Management version 6.2 and later can be used to provision from within vSphere Client Virtual Machine File System (VMFS) datastores and Raw Device Mapping volumes on XtremIO. Furthermore, Dell EMC VSI Storage Viewer version 6.2 and later extends the vSphere Client to facilitate the discovery and identification of XtremIO storage devices that are allocated to VMware ESXi hosts and virtual machines.

For more information about using these two vSphere Client plug-ins, see the *VSI Unified Storage Management product guide* and the *VSI Storage Viewer product guide*.

Configuring the XtremIO Cluster Connectivity Mode with ESXi Hosts

The Connectivity mode in XtremIO version XIOS 4.0.10-33 or later allows you to preset the response of the XtremIO cluster when a volume becomes unavailable to the ESXi server.

The Connectivity mode setting defines the action to be taken on the XtremIO cluster when storage connectivity to the ESXi host is lost. There are two possible modes of action:

- All Paths Down (APD) - Refers to a condition where all paths to the storage device are unexpectedly lost and the cluster does not respond to any incoming I/O requests. In this case, the VMkernel core storage stack assumes that the device will reappear (the loss is considered to be temporary).
- Permanent Device Loss (PDL) - Refers to a condition where all paths to the storage device are unexpectedly lost and the cluster responds to all incoming I/O requests with a SCSI error (resembling the behavior when deleting or unmapping a volume). In this case, the VMkernel core storage stack assumes that the device will never reappear (the loss is considered as permanent).

The following table summarizes the pros and cons of the two modes of action:

Mode of Action	Pros	Cons
All Paths Down (APD)	Fast recovery from a temporary connectivity disruption or a high-availability event (for example: cluster software upgrade).	I/O queues may overload and freeze the ESXi host in an I/O intensive environment.
Permanent Device Loss (PDL)	Fast recovery from a permanent device removal such as deletion or unmapping of a volume.	Complex recovery if there is a temporary disconnection that may require an ESXi host reboot.

When using vSphere with XtremIO, Dell EMC recommends setting the Connectivity mode to APD, as it is appropriate for most installations.

NOTE: With XIOS 4.0.25-22 or later, the default value for the XtremIO Connectivity mode with new cluster installations is changed from PDL to APD. With XIOS 4.0.15-24 or earlier, PDL is the default value for the XtremIO Connectivity mode.

The Connectivity mode setting is only applicable to ESXi hosts with:

- vSphere versions 5.x or 6.x
- Initiators set with an ESXi Operating System

For details about how to adjust this cluster setting, see the *XtremIO Storage Array User Guide*.

NOTE: When performing a cluster software upgrade (NDU) to XIOS version 4.0.10-33 or later, Dell EMC recommends adjusting the Connectivity mode to APD. Include this as part of the Service Request that is submitted for the NDU and provide Dell EMC KB 483391 to expedite the processing and handling.

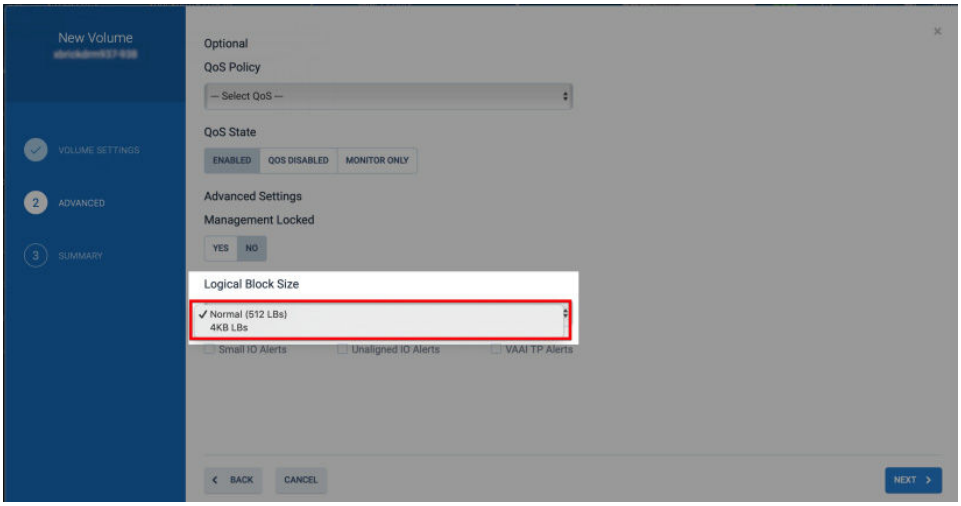
NOTE: If, prior to the NDU, the cluster is running a XIOS version earlier than 4.0.10-33, perform the NDU. After completing the NDU, Dell EMC recommends adjusting the Connectivity mode of the cluster to APD.

Disk formatting

When creating volumes in XtremIO for a vSphere host, consider the following:

- Disk logical block size - The only logical block (LB) size that is supported by vSphere for presenting to ESXi volumes is 512 bytes.

  The following figure demonstrates formatting an XtremIO Volume using the WebUI:



  NOTE: For details on formatting a newly created Volume (using either the WebUI or the GUI interfaces), see the *XtremIO Storage Array User Guide* that corresponds to the version running on your XtremIO cluster.

- Disk alignment - Unaligned disk partitions may substantially impact I/O to the disk. With vSphere, datastores and virtual drives are aligned by default as they are created. No further action is required to align these in ESXi.

  With virtual machine disk partitions within the virtual drive, alignment is determined by the guest operating system. For virtual machines that are not aligned, consider using tools such as UBERalign to realign the disk partitions as required.

Presenting XtremIO Volumes to the ESXi Host

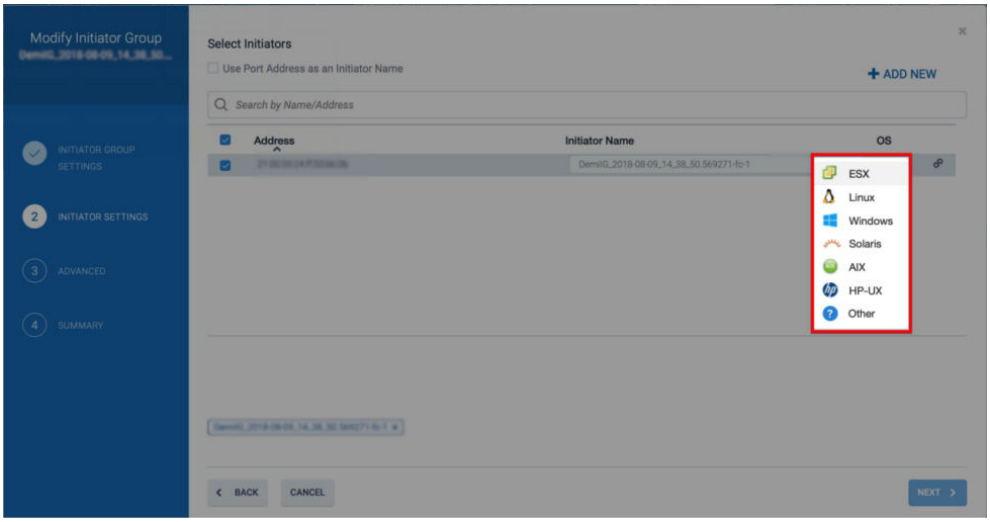
NOTE: Dell EMC does not recommend using host-level software for deduplication, compression, and encryption of data stored in XtremIO. These operations are performed natively on the XtremIO cluster, thus releasing host computing resources. Furthermore, using these operations may severely impact data efficiency on the XtremIO cluster, as they can lead to a higher physical or logical capacity utilization.

NOTE: The information in this section applies only to XtremIO version 4.0 and later.

NOTE: When using iSCSI software initiator with ESXi and XtremIO storage, Dell EMC recommends using only lower case characters in the IQN to correctly present the XtremIO volumes to ESXi. For more details, see the VMware Knowledge Base article 2017582.

When adding Initiator Groups and Initiators to allow ESXi hosts to access XtremIO volumes, specify ESXi as the operating system for newly created Initiators.

The following figure demonstrates setting the Operating System field for an Initiator using the WebUI:



NOTE: Setting the Initiator's Operating System is required for optimal interoperability and stability of the host with XtremIO storage. You can adjust the setting while the host is online and connected to the XtremIO cluster with no I/O impact.

NOTE: See the *XtremIO Storage Array User Guide* that corresponds to the version running on your XtremIO cluster.

Following a cluster upgrade from XtremIO version 3.0.x to version 4.0 or later, modify the operating system for each initiator that is connected to an ESXi host.

Creating a File System

NOTE: File system configuration and management are out of the scope of this document.

Dell EMC recommends creating the file system using its default block size. Using a non-default block size may lead to unexpected behavior. For details, see your operating system and file system documentation.

Using LUN 0 with XtremIO Storage

This section details the considerations for using LUN 0 with vSphere.

- In XtremIO version 4.0.0 or later, volumes are numbered by default starting from LUN ID 1.
- If LUN ID 0 is required for an XtremIO volume (For example, boot from SAN, PowerPath/VE), manually adjust the LUN ID to 0. Restart the ESXi host if a rescan fails to locate this volume.
- When a cluster is updated from XtremIO version 3.0.x to 4.0.x, an XtremIO volume with a LUN ID 0 remains accessible after the upgrade.
- With XtremIO version 4.0.0 or later, no further action is required if volumes are numbered starting from LUN ID 1.

VMware Paravirtual SCSI Controllers

For optimal resource utilization of virtual machines with XtremIO, Dell EMC recommends configuring virtual machines with Paravirtualized SCSI controllers. VMware Paravirtual SCSI controllers are high-performance storage controllers that can provide higher throughput and lower CPU usage. These controllers are best suited for high-performance storage environments.

NOTE: Virtual machines with paravirtualized SCSI controllers destined to be part of an MSCS cluster are supported with vSphere 5.5 update 3 or later and vSphere 6.0 or later.

For more details about configuring virtual machines with Paravirtualized SCSI controllers, see the *vSphere Virtual Machine Administration Guide* in vSphere documentation.

Virtual Machine formatting

About this task

NOTE: Dell EMC recommends not to use host-level software for deduplication, compression, and encryption of data stored on the XtremIO cluster. It is also not recommended to use the Virtual Machine Encryption feature with vSphere 6.5 and XtremIO. Encryption is performed natively on the XtremIO cluster, thus releasing host computing resources.

For optimal space utilization with vSphere 6.x, Dell EMC recommends formatting virtual machines on XtremIO storage, using Thin Provision. Using this format, in-guest space reclamation is available, provided the following requirements are fulfilled:

- Thin virtual drives
- VM hardware version 11
- ESXi 6.x
- EnableBlockDelete set to 1

- Guest Operating System support of UNMAP

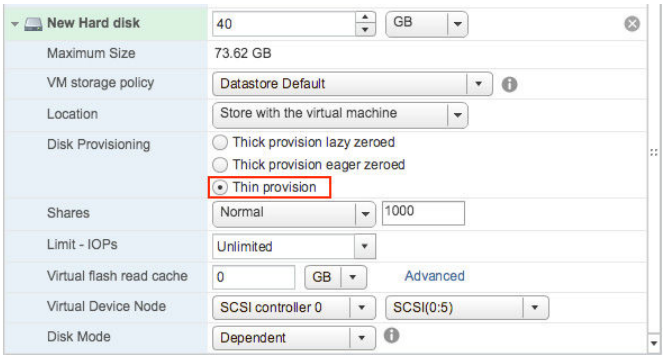
NOTE: Some guest Operating Systems that support unmapping of blocks, such as Linux-based systems, do not generate UNMAP commands on virtual drives in vSphere 6.0. This occurs because the level of SCSI support for ESXi 6.0 virtual drives is SCSI-2, while Linux expects 5 or higher for SPC-4 standard. This limitation prevents the generation of UNMAP commands until the virtual drives can claim support for at least SPC-4 SCSI commands.

For details on virtual volumes and UNMAP, see the VMware Knowledge Base.

To format a virtual machine using the Thin Provision option:

Steps

1. From vSphere Web Client, launch the *Create New Virtual Machine* wizard.
2. Proceed, using the wizard, up to the **2f Customize Hardware** screen.
3. In the **Customize Hardware** screen, click **Virtual Hardware**.
4. Toggle to the **New Hard Disk** option.
5. Select the **Thin Provision** option to format the virtual machine's virtual drive.



For details about migrating a virtual machine from Thick provision to Thin provision, see the VMware Knowledge Base article 2014832.

Virtual Machine formatting with Thick-Provision VMDK

When optimal performance is required or when vSphere version is lower than 6.x, format virtual machines on XtremIO storage, using Thick Provision Eager Zeroed. Using this format, the required space for the virtual machine is allocated and zeroed on creation time. However, with native XtremIO data reduction, thin provisioning, and VAAI support, no actual physical capacity allocation occurs.

About this task

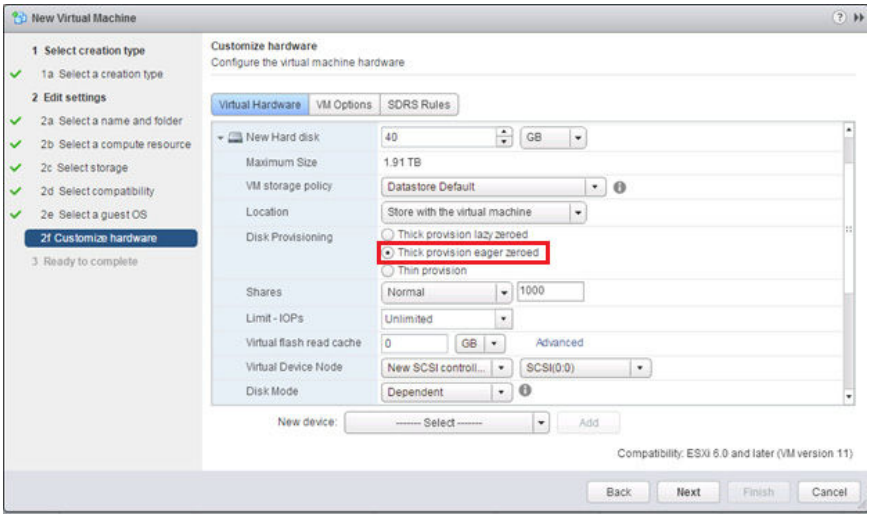
Thick Provision Eager Zeroed format advantages are:

- Logical space is allocated and zeroed on virtual machine provisioning time, rather than scattered, with each I/O sent by the virtual machine to the disk (when Thick Provision Lazy Zeroed format is used).
- Thin provisioning is managed in the XtremIO Storage Array rather than in the ESXi host (when Thin Provision format is used).

To format a virtual machine using Thick Provision Eager Zeroed:

Steps

1. From vSphere Web Client, launch the *Create New Virtual Machine* wizard.
2. Proceed using the wizard up to the **2f Customize Hardware** screen.
3. In the **Customize Hardware** screen, click **Virtual Hardware**.
4. Toggle to the **New Hard Disk** option.
5. Select the **Thick Provision Eager Zeroed** to format the virtual machine's virtual drive.

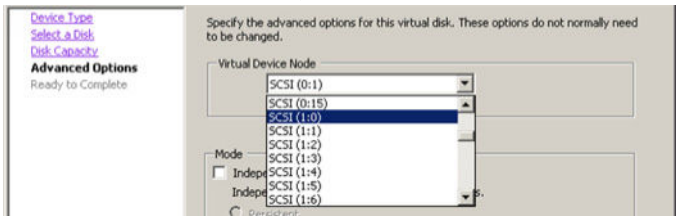


6. Proceed using the wizard to complete creating the virtual machine.

Virtual machine guest operating system settings

This section details the recommended settings and considerations for virtual machines guest operating system.

- LUN Queue Depth - For optimal virtual machine operation, configure the virtual machine guest operating system to use the maximum queue depth of the virtual SCSI controller. For details on adjusting the guest operating system LUN queue depth, see the VMware Knowledge Base article 2053145.
- RDM volumes in Guest operating system - Span RDM volumes that are used by the virtual machine, across SCSI controllers to prevent a bottleneck on a single SCSI controller.



- RDM volumes in guest operating system used for Microsoft Cluster (MSCS) - ESXi hosts with visibility to RDM volumes that are used by Microsoft Cluster (MSCS), may take a long time to start or to perform LUN rescan. For the required settings on the RDM volumes, see the VMware Knowledge Base article 1016106.

Space Reclamation

This section provides a comprehensive list of capacity management steps for achieving optimal capacity utilization on the XtremIO array, when connected to an ESXi host.

Data space reclamation helps to achieve optimal XtremIO capacity utilization. Space reclamation is a vSphere function that enables the reclamation of used space. This is done by sending zeros to a specific address of the volume, after the file system notifies that the address space was deleted.

Unlike traditional operating systems, ESXi is a hypervisor, running guest operating systems on its file-system (VMFS). As a result, space reclamation is divided into guest operating system and ESXi levels.

ESXi level space reclamation should be run only when deleting multiple VMs, and space is reclaimed from the ESXi datastore. Guest level space reclamation should be run as a periodic maintenance procedure to achieve optimal capacity savings.

The following figure displays a scenario in which VM2 is deleted while VM1 and VM3 remain:

Space Reclamation at Guest Level

On VSI environments, every virtual server should be treated as a unique object. When using VMDK devices, T10 trim commands are blocked. Therefore, you must run space reclamation manually. RDM devices pass through T10 trim commands.

There are two types of VDI provisioning that differ by their space reclamation guidelines:

- Temporary desktop (Linked Clones) - Temporary desktops are deleted once the users log off. Therefore, running space reclamation on the guest operating system is not relevant, and only ESXi level space reclamation should be used.
- Persistent desktop (Full Clones) - Persistent desktop contains long-term user data. Therefore, space reclamation should be run on guest operating system level first, and only then on ESXi level.

On large-scale VSI/VDI environments, divide the VMs into groups to avoid overloading the SAN fabric.

Space Reclamation at ESXi Level

- ESXi 5.1 and below

  In versions before ESXi 5.5, the `vmkfstools` command is used for space-reclamation. This command supports datastores up to 2 TB.

  The following example describes running `vmkfstool` on a datastore XtremIO_DS_1 with 1% free space to enable user writes:

  ```
  # cd /vmfs/volumes/XtremIO_DS_1
  # vmkfstools -y 99
  ```

  Vmfs reclamation may fail due to T10 commands blocking (VPLEX). In such cases, you must apply a manual copy of zeroes to the relevant free space.

  The following example describes running a manual script on X41-VMFS-3 datastore (see ESXi Space Reclamation Script):

  ```
  # ./reclaim_space.sh X41-VMFS-3
  ```

  NOTE: The datastore name cannot include spaces.

- ESXi 5.5 and 6.0

  ESXi 5.5 introduces a new command for space reclamation and supports datastores larger than 2 TB.

  The following example describes running space reclamation on a datastore XtremIO_DS_1:

  ```
  # esxicli storage vmfs unmap --volume-label=XtremIO_DS_1
  --reclaim-unit=20000
  ```

  The *reclaim-unit* argument is an optional argument, indicating the number of vmfs blocks to UNMAP per iteration.

  NOTE: For details on the reclaim-unit argument setting for specific environments, see the VMware Knowledge Base article 2057513.

  Vmfs reclamation may fail due to T10 commands blocking (VPLEX). In such cases, you must apply a manual copy of zeroes to the relevant free space.

  The following example describes running a manual script on X41-VMFS-3 datastore (see ESXi Space Reclamation Script):

  ```
  # ./reclaim_space.sh X41-VMFS-3
  ```

  NOTE: The datastore name cannot include spaces.

- ESXi 6.5

  Starting from ESXi version 6.5, automatic space reclamation at ESXi level is supported and enabled on VMFS6 datastores by default.

Checking the status of automatic space reclamation settings is done at the datastore level as follows:

```
# esxicli storage vmfs reclaim config get -l=VMFS_label|-u=VMFS_uuid
# esxicli storage vmfs reclaim config get -l my_datastore
Reclaim Granularity: 1048576 Bytes
Reclaim Priority: low
```

Space reclamation priority defines how blocks, which are deleted from a VMFS6 datastore, are reclaimed on a LUN backing the datastore. The LUN performs the space reclamation operation at a low rate by default.

Using `esxcli`, you can also set the priority to **`medium`** or **`high`**.

```
# esxcli storage vmfs reclaim config set -l my_datastore -p high
```

For more information, see the documentation at the VMware website.

ESXi Space Reclamation Script

The following example describes an ESXi space reclamation script usage and the ESXi space reclamation script:

```
# python linux-reclaim.py --help
usage: esx-reclamation.py [-c <cluster_name>]
optional arguments:
-h, --help show this help message and exit
-v, --version show program's version number and exit
-t, --thread_number Threads Amount
-m, --mount_point Mount Point Name
# python esx-reclamation.py --thread_number 4 --mount_point
XBR143-VMFS-01
#!/usr/bin/env python
from__future__import print_function
import time
import threading
import subprocess
import argparse
import re
class myThread (threading.Thread):
def __init__(self, name, count, running):
threading.Thread.__init__(self)
self.name = name
self.count = count
self.running = running
def run(self):
dd_something(self.name, self.count, self.running)
def dd_something(name, count, running):
execute('/bin/dd count={1} bs=131072 if=/dev/zero
of={0}/xtremio_file-{2}.zf conv=fxync'.format(name, count, running))
def execute(command):
return subprocess.Popen(command, shell=True,
stdout=subprocess.PIPE,
stderr=subprocess.PIPE).communicate()[0].splitlines()
def get_mount_points():
return execute('df -m')
threadLock = threading.Lock()
def main():
parser = argparse.ArgumentParser(usage='%(prog)s [-c
<cluster_name>]')
parser.add_argument('-v', '--version', action='version',
version='%(prog)s 1.02')
parser.add_argument('-t', '--thread_number', action='store',
dest='thread_number', default='1', required=False, help='Threads
Amount', metavar='')
parser.add_argument('-m', '--mount_point', action='store',
dest='mount_point', default=None, required=True, help='Mount Point
Name', metavar='')
args = parser.parse_args()
thread_number = int(args.thread_number)
mount_point = args.mount_point
print('Starting Main Thread {0}'.format(time.ctime(time.time())))
threads = []
for entry in get_mount_points():
if mount_point in entry:
```

```
filesystem, mblocks, used, available, usep, mounted =
map(str.strip, re.split(" +", entry))
for i in xrange(thread_number):
i = myThread(mounted, int(int(available) * 0.95 /
thread_number) *8, str(i))
i.start()
threads.append(i)
for t in threads:
t.join()
execute('/bin/rm -rf {0}/xtremio_file-*.zf'.format(mounted))
print('Exiting Main Thread {0}'.format(time.ctime(time.time())))
if __name__ == "__main__":
exit(main())
```

NOTE: Increasing percentage leads to elevated precision, it may increase the probability of receiving a *no free space* SCSI error during the reclamation.

Creating Copies of VMFS Datastores Using XtremIO snapshots

Snapshots are instantaneous copy images of volume data. A data snapshot captures the state of the data exactly as it is displayed at the specific point in time on which the snapshot was created. This enables users to save the Volume data state and then access the specific volume data as required, even after the source volume was changed.

You can create the following snapshot types:

- Protection - A read-only Snapshot that can be used to restore data
- Copy - A writable Snapshot that can be refreshed with new data

For more details on using XtremIO snapshots, see the *XtremIO Storage Array User Guide* for the version the cluster is running.

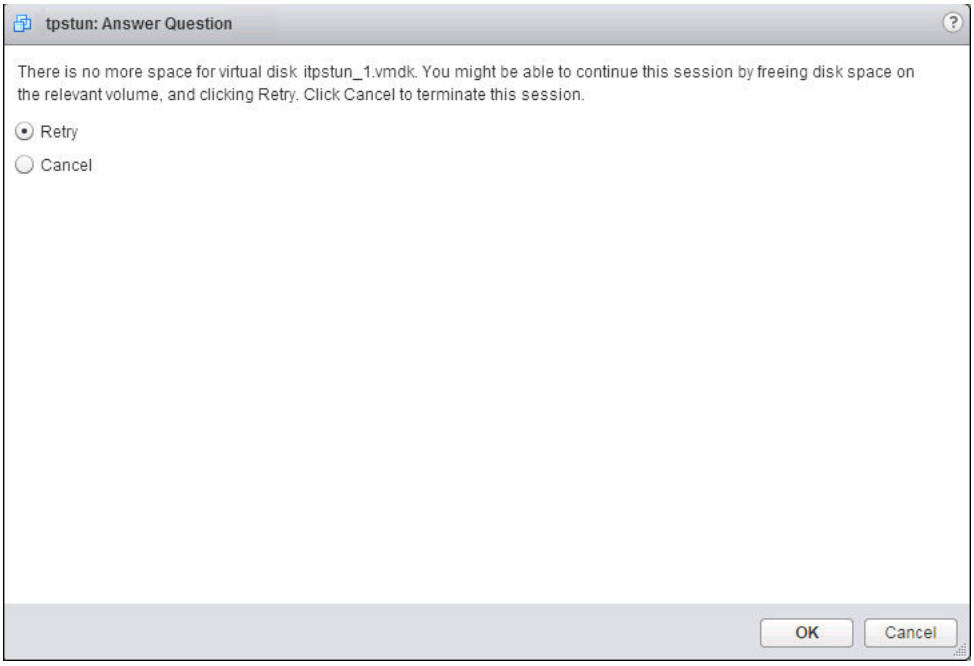
NOTE:

When using XtremIO snapshots to create multiple copies of a VMFS datastore, simultaneous resignaturing of multiple copies of a VMFS datastore fail with an error. For further details on this vSphere limitation, see the VMware Knowledge Base article 1028720.

Out of Space VM Suspend and Notification with Thin Provisioning (TPSTUN)

TPSTUN is a VAAI primitive that enables the array to notify vSphere when a LUN is running out of space due to thin provisioning overcommit. The command causes all virtual machines on that LUN to be suspended. XtremIO supports this VAAI primitive.

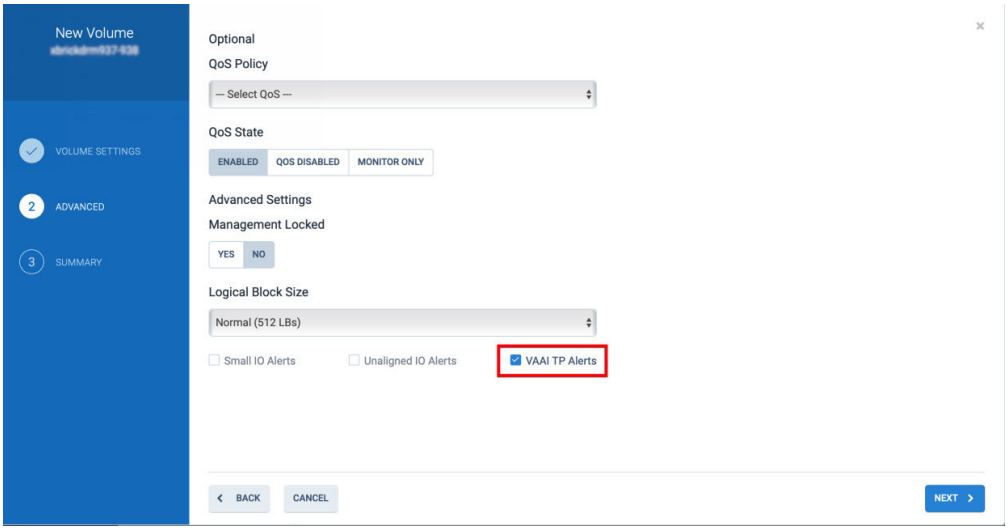
When a virtual machine that is provisioned on XtremIO storage (with VAAI enabled on the ESXi server) is approaching full capacity, it becomes suspended and the following message is displayed:



The VMware administrator can resolve the out-of-space situation on the XtremIO cluster, and prevent the guest operating system in the VMs from being unresponsive.

Also, it is possible to receive advanced notifications when nearing thin-provisioning overcommit on the XtremIO cluster. To set these advanced notifications, use the VAAI TP limit setting on the XtremIO cluster that is connected to vSphere. This setting enables you to define a VAAI thin provisioning soft limit (TPST) on XtremIO volumes. The limit is set as a percentage of the storage capacity and can be as high as 100%. The set limit is used as the threshold that triggers an advanced warning on the SCSI interface for specific XtremIO volumes, with set VAAI limits (VAAI soft limit).

The following figure demonstrates enabling VAAI TP alerts on an XtremIO Volume using the WebUI:



For further details on setting VAAI TP limit, see the *XtremIO Storage Array User Guide*.

Configuring boot from SAN with XtremIO

This section provides instructions for configuring boot from SAN with XtremIO.

This section includes the following topics:

- Boot from SAN Best Practices
- Configuring the Qlogic HBA BIOS for SAN Boot
- Configuring the Emulex HBA BIOS for SAN Boot
- Configuring Cisco UCS for SAN Boot

NOTE: Windows physical host UEFI boot from SAN with an XtremIO cluster is not supported for all XIOS versions. This is due to an I/O size mismatch between Windows and XtremIO for UEFI boot from SAN that cannot be adjusted in Windows.

Boot from SAN best practices

Consider the following points when using boot from SAN:

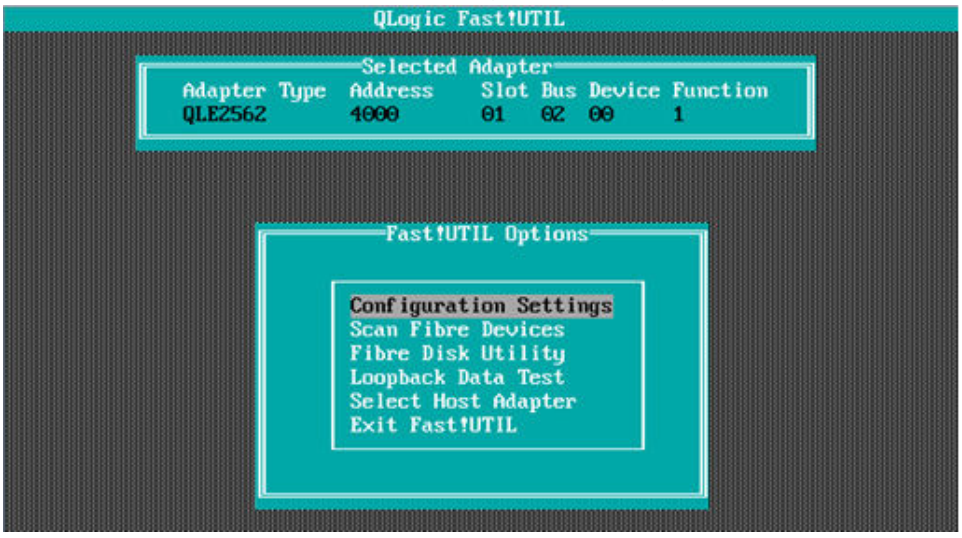
- If there are multiple HBAs on the server, use the lowest numbered PCI slot in the host.
- Make sure that only one volume, the boot device, is presented to the host. The LUN ID for this volume should be 0. More volumes can be added after operating system installation on the host.
- Only a single path is configured from the host HBA to the boot device.

Configuring the Qlogic HBA BIOS for SAN Boot

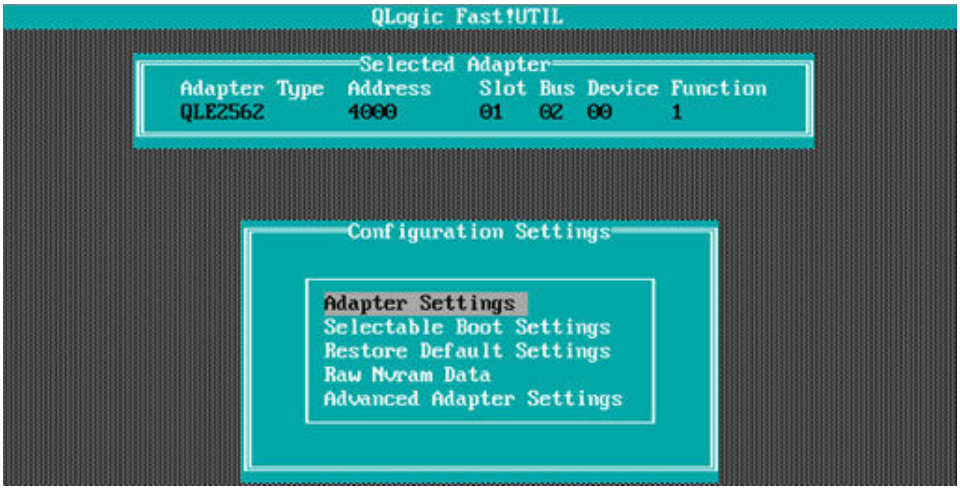
When the BIOS is installed and enabled, it needs to be configured to enable booting from the SAN. To configure the BIOS:

Steps

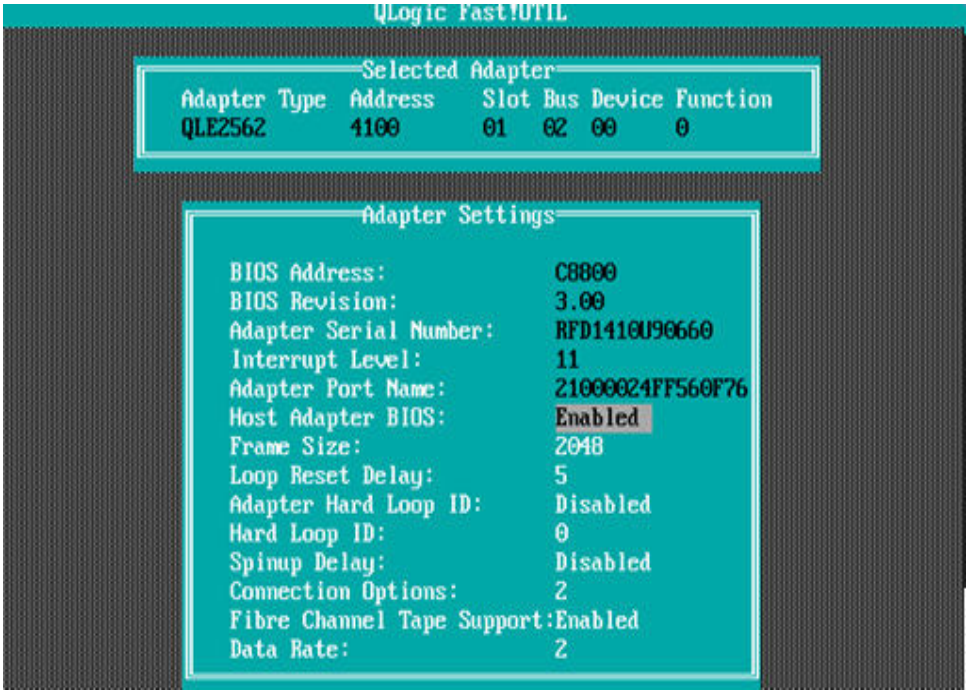
1. Connect the XtremIO Storage Controller port to the adapter in the lowest-numbered PCI slot in the server (for example, if there are three adapters in the system, which is located in slots 2, 4 and 5, connect the cable to the adapter in slot 2). Do not connect cables to the other adapters.
2. Boot the server and press **Ctrl-Q** when the Qlogic banner is displayed. The banner display corresponds to the BIOS revision pertinent to the adapters that are installed. For specific adapters and BIOS revisions, see the Dell EMC Online Support.
3. When Fast!UTIL loads, a list of addresses that are occupied by the adapters are displayed. Select the adapter from the list and press **Enter**; the Fast!UTIL Options menu displays.
4. From the Fast!UTIL Options menu, select **Configuration Settings** and press **Enter**.



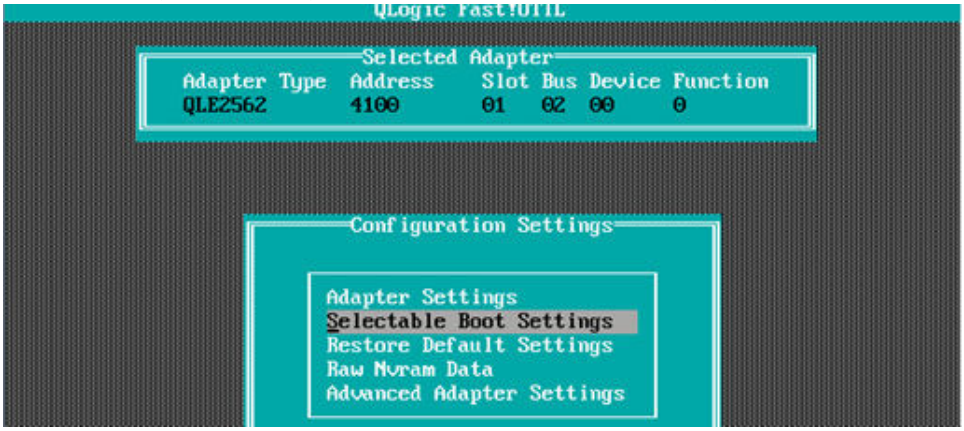
5. From the Configuration Settings menu, select **Adapter Settings** and press **Enter**.



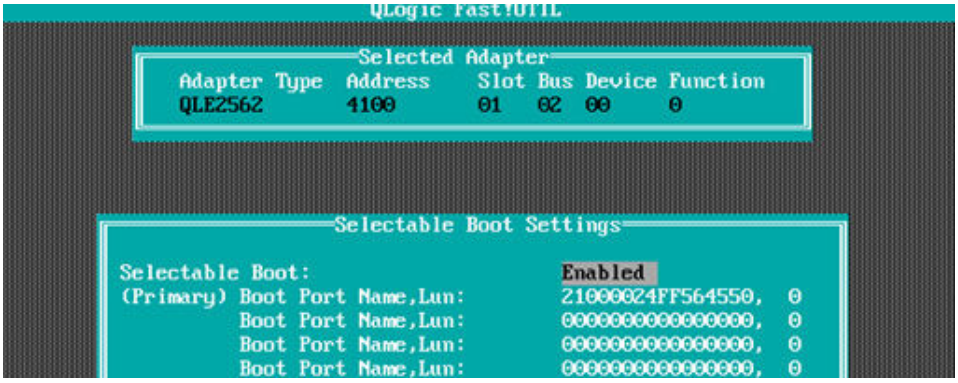
6. From the Host Adapter menu, select **Host Adapter BIOS** and press **Enter** to enable it (if it is not already enabled).



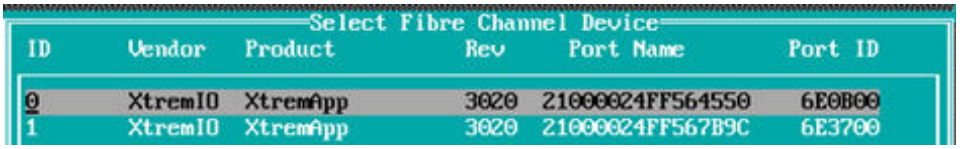
7. Press **ESC** to exit the Adapter Settings menu.
8. From the Configuration Setting menu, select **Selectable Boot Settings** and press **Enter**.



9. From the Selectable Boot Settings menu, select **Selectable Boot Device** and press **Enter** to enable this option (if it is not already enabled).



10. Select **Current Boot Node Name** and press **Enter**. The adapter scans for attached storage devices which are then displayed on the screen.
11. Select the XtremIO Storage Array port from which the system boots. The entry of the port is similar to the following example:



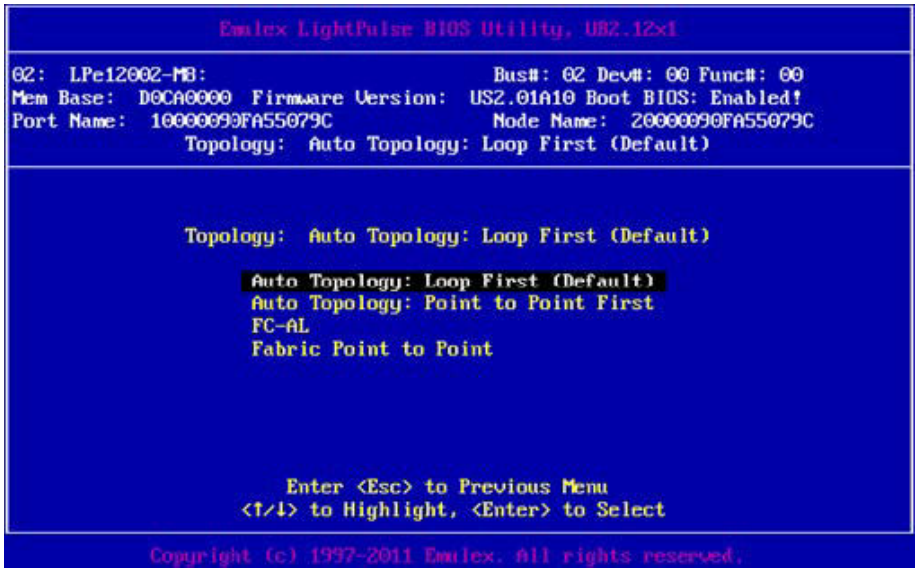
12. From the displayed list of LUNs, select the LUN to be used as the boot device and press **Enter**.
13. In the Selectable Boot Settings menu, press **ESC** to return to the Configuration Settings menu.
14. In the Configuration Settings menu, press **ESC** to return to the Options menu.
15. Select **Save Changes** and press **Enter**.
16. Press **ESC** to exit the Fast!UTIL menu.
17. Reboot the host.

Configuring the Emulex HBA BIOS for SAN Boot

When the boot BIOS is installed and enabled, it needs to be configured to enable booting from the SAN. To configure the boot BIOS:

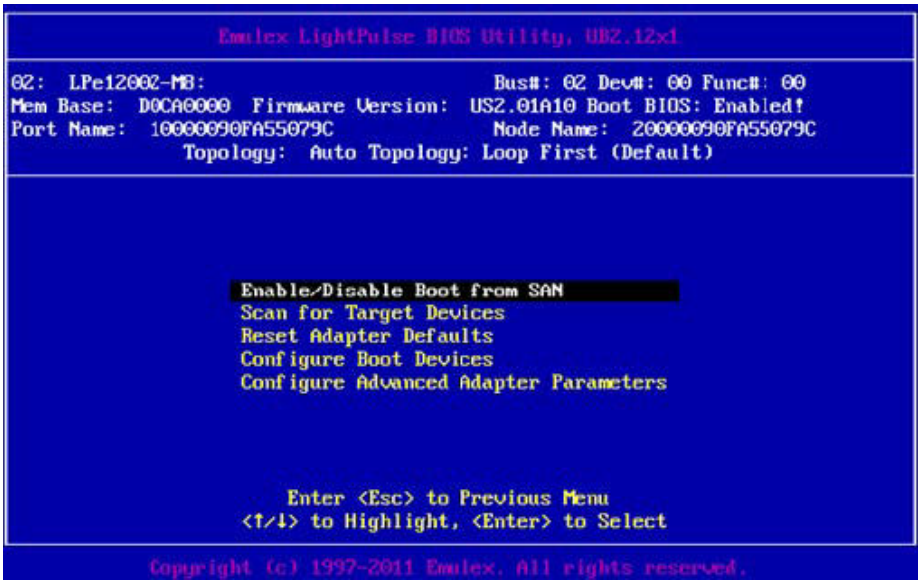
Steps

1. Connect the XtremIO Storage Controller port to the adapter in the lowest-numbered PCI slot in the server (for example, if there are three adapters in the system, located in slots 2, 4 and 5, connect the cable to the adapter in slot 2). Do not connect the cable to the other adapters.
2. Boot the server and press **ALT-E** when the Emulex banner opens to display the BIOS setup utility. The banner display corresponds to the BIOS revision pertinent to the adapters installed. For specific adapters and BIOS revisions, see Dell EMC Online Support.
3. From the displayed list of adapters with the boot BIOS installed, select one HBA port and press **Enter**.
4. Verify that the displayed topology is *Auto Topology: Loop First (Default)*. If a different topology is displayed, perform the following steps:
   a. Select **Configure Advanced Adapter Parameters**.
   b. Select **Topology Selection**.
   c. From the Topology Selection menu, select **Auto Topology: Loop First (Default)** and press **Enter**.



d. Press **ESC** and then press **ESC** again to return to the main screen.

5. On the main page, select **Enable/Disable Boot from SAN** and press **Enter**.

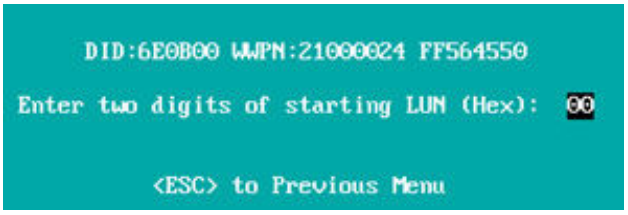


6. If the BIOS is listed as *Disabled*, select **1** to enable it and press **Enter**.
7. Press **ESC** to return to the main screen.
8. From the main screen, select **Configure Boot Devices** and press **Enter**; the screen that opens displays a list of devices available for use as a boot LUN. Initially, the values are set to zero. However, once the information regarding the boot LUN is obtained, it is logged in the field.
9. At the prompt, select the boot entry **1** (this entry is used to enter the information regarding the target boot device).

10. In the next screen, select **00** to clear the selected priority slot, or select the device from which the host will reboot.

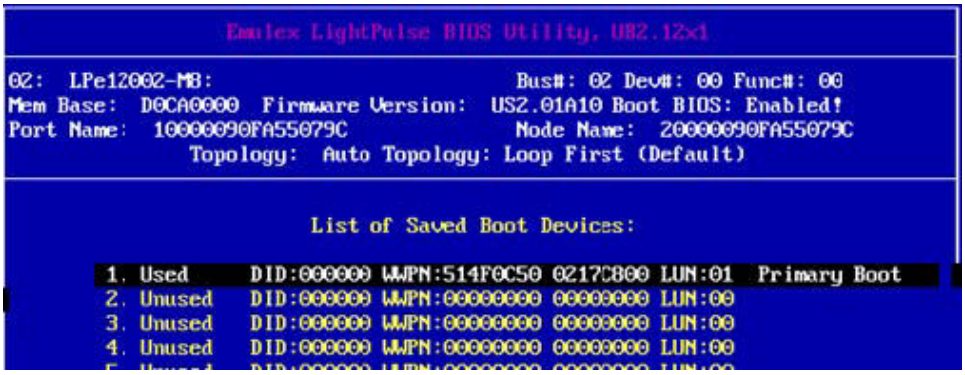
In the following example, entries 01 and 02 refer to devices from XtremIO version 4.0 code, while entries 03 and 04 refer to devices from XtremIO 2.4 code.

11. Select the boot device (for example 01); the displayed list also provides the LUN number of the first LUN visible on the device.
12. Select the LUN that you want to boot from and press **Enter**.



13. When the next screen opens, press **Enter**.
14. When prompted to select whether to boot through WWPN or DID, select **Boot this device via WWPN**.

15. Press **Enter** to return to the List of Saved Boot Devices screen; the device with the corresponding target WWN is displayed.



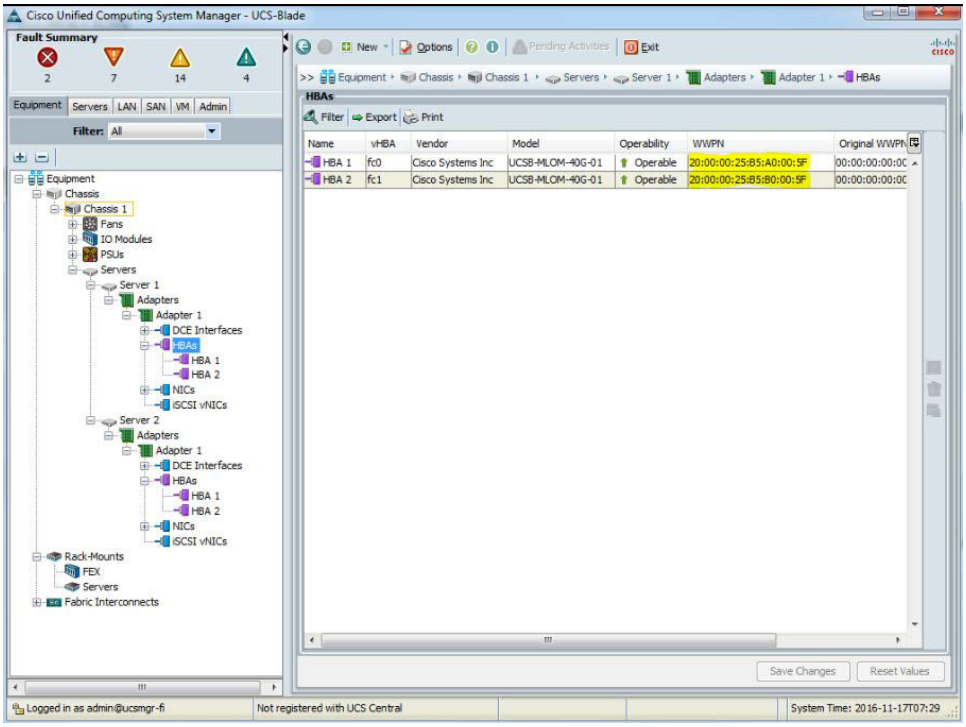
16. Press **ESC** to exit from the menu.
17. Save the configuration and reboot the system.
18. Reboot the host.
    The HBA is configured to boot from the desired LUN on the X-Brick.

Configuring Cisco UCS for SAN Boot

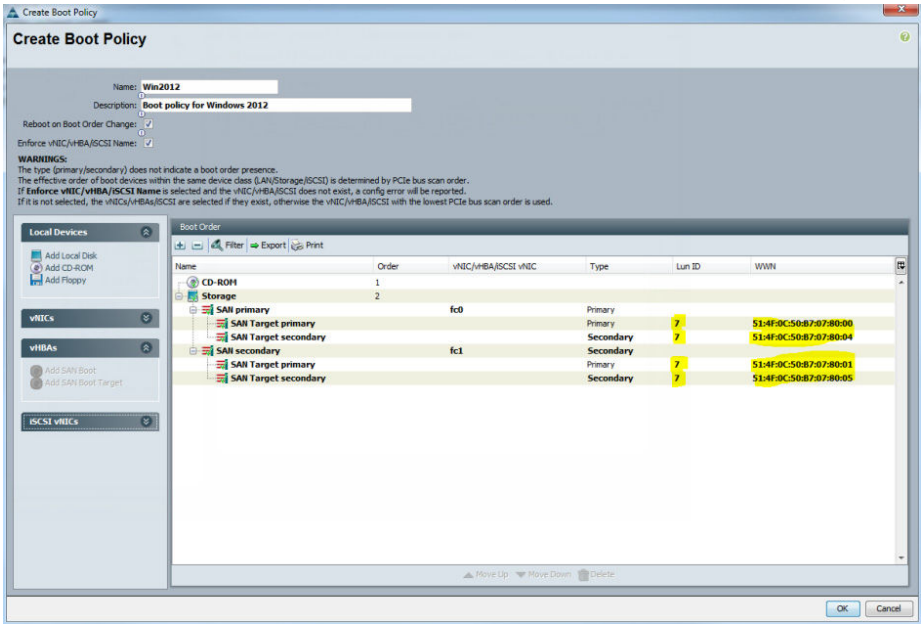
To configure Cisco UCS for SAN boot:

Steps

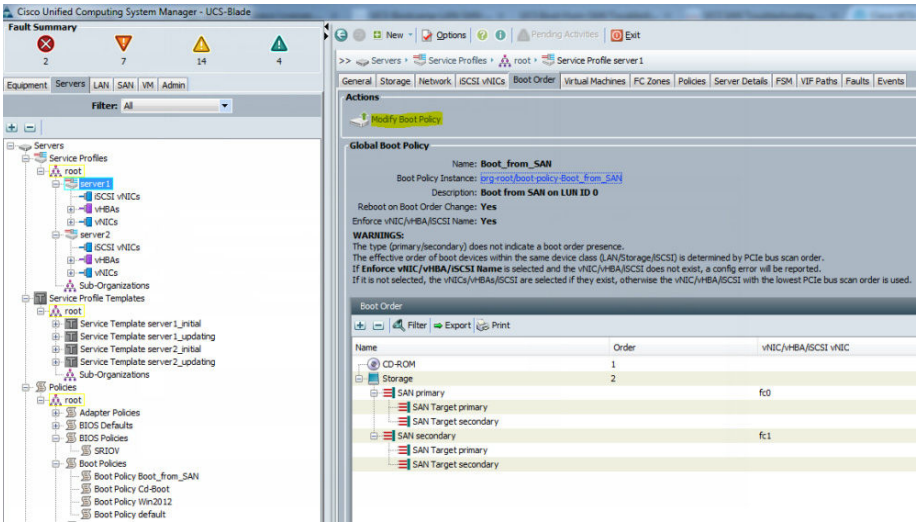
1. Determine the WWPN of HBAs on UCS server.



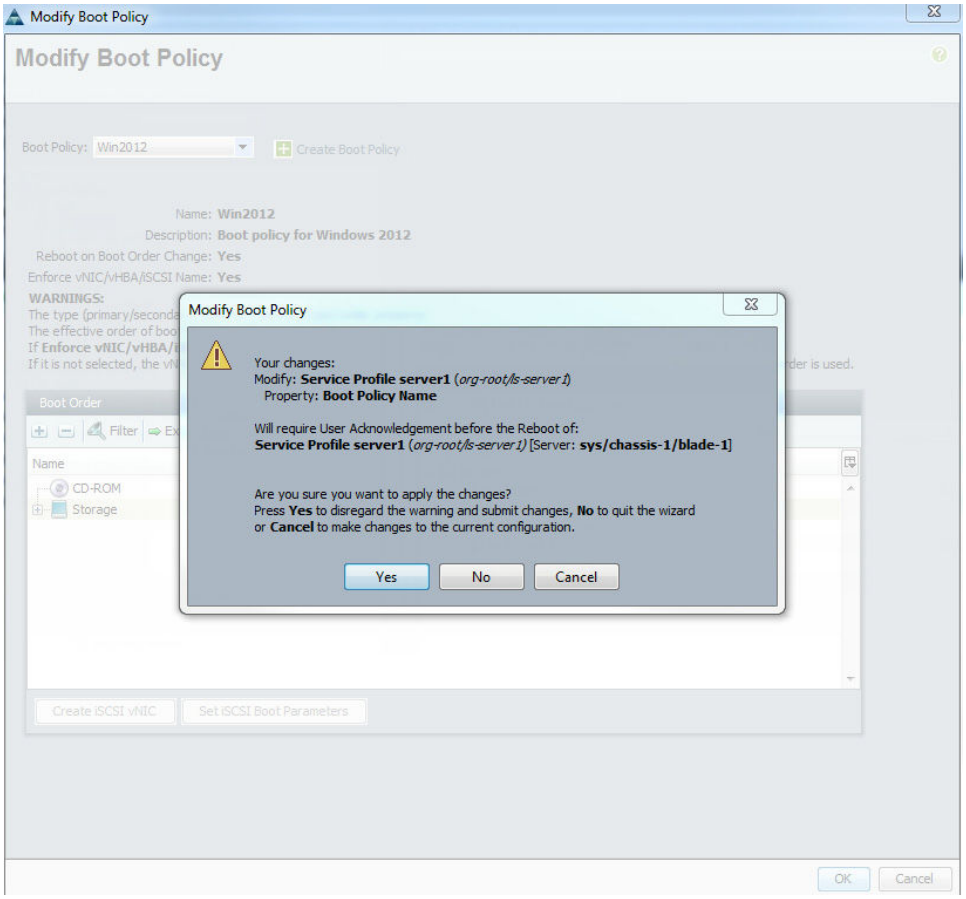
2. Zone each initiator with XtremIO targets on the FC switch (Brocade or Cisco), or connect SC-FC and UCS fabric directly.
3. On XtremIO, create an Initiator Group consisting of the UCS Initiators to be connected to XtremIO array.
4. On XtremIO, create a volume that will be used for booting.
5. Map the created volume to the previously created Initiator Group.
6. Create a Boot Policy consisting of the XtremIO targets (WWN) which are zoned together with UCS initiators.



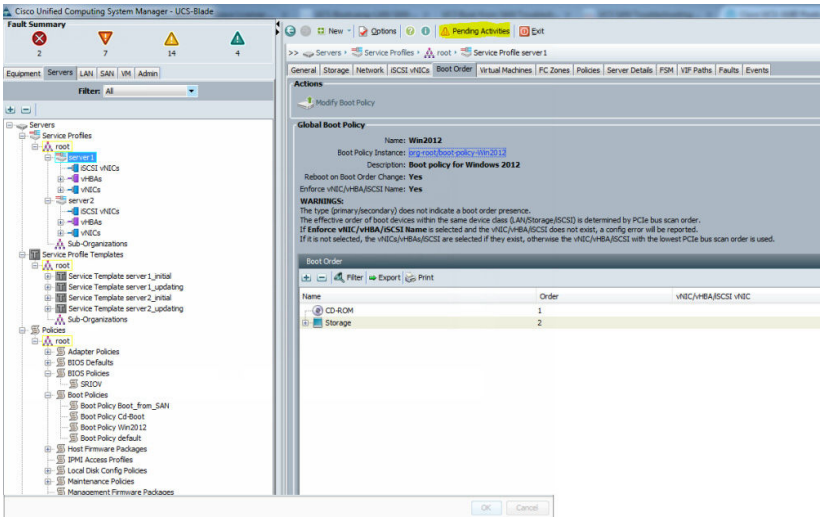
7. Modify the boot policy on the UCS server and select the previously created boot policy.



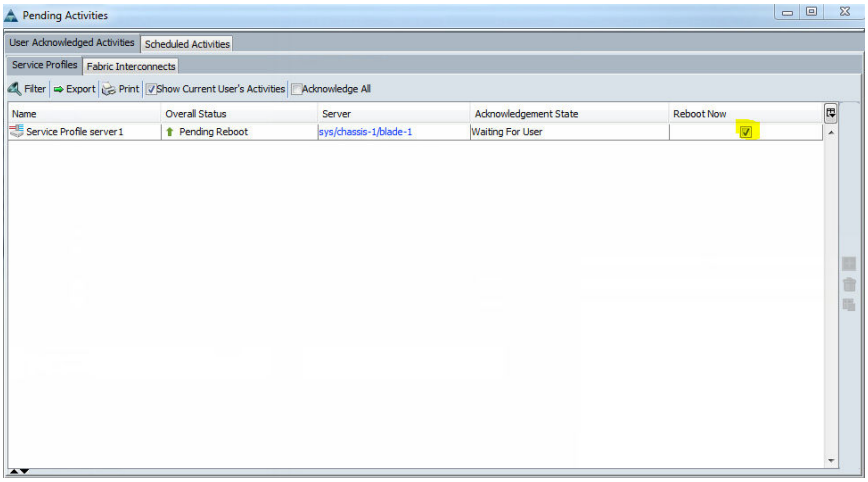
8. Submit the changes.



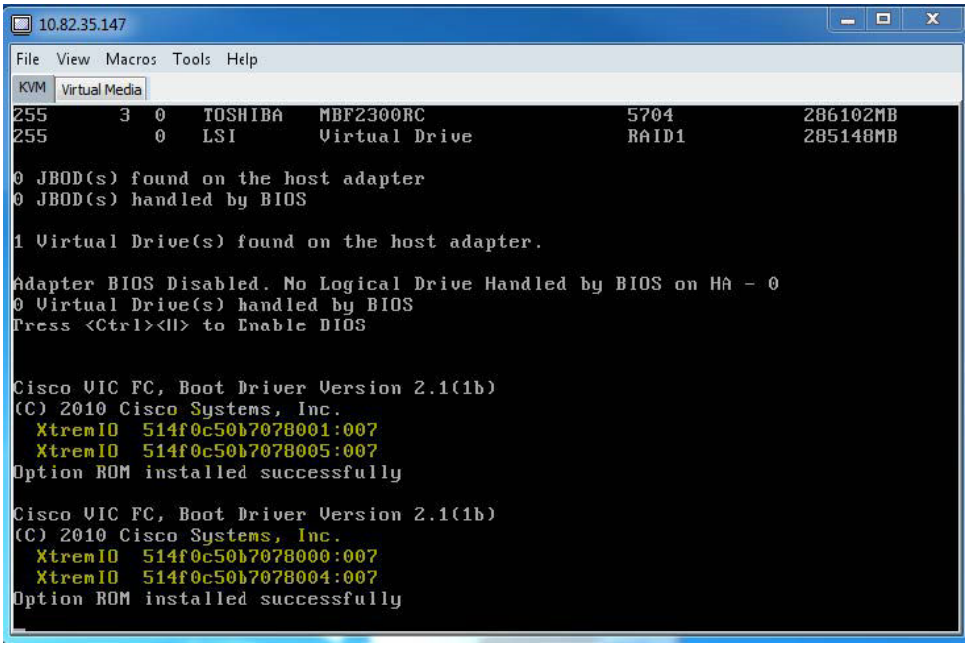
9. Once submitted, click **Pending Activities** and confirm Reboot of Service Profile.



10. Select the **Reboot Now** option and click **Yes** or **Apply** to confirm.



11. During the server reboot, the XtremIO volume should be identified by the Cisco VIC FC Boot Driver.



12. To install any supported operating system on this volume, select a LAN to install, using PXE or CD-ROM, as the first boot option. See the Cisco Compatibilty Matrix to understand the compatible components.

Executing the ESXi Host Validation Script on ESXi Hosts

After all ESXi hosts connected to XtremIO are configured according to the recommendations provided in this document, run the ESXi Host Validation Script (HVS) on these hosts. This validates that the host configuration is aligned with these recommendations.

For details on using the ESXi HVS, see the Dell EMC Knowledge Base article 498027.

5

Operating System-Specific Features

Topics:

Virtual Volumes

Virtual Volumes (vVols) have enhanced vSphere 6.x. vVols changes the way storage is presented, managed, and consumed, making the storage system VM-centric, as shown in following figure.

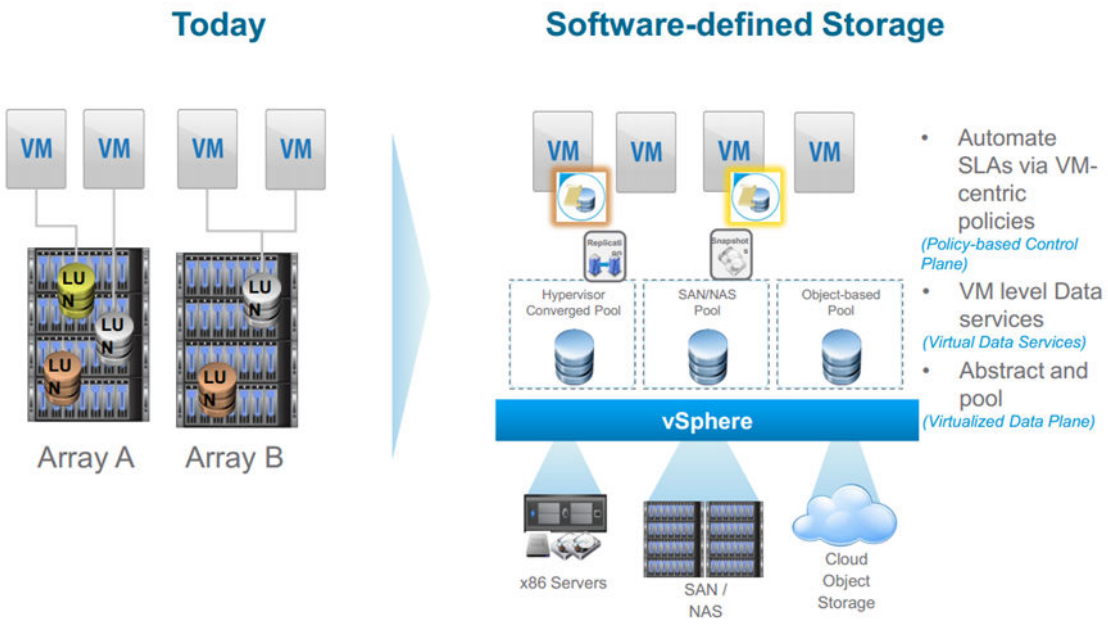


Figure 16. VM example

VVol is part of VMware's Software Defined Storage, which is split between the control plane with Virtual Data Services (all policy driven) and the data plane with Virtual Data Plane, which is where the data is stored.

All storage is LUN-centric or volume-centric, especially in snapshots, clones, and replication. VVols makes the storage VM-centric. With vVol, you can offload data operations to the storage arrays. However, vVols makes storage arrays aware of individual Virtual Machine Disk (VMDK) files.

The following concepts are introduced to provide the management capabilities of vVols:

- Vendor Provider (VP) - management of data operations

  VP is a plug-in that storage vendors write. The VP uses out-of-band management APIs, VASA (updated to version 2.0). The VP exports storage array capabilities and presents them to vSphere through the VASA APIs.
- Storage Containers (SC) - management of data capacity

SCs are chunks of physical storage in which you create and logically group vVols. SCs were previously datastores. SCs are based on the grouping of VMDKs onto which application-specific SLAs are translated to capabilities through the VASA APIs. Hardware capacity limits SC capacity. You need at least one SC per storage system but you can have multiple SCs per array. Storage array capabilities such as replication, encryption, and snapshot retention are assigned to an SC to create different SCs for different grouping of requirements. If you have, for instance, physical storage with SSD disks and another with SAS disks, you would present these as different SCs.

- Protocol Endpoints (PE) - management of access control and communications

  PEs are the access points from the hosts to the storage systems, which storage administrators create. PEs administer all paths and policies. PEs are compliant with FC, iSCSI, and NFS. They are intended to replace the concept of LUNs and mount points.

  vVols are *bound* and *unbound* to a PE. vCenter initiates the *bind* or *unbind* operation. You can apply existing multipath policies and NFS topology requirements to the PE.

  Storage administrators no longer need to configure LUNs or NFS shares. They set up a single I/O access PE for the array to set up a data path from VMs to vVols, and create SCs based on groupings of required capabilities and physical storage. Then the VM administrator creates vVols in these SCs through the vCenter.

VMAX3/PowerMax vVol

See the Using VMware Virtual Volumes with Dell EMC VMAX and PowerMax Guide for instructions to create and integrate vVol datastore from VMAX to ESXi.

See the VMware Compatibility Guide for the certified VASA Provider version and download URL.

Policy-based management

Policies are set based on the need for capacity, performance, and availability of an application. These are capabilities that the array advertises through the vStorage APIs for Storage Awareness (VASA) APIs.

Define the desired policies and then assign the VVols to policies. Then the external storage array automates control of where these VVols ultimately are located, replicates or encrypts them, and manages the service levels. This gives you per-VM storage service levels coming from a single self-tuning datastore. You can define performance rules as part of a VVol. For example; you can specify minimum and maximum ReadOPs, WriteOPs, ReadLatency, and WriteLatency.

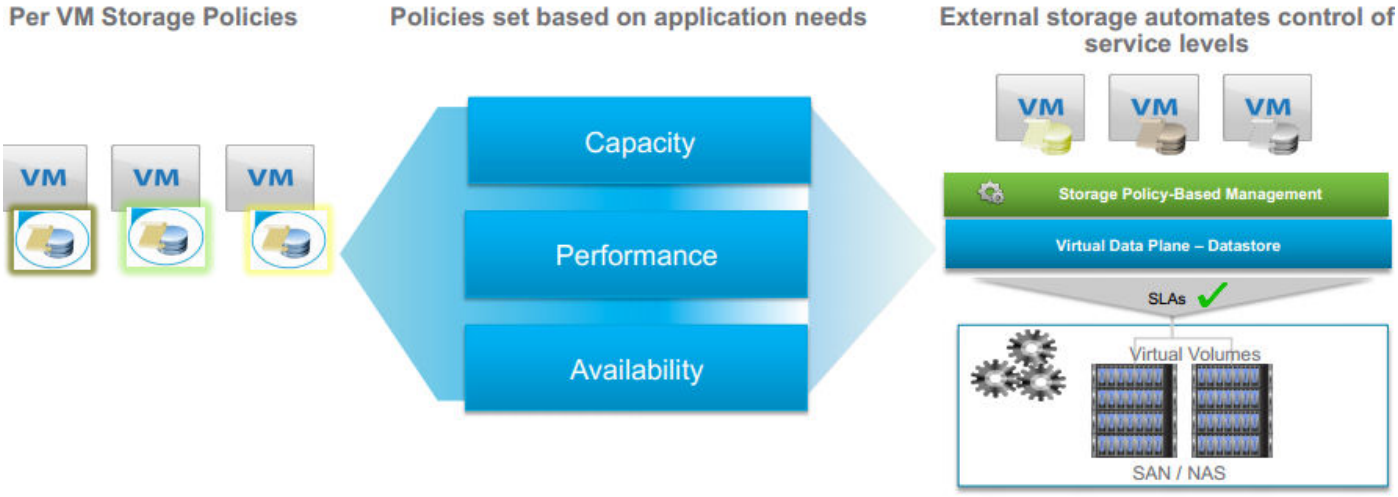


Figure 17. Policy-based management example

To understand how these policies are implemented, consider Virtual SAN (VSAN). VSAN creates a single datastore and then manages capacity, performance, and availability per-VM using policies within this single datastore. VSAN is implemented differently from VVols. VVols use VASA 2.0 to communicate with the VASA Provider of an array to manage VVols on the array. However, VSAN uses its own APIs to manage virtual drives. Storage Policy Based Management is used by both VVols and VSAN to present and use storage-specific capabilities.

A VM will have a number of VVols which could have different policies associated with them, if needed. When a single VVol is created, it contains the following VM configurations:

- One VVol for every virtual drive

- One VVol for swap
- One VVol per disk snapshot
- One VVol per memory snapshot

When you snapshot a VM using the Web Client, it is translated into simultaneous snapshots of all the virtual drives of the VMs together, with a snapshot of the memory of the VM, if requested. However, with the API, you can take a snapshot of the individual VMDK VVols.

Fault tolerance support for 4 vCPUs

Fault Tolerance (FT) supports VMs with up to 4 x vCPUs and 64 GB RAM. Symmetric MultiProcessing Fault Tolerance (SMP-FT) works differently from FT for single CPUs.

There is a new, Fast Checkpointing mechanism to keep the primary and secondary VMs synchronized. Previously a Record-Replay sync mechanism was used, but the new Fast Checkpointing has enabled FT to expand beyond 1 x vCPU. Record-Replay kept a secondary VM in *virtual lockstep* with the primary VM. With Fast Checkpointing, the primary and secondary VM run the same instruction stream simultaneously, making it faster. If the FT network latency is too high for the primary and secondary VMs to stay synchronized, the primary VM is slowed down. You can also hot-configure FT.

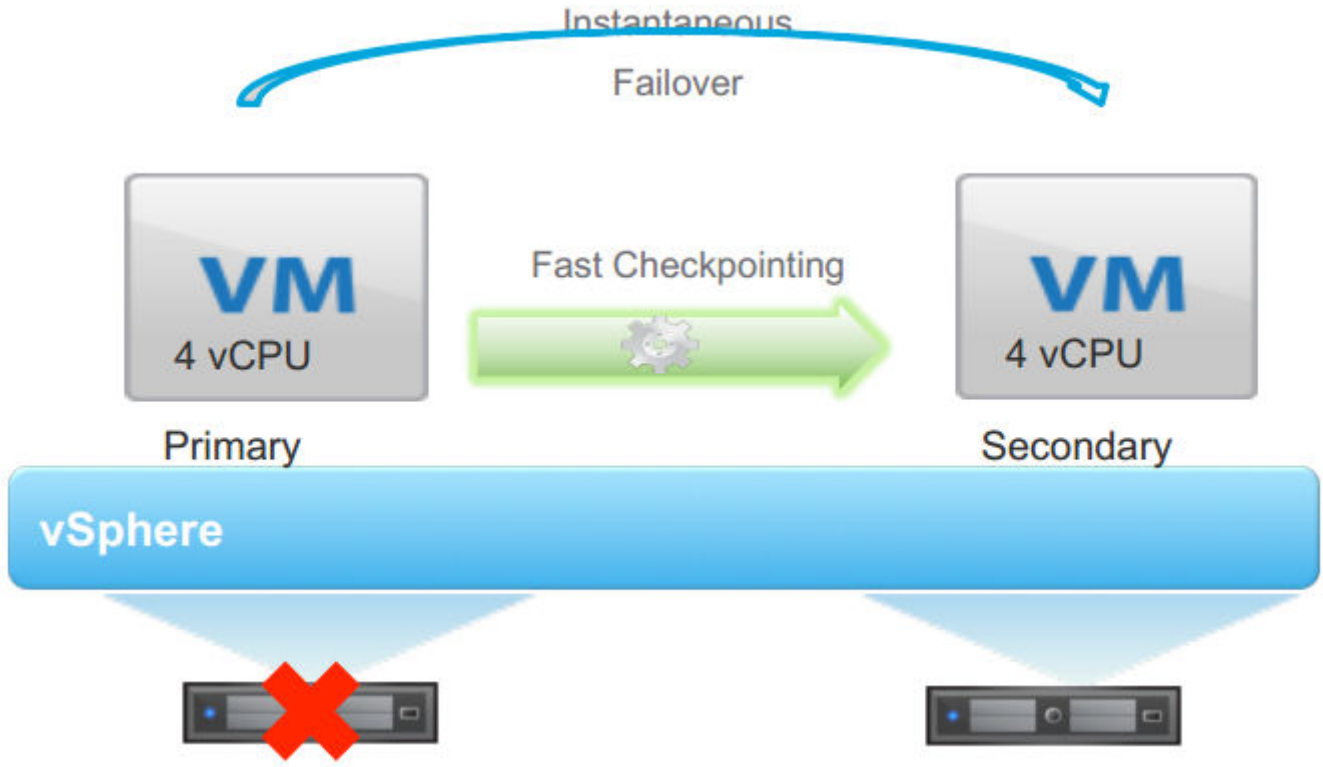


Figure 18. Fault tolerance example

There are restrictions on the number of FT protected VMs. You can have a maximum of four FT protected VMs or eight FT protected vCPUs per host, whichever limit is reached first. The maximum includes both primary and secondary VMs. A 10 Gb dedicated NIC is a requirement for SMP-FT, but you may be able to share bandwidth using NIOC with FT given priority. You cannot use hot-add for RAM or CPU with FT in vSphere 6.0.

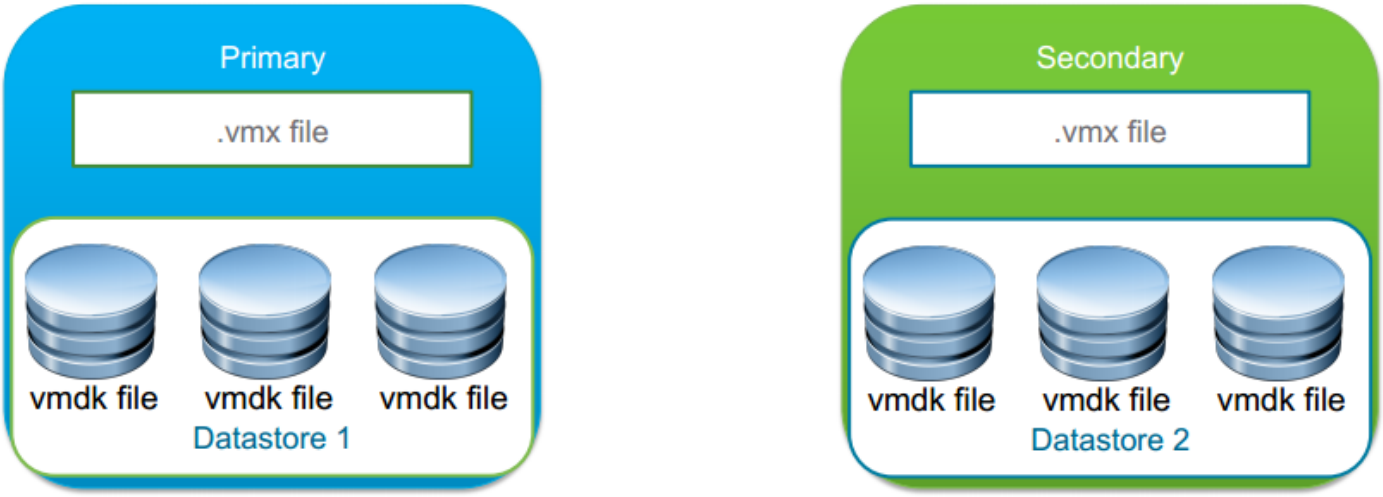


Figure 19. Creation of secondary VMX file with SMP-FT

With SMP-FT, two complete VMs are created. The secondary VM is a separate VM (including storage for the first time) and it can now be on a different datastore from the primary. This means FT can protect a VM not only against a host failure, but also against a datastore issue/failure as well with zero downtime. As a separate, secondary VM is created, it uses additional I/O. You need shared storage for the tiebreaker file, but the .vmx and .vmdk files do not have to be on shared storage; they could be local disks. In previous vSphere versions, you were required to create FT VMs with Eager Zero Thick disks but with vSphere 6.0, all disk types are supported. Para Virtualization Devices are also supported.

vMotion is also supported for both primary and secondary VMs. DRS initial placement chooses where the VMs are placed upon system start-up. DRS does not balance FT VMs but can balance other non-FT VMs. If the primary host fails, HA (High-availability) cluster is FT aware and takes over the functionality on the secondary. If either the primary or secondary host fails, FT automatically starts to create another secondary VM. During a host reboot, a new secondary VM is created automatically and started on another host.

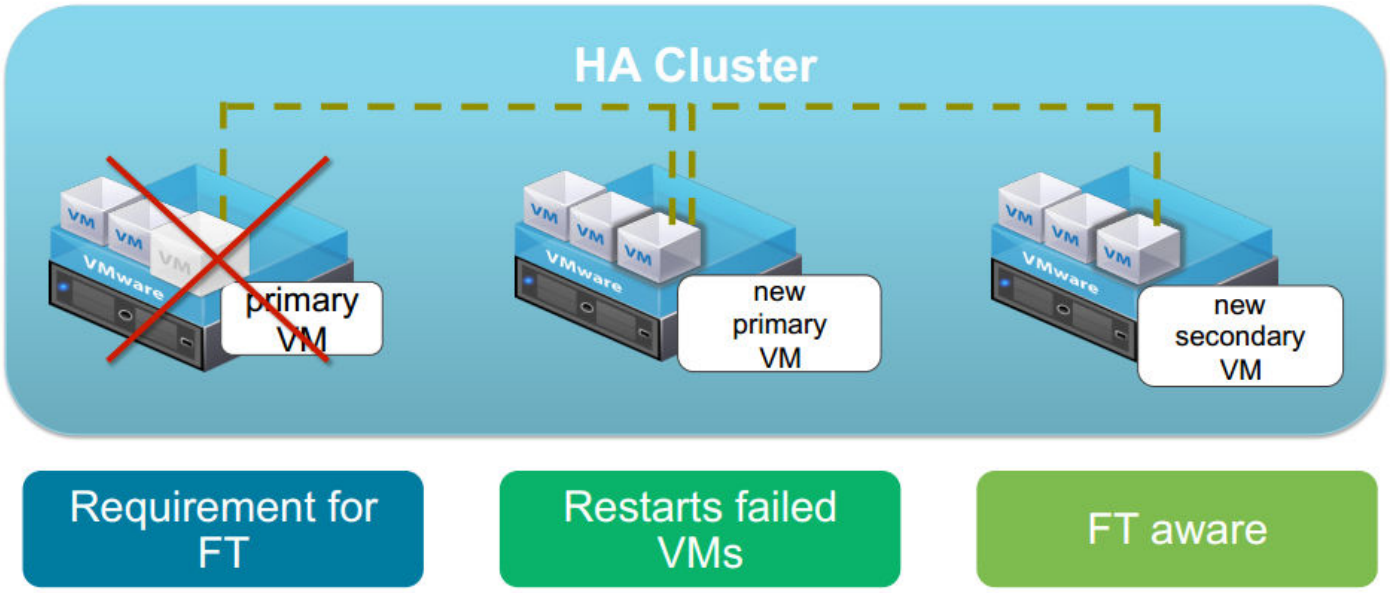


Figure 20. FT creates secondary VM when primary fails

Storage vMotion is not available for FT with multiple vCPUs and vCloud Director; vSphere Replication, VSAN/vVols, and vFlash do not support SMP-FT.

SMP-FT has made some improvements to backup, such as vStorage APIs for Data Protection (VADP) support and snapshots which were not available in vSphere 5.5.

There is a performance penalty of SMP-FT to the guest depending on workload. This could be between 10 and 30 percent and the Host CPU impact is minimal.

Long-distance vMotion

Long-distance vMotion now supports 100-millisecond roundtrip.

In vSphere 5.5, vMotion is supported within a single cluster and across clusters within the same Datacenter and vCenter. With vSphere 6.0, vMotion is expanded to include vMotion across vCenters, across virtual switches, across long distances, and routed vMotion networks aligning vMotion capabilities with larger data center environments.

vMotion across vCenters simultaneously changes compute, storage, networks, and management. This uses vMotion with unshared storage and supports local, metro, and cross-continental distances.

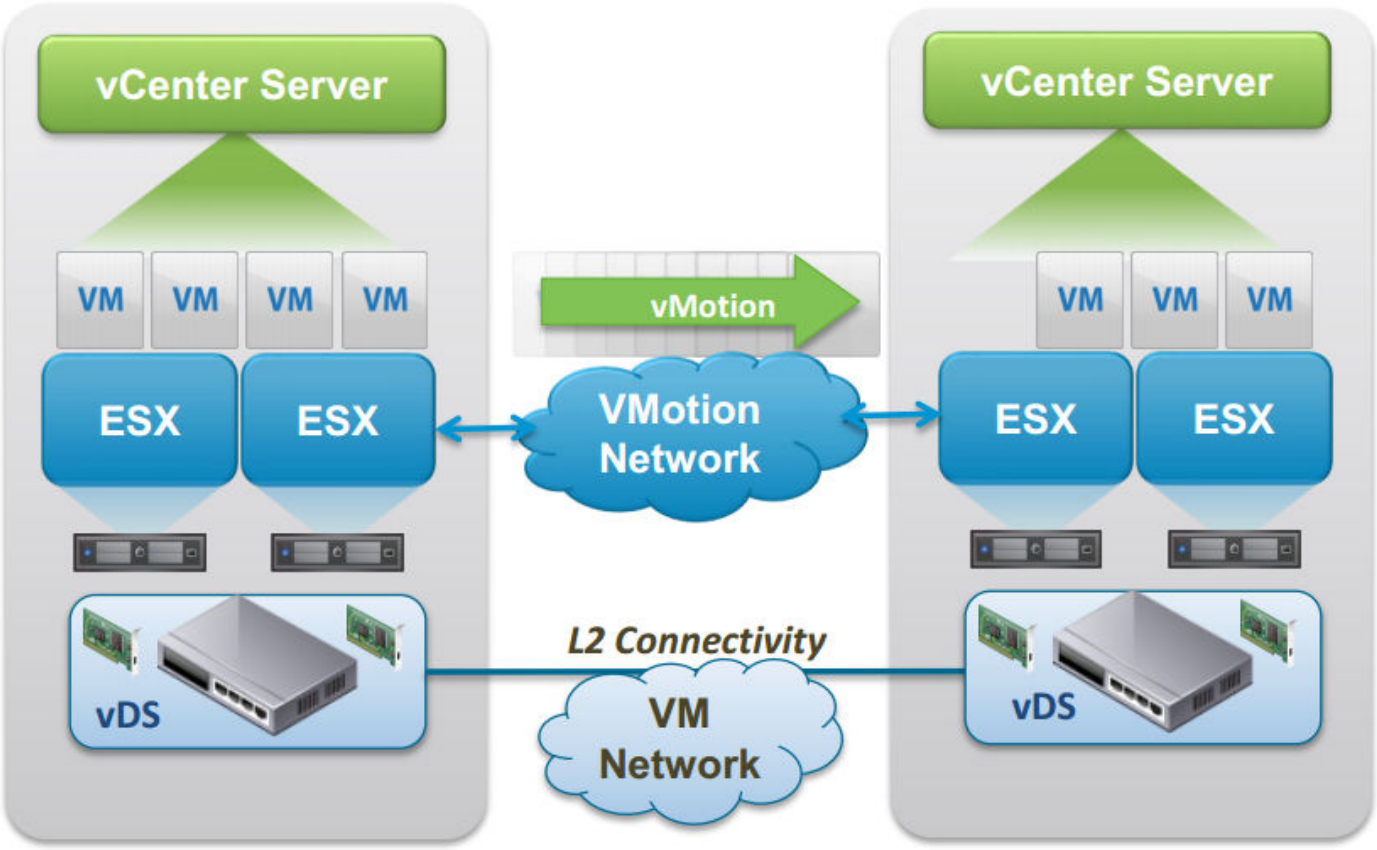


Figure 21. vMotion Connectivity across vCenters

If you use the GUI to initiate vMotion, you need the same Single Sign-On (SSO) domain for both vCenters. The VM UUID can be maintained across vCenter Server instances, but with the API, you can have a different SSO domain. VM historical data such as Events, Alarms, and Task History is preserved. Performance Data is preserved once the VM is moved but is not aggregated in the vCenter UI. The information can still be accessed using third-party tools or the API using the VM instance ID, which remains across vCenters.

When a VM moves across vCenters, HA properties are preserved and DRS anti-affinity rules are honored. The standard vMotion compatibility checks are run. You need 250 Mbps network bandwidth per vMotion operation.

Another new function is the ability to vMotion or clone powered-off VMs across vCenters. This uses the VMware Network File Copy (NFC) protocol.

vMotion previously could only occur within a network that a single virtual switch manages, either a Virtual Standard Switch (VSS) or Virtual Distributed Switch (VDS). vMotion across vCenters enable VMs to vMotion to a network that a different virtual switch manages. These include:

- From VSS to VSS
- From VSS to VDS
- From VDS to VDS

You cannot vMotion from a VDS to a VSS. VDS port metadata is transferred and the cross vCenter vMotion is still transparent to the guest operating system. You need Layer 2 VM network connectivity.

In vSphere 5.5, vMotion requires Layer 2 connectivity for the vMotion network. vSphere 6.x enables VMs to vMotion using routed vMotion networks.

Another addition in vSphere 6.x is the ability to support long-distance vMotion. It can support cross-continental US distances with up to 100+millisecond RTTs while still maintaining standard vMotion guarantees. Use cases are:

- Disaster avoidance
- SRM and disaster avoidance testing
- Multisite load balancing and capacity utilization
- Follow-the-sun scenarios
- You can also use Long-distance vMotion to live-move VMs onto vSphere-based public clouds, including VMware VCHS (now called vCloud Air).

In long-distance vMotion, a Layer 2 connection is required for the VM network in both source and destination. The same VM IP address must be available at the destination. vCenters must connect through Layer 3, and the vMotion network can now be a Layer 3 connection. The vMotion network can be secure either by being dedicated or encrypted (VM memory is copied across this network).

vMotion includes over a VMs CPU and Memory, but storage needs also to be considered if you are moving VMs across sites and arrays. There are various storage replication architectures to enable this. Active-Active replication over a shared site, as with a metro cluster, is displayed as shared storage to a VM, so this works like classic vMotion. For geo-distance vMotion, where storage Active-Active replication is not possible, VVols are required, creating a use case for VVols.

- vMotion across vCenters - vMotion using routed vMotion networks and vMotion across virtual switches.
- Using VMware NSX, network properties are vMotioned as well, when using long-distance vMotion.

Virtual Datacenters

In vSphere 6.x, a Virtual Datacenter aggregates compute clusters, storage clusters, network, and policies.

A virtual Datacenter can aggregate resources across multiple clusters within a single vCenter Server into a single large pool of capacity. This benefits large deployments such as VDI, where you have multiple clusters with similar network and storage connections. Now you can group them together. Within this single pool of capacity, the Virtual Datacenter automates VM initial placement by deciding into which cluster the VM should be placed based on capacity and capability.

You can then create VM placement and storage policies and associate these policies with specific clusters or hosts as well as the datastores to which they are connected. This policy could be a policy to store SQL VMs on a subset of hosts within a particular cluster for licensing reasons. You can then monitor adherence to these policies and automatically rectify any issues. When you deploy a VM, you would select from various policies. The Virtual Datacenter then decides where a VM must be placed, based on those policies. This helps reduce the administrator's workload.

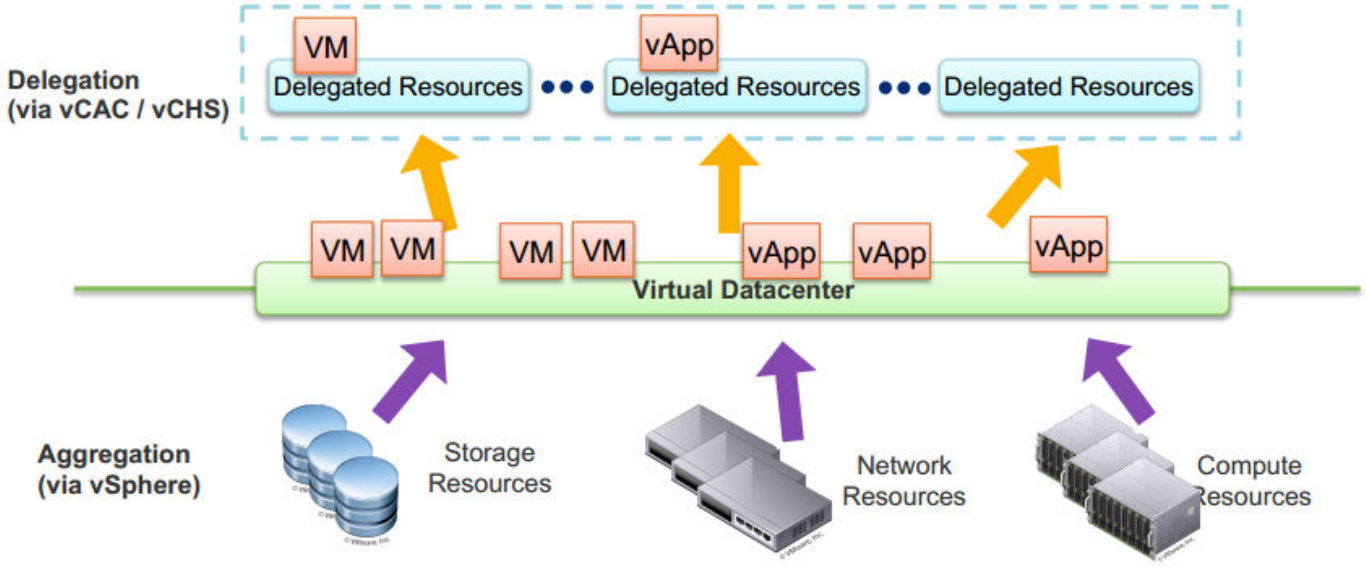


Figure 22. Virtual Datacenter delegates allocation of resources

Virtual Datacenters require Distributed Resource Scheduler (DRS)-enabled clusters to handle the initial placement. Individual hosts cannot be added. You can remove a host from a cluster within a Virtual Datacenter by putting it in maintenance mode. All VMs stay within the Virtual Datacenters, moving to other hosts in the cluster. If you need to remove a cluster or turn off DRS for any reason and are not able to use Partially Automated Mode, you should remove the cluster from the Virtual Datacenter. The VMs stay in the cluster but no longer

perform VM placement policy monitoring checks until the cluster rejoins a Virtual Datacenter. You can manually vMotion VMs to other clusters within the VDC before removing a cluster.

Platform Service Controller

Platform Service Controller (PSC) can be used as a shared component by vCenter, vCOPs, vCloud Director, vCloud Automation Center.

The PSC contains the following functionality:

- SSO
- Licensing
- Certificate Authority
- Certificate Store
- Service (Product) Registration

The Certificate Authority and Certificate Store are new components. The VMware Certificate Authority (VMCA) can act as a root certificate authority, either managing its own certificates or handling certificates from an external Certificate Authority. VMCA provides each ESXi host with a signed certificate when it is added to vCenter as part of installation or upgrade. You can view and manage these certificates from the vSphere Web Client and manage the full certificate life-cycle workflow.

Service (Product) Registration is a component where all the other services are registered. It is the lookup service in vSphere and it shows all the services that are running in the system.

The PSC is built into vCenter and runs as a vPostgres database so that there is no additional database need. It runs in both the Windows and appliance versions. Since PSCs self-replicate and do not use ADAM, they can replicate between Windows and appliance vCenters.

You can either have the PSC embedded within vCenter Server or run it as an external component to vCenter Server.

If you are running 8 or fewer vCenters, it is best to use PSC embedded with vCenter. vCenter connects only to its internal PSC.

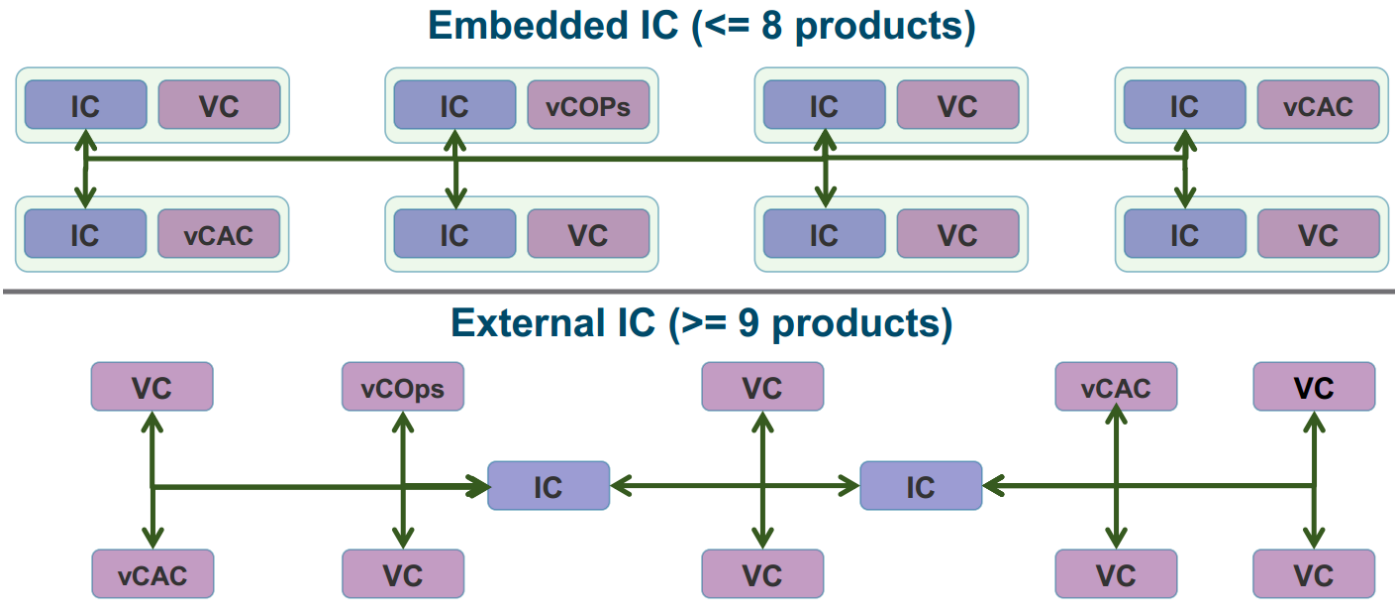


Figure 23. PSC Service shared by many vCenters

NOTE: The PSC is a new service and separate from the Inventory Service, which handles inventory management of VMs.

vCenter Server Appliance

The vCenter Server Appliance (VCSA) has also been improved. With vSphere 5.1, you can manage 100 hosts and 3,000 powered-on VMs. vSphere 6 permits 1,000 hosts and 10,000 powered-on VMs.

vSphere Web Client

The vSphere Web Client is the primary method for system administrators and end users to interact with the virtual data center environment created by VMware vSphere. vSphere manages a collection of objects that make up the virtual data center, including hosts, clusters, virtual machines, data storage, and networking resources.

The vSphere Web Client is a Web browser-based application that you can use to manage, monitor, and administer the objects that make up your virtualized data center. You can use the vSphere Web Client to observe and modify the vSphere environment in the following ways:

- Viewing health, status, and performance information about vSphere objects
- Issuing management and administration commands to vSphere objects
- Creating, configuring, provisioning, or deleting vSphere objects

You can extend vSphere in different ways to create a solution for your unique IT infrastructure. You can extend the vSphere Web Client with additional GUI features to support these new capabilities, with which you can manage and monitor your unique vSphere environment.