

# vSphere Storage

Update 2

Modified 15 AUG 2020

VMware vSphere 5.5

VMware ESXi 5.5

vCenter Server 5.5

You can find the most up-to-date technical documentation on the VMware website at:

<https://docs.vmware.com/>

**VMware, Inc.**  
3401 Hillview Ave.  
Palo Alto, CA 94304  
[www.vmware.com](http://www.vmware.com)

Copyright © 2009-2020 VMware, Inc. All rights reserved. [Copyright and trademark information.](#)

# Contents

About vSphere Storage 13

Updated Information 14

## 1 Introduction to Storage 15

Storage Virtualization 15

Types of Physical Storage 16

Local Storage 16

Networked Storage 17

Target and Device Representations 20

Storage Device Characteristics 21

Display Storage Devices for a Host in the vSphere Web Client 22

Display Storage Devices for an Adapter in the vSphere Web Client 23

Supported Storage Adapters 23

Storage Adapter Characteristics 23

View Storage Adapters Information in the vSphere Web Client 24

Datastore Characteristics 24

Display Datastore Information in the vSphere Web Client 26

List Datastores for an Infrastructure Object 27

How Virtual Machines Access Storage 28

Comparing Types of Storage 28

## 2 Overview of Using ESXi with a SAN 30

ESXi and SAN Use Cases 31

Specifics of Using SAN Storage with ESXi 32

ESXi Hosts and Multiple Storage Arrays 32

Making LUN Decisions 32

Use the Predictive Scheme to Make LUN Decisions 33

Use the Adaptive Scheme to Make LUN Decisions 33

Choosing Virtual Machine Locations 34

Layered Applications 35

Array-Based (Third-Party) Solution 35

File-Based (VMFS) Solution 35

Third-Party Management Applications 36

SAN Storage Backup Considerations 36

Using Third-Party Backup Packages 37

## 3 Using ESXi with Fibre Channel SAN 39

Fibre Channel SAN Concepts	39
Ports in Fibre Channel SAN	40
Fibre Channel Storage Array Types	40
Using Zoning with Fibre Channel SANs	41
How Virtual Machines Access Data on a Fibre Channel SAN	41
<b>4 Configuring Fibre Channel Storage</b>	<b>43</b>
ESXi Fibre Channel SAN Requirements	43
ESXi Fibre Channel SAN Restrictions	44
Setting LUN Allocations	44
Setting Fibre Channel HBAs	44
Installation and Setup Steps	45
N-Port ID Virtualization	45
How NPIV-Based LUN Access Works	45
Requirements for Using NPIV	46
NPIV Capabilities and Limitations	47
Assign WWNs to Virtual Machines	47
Modify WWN Assignments	48
<b>5 Configuring Fibre Channel over Ethernet</b>	<b>50</b>
Fibre Channel over Ethernet Adapters	50
Configuration Guidelines for Software FCoE	51
Set Up Networking for Software FCoE in the vSphere Web Client	52
Add Software FCoE Adapters in the vSphere Web Client	53
<b>6 Modifying Fibre Channel Storage for ESXi</b>	<b>54</b>
Testing ESXi SAN Configurations	54
General Setup Considerations for Fibre Channel SAN Arrays	55
EMC CLARiiON Storage Systems in Fibre Channel SAN Configurations	56
EMC Symmetrix Storage Systems	56
IBM Systems Storage 8000 and IBM ESS800	57
HP StorageWorks Storage Systems	57
HP StorageWorks EVA	57
HP StorageWorks XP	58
Hitachi Data Systems Storage	58
NetApp Storage	58
LSI-Based Storage Systems	59
<b>7 Booting ESXi from Fibre Channel SAN</b>	<b>60</b>
Boot from SAN Benefits	60
Boot from Fibre Channel SAN Requirements and Considerations	61

Getting Ready for Boot from SAN	61
Configure SAN Components and Storage System	62
Configure Storage Adapter to Boot from SAN	63
Set Up Your System to Boot from Installation Media	63
Configure Emulex HBA to Boot from SAN	63
Enable the BootBIOS Prompt	64
Enable the BIOS	64
Configure QLogic HBA to Boot from SAN	65
<b>8 Booting ESXi with Software FCoE</b>	<b>67</b>
Requirements and Considerations for Software FCoE Boot	67
Best Practices for Software FCoE Boot	68
Set Up Software FCoE Boot	68
Configure Software FCoE Boot Parameters	69
Install and Boot ESXi from Software FCoE LUN	69
Troubleshooting Installation and Boot from Software FCoE	70
<b>9 Best Practices for Fibre Channel Storage</b>	<b>71</b>
Preventing Fibre Channel SAN Problems	71
Disable Automatic Host Registration in the vSphere Web Client	72
Optimizing Fibre Channel SAN Storage Performance	73
Storage Array Performance	73
Server Performance with Fibre Channel	73
Fibre Channel SAN Configuration Checklist	74
<b>10 Using ESXi with iSCSI SAN</b>	<b>76</b>
iSCSI SAN Concepts	76
Ports in the iSCSI SAN	77
iSCSI Naming Conventions	77
iSCSI Initiators	78
Establishing iSCSI Connections	79
iSCSI Storage System Types	80
Discovery, Authentication, and Access Control	80
Error Correction	81
How Virtual Machines Access Data on an iSCSI SAN	82
<b>11 Configuring iSCSI Adapters and Storage</b>	<b>84</b>
ESXi iSCSI SAN Requirements	85
ESXi iSCSI SAN Restrictions	85
Setting LUN Allocations for iSCSI	86
Network Configuration and Authentication	86

Setting Up Independent Hardware iSCSI Adapters	86
View Independent Hardware iSCSI Adapters in the vSphere Web Client	87
Change IP Address and DNS Settings in the vSphere Web Client	88
Configuring Dependent Hardware iSCSI Adapters	88
Dependent Hardware iSCSI Considerations	89
View Dependent Hardware iSCSI Adapters in the vSphere Web Client	89
Determine Association Between iSCSI and Network Adapters in the vSphere Web Client	90
Configuring Software iSCSI Adapter	90
Activate the Software iSCSI Adapter in the vSphere Web Client	91
Disable Software iSCSI Adapter in the vSphere Web Client	92
Modify General Properties for iSCSI Adapters	92
Setting Up iSCSI Network	93
Guidelines for Using iSCSI Port Binding in ESXi	96
Create Network Connections for iSCSI in the vSphere Web Client	96
Managing iSCSI Network	101
iSCSI Network Troubleshooting	101
Using Jumbo Frames with iSCSI	102
Enable Jumbo Frames for Software and Dependent Hardware iSCSI	102
Enable Jumbo Frames for Independent Hardware iSCSI	103
Configuring Discovery Addresses for iSCSI Adapters	103
Set Up Dynamic Discovery in the vSphere Web Client	103
Set Up Static Discovery in the vSphere Web Client	104
Remove Static Targets in the vSphere Web Client	104
Configuring CHAP Parameters for iSCSI Adapters	105
Choosing CHAP Authentication Method	105
Set Up CHAP for iSCSI Adapter in the vSphere Web Client	106
Set Up CHAP for Target in the vSphere Web Client	107
Disable CHAP	109
Configuring Advanced Parameters for iSCSI	109
Configure Advanced Parameters for iSCSI in the vSphere Web Client	111
iSCSI Session Management	111
Review iSCSI Sessions	112
Add iSCSI Sessions	112
Remove iSCSI Sessions	113

## 12 Modifying iSCSI Storage Systems for ESXi 115

Testing ESXi iSCSI SAN Configurations	115
General Considerations for iSCSI SAN Storage Systems	116
EMC CLARiiON Storage Systems	117
EMC Symmetrix Storage Systems	117
Enable HP StorageWorks MSA1510i to Communicate with ESXi	118

HP StorageWorks EVA Storage Systems	119
NetApp Storage Systems	119
Dell EqualLogic Storage Systems	120
HP StorageWorks SAN/iQ Storage Systems	121
iSCSI Targets in vApps	121

### **13 Booting from iSCSI SAN 122**

General Boot from iSCSI SAN Recommendations	122
Prepare the iSCSI SAN	123
Configure Independent Hardware iSCSI Adapter for SAN Boot	124
Configure iSCSI Boot Settings	124
iBFT iSCSI Boot Overview	125
iBFT iSCSI Boot Considerations	126
Configuring iBFT Boot from SAN	126
Networking Best Practices	129
Change iBFT iSCSI Boot Settings	130
Troubleshooting iBFT iSCSI Boot	130

### **14 Best Practices for iSCSI Storage 132**

Preventing iSCSI SAN Problems	132
Optimizing iSCSI SAN Storage Performance	133
Storage System Performance	133
Server Performance with iSCSI	134
Network Performance	135
Checking Ethernet Switch Statistics	137
iSCSI SAN Configuration Checklist	137

### **15 Managing Storage Devices 139**

Understanding Storage Device Naming	139
Rename Storage Devices	140
Storage Refresh and Rescan Operations	141
Perform Storage Rescan in the vSphere Web Client	142
Perform Adapter Rescan in the vSphere Web Client	142
Change the Number of Scanned LUNs in the vSphere Web Client	142
Identifying Device Connectivity Problems	143
Detecting PDL Conditions	143
Performing Planned Storage Device Removal	144
Recovering From PDL Conditions	146
Handling Transient APD Conditions	146
Check the Connection Status of a Storage Device	148
PDL Conditions and High Availability	149

## 16 Working with Datastores 152

- Understanding VMFS Datastores 153
  - How VMFS5 Differs from VMFS3 153
  - VMFS Datastores and Storage Disk Formats 154
  - VMFS Datastores as Repositories 155
  - Sharing a VMFS Datastore Across Hosts 155
  - VMFS Metadata Updates 156
  - VMFS Locking Mechanisms 157
- Understanding NFS Datastores 158
  - NFS Datastores as Repositories for Commonly Used Files 158
  - Using Layer 3 Routed Connections to Access NFS Storage 159
- Creating Datastores 159
  - Create a VMFS Datastore in the vSphere Web Client 160
  - Create NFS Datastore in the vSphere Web Client 161
- Change Datastore Name in the vSphere Web Client 162
- Managing Duplicate VMFS Datastores 162
  - Keep Existing Datastore Signature in the vSphere Web Client 163
  - Resignature a VMFS Datastore Copy in the vSphere Web Client 163
- Upgrading VMFS Datastores 164
  - Upgrade a Datastore to VMFS5 in the vSphere Web Client 165
- Increasing VMFS Datastore Capacity 166
  - Increase VMFS Datastore Capacity in the vSphere Web Client 166
- Unmount VMFS or NFS Datastores 167
- Mount VMFS Datastores 168
- Remove VMFS Datastores in the vSphere Web Client 169
- Storage Filtering 169
  - Turn Off Storage Filters 170
- Set Up Dynamic Disk Mirroring 171
- Collecting Diagnostic Information for ESXi Hosts on a Storage Device 172
  - Set Up a Device Partition as Core Dump Location 172
  - Set Up a File as Core Dump Location 174
- Checking Metadata Consistency with VOMA 175
- Configuring VMFS Pointer Block Cache 176
  - Set Advanced Host Attributes in the vSphere Web Client 177

## 17 Raw Device Mapping 178

- About Raw Device Mapping 178
  - Benefits of Raw Device Mapping 179
  - RDM Considerations and Limitations 182
- Raw Device Mapping Characteristics 182
  - RDM Virtual and Physical Compatibility Modes 182



Dynamic Name Resolution	183
Raw Device Mapping with Virtual Machine Clusters	183
Comparing Available SCSI Device Access Modes	184
Create Virtual Machines with RDMs in the vSphere Web Client	184
Manage Paths for a Mapped LUN in the vSphere Web Client	186
<b>18 Working with Solid State Disks</b>	<b>187</b>
Benefits of SSD Enablement	188
Auto-Detection of SSD Devices	188
Tag Devices as SSD	188
Untag an SSD Device	189
Untag an Automatically Detected SSD Device	190
Tag Devices as Local	191
Monitor SSD Devices	192
Identify SSD Devices	193
Identifying a Virtual SSD Device	193
Best Practices for SSD Devices	194
Estimate Lifetime of Flash Devices	194
Swapping to Host Cache	195
Configure the Host Cache in the vSphere Web Client	195
Local SSDs Are Undetectable	196
<b>19 Working with Virtual SAN</b>	<b>197</b>
About Virtual SAN	197
Virtual SAN and Object-Based Storage	199
Characteristics of Virtual SAN	199
About Virtual SAN Datastores	200
Requirements for Virtual SAN	201
Virtual SAN Networking Requirements and Best Practices	202
Using Virtual SAN and vSphere HA	204
Enabling Virtual SAN	205
Create a Virtual SAN Cluster	206
Enable Virtual SAN on Existing Clusters	207
Assign a License Key to a Virtual SAN Cluster	208
View Virtual SAN Datastore	208
Disable Virtual SAN	209
Managing Disk Groups	209
Use Semi-Automatic Method to Claim Disks for Virtual SAN	210
Use Manual Method to Claim Disks for Virtual SAN	211
Add a Disk to the Disk Group	211
Remove Disks from Virtual SAN	212

Monitoring Virtual SAN	212
Monitor the Virtual SAN Cluster	213
Monitor Disks that Participate in Virtual SAN Datastores	213
Managing Virtual SAN	213
Expanding a Virtual SAN Cluster	214
Place a Member of Virtual SAN Cluster in Maintenance Mode	217
Virtual SAN and Storage Policy-Based Management	217
View Virtual SAN Storage Providers	219
Define a Virtual Machine Storage Policy for Virtual SAN	220
<b>20 About Virtual Machine Storage Policies</b>	<b>222</b>
Understanding Storage Capabilities	222
Assign Tags to Datastores	223
Understanding Virtual Machine Storage Policies	224
About Rule Sets	225
Enable Virtual Machine Storage Policies	226
Define a Storage Policy for a Virtual Machine	227
Add or Edit Tag-Based Rules	228
Edit a Storage Policy	228
Delete a Virtual Machine Storage Policy	229
Apply a Storage Policy to a Virtual Machine	229
Monitor Storage Compliance for Virtual Machines	230
Check Storage Compliance for a Virtual Machine	231
Find a Compatible Storage Resource for a Noncompliant Virtual Machine	232
Reapply Virtual Machine Storage Policy	233
<b>21 About VMware vSphere Flash Read Cache</b>	<b>234</b>
DRS Support for Flash Read Cache	235
vSphere High Availability Support for Flash Read Cache	235
Using Virtual Flash Resource to Configure a Swap Cache	235
Virtual Flash Volume Limits	236
Set Up Virtual Flash Resource	236
Set Up Multiple Virtual Flash Resources	236
Configure Host Swap Cache with Virtual Flash Resource	237
Configure Flash Read Cache for a Virtual Machine	237
Remove Virtual Flash Resource	238
Migrate Virtual Machines with Flash Read Cache	239
Using esxcli Commands with Virtual Flash	240
<b>22 VMkernel and Storage</b>	<b>241</b>
Storage APIs	242

**23 Understanding Multipathing and Failover 244**

- [Failover with Fibre Channel 245](#)
- [Host-Based Failover with iSCSI 245](#)
- [Array-Based Failover with iSCSI 247](#)
- [Path Failover and Virtual Machines 248](#)
  - [Set Timeout on Windows Guest OS 249](#)
- [Managing Multiple Paths 249](#)
- [VMware Multipathing Module 251](#)
  - [VMware SATPs 251](#)
  - [VMware PSPs 252](#)
  - [VMware NMP Flow of I/O 253](#)
- [Path Scanning and Claiming 253](#)
  - [Viewing the Paths Information 254](#)
  - [View Datastore Paths in the vSphere Web Client 255](#)
  - [View Storage Device Paths in the vSphere Web Client 255](#)
  - [Setting a Path Selection Policy 256](#)
  - [Change the Path Selection Policy in the vSphere Web Client 256](#)
  - [Disable Paths in the vSphere Web Client 257](#)
- [Managing Storage Paths and Multipathing Plug-Ins 257](#)
  - [Multipathing Considerations 257](#)
  - [List Multipathing Claim Rules for the Host 258](#)
  - [Display Multipathing Modules 260](#)
  - [Display SATPs for the Host 260](#)
  - [Display NMP Storage Devices 261](#)
  - [Add Multipathing Claim Rules 261](#)
  - [Delete Multipathing Claim Rules 263](#)
  - [Mask Paths 264](#)
  - [Unmask Paths 266](#)
  - [Define NMP SATP Rules 266](#)

**24 Storage Hardware Acceleration 269**

- [Hardware Acceleration Benefits 269](#)
- [Hardware Acceleration Requirements 270](#)
- [Hardware Acceleration Support Status 270](#)
- [Hardware Acceleration for Block Storage Devices 270](#)
  - [Disable Hardware Acceleration for Block Storage Devices in the vSphere Web Client 271](#)
  - [Managing Hardware Acceleration on Block Storage Devices 272](#)
- [Hardware Acceleration on NAS Devices 277](#)
  - [Install NAS Plug-In 277](#)
  - [Uninstall NAS Plug-Ins 278](#)
  - [Update NAS Plug-Ins 279](#)

Verify Hardware Acceleration Status for NAS	279
Hardware Acceleration Considerations	280
<b>25 Storage Thin Provisioning</b>	<b>281</b>
Storage Over-Subscription	281
Virtual Disk Thin Provisioning	281
About Virtual Disk Provisioning Policies	282
Create Thin Provisioned Virtual Disks in the vSphere Web Client	283
View Virtual Machine Storage Resources in the vSphere Web Client	284
Determine the Disk Format of a Virtual Machine in the vSphere Web Client	284
Inflate Thin Virtual Disks in the vSphere Web Client	285
Handling Datastore Over-Subscription	285
Array Thin Provisioning and VMFS Datastores	285
Space Usage Monitoring	286
Identify Thin-Provisioned Storage Devices	287
Reclaim Accumulated Storage Space	288
<b>26 Using Storage Providers</b>	<b>289</b>
Storage Providers and Storage Data Representation	290
Storage Provider Requirements and Considerations	290
Storage Status Reporting	291
Register Storage Providers	291
Securing Communication with Storage Providers	292
View Storage Provider Information in the vSphere Web Client	293
Unregister Storage Providers in the vSphere Web Client	293
Update Storage Providers in the vSphere Web Client	294
<b>27 Using vmkfstools</b>	<b>295</b>
vmkfstools Command Syntax	295
vmkfstools Options	296
-v Suboption	297
File System Options	297
Virtual Disk Options	299
Storage Device Options	306

# About vSphere Storage

*vSphere Storage* describes storage options available to VMware® ESXi and explains how to configure your ESXi system so that it can use and manage different types of storage. In addition, *vSphere Storage* explicitly concentrates on Fibre Channel and iSCSI storage area networks (SANs) as storage options and discusses specifics of using ESXi in Fibre Channel and iSCSI environments.

## Intended Audience

This information is for experienced system administrators who are familiar with virtual machine technology, datacenter operations, and SAN storage concepts.

# Updated Information

This *vSphere Storage* is updated with each release of the product or when necessary.

This table provides the update history of the *vSphere Storage*.

Revision	Description
15 AUG 2020	At VMware, we value inclusion. To foster this principle within our customer, partner, and internal community, we are replacing some of the terminology in our content. We have updated this guide to remove instances of non-inclusive language.
10 DEC 2018	Minor revisions.
20 MAR 2018	Minor revisions.
12 MAR 2018	Minor revisions.
EN-001523-05	<ul style="list-style-type: none"><li>■ <a href="#">Estimate Lifetime of Flash Devices</a> has been updated to include additional details.</li><li>■ <a href="#">Storage Filtering</a> has been updated to correct the value of the Same Hosts and Transports Filter. The correct value is <code>config.vpxd.filter.sameHostsAndTransportsFilter</code>.</li></ul>
EN-001523-04	<a href="#">Migrate Virtual Machines with Flash Read Cache</a> has been corrected to match the vSphere Web Client configuration options.
EN-001523-03	Minor revisions.
EN-001523-02	<ul style="list-style-type: none"><li>■ <a href="#">About Virtual SAN</a> and <a href="#">Requirements for Virtual SAN</a> have been updated to state more clearly that each disk group can include only one SSD and one or multiple HDDs.</li><li>■ <a href="#">Requirements for Virtual SAN</a> and <a href="#">Managing Disk Groups</a> have been updated to state more clearly that SSD capacity should be at least 10 percent of the total consumed HDD in each disk group.</li><li>■ <a href="#">Sharing a VMFS Datastore Across Hosts</a> now references <i>Configuration Maximums</i> as a reliable source of information about the maximum number of hosts that can connect to a single VMFS datastore.</li><li>■ <a href="#">Understanding NFS Datastores</a> now states that ESXi does not support IPv6 with NFS.</li></ul>
EN-001523-01	<ul style="list-style-type: none"><li>■ <a href="#">Network Configuration and Authentication</a> contained an incorrect requirement for physical network adapters used for software or dependent hardware iSCSI to be on the same subnet as an iSCSI target. This requirement is valid only when you use port binding. The statement has been removed from the topic.</li><li>■ <a href="#">Collecting Diagnostic Information for ESXi Hosts on a Storage Device</a> has been updated to clarify core dump file information.</li></ul>
EN-001523	Initial release.

# Introduction to Storage

# 1

This introduction describes storage options available in vSphere and explains how to configure your host so that it can use and manage different types of storage.

This chapter includes the following topics:

- [Storage Virtualization](#)
- [Types of Physical Storage](#)
- [Target and Device Representations](#)
- [Storage Device Characteristics](#)
- [Supported Storage Adapters](#)
- [Datastore Characteristics](#)
- [How Virtual Machines Access Storage](#)
- [Comparing Types of Storage](#)

## Storage Virtualization

ESXi provides host-level storage virtualization, which logically abstracts the physical storage layer from virtual machines.

An ESXi virtual machine uses a virtual disk to store its operating system, program files, and other data associated with its activities. A virtual disk is a large physical file, or a set of files, that can be copied, moved, archived, and backed up as easily as any other file. You can configure virtual machines with multiple virtual disks.

To access virtual disks, a virtual machine uses virtual SCSI controllers. These virtual controllers include BusLogic Parallel, LSI Logic Parallel, LSI Logic SAS, and VMware Paravirtual. These controllers are the only types of SCSI controllers that a virtual machine can see and access.

Each virtual disk resides on a datastore that is deployed on physical storage. From the standpoint of the virtual machine, each virtual disk appears as if it were a SCSI drive connected to a SCSI controller. Whether the actual physical storage is being accessed through storage or network adapters on the host is typically transparent to the guest operating system and to applications running on the virtual machine.

In addition to virtual disks, vSphere offers a mechanism called raw device mapping (RDM). RDM is useful when a guest operation system inside a virtual machine requires direct access to a storage device. For information about RDMs, see [Chapter 17 Raw Device Mapping](#).

## Types of Physical Storage

The ESXi storage management process starts with storage space that your storage administrator preallocates on different storage systems.

ESXi supports the following types of storage:

### Local Storage

Stores virtual machine files on internal or directly connected external storage disks.

### Networked Storage

Stores virtual machine files on external storage disks or arrays attached to your host through a direct connection or through a high-speed network.

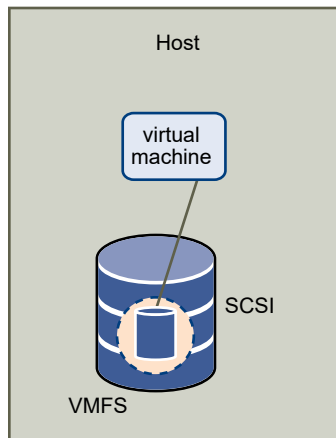
## Local Storage

Local storage can be internal hard disks located inside your ESXi host, or it can be external storage systems located outside and connected to the host directly through protocols such as SAS or SATA.

Local storage does not require a storage network to communicate with your host. You need a cable connected to the storage unit and, when required, a compatible HBA in your host.

The following illustration depicts a virtual machine using local SCSI storage.

Figure 1-1. Local Storage



In this example of a local storage topology, the host uses a single connection to a storage disk. On that disk, you can create a VMFS datastore, which you use to store virtual machine disk files.



Although this storage configuration is possible, it is not a recommended topology. Using single connections between storage arrays and hosts creates single points of failure (SPOF) that can cause interruptions when a connection becomes unreliable or fails. However, because the majority of local storage devices do not support multiple connections, you cannot use multiple paths to access local storage.

ESXi supports a variety of local storage devices, including SCSI, IDE, SATA, USB, and SAS storage systems. Regardless of the type of storage you use, your host hides a physical storage layer from virtual machines.

---

**Note** You cannot use IDE/ATA or USB drives to store virtual machines.

---

Local storage does not support sharing across multiple hosts. Only one host has access to a datastore on a local storage device. As a result, although you can use local storage to create virtual machines, it prevents you from using VMware features that require shared storage, such as HA and vMotion.

However, if you use a cluster of hosts that have just local storage devices, you can implement Virtual SAN or vSphere Storage Appliance. Both technologies transform local storage resources into software-defined shared storage and allow you to use features that require shared storage. For details, see [Chapter 19 Working with Virtual SAN](#) or the *vSphere Storage Appliance* documentation.

## Networked Storage

Networked storage consists of external storage systems that your ESXi host uses to store virtual machine files remotely. Typically, the host accesses these systems over a high-speed storage network.

Networked storage devices are shared. Datastores on networked storage devices can be accessed by multiple hosts concurrently. ESXi supports multiple networked storage technologies.

In addition to traditional networked storage that this topic covers, VMware supports virtualized shared storage technologies. These technologies transform the internal storage resources of your ESXi hosts into shared storage that provides such capabilities as High Availability and vMotion for virtual machines. For details, see [Chapter 19 Working with Virtual SAN](#) or the *vSphere Storage Appliance* documentation.

---

**Note** The same LUN cannot be presented to an ESXi host or multiple hosts through different storage protocols. To access the LUN, hosts must always use a single protocol, for example, either Fibre Channel only or iSCSI only.

---

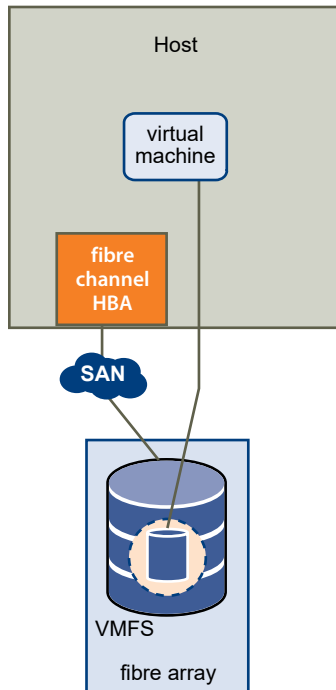
### Fibre Channel (FC)

Stores virtual machine files remotely on an FC storage area network (SAN). FC SAN is a specialized high-speed network that connects your hosts to high-performance storage devices. The network uses Fibre Channel protocol to transport SCSI traffic from virtual machines to the FC SAN devices.

To connect to the FC SAN, your host should be equipped with Fibre Channel host bus adapters (HBAs). Unless you use Fibre Channel direct connect storage, you need Fibre Channel switches to route storage traffic. If your host contains FCoE (Fibre Channel over Ethernet) adapters, you can connect to your shared Fibre Channel devices by using an Ethernet network.

Fibre Channel Storage depicts virtual machines using Fibre Channel storage.

**Figure 1-2. Fibre Channel Storage**



In this configuration, a host connects to a SAN fabric, which consists of Fibre Channel switches and storage arrays, using a Fibre Channel adapter. LUNs from a storage array become available to the host. You can access the LUNs and create datastores for your storage needs. The datastores use the VMFS format.

For specific information on setting up the Fibre Channel SAN, see [Chapter 3 Using ESXi with Fibre Channel SAN](#).

## Internet SCSI (iSCSI)

Stores virtual machine files on remote iSCSI storage devices. iSCSI packages SCSI storage traffic into the TCP/IP protocol so that it can travel through standard TCP/IP networks instead of the specialized FC network. With an iSCSI connection, your host serves as the initiator that communicates with a target, located in remote iSCSI storage systems.

ESXi offers the following types of iSCSI connections:

### Hardware iSCSI

Your host connects to storage through a third-party adapter capable of offloading the iSCSI and network processing. Hardware adapters can be dependent and independent.

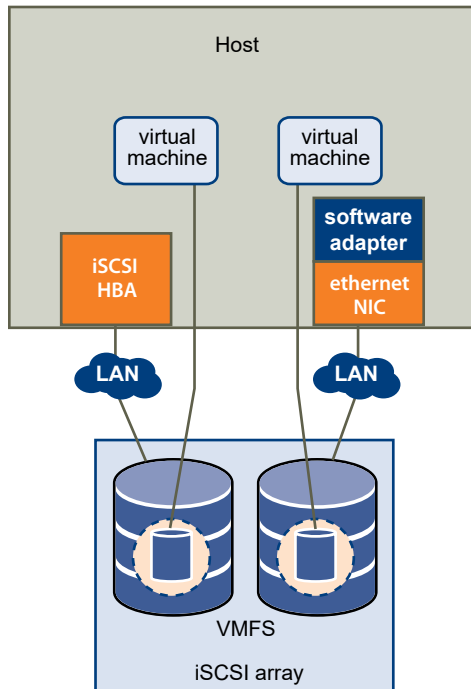
### Software iSCSI

Your host uses a software-based iSCSI initiator in the VMkernel to connect to storage. With this type of iSCSI connection, your host needs only a standard network adapter for network connectivity.

You must configure iSCSI initiators for the host to access and display iSCSI storage devices.

iSCSI Storage depicts different types of iSCSI initiators.

**Figure 1-3. iSCSI Storage**



In the left example, the host uses the hardware iSCSI adapter to connect to the iSCSI storage system.

In the right example, the host uses a software iSCSI adapter and an Ethernet NIC to connect to the iSCSI storage.

iSCSI storage devices from the storage system become available to the host. You can access the storage devices and create VMFS datastores for your storage needs.

For specific information on setting up the iSCSI SAN, see [Chapter 10 Using ESXi with iSCSI SAN](#).

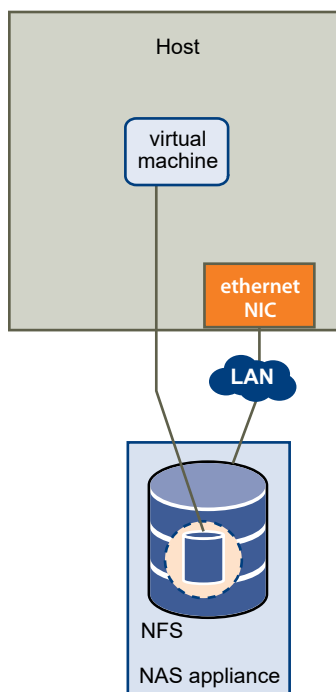
## Network-attached Storage (NAS)

Stores virtual machine files on remote file servers accessed over a standard TCP/IP network. The NFS client built into ESXi uses Network File System (NFS) protocol version 3 to communicate with the NAS/NFS servers. For network connectivity, the host requires a standard network adapter.

**Note** ESXi does not support the delegate user functionality that enables access to NFS volumes using non-root credentials.

NFS Storage depicts a virtual machine using the NFS volume to store its files. In this configuration, the host connects to the NFS server, which stores the virtual disk files, through a regular network adapter.

Figure 1-4. NFS Storage



## Shared Serial Attached SCSI (SAS)

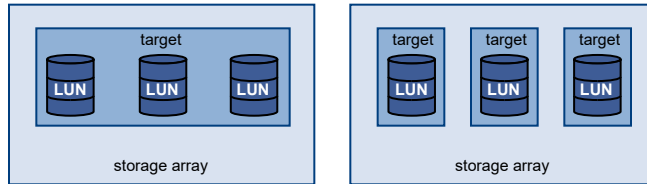
Stores virtual machines on direct-attached SAS storage systems that offer shared access to multiple hosts. This type of access permits multiple hosts to access the same VMFS datastore on a LUN.

## Target and Device Representations

In the ESXi context, the term target identifies a single storage unit that the host can access. The terms device and LUN describe a logical volume that represents storage space on a target. Typically, the terms device and LUN, in the ESXi context, mean a storage volume presented to the host from a storage target and available for formatting.

Different storage vendors present the storage systems to ESXi hosts in different ways. Some vendors present a single target with multiple storage devices or LUNs on it, while others present multiple targets with one LUN each.

**Figure 1-5. Target and LUN Representations**



In this illustration, three LUNs are available in each configuration. In one case, the host sees one target, but that target has three LUNs that can be used. Each LUN represents an individual storage volume. In the other example, the host sees three different targets, each having one LUN.

Targets that are accessed through the network have unique names that are provided by the storage systems. The iSCSI targets use iSCSI names, while Fibre Channel targets use World Wide Names (WWNs).

---

**Note** ESXi does not support accessing the same LUN through different transport protocols, such as iSCSI and Fibre Channel.

---

A device, or LUN, is identified by its UUID name. If a LUN is shared by multiple hosts, it must be presented to all hosts with the same UUID.

## Storage Device Characteristics

You can display all storage devices or LUNs available to the host, including all local and networked devices. If you use third-party multipathing plug-ins, the storage devices available through the plug-ins also appear on the list.

For each storage adapter, you can display a separate list of storage devices available for this adapter.

Generally, when you review storage devices, you see the following information.

**Table 1-1. Storage Device Information**

Storage Device Information	Description
Name	Also called Display Name. It is a name that the ESXi host assigns to the device based on the storage type and manufacturer. You can change this name to a name of your choice.
Identifier	A universally unique identifier that is intrinsic to the device.
Operational State	Indicates whether the device is mounted or unmounted. For details, see <a href="#">Detach Storage Devices</a> .
LUN	Logical Unit Number (LUN) within the SCSI target. The LUN number is provided by the storage system. If a target has only one LUN, the LUN number is always zero (0).

Table 1-1. Storage Device Information (continued)

Storage Device Information	Description
Type	Type of device, for example, disk or CD-ROM.
Drive Type	Information about whether the device is a solid-state drive (SSD) or a regular non-SSD hard drive. For details, see <a href="#">Chapter 18 Working with Solid State Disks</a> .
Transport	Transportation protocol your host uses to access the device.
Capacity	Total capacity of the storage device.
Owner	The plug-in, such as the NMP or a third-party plug-in, that the host uses to manage paths to the storage device. For details, see <a href="#">Managing Multiple Paths</a> .
Hardware Acceleration	Information about whether the storage device assists the host with virtual machine management operations. The status can be Supported, Not Supported, or Unknown. For details, see <a href="#">Chapter 24 Storage Hardware Acceleration</a> .
Location	A path to the storage device in the /vmfs/devices/ directory.
Partition Format	A partition scheme used by the storage device. It could be of a master boot record (MBR) or GUID partition table (GPT) format. The GPT devices can support datastores greater than 2TB. For more information, see <a href="#">VMFS Datastores and Storage Disk Formats</a> .
Partitions	Primary and logical partitions, including a VMFS datastore, if configured.
Multipathing Policies (VMFS datastores)	Path Selection Policy and Storage Array Type Policy the host uses to manage paths to storage. For more information, see <a href="#">Chapter 23 Understanding Multipathing and Failover</a> .
Paths (VMFS datastores)	Paths used to access storage and their status.

## Display Storage Devices for a Host in the vSphere Web Client

Display all storage devices available to a host. If you use any third-party multipathing plug-ins, the storage devices available through the plug-ins also appear on the list.

The Storage Devices view allows you to list the hosts' storage devices, analyze their information, and modify properties.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Devices**.

All storage devices available to the host are listed under Storage Devices.

- 4 To view details for a specific device, select the device from the list.

- 5 Use tabs under Device Details to access additional information and modify properties for the selected device.

Tab	Description
<b>Properties</b>	View device properties and characteristics. View and modify multipathing policies for the device.
<b>Paths</b>	Display paths available for the device. Disable or enable a selected path.

## Display Storage Devices for an Adapter in the vSphere Web Client

Display a list of storage devices accessible through a specific storage adapter on the host.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click Storage Adapters.  
All storage adapters installed on the host are listed under Storage Adapters.
- 4 Select the adapter from the list and click the **Devices** tab.  
Storage devices that the host can access through the adapter are displayed.

## Supported Storage Adapters

Storage adapters provide connectivity for your ESXi host to a specific storage unit or network.

ESXi supports different classes of adapters, including SCSI, iSCSI, RAID, Fibre Channel, Fibre Channel over Ethernet (FCoE), and Ethernet. ESXi accesses the adapters directly through device drivers in the VMkernel.

Depending on the type of storage you use, you might need to enable and configure a storage adapter on your host.

For information on setting up software FCoE adapters, see [Chapter 5 Configuring Fibre Channel over Ethernet](#).

For information on configuring different types of iSCSI adapters, see [Chapter 11 Configuring iSCSI Adapters and Storage](#).

## Storage Adapter Characteristics

The host uses storage adapters to access different storage devices. You can display details for the available storage adapters and review their information.

You must enable certain adapters, for example software iSCSI or FCoE, before you can view their information.

**Table 1-2. Storage Adapter Information**

<b>Adapter Information</b>	<b>Description</b>
Model	Model of the adapter.
Targets (Fibre Channel and SCSI)	Number of targets accessed through the adapter.
Connected Targets (iSCSI)	Number of connected targets on an iSCSI adapter.
WWN (Fibre Channel)	World Wide Name formed according to Fibre Channel standards that uniquely identifies the FC adapter.
iSCSI Name (iSCSI)	Unique name formed according to iSCSI standards that identifies the iSCSI adapter.
iSCSI Alias (iSCSI)	A friendly name used instead of the iSCSI name.
IP Address (independent hardware iSCSI)	Address assigned to the iSCSI HBA.
Devices	All storage devices or LUNs the adapter can access.
Paths	All paths the adapter uses to access storage devices.
Properties	Link that indicates that the adapter requires additional configuration. iSCSI and FCoE adapters display this link.

## View Storage Adapters Information in the vSphere Web Client

Display storage adapters that your host uses and review their information.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters**.
- 4 To view details for a specific adapter, select the adapter from the list.

## Datastore Characteristics

Datastores are logical containers, analogous to file systems, that hide specifics of each storage device and provide a uniform model for storing virtual machine files. You can display all datastores available to your hosts and analyze their properties.

Datastores are added to vCenter Server in the following ways:

- You can create VMFS or NFS datastores using the New Datastore wizard. A Virtual SAN datastore is automatically created when you enable Virtual SAN.
- When you add a host to vCenter Server, all datastores on the host are added to vCenter Server.

The following table describes datastore details that you can see when you review datastores. Certain characteristic might not be available or applicable to all types of datastores.



Table 1-3. Datastore Information

Datastore Information	Applicable Datastore Type	Description
Name	VMFS NFS Virtual SAN	Editable name that you assign to a datastore. For information on renaming a datastore, see <a href="#">Change Datastore Name in the vSphere Web Client</a> .
File System Type	VMFS NFS Virtual SAN	File system that the datastore uses. For information about VMFS and NFS datastores and how to manage them, see <a href="#">Chapter 16 Working with Datastores</a> . For information about Virtual SAN (vsan) datastores, see <a href="#">Chapter 19 Working with Virtual SAN</a> .
Device Backing	VMFS NFS Virtual SAN	Information about underlying storage, such as a storage device on which the datastore is deployed (VMFS), server and folder (NFS), or disk groups (Virtual SAN) .
Extents	VMFS	Individual extents that the datastore spans and their capacity.
Drive Type	VMFS NFS	Type of underlying storage device, a Solid State Drive (SSD) or a regular non-SSD hard drive. For details, see <a href="#">Chapter 18 Working with Solid State Disks</a> .
Capacity	VMFS NFS Virtual SAN	Includes total capacity, provisioned space, and free space.
Location (Mount Point)	VMFS NFS Virtual SAN	A path to the datastore in the /vmfs/volumes/ directory.
Server	NFS	Name or IP address of a NAS server.
Folder	NFS	Name of a mounted folder.
Datastore Capabilities	VMFS <b>Note</b> A multi-extent VMFS datastore assumes capabilities of only one of its extents. NFS Virtual SAN	Storage capabilities reported by physical storage and inherited by the datastores. You cannot modify them.
Storage I/O Control	VMFS NFS	Information on whether cluster-wide storage I/O prioritization is enabled. See the <i>vSphere Resource Management</i> documentation.
Hardware Acceleration	VMFS NFS Virtual SAN	Information on whether the underlying storage device supports hardware acceleration. The status can be Supported, Not Supported, or Unknown. For details, see <a href="#">Chapter 24 Storage Hardware Acceleration</a> .
Tags	VMFS NFS Virtual SAN	Datastore capabilities that you define and associate with datastores in a form of tags. For information, see <a href="#">Understanding Storage Capabilities</a> .

Table 1-3. Datastore Information (continued)

Datastore Information	Applicable Datastore Type	Description
Connectivity	VMFS NFS	Hosts where the datastore is mounted.
Multipathing Policies	VMFS	Path selection policy the host uses to access storage. For more information, see <a href="#">Chapter 23 Understanding Multipathing and Failover</a> .
Paths	VMFS	Number of paths used to access storage and their status.






## Display Datastore Information in the vSphere Web Client

Access the Datastores view with the vSphere Web Client navigator. The Datastores view that you display through the navigator lets you list all datastores available in the vSphere infrastructure inventory, analyze the information, and modify properties. You can also use the view to create datastores.

To list datastores for a particular parent object, such as a datacenter, cluster, or host, see [List Datastores for an Infrastructure Object](#).

### Procedure

- 1 From the vSphere Web Client Home, click **vCenter**.
- 2 Under Inventory Lists, click the **Datastores** category.  
Datastores that are available in the inventory appear in the center Datastores panel.
- 3 Use the icons to create a datastore or to perform basic tasks for a selected datastore.

Icon	Description
	Create a datastore.
	Increase datastore capacity.
	Mount a datastore to certain hosts.
	Remove a datastore.
	Unmount a datastore from certain hosts.

- 4 To view specific datastore details, double-click a selected datastore.
- 5 Use tabs to access additional information and modify datastore properties.

Tab	Description
<b>Getting Started</b>	View introductory information and access basic actions.
<b>Summary</b>	View statistics and configuration for the selected datastore.

Tab	Description
<b>Monitor</b>	View alarms, performance data, resource allocation, events, and other status information for the datastore.
<b>Manage</b>	View and modify datastore properties, alarm definitions, tags, and permissions. Use this tab to access storage devices that back the datastore, and to view and edit multipathing details for the datastore devices.
<b>Related Objects</b>	View objects related to the datastore. The objects include virtual machines that reside on the datastore and hosts where the datastore is mounted.

## List Datastores for an Infrastructure Object

Display datastores for a specific parent object, such as a datacenter, cluster, or host.






### Procedure

- 1 From the vSphere Web Client Home, click **vCenter**.
- 2 Use the vSphere Web Client object navigator to browse to an object that is a valid parent object of a datastore, such as a datacenter, cluster, or host.

- 3 Click the **Related Objects** tab and click **Datastores**.

If any datastores are configured for this object, they appear in the center Datastores panel.

- 4 Use the icons to create a datastore or to perform basic tasks for a selected datastore.

Icon	Description
	Create a datastore.
	Increase datastore capacity.
	Mount a datastore to certain hosts.
	Remove a datastore.
	Unmount a datastore from certain hosts.

- 5 Use tabs to access additional information and modify datastore properties.

Tab	Description
<b>Getting Started</b>	View introductory information and access basic actions.
<b>Summary</b>	View statistics and configuration for the selected datastore.
<b>Monitor</b>	View alarms, performance data, resource allocation, events, and other status information for the datastore.
<b>Manage</b>	View and modify datastore properties, alarm definitions, tags, and permissions. Use this tab to access storage devices that back the datastore, and to view and edit multipathing details for the datastore devices.
<b>Related Objects</b>	View objects related to the datastore. The objects include virtual machines that reside on the datastore and hosts where the datastore is mounted.

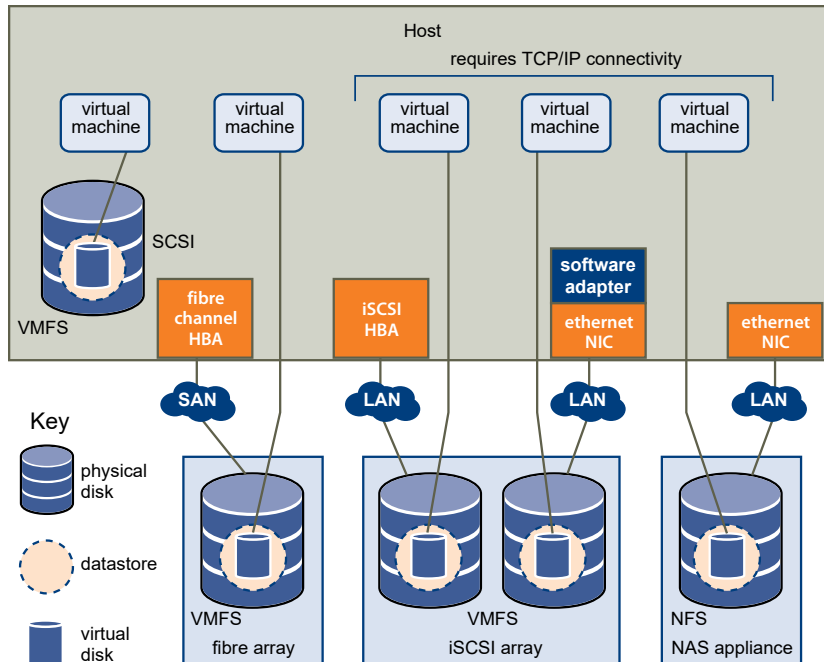
## How Virtual Machines Access Storage

When a virtual machine communicates with its virtual disk stored on a datastore, it issues SCSI commands. Because datastores can exist on various types of physical storage, these commands are encapsulated into other forms, depending on the protocol that the ESXi host uses to connect to a storage device.

ESXi supports Fibre Channel (FC), Internet SCSI (iSCSI), Fibre Channel over Ethernet (FCoE), and NFS protocols. Regardless of the type of storage device your host uses, the virtual disk always appears to the virtual machine as a mounted SCSI device. The virtual disk hides a physical storage layer from the virtual machine's operating system. This allows you to run operating systems that are not certified for specific storage equipment, such as SAN, inside the virtual machine.

The following graphic depicts five virtual machines using different types of storage to illustrate the differences between each type.

**Figure 1-6. Virtual machines accessing different types of storage**



**Note** This diagram is for conceptual purposes only. It is not a recommended configuration.

You can use maps on the **Storage Views** tab to visually represent and understand the relationships between virtual machines on your host and all available virtual and physical storage resources. For more information, see the *vSphere Monitoring and Performance* documentation.

## Comparing Types of Storage

Whether certain vSphere functionality is supported might depend on the storage technology that you use.

The following table compares networked storage technologies that ESXi supports.

**Table 1-4. Networked Storage that ESXi Supports**

Technology	Protocols	Transfers	Interface
Fibre Channel	FC/SCSI	Block access of data/LUN	FC HBA
Fibre Channel over Ethernet	FCoE/SCSI	Block access of data/LUN	<ul style="list-style-type: none"> <li>■ Converged Network Adapter (hardware FCoE)</li> <li>■ NIC with FCoE support (software FCoE)</li> </ul>
iSCSI	IP/SCSI	Block access of data/LUN	<ul style="list-style-type: none"> <li>■ iSCSI HBA or iSCSI-enabled NIC (hardware iSCSI)</li> <li>■ Network adapter (software iSCSI)</li> </ul>
NAS	IP/NFS	File (no direct LUN access)	Network adapter

The following table compares the vSphere features that different types of storage support.

**Table 1-5. vSphere Features Supported by Storage**

Storage Type	Boot VM	vMotion	Datastore	RDM	VM Cluster	VMware HA and DRS	Storage APIs - Data Protection
Local Storage	Yes	No	VMFS	No	Yes	No	Yes
Fibre Channel	Yes	Yes	VMFS	Yes	Yes	Yes	Yes
iSCSI	Yes	Yes	VMFS	Yes	Yes	Yes	Yes
NAS over NFS	Yes	Yes	NFS	No	No	Yes	Yes

**Note** Local storage supports a cluster of virtual machines on a single host (also known as a cluster in a box). A shared virtual disk is required. For more information about this configuration, see the *vSphere Resource Management* documentation.

# Overview of Using ESXi with a SAN

## 2

Using ESXi with a SAN improves flexibility, efficiency, and reliability. Using ESXi with a SAN also supports centralized management, failover, and load balancing technologies.

The following are benefits of using ESXi with a SAN:

- You can store data securely and configure multiple paths to your storage, eliminating a single point of failure.
- Using a SAN with ESXi systems extends failure resistance to the server. When you use SAN storage, all applications can instantly be restarted on another host after the failure of the original host.
- You can perform live migration of virtual machines using VMware vMotion.
- Use VMware High Availability (HA) in conjunction with a SAN to restart virtual machines in their last known state on a different server if their host fails.
- Use VMware Fault Tolerance (FT) to replicate protected virtual machines on two different hosts. Virtual machines continue to function without interruption on the secondary host if the primary one fails.
- Use VMware Distributed Resource Scheduler (DRS) to migrate virtual machines from one host to another for load balancing. Because storage is on a shared SAN array, applications continue running seamlessly.
- If you use VMware DRS clusters, put an ESXi host into maintenance mode to have the system migrate all running virtual machines to other ESXi hosts. You can then perform upgrades or other maintenance operations on the original host.

The portability and encapsulation of VMware virtual machines complements the shared nature of this storage. When virtual machines are located on SAN-based storage, you can quickly shut down a virtual machine on one server and power it up on another server, or suspend it on one server and resume operation on another server on the same network. This ability allows you to migrate computing resources while maintaining consistent shared access.

This chapter includes the following topics:

- [ESXi and SAN Use Cases](#)
- [Specifics of Using SAN Storage with ESXi](#)

- [ESXi Hosts and Multiple Storage Arrays](#)
- [Making LUN Decisions](#)
- [Choosing Virtual Machine Locations](#)
- [Layered Applications](#)
- [Third-Party Management Applications](#)
- [SAN Storage Backup Considerations](#)

## ESXi and SAN Use Cases

When used with a SAN, ESXi can benefit from multiple vSphere features, including Storage vMotion, Distributed Resource Scheduler (DRS), High Availability, and so on.

Using ESXi in conjunction with a SAN is effective for the following tasks:

### **Storage consolidation and simplification of storage layout**

If you are working with multiple hosts, and each host is running multiple virtual machines, the storage on the hosts is no longer sufficient and external storage is required. Choose a SAN for external storage to provide a simpler system architecture along with other benefits.

### **Maintenance with zero downtime**

When performing ESXi host or infrastructure maintenance, use vMotion to migrate virtual machines to other host. If shared storage is on the SAN, you can perform maintenance without interruptions to the users of the virtual machines. Virtual machine working processes continue throughout a migration.

### **Load balancing**

You can add a host to a DRS cluster, and the host's resources become part of the cluster's resources. The distribution and usage of CPU and memory resources for all hosts and virtual machines in the cluster are continuously monitored. DRS compares these metrics to an ideal resource utilization. Ideal utilization takes into account the attributes of the cluster's resource pools and virtual machines, the current demand, and the imbalance target. It then performs (or recommends) virtual machine migrations accordingly.

### **Disaster recovery**

You can use VMware High Availability to configure multiple ESXi hosts as a cluster to provide rapid recovery from outages and cost-effective high availability for applications running in virtual machines.

### **Simplified array migrations and storage upgrades**

When you purchase new storage systems or arrays, use Storage vMotion to perform live automated migration of virtual machine disk files from existing storage to their new destination without interruptions to the users of the virtual machines.

## Specifics of Using SAN Storage with ESXi

Using a SAN in conjunction with an ESXi host differs from traditional SAN usage in a variety of ways.

When you use SAN storage with ESXi, keep in mind the following considerations:

- You cannot use SAN administration tools to directly access operating systems of virtual machines that use the storage. With traditional tools, you can monitor only the VMware ESXi operating system. You use the vSphere Web Client to monitor virtual machines.
- The HBA visible to the SAN administration tools is part of the ESXi system, not part of the virtual machine.
- Typically, your ESXi system performs multipathing for you.

## ESXi Hosts and Multiple Storage Arrays

An ESXi host can access storage devices presented from multiple storage arrays, including arrays from different vendors.

When you use multiple arrays from different vendors, the following considerations apply:

- If your host uses the same Storage Array Type Plugin (SATP) for multiple arrays, be careful when you need to change the default Path Selection Policy (PSP) for that SATP. The change will apply to all arrays. For information on SATPs and PSPs, see [Chapter 23 Understanding Multipathing and Failover](#).
- Some storage arrays make recommendations on queue depth and other settings. Typically, these settings are configured globally at the ESXi host level. Making a change for one array impacts other arrays that present LUNs to the host. For information on changing queue depth, see the VMware knowledge base article at <http://kb.vmware.com/kb/1267>.
- Use single-initiator-single-target zoning when zoning ESXi hosts to Fibre Channel arrays. With this type of configuration, fabric related events that occur on one array do not impact other arrays. For more information about zoning, see [Using Zoning with Fibre Channel SANs](#).

## Making LUN Decisions

You must plan how to set up storage for your ESXi systems before you format LUNs with VMFS datastores.

When you make your LUN decision, keep in mind the following considerations:

- Each LUN should have the correct RAID level and storage characteristic for the applications running in virtual machines that use the LUN.
- Each LUN must contain only one VMFS datastore.
- If multiple virtual machines access the same VMFS, use disk shares to prioritize virtual machines.



You might want fewer, larger LUNs for the following reasons:

- More flexibility to create virtual machines without asking the storage administrator for more space.
- More flexibility for resizing virtual disks, doing snapshots, and so on.
- Fewer VMFS datastores to manage.

You might want more, smaller LUNs for the following reasons:

- Less wasted storage space.
- Different applications might need different RAID characteristics.
- More flexibility, as the multipathing policy and disk shares are set per LUN.
- Use of Microsoft Cluster Service requires that each cluster disk resource is in its own LUN.
- Better performance because there is less contention for a single volume.

When the storage characterization for a virtual machine is not available, there is often no simple method to determine the number and size of LUNs to provision. You can experiment using either a predictive or adaptive scheme.

## Use the Predictive Scheme to Make LUN Decisions

When setting up storage for ESXi systems, before creating VMFS datastores, you must decide on the size and number of LUNs to provision. You can experiment using the predictive scheme.

### Procedure

- 1 Provision several LUNs with different storage characteristics.
- 2 Create a VMFS datastore on each LUN, labeling each datastore according to its characteristics.
- 3 Create virtual disks to contain the data for virtual machine applications in the VMFS datastores created on LUNs with the appropriate RAID level for the applications' requirements.
- 4 Use disk shares to distinguish high-priority from low-priority virtual machines.

---

**Note** Disk shares are relevant only within a given host. The shares assigned to virtual machines on one host have no effect on virtual machines on other hosts.

---

- 5 Run the applications to determine whether virtual machine performance is acceptable.

## Use the Adaptive Scheme to Make LUN Decisions

When setting up storage for ESXi hosts, before creating VMFS datastores, you must decide on the number and size of LUNs to provision. You can experiment using the adaptive scheme.

### Procedure

- 1 Provision a large LUN (RAID 1+0 or RAID 5), with write caching enabled.
- 2 Create a VMFS on that LUN.
- 3 Create four or five virtual disks on the VMFS.
- 4 Run the applications to determine whether disk performance is acceptable.

### Results

If performance is acceptable, you can place additional virtual disks on the VMFS. If performance is not acceptable, create a new, large LUN, possibly with a different RAID level, and repeat the process. Use migration so that you do not lose virtual machines data when you recreate the LUN.

## Choosing Virtual Machine Locations

When you're working on optimizing performance for your virtual machines, storage location is an important factor. A trade-off always exists between expensive storage that offers high performance and high availability and storage with lower cost and lower performance.

Storage can be divided into different tiers depending on a number of factors:

- High Tier. Offers high performance and high availability. Might offer built-in snapshots to facilitate backups and point-in-time (PiT) restorations. Supports replication, full storage processor redundancy, and SAS drives. Uses high-cost spindles.
- Mid Tier. Offers mid-range performance, lower availability, some storage processor redundancy, and SCSI or SAS drives. May offer snapshots. Uses medium-cost spindles.
- Lower Tier. Offers low performance, little internal storage redundancy. Uses low end SCSI drives or SATA (serial low-cost spindles).

Not all applications need to be on the highest-performance, most-available storage—at least not throughout their entire life cycle.

---

**Note** If you need some of the functionality of the high tier, such as snapshots, but do not want to pay for it, you might be able to achieve some of the high-performance characteristics in software. For example, you can create snapshots in software.

---

When you decide where to place a virtual machine, ask yourself these questions:

- How critical is the virtual machine?
- What are its performance and availability requirements?
- What are its PiT restoration requirements?
- What are its backup requirements?
- What are its replication requirements?

A virtual machine might change tiers throughout its life cycle because of changes in criticality or changes in technology that push higher-tier features to a lower tier. Criticality is relative and might change for a variety of reasons, including changes in the organization, operational processes, regulatory requirements, disaster planning, and so on.

## Layered Applications

SAN administrators customarily use specialized array-based software for backup, disaster recovery, data mining, forensics, and configuration testing.

Storage providers typically supply two types of advanced services for their LUNs: snapshotting and replication.

- Snapshotting creates space with efficient copies of LUNs that share common blocks of data. In general, snapshotting is used locally on the same storage systems as the primary LUN for quick backups, application testing, forensics, or data mining.
- Replication creates full copies of LUNs. Replicas are usually made to separate storage systems, possibly separate sites to protect against major outages that incapacitate or destroy an entire array or site.

When you use an ESXi system in conjunction with a SAN, you must decide whether array-based or host-based tools are more suitable for your particular situation.

## Array-Based (Third-Party) Solution

When you use an ESXi system in conjunction with a SAN, you must decide whether array-based tools are more suitable for your particular situation.

When you consider an array-based solution, keep in mind the following points:

- Array-based solutions usually result in more comprehensive statistics. With RDMs, data always takes the same path, which results in easier performance management.
- Security is more transparent to the storage administrator when you use an RDM and an array-based solution because with RDMs, virtual machines more closely resemble physical machines.
- If you use an array-based solution, physical compatibility RDMs are often used for the storage of virtual machines. If you do not intend to use RDMs, check the storage vendor documentation to see if operations on LUNs with VMFS volumes are supported. If you use array operations on VMFS LUNs, carefully read the section on resignaturing.

## File-Based (VMFS) Solution

When you use an ESXi system in conjunction with a SAN, you must decide whether file-based tools are more suitable for your particular situation.

When you consider a file-based solution that uses VMware tools and VMFS instead of the array tools, be aware of the following points:

- Using VMware tools and VMFS is better for provisioning. One large LUN is allocated and multiple .vmdk files can be placed on that LUN. With an RDM, a new LUN is required for each virtual machine.
- Snapshotting is included with your ESXi host at no extra cost.
- Using VMFS is easier for ESXi administrators.
- ESXi administrators who use the file-based solution are more independent from the SAN administrator.

## Third-Party Management Applications

You can use third-party management applications in conjunction with your ESXi host.

Most SAN hardware is packaged with storage management software. In many cases, this software is a web application that can be used with any web browser connected to your network. In other cases, this software typically runs on the storage system or on a single server, independent of the servers that use the SAN for storage.

Use this third-party management software for the following tasks:

- Storage array management, including LUN creation, array cache management, LUN mapping, and LUN security.
- Setting up replication, check points, snapshots, or mirroring.

If you decide to run the SAN management software on a virtual machine, you gain the benefits of running a virtual machine, including failover using vMotion and VMware HA. Because of the additional level of indirection, however, the management software might not be able to see the SAN. In this case, you can use an RDM.

---

**Note** Whether a virtual machine can run management software successfully depends on the particular storage system.

---

## SAN Storage Backup Considerations

Having a proper backup strategy is one of the most important aspects of SAN management. In the SAN environment, backups have two goals. The first goal is to archive online data to offline media. This process is repeated periodically for all online data on a time schedule. The second goal is to provide access to offline data for recovery from a problem. For example, database recovery often requires retrieval of archived log files that are not currently online.

Scheduling a backup depends on a number of factors:

- Identification of critical applications that require more frequent backup cycles within a given period of time.

- Recovery point and recovery time goals. Consider how precise your recovery point needs to be, and how long you are willing to wait for it.
- The rate of change (RoC) associated with the data. For example, if you are using synchronous/asynchronous replication, the RoC affects the amount of bandwidth required between the primary and secondary storage devices.
- Overall impact on SAN environment, storage performance (while backing up), and other applications.
- Identification of peak traffic periods on the SAN (backups scheduled during those peak periods can slow the applications and the backup process).
- Time to schedule all backups within the datacenter.
- Time it takes to back up an individual application.
- Resource availability for archiving data; usually offline media access (tape).

Include a recovery-time objective for each application when you design your backup strategy. That is, consider the time and resources necessary to perform a backup. For example, if a scheduled backup stores so much data that recovery requires a considerable amount of time, examine the scheduled backup. Perform the backup more frequently, so that less data is backed up at a time and the recovery time decreases.

If a particular application requires recovery within a certain time frame, the backup process needs to provide a time schedule and specific data processing to meet this requirement. Fast recovery can require the use of recovery volumes that reside on online storage to minimize or eliminate the need to access slow offline media for missing data components.

## Using Third-Party Backup Packages

You can use third-party backup solutions to protect system, application, and user data in your virtual machines.

VMware offers the Storage APIs - Data Protection to work in conjunction with third-party products. When using the APIs, third-party software can perform backups without loading ESXi hosts with the processing of backup tasks.

The third-party products using the Storage APIs - Data Protection can perform the following backup tasks:

- Perform full, differential, and incremental image backup and restore of virtual machines.
- Perform file-level backup of virtual machines that use supported Windows and Linux operating systems.
- Ensure data consistency by using Microsoft Volume Shadow Copy Services (VSS) for virtual machines that run supported Microsoft Windows operating systems.

Because the Storage APIs - Data Protection leverage the snapshot capabilities of VMFS, backups that you can perform do not require downtime for virtual machines. These backups are nondisruptive, can be performed at any time, and do not need extended backup windows.

For information about the Storage APIs - Data Protection and integration with backup products, see the VMware Web site or contact your backup vendor.

# Using ESXi with Fibre Channel SAN

# 3

When you set up ESXi hosts to use FC SAN storage arrays, special considerations are necessary. This section provides introductory information about how to use ESXi with a FC SAN array.

This chapter includes the following topics:

- [Fibre Channel SAN Concepts](#)
- [Using Zoning with Fibre Channel SANs](#)
- [How Virtual Machines Access Data on a Fibre Channel SAN](#)

## Fibre Channel SAN Concepts

If you are an ESXi administrator planning to set up hosts to work with SANs, you must have a working knowledge of SAN concepts. You can find information about SANs in print and on the Internet. Because this industry changes constantly, check these resources frequently.

If you are new to SAN technology, familiarize yourself with the basic terminology.

A storage area network (SAN) is a specialized high-speed network that connects computer systems, or host servers, to high performance storage subsystems. The SAN components include host bus adapters (HBAs) in the host servers, switches that help route storage traffic, cables, storage processors (SPs), and storage disk arrays.

A SAN topology with at least one switch present on the network forms a SAN fabric.

To transfer traffic from host servers to shared storage, the SAN uses the Fibre Channel (FC) protocol that packages SCSI commands into Fibre Channel frames.

To restrict server access to storage arrays not allocated to that server, the SAN uses zoning. Typically, zones are created for each group of servers that access a shared group of storage devices and LUNs. Zones define which HBAs can connect to which SPs. Devices outside a zone are not visible to the devices inside the zone.

Zoning is similar to LUN masking, which is commonly used for permission management. LUN masking is a process that makes a LUN available to some hosts and unavailable to other hosts.

When transferring data between the host server and storage, the SAN uses a technique known as multipathing. Multipathing allows you to have more than one physical path from the ESXi host to a LUN on a storage system.

Generally, a single path from a host to a LUN consists of an HBA, switch ports, connecting cables, and the storage controller port. If any component of the path fails, the host selects another available path for I/O. The process of detecting a failed path and switching to another is called path failover.

## Ports in Fibre Channel SAN

In the context of this document, a port is the connection from a device into the SAN. Each node in the SAN, such as a host, a storage device, or a fabric component has one or more ports that connect it to the SAN. Ports are identified in a number of ways.

### WWPN (World Wide Port Name)

A globally unique identifier for a port that allows certain applications to access the port. The FC switches discover the WWPN of a device or host and assign a port address to the device.

### Port\_ID (or port address)

Within a SAN, each port has a unique port ID that serves as the FC address for the port. This unique ID enables routing of data through the SAN to that port. The FC switches assign the port ID when the device logs in to the fabric. The port ID is valid only while the device is logged on.

When N-Port ID Virtualization (NPIV) is used, a single FC HBA port (N-port) can register with the fabric by using several WWPNs. This method allows an N-port to claim multiple fabric addresses, each of which appears as a unique entity. When ESXi hosts use a SAN, these multiple, unique identifiers allow the assignment of WWNs to individual virtual machines as part of their configuration.

## Fibre Channel Storage Array Types

ESXi supports different storage systems and arrays.

The types of storage that your host supports include active-active, active-passive, and ALUA-compliant.

### Active-active storage system

Allows access to the LUNs simultaneously through all the storage ports that are available without significant performance degradation. All the paths are active at all times, unless a path fails.

### Active-passive storage system

A system in which one storage processor is actively providing access to a given LUN. The other processors act as backup for the LUN and can be actively providing access to other LUN I/O. I/O can be successfully sent only to an active port for a given LUN. If access through



the active storage port fails, one of the passive storage processors can be activated by the servers accessing it.

### **Asymmetrical storage system**

Supports Asymmetric Logical Unit Access (ALUA). ALUA-complaint storage systems provide different levels of access per port. ALUA allows hosts to determine the states of target ports and prioritize paths. The host uses some of the active paths as primary while others as secondary.

## **Using Zoning with Fibre Channel SANs**

Zoning provides access control in the SAN topology. Zoning defines which HBAs can connect to which targets. When you configure a SAN by using zoning, the devices outside a zone are not visible to the devices inside the zone.

Zoning has the following effects:

- Reduces the number of targets and LUNs presented to a host.
- Controls and isolates paths in a fabric.
- Can prevent non-ESXi systems from accessing a particular storage system, and from possibly destroying VMFS data.
- Can be used to separate different environments, for example, a test from a production environment.

With ESXi hosts, use a single-initiator zoning or a single-initiator-single-target zoning. The latter is a preferred zoning practice. Using the more restrictive zoning prevents problems and misconfigurations that can occur on the SAN.

For detailed instructions and best zoning practices, contact storage array or switch vendors.

## **How Virtual Machines Access Data on a Fibre Channel SAN**

ESXi stores a virtual machine's disk files within a VMFS datastore that resides on a SAN storage device. When virtual machine guest operating systems issue SCSI commands to their virtual disks, the SCSI virtualization layer translates these commands to VMFS file operations.

When a virtual machine interacts with its virtual disk stored on a SAN, the following process takes place:

- 1 When the guest operating system in a virtual machine reads or writes to a SCSI disk, it issues SCSI commands to the virtual disk.
- 2 Device drivers in the virtual machine's operating system communicate with the virtual SCSI controllers.
- 3 The virtual SCSI controller forwards the command to the VMkernel.

- 4 The VMkernel performs the following tasks.
  - a Locates the file in the VMFS volume that corresponds to the guest virtual machine disk.
  - b Maps the requests for the blocks on the virtual disk to blocks on the appropriate physical device.
  - c Sends the modified I/O request from the device driver in the VMkernel to the physical HBA.
- 5 The physical HBA performs the following tasks.
  - a Packages the I/O request according to the rules of the FC protocol.
  - b Transmits the request to the SAN.
- 6 Depending on a port the HBA uses to connect to the fabric, one of the SAN switches receives the request and routes it to the storage device that the host wants to access.

# Configuring Fibre Channel Storage

# 4

When you use ESXi systems with SAN storage, specific hardware and system requirements exist.

This chapter includes the following topics:

- [ESXi Fibre Channel SAN Requirements](#)
- [Installation and Setup Steps](#)
- [N-Port ID Virtualization](#)

## ESXi Fibre Channel SAN Requirements

In preparation for configuring your SAN and setting up your ESXi system to use SAN storage, review the requirements and recommendations.

- Make sure that the SAN storage hardware and firmware combinations you use are supported in conjunction with ESXi systems. For an up-to-date list, see the *vSphere Compatibility Guide*.
- Configure your system to have only one VMFS volume per LUN.
- Unless you are using diskless servers, do not set up the diagnostic partition on a SAN LUN.  
In the case of diskless servers that boot from a SAN, a shared diagnostic partition is appropriate.
- Use RDMs to access raw disks. For information, see [Chapter 17 Raw Device Mapping](#).
- For multipathing to work properly, each LUN must present the same LUN ID number to all ESXi hosts.
- Make sure the storage device driver specifies a large enough queue. You can set the queue depth for the physical HBA during system setup. For information on changing queue depth for HBAs and virtual machines, see the *vSphere Troubleshooting* documentation.
- On virtual machines running Microsoft Windows, increase the value of the SCSI TimeoutValue parameter to 60. This increase allows Windows to better tolerate delayed I/O resulting from path failover. For information, see [Set Timeout on Windows Guest OS](#).

## ESXi Fibre Channel SAN Restrictions

When you use ESXi with a SAN, certain restrictions apply.

- ESXi does not support FC connected tape devices.
- You cannot use multipathing software inside a virtual machine to perform I/O load balancing to a single physical LUN. However, when your Microsoft Windows virtual machine uses dynamic disks, this restriction does not apply. For information about configuring dynamic disks, see [Set Up Dynamic Disk Mirroring](#).

## Setting LUN Allocations

This topic provides general information about how to allocate LUNs when your ESXi works in conjunction with SAN.

When you set LUN allocations, be aware of the following points:

### Storage provisioning

To ensure that the ESXi system recognizes the LUNs at startup time, provision all LUNs to the appropriate HBAs before you connect the SAN to the ESXi system.

VMware recommends that you provision all LUNs to all ESXi HBAs at the same time. HBA failover works only if all HBAs see the same LUNs.

For LUNs that will be shared among multiple hosts, make sure that LUN IDs are consistent across all hosts. For example, LUN 5 should be mapped to host 1, host 2, and host 3 as LUN 5.

### vMotion and VMware DRS

When you use vCenter Server and vMotion or DRS, make sure that the LUNs for the virtual machines are provisioned to all ESXi hosts. This provides the most ability to move virtual machines.

### Active-active compared to active-passive arrays

When you use vMotion or DRS with an active-passive SAN storage device, make sure that all ESXi systems have consistent paths to all storage processors. Not doing so can cause path thrashing when a vMotion migration occurs.

For active-passive storage arrays not listed in Storage/SAN Compatibility, VMware does not support storage port failover. In those cases, you must connect the server to the active port on the storage array. This configuration ensures that the LUNs are presented to the ESXi host.

## Setting Fibre Channel HBAs

Typically, FC HBAs that you use on your ESXi host work correctly with the default configuration settings.

You should follow the configuration guidelines provided by your storage array vendor. During FC HBA setup, consider the following issues.

- Do not mix FC HBAs from different vendors in a single host. Having different models of the same HBA is supported, but a single LUN cannot be accessed through two different HBA types, only through the same type.
- Ensure that the firmware level on each HBA is the same.
- Set the timeout value for detecting a failover. To ensure optimal performance, do not change the default value.
- ESXi supports 16 GB end-to-end Fibre Channel connectivity.

## Installation and Setup Steps

This topic provides an overview of installation and setup steps that you need to follow when configuring your SAN environment to work with ESXi.

Follow these steps to configure your ESXi SAN environment.

- 1 Design your SAN if it is not already configured. Most existing SANs require only minor modification to work with ESXi.
- 2 Check that all SAN components meet requirements.
- 3 Perform any necessary storage array modification.

Most vendors have vendor-specific documentation for setting up a SAN to work with VMware ESXi.

- 4 Set up the HBAs for the hosts you have connected to the SAN.
- 5 Install ESXi on the hosts.
- 6 Create virtual machines and install guest operating systems.
- 7 (Optional) Set up your system for VMware HA failover or for using Microsoft Clustering Services.
- 8 Upgrade or modify your environment as needed.

## N-Port ID Virtualization

N-Port ID Virtualization (NPIV) is an ANSI T11 standard that describes how a single Fibre Channel HBA port can register with the fabric using several worldwide port names (WWPNs). This allows a fabric-attached N-port to claim multiple fabric addresses. Each address appears as a unique entity on the Fibre Channel fabric.

## How NPIV-Based LUN Access Works

NPIV enables a single FC HBA port to register several unique WWNs with the fabric, each of which can be assigned to an individual virtual machine.

SAN objects, such as switches, HBAs, storage devices, or virtual machines can be assigned World Wide Name (WWN) identifiers. WWNs uniquely identify such objects in the Fibre Channel fabric. When virtual machines have WWN assignments, they use them for all RDM traffic, so the LUNs pointed to by any of the RDMs on the virtual machine must not be masked against its WWNs. When virtual machines do not have WWN assignments, they access storage LUNs with the WWNs of their host's physical HBAs. By using NPIV, however, a SAN administrator can monitor and route storage access on a per virtual machine basis. The following section describes how this works.

When a virtual machine has a WWN assigned to it, the virtual machine's configuration file (.vmx) is updated to include a WWN pair (consisting of a World Wide Port Name, WWPN, and a World Wide Node Name, WWNN). As that virtual machine is powered on, the VMkernel instantiates a virtual port (VPORT) on the physical HBA which is used to access the LUN. The VPORT is a virtual HBA that appears to the FC fabric as a physical HBA, that is, it has its own unique identifier, the WWN pair that was assigned to the virtual machine. Each VPORT is specific to the virtual machine, and the VPORT is destroyed on the host and it no longer appears to the FC fabric when the virtual machine is powered off. When a virtual machine is migrated from one host to another, the VPORT is closed on the first host and opened on the destination host.

If NPIV is enabled, WWN pairs (WWPN & WWNN) are specified for each virtual machine at creation time. When a virtual machine using NPIV is powered on, it uses each of these WWN pairs in sequence to try to discover an access path to the storage. The number of VPORTs that are instantiated equals the number of physical HBAs present on the host. A VPORT is created on each physical HBA that a physical path is found on. Each physical path is used to determine the virtual path that will be used to access the LUN. Note that HBAs that are not NPIV-aware are skipped in this discovery process because VPORTs cannot be instantiated on them.

## Requirements for Using NPIV

If you plan to enable NPIV on your virtual machines, you should be aware of certain requirements.

The following requirements exist:

- NPIV can be used only for virtual machines with RDM disks. Virtual machines with regular virtual disks use the WWNs of the host's physical HBAs.
- HBAs on your host must support NPIV.

For information, see the *vSphere Compatibility Guide* and refer to your vendor documentation.

- Use HBAs of the same type, either all QLogic or all Emulex. VMware does not support heterogeneous HBAs on the same host accessing the same LUNs.
- If a host uses multiple physical HBAs as paths to the storage, zone all physical paths to the virtual machine. This is required to support multipathing even though only one path at a time will be active.
- Make sure that physical HBAs on the host have access to all LUNs that are to be accessed by NPIV-enabled virtual machines running on that host.

- The switches in the fabric must be NPIV-aware.
- When configuring a LUN for NPIV access at the storage level, make sure that the NPIV LUN number and NPIV target ID match the physical LUN and Target ID.

## NPIV Capabilities and Limitations

Learn about specific capabilities and limitations of the use of NPIV with ESXi.

ESXi with NPIV supports the following items:

- NPIV supports vMotion. When you use vMotion to migrate a virtual machine it retains the assigned WWN.  
  
If you migrate an NPIV-enabled virtual machine to a host that does not support NPIV, VMkernel reverts to using a physical HBA to route the I/O.
- If your FC SAN environment supports concurrent I/O on the disks from an active-active array, the concurrent I/O to two different NPIV ports is also supported.

When you use ESXi with NPIV, the following limitations apply:

- Because the NPIV technology is an extension to the FC protocol, it requires an FC switch and does not work on the direct attached FC disks.
- When you clone a virtual machine or template with a WWN assigned to it, the clones do not retain the WWN.
- NPIV does not support Storage vMotion.
- Disabling and then re-enabling the NPIV capability on an FC switch while virtual machines are running can cause an FC link to fail and I/O to stop.

## Assign WWNs to Virtual Machines

Assign WWN settings to virtual machine with an RDM disk.

You can create from 1 to 16 WWN pairs, which can be mapped to the first 1 to 16 physical FC HBAs on the host.

### Prerequisites

Create a virtual machine with an RDM disk. See [Create Virtual Machines with RDMs in the vSphere Web Client](#).

### Procedure

- 1 In the vSphere Web Client, browse to the virtual machine.
- 2 Right-click the virtual machine and select **Edit Settings**.
- 3 Click **VM Options**.
- 4 Click the Fibre Channel NPIV triangle to expand the NPIV options.
- 5 Deselect the **Temporarily Disable NPIV for this virtual machine** check box.

**6** Select **Generate new WWNs**.**7** Specify the number of WWNNs and WWPNS.

A minimum of 2 WWPNS are needed to support failover with NPIV. Typically only 1 WWNN is created for each virtual machine.

**Results**

The host creates WWN assignments for the virtual machine.

**What to do next**

Register newly created WWNs in the fabric so that the virtual machine is able to log in to the switch, and assign storage LUNs to the WWNs.

## Modify WWN Assignments

You can modify WWN assignments for a virtual machine with an RDM.

Typically, you do not need to change existing WWN assignments on your virtual machine. In certain circumstances, for example, when manually assigned WWNs are causing conflicts on the SAN, you might need to change or remove WWNs.

**Prerequisites**

Make sure to power off the virtual machine if you want to edit the existing WWNs.

Before you begin, ensure that your SAN administrator has provisioned the storage LUN ACL to allow the virtual machine's ESXi host to access it.

**Procedure**

- 1 Open the Virtual Machine Properties dialog box by clicking the **Edit Settings** link for the selected virtual machine.
- 2 Click the **Options** tab and select **Fibre Channel NPIV**.  
The Virtual Machine Properties dialog box opens.
- 3 Edit the WWN assignments by selecting one of the following options:

Option	Description
<b>Temporarily disable NPIV for this virtual machine</b>	Disable the WWN assignments for the virtual machine.
<b>Leave unchanged</b>	The existing WWN assignments are retained. The read-only WWN Assignments section of this dialog box displays the node and port values of any existing WWN assignments.



Option	Description
<b>Generate new WWNs</b>	New WWNs are generated and assigned to the virtual machine, overwriting any existing WWNs (those of the HBA itself are unaffected).
<b>Remove WWN assignment</b>	The WWNs assigned to the virtual machine are removed and it uses the HBA WWNs to access the storage LUN. This option is not available if you are creating a new virtual machine.

- 4 Click **OK** to save your changes.

# Configuring Fibre Channel over Ethernet

# 5

To access Fibre Channel storage, an ESXi host can use the Fibre Channel over Ethernet (FCoE) protocol.

The FCoE protocol encapsulates Fibre Channel frames into Ethernet frames. As a result, your host does not need special Fibre Channel links to connect to Fibre Channel storage, but can use 10Gbit lossless Ethernet to deliver Fibre Channel traffic.

This chapter includes the following topics:

- [Fibre Channel over Ethernet Adapters](#)
- [Configuration Guidelines for Software FCoE](#)
- [Set Up Networking for Software FCoE in the vSphere Web Client](#)
- [Add Software FCoE Adapters in the vSphere Web Client](#)

## Fibre Channel over Ethernet Adapters

To use Fibre Channel over Ethernet (FCoE), you need to install FCoE adapters on your host.

The adapters that VMware supports generally fall into two categories, hardware FCoE adapters and software FCoE adapters that use the native FCoE stack in ESXi.

### Hardware FCoE Adapters

This category includes completely offloaded specialized Converged Network Adapters (CNAs) that contain network and Fibre Channel functionalities on the same card.

When such adapter is installed, your host detects and can use both CNA components. In the client, the networking component appears as a standard network adapter (vmnic) and the Fibre Channel component as a FCoE adapter (vmhba). You do not need to configure the hardware FCoE adapter to be able to use it.

## Software FCoE Adapters

A software FCoE adapter uses the native FCoE protocol stack in ESXi for the protocol processing. The software FCoE adapter is used with a NIC that offers Data Center Bridging (DCB) and I/O offload capabilities. Intel X520 is an example of such NIC. For information on NICs supporting software FCoE, see the *vSphere Compatibility Guide*.

For the software FCoE adapter, you must properly configure networking and then activate the adapter.

---

**Note** The number of software FCoE adapters you activate corresponds to the number of physical NIC ports. ESXi supports a maximum of four software FCoE adapters on one host.

---

## Configuration Guidelines for Software FCoE

When setting up your network environment to work with ESXi software FCoE, follow the guidelines and best practices that VMware offers.

### Network Switch Guidelines

Follow these guidelines when you configure a network switch for software FCoE environment:

- On the ports that communicate with your ESXi host, disable the Spanning Tree Protocol (STP). Having the STP enabled might delay the FCoE Initialization Protocol (FIP) response at the switch and cause an all paths down (APD) condition.

The FIP is a protocol that FCoE uses to discover and initialize FCoE entities on the Ethernet.

- Turn on Priority-based Flow Control (PFC) and set it to AUTO.

VMware recommends that you use the following firmware on the FCoE switch:

- Cisco Nexus 5000: version 4.1(3)N2 or higher.
- Brocade FCoE switch: version 6.3.1 or higher.

### Network Adapter Best Practices

If you plan to enable software FCoE adapters to work with network adapters, specific considerations apply.

- Make sure that the latest microcode is installed on the FCoE network adapter.
- If the network adapter has multiple ports, when configuring networking, add each port to a separate vSwitch. This practice helps you to avoid an APD condition when a disruptive event, such as an MTU change, occurs.
- Do not move a network adapter port from one vSwitch to another when FCoE traffic is active. If you need to make this change, reboot your host afterwards.
- If you changed the vSwitch for a network adapter port and caused a failure, moving the port back to the original vSwitch resolves the problem.

# Set Up Networking for Software FCoE in the vSphere Web Client

Before you activate the software FCoE adapters, you need to create VMkernel network adapters for all physical FCoE NICs installed on your host.

This procedure explains how to create a single VMkernel network adapter connected to a single FCoE physical network adapter through a vSphere standard switch. If your host has multiple network adapters or multiple ports on the adapter, connect each FCoE NIC to a separate standard switch. For more information, see the *vSphere Networking* documentation.

---

**Note** ESXi supports the maximum of four network adapter ports used for software FCoE.

---

## Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click **Actions**, and select **All vCenter Actions > Add Networking**.
- 3 Select **VMkernel Network Adapter**, and click **Next**.
- 4 Select **New standard switch** to create a vSphere standard switch.
- 5 Under Unclaimed Adapters, select the network adapter (vmnic#) that supports FCoE and click **Assign**.

Make sure to assign the adapter to Active Adapters.

- 6 Enter a network label.

Network label is a friendly name that identifies the VMkernel adapter that you are creating, for example, FCoE.

- 7 Specify a VLAN ID and click **Next**.

Because FCoE traffic requires an isolated network, make sure that the VLAN ID you enter is different from the one used for regular networking on your host. For more information, see the *vSphere Networking* documentation.

- 8 After completing configuration, review the information and click **Finish**.

## Results

You have created the virtual VMkernel adapter for the physical FCoE network adapter installed on your host.

---

**Note** To avoid FCoE traffic disruptions, do not remove the FCoE network adapter (vmnic#) from the vSphere standard switch after you set up FCoE networking.

---

## Add Software FCoE Adapters in the vSphere Web Client

You must activate software FCoE adapters so that your host can use them to access Fibre Channel storage.

The number of software FCoE adapters you can activate corresponds to the number of physical FCoE NIC ports on your host. ESXi supports the maximum of four software FCoE adapters on one host.

### Prerequisites

Set up networking for the software FCoE adapter.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters**, and click **Add**.
- 4 Select **Software FCoE Adapter**.
- 5 On the Add Software FCoE Adapter dialog box, select an appropriate vmnic from the drop-down list of physical network adapters.

Only those adapters that are not yet used for FCoE traffic are listed.

- 6 Click **OK**.

The software FCoE adapter appears on the list of storage adapters.

### Results

After you activate the software FCoE adapter, you can view its properties. If you do not use the adapter, you can remove it from the list of adapters.

# Modifying Fibre Channel Storage for ESXi

# 6

This section discusses many of the storage devices supported in conjunction with VMware ESXi. For each device, it lists the major known potential issues, points to vendor-specific information (if available), and includes information from VMware knowledge base articles.

---

**Note** Information related to specific storage devices is updated only with each release. New information might already be available. Consult the most recent Storage/SAN Compatibility, check with your storage array vendor, and explore the VMware knowledge base articles.

---

This chapter includes the following topics:

- [Testing ESXi SAN Configurations](#)
- [General Setup Considerations for Fibre Channel SAN Arrays](#)
- [EMC CLARiiON Storage Systems in Fibre Channel SAN Configurations](#)
- [EMC Symmetrix Storage Systems](#)
- [IBM Systems Storage 8000 and IBM ESS800](#)
- [HP StorageWorks Storage Systems](#)
- [Hitachi Data Systems Storage](#)
- [NetApp Storage](#)
- [LSI-Based Storage Systems](#)

## Testing ESXi SAN Configurations

ESXi supports a variety of SAN storage systems in different configurations. Generally, VMware tests ESXi with supported storage systems for basic connectivity, HBA failover, and so on.

Not all storage devices are certified for all features and capabilities of ESXi, and vendors might have specific positions of support with regard to ESXi.

### Basic connectivity

Tests whether ESXi can recognize and operate with the storage array. This configuration does not allow for multipathing or any type of failover.

**HBA failover**

The server is equipped with multiple HBAs connecting to one or more SAN switches. The server is robust to HBA and switch failure only.

**Storage port failover**

The server is attached to multiple storage ports and is robust to storage port failures and switch failures.

**Boot from SAN**

The host boots from a LUN configured on the SAN rather than from the server itself.

**Direct connect**

The server connects to the array without using switches. For all other tests, a fabric connection is used. FC Arbitrated Loop (AL) is not supported.

**Clustering**

The system is tested with Microsoft Cluster Service running in the virtual machine.

## General Setup Considerations for Fibre Channel SAN Arrays

When you prepare your FC SAN storage to work with ESXi, you must follow specific general requirements that apply to all storage arrays.

For all storage arrays, make sure that the following requirements are met:

- LUNs must be presented to each HBA of each host with the same LUN ID number.  
Because instructions on how to configure identical SAN LUN IDs are vendor specific, consult your storage array documentation for more information.
- Unless specified for individual storage arrays, set the host type for LUNs presented to ESXi to Linux, Linux Cluster, or, if available, to vmware or esx.
- If you are using vMotion, DRS, or HA, make sure that both source and target hosts for virtual machines can see the same LUNs with identical LUN IDs.

SAN administrators might find it counterintuitive to have multiple hosts see the same LUNs because they might be concerned about data corruption. However, VMFS prevents multiple virtual machines from writing to the same file at the same time, so provisioning the LUNs to all required ESXi system is appropriate.

## EMC CLARiiON Storage Systems in Fibre Channel SAN Configurations

EMC CLARiiON storage systems work with ESXi hosts in SAN configurations.

Basic configuration includes the following steps:

- 1 Installing and configuring the storage device.
- 2 Configuring zoning at the switch level.
- 3 Creating RAID groups.
- 4 Creating and binding LUNs.
- 5 Registering the servers connected to the SAN. By default, the host automatically performs this step.
- 6 Creating storage groups that contain the servers and LUNs.

Use the EMC storage management software to perform configuration. For information, see the EMC documentation.

ESXi automatically sends the host's name and IP address to the array and registers the host with the array. You are no longer required to perform host registration manually. However, if you prefer to use storage management software, such as EMC Navisphere, to perform manual registration, turn off the ESXi auto-registration. Turning it off helps you avoid overwriting the manual user registration. For information, see [Disable Automatic Host Registration in the vSphere Web Client](#).

Because this array is an active-passive disk array, the following general considerations apply.

- The default multipathing policy for CLARiiON arrays that do not support ALUA is Most Recently Used. For CLARiiON arrays that support ALUA, the default multipathing policy is VMW\_PSP\_FIXED. The ESXi system sets the default policy when it identifies the array.
- To use boot from SAN, make sure that the active SP is chosen for the boot LUN's target in the HBA BIOS.

---

**Important** For ESXi to support EMC CLARiiON with ALUA, check the HCLs to make sure that you use the correct firmware version on the storage array. For additional information, contact your storage vendor.

---

## EMC Symmetrix Storage Systems

EMC Symmetrix storage systems work with ESXi hosts in FC SAN configurations. Generally, you use the EMC software to perform configurations.

The following settings are required on the Symmetrix networked storage system. For more information, see the EMC documentation.

- Common serial number (C)



- Auto negotiation (EAN) enabled
- Fibrepath enabled on this port (VCM)
- SCSI 3 (SC3) set enabled
- Unique world wide name (UWN)
- SPC-2 (Decal) (SPC2) SPC-2 flag is required

The ESXi host considers any LUNs from a Symmetrix storage array with a capacity of 50MB or less as management LUNs. These LUNs are also known as pseudo or gatekeeper LUNs. These LUNs appear in the EMC Symmetrix Management Interface and should not be used to hold data.

## IBM Systems Storage 8000 and IBM ESS800

The IBM Systems Storage 8000 and IBM ESS800 systems use an active-active array that does not need special configuration in conjunction with VMware ESXi.

The following considerations apply when you use these systems:

- Automatic resignaturing is not supported for these systems.
- To use RDMs successfully, a given LUN must be presented with the same LUN ID to every ESXi host in the cluster.
- In the ESS800 Configuration Management tool, select **Use same ID for LUN in source and target**.
- If you are configuring the host to use boot from SAN from these arrays, disable the internal fibre port for the corresponding blade until installation is finished.

## HP StorageWorks Storage Systems

This section includes configuration information for the different HP StorageWorks storage systems.

For additional information, see the HP ActiveAnswers section on VMware ESXi at the HP web site.

### HP StorageWorks EVA

To use an HP StorageWorks EVA system with ESXi, you must configure the correct host mode type.

Set the connection type to Custom when you present a LUN to the host. The value is one of the following:

- For EVA4000/6000/8000 active-active arrays with firmware below 5.031, use the host mode type 000000202200083E.
- For EVA4000/6000/8000 active-active arrays with firmware 5.031 and above, use the host mode type VMware.

Otherwise, EVA systems do not require special configuration changes to work with an ESXi system.

See the VMware Infrastructure, HP StorageWorks Best Practices at the HP Web site.

## HP StorageWorks XP

For HP StorageWorks XP, you need to set the host mode to specific parameters.

- On XP128/1024/10000/12000, set the host mode to Windows (0x0C).
- On XP24000/20000, set the host mode to 0x01.

## Hitachi Data Systems Storage

This section introduces the setup for Hitachi Data Systems storage. This storage solution is also available from Sun and as HP XP storage.

### LUN masking

To mask LUNs on an ESXi host, use the HDS Storage Navigator software for best results.

### Microcode and configurations

Check with your HDS representative for exact configurations and microcode levels needed for interoperability with ESXi. If your microcode is not supported, interaction with ESXi is usually not possible.

### Modes

The modes you set depend on the model you are using, for example:

- 9900 and 9900v uses Netware host mode.
- 9500v series uses Hostmode1: standard and Hostmode2: SUN Cluster.

Check with your HDS representative for host mode settings for the models not listed here.

## NetApp Storage

When configuring a NetApp storage device, first set the appropriate LUN type and initiator group type for the storage array.

### LUN type

VMware (if VMware type is not available, use Linux).

### Initiator group type

VMware (if VMware type is not available, use Linux).

You must then provision storage.

## LSI-Based Storage Systems

During ESXi installation, do not present the management LUN, also known as access LUN, from the LSI-based arrays to the host.

Otherwise, ESXi installation might fail.

# Booting ESXi from Fibre Channel SAN

## 7

When you set up your host to boot from a SAN, your host's boot image is stored on one or more LUNs in the SAN storage system. When the host starts, it boots from the LUN on the SAN rather than from its local disk.

ESXi supports booting through a Fibre Channel host bus adapter (HBA) or a Fibre Channel over Ethernet (FCoE) converged network adapter (CNA).

This chapter includes the following topics:

- [Boot from SAN Benefits](#)
- [Boot from Fibre Channel SAN Requirements and Considerations](#)
- [Getting Ready for Boot from SAN](#)
- [Configure Emulex HBA to Boot from SAN](#)
- [Configure QLogic HBA to Boot from SAN](#)

## Boot from SAN Benefits

Boot from SAN can provide numerous benefits to your environment. However, in certain cases, you should not use boot from SAN for ESXi hosts. Before you set up your system for boot from SAN, decide whether it is appropriate for your environment.

---

**Caution** When you use boot from SAN with multiple ESXi hosts, each host must have its own boot LUN. If you configure multiple hosts to share the same boot LUN, ESXi image corruption is likely to occur.

---

If you use boot from SAN, the benefits for your environment will include the following:

- Cheaper servers. Servers can be more dense and run cooler without internal storage.
- Easier server replacement. You can replace servers and have the new server point to the old boot location.
- Less wasted space. Servers without local disks often take up less space.

- Easier backup processes. You can backup the system boot images in the SAN as part of the overall SAN backup procedures. Also, you can use advanced array features such as snapshots on the boot image.
- Improved management. Creating and managing the operating system image is easier and more efficient.
- Better reliability. You can access the boot disk through multiple paths, which protects the disk from being a single point of failure.

## Boot from Fibre Channel SAN Requirements and Considerations

Your ESXi boot configuration must meet specific requirements.

**Table 7-1. Boot from SAN Requirements**

Requirement	Description
ESXi system requirements	Follow vendor recommendation for the server booting from a SAN.
Adapter requirements	Enable and correctly configure the adapter, so it can access the boot LUN. See your vendor documentation.
Access control	<ul style="list-style-type: none"> <li>■ Each host must have access to its own boot LUN only, not the boot LUNs of other hosts. Use storage system software to make sure that the host accesses only the designated LUNs.</li> <li>■ Multiple servers can share a diagnostic partition. You can use array specific LUN masking to achieve this.</li> </ul>
Multipathing support	Multipathing to a boot LUN on active-passive arrays is not supported because the BIOS does not support multipathing and is unable to activate a standby path.
SAN considerations	SAN connections must be through a switched topology if the array is not certified for direct connect topology. If the array is certified for direct connect topology, the SAN connections can be made directly to the array. Boot from SAN is supported for both switched topology and direct connect topology if these topologies for the specific array are certified.
Hardware- specific considerations	If you are running an IBM eServer BladeCenter and use boot from SAN, you must disable IDE drives on the blades.

## Getting Ready for Boot from SAN

When you set up your boot from SAN environment, you perform a number of tasks.

This section describes the generic boot-from-SAN enablement process on the rack mounted servers. For information on enabling boot from SAN on Cisco Unified Computing System FCoE blade servers, refer to Cisco documentation.

### Procedure

#### 1 Configure SAN Components and Storage System

Before you set up your ESXi host to boot from a SAN LUN, configure SAN components and a storage system.

#### 2 Configure Storage Adapter to Boot from SAN

When you set up your host to boot from SAN, you enable the boot adapter in the host BIOS. You then configure the boot adapter to initiate a primitive connection to the target boot LUN.

#### 3 Set Up Your System to Boot from Installation Media

When setting up your host to boot from SAN, you first boot the host from the VMware installation media. To achieve this, you need to change the system boot sequence in the BIOS setup.

## Configure SAN Components and Storage System

Before you set up your ESXi host to boot from a SAN LUN, configure SAN components and a storage system.

Because configuring the SAN components is vendor specific, refer to the product documentation for each item.

### Procedure

#### 1 Connect network cable, referring to any cabling guide that applies to your setup.

Check the switch wiring, if there is any.

#### 2 Configure the storage array.

- a From the SAN storage array, make the ESXi host visible to the SAN. This process is often referred to as creating an object.
- b From the SAN storage array, set up the host to have the WWPNs of the host's adapters as port names or node names.
- c Create LUNs.
- d Assign LUNs.
- e Record the IP addresses of the switches and storage arrays.
- f Record the WWPN for each SP.

---

**Caution** If you use scripted installation to install ESXi in boot from SAN mode, you need to take special steps to avoid unintended data loss.

---

## Configure Storage Adapter to Boot from SAN

When you set up your host to boot from SAN, you enable the boot adapter in the host BIOS. You then configure the boot adapter to initiate a primitive connection to the target boot LUN.

### Prerequisites

Determine the WWPN for the storage adapter.

### Procedure

- ◆ Configure the storage adapter to boot from SAN.

Because configuring boot adapters is vendor specific, refer to your vendor documentation.

## Set Up Your System to Boot from Installation Media

When setting up your host to boot from SAN, you first boot the host from the VMware installation media. To achieve this, you need to change the system boot sequence in the BIOS setup.

Because changing the boot sequence in the BIOS is vendor specific, refer to vendor documentation for instructions. The following procedure explains how to change the boot sequence on an IBM host.

### Procedure

- 1 During your system power up, enter the system BIOS Configuration/Setup Utility.
- 2 Select **Startup Options** and press Enter.
- 3 Select **Startup Sequence Options** and press Enter.
- 4 Change the **First Startup Device** to **[CD-ROM]**.

### Results

You can now install ESXi.

## Configure Emulex HBA to Boot from SAN

Configuring the Emulex HBA BIOS to boot from SAN includes enabling the BootBIOS prompt and enabling BIOS.

### Procedure

- 1 [Enable the BootBIOS Prompt](#)

When you configure the Emulex HBA BIOS to boot ESXi from SAN, you need to enable the BootBIOS prompt.

- 2 [Enable the BIOS](#)

When you configure the Emulex HBA BIOS to boot ESXi from SAN, you need to enable BIOS.

## Enable the BootBIOS Prompt

When you configure the Emulex HBA BIOS to boot ESXi from SAN, you need to enable the BootBIOS prompt.

### Procedure

- 1 Run `lputil`.
- 2 Select **3. Firmware Maintenance**.
- 3 Select an adapter.
- 4 Select **6. Boot BIOS Maintenance**.
- 5 Select **1. Enable Boot BIOS**.

## Enable the BIOS

When you configure the Emulex HBA BIOS to boot ESXi from SAN, you need to enable BIOS.

### Procedure

- 1 Reboot the host.
- 2 To configure the adapter parameters, press ALT+E at the Emulex prompt and follow these steps.
  - a Select an adapter (with BIOS support).
  - b Select **2. Configure This Adapter's Parameters**.
  - c Select **1. Enable or Disable BIOS**.
  - d Select **1** to enable BIOS.
  - e Select **x** to exit and **Esc** to return to the previous menu.
- 3 To configure the boot device, follow these steps from the Emulex main menu.
  - a Select the same adapter.
  - b Select **1. Configure Boot Devices**.
  - c Select the location for the Boot Entry.
  - d Enter the two-digit boot device.
  - e Enter the two-digit (HEX) starting LUN (for example, **08**).
  - f Select the boot LUN.
  - g Select **1. WWPN**. (Boot this device using WWPN, not DID).
  - h Select **x** to exit and **Y** to reboot.
- 4 Boot into the system BIOS and move Emulex first in the boot controller sequence.
- 5 Reboot and install on a SAN LUN.



## Configure QLogic HBA to Boot from SAN

This sample procedure explains how to configure the QLogic HBA to boot ESXi from SAN. The procedure involves enabling the QLogic HBA BIOS, enabling the selectable boot, and selecting the boot LUN.

### Procedure

- 1 While booting the server, press **Ctrl+Q** to enter the Fast!UTIL configuration utility.
- 2 Perform the appropriate action depending on the number of HBAs.

Option	Description
<b>One HBA</b>	If you have only one host bus adapter (HBA), the Fast!UTIL Options page appears. Skip to <a href="#">Step Step 3</a> .
<b>Multiple HBAs</b>	<p>If you have more than one HBA, select the HBA manually.</p> <ol style="list-style-type: none"> <li>a In the Select Host Adapter page, use the arrow keys to position the cursor on the appropriate HBA.</li> <li>b Press <b>Enter</b>.</li> </ol>

- 3 In the Fast!UTIL Options page, select **Configuration Settings** and press **Enter**.
- 4 In the Configuration Settings page, select **Adapter Settings** and press **Enter**.
- 5 Set the BIOS to search for SCSI devices.
  - a In the Host Adapter Settings page, select **Host Adapter BIOS**.
  - b Press **Enter** to toggle the value to Enabled.
  - c Press **Esc** to exit.
- 6 Enable the selectable boot.
  - a Select **Selectable Boot Settings** and press **Enter**.
  - b In the Selectable Boot Settings page, select **Selectable Boot**.
  - c Press **Enter** to toggle the value to **Enabled**.
- 7 Use the cursor keys to select the Boot Port Name entry in the list of storage processors (SPs) and press **Enter** to open the Select Fibre Channel Device screen.
- 8 Use the cursor keys to select the specific SP and press **Enter**.

If you are using an active-passive storage array, the selected SP must be on the preferred (active) path to the boot LUN. If you are not sure which SP is on the active path, use your storage array management software to find out. The target IDs are created by the BIOS and might change with each reboot.

- 9 Perform the appropriate action depending on the number of LUNs attached to the SP.

Option	Description
One LUN	The LUN is selected as the boot LUN. You do not need to enter the Select LUN screen.
Multiple LUNs	Select LUN screen opens. Use the cursor to select the boot LUN, then press <b>Enter</b> .

- 10 If any remaining storage processors show in the list, press **C** to clear the data.
- 11 Press **Esc** twice to exit and press **Enter** to save the setting.

# Booting ESXi with Software FCoE

# 8

ESXi supports boot from FCoE capable network adapters.

When you install and boot ESXi from an FCoE LUN, the host can use a VMware software FCoE adapter and a network adapter with FCoE capabilities. The host does not require a dedicated FCoE HBA.

You perform most configurations through the option ROM of your network adapter. The network adapters must support one of the following formats, which communicate parameters about an FCoE boot device to VMkernel.

- FCoE Boot Firmware Table (FBFT). FBFT is Intel propriety.
- FCoE Boot Parameter Table (FBPT). FBPT is defined by VMware for third-party vendors to implement software FCoE boot.

The configuration parameters are set in the option ROM of your adapter. During an ESXi installation or a subsequent boot, these parameters are exported in to system memory in either FBFT format or FBPT format. The VMkernel can read the configuration settings and use them to access the boot LUN.

This chapter includes the following topics:

- [Requirements and Considerations for Software FCoE Boot](#)
- [Best Practices for Software FCoE Boot](#)
- [Set Up Software FCoE Boot](#)
- [Troubleshooting Installation and Boot from Software FCoE](#)

## Requirements and Considerations for Software FCoE Boot

When you boot the ESXi host from SAN using software FCoE, certain requirements and considerations apply.

### Requirements

- ESXi 5.1 or later.

- The network adapter must have the following capabilities:
  - Be FCoE capable.
  - Support ESXi 5.x open FCoE stack.
  - Contain FCoE boot firmware which can export boot information in FBFT format or FBPT format.

## Considerations

- You cannot change software FCoE boot configuration from within ESXi.
- Coredump is not supported on any software FCoE LUNs, including the boot LUN.
- Multipathing is not supported at pre-boot.
- Boot LUN cannot be shared with other hosts even on shared storage.

## Best Practices for Software FCoE Boot

VMware recommends several best practices when you boot your system from a software FCoE LUN.

- Make sure that the host has access to the entire boot LUN. The boot LUN cannot be shared with other hosts even on shared storage.
- If you use Intel 10 Gigabit Ethernet Controller (Niantec) with a Cisco switch, configure the switch port in the following way:
  - Enable the Spanning Tree Protocol (STP).
  - Turn off `switchport trunk native vlan` for the VLAN used for FCoE.

## Set Up Software FCoE Boot

Your ESXi host can boot from a FCoE LUN using the software FCoE adapter a network adapter.

When you configure your host for a software FCoE boot, you perform a number of tasks.

### Prerequisites

The network adapter has the following capabilities:

- Support partial FCoE offload (software FCoE).
- Contain either a FCoE Boot Firmware Table (FBFT) or a FCoE Boot Parameter Table (FBPT).

For information about network adapters that support software FCoE boot, see the *vSphere Compatibility Guide*.

## Procedure

### 1 Configure Software FCoE Boot Parameters

To support a software FCoE boot process, a network adapter on your host must have a specially configured FCoE boot firmware. When you configure the firmware, you enable the adapter for the software FCoE boot and specify the boot LUN parameters.

### 2 Install and Boot ESXi from Software FCoE LUN

When you set up your system to boot from a software FCoE LUN, you install the ESXi image to the target LUN. You can then boot your host from that LUN.

## Configure Software FCoE Boot Parameters

To support a software FCoE boot process, a network adapter on your host must have a specially configured FCoE boot firmware. When you configure the firmware, you enable the adapter for the software FCoE boot and specify the boot LUN parameters.

## Procedure

- ◆ In the option ROM of the network adapter, specify software FCoE boot parameters.

These parameters include boot target, boot LUN, VLAN ID, and so on.

Because configuring the network adapter is vendor specific, review your vendor documentation for instructions.

## Install and Boot ESXi from Software FCoE LUN

When you set up your system to boot from a software FCoE LUN, you install the ESXi image to the target LUN. You can then boot your host from that LUN.

## Prerequisites

- Configure the option ROM of the network adapter to point to a target LUN that you want to use as the boot LUN. Make sure that you have information about the bootable LUN.
- Change the boot order in the system BIOS to the following sequence:
  - a The network adapter that you use for the software FCoE boot.
  - b The ESXi installation media.

See the vendor documentation for your system.

## Procedure

- 1 Start an interactive installation from the ESXi installation CD/DVD.

The ESXi installer verifies that FCoE boot is enabled in the BIOS and, if needed, creates a standard virtual switch for the FCoE capable network adapter. The name of the vSwitch is VMware\_FCoE\_vSwitch. The installer then uses preconfigured FCoE boot parameters to discover and display all available FCoE LUNs.

- 2 On the **Select a Disk** screen, select the software FCoE LUN that you specified in the boot parameter setting.

If the boot LUN does not appear in this menu, make sure that you correctly configured boot parameters in the option ROM of the network adapter.

- 3 Follow the prompts to complete the installation.

- 4 Reboot the host.

- 5 Change the boot order in the system BIOS so that the FCoE boot LUN is the first bootable device.

ESXi continues booting from the software FCoE LUN until it is ready to be used.

## What to do next

If needed, you can rename and modify the VMware\_FCoE\_vSwitch that the installer automatically created. Make sure that the Cisco Discovery Protocol (CDP) mode is set to Listen or Both.

# Troubleshooting Installation and Boot from Software FCoE

If the installation or boot of ESXi from a software FCoE LUN fails, you can use several troubleshooting methods.

## Problem

When you install or boot ESXi from FCoE storage using a VMware software FCoE adapter and a network adapter with partial FCoE offload capabilities, the installation or the boot process fails.

## Solution

- Make sure that you correctly configured boot parameters in the option ROM of the FCoE network adapter.
- During installation, monitor the BIOS of the FCoE network adapter for any errors.
- If possible, check the VMkernel log for errors.
- Use the `esxcli` command to verify whether the boot LUN is present.

```
esxcli conn_options hardware bootdevice list
```

# Best Practices for Fibre Channel Storage

## 9

When using ESXi with Fibre Channel SAN, follow best practices that VMware offers to avoid performance problems.

The vSphere Web Client offer extensive facilities for collecting performance information. The information is graphically displayed and frequently updated.

You can also use the `resxtop` or `esxtop` command-line utilities. The utilities provide a detailed look at how ESXi uses resources in real time. For more information, see the *vSphere Resource Management* documentation.

Check with your storage representative if your storage system supports Storage API - Array Integration hardware acceleration features. If it does, refer to your vendor documentation for information on how to enable hardware acceleration support on the storage system side. For more information, see [Chapter 24 Storage Hardware Acceleration](#).

This chapter includes the following topics:

- [Preventing Fibre Channel SAN Problems](#)
- [Disable Automatic Host Registration in the vSphere Web Client](#)
- [Optimizing Fibre Channel SAN Storage Performance](#)
- [Fibre Channel SAN Configuration Checklist](#)

## Preventing Fibre Channel SAN Problems

When using ESXi in conjunction with a Fibre Channel SAN, you must follow specific guidelines to avoid SAN problems.

You should observe these tips for preventing problems with your SAN configuration:

- Place only one VMFS datastore on each LUN.
- Do not change the path policy the system sets for you unless you understand the implications of making such a change.
- Document everything. Include information about zoning, access control, storage, switch, server and FC HBA configuration, software and firmware versions, and storage cable plan.

- Plan for failure:
  - Make several copies of your topology maps. For each element, consider what happens to your SAN if the element fails.
  - Cross off different links, switches, HBAs and other elements to ensure you did not miss a critical failure point in your design.
- Ensure that the Fibre Channel HBAs are installed in the correct slots in the host, based on slot and bus speed. Balance PCI bus load among the available busses in the server.
- Become familiar with the various monitor points in your storage network, at all visibility points, including host's performance charts, FC switch statistics, and storage performance statistics.
- Be cautious when changing IDs of the LUNs that have VMFS datastores being used by your ESXi host. If you change the ID, the datastore becomes inactive and its virtual machines fail. You can resignature the datastore to make it active again. See [Managing Duplicate VMFS Datastores](#).

If there are no running virtual machines on the VMFS datastore, after you change the ID of the LUN, you must use rescan to reset the ID on your host. For information on using rescan, see [Storage Refresh and Rescan Operations](#).

## Disable Automatic Host Registration in the vSphere Web Client

When you use EMC CLARiiON or InVista arrays for storage, it is required that the hosts register with the arrays. ESXi performs automatic host registration by sending the host's name and IP address to the array. If you prefer to perform manual registration using storage management software, disable the ESXi auto-registration feature.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Settings**.
- 3 Under System, click **Advanced System Settings**.
- 4 Under Advanced System Settings, select the **Disk.EnableNaviReg** parameter and click the Edit icon.
- 5 Change the value to 0.

### Results

This disables the automatic host registration enabled by default.



# Optimizing Fibre Channel SAN Storage Performance

Several factors contribute to optimizing a typical SAN environment.

If the environment is properly configured, the SAN fabric components (particularly the SAN switches) are only minor contributors because of their low latencies relative to servers and storage arrays. Make sure that the paths through the switch fabric are not saturated, that is, that the switch fabric is running at the highest throughput.

## Storage Array Performance

Storage array performance is one of the major factors contributing to the performance of the entire SAN environment.

If there are issues with storage array performance, be sure to consult your storage array vendor's documentation for any relevant information.

Follow these general guidelines to improve the array performance in the vSphere environment:

- When assigning LUNs, remember that each LUN is accessed by a number of hosts, and that a number of virtual machines can run on each host. One LUN used by a host can service I/O from many different applications running on different operating systems. Because of this diverse workload, the RAID group containing the ESXi LUNs should not include LUNs used by other servers that are not running ESXi.
- Make sure read/write caching is enabled.
- SAN storage arrays require continual redesign and tuning to ensure that I/O is load balanced across all storage array paths. To meet this requirement, distribute the paths to the LUNs among all the SPs to provide optimal load balancing. Close monitoring indicates when it is necessary to rebalance the LUN distribution.

Tuning statically balanced storage arrays is a matter of monitoring the specific performance statistics (such as I/O operations per second, blocks per second, and response time) and distributing the LUN workload to spread the workload across all the SPs.

---

**Note** Dynamic load balancing is not currently supported with ESXi.

---

## Server Performance with Fibre Channel

You must consider several factors to ensure optimal server performance.

Each server application must have access to its designated storage with the following conditions:

- High I/O rate (number of I/O operations per second)
- High throughput (megabytes per second)
- Minimal latency (response times)

Because each application has different requirements, you can meet these goals by choosing an appropriate RAID group on the storage array. To achieve performance goals:

- Place each LUN on a RAID group that provides the necessary performance levels. Pay attention to the activities and resource utilization of other LUNS in the assigned RAID group. A high-performance RAID group that has too many applications doing I/O to it might not meet performance goals required by an application running on the ESXi host.
- Make sure that each server has a sufficient number of HBAs to allow maximum throughput for all the applications hosted on the server for the peak period. I/O spread across multiple HBAs provide higher throughput and less latency for each application.
- To provide redundancy in the event of HBA failure, make sure the server is connected to a dual redundant fabric.
- When allocating LUNs or RAID groups for ESXi systems, multiple operating systems use and share that resource. As a result, the performance required from each LUN in the storage subsystem can be much higher if you are working with ESXi systems than if you are using physical machines. For example, if you expect to run four I/O intensive applications, allocate four times the performance capacity for the ESXi LUNs.
- When using multiple ESXi systems in conjunction with vCenter Server, the performance needed from the storage subsystem increases correspondingly.
- The number of outstanding I/Os needed by applications running on an ESXi system should match the number of I/Os the HBA and storage array can handle.

## Fibre Channel SAN Configuration Checklist

This topic provides a checklist of special setup requirements for different storage arrays and ESXi hosts.

**Table 9-1. Multipathing Setup Requirements**

Component	Comments
All storage arrays	Write cache must be disabled if not battery backed.
Topology	No single failure should cause both HBA and SP failover, especially with active-passive storage arrays.
IBM TotalStorage DS 4000 (formerly FastT)	Host type must be LNXCL or VMware in later versions. AVT (Auto Volume Transfer) is disabled in this host mode.
HDS 99xx and 95xxV family	HDS 9500V family (Thunder) requires two host modes: <ul style="list-style-type: none"> <li>■ Host Mode 1: Standard.</li> <li>■ Host Mode 2: Sun Cluster</li> </ul> HDS 99xx family (Lightning) and HDS Tabma (USP) require host mode set to Netware.
EMC Symmetrix	Enable the SPC2 and SC3 settings. Contact EMC for the latest settings.
EMC Clariion	Set the EMC Clariion failover mode to 1 or 4. Contact EMC for details.
HP MSA	Host type must be Linux. Set the connection type for each HBA port to Linux.

Table 9-1. Multipathing Setup Requirements (continued)

Component	Comments
HP EVA	For EVA4000/6000/8000 firmware 5.031 and above, set the host type to VMware. Otherwise, set the host mode type to Custom. The value is: 000000202200083E.
HP XP	<ul style="list-style-type: none"> <li>■ On XP128/1024/10000/12000, set the host mode to Windows (0x0C).</li> <li>■ On XP24000/20000, set the host mode to 0x01.</li> </ul>
NetApp	No specific requirements
ESXi Configuration	<ul style="list-style-type: none"> <li>■ For all LUNs hosting clustered disks on active-passive arrays, use the Most Recently Used PSP.</li> <li>■ For LUNs on active-active arrays, you can use the Most Recently Used or Fixed PSP.</li> <li>■ With either active-passive or active-active arrays, you can use the Round Robin PSP.</li> <li>■ All FC HBAs must be of the same model.</li> <li>■ Set the following Software Advanced Settings for the host: Disk.UseLunReset to 1.</li> </ul>

# Using ESXi with iSCSI SAN

# 10

You can use ESXi in conjunction with a storage area network (SAN), a specialized high-speed network that connects computer systems to high-performance storage subsystems. Using ESXi together with a SAN provides storage consolidation, improves reliability, and helps with disaster recovery.

To use ESXi effectively with a SAN, you must have a working knowledge of ESXi systems and SAN concepts. Also, when you set up ESXi hosts to use Internet SCSI (iSCSI) SAN storage systems, you must be aware of certain special considerations that exist.

This chapter includes the following topics:

- [iSCSI SAN Concepts](#)
- [How Virtual Machines Access Data on an iSCSI SAN](#)

## iSCSI SAN Concepts

If you are an administrator who plans to set up ESXi hosts to work with iSCSI SANs, you must have a working knowledge of iSCSI concepts.

iSCSI SANs use Ethernet connections between computer systems, or host servers, and high performance storage subsystems. The SAN components include iSCSI host bus adapters (HBAs) or Network Interface Cards (NICs) in the host servers, switches and routers that transport the storage traffic, cables, storage processors (SPs), and storage disk systems.

iSCSI SAN uses a client-server architecture. The client, called iSCSI initiator, operates on your host. It initiates iSCSI sessions by issuing SCSI commands and transmitting them, encapsulated into iSCSI protocol, to a server. The server is known as an iSCSI target. The iSCSI target represents a physical storage system on the network. It can also be provided by a virtual iSCSI SAN, for example, an iSCSI target emulator running in a virtual machine. The iSCSI target responds to the initiator's commands by transmitting required iSCSI data.

## iSCSI Multipathing

When transferring data between the host server and storage, the SAN uses a technique known as multipathing. Multipathing allows you to have more than one physical path from the ESXi host to a LUN on a storage system.

Generally, a single path from a host to a LUN consists of an iSCSI adapter or NIC, switch ports, connecting cables, and the storage controller port. If any component of the path fails, the host selects another available path for I/O. The process of detecting a failed path and switching to another is called path failover.

For more information on multipathing, see [Chapter 23 Understanding Multipathing and Failover](#).

## Ports in the iSCSI SAN

A single discoverable entity on the iSCSI SAN, such as an initiator or a target, represents an iSCSI node. Each node has one or more ports that connect it to the SAN.

iSCSI ports are end-points of an iSCSI session. Each node can be identified in a number of ways.

### IP Address

Each iSCSI node can have an IP address associated with it so that routing and switching equipment on your network can establish the connection between the server and storage. This address is just like the IP address that you assign to your computer to get access to your company's network or the Internet.

### iSCSI Name

A worldwide unique name for identifying the node. iSCSI uses the iSCSI Qualified Name (IQN) and Extended Unique Identifier (EUI).

By default, ESXi generates unique iSCSI names for your iSCSI initiators, for example, `iqn.1998-01.com.vmware:iscsitestox-68158ef2`. Usually, you do not have to change the default value, but if you do, make sure that the new iSCSI name you enter is worldwide unique.

### iSCSI Alias

A more manageable name for an iSCSI device or port used instead of the iSCSI name. iSCSI aliases are not unique and are intended to be just a friendly name to associate with a port.

## iSCSI Naming Conventions

iSCSI uses a special unique name to identify an iSCSI node, either target or initiator. This name is similar to the WorldWide Name (WWN) associated with Fibre Channel devices and is used as a way to universally identify the node.

iSCSI names are formatted in two different ways. The most common is the IQN format.

For more details on iSCSI naming requirements and string profiles, see RFC 3721 and RFC 3722 on the IETF Web site.

### iSCSI Qualified Name (IQN) Format

The IQN format takes the form `iqn.yyyy-mm.naming-authority:unique name`, where:

- `yyyy-mm` is the year and month when the naming authority was established.

- *naming-authority* is usually reverse syntax of the Internet domain name of the naming authority. For example, the `iscsi.vmware.com` naming authority could have the iSCSI qualified name form of `iqn.1998-01.com.vmware.iscsi`. The name indicates that the `vmware.com` domain name was registered in January of 1998, and `iscsi` is a subdomain, maintained by `vmware.com`.
- *unique name* is any name you want to use, for example, the name of your host. The naming authority must make sure that any names assigned following the colon are unique, such as:
  - `iqn.1998-01.com.vmware.iscsi:name1`
  - `iqn.1998-01.com.vmware.iscsi:name2`
  - `iqn.1998-01.com.vmware.iscsi:name999`

## Enterprise Unique Identifier (EUI) Format

The EUI format takes the form `eui.16 hex digits`.

For example, `eui.0123456789ABCDEF`.

The 16-hexadecimal digits are text representations of a 64-bit number of an IEEE EUI (extended unique identifier) format. The top 24 bits are a company ID that IEEE registers with a particular company. The lower 40 bits are assigned by the entity holding that company ID and must be unique.

## iSCSI Initiators

To access iSCSI targets, your host uses iSCSI initiators. The initiators transport SCSI requests and responses, encapsulated into the iSCSI protocol, between the host and the iSCSI target.

Your host supports different types of initiators.

For information on configuring and using iSCSI adapters, see [Chapter 11 Configuring iSCSI Adapters and Storage](#).

### Software iSCSI Adapter

A software iSCSI adapter is a VMware code built into the VMkernel. It allows your host to connect to the iSCSI storage device through standard network adapters. The software iSCSI adapter handles iSCSI processing while communicating with the network adapter. With the software iSCSI adapter, you can use iSCSI technology without purchasing specialized hardware.

### Hardware iSCSI Adapter

A hardware iSCSI adapter is a third-party adapter that offloads iSCSI and network processing from your host. Hardware iSCSI adapters are divided into categories.

#### Dependent Hardware iSCSI Adapter

Depends on VMware networking, and iSCSI configuration and management interfaces provided by VMware.

This type of adapter can be a card that presents a standard network adapter and iSCSI offload functionality for the same port. The iSCSI offload functionality depends on the host's network configuration to obtain the IP, MAC, and other parameters used for iSCSI sessions. An example of a dependent adapter is the iSCSI licensed Broadcom 5709 NIC.

### Independent Hardware iSCSI Adapter

Implements its own networking and iSCSI configuration and management interfaces.

An example of an independent hardware iSCSI adapter is a card that either presents only iSCSI offload functionality or iSCSI offload functionality and standard NIC functionality. The iSCSI offload functionality has independent configuration management that assigns the IP, MAC, and other parameters used for the iSCSI sessions. An example of a independent adapter is the QLogic QLA4052 adapter.

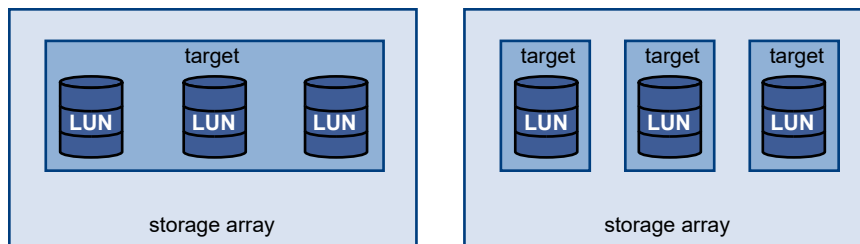
Hardware iSCSI adapters might need to be licensed. Otherwise, they will not appear in the client or vSphere CLI. Contact your vendor for licensing information.

## Establishing iSCSI Connections

In the ESXi context, the term target identifies a single storage unit that your host can access. The terms storage device and LUN describe a logical volume that represents storage space on a target. Typically, the terms device and LUN, in the ESXi context, mean a SCSI volume presented to your host from a storage target and available for formatting.

Different iSCSI storage vendors present storage to servers in different ways. Some vendors present multiple LUNs on a single target, while others present multiple targets with one LUN each. While the way the storage is used by ESXi is similar, the way the information is presented through administrative tools is different.

**Figure 10-1. Target Compared to LUN Representations**



Three LUNs are available in each of these configurations. In the first case, the host detects one target but that target has three LUNs that can be used. Each of the LUNs represents individual storage volume. In the second case, the host detects three different targets, each having one LUN.

Host-based iSCSI initiators establish connections to each target. Storage systems with a single target containing multiple LUNs have traffic to all the LUNs on a single connection. With a system that has three targets with one LUN each, a host uses separate connections to the three LUNs. This information is useful when you are trying to aggregate storage traffic on multiple connections from the host with multiple iSCSI HBAs, where traffic for one target can be set to a particular HBA, while traffic for another target can use a different HBA.

## iSCSI Storage System Types

ESXi supports different storage systems and arrays.

The types of storage that your host supports include active-active, active-passive, and ALUA-compliant.

### Active-active storage system

Allows access to the LUNs simultaneously through all the storage ports that are available without significant performance degradation. All the paths are active at all times, unless a path fails.

### Active-passive storage system

A system in which one storage processor is actively providing access to a given LUN. The other processors act as backup for the LUN and can be actively providing access to other LUN I/O. I/O can be successfully sent only to an active port for a given LUN. If access through the active storage port fails, one of the passive storage processors can be activated by the servers accessing it.

### Asymmetrical storage system

Supports Asymmetric Logical Unit Access (ALUA). ALUA-complaint storage systems provide different levels of access per port. ALUA allows hosts to determine the states of target ports and prioritize paths. The host uses some of the active paths as primary while others as secondary.

### Virtual port storage system

Allows access to all available LUNs through a single virtual port. These are active-active storage devices, but hide their multiple connections though a single port. ESXi multipathing does not make multiple connections from a specific port to the storage by default. Some storage vendors supply session managers to establish and manage multiple connections to their storage. These storage systems handle port failover and connection balancing transparently. This is often referred to as transparent failover.

## Discovery, Authentication, and Access Control

You can use several mechanisms to discover your storage and to limit access to it.

You must configure your host and the iSCSI storage system to support your storage access control policy.



## Discovery

A discovery session is part of the iSCSI protocol, and it returns the set of targets you can access on an iSCSI storage system. The two types of discovery available on ESXi are dynamic and static. Dynamic discovery obtains a list of accessible targets from the iSCSI storage system, while static discovery can only try to access one particular target by target name and address.

For more information, see [Configuring Discovery Addresses for iSCSI Adapters](#).

## Authentication

iSCSI storage systems authenticate an initiator by a name and key pair. ESXi supports the CHAP protocol, which VMware recommends for your SAN implementation. To use CHAP authentication, the ESXi host and the iSCSI storage system must have CHAP enabled and have common credentials.

For information on enabling CHAP, see [Configuring CHAP Parameters for iSCSI Adapters](#).

## Access Control

Access control is a policy set up on the iSCSI storage system. Most implementations support one or more of three types of access control:

- By initiator name
- By IP address
- By the CHAP protocol

Only initiators that meet all rules can access the iSCSI volume.

Using only CHAP for access control can slow down rescans because the ESXi host can discover all targets, but then fails at the authentication step. iSCSI rescans work faster if the host discovers only the targets it can authenticate.

## Error Correction

To protect the integrity of iSCSI headers and data, the iSCSI protocol defines error correction methods known as header digests and data digests.

Both parameters are disabled by default, but you can enable them. These digests pertain to, respectively, the header and SCSI data being transferred between iSCSI initiators and targets, in both directions.

Header and data digests check the end-to-end, noncryptographic data integrity beyond the integrity checks that other networking layers provide, such as TCP and Ethernet. They check the entire communication path, including all elements that can change the network-level traffic, such as routers, switches, and proxies.

The existence and type of the digests are negotiated when an iSCSI connection is established. When the initiator and target agree on a digest configuration, this digest must be used for all traffic between them.

Enabling header and data digests does require additional processing for both the initiator and the target and can affect throughput and CPU use performance.

---

**Note** Systems that use Intel Nehalem processors offload the iSCSI digest calculations, thus reducing the impact on performance.

---

For information on enabling header and data digests, see [Configuring Advanced Parameters for iSCSI](#).

## How Virtual Machines Access Data on an iSCSI SAN

ESXi stores a virtual machine's disk files within a VMFS datastore that resides on a SAN storage device. When virtual machine guest operating systems issue SCSI commands to their virtual disks, the SCSI virtualization layer translates these commands to VMFS file operations.

When a virtual machine interacts with its virtual disk stored on a SAN, the following process takes place:

- 1 When the guest operating system in a virtual machine reads or writes to SCSI disk, it issues SCSI commands to the virtual disk.
- 2 Device drivers in the virtual machine's operating system communicate with the virtual SCSI controllers.
- 3 The virtual SCSI controller forwards the command to the VMkernel.
- 4 The VMkernel performs the following tasks.
  - a Locates the file, which corresponds to the guest virtual machine disk, in the VMFS volume.
  - b Maps the requests for the blocks on the virtual disk to blocks on the appropriate physical device.
  - c Sends the modified I/O request from the device driver in the VMkernel to the iSCSI initiator (hardware or software).
- 5 If the iSCSI initiator is a hardware iSCSI adapter (both independent or dependent), the adapter performs the following tasks.
  - a Encapsulates I/O requests into iSCSI Protocol Data Units (PDUs).
  - b Encapsulates iSCSI PDUs into TCP/IP packets.
  - c Sends IP packets over Ethernet to the iSCSI storage system.
- 6 If the iSCSI initiator is a software iSCSI adapter, the following takes place.
  - a The iSCSI initiator encapsulates I/O requests into iSCSI PDUs.
  - b The initiator sends iSCSI PDUs through TCP/IP connections.
  - c The VMkernel TCP/IP stack relays TCP/IP packets to a physical NIC.
  - d The physical NIC sends IP packets over Ethernet to the iSCSI storage system.

- 7 Depending on which port the iSCSI initiator uses to connect to the network, Ethernet switches and routers carry the request to the storage device that the host wants to access.

# Configuring iSCSI Adapters and Storage

# 11

Before ESXi can work with a SAN, you must set up your iSCSI adapters and storage.

To do this, you must first observe certain basic requirements and then follow best practices for installing and setting up hardware or software iSCSI adapters to access the SAN.

The following table lists the iSCSI adapters (vmhbases) that ESXi supports and indicates whether VMkernel networking configuration is required.

**Table 11-1. Supported iSCSI adapters**

iSCSI Adapter (vmhba)	Description	VMkernel Networking
Software	Uses standard NICs to connect your host to a remote iSCSI target on the IP network .	Required
Independent Hardware	Third-party adapter that offloads the iSCSI and network processing and management from your host.	Not required
Dependent Hardware	Third-party adapter that depends on VMware networking and iSCSI configuration and management interfaces.	Required

After you set up the iSCSI adapters, you can create a datastore on iSCSI storage. For details on how to create and manage datastores, see [Creating Datastores](#).

This chapter includes the following topics:

- [ESXi iSCSI SAN Requirements](#)
- [ESXi iSCSI SAN Restrictions](#)
- [Setting LUN Allocations for iSCSI](#)
- [Network Configuration and Authentication](#)
- [Setting Up Independent Hardware iSCSI Adapters](#)
- [Configuring Dependent Hardware iSCSI Adapters](#)
- [Configuring Software iSCSI Adapter](#)

- [Modify General Properties for iSCSI Adapters](#)
- [Setting Up iSCSI Network](#)
- [Using Jumbo Frames with iSCSI](#)
- [Configuring Discovery Addresses for iSCSI Adapters](#)
- [Configuring CHAP Parameters for iSCSI Adapters](#)
- [Configuring Advanced Parameters for iSCSI](#)
- [iSCSI Session Management](#)

## ESXi iSCSI SAN Requirements

You must meet several requirements for your ESXi host to work properly with a SAN.

- Verify that your SAN storage hardware and firmware combinations are supported in conjunction with ESXi systems. For an up-to-date list, see *vSphere Compatibility Guide*.
- Configure your system to have only one VMFS datastore for each LUN.
- Unless you are using diskless servers, set up a diagnostic partition on a local storage. If you have diskless servers that boot from iSCSI SAN, see [General Boot from iSCSI SAN Recommendations](#) for information about diagnostic partitions with iSCSI.
- Use RDMs for access to any raw disk. For information, see [Chapter 17 Raw Device Mapping](#).
- Set the SCSI controller driver in the guest operating system to a large enough queue. For information on changing queue depth for iSCSI adapters and virtual machines, see *vSphere Troubleshooting*.
- On virtual machines running Microsoft Windows, increase the value of the SCSI TimeoutValue parameter to allow Windows to better tolerate delayed I/O resulting from path failover. For information, see [Set Timeout on Windows Guest OS](#).

## ESXi iSCSI SAN Restrictions

A number of restrictions exist when you use ESXi with an iSCSI SAN.

- ESXi does not support iSCSI-connected tape devices.
- You cannot use virtual-machine multipathing software to perform I/O load balancing to a single physical LUN.
- ESXi does not support multipathing when you combine independent hardware adapters with either software or dependent hardware adapters.
- ESXi does not support IPv6 with software iSCSI and dependent hardware iSCSI.

## Setting LUN Allocations for iSCSI

When preparing your ESXi system to use iSCSI SAN storage you need to set LUN allocations.

Note the following points:

- **Storage Provisioning.** To ensure that the host recognizes LUNs at startup time, configure all iSCSI storage targets so that your host can access them and use them. Also, configure your host so that it can discover all available iSCSI targets.
- **vMotion and VMware DRS.** When you use vCenter Server and vMotion or DRS, make sure that the LUNs for the virtual machines are provisioned to all hosts. This configuration provides the greatest freedom in moving virtual machines.
- **Active-active versus active-passive arrays.** When you use vMotion or DRS with an active-passive SAN storage device, make sure that all hosts have consistent paths to all storage processors. Not doing so can cause path thrashing when a vMotion migration occurs.

For active-passive storage arrays not listed in Storage/SAN Compatibility, VMware does not support storage-port failover. You must connect the server to the active port on the storage system. This configuration ensures that the LUNs are presented to the host.

## Network Configuration and Authentication

Before your ESXi host can discover iSCSI storage, the iSCSI initiators must be configured and authentication might have to be set up.

- For software iSCSI and dependent hardware iSCSI, networking for the VMkernel must be configured. You can verify the network configuration by using the `vmkping` utility.
- For independent hardware iSCSI, network parameters, such as IP address, subnet mask, and default gateway must be configured on the HBA.
- Check and change the default initiator name if necessary.
- The dynamic discovery address or static discovery address and target name of the storage system must be set. For software iSCSI and dependent hardware iSCSI, the address should be pingable using `vmkping`.
- For CHAP authentication, enable it on the initiator and the storage system side. After authentication is enabled, it applies for all of the targets that are not yet discovered, but does not apply to targets that are already discovered. After the discovery address is set, the new targets discovered are exposed and can be used at that point.

For details on how to use the `vmkping` command, search the VMware Knowledge Base.

## Setting Up Independent Hardware iSCSI Adapters

An independent hardware iSCSI adapter is a specialized third-party adapter capable of accessing iSCSI storage over TCP/IP. This iSCSI adapter handles all iSCSI and network processing and management for your ESXi system.

The setup and configuration process for the independent hardware iSCSI adapters involves these steps:

- 1 Check whether the adapter needs to be licensed.  
See your vendor documentation.
- 2 Install the adapter.  
For installation information and information on firmware updates, see vendor documentation.
- 3 Verify that the adapter is installed correctly.  
See [View Independent Hardware iSCSI Adapters in the vSphere Web Client](#).
- 4 Configure discovery information.  
See [Configuring Discovery Addresses for iSCSI Adapters](#).
- 5 (Optional) Configure CHAP parameters.  
See [Configuring CHAP Parameters for iSCSI Adapters](#).
- 6 (Optional) Enable Jumbo Frames.  
See [Using Jumbo Frames with iSCSI](#).

## View Independent Hardware iSCSI Adapters in the vSphere Web Client

View an independent hardware iSCSI adapter to verify that it is correctly installed and ready for configuration.

After you install an independent hardware iSCSI adapter on a host, it appears on the list of storage adapters available for configuration. You can view its properties.

### Prerequisites

Required privilege: **Host.Configuration.Storage Partition Configuration**

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters**.

If installed, the hardware iSCSI adapter appears on the list of storage adapters.

- 4 Select the adapter to view.

The default details for the adapter appear, including the model, iSCSI name, iSCSI alias, IP address, and target and paths information.

## Change IP Address and DNS Settings in the vSphere Web Client

After you install an independent hardware iSCSI adapter, you might need to change its default IP settings so that the adapter is configured properly for the IP SAN.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters**, and select the adapter (vmhba#) to configure.
- 4 Under Adapter Details, click the **Properties** tab and click **Edit** in the IP Address and DNS Configuration panel.
- 5 Change the default IP settings.

Work with your network administrator to determine the correct IP setting for the adapter.

## Configuring Dependent Hardware iSCSI Adapters

A dependent hardware iSCSI adapter is a third-party adapter that depends on VMware networking, and iSCSI configuration and management interfaces provided by VMware.

An example of a dependent iSCSI adapter is a Broadcom 5709 NIC. When installed on a host, it presents its two components, a standard network adapter and an iSCSI engine, to the same port. The iSCSI engine appears on the list of storage adapters as an iSCSI adapter (vmhba). Although the iSCSI adapter is enabled by default, to make it functional, you must first connect it, through a virtual VMkernel adapter (vmk), to a physical network adapter (vmnic) associated with it. You can then configure the iSCSI adapter.

After you configure the dependent hardware iSCSI adapter, the discovery and authentication data are passed through the network connection, while the iSCSI traffic goes through the iSCSI engine, bypassing the network.

The entire setup and configuration process for the dependent hardware iSCSI adapters involves these steps:

- 1 View the dependent hardware adapters.

See [View Dependent Hardware iSCSI Adapters in the vSphere Web Client](#).

If your dependent hardware adapters do not appear on the list of storage adapters, check whether they need to be licensed. See your vendor documentation.

- 2 Determine the association between the dependent hardware adapters and physical NICs.

See [Determine Association Between iSCSI and Network Adapters in the vSphere Web Client](#).

Make sure to note the names of the corresponding physical NICs. For example, the vmhba33 adapter corresponds to vmnic1 and vmhba34 corresponds to vmnic2.

- 3 Configure networking for iSCSI.



See [Setting Up iSCSI Network](#).

Configuring the network involves creating a VMkernel interface for each physical network adapter and associating the interface with an appropriate iSCSI adapter.

- 4 (Optional) Enable Jumbo Frames.

See [Using Jumbo Frames with iSCSI](#).

- 5 Configure discovery information.

See [Configuring Discovery Addresses for iSCSI Adapters](#).

- 6 (Optional) Configure CHAP parameters.

See [Configuring CHAP Parameters for iSCSI Adapters](#).

## Dependent Hardware iSCSI Considerations

When you use dependent hardware iSCSI adapters with ESXi, certain considerations apply.

- When you use any dependent hardware iSCSI adapter, performance reporting for a NIC associated with the adapter might show little or no activity, even when iSCSI traffic is heavy. This behavior occurs because the iSCSI traffic bypasses the regular networking stack.
- If you use a third-party virtual switch, for example Cisco Nexus 1000V DVS, disable automatic pinning. Use manual pinning instead, making sure to connect a VMkernel adapter (vmk) to an appropriate physical NIC (vmnic). For information, refer to your virtual switch vendor documentation.
- The Broadcom iSCSI adapter performs data reassembly in hardware, which has a limited buffer space. When you use the Broadcom iSCSI adapter in a congested network or under heavy load, enable flow control to avoid performance degradation.

Flow control manages the rate of data transmission between two nodes to prevent a fast sender from overrunning a slow receiver. For best results, enable flow control at the end points of the I/O path, at the hosts and iSCSI storage systems.

To enable flow control for the host, use the `esxcli system module parameters` command. For details, see the VMware knowledge base article at <http://kb.vmware.com/kb/1013413>

- Broadcom iSCSI adapters do not support IPv6.

## View Dependent Hardware iSCSI Adapters in the vSphere Web Client

View a dependent hardware iSCSI adapter to verify that it is correctly loaded.

If installed, the dependent hardware iSCSI adapter (vmhba#) appears on the list of storage adapters under such category as, for example, Broadcom iSCSI Adapter. If the dependent hardware adapter does not appear on the list of storage adapters, check whether it needs to be licensed. See your vendor documentation.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters**.
- 4 Select the adapter (vmhba#) to view.

The default details for the adapter appear, including the iSCSI name, iSCSI alias, and the status.

### What to do next

Although the dependent iSCSI adapter is enabled by default, to make it functional, you must set up networking for the iSCSI traffic and bind the adapter to the appropriate VMkernel iSCSI port. You then configure discovery addresses and CHAP parameters.

## Determine Association Between iSCSI and Network Adapters in the vSphere Web Client

You create network connections to bind dependent iSCSI and physical network adapters. To create the connections correctly, you must determine the name of the physical NIC with which the dependent hardware iSCSI adapter is associated.

### Prerequisites

In the vSphere Web Client, browse to the dependent hardware iSCSI adapter (vmhba#). See [View Dependent Hardware iSCSI Adapters in the vSphere Web Client](#).

### Procedure

- 1 Select the iSCSI adapter (vmhba#) and click the **Network Port Binding** tab under Adapter Details.
- 2 Click **Add**.

The network adapter (vmnic#) that corresponds to the dependent iSCSI adapter is listed in the Physical Network Adapter column.

### What to do next

If the VMkernel Adapter column is empty, create a VMkernel adapter (vmk#) for the physical network adapter (vmnic#) and then bind them to the associated dependent hardware iSCSI. See [Setting Up iSCSI Network](#).

## Configuring Software iSCSI Adapter

With the software-based iSCSI implementation, you can use standard NICs to connect your host to a remote iSCSI target on the IP network. The software iSCSI adapter that is built into ESXi facilitates this connection by communicating with the physical NICs through the network stack.

Before you can use the software iSCSI adapter, you must set up networking, activate the adapter, and configure parameters such as discovery addresses and CHAP.

---

**Note** Designate a separate network adapter for iSCSI. Do not use iSCSI on 100Mbps or slower adapters.

---

The software iSCSI adapter configuration workflow includes these steps:

- 1 Activate the software iSCSI adapter.

See [Activate the Software iSCSI Adapter in the vSphere Web Client](#).

- 2 Configure networking for iSCSI.

See [Setting Up iSCSI Network](#).

Configuring the network involves creating a VMkernel interface for each physical network adapter that you use for iSCSI and associating all interfaces with the software iSCSI adapter.

- 3 (Optional) Enable Jumbo Frames.

See [Using Jumbo Frames with iSCSI](#).

- 4 Configure discovery information.

See [Configuring Discovery Addresses for iSCSI Adapters](#).

- 5 (Optional) Configure CHAP parameters.

See [Configuring CHAP Parameters for iSCSI Adapters](#).

## Activate the Software iSCSI Adapter in the vSphere Web Client

You must activate your software iSCSI adapter so that your host can use it to access iSCSI storage.

You can activate only one software iSCSI adapter.

### Prerequisites

Required privilege: **Host.Configuration.Storage Partition Configuration**

---

**Note** If you boot from iSCSI using the software iSCSI adapter, the adapter is enabled and the network configuration is created at the first boot. If you disable the adapter, it is reenabled each time you boot the host.

---

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters**, and click **Add**.
- 4 Select **Software iSCSI Adapter** and confirm that you want to add the adapter.

## Results

The software iSCSI adapter (vmhba#) is enabled and appears on the list of storage adapters. After enabling the adapter, the host assigns the default iSCSI name to it. If you need to change the default name, follow iSCSI naming conventions.

## What to do next

Select the adapter and use the Adapter Details section to complete configuration.

## Disable Software iSCSI Adapter in the vSphere Web Client

If you do not need the software iSCSI adapter, you can disable it.

Disabling the software iSCSI adapter marks it for removal. The adapter is removed from the host on the next host reboot. After removal, all virtual machines and other data on the storage devices associated with this adapter become inaccessible to the host.

## Prerequisites

Required privilege: **Host.Configuration.Storage Partition Configuration**

## Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters** and select the software iSCSI adapter from the list.
- 4 Under Adapter Details, click **Disable** and confirm that you want to disable the adapter.

The status indicates that the adapter is disabled.

- 5 Reboot the host.

After reboot, the adapter no longer appears on the list of storage adapters.

## Results

The iSCSI software adapter is no longer available and storage devices associated with it are inaccessible. You can later activate the adapter.

## Modify General Properties for iSCSI Adapters

You can change the default iSCSI name and alias assigned to your iSCSI adapters. For the independent hardware iSCSI adapters, you can also change the default IP settings.

---

**Important** When you modify any default properties for your iSCSI adapters, make sure to correctly format their names and IP addresses.

---

## Prerequisites

Required privilege: **Host .Configuration.Storage Partition Configuration**

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters**, and select the adapter (vmhba#) to configure.
- 4 Under Adapter Details, click the **Properties** tab and click **Edit** in the General panel.
- 5 To change the default iSCSI name for your adapter, enter the new name.  
  
Make sure the name you enter is worldwide unique and properly formatted or some storage devices might not recognize the iSCSI adapter.
- 6 (Optional) Enter the iSCSI alias.  
  
The alias is a name that you use to identify the iSCSI adapter.

### Results

If you change the iSCSI name, it will be used for new iSCSI sessions. For existing sessions, new settings will not be used until logout and re-login.

## Setting Up iSCSI Network

Software and dependent hardware iSCSI adapters depend on VMkernel networking. If you use the software or dependent hardware iSCSI adapters, you must configure connections for the traffic between the iSCSI component and the physical network adapters.

Configuring the network connection involves creating a virtual VMkernel adapter for each physical network adapter. You then associate the VMkernel adapter with an appropriate iSCSI adapter. This process is called port binding.

For specific considerations on when and how to use network connections with software iSCSI, see the VMware knowledge base article at <http://kb.vmware.com/kb/2038869>.

## Multiple Network Adapters in iSCSI Configuration

If your host has more than one physical network adapter for software and dependent hardware iSCSI, use the adapters for multipathing.

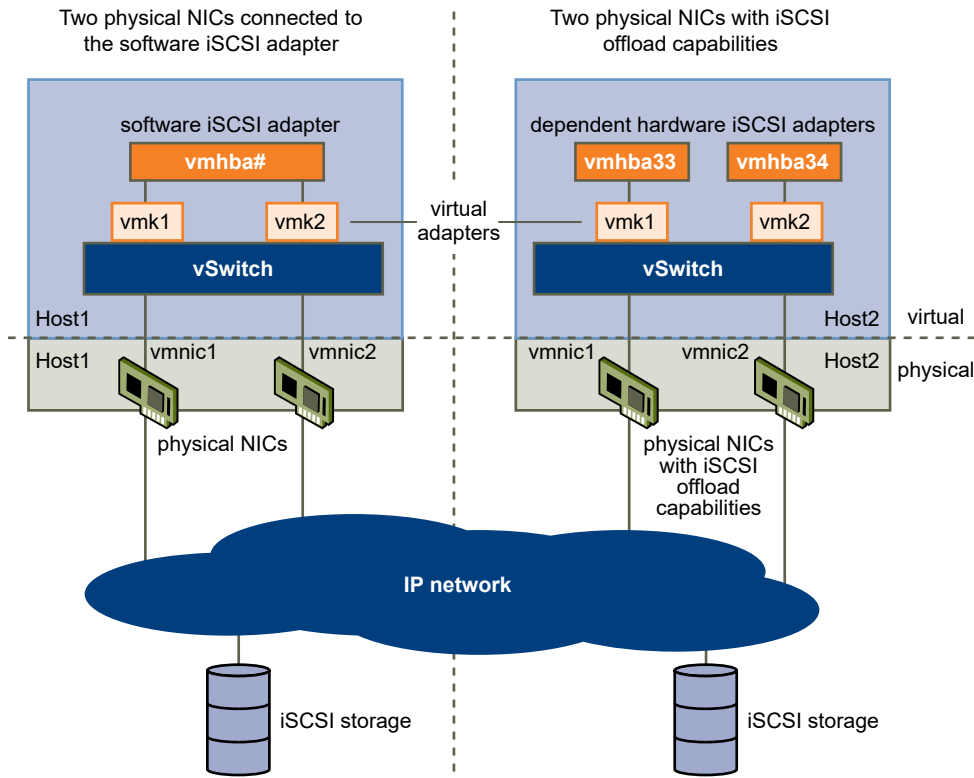
You can connect the software iSCSI adapter with any physical NICs available on your host. The dependent iSCSI adapters must be connected only to their own physical NICs.

---

**Note** Physical NICs must be on the same subnet as the iSCSI storage system they connect to.

---

Figure 11-1. Networking with iSCSI



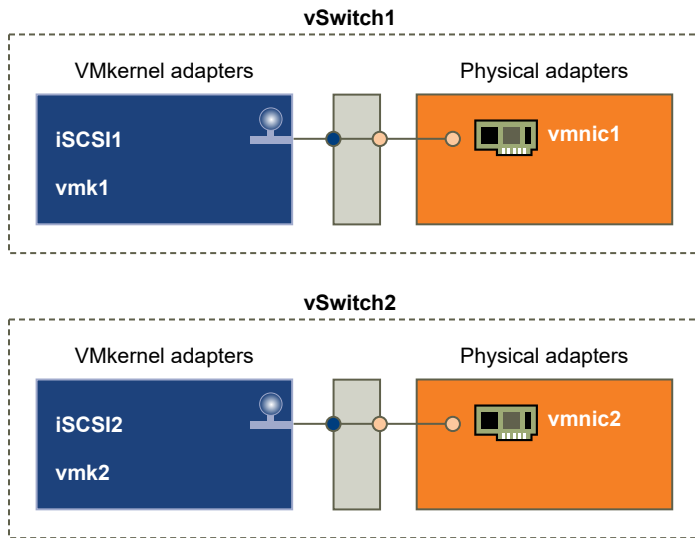
The iSCSI adapter and physical NIC connect through a virtual VMkernel adapter, also called virtual network adapter or VMkernel port. You create a VMkernel adapter (vmk) on a vSphere switch (vSwitch) using 1:1 mapping between each virtual and physical network adapter.

One way to achieve the 1:1 mapping when you have multiple NICs, is to designate a separate vSphere switch for each virtual-to-physical adapter pair.

**Note** If you use separate vSphere switches, you must connect them to different IP subnets. Otherwise, VMkernel adapters might experience connectivity problems and the host will fail to discover iSCSI LUNs.

The following examples show configurations that use vSphere standard switches, but you can use distributed switches as well. For more information about vSphere distributed switches, see the *vSphere Networking* documentation.

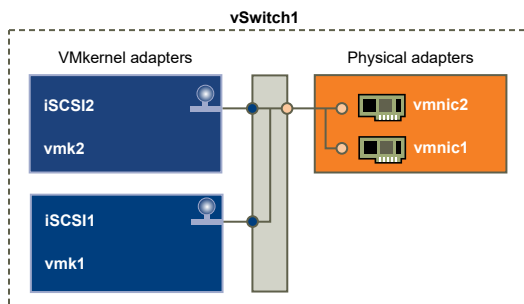
Figure 11-2. 1:1 adapter mapping on separate vSphere standard switches



An alternative is to add all NICs and VMkernel adapters to a single vSphere standard switch. In this case, you must override the default network setup and make sure that each VMkernel adapter maps to only one corresponding active physical adapter.

**Note** You must use the single vSwitch configuration if VMkernel adapters are on the same subnet.

Figure 11-3. 1:1 adapter mapping on a single vSphere standard switch



The following table summarises the iSCSI networking configuration discussed in this topic.

Table 11-2. Networking configuration for iSCSI

iSCSI Adapters	VMkernel Adapters (Ports)	Physical Adapters (NICs)
Software iSCSI		
vmhba32	vmk1	vmnic1
	vmk2	vmnic2
Dependent Hardware iSCSI		
vmhba33	vmk1	vmnic1
vmhba34	vmk2	vmnic2

## Guidelines for Using iSCSI Port Binding in ESXi

You can use multiple VMkernel adapters bound to iSCSI to have multiple paths to an iSCSI array that broadcasts a single IP address.

When you use port binding for multipathing, follow these guidelines:

- iSCSI ports of the array target must reside in the same broadcast domain and IP subnet as the VMkernel adapters.
- All VMkernel adapters used for iSCSI port binding must reside in the same broadcast domain and IP subnet.
- All VMkernel adapters used for iSCSI connectivity must reside in the same virtual switch.
- Port binding does not support network routing.

Do not use port binding when any of the following conditions exist:

- Array target iSCSI ports are in a different broadcast domain and IP subnet.
- VMkernel adapters used for iSCSI connectivity exist in different broadcast domains, IP subnets, or use different virtual switches.
- Routing is required to reach the iSCSI array.

## Create Network Connections for iSCSI in the vSphere Web Client

Configure connections for the traffic between the software or dependent hardware iSCSI adapters and the physical network adapters.

The following tasks discuss the iSCSI network configuration with a vSphere standard switch.

If you use a vSphere distributed switch with multiple uplink ports, for port binding, create a separate distributed port group per each physical NIC. Then set the team policy so that each distributed port group has only one active uplink port. For detailed information on vSphere distributed switches, see the *vSphere Networking* documentation.

### Procedure

#### 1 [Create a Single VMkernel Adapter for iSCSI in the vSphere Web Client](#)

Connect the VMkernel, which runs services for iSCSI storage, to a physical network adapter.

#### 2 [Create Additional VMkernel Adapters for iSCSI in the vSphere Web Client](#)

Use this task if you have two or more physical network adapters for iSCSI and you want to connect all of your physical adapters to a single vSphere standard switch. In this task, you add the physical adapters and VMkernel adapters to an existing vSphere standard switch.

#### 3 [Change Network Policy for iSCSI in the vSphere Web Client](#)

If you use a single vSphere standard switch to connect multiple VMkernel adapters to multiple network adapters, set up network policy so that only one physical network adapter is active for each VMkernel adapter.



#### 4 [Bind iSCSI and VMkernel Adapters in the vSphere Web Client](#)

Bind an iSCSI adapter with a VMkernel adapter.

#### 5 [Review Port Binding Details in the vSphere Web Client](#)

Review networking details of the VMkernel adapter that is bound to the iSCSI adapter.

### Create a Single VMkernel Adapter for iSCSI in the vSphere Web Client

Connect the VMkernel, which runs services for iSCSI storage, to a physical network adapter.

#### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click **Actions**, and select **All vCenter Actions > Add Networking**.
- 3 Select **VMkernel Network Adapter**, and click **Next**.
- 4 Select **New standard switch** to create a vSphere standard switch.
- 5 Click the **Add adapters** icon, and select the network adapter (vmnic#) to use for iSCSI.

Make sure to assign the adapter to Active Adapters.

---

**Important** If you are creating a VMkernel adapter for dependent hardware iSCSI, select the network adapter that corresponds to the iSCSI component. See [Determine Association Between iSCSI and Network Adapters in the vSphere Web Client](#).

---

- 6 Enter a network label.

A network label is a friendly name that identifies the VMkernel adapter that you are creating, for example, iSCSI.

- 7 Specify the IP settings.
- 8 Review the information and click **Finish**.

#### Results

You created the virtual VMkernel adapter (vmk#) for a physical network adapter (vmnic#) on your host.

#### What to do next

If your host has one physical network adapter for iSCSI traffic, you must bind the virtual adapter that you created to the iSCSI adapter.

If you have multiple network adapters, create additional VMkernel adapters and then perform iSCSI binding. The number of virtual adapters must correspond to the number of physical adapters on the host.

## Create Additional VMkernel Adapters for iSCSI in the vSphere Web Client

Use this task if you have two or more physical network adapters for iSCSI and you want to connect all of your physical adapters to a single vSphere standard switch. In this task, you add the physical adapters and VMkernel adapters to an existing vSphere standard switch.

### Prerequisites

Create a vSphere standard switch that maps an iSCSI VMkernel adapter to a single physical network adapter designated for iSCSI traffic.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Networking**.
- 3 Click **Virtual Switches**, and select the vSphere switch that you want to modify from the list.
- 4 Connect additional network adapters to the switch.

- a Click the **Add host networking** icon.
- b Select **Physical Network Adapters**, and click **Next**.
- c Make sure that you are using the existing switch, and click **Next**.
- d Click the **Add adapters** icon, and select one or more network adapters (vmnic#) to use for iSCSI.

With dependent hardware iSCSI adapters, select only those NICs that have a corresponding iSCSI component.

- e Complete configuration, and click **Finish**.
- 5 Create iSCSI VMkernel adapters for all physical network adapters that you added.

The number of VMkernel interfaces must correspond to the number of physical network adapters on the vSphere standard switch.

- a Click the **Add host networking** icon.
- b Select **VMkernel Network Adapter**, and click **Next**.
- c Make sure that you are using the existing switch, and click **Next**.
- d Complete configuration, and click **Finish**.

### What to do next

Change the network policy for all VMkernel adapters, so that only one physical network adapter is active for each VMkernel adapter. You can then bind the iSCSI VMkernel adapters to the software iSCSI or dependent hardware iSCSI adapters.

## Change Network Policy for iSCSI in the vSphere Web Client

If you use a single vSphere standard switch to connect multiple VMkernel adapters to multiple network adapters, set up network policy so that only one physical network adapter is active for each VMkernel adapter.

By default, for each VMkernel adapter on the vSphere standard switch, all network adapters appear as active. You must override this setup, so that each VMkernel adapter maps to only one corresponding active physical. For example, vmk1 maps to vmnic1, vmk2 maps to vmnic2, and so on.

### Prerequisites

Create a vSphere standard switch that connects VMkernel with physical network adapters designated for iSCSI traffic. The number of VMkernel adapters must correspond to the number of physical adapters on the vSphere standard switch.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Networking**.
- 3 Click **Virtual Switches**, and select the vSphere switch that you want to modify from the list.
- 4 On the vSwitch diagram, select the VMkernel adapter and click the **Edit Settings** icon.
- 5 On the **Edit Settings** wizard, click **Teaming and Failover** and click **Override** under Failover Order.
- 6 Designate only one physical adapter as active and move all remaining adapters to the **Unused Adapters** category.
- 7 Repeat [Step 4](#) through [Step 6](#) for each iSCSI VMkernel interface on the vSphere standard switch.

### Example: iSCSI Network Policy

The following table illustrates the proper iSCSI mapping where only one physical network adapter is active for each VMkernel adapter.

VMkernel Adapter (vmk#)	Physical Network Adapter (vmnic#)
vmk1	<b>Active Adapters</b>
	vmnic1
	<b>Unused Adapters</b>
	vmnic2
vmk2	<b>Active Adapters</b>
	vmnic2
	<b>Unused Adapters</b>
	vmnic1

## What to do next

After you perform this task, bind the virtual VMkernel adapters to the software iSCSI or dependent hardware iSCSI adapters.

## Bind iSCSI and VMkernel Adapters in the vSphere Web Client

Bind an iSCSI adapter with a VMkernel adapter.

### Prerequisites

Create a virtual VMkernel adapter for each physical network adapter on your host. If you use multiple VMkernel adapters, set up the correct network policy.

Required privilege: **Host.Configuration.Storage Partition Configuration**

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters** and select the software or dependent iSCSI adapter to configure from the list.
- 4 Under Adapter Details, click the **Network Port Binding** tab and click **Add**.
- 5 Select a VMkernel adapter to bind with the iSCSI adapter.

---

**Note** Make sure that the network policy for the VMkernel adapter is compliant with the binding requirements.

---

You can bind the software iSCSI adapter to one or more VMkernel adapters. For a dependent hardware iSCSI adapter, only one VMkernel adapter associated with the correct physical NIC is available.

- 6 Click **OK**.

### Results

The network connection appears on the list of VMkernel port bindings for the iSCSI adapter.

## Review Port Binding Details in the vSphere Web Client

Review networking details of the VMkernel adapter that is bound to the iSCSI adapter.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters** and select the software or dependent iSCSI adapter from the list.
- 4 Under Adapter Details, click the **Network Port Binding** tab and click **View Details**.

- 5 Review the VMkernel adapter information by switching between available tabs.

## Managing iSCSI Network

Special consideration apply to network adapters, both physical and VMkernel, that are associated with an iSCSI adapter.

After you create network connections for iSCSI, an iSCSI indicator on a number of Networking dialog boxes becomes enabled. This indicator shows that a particular virtual or physical network adapter is iSCSI-bound. To avoid disruptions in iSCSI traffic, follow these guidelines and considerations when managing iSCSI-bound virtual and physical network adapters:

- Make sure that the VMkernel network adapters are assigned addresses on the same subnet as the iSCSI storage portal they connect to.
- iSCSI adapters using VMkernel adapters are not able to connect to iSCSI ports on different subnets, even if those ports are discovered by the iSCSI adapters.
- When using separate vSphere switches to connect physical network adapters and VMkernel adapters, make sure that the vSphere switches connect to different IP subnets.
- If VMkernel adapters are on the same subnet, they must connect to a single vSwitch.
- If you migrate VMkernel adapters to a different vSphere switch, move associated physical adapters.
- Do not make configuration changes to iSCSI-bound VMkernel adapters or physical network adapters.
- Do not make changes that might break association of VMkernel adapters and physical network adapters. You can break the association if you remove one of the adapters or the vSphere switch that connects them, or change the 1:1 network policy for their connection.

## iSCSI Network Troubleshooting

A warning sign indicates non-compliant port group policy for an iSCSI-bound VMkernel adapter.

### Problem

The VMkernel adapter's port group policy is considered non-compliant in the following cases:

- The VMkernel adapter is not connected to an active physical network adapter.
- The VMkernel adapter is connected to more than one physical network adapter.
- The VMkernel adapter is connected to one or more standby physical adapters.
- The active physical adapter is changed.

### Solution

Follow the steps in [Change Network Policy for iSCSI in the vSphere Web Client](#) to set up the correct network policy for the iSCSI-bound VMkernel adapter.

## Using Jumbo Frames with iSCSI

ESXi supports the use of Jumbo Frames with iSCSI.

Jumbo Frames are Ethernet frames with the size that exceeds 1500 Bytes. The maximum transmission unit (MTU) parameter is typically used to measure the size of Jumbo Frames. ESXi allows Jumbo Frames with the MTU up to 9000 Bytes.

When you use Jumbo Frames for iSCSI traffic, the following considerations apply:

- The network must support Jumbo Frames end-to-end for Jumbo Frames to be effective.
- Check with your vendors to ensure your physical NICs and iSCSI HBAs support Jumbo Frames.
- To set up and verify physical network switches for Jumbo Frames, consult your vendor documentation.

The following table explains the level of support that ESXi provides to Jumbo Frames.

**Table 11-3. Support of Jumbo Frames**

Type of iSCSI Adapters	Jumbo Frames Support
Software iSCSI	Supported
Dependent Hardware iSCSI	Supported. Check with vendor.
Independent Hardware iSCSI	Supported. Check with vendor.

## Enable Jumbo Frames for Software and Dependent Hardware iSCSI

To enable Jumbo Frames for software and dependent hardware iSCSI adapters in the vSphere Web Client, change the default value of the maximum transmission units (MTU) parameter.

You change the MTU parameter on the vSphere switch that you use for iSCSI traffic. For more information, see the *vSphere Networking* documentation.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Networking**.
- 3 Click **Virtual Switches**, and select the vSphere switch that you want to modify from the list.
- 4 Click the **Edit Settings** icon.
- 5 On the Properties page, change the MTU parameter.

This step sets the MTU for all physical NICs on that standard switch. The MTU value should be set to the largest MTU size among all NICs connected to the standard switch. ESXi supports the MTU size up to 9000 Bytes.

## Enable Jumbo Frames for Independent Hardware iSCSI

To enable Jumbo Frames for independent hardware iSCSI adapters in the vSphere Web Client, change the default value of the maximum transmission units (MTU) parameter.

Use the Advanced Options settings to change the MTU parameter for the iSCSI HBA.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters**, and select the independent hardware iSCSI adapter from the list of adapters.
- 4 Under Adapter Details, click the **Advanced Options** tab and click **Edit**.
- 5 Change the value of the MTU parameter.

ESXi supports the MTU size up to 9000 Bytes.

## Configuring Discovery Addresses for iSCSI Adapters

Set up target discovery addresses so that the iSCSI adapter can determine which storage resource on the network is available for access.

The ESXi system supports these discovery methods:

### Dynamic Discovery

Also known as SendTargets discovery. Each time the initiator contacts a specified iSCSI server, the initiator sends the SendTargets request to the server. The server responds by supplying a list of available targets to the initiator. The names and IP addresses of these targets appear on the **Static Discovery** tab. If you remove a static target added by dynamic discovery, the target might be returned to the list the next time a rescan happens, the HBA is reset, or the host is rebooted.

### Static Discovery

The initiator does not have to perform any discovery. The initiator has a list of targets it can contact and uses their IP addresses and target names to communicate with them.

## Set Up Dynamic Discovery in the vSphere Web Client

With Dynamic Discovery, each time the initiator contacts a specified iSCSI storage system, it sends the SendTargets request to the system. The iSCSI system responds by supplying a list of available targets to the initiator.

When you set up Dynamic Discovery, you can only add a new iSCSI system. You cannot change the IP address, DNS name, or port number of an existing iSCSI system. To make changes, remove the existing system and add a new one.

### Prerequisites

Required privilege: **Host.Configuration.Storage Partition Configuration**

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters** and select the iSCSI adapter to configure from the list.
- 4 Under Adapter Details, click the **Targets** tab and click **Dynamic Discovery**.
- 5 Click **Add**.
- 6 Type the IP address or DNS name of the storage system and click **OK**.
- 7 Rescan the iSCSI adapter.

### Results

After establishing the SendTargets session with the iSCSI system, you host populates the Static Discovery list with all newly discovered targets.

## Set Up Static Discovery in the vSphere Web Client

With iSCSI initiators, in addition to the dynamic discovery method, you can use static discovery and manually enter information for the targets.

When you set up Static Discovery, you can only add new iSCSI targets. You cannot change the IP address, DNS name, iSCSI target name, or port number of an existing target. To make changes, remove the existing target and add a new one.

### Prerequisites

Required privilege: **Host.Configuration.Storage Partition Configuration**

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters** and select the iSCSI adapter to configure from the list.
- 4 Under Adapter Details, click the **Targets** and click **Static Discovery**.
- 5 Click **Add**.
- 6 Enter the target's information and click **OK**.
- 7 Rescan the iSCSI adapter.

## Remove Static Targets in the vSphere Web Client

Remove targets that appear on the Static Discovery list.



## Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters** and select the iSCSI adapter to modify from the list.
- 4 Under Adapter Details, click the **Targets** and click **Static Discovery**.
- 5 Select the static target to remove and click **Remove**.
- 6 Rescan the iSCSI adapter.

If the static target was dynamically discovered, you need to remove it from the storage system before performing the rescan. Otherwise, your host will automatically discover and add the target to the list of static targets when you rescan the adapter.

## Configuring CHAP Parameters for iSCSI Adapters

Because the IP networks that the iSCSI technology uses to connect to remote targets do not protect the data they transport, you must ensure security of the connection. One of the protocols that iSCSI implements is the Challenge Handshake Authentication Protocol (CHAP), which verifies the legitimacy of initiators that access targets on the network.

CHAP uses a three-way handshake algorithm to verify the identity of your host and, if applicable, of the iSCSI target when the host and target establish a connection. The verification is based on a predefined private value, or CHAP secret, that the initiator and target share.

ESXi supports CHAP authentication at the adapter level. In this case, all targets receive the same CHAP name and secret from the iSCSI initiator. For software and dependent hardware iSCSI adapters, ESXi also supports per-target CHAP authentication, which allows you to configure different credentials for each target to achieve greater level of security.

## Choosing CHAP Authentication Method

ESXi supports unidirectional CHAP for all types of iSCSI initiators, and bidirectional CHAP for software and dependent hardware iSCSI.

Before configuring CHAP, check whether CHAP is enabled at the iSCSI storage system and check the CHAP authentication method the system supports. If CHAP is enabled, enable it for your initiators, making sure that the CHAP authentication credentials match the credentials on the iSCSI storage.

ESXi supports the following CHAP authentication methods:

### Unidirectional CHAP

In unidirectional CHAP authentication, the target authenticates the initiator, but the initiator does not authenticate the target.

### Bidirectional CHAP

In bidirectional CHAP authentication, an additional level of security enables the initiator to authenticate the target. VMware supports this method for software and dependent hardware iSCSI adapters only.

For software and dependent hardware iSCSI adapters, you can set unidirectional CHAP and bidirectional CHAP for each adapter or at the target level. Independent hardware iSCSI supports CHAP only at the adapter level.

When you set the CHAP parameters, specify a security level for CHAP.

**Note** When you specify the CHAP security level, how the storage array responds depends on the array's CHAP implementation and is vendor specific. For information on CHAP authentication behavior in different initiator and target configurations, consult the array documentation.

**Table 11-4. CHAP Security Level**

CHAP Security Level	Description	Supported
None	The host does not use CHAP authentication. Select this option to disable authentication if it is currently enabled.	Software iSCSI Dependent hardware iSCSI Independent hardware iSCSI
Use unidirectional CHAP if required by target	The host prefers a non-CHAP connection, but can use a CHAP connection if required by the target.	Software iSCSI Dependent hardware iSCSI
Use unidirectional CHAP unless prohibited by target	The host prefers CHAP, but can use non-CHAP connections if the target does not support CHAP.	Software iSCSI Dependent hardware iSCSI Independent hardware iSCSI
Use unidirectional CHAP	The host requires successful CHAP authentication. The connection fails if CHAP negotiation fails.	Software iSCSI Dependent hardware iSCSI Independent hardware iSCSI
Use bidirectional CHAP	The host and the target support bidirectional CHAP.	Software iSCSI Dependent hardware iSCSI

## Set Up CHAP for iSCSI Adapter in the vSphere Web Client

When you set up CHAP name and secret at the iSCSI adapter level, all targets receive the same parameters from the adapter. By default, all discovery addresses or static targets inherit CHAP parameters that you set up at the adapter level.

The CHAP name should not exceed 511 alphanumeric characters and the CHAP secret should not exceed 255 alphanumeric characters. Some adapters, for example the QLogic adapter, might have lower limits, 255 for the CHAP name and 100 for the CHAP secret.

### Prerequisites

- Before setting up CHAP parameters for software or dependent hardware iSCSI, determine whether to configure unidirectional or bidirectional CHAP. Independent hardware iSCSI adapters do not support bidirectional CHAP.

- Verify CHAP parameters configured on the storage side. Parameters that you configure must match the ones on the storage side.
- Required privilege: **Host.Configuration.Storage Partition Configuration**

#### Procedure

- 1 Display storage adapters and select the iSCSI adapter to configure.
- 2 Under Adapter Details, click the **Properties** tab and click **Edit** in the Authentication panel.
- 3 Specify authentication method.
  - **None**
  - **Use unidirectional CHAP if required by target**
  - **Use unidirectional CHAP unless prohibited by target**
  - **Use unidirectional CHAP**
  - **Use bidirectional CHAP.** To configure bidirectional CHAP, you must select this option.
- 4 Specify the outgoing CHAP name.
 

Make sure that the name you specify matches the name configured on the storage side.

  - To set the CHAP name to the iSCSI adapter name, select **Use initiator name**.
  - To set the CHAP name to anything other than the iSCSI initiator name, deselect **Use initiator name** and type a name in the **Name** text box.
- 5 Enter an outgoing CHAP secret to be used as part of authentication. Use the same secret that you enter on the storage side.
- 6 If configuring bidirectional CHAP, specify incoming CHAP credentials.
 

Make sure to use different secrets for the outgoing and incoming CHAP.
- 7 Click **OK**.
- 8 Rescan the iSCSI adapter.

#### Results

If you change the CHAP parameters, they are used for new iSCSI sessions. For existing sessions, new settings are not used until you log out and log in again.

## Set Up CHAP for Target in the vSphere Web Client

If you use software and dependent hardware iSCSI adapters, you can configure different CHAP credentials for each discovery address or static target.

The CHAP name should not exceed 511 and the CHAP secret 255 alphanumeric characters.

## Prerequisites

- Before setting up CHAP parameters for software or dependent hardware iSCSI, determine whether to configure unidirectional or bidirectional CHAP.
- Verify CHAP parameters configured on the storage side. Parameters that you configure must match the ones on the storage side.
- Access storage adapters.
- Required privilege: **Host.Configuration.Storage Partition Configuration**

## Procedure

- 1 Select the iSCSI adapter to configure, and click the **Targets** tab under Adapter Details.
- 2 Click either **Dynamic Discovery** or **Static Discovery**.
- 3 From the list of available targets, select a target to configure and click **Authentication**.
- 4 Deselect **Inherit settings from parent** and specify authentication method.
  - **None**
  - **Use unidirectional CHAP if required by target**
  - **Use unidirectional CHAP unless prohibited by target**
  - **Use unidirectional CHAP**
  - **Use bidirectional CHAP.** To configure bidirectional CHAP, you must select this option.
- 5 Specify the outgoing CHAP name.
 

Make sure that the name you specify matches the name configured on the storage side.

  - To set the CHAP name to the iSCSI adapter name, select **Use initiator name**.
  - To set the CHAP name to anything other than the iSCSI initiator name, deselect **Use initiator name** and type a name in the **Name** text box.
- 6 Enter an outgoing CHAP secret to be used as part of authentication. Use the same secret that you enter on the storage side.
- 7 If configuring bi-directional CHAP, specify incoming CHAP credentials.
 

Make sure to use different secrets for the outgoing and incoming CHAP.
- 8 Click **OK**.
- 9 Rescan the iSCSI adapter.

## Results

If you change the CHAP parameters, they are used for new iSCSI sessions. For existing sessions, new settings are not used until you log out and login again.

## Disable CHAP

You can disable CHAP if your storage system does not require it.

If you disable CHAP on a system that requires CHAP authentication, existing iSCSI sessions remain active until you reboot your host, end the session through the command line, or the storage system forces a logout. After the session ends, you can no longer connect to targets that require CHAP.

Required privilege: **Host.Configuration.Storage Partition Configuration**

### Procedure

- 1 Open the CHAP Credentials dialog box.
- 2 For software and dependent hardware iSCSI adapters, to disable just the mutual CHAP and leave the one-way CHAP, select **Do not use CHAP** in the Mutual CHAP area.
- 3 To disable one-way CHAP, select **Do not use CHAP** in the CHAP area.  
 The mutual CHAP, if set up, automatically turns to **Do not use CHAP** when you disable the one-way CHAP.
- 4 Click **OK**.

## Configuring Advanced Parameters for iSCSI

You might need to configure additional parameters for your iSCSI initiators. For example, some iSCSI storage systems require ARP (Address Resolution Protocol) redirection to move iSCSI traffic dynamically from one port to another. In this case, you must activate ARP redirection on your host.

The following table lists advanced iSCSI parameters that you can configure using the vSphere Web Client. In addition, you can use the vSphere CLI commands to configure some of the advanced parameters. For information, see the *Getting Started with vSphere Command-Line Interfaces* documentation.

---

**Important** Do not make any changes to the advanced iSCSI settings unless you are directed by VMware support or Storage Vendors.

---

Table 11-5. Additional Parameters for iSCSI Initiators

Advanced Parameter	Description	Configurable On
Header Digest	Increases data integrity. When header digest is enabled, the system performs a checksum over each iSCSI Protocol Data Unit's (PDU's) header part and verifies using the CRC32C algorithm.	Software iSCSI Dependent Hardware iSCSI
Data Digest	Increases data integrity. When data digest is enabled, the system performs a checksum over each PDU's data part and verifies using the CRC32C algorithm.  <b>Note</b> Systems that use Intel Nehalem processors offload the iSCSI digest calculations for software iSCSI, thus reducing the impact on performance.	Software iSCSI Dependent Hardware iSCSI
Maximum Outstanding R2T	Defines the R2T (Ready to Transfer) PDUs that can be in transition before an acknowledge PDU is received.	Software iSCSI Dependent Hardware iSCSI
First Burst Length	Specifies the maximum amount of unsolicited data an iSCSI initiator can send to the target during the execution of a single SCSI command, in bytes.	Software iSCSI Dependent Hardware iSCSI
Maximum Burst Length	Maximum SCSI data payload in a Data-In or a solicited Data-Out iSCSI sequence, in bytes.	Software iSCSI Dependent Hardware iSCSI
Maximum Receive Data Segment Length	Maximum data segment length, in bytes, that can be received in an iSCSI PDU.	Software iSCSI Dependent Hardware iSCSI
Session Recovery Timeout	Specifies the amount of time, in seconds, that can lapse while a session recovery is performed. If the timeout exceeds its limit, the iSCSI initiator terminates the session.	Software iSCSI Dependent Hardware iSCSI
No-Op Interval	Specifies the time interval, in seconds, between NOP-Out requests sent from your iSCSI initiator to an iSCSI target. The NOP-Out requests serve as the ping mechanism to verify that a connection between the iSCSI initiator and the iSCSI target is active.	Software iSCSI Dependent Hardware iSCSI
No-Op Timeout	Specifies the amount of time, in seconds, that can lapse before your host receives a NOP-In message. The message is sent by the iSCSI target in response to the NOP-Out request. When the no-op timeout limit is exceeded, the initiator terminates the current session and starts a new one.	Software iSCSI Dependent Hardware iSCSI
ARP Redirect	Allows storage systems to move iSCSI traffic dynamically from one port to another. ARP is required by storage systems that do array-based failover.	Software iSCSI Dependent Hardware iSCSI Independent Hardware iSCSI
Delayed ACK	Allows systems to delay acknowledgment of received data packets.	Software iSCSI Dependent Hardware iSCSI

## Configure Advanced Parameters for iSCSI in the vSphere Web Client

The advanced iSCSI settings control such parameters as header and data digest, ARP redirection, delayed ACK, and so on.

---

**Caution** Do not make any changes to the advanced iSCSI settings unless you are working with the VMware support team or otherwise have thorough information about the values to provide for the settings.

---

### Prerequisites

Required privilege: **Host.Configuration.Storage Partition Configuration**

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters**, and select the adapter (vmhba#) to configure.
- 4 Configure advanced parameters.
  - To configure advanced parameters at the adapter level, under Adapter Details, click the **Advanced Options** tab and click **Edit**.
  - Configure advanced parameters at the target level.
    - a Click the **Targets** tab and click either **Dynamic Discovery** or **Static Discovery**.
    - b From the list of available targets, select a target to configure and click **Advanced Options**.
- 5 Enter any required values for the advanced parameters you want to modify.

## iSCSI Session Management

To communicate with each other, iSCSI initiators and targets establish iSCSI sessions. You can review and manage iSCSI sessions using vSphere CLI.

By default, software iSCSI and dependent hardware iSCSI initiators start one iSCSI session between each initiator port and each target port. If your iSCSI initiator or target have more than one port, your host can have multiple sessions established. The default number of sessions for each target equals the number of ports on the iSCSI adapter times the number of target ports.

Using vSphere CLI, you can display all current sessions to analyze and debug them. To create more paths to storage systems, you can increase the default number of sessions by duplicating existing sessions between the iSCSI adapter and target ports.

You can also establish a session to a specific target port. This can be useful if your host connects to a single-port storage system that, by default, presents only one target port to your initiator, but can redirect additional sessions to a different target port. Establishing a new session between your iSCSI initiator and another target port creates an additional path to the storage system.

The following considerations apply to iSCSI session management:

- Some storage systems do not support multiple sessions from the same initiator name or endpoint. Attempts to create multiple sessions to such targets can result in unpredictable behavior of your iSCSI environment.
- Storage vendors can provide automatic session managers. Using the automatic session manages to add or delete sessions, does not guarantee lasting results and can interfere with the storage performance.

## Review iSCSI Sessions

Use the vCLI command to display iSCSI sessions between an iSCSI adapter and a storage system.

In the procedure, **--server=server\_name** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

### Procedure

- ◆ To list iSCSI sessions, run the following command:

```
esxcli --server=server_name iscsi session list
```

The command takes these options:

Option	Description
<b>-A --adapter=str</b>	The iSCSI adapter name, for example, vmhba34.
<b>-s --isid=str</b>	The iSCSI session identifier.
<b>-n --name=str</b>	The iSCSI target name, for example, iqn.X.

## Add iSCSI Sessions

Use the vCLI to add an iSCSI session for a target you specify or to duplicate an existing session. By duplicating sessions, you increase the default number of sessions and create additional paths to storage systems.

In the procedure, **--server=server\_name** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.



## Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

## Procedure

- ◆ To add or duplicate an iSCSI session, run the following command:

```
esxcli --server=server_name iscsi session add
```

The command takes these options:

Option	Description
<b>-A --adapter=<i>str</i></b>	The iSCSI adapter name, for example, vmhba34. This option is required.
<b>-s --isid=<i>str</i></b>	The ISID of a session to duplicate. You can find it by listing all sessions.
<b>-n --name=<i>str</i></b>	The iSCSI target name, for example, iqn.X.

## What to do next

Rescan the iSCSI adapter.

## Remove iSCSI Sessions

Use the vCLI command to remove an iSCSI session between an iSCSI adapter and a target.

In the procedure, **--server=*server\_name*** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

## Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

## Procedure

- ◆ To remove a session, run the following command:

```
esxcli --server=server_name iscsi session remove
```

The command takes these options:

Option	Description
<b>-A --adapter=<i>str</i></b>	The iSCSI adapter name, for example, vmhba34. This option is required.
<b>-s --isid=<i>str</i></b>	The ISID of a session to remove. You can find it by listing all session.
<b>-n --name=<i>str</i></b>	The iSCSI target name, for example, iqn.X.

### What to do next

Rescan the iSCSI adapter.

# Modifying iSCSI Storage Systems for ESXi

# 12

After you configure your iSCSI initiators and storage, you might need to modify your storage system to ensure that it works properly with your ESXi implementation.

This section discusses many of the iSCSI storage systems supported in conjunction with VMware ESXi. For each device, it lists major known potential issues, points to vendor-specific information (if available), or includes information from VMware knowledge base articles.

---

**Note** Information in this section is updated only with each release. New information might already be available. Also, other iSCSI storage systems are supported but are not covered in this chapter. Consult the most recent *Storage/SAN Compatibility*, check with your storage vendor, and explore the VMware knowledge base articles.

---

This chapter includes the following topics:

- [Testing ESXi iSCSI SAN Configurations](#)
- [General Considerations for iSCSI SAN Storage Systems](#)
- [EMC CLARiiON Storage Systems](#)
- [EMC Symmetrix Storage Systems](#)
- [Enable HP StorageWorks MSA1510i to Communicate with ESXi](#)
- [HP StorageWorks EVA Storage Systems](#)
- [NetApp Storage Systems](#)
- [Dell EqualLogic Storage Systems](#)
- [HP StorageWorks SAN/iQ Storage Systems](#)
- [iSCSI Targets in vApps](#)

## Testing ESXi iSCSI SAN Configurations

ESXi supports a variety of SAN storage systems in different configurations. Generally, VMware tests ESXi with supported storage systems for basic connectivity, HBA failover, and so on.

Not all storage devices are certified for all features and capabilities of ESXi, and vendors might have specific positions of support with regard to ESXi.

VMware tests ESXi with storage systems in the following configurations:

### **Basic Connectivity**

Tests whether ESXi can recognize and operate with the storage system. This configuration does not allow for multipathing or any type of failover.

### **iSCSI Failover**

The server is equipped with multiple iSCSI HBAs or NICs. The server is robust to HBA or NIC failure.

### **Storage Port Failover**

The server is attached to multiple storage ports and is robust to storage port failures and switch failures.

### **Bootting from a SAN**

The host successfully boots from a LUN configured on the iSCSI SAN.

## **General Considerations for iSCSI SAN Storage Systems**

When you prepare your iSCSI SAN storage system to work with ESXi, you need to follow specific general requirements that apply to all storage systems.

For all storage systems, the following general requirements exist:

- LUNs must be presented to each HBA of each host with the same LUN ID number. If different numbers are used, the ESXi hosts do not recognize different paths to the same LUN. Because instructions on how to configure identical SAN LUN IDs are vendor-specific, consult your storage documentation for more information.
- Unless specified for individual storage systems discussed in this chapter, set the host type for LUNs presented to ESXi to `Linux` or `Linux Cluster`, if applicable to your storage system. The method ESXi uses to access the storage system is most compatible with Linux access, however, this can vary depending on the storage system you are using.
- If you are using vMotion, DRS, or HA, make sure that source and target hosts for virtual machines can see the same LUNs with identical LUN IDs. SAN administrators might find it counterintuitive to have multiple hosts see the same LUNs because they might be concerned about data corruption. However, VMFS prevents multiple virtual machines from writing to the same file at the same time, so provisioning the LUNs to all required ESXi system is appropriate.
- If you do not have CHAP authentication set up on the LUNs that are being accessed, you must also disable CHAP on the ESXi host. Otherwise, authentication of the storage system fails, although the LUNs have no CHAP requirement.

## EMC CLARiiON Storage Systems

EMC CLARiiON storage systems work with ESXi hosts in iSCSI SAN configurations. Generally, you use the EMC software to perform configurations.

This is an active-passive disk array, so any related issues that apply to all active-passive disk arrays are relevant. In addition, keep in mind the following:

- To avoid the possibility of path thrashing, the default multipathing policy is Most Recently Used, not Fixed. The ESXi system sets the default policy when it identifies the storage system.
- To boot from a SAN, choose the active storage processor for the boot LUN's target in the HBA BIOS.
- For ESXi to support EMC CLARiiON with ALUA, check the HCLs to make sure that you use the correct firmware version on the storage array. For additional information, contact your storage vendor.
- Certain EMC CLARiiON CX3 models might not behave appropriately during periods of network congestion. For information on how to address this behavior, see the VMware knowledge base article at <http://kb.vmware.com/kb/1002598>.

## EMC Symmetrix Storage Systems

To work with ESXi, EMC Symmetrix storage systems require certain specific settings. Use EMC software to configure the storage system.

The following settings are required for ESXi operations on the Symmetrix networked storage system:

- Common serial number (C)
- Auto negotiation (EAN) enabled
- SCSI 3 (SC3) set (enabled)
- Unique world wide name (UWN)
- SPC-2 (Decal) (SPC2) SPC-2 flag is required

---

**Note** The ESXi host considers any LUNs from a Symmetrix storage system that have a capacity of 50MB or less as management LUNs. These LUNs are also known as pseudo or gatekeeper LUNs. These LUNs appear in the EMC Symmetrix Management Interface and should not be used to hold data.

---

# Enable HP StorageWorks MSA1510i to Communicate with ESXi

This section describes the setup and configuration steps needed to allow an HP StorageWorks MSA1510i storage system to communicate with ESXi hosts.

## Procedure

- 1 Install, connect, and power up the network devices as detailed in the vendor installation document.
- 2 Obtain the IP address assigned to the MSA1510i controller management port.
  - a Scroll through the messages on the LCD panel until the following message appears: 603 Port MA0 IP *address*
  - b Record the management port IP address that appears in **Basic MSA1510i information**.
- 3 From the server or a workstation on the MSA1510i LAN segment, open a Web browser and enter the address obtained in [Step 2](#).
- 4 When prompted, enter the default access permissions.
  - User name: root
  - Password: root
- 5 When prompted, set a unique user name and password.
- 6 Using the wizard, complete the following actions.

Option	Description
<b>Storage configuration</b>	<ol style="list-style-type: none"> <li>a Set the Fault Tolerant mode (RAID mode).</li> <li>b Assign a spare disk for appropriate RAID level.</li> </ol>
<b>iSCSI configuration (configure an iSCSI portal)</b>	<ol style="list-style-type: none"> <li>a Select a data port.</li> <li>b Assign an IP address to the data port.</li> <li>c VLANs are set up on the switch and are used as one method of controlling access to the storage. If you are using VLANs, enter the VLAN ID to use (0 = not used).</li> <li>d The wizard suggests a default iSCSI Target Name and iSCSI Target Alias. Accept the default or enter user-defined values.</li> </ol> <p><b>Note</b> To configure the remaining data ports, complete the Initial System Configuration Wizard process, and then use tasks available on the <b>Configure</b> tab.</p>
<b>Login settings</b>	Enter login settings.
<b>Management settings</b>	Enter management settings.

- 7 Click **Finish** to apply the configuration settings.

## Results

---

**Note** Wizards are available for basic configuration tasks only. Use the **Manage** and **Configure** tabs to view and change your configuration.

---

### What to do next

After initial setup, perform the following tasks to complete the configuration:

- Create an array.
- Create a logical drive.
- Create a target.
- Create a portal group.
- Associate or assign the portals created using the wizard with the portal group created.
- Map logical drives to the target.
- Add initiators (initiator IQN name and alias).
- Update the ACLs of the logical drives to provide access to initiators (select the list of initiators to access the logical drive).

## HP StorageWorks EVA Storage Systems

The two types of HP StorageWorks EVA systems are EVA\_GL, an active-passive system, and EVA\_XL, an active-active system. For the systems to work with ESXi, certain specific settings are required.

Set the connection type to **Custom** when you present a LUN to an ESXi host. The value is one of the following:

- For HP EVAgl 3000/5000 (active-passive), use the 000000002200282E host mode type.
- For HP EVAgl firmware 4.001 (active-active firmware for GL series) and above, use the VMware host mode type.
- For EVA4000/6000/8000 active-active arrays with firmware earlier than 5.031, use the 000000202200083E host mode type.
- For EVA4000/6000/8000 active-active arrays with firmware 5.031 and later, use the VMware host mode type.

Otherwise, EVA systems do not require special configuration changes to work with an ESXi system.

## NetApp Storage Systems

For NetApp storage systems to communicate within an ESXi environment, you must perform specific configuration steps.

For additional documentation on NetApp and VMware best practices and SAN solutions, search the NetApp web page.

**Table 12-1. Configuration Steps**

Configuration Step	Description
Disable ALUA for 7-mode.	If any of your iSCSI initiators are a part of an initiator group (igroup), disable ALUA on the 7-mode NetApp controller. Clustered ONTAP is ALUA only.
Set up multipathing.	When you set up multipathing between two iSCSI HBAs and multiple ports on a NetApp storage system, give each HBA a different iSCSI initiator name.  The NetApp storage system only permits one connection for each target and each initiator. Attempts to make additional connections cause the first connection to drop. Therefore, a single HBA should not attempt to connect to multiple IP addresses associated with the same NetApp target.
Set LUN type and initiator group type.	Set the appropriate LUN type and initiator group type for the storage system: <ul style="list-style-type: none"> <li>■ LUN type – VMware (if VMware type is not available, use Linux).</li> <li>■ Initiator group type – VMware (if VMware type is not available, use Linux).</li> </ul>
Provision storage.	Use either FilerView or CLI.

## Dell EqualLogic Storage Systems

When setting up your EqualLogic storage systems to work in an ESXi implementation, you must address certain specific issues.

Follow these requirements:

- You do not need any special multipathing setup. EqualLogic storage systems support storage-processor failover that is transparent to iSCSI. Multiple iSCSI HBAs or NICs can connect to the same target or LUN on the storage side. EqualLogic provides a multipathing extension module that can be installed on ESXi to provide additional capabilities.

For information about the custom multipathing configuration packages, contact Dell EqualLogic.

- Enable ARP redirection for ESXi hardware iSCSI HBAs. See [Configuring Advanced Parameters for iSCSI](#).
- EqualLogic storage systems impose a maximum limit of 1024 iSCSI connections per storage pool and 2048 connections per storage group.
- Extend the iSCSI login timeout on the ESXi host to 15 - 30 seconds. See [Configuring Advanced Parameters for iSCSI](#).

For more information about configuring and using EqualLogic storage systems, see the vendor's documentation and the VMware knowledge base article at <http://kb.vmware.com/kb/2007829>.



## HP StorageWorks SAN/iQ Storage Systems

HP StorageWorks (formerly LeftHand) SAN/iQ systems support ESXi iSCSI connections from a software initiator and hardware initiator.

When configuring SAN/iQ, enable automatic volume resignaturing for SAN/iQ storage devices to allow access to SAN/iQ snapshots and remote copies.

For more information on configuring HP StorageWorks storage for VMware vSphere, see the vendor documentation related to VMware.

Basic configuration steps include several tasks.

- 1 Install SAN/iQ storage nodes.
- 2 Create SAN/iQ management groups and clusters.
- 3 Create volumes.
- 4 Assign volumes to authentication groups and volume lists.
- 5 Enable ARP redirection on hardware iSCSI HBAs.

As a best practice, configure virtual IP load balancing in SAN/iQ for all ESXi authentication groups.

## iSCSI Targets in vApps

If you use an iSCSI target in a virtual appliance, for example HP LeftHand P4000 VSA, the host should connect to the target through the software iSCSI adapter rather than a hardware iSCSI adapter.

# Booting from iSCSI SAN

# 13

When you set up your host to boot from a SAN, your host's boot image is stored on one or more LUNs in the SAN storage system. When the host starts, it boots from the LUN on the SAN rather than from its local disk.

You can use boot from the SAN if you do not want to handle maintenance of local storage or have diskless hardware configurations, such as blade systems.

ESXi supports different methods of booting from the iSCSI SAN.

**Table 13-1. Boot from iSCSI SAN support**

Independent Hardware iSCSI	Software iSCSI and Dependent Hardware iSCSI
Configure the iSCSI HBA to boot from the SAN. For information on configuring the HBA, see <a href="#">Configure Independent Hardware iSCSI Adapter for SAN Boot</a>	Use the network adapter that supports the iBFT. For information, see <a href="#">iBFT iSCSI Boot Overview</a> .

This chapter includes the following topics:

- [General Boot from iSCSI SAN Recommendations](#)
- [Prepare the iSCSI SAN](#)
- [Configure Independent Hardware iSCSI Adapter for SAN Boot](#)
- [iBFT iSCSI Boot Overview](#)

## General Boot from iSCSI SAN Recommendations

If you plan to set up and use an iSCSI LUN as the boot device for your host, you need to follow certain general guidelines.

The following guidelines apply to booting from independent hardware iSCSI and iBFT.

- Review any vendor recommendations for the hardware you use in your boot configuration.
- For installation prerequisites and requirements, review *vSphere Installation and Setup*.
- Use static IP addresses to reduce the chances of DHCP conflicts.
- Use different LUNs for VMFS datastores and boot partitions.

- Configure proper ACLs on your storage system.
  - The boot LUN should be visible only to the host that uses the LUN. No other host on the SAN should be permitted to see that boot LUN.
  - If a LUN is used for a VMFS datastore, it can be shared by multiple hosts. ACLs on the storage systems can allow you to do this.
- Configure a diagnostic partition.
  - With independent hardware iSCSI only, you can place the diagnostic partition on the boot LUN. If you configure the diagnostic partition in the boot LUN, this LUN cannot be shared across multiple hosts. If a separate LUN is used for the diagnostic partition, it can be shared by multiple hosts.
  - If you boot from SAN using iBFT, you cannot set up a diagnostic partition on a SAN LUN. To collect your host's diagnostic information, use the vSphere ESXi Dump Collector on a remote server. For information about the ESXi Dump Collector, see *vSphere Installation and Setup* and *vSphere Networking*.

## Prepare the iSCSI SAN

Before you configure your host to boot from an iSCSI LUN, prepare and configure your storage area network.

---

**Caution** If you use scripted installation to install ESXi when booting from a SAN, you must take special steps to avoid unintended data loss.

---

### Procedure

- 1 Connect network cables, referring to any cabling guide that applies to your setup.
- 2 Ensure IP connectivity between your storage system and server.
 

This includes proper configuration of any routers or switches on your storage network. Storage systems must be able to ping the iSCSI adapters in your hosts.
- 3 Configure the storage system.
  - a Create a volume (or LUN) on the storage system for your host to boot from.
  - b Configure the storage system so that your host has access to the assigned LUN.
 

This could involve updating ACLs with the IP addresses, iSCSI names, and the CHAP authentication parameter you use on your host. On some storage systems, in addition to providing access information for the ESXi host, you must also explicitly associate the assigned LUN with the host.
  - c Ensure that the LUN is presented to the host correctly.

- d Ensure that no other system has access to the configured LUN.
- e Record the iSCSI name and IP addresses of the targets assigned to the host.

You must have this information to configure your iSCSI adapters.

## Configure Independent Hardware iSCSI Adapter for SAN Boot

If your ESXi host uses an independent hardware iSCSI adapter, such as QLogic HBA, you need to configure the adapter to boot from the SAN.

This procedure discusses how to enable the QLogic iSCSI HBA to boot from the SAN. For more information and more up-to-date details about QLogic adapter configuration settings, see the QLogic web site.

### Prerequisites

Because you first need to boot from the VMware installation media, set up your host to boot from CD/DVD-ROM. To achieve this, change the system boot sequence in your system BIOS setup.

### Procedure

- 1 Insert the installation CD/DVD in the CD/DVD-ROM drive and reboot the host.
- 2 Use the BIOS to set the host to boot from the CD/DVD-ROM drive first.
- 3 During server POST, press Ctrl+q to enter the QLogic iSCSI HBA configuration menu.
- 4 Select the I/O port to configure.  
By default, the Adapter Boot mode is set to Disable.
- 5 Configure the HBA.
  - a From the **Fast!UTIL Options** menu, select **Configuration Settings > Host Adapter Settings**.
  - b Configure the following settings for your host adapter: initiator IP address, subnet mask, gateway, initiator iSCSI name, and CHAP (if required).
- 6 Configure iSCSI settings.  
See [Configure iSCSI Boot Settings](#).
- 7 Save your changes and restart the system.

## Configure iSCSI Boot Settings

When setting up your ESXi host to boot from iSCSI, you need to configure iSCSI boot settings.

### Procedure

- 1 From the **Fast!UTIL Options** menu, select **Configuration Settings > iSCSI Boot Settings**.

- 2 Before you can set SendTargets, set Adapter Boot mode to **Manual**.
- 3 Select **Primary Boot Device Settings**.
  - a Enter the discovery **Target IP** and **Target Port**.
  - b You can leave the **Boot LUN** and **iSCSI Name** fields blank if only one iSCSI target and one LUN are at the specified address to boot from. Otherwise, you must specify these fields to ensure that you do not boot from a volume for some other system. After the target storage system is reached, these fields will be populated after a rescan.
  - c Save changes.
- 4 From the **iSCSI Boot Settings** menu, select the primary boot device. An auto rescan of the HBA is made to find new target LUNS.
- 5 Select the iSCSI target.

---

**Note** If more than one LUN exists within the target, you can choose a specific LUN ID by pressing **Enter** after you locate the iSCSI device.

---

- 6 Return to the **Primary Boot Device Setting** menu. After the rescan, the **Boot LUN** and **iSCSI Name** fields are populated. Change the value of **Boot LUN** to the desired LUN ID.

## iBFT iSCSI Boot Overview

ESXi hosts can boot from an iSCSI SAN using the software or dependent hardware iSCSI adapters and network adapters.

To deploy ESXi and boot from the iSCSI SAN, the host must have an iSCSI boot capable network adapter that supports the iSCSI Boot Firmware Table (iBFT) format. The iBFT is a method of communicating parameters about the iSCSI boot device to an operating system.

Before installing ESXi and booting from the iSCSI SAN, configure the networking and iSCSI boot parameters on the network adapter and enable the adapter for the iSCSI boot. Because configuring the network adapter is vendor specific, review your vendor documentation for instructions.

When you first boot from iSCSI, the iSCSI boot firmware on your system connects to an iSCSI target. If login is successful, the firmware saves the networking and iSCSI boot parameters in the iBFT and stores the table in the system's memory. The system uses this table to configure its own iSCSI connection and networking and to start up.

The following list describes the iBFT iSCSI boot sequence.

- 1 When restarted, the system BIOS detects the iSCSI boot firmware on the network adapter.
- 2 The iSCSI boot firmware uses the preconfigured boot parameters to connect with the specified iSCSI target.

- 3 If the connection to the iSCSI target is successful, the iSCSI boot firmware writes the networking and iSCSI boot parameters in to the iBFT and stores the table in the system memory.

---

**Note** The system uses this table to configure its own iSCSI connection and networking and to start up.

---

- 4 The BIOS boots the boot device.
- 5 The VMkernel starts loading and takes over the boot operation.
- 6 Using the boot parameters from the iBFT, the VMkernel connects to the iSCSI target.
- 7 After the iSCSI connection is established, the system boots.

## iBFT iSCSI Boot Considerations

When you boot the ESXi host from iSCSI using iBFT-enabled network adapters, certain considerations apply.

The iBFT iSCSI boot does not support the following items:

- IPv6
- Failover for the iBFT-enabled network adapters

---

**Note** Update your NIC's boot code and iBFT firmware using vendor supplied tools before trying to install and boot VMware ESXi. Consult vendor documentation and VMware HCL for supported boot code and iBFT firmware versions for VMware ESXi iBFT boot. The boot code and iBFT firmware released by vendors prior to the ESXi 4.1 release might not work.

---

After you set up your host to boot from iBFT iSCSI, the following restrictions apply:

- You cannot disable the software iSCSI adapter. If the iBFT configuration is present in the BIOS, the host re-enables the software iSCSI adapter during each reboot.

---

**Note** If you do not use the iBFT-enabled network adapter for the iSCSI boot and do not want the software iSCSI adapter to be always enabled, remove the iBFT configuration from the network adapter.

---

- You cannot remove the iBFT iSCSI boot target using the vSphere Web Client. The target appears on the list of adapter static targets.

## Configuring iBFT Boot from SAN

You can boot from the iSCSI SAN using the software iSCSI adapter or a dependent hardware iSCSI adapter and a network adapter. The network adapter must support iBFT.

When you set up your host to boot with iBFT, you perform a number of tasks.

## Procedure

### 1 Configure iSCSI Boot Parameters

To begin an iSCSI boot process, a network adapter on your host must have a specially configured iSCSI boot firmware. When you configure the firmware, you specify the networking and iSCSI parameters and enable the adapter for the iSCSI boot.

### 2 Change Boot Sequence in BIOS

When setting up your host to boot from iBFT iSCSI, change the boot sequence to force your host to boot in an appropriate order.

### 3 Install ESXi to iSCSI Target

When setting up your host to boot from iBFT iSCSI, install the ESXi image to the target LUN.

### 4 Boot ESXi from iSCSI Target

After preparing the host for an iBFT iSCSI boot and copying the ESXi image to the iSCSI target, perform the actual boot.

## Configure iSCSI Boot Parameters

To begin an iSCSI boot process, a network adapter on your host must have a specially configured iSCSI boot firmware. When you configure the firmware, you specify the networking and iSCSI parameters and enable the adapter for the iSCSI boot.

Configuration on the network adapter can be dynamic or static. If you use the dynamic configuration, you indicate that all target and initiator boot parameters are acquired using DHCP. For the static configuration, you manually enter data that includes your host's IP address and initiator IQN, and the target parameters.

## Procedure

- ◆ On the network adapter that you use for the boot from iSCSI, specify networking and iSCSI parameters.

Because configuring the network adapter is vendor specific, review your vendor documentation for instructions.

## Change Boot Sequence in BIOS

When setting up your host to boot from iBFT iSCSI, change the boot sequence to force your host to boot in an appropriate order.

Change the BIOS boot sequence to the following sequence:

- iSCSI
- DVD-ROM

Because changing the boot sequence in the BIOS is vendor specific, refer to vendor documentation for instructions. The following sample procedure explains how to change the boot sequence on a Dell host with a Broadcom network adapter.

#### Procedure

- 1 Turn on the host.
- 2 During Power-On Self-Test (POST), press F2 to enter the BIOS Setup.
- 3 In the BIOS Setup, select **Boot Sequence** and press Enter.
- 4 In the Boot Sequence menu, arrange the bootable items so that iSCSI precedes the DVD-ROM.
- 5 Press Esc to exit the Boot Sequence menu.
- 6 Press Esc to exit the BIOS Setup.
- 7 Select **Save Changes** and click **Exit** to exit the BIOS Setup menu.

### Install ESXi to iSCSI Target

When setting up your host to boot from iBFT iSCSI, install the ESXi image to the target LUN.

#### Prerequisites

- Configure iSCSI boot firmware on your boot NIC to point to the target LUN that you want to use as the boot LUN.
- Change the boot sequence in the BIOS so that iSCSI precedes the DVD-ROM.
- If you use Broadcom adapters, set **Boot to iSCSI target** to **Disabled**.

#### Procedure

- 1 Insert the installation media in the CD/DVD-ROM drive and restart the host.
- 2 When the installer starts, follow the typical installation procedure.
- 3 When prompted, select the iSCSI LUN as the installation target.  
The installer copies the ESXi boot image to the iSCSI LUN.
- 4 After the system restarts, remove the installation DVD.

### Boot ESXi from iSCSI Target

After preparing the host for an iBFT iSCSI boot and copying the ESXi image to the iSCSI target, perform the actual boot.

#### Prerequisites

- Configure the iSCSI boot firmware on your boot NIC to point to the boot LUN.
- Change the boot sequence in the BIOS so that iSCSI precedes the boot device.
- If you use Broadcom adapters, set **Boot to iSCSI target** to **Enabled**



## Procedure

### 1 Restart the host.

The host boots from the iSCSI LUN using iBFT data. During the first boot, the iSCSI initialization script sets up default networking. The network setup is persistent after subsequent reboots.

### 2 (Optional) Adjust networking configuration using the vSphere Web Client.

## Networking Best Practices

To boot the ESXi host from iSCSI using iBFT, you must properly configure networking.

To achieve greater security and better performance, have redundant network adapters on the host.

How you set up all the network adapters depends on whether your environment uses shared or isolated networks for the iSCSI traffic and host management traffic.

### Shared iSCSI and Management Networks

Configure the networking and iSCSI parameters on the first network adapter on the host. After the host boots, you can add secondary network adapters to the default port group.

### Isolated iSCSI and Management Networks

When you configure isolated iSCSI and management networks, follow these guidelines to avoid bandwidth problems.

- Your isolated networks must be on different subnets.
- If you use VLANs to isolate the networks, they must have different subnets to ensure that routing tables are properly set up.
- VMware recommends that you configure the iSCSI adapter and target to be on the same subnet. If you set up the iSCSI adapter and target on different subnets, the following restrictions apply:
  - The default VMkernel gateway must be able to route both the management and iSCSI traffic.
  - After you boot your host, you can use the iBFT-enabled network adapter only for iBFT. You cannot use the adapter for other iSCSI traffic.
- Use the first physical network adapter for the management network.
- Use the second physical network adapter for the iSCSI network. Make sure to configure the iBFT.
- After the host boots, you can add secondary network adapters to both the management and iSCSI networks.

## Change iBFT iSCSI Boot Settings

If settings, such as the IQN name, IP address, and so on, change on the iSCSI storage or your host, update the iBFT. This task assumes that the boot LUN and the data stored on the LUN remain intact.

### Procedure

- 1 Shut down the ESXi host.
- 2 Change iSCSI storage settings.
- 3 Update the iBFT on the host with the new settings.
- 4 Restart the host.

The host boots using the new information stored in the iBFT.

## Troubleshooting iBFT iSCSI Boot

The topics in this section help you to identify and solve problems you might encounter when using iBFT iSCSI boot.

### Loss of System's Gateway Causes Loss of Network Connectivity

You lose network connectivity when you delete a port group associated with the iBFT network adapter.

#### Problem

A loss of network connectivity occurs after you delete a port group.

#### Cause

When you specify a gateway in the iBFT-enabled network adapter during ESXi installation, this gateway becomes the system's default gateway. If you delete the port group associated with the network adapter, the system's default gateway is lost. This action causes the loss of network connectivity.

#### Solution

Do not set an iBFT gateway unless it is required. If the gateway is required, after installation, manually set the system's default gateway to the one that the management network uses.

### Changing iSCSI Boot Parameters Causes ESXi to Boot in Stateless Mode

Changing iSCSI boot parameters on the network adapter after the first boot does not update the iSCSI and networking configuration on the ESXi host.

#### Problem

If you change the iSCSI boot parameters on the network adapter after the first ESXi boot from iSCSI, the host will boot in a stateless mode.

**Cause**

The firmware uses the updated boot configuration and is able to connect to the iSCSI target and load the ESXi image. However, when loaded, the system does not pick up the new parameters, but continues to use persistent networking and iSCSI parameters from the previous boot. As a result, the host cannot connect to the target and boots in the stateless mode.

**Solution**

- 1** Use the vSphere Web Client to connect to the ESXi host.
- 2** Re-configure the iSCSI and networking on the host to match the iBFT parameters.
- 3** Perform a rescan.

# Best Practices for iSCSI Storage

# 14

When using ESXi with the iSCSI SAN, follow best practices that VMware offers to avoid problems.

Check with your storage representative if your storage system supports Storage API - Array Integration hardware acceleration features. If it does, refer to your vendor documentation for information on how to enable hardware acceleration support on the storage system side. For more information, see [Chapter 24 Storage Hardware Acceleration](#).

This chapter includes the following topics:

- [Preventing iSCSI SAN Problems](#)
- [Optimizing iSCSI SAN Storage Performance](#)
- [Checking Ethernet Switch Statistics](#)
- [iSCSI SAN Configuration Checklist](#)

## Preventing iSCSI SAN Problems

When using ESXi in conjunction with a SAN, you must follow specific guidelines to avoid SAN problems.

You should observe these tips for avoiding problems with your SAN configuration:

- Place only one VMFS datastore on each LUN. Multiple VMFS datastores on one LUN is not recommended.
- Do not change the path policy the system sets for you unless you understand the implications of making such a change.
- Document everything. Include information about configuration, access control, storage, switch, server and iSCSI HBA configuration, software and firmware versions, and storage cable plan.
- Plan for failure:
  - Make several copies of your topology maps. For each element, consider what happens to your SAN if the element fails.

- Cross off different links, switches, HBAs and other elements to ensure you did not miss a critical failure point in your design.
- Ensure that the iSCSI HBAs are installed in the correct slots in the ESXi host, based on slot and bus speed. Balance PCI bus load among the available busses in the server.
- Become familiar with the various monitor points in your storage network, at all visibility points, including ESXi performance charts, Ethernet switch statistics, and storage performance statistics.
- Be cautious when changing IDs of the LUNs that have VMFS datastores being used by your host. If you change the ID, virtual machines running on the VMFS datastore will fail.

If there are no running virtual machines on the VMFS datastore, after you change the ID of the LUN, you must use rescan to reset the ID on your host. For information on using rescan, see [Storage Refresh and Rescan Operations](#).

- If you need to change the default iSCSI name of your iSCSI adapter, make sure the name you enter is worldwide unique and properly formatted. To avoid storage access problems, never assign the same iSCSI name to different adapters, even on different hosts.

## Optimizing iSCSI SAN Storage Performance

Several factors contribute to optimizing a typical SAN environment.

If the network environment is properly configured, the iSCSI components provide adequate throughput and low enough latency for iSCSI initiators and targets. If the network is congested and links, switches or routers are saturated, iSCSI performance suffers and might not be adequate for ESXi environments.

### Storage System Performance

Storage system performance is one of the major factors contributing to the performance of the entire iSCSI environment.

If issues occur with storage system performance, consult your storage system vendor's documentation for any relevant information.

When you assign LUNs, remember that you can access each shared LUN through a number of hosts, and that a number of virtual machines can run on each host. One LUN used by the ESXi host can service I/O from many different applications running on different operating systems. Because of this diverse workload, the RAID group that contains the ESXi LUNs should not include LUNs that other hosts use that are not running ESXi for I/O intensive applications.

Enable read caching and write caching.

Load balancing is the process of spreading server I/O requests across all available SPs and their associated host server paths. The goal is to optimize performance in terms of throughput (I/O per second, megabytes per second, or response times).

SAN storage systems require continual redesign and tuning to ensure that I/O is load balanced across all storage system paths. To meet this requirement, distribute the paths to the LUNs among all the SPs to provide optimal load balancing. Close monitoring indicates when it is necessary to manually rebalance the LUN distribution.

Tuning statically balanced storage systems is a matter of monitoring the specific performance statistics (such as I/O operations per second, blocks per second, and response time) and distributing the LUN workload to spread the workload across all the SPs.

## Server Performance with iSCSI

You must consider several factors to ensure optimal server performance.

Each server application must have access to its designated storage with the following conditions:

- High I/O rate (number of I/O operations per second)
- High throughput (megabytes per second)
- Minimal latency (response times)

Because each application has different requirements, you can meet these goals by choosing an appropriate RAID group on the storage system. To achieve performance goals, perform the following tasks:

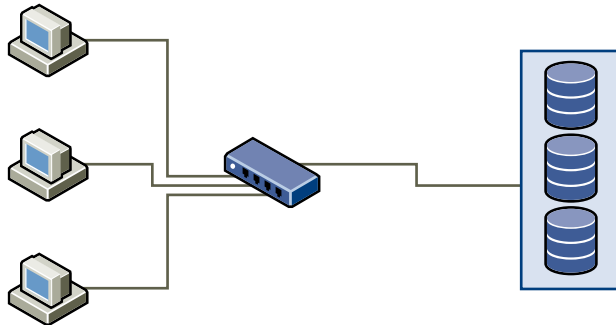
- Place each LUN on a RAID group that provides the necessary performance levels. Pay attention to the activities and resource utilization of other LUNS in the assigned RAID group. A high-performance RAID group that has too many applications doing I/O to it might not meet performance goals required by an application running on the ESXi host.
- Provide each server with a sufficient number of network adapters or iSCSI hardware adapters to allow maximum throughput for all the applications hosted on the server for the peak period. I/O spread across multiple ports provides higher throughput and less latency for each application.
- To provide redundancy for software iSCSI, make sure the initiator is connected to all network adapters used for iSCSI connectivity.
- When allocating LUNs or RAID groups for ESXi systems, multiple operating systems use and share that resource. As a result, the performance required from each LUN in the storage subsystem can be much higher if you are working with ESXi systems than if you are using physical machines. For example, if you expect to run four I/O intensive applications, allocate four times the performance capacity for the ESXi LUNs.
- When using multiple ESXi systems in conjunction with vCenter Server, the performance needed from the storage subsystem increases correspondingly.
- The number of outstanding I/Os needed by applications running on an ESXi system should match the number of I/Os the SAN can handle.

## Network Performance

A typical SAN consists of a collection of computers connected to a collection of storage systems through a network of switches. Several computers often access the same storage.

Single Ethernet Link Connection to Storage shows several computer systems connected to a storage system through an Ethernet switch. In this configuration, each system is connected through a single Ethernet link to the switch, which is also connected to the storage system through a single Ethernet link. In most configurations, with modern switches and typical traffic, this is not a problem.

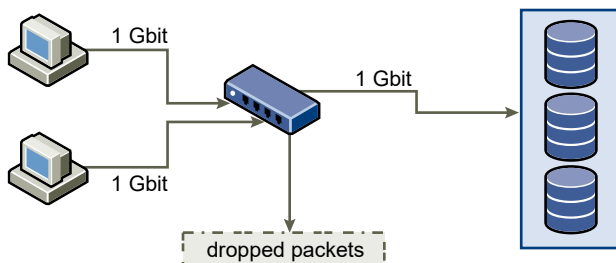
**Figure 14-1. Single Ethernet Link Connection to Storage**



When systems read data from storage, the maximum response from the storage is to send enough data to fill the link between the storage systems and the Ethernet switch. It is unlikely that any single system or virtual machine gets full use of the network speed, but this situation can be expected when many systems share one storage device.

When writing data to storage, multiple systems or virtual machines might attempt to fill their links. As Dropped Packets shows, when this happens, the switch between the systems and the storage system has to drop data. This happens because, while it has a single connection to the storage device, it has more traffic to send to the storage system than a single link can carry. In this case, the switch drops network packets because the amount of data it can transmit is limited by the speed of the link between it and the storage system.

**Figure 14-2. Dropped Packets**



Recovering from dropped network packets results in large performance degradation. In addition to time spent determining that data was dropped, the retransmission uses network bandwidth that could otherwise be used for current transactions.

iSCSI traffic is carried on the network by the Transmission Control Protocol (TCP). TCP is a reliable transmission protocol that ensures that dropped packets are retried and eventually reach their destination. TCP is designed to recover from dropped packets and retransmits them quickly and seamlessly. However, when the switch discards packets with any regularity, network throughput suffers significantly. The network becomes congested with requests to resend data and with the resent packets, and less data is actually transferred than in a network without congestion.

Most Ethernet switches can buffer, or store, data and give every device attempting to send data an equal chance to get to the destination. This ability to buffer some transmissions, combined with many systems limiting the number of outstanding commands, allows small bursts from several systems to be sent to a storage system in turn.

If the transactions are large and multiple servers are trying to send data through a single switch port, a switch's ability to buffer one request while another is transmitted can be exceeded. In this case, the switch drops the data it cannot send, and the storage system must request retransmission of the dropped packet. For example, if an Ethernet switch can buffer 32KB on an input port, but the server connected to it thinks it can send 256KB to the storage device, some of the data is dropped.

Most managed switches provide information on dropped packets, similar to the following:

```

*: interface is up
IHQ: pkts in input hold queue      IQD: pkts dropped from input queue
OHQ: pkts in output hold queue    OQD: pkts dropped from output queue
RXBS: rx rate (bits/sec)          RXPS: rx rate (pkts/sec)
TXBS: tx rate (bits/sec)          TXPS: tx rate (pkts/sec)
TRTL: throttle count

```

**Table 14-1. Sample Switch Information**

Interface	IHQ	IQD	OHQ	OQD	RXBS	RXPS	TXBS	TXPS	TRTL
* GigabitEthernet0/1	3	9922	0	0	476303000	62273	47784000	63677	0

In this example from a Cisco switch, the bandwidth used is 476303000 bits/second, which is less than half of wire speed. In spite of this, the port is buffering incoming packets and has dropped quite a few packets. The final line of this interface summary indicates that this port has already dropped almost 10,000 inbound packets in the IQD column.

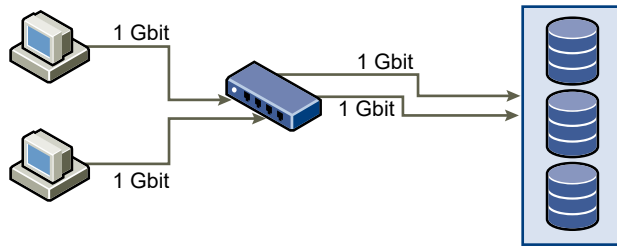
Configuration changes to avoid this problem involve making sure several input Ethernet links are not funneled into one output link, resulting in an oversubscribed link. When a number of links transmitting near capacity are switched to a smaller number of links, oversubscription is a possibility.

Generally, applications or systems that write a lot of data to storage, such as data acquisition or transaction logging systems, should not share Ethernet links to a storage device. These types of applications perform best with multiple connections to storage devices.



Multiple Connections from Switch to Storage shows multiple connections from the switch to the storage.

**Figure 14-3. Multiple Connections from Switch to Storage**



Using VLANs or VPNs does not provide a suitable solution to the problem of link oversubscription in shared configurations. VLANs and other virtual partitioning of a network provide a way of logically designing a network, but do not change the physical capabilities of links and trunks between switches. When storage traffic and other network traffic end up sharing physical connections, as they would with a VPN, the possibility for oversubscription and lost packets exists. The same is true of VLANs that share interswitch trunks. Performance design for a SANs must take into account the physical limitations of the network, not logical allocations.

## Checking Ethernet Switch Statistics

Many Ethernet switches provide different methods for monitoring switch health.

Switches that have ports operating near maximum throughput much of the time do not provide optimum performance. If you have ports in your iSCSI SAN running near the maximum, reduce the load. If the port is connected to an ESXi system or iSCSI storage, you can reduce the load by using manual load balancing.

If the port is connected between multiple switches or routers, consider installing additional links between these components to handle more load. Ethernet switches also commonly provide information about transmission errors, queued packets, and dropped Ethernet packets. If the switch regularly reports any of these conditions on ports being used for iSCSI traffic, performance of the iSCSI SAN will be poor.

## iSCSI SAN Configuration Checklist

This topic provides a checklist of special setup requirements for different storage systems and ESXi hosts.

**Table 14-2. iSCSI SAN Configuration Requirements**

Component	Comments
All storage systems	Write cache must be disabled if not battery backed.
Topology	No single failure should cause HBA and SP failover, especially with active-passive storage arrays.
EMC Symmetrix	Enable the SPC2 and SC3 settings. Contact EMC for the latest settings.

Table 14-2. iSCSI SAN Configuration Requirements (continued)

Component	Comments
EMC Clariion	Set the EMC Clariion failover mode to 1 or 4. Contact EMC for details.
HP MSA	No specific requirements
HP EVA	<p>For EVA3000/5000 firmware 4.001 and later, and EVA4000/6000/8000 firmware 5.031 and later, set the host type to VMware.</p> <p>Otherwise, set the host mode type to Custom. The value is:</p> <ul style="list-style-type: none"> <li>■ EVA3000/5000 firmware 3.x: 000000002200282E</li> <li>■ EVA4000/6000/8000: 000000202200083E</li> </ul>
NetApp	If any of your iSCSI initiators are a part of an initiator group (igroup), disable ALUA on the NetApp array.
Dell EqualLogic	Make sure ARP Redirect is enabled on independent hardware iSCSI adapters.
HP StorageWorks (formerly LeftHand)	Make sure ARP Redirect is enabled on independent hardware iSCSI adapters.
ESXi Configuration	<p>Set the following Advanced Settings for the ESXi host: <b>Disk.UseLunReset</b> to <b>1</b>.</p> <p>A multipathing policy of Most Recently Used must be set for all LUNs hosting clustered disks for active-passive arrays. A multipathing policy of Most Recently Used or Fixed may be set for LUNs on active-active arrays.</p> <p>Allow ARP redirection if the storage system supports transparent failover.</p>

# Managing Storage Devices

# 15

Manage local and networked storage device that your ESXi host has access to.

This chapter includes the following topics:

- [Understanding Storage Device Naming](#)
- [Storage Refresh and Rescan Operations](#)
- [Identifying Device Connectivity Problems](#)

## Understanding Storage Device Naming

Each storage device, or LUN, is identified by several names.

### Device Identifiers

Depending on the type of storage, the ESXi host uses different algorithms and conventions to generate an identifier for each storage device.

#### SCSI INQUIRY identifiers.

The host uses the SCSI INQUIRY command to query a storage device and uses the resulting data, in particular the Page 83 information, to generate a unique identifier. Device identifiers that are based on Page 83 are unique across all hosts, persistent, and have one of the following formats:

- *naa.number*
- *t10.number*
- *eui.number*

These formats follow the T10 committee standards. See the SCSI-3 documentation on the T10 committee Web site.

#### Path-based identifier.

When the device does not provide the Page 83 information, the host generates an *mpx.path* name, where *path* represents the first path to the device, for example, *mpx.vmhba1:CO:T1:L3*. This identifier can be used in the same way as the SCSI INQUIRY identifies.

The mp<sub>x</sub>. identifier is created for local devices on the assumption that their path names are unique. However, this identifier is neither unique nor persistent and could change after every boot.

Typically, the path to the device has the following format:

*vmhbaAdapter:CChannel:TTarget:LLUN*

- *vmhbaAdapter* is the name of the storage adapter. The name refers to the physical adapter on the host, not to the SCSI controller used by the virtual machines.
- *CChannel* is the storage channel number.  
  
Software iSCSI adapters and dependent hardware adapters use the channel number to show multiple paths to the same target.
- *TTarget* is the target number. Target numbering is determined by the host and might change if the mappings of targets visible to the host change. Targets that are shared by different hosts might not have the same target number.
- *LLUN* is the LUN number that shows the position of the LUN within the target. The LUN number is provided by the storage system. If a target has only one LUN, the LUN number is always zero (0).

For example, *vmhba1:C0:T3:L1* represents LUN1 on target 3 accessed through the storage adapter *vmhba1* and channel 0.

## Legacy Identifier

In addition to the SCSI INQUIRY or mp<sub>x</sub>. identifiers, for each device, ESXi generates an alternative legacy name. The identifier has the following format:

*vm<sub>l</sub>.number*

The legacy identifier includes a series of digits that are unique to the device and can be derived in part from the Page 83 information, if it is available. For nonlocal devices that do not support Page 83 information, the *vm<sub>l</sub>*. name is used as the only available unique identifier.

## Example: Displaying Device Names in the vSphere CLI

You can use the **esxcli --server=*server\_name* storage core device list** command to display all device names in the vSphere CLI. The output is similar to the following example:

```
# esxcli --server=server_name storage core device list
naa.number
  Display Name: DGC Fibre Channel Disk(naa.number)
  ...
  Other UUIDs:vml.number
```

## Rename Storage Devices

You can change the display name of a storage device. The display name is assigned by the ESXi host based on the storage type and manufacturer.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Devices**.
- 4 Select the device to rename and click **Rename**.
- 5 Change the device name to a friendly name.

## Storage Refresh and Rescan Operations

The refresh operation for datastores, storage devices, and storage adapters updates the lists and storage information displayed in the vSphere Web Client. For example, it updates such information as the datastore capacity. When you perform storage management tasks or make changes in the SAN configuration, you might need to rescan your storage.

When you perform VMFS datastore management operations, such as creating a VMFS datastore or RDM, adding an extent, and increasing or deleting a VMFS datastore, your host or the vCenter Server automatically rescans and updates your storage. You can disable the automatic rescan feature by turning off the Host Rescan Filter. See [Turn Off Storage Filters](#).

In certain cases, you need to perform a manual rescan. You can rescan all storage available to your host, or, if you are using the vCenter Server, to all hosts in a folder, cluster, and datacenter.

If the changes you make are isolated to storage connected through a specific adapter, perform a rescan for this adapter.

Perform the manual rescan each time you make one of the following changes.

- Zone a new disk array on a SAN.
- Create new LUNs on a SAN.
- Change the path masking on a host.
- Reconnect a cable.
- Change CHAP settings (iSCSI only).
- Add or remove discovery or static addresses (iSCSI only).
- Add a single host to the vCenter Server after you have edited or removed from the vCenter Server a datastore shared by the vCenter Server hosts and the single host.

---

**Important** If you rescan when a path is unavailable, the host removes the path from the list of paths to the device. The path reappears on the list as soon as it becomes available and starts working again.

---

## Perform Storage Rescan in the vSphere Web Client

When you make changes in your SAN configuration, you might need to rescan your storage. You can rescan all storage available to your host, cluster, or datacenter. If the changes you make are isolated to storage accessed through a specific host, perform rescan for only this host.

### Procedure

- 1 In the vSphere Web Client object navigator, select a host, a cluster, a datacenter, or a folder that contains hosts.
- 2 Select **Actions > All vCenter Actions > Rescan Storage**.
- 3 Specify extent of rescan.

Option	Description
<b>Scan for New Storage Devices</b>	Rescan all adapters to discover new storage devices. If new devices are discovered, they appear in the device list.
<b>Scan for New VMFS Volumes</b>	Rescan all storage devices to discover new datastores that have been added since the last scan. Any new datastores appear in the datastore list.

## Perform Adapter Rescan in the vSphere Web Client

When you make changes in your SAN configuration and these changes are isolated to storage accessed through a specific adapter, perform rescan for only this adapter.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Adapters**, and select the adapter to rescan from the list.
- 4 Click **Rescan Adapter**.

## Change the Number of Scanned LUNs in the vSphere Web Client

By default, the VMkernel scans for LUN 0 to LUN 255 for every target (a total of 256 LUNs). You can modify the **Disk.MaxLUN** parameter to improve LUN discovery speed.

---

**Important** You cannot discover LUNs with a LUN ID number that is greater than 255.

---

Reducing the value can shorten rescan time and boot time. However, the time to rescan LUNs might depend on other factors, including the type of storage system and whether sparse LUN support is enabled.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Settings**.

- 3 Under System, click **Advanced System Settings**.
- 4 Under Advanced System Settings, select **Disk.MaxLUN** and click the **Edit** icon.
- 5 Change the existing value to the value of your choice, and click **OK**.

The value you enter specifies the LUN after the last one you want to discover.

For example, to discover LUNs from 0 through 31, set **Disk.MaxLUN** to 32.

## Identifying Device Connectivity Problems

When your ESXi host experiences a problem while connecting to a storage device, the host treats the problem as permanent or temporary depending on certain factors.

Storage connectivity problems are caused by a variety of reasons. Although ESXi cannot always determine the reason for a storage device or its paths being unavailable, the host differentiates between a permanent device loss (PDL) state of the device and a transient all paths down (APD) state of storage.

### Permanent Device Loss (PDL)

A condition that occurs when a storage device permanently fails or is administratively removed or excluded. It is not expected to become available. When the device becomes permanently unavailable, ESXi receives appropriate sense codes or a login rejection from storage arrays, and is able to recognize that the device is permanently lost.

### All Paths Down (APD)

A condition that occurs when a storage device becomes inaccessible to the host and no paths to the device are available. ESXi treats this as a transient condition because typically the problems with the device are temporary and the device is expected to become available again.

## Detecting PDL Conditions

A storage device is considered to be in the permanent device loss (PDL) state when it becomes permanently unavailable to your ESXi host.

Typically, the PDL condition occurs when a device is unintentionally removed, or its unique ID changes, or when the device experiences an unrecoverable hardware error.

When the storage array determines that the device is permanently unavailable, it sends SCSI sense codes to the ESXi host. The sense codes allow your host to recognize that the device has failed and register the state of the device as PDL. The sense codes must be received on all paths to the device for the device to be considered permanently lost.

After registering the PDL state of the device, the host stops attempts to reestablish connectivity or to issue commands to the device to avoid becoming blocked or unresponsive.

The vSphere Web Client displays the following information for the device:

- The operational state of the device changes to **Lost Communication**.

- All paths are shown as Dead.
- Datastores on the device are grayed out.

The host automatically removes the PDL device and all paths to the device if no open connections to the device exist, or after the last connection closes. You can disable the automatic removal of paths by setting the advanced host parameter `Disk.AutoremoveOnPDL` to 0. See [Set Advanced Host Attributes in the vSphere Web Client](#).

If the device returns from the PDL condition, the host can discover it, but treats it as a new device. Data consistency for virtual machines on the recovered device is not guaranteed.

---

**Note** The host cannot detect PDL conditions and continues to treat the device connectivity problems as APD when a storage device permanently fails in a way that does not return appropriate SCSI sense codes or iSCSI login rejection.

---

## Permanent Device Loss and SCSI Sense Codes

The following VMkernel log example of a SCSI sense code indicates that the device is in the PDL state.

```
H:0x0 D:0x2 P:0x0 Valid sense data: 0x5 0x25 0x0 or Logical Unit Not Supported
```

For information about SCSI sense codes, see *Troubleshooting Storage in vSphere Troubleshooting*.

## Permanent Device Loss and iSCSI

In the case of iSCSI arrays with a single LUN per target, PDL is detected through iSCSI login failure. An iSCSI storage array rejects your host's attempts to start an iSCSI session with a reason `Target Unavailable`. As with the sense codes, this response must be received on all paths for the device to be considered permanently lost.

## Permanent Device Loss and Virtual Machines

After registering the PDL state of the device, the host terminates all I/O from virtual machines. vSphere HA can detect PDL and restart failed virtual machines. For more information, see [PDL Conditions and High Availability](#).

## Performing Planned Storage Device Removal

When a storage device is malfunctioning, you can avoid permanent device loss (PDL) conditions or all paths down (APD) conditions and perform a planned removal and reconnection of a storage device.

Planned device removal is an intentional disconnection of a storage device. You might also plan to remove a device for such reasons as upgrading your hardware or reconfiguring your storage devices. When you perform an orderly removal and reconnection of a storage device, you complete a number of tasks.

- 1 Migrate virtual machines from the device you plan to detach.



See the *vCenter Server and Host Management* documentation.

- 2 Unmount the datastore deployed on the device.

See [Unmount VMFS or NFS Datastores](#).

- 3 Detach the storage device.

See [Detach Storage Devices](#).

- 4 For an iSCSI device with a single LUN per target, delete the static target entry from each iSCSI HBA that has a path to the storage device.

See [Remove Static Targets in the vSphere Web Client](#).

- 5 Perform any necessary reconfiguration of the storage device by using the array console.

- 6 Reattach the storage device.

See [Attach Storage Devices](#).

- 7 Mount the datastore and restart the virtual machines. See [Mount VMFS Datastores](#).

## Detach Storage Devices

Safely detach a storage device from your host.

You might need to detach the device to make it inaccessible to your host, when, for example, you perform a hardware upgrade on the storage side.

### Prerequisites

- The device does not contain any datastores.
- No virtual machines use the device as an RDM disk.
- The device does not contain a diagnostic partition or a scratch partition.

### Procedure

- 1 In the vSphere Web Client, display storage devices.
- 2 Select the device to detach and click the **Detach** icon.

### Results

The device becomes inaccessible. The operational state of the device changes to Unmounted.

### What to do next

If multiple hosts share the device, detach the device from each host.

## Attach Storage Devices

Reattach a storage device that you previously detached.

### Procedure

- 1 In the vSphere Web Client, display storage devices.

- 2 Select the detached storage device and click the **Attach** icon.

### Results

The device becomes accessible.

## Recovering From PDL Conditions

An unplanned permanent device loss (PDL) condition occurs when a storage device becomes permanently unavailable without being properly detached from the ESXi host.

The following items in the vSphere Web Client indicate that the device is in the PDL state:

- The datastore deployed on the device is unavailable.
- Operational state of the device changes to Lost Communication.
- All paths are shown as Dead.
- A warning about the device being permanently inaccessible appears in the VMkernel log file.

To recover from the unplanned PDL condition and remove the unavailable device from the host, you need to perform a number of tasks.

- 1 Power off and unregister all virtual machines that are running on the datastore affected by the PDL condition.

- 2 Unmount the datastore.

See [Unmount VMFS or NFS Datastores](#).

- 3 Perform a rescan on all ESXi hosts that had access to the device.

See [Perform Storage Rescan in the vSphere Web Client](#).

---

**Note** If the rescan is not successful and the host continues to list the device, some pending I/O or active references to the device might still exist. Check for virtual machines, templates, ISO images, raw device mappings, and so on that might still have an active reference to the device or datastore.

---

## Handling Transient APD Conditions

A storage device is considered to be in the all paths down (APD) state when it becomes unavailable to your ESXi host for an unspecified period of time.

The reasons for an APD state can be, for example, a failed switch or a disconnected storage cable.

In contrast with the permanent device loss (PDL) state, the host treats the APD state as transient and expects the device to be available again.

The host indefinitely continues to retry issued commands in an attempt to reestablish connectivity with the device. If the host's commands fail the retries for a prolonged period of time, the host and its virtual machines might be at risk of having performance problems and potentially becoming unresponsive.

To avoid these problems, your host uses a default APD handling feature. When a device enters the APD state, the system immediately turns on a timer and allows your host to continue retrying nonvirtual machine commands for a limited time period.

By default, the APD timeout is set to 140 seconds, which is typically longer than most devices need to recover from a connection loss. If the device becomes available within this time, the host and its virtual machine continue to run without experiencing any problems.

If the device does not recover and the timeout ends, the host stops its attempts at retries and terminates any nonvirtual machine I/O. Virtual machine I/O will continue retrying. The vSphere Web Client displays the following information for the device with the expired APD timeout:

- The operational state of the device changes to **Dead** or **Error**.
- All paths are shown as **Dead**.
- Datastores on the device are dimmed.

Even though the device and datastores are unavailable, virtual machines remain responsive. You can power off the virtual machines or migrate them to a different datastore or host.

If later one or more device paths becomes operational, subsequent I/O to the device is issued normally and all special APD treatment ends.

## Disable Storage APD Handling

The storage all paths down (APD) handling on your ESXi host is enabled by default. When it is enabled, the host continues to retry nonvirtual machine I/O commands to a storage device in the APD state for a limited time period. When the time period expires, the host stops its retry attempts and terminates any nonvirtual machine I/O. You can disable the APD handling feature on your host.

If you disable the APD handling, the host will indefinitely continue to retry issued commands in an attempt to reconnect to the APD device. Continuing to retry is the same behavior as in ESXi version 5.0. This behavior might cause virtual machines on the host to exceed their internal I/O timeout and become unresponsive or fail. The host might become disconnected from vCenter Server.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Settings**.
- 3 Under System, click **Advanced System Settings**.
- 4 Under Advanced System Settings, select the **Misc.APDHandlingEnable** parameter and click the Edit icon.
- 5 Change the value to 0.

## Results

If you disabled the APD handling, you can reenable it when a device enters the APD state. The internal APD handling feature turns on immediately and the timer starts with the current timeout value for each device in APD.

## Change Timeout Limits for Storage APD

The timeout parameter controls how many seconds the ESXi host will retry nonvirtual machine I/O commands to a storage device in an all paths down (APD) state. If needed, you can change the default timeout value.

The timer starts immediately after the device enters the APD state. When the timeout expires, the host marks the APD device as unreachable and fails any pending or new nonvirtual machine I/O. Virtual machine I/O will continue to be retried.

The default timeout parameter on your host is 140 seconds. You can increase the value of the timeout if, for example, storage devices connected to your ESXi host take longer than 140 seconds to recover from a connection loss.

---

**Note** If you change the timeout value while an APD is in progress, it will not effect the timeout for that APD.

---

## Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Settings**.
- 3 Under System, click **Advanced System Settings**.
- 4 Under Advanced System Settings, select the **Misc.APDTimeout** parameter and click the Edit icon.
- 5 Change the default value.

You can enter a value between 20 and 99999 seconds.

## Check the Connection Status of a Storage Device

Use the `esxcli` command to verify the connection status of a particular storage device.

In the procedure, `--server=server_name` specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

## Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

## Procedure

- 1 Run the **esxcli --server=*server\_name* storage core device list -d=*device\_ID*** command.
- 2 Check the connection status in the Status: field.
  - on - Device is connected.
  - dead - Device has entered the APD state. The APD timer starts.
  - dead timeout - The APD timeout has expired.
  - not connected - Device is in the PDL state.

## PDL Conditions and High Availability

When a datastore enters a Permanent Device Loss (PDL) state, High Availability (HA) can power off virtual machines and restart them later. VMware offers advanced options to regulate the power off and restart operations for virtual machines. In the majority of environments, default settings do not need to be changed. However, depending on your specific environment you might need to modify some of these options.

**Table 15-1. Advanced Parameters to Regulate PDL**

Parameter	Type of Parameter	Description
scsi#:#.terminateVMonPDL ide#:#.terminateVMonPDL sata#:#.terminateVMonPDL	Advanced Virtual Machine Option. To configure, see <a href="#">Edit Configuration File Parameters</a> .	Default value is FALSE. When TRUE, this parameter powers off the virtual machine if the specified device that backs up the virtual machine's datastore enters the PDL state. HA will not restart this virtual machine. This setting overrides <code>disk.terminateVMonPDLDefault</code> . When set to DEFAULT, <code>disk.terminateVMonPDLDefault</code> is used.
<code>disk.terminateVMonPDLDefault</code>	Advanced Virtual Machine Option. To configure, see <a href="#">Edit Configuration File Parameters</a> .	Default value is FALSE. When TRUE, this parameter powers off the virtual machine if any device that backs up the virtual machine's datastore enters the PDL state. HA will not restart this virtual machine. This parameter can be overridden on individual device basis by the appropriate <code>{scsi, ide, sata}#:#.terminateVMonPDL</code> parameter. When set to DEFAULT, <code>VMkernel.Boot.terminateVMonPDL</code> is used.

Table 15-1. Advanced Parameters to Regulate PDL (continued)

Parameter	Type of Parameter	Description
VMkernel.Boot.terminateVMOnPDL	Advanced VMkernel Option. To configure, see <a href="#">Set Advanced Host Attributes in the vSphere Web Client</a>	Default value is FALSE. When set to TRUE, this parameter powers off all virtual machines on the system when storage that they are using enters the PDL state. Setting can be overridden for each virtual machine by the appropriate {scsi, ide, sata}#:#.terminateVMOnPDL or disk.terminateVMOnPDLDefault parameter. Can be set only to TRUE or FALSE.  <b>Note</b> After you make changes to VMkernel.Boot.terminateVMOnPDL, reboot the host.
das.maskCleanShutdownEnabled	HA Advanced Option. To configure, see <a href="http://kb.vmware.com/kb/2033250">http://kb.vmware.com/kb/2033250</a> .	This option is set to TRUE by default. It allows HA to restart virtual machines that were powered off while the PDL condition was in progress. When this option is set to true, HA restarts all virtual machines, including those that were intentionally powered off by a user.

## Edit Configuration File Parameters

You can change or add virtual machine configuration parameters when instructed by a VMware technical support representative, or if you see VMware documentation that instructs you to add or change a parameter to fix a problem with your system.

**Important** Changing or adding parameters when a system does not have problems might lead to decreased system performance and instability.

The following conditions apply:

- To change a parameter, you change the existing value for the keyword/value pair. For example, if you start with the keyword/value pair, keyword/value, and change it to keyword/value2, the result is keyword=value2.
- You cannot delete a configuration parameter entry.

**Caution** You must assign a value to configuration parameter keywords. If you do not assign a value, the keyword can return a value of 0, false, or disable, which can result in a virtual machine that cannot power on.

### Procedure

- 1 Right-click a virtual machine in the inventory and select **Edit Settings**.
- 2 Click the **VM Options** tab and expand **Advanced**.

- 3 Click **Edit Configuration**.
- 4 (Optional) To add a parameter, click **Add Row** and type a name and value for the parameter.
- 5 (Optional) To change a parameter, type a new value in the **Value** text box for that parameter.
- 6 Click **OK**.

## Set Advanced Host Attributes in the vSphere Web Client

You can set advanced attributes for a host.

---

**Caution** Changing advanced options is considered unsupported unless VMware technical support or a KB article instruct you to do so. In all other cases, changing these options is considered unsupported. In most cases, the default settings produce the optimum result.

---

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab and click **Settings**.
- 3 Under **System**, click **Advanced System Settings**.
- 4 In Advanced Settings, select the appropriate item.
- 5 Click the **Edit** button to edit the value.
- 6 Click **OK**.

# Working with Datastores

# 16

Datastores are logical containers, analogous to file systems, that hide specifics of physical storage and provide a uniform model for storing virtual machine files. Datastores can also be used for storing ISO images, virtual machine templates, and floppy images.

Depending on the type of storage you use, datastores can be of the following types:

- VMFS datastores that are backed by the Virtual Machine File System format. See [Understanding VMFS Datastores](#).
- NFS datastores that are backed by the Network File System NFS format. See [Understanding NFS Datastores](#).
- Virtual SAN datastores. See [Chapter 19 Working with Virtual SAN](#).

After creating datastores, you can organize them in different ways. For example, you can group them into folders according to business practices. This allows you to assign the same permissions and alarms on the datastores in the group at one time.

You can also add datastores to datastore clusters. A datastore cluster is a collection of datastores with shared resources and a shared management interface. When you create a datastore cluster, you can use Storage DRS to manage storage resources. For information about datastore clusters, see the *vSphere Resource Management* documentation.

This chapter includes the following topics:

- [Understanding VMFS Datastores](#)
- [Understanding NFS Datastores](#)
- [Creating Datastores](#)
- [Change Datastore Name in the vSphere Web Client](#)
- [Managing Duplicate VMFS Datastores](#)
- [Upgrading VMFS Datastores](#)
- [Increasing VMFS Datastore Capacity](#)
- [Unmount VMFS or NFS Datastores](#)
- [Mount VMFS Datastores](#)



- [Remove VMFS Datastores in the vSphere Web Client](#)
- [Storage Filtering](#)
- [Set Up Dynamic Disk Mirroring](#)
- [Collecting Diagnostic Information for ESXi Hosts on a Storage Device](#)
- [Checking Metadata Consistency with VOMA](#)
- [Configuring VMFS Pointer Block Cache](#)

## Understanding VMFS Datastores

To store virtual disks, ESXi uses datastores, which are logical containers that hide specifics of physical storage from virtual machines and provide a uniform model for storing virtual machine files. Datastores that you deploy on block storage devices use the vSphere VMFS format, a special high-performance file system format that is optimized for storing virtual machines.

Several versions of the VMFS file system have been released since its introduction. The following table shows host-to-VMFS version relationships.

**Table 16-1. Host access to VMFS version**

VMFS	ESX/ESXi 3.x host	ESX/ESXi 4.x host	ESXi 5.x host
VMFS2	RO	RO	N
VMFS3	RW	RW	RW
VMFS5	N	N	RW

- RW: Complete read and write support. You can create and power on virtual machines.
- RO: Read only support. You cannot create or power on virtual machines.
- N: No access. ESXi 5.x does not support VMFS2. If your datastore was formatted with VMFS2, first upgrade the datastore to VMFS3 using legacy hosts and the vSphere Client.

Use the vSphere Web Client to set up a VMFS datastore in advance on a block-based storage device that your ESXi host discovers. A VMFS datastore can be extended to span several physical storage extents, including SAN LUNs and local storage. This feature allows you to pool storage and gives you flexibility in creating the datastore necessary for your virtual machines.

You can increase the capacity of a datastore while virtual machines are running on the datastore. This ability lets you add new space to your VMFS datastores as your virtual machine requires it. VMFS is designed for concurrent access from multiple physical machines and enforces the appropriate access controls on virtual machine files.

## How VMFS5 Differs from VMFS3

VMFS5 provides many improvements in scalability and performance over the previous version.

VMFS5 has the following improvements:

- Greater than 2TB storage devices for each VMFS extent.

- Support of virtual machines with large capacity virtual disks, or disks greater than 2TB. Because VMFS3 datastores do not support large capacity virtual disks, you cannot move virtual disks greater than 2TB from a VMFS5 datastore to a VMFS3 datastore.
- Increased resource limits such as file descriptors.
- Standard 1MB file system block size with support of 2TB virtual disks.
- Greater than 2TB disk size for RDMs.
- Support of small files of 1KB.
- Any file located on a VMFS5 datastore, new or upgraded from VMFS3, can be opened in a shared mode by a maximum of 32 hosts. VMFS3 continues to support 8 hosts or fewer for file sharing. This affects VMware products that use linked clones, such as View Manager.
- Scalability improvements on storage devices that support hardware acceleration. For information, see [Chapter 24 Storage Hardware Acceleration](#).
- Default use of hardware assisted locking, also called atomic test and set (ATS) locking, on storage devices that support hardware acceleration. For information about how to turn off ATS locking, see [Turn off ATS Locking](#).
- Ability to reclaim physical storage space on thin provisioned storage devices. For information, see [Array Thin Provisioning and VMFS Datastores](#).
- Online upgrade process that upgrades existing datastores without disrupting hosts or virtual machines that are currently running. For information, see [Upgrading VMFS Datastores](#).

For information about block size limitations of a VMFS datastore, see the VMware knowledge base article at <http://kb.vmware.com/kb/1003565>.

## VMFS Datastores and Storage Disk Formats

Storage devices that your host supports can use either the master boot record (MBR) format or the GUID partition table (GPT) format.

With ESXi 5.0 and later, if you create a new VMFS5 datastore, the device is formatted with GPT. The GPT format enables you to create datastores larger than 2TB and up to 64TB for a single extent.

VMFS3 datastores continue to use the MBR format for their storage devices. Consider the following items when you work with VMFS3 datastores:

- For VMFS3 datastores, the 2TB limit still applies, even when the storage device has a capacity of more than 2TB. To be able to use the entire storage space, upgrade a VMFS3 datastore to VMFS5. Conversion of the MBR format to GPT happens only after you expand the datastore to a size larger than 2TB.
- When you upgrade a VMFS3 datastore to VMFS5, the datastore uses the MBR format. Conversion to GPT happens only after you expand the datastore to a size larger than 2TB.

- When you upgrade a VMFS3 datastore, remove from the storage device any partitions that ESXi does not recognize, for example, partitions that use the EXT2 or EXT3 formats. Otherwise, the host cannot format the device with GPT and the upgrade fails.
- You cannot expand a VMFS3 datastore on devices that have the GPT partition format.

## VMFS Datastores as Repositories

ESXi can format SCSI-based storage devices as VMFS datastores. VMFS datastores primarily serve as repositories for virtual machines.

With VMFS5, you can have up to 256 VMFS datastores per host, with the maximum size of 64TB. The required minimum size for a VMFS datastore is 1.3GB, however, the recommended minimum size is 2GB.

---

**Note** Always have only one VMFS datastore for each LUN.

---

You can store multiple virtual machines on the same VMFS datastore. Each virtual machine, encapsulated in a set of files, occupies a separate single directory. For the operating system inside the virtual machine, VMFS preserves the internal file system semantics, which ensures correct application behavior and data integrity for applications running in virtual machines.

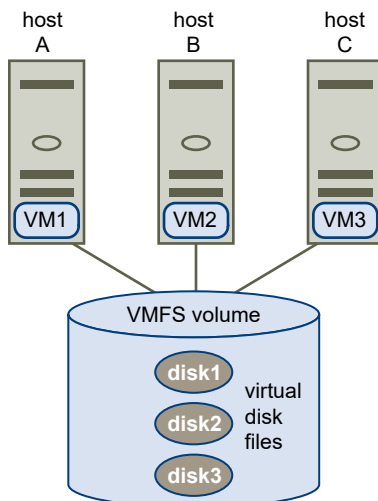
When you run multiple virtual machines, VMFS provides specific locking mechanisms for virtual machine files, so that virtual machines can operate safely in a SAN environment where multiple ESXi hosts share the same VMFS datastore.

In addition to virtual machines, the VMFS datastores can store other files, such as virtual machine templates and ISO images.

## Sharing a VMFS Datastore Across Hosts

As a cluster file system, VMFS lets multiple ESXi hosts access the same VMFS datastore concurrently.

**Figure 16-1. Sharing a VMFS Datastore Across Hosts**



For information on the maximum number of hosts that can connect to a single VMFS datastore, see the *Configuration Maximums* document.

To ensure that multiple hosts do not access the same virtual machine at the same time, VMFS provides on-disk locking.

Sharing the VMFS volume across multiple hosts offers several advantages, for example, the following:

- You can use VMware Distributed Resource Scheduling (DRS) and VMware High Availability (HA).

You can distribute virtual machines across different physical servers. That means you run a mix of virtual machines on each server, so that not all experience high demand in the same area at the same time. If a server fails, you can restart virtual machines on another physical server. If the failure occurs, the on-disk lock for each virtual machine is released. For more information about VMware DRS, see the *vSphere Resource Management* documentation. For information about VMware HA, see the *vSphere Availability* documentation.

- You can use vMotion to migrate running virtual machines from one physical server to another. For information about migrating virtual machines, see the *vCenter Server and Host Management* documentation.

To create a shared datastore, mount the datastore on those ESXi hosts that require the datastore access. See [Mount VMFS Datastores](#).

## VMFS Metadata Updates

A VMFS datastore holds virtual machine files, directories, symbolic links, RDM descriptor files, and so on. The datastore also maintains a consistent view of all the mapping information for these objects. This mapping information is called metadata.

Metadata is updated each time you perform datastore or virtual machine management operations. Examples of operations requiring metadata updates include the following:

- Creating, growing, or locking a virtual machine file
- Changing a file's attributes
- Powering a virtual machine on or off
- Creating or deleting a VMFS datastore
- Expanding a VMFS datastore
- Creating a template
- Deploying a virtual machine from a template
- Migrating a virtual machine with vMotion

When metadata changes are made in a shared storage environment, VMFS uses special locking mechanisms to protect its data and prevent multiple hosts from concurrently writing to the metadata.

## VMFS Locking Mechanisms

In a shared storage environment, when multiple hosts access the same VMFS datastore, specific locking mechanisms are used. These locking mechanisms prevent multiple hosts from concurrently writing to the metadata and ensure that no data corruption occurs.

VMFS supports SCSI reservations and atomic test and set (ATS) locking.

### SCSI Reservations

VMFS uses SCSI reservations on storage devices that do not support hardware acceleration. SCSI reservations lock an entire storage device while an operation that requires metadata protection is performed. After the operation completes, VMFS releases the reservation and other operations can continue. Because this lock is exclusive, excessive SCSI reservations by a host can cause performance degradation on other hosts that are accessing the same VMFS. For information about how to reduce SCSI reservations, see the *vSphere Troubleshooting* documentation.

### Atomic Test and Set (ATS)

For storage devices that support hardware acceleration, VMFS uses the ATS algorithm, also called hardware assisted locking. In contrast with SCSI reservations, ATS supports discrete locking per disk sector. For information about hardware acceleration, see [Chapter 24 Storage Hardware Acceleration](#).

Mechanisms that VMFS uses to apply different types of locking depend on the VMFS version.

**Table 16-2. Use of ATS Locking on Devices with Hardware Acceleration Support**

Storage Devices	New VMFS5	Upgraded VMFS5	VMFS3
Single extent	ATS only	ATS, but can revert to SCSI reservations	ATS, but can revert to SCSI reservations
Multiple extents	Spans only over ATS-capable devices	ATS except when locks on non-head	ATS except when locks on non-head

In certain cases, you might need to turn off the ATS-only setting for a new VMFS5 datastore. For information, see [Turn off ATS Locking](#).

### Turn off ATS Locking

When you create a VMFS5 datastore on a device that supports atomic test and set (ATS) locking, the datastore is set to the ATS-only mode. In certain circumstances, you might need to turn off the ATS mode setting.

Turn off the ATS setting when, for example, your storage device is downgraded or firmware updates fail and the device no longer supports hardware acceleration. The option that you use to turn off the ATS setting is available only through the ESXi Shell. For more information, see the *Getting Started with vSphere Command-Line Interfaces*.

## Procedure

- ◆ To turn off the ATS setting, run the following command:

```
vmkfstools --configATSOnly 0 device
```

The *device* parameter is the path to the head extent device on which VMFS5 was deployed. Use the following format:

```
/vmfs/devices/disks/disk_ID:P
```

---

**Note** It is sufficient to run this command on one of the ESXi hosts sharing the VMFS5 datastore. Other hosts automatically recognize the change.

---

## Understanding NFS Datastores

ESXi can access a designated NFS volume located on a NAS server, mount the volume, and use it for its storage needs. You can use NFS volumes to store and boot virtual machines in the same way that you use VMFS datastores.

ESXi supports NFS version 3. With NFS, the following storage capabilities are supported:

- vMotion
- VMware DRS and VMware HA
- ISO images, which are presented as CD-ROMs to virtual machines
- Virtual machine snapshots
- Virtual machines with large capacity virtual disks, or disks greater than 2TB.

When you work with NFS storage, the following considerations apply:

- The maximum size of NFS datastores depends on the support that an NFS server provides. ESXi does not impose any limits on the NFS datastore size.
- NFS does not support IPv6.
- If you use non-ASCII characters to name datastores and virtual machines, make sure that the underlying NFS server offers internationalization support. If the server does not support international characters, use only ASCII characters, otherwise unpredictable failures might occur.

## NFS Datastores as Repositories for Commonly Used Files

In addition to storing virtual disks on NFS datastores, you can also use NFS as a central repository for ISO images, virtual machine templates, and so on.

To use NFS as a shared repository, you create a directory on the NFS server and then mount it as a datastore on all hosts. If you use the datastore for ISO images, you can connect the virtual machine's CD-ROM device to an ISO file on the datastore and install a guest operating system from the ISO file.

---

**Note** If the underlying NFS volume, on which the files are stored, is read-only, make sure that the volume is exported as a read-only share by the NFS server, or configure it as a read-only datastore on the ESXi host. Otherwise, the host considers the datastore to be read-write and might not be able to open the files.

---

## Using Layer 3 Routed Connections to Access NFS Storage

When you use Layer 3 (L3) routed connections to access NFS storage, consider certain requirements and restrictions.

Ensure that your environment meets the following requirements:

- Use Cisco's Hot Standby Router Protocol (HSRP) in IP Router. If you are using non-Cisco router, be sure to use Virtual Router Redundancy Protocol (VRRP) instead.
- Use Quality of Service (QoS) to prioritize NFS L3 traffic on networks with limited bandwidths, or on networks that experience congestion. See your router documentation for details.
- Follow Routed NFS L3 best practices recommended by storage vendor. Contact your storage vendor for details.
- Disable Network I/O Resource Management (NetIORM).
- If you are planning to use systems with top-of-rack switches or switch-dependent I/O device partitioning, contact your system vendor for compatibility and support.

In an L3 environment the following restrictions apply:

- The environment does not support VMware Site Recovery Manager.
- The environment supports only NFS protocol. Do not use other storage protocols such as FCoE over the same physical network.
- The NFS traffic in this environment does not support IPv6.
- The NFS traffic in this environment can be routed only over a LAN. Other environments such as WAN are not supported.

## Creating Datastores

Depending on the type of storage you use and your storage needs, you can create a VMFS or NFS datastore.

You use the New Datastore wizard to create your datastores. You also use this wizard to manage VMFS datastore copies. For information, see [Managing Duplicate VMFS Datastores](#).

A Virtual SAN datastore is automatically created when you enable Virtual SAN. For information, see [Chapter 19 Working with Virtual SAN](#).

## Create a VMFS Datastore in the vSphere Web Client

VMFS datastores serve as repositories for virtual machines. You can set up VMFS datastores on any SCSI-based storage devices that the host discovers, including Fibre Channel, iSCSI, and local storage devices.

### Prerequisites

- Install and configure any adapters that your storage requires. Rescan the adapters to discover newly added storage devices.
- Browse to Datastores in the vSphere Web Client navigator. See [Display Datastore Information in the vSphere Web Client](#).

### Procedure

- 1 Click the Create a New Datastore icon.
- 2 Type the datastore name and if required, select the placement location for the datastore.  
The vSphere Web Client enforces a 42 character limit for the datastore name.
- 3 Select VMFS as the datastore type.
- 4 Select the device to use for your datastore.

**Important** The device you select must not have any values displayed in the Snapshot Volume column. If a value is present, the device contains a copy of an existing VMFS datastore. For information on managing datastore copies, see [Managing Duplicate VMFS Datastores](#).

- 5 Select the VMFS version.

Option	Description
<b>VMFS 5</b>	Enables more than 2TB LUN support.
<b>VMFS 3</b>	Allows ESX/ESXi hosts version lower than 5.0 to access the datastore.

- 6 Specify partition configuration.

Option	Description
<b>Use all available partitions</b>	Dedicates the entire disk to a single VMFS datastore. If you select this option, all file systems and data currently stored on this device are destroyed.
<b>Use free space</b>	Deploys a VMFS datastore in the remaining free space of the disk.



- 7 (Optional) If the space allocated for the datastore is excessive for your purposes, adjust the capacity values in the Datastore Size field.

By default, the entire free space on the storage device is allocated.

- 8 For the VMFS3 version, select the maximum file size.
- 9 In the Ready to Complete page, review the datastore configuration information and click **Finish**.

## Results

The datastore on the SCSI-based storage device is created. It is available to all hosts that have access to the device.

## Create NFS Datastore in the vSphere Web Client

You can use the **New Datastore** wizard to mount an NFS volume.

### Prerequisites

- Configure VMkernel networking. NFS requires network connectivity to access data stored on remote servers. See the *vSphere Networking* documentation.
- Browse to Datastores in the vSphere Web Client navigator. See [Display Datastore Information in the vSphere Web Client](#).

### Procedure

- 1 Click the Create a New Datastore icon.
- 2 Type the datastore name and if required, select the placement location for the datastore.  
The vSphere Web Client enforces a 42 character limit for the datastore name.
- 3 Select NFS as the datastore type.
- 4 Type the server name and the mount point folder name.

For the server name, you can enter an IP address, a DNS name, or an NFS UUID.

---

**Note** When you mount the same NFS volume on different hosts, make sure that the server and folder names are identical across the hosts. If the names do not match, the hosts see the same NFS volume as two different datastores. This error might result in a failure of such features as vMotion. An example of such discrepancy is if you enter **filer** as the server name on one host and **filer.domain.com** on the other.

---

- 5 (Optional) Select **Mount NFS read only** if the volume is exported as read-only by the NFS server.
- 6 If you are creating a datastore at the datacenter or cluster level, select hosts that require access to the datastore.
- 7 Review the configuration options and click **Finish**.

## Change Datastore Name in the vSphere Web Client

You can change the name of an existing VMFS, NFS, or Virtual SAN datastore. You can rename the datastore that has virtual machines running on it without any negative impact.

### Prerequisites

Browse to Datastores in the vSphere Web Client navigator. See [Display Datastore Information in the vSphere Web Client](#).

### Procedure

- 1 Right-click the datastore to rename.
- 2 Select **Rename** from the right-click menu.
- 3 Type a new datastore name.

The vSphere Web Client enforces a 42 character limit for the datastore name.

### Results

The new name appears on all hosts that have access to the datastore.

## Managing Duplicate VMFS Datastores

When a storage device contains a VMFS datastore copy, you can mount the datastore with the existing signature or assign a new signature.

Each VMFS datastore created in a storage disk has a unique signature, also called UUID, that is stored in the file system superblock. When the storage disk is replicated or its snapshot is taken on the storage side, the resulting disk copy is identical, byte-for-byte, with the original disk. As a result, if the original storage disk contains a VMFS datastore with UUID X, the disk copy appears to contain an identical VMFS datastore, or a VMFS datastore copy, with exactly the same UUID X.

In addition to LUN snapshotting and replication, the following storage device operations might cause ESXi to mark the existing datastore on the device as a copy of the original datastore:

- LUN ID changes
- SCSI device type changes, for example, from SCSI-2 to SCSI-3
- SPC-2 compliancy enablement

ESXi can detect the VMFS datastore copy and display it in the vSphere Web Client. You have an option of mounting the datastore copy with its original UUID or changing the UUID to resignature the datastore.

Whether you chose resignaturing or mounting without resignaturing depends on how the LUNs are masked in the storage environment. If your hosts are able to see both copies of the LUN, then resignaturing is the recommended method. Otherwise, mounting is an option.

## Keep Existing Datastore Signature in the vSphere Web Client

If you do not need to resignature a VMFS datastore copy, you can mount it without changing its signature.

You can keep the signature if, for example, you maintain synchronized copies of virtual machines at a secondary site as part of a disaster recovery plan. In the event of a disaster at the primary site, you mount the datastore copy and power on the virtual machines at the secondary site.

### Prerequisites

- Perform a storage rescan on your host to update the view of storage devices presented to the host.
- Unmount the original VMFS datastore that has the same UUID as the copy you plan to mount. You can mount the VMFS datastore copy only if it does not collide with the original VMFS datastore.
- Browse to Datastores in the vSphere Web Client navigator. See [Display Datastore Information in the vSphere Web Client](#).

### Procedure

- 1 Click the Create a New Datastore icon.
- 2 Type the datastore name and if required, select the placement location for the datastore.
- 3 Select VMFS as the datastore type.
- 4 From the list of storage devices, select the device that has a specific value displayed in the Snapshot Volume column.

The value present in the Snapshot Volume column indicates that the device is a copy that contains a copy of an existing VMFS datastore.

- 5 Under Mount Options, select **Keep Existing Signature**.
- 6 Review the datastore configuration information and click **Finish**.

### What to do next

If you later want to resignature the mounted datastore, you must unmount it first.

## Resignature a VMFS Datastore Copy in the vSphere Web Client

Use datastore resignaturing if you want to retain the data stored on the VMFS datastore copy.

When resignaturing a VMFS copy, ESXi assigns a new signature (UUID) to the copy, and mounts the copy as a datastore distinct from the original. All references to the original signature from virtual machine configuration files are updated.

When you perform datastore resignaturing, consider the following points:

- Datastore resignaturing is irreversible.

- After resignaturing, the storage device replica that contained the VMFS copy is no longer treated as a replica.
- A spanned datastore can be resignatured only if all its extents are online.
- The resignaturing process is crash and fault tolerant. If the process is interrupted, you can resume it later.
- You can mount the new VMFS datastore without a risk of its UUID conflicting with UUIDs of any other datastore, such as an ancestor or child in a hierarchy of storage device snapshots.

#### Prerequisites

- Unmount the datastore copy.
- Perform a storage rescan on your host to update the view of storage devices presented to the host.
- Browse to Datastores in the vSphere Web Client navigator. See [Display Datastore Information in the vSphere Web Client](#).

#### Procedure

- 1 Click the Create a New Datastore icon.
- 2 Type the datastore name and if required, select the placement location for the datastore.
- 3 Select VMFS as the datastore type.
- 4 From the list of storage devices, select the device that has a specific value displayed in the Snapshot Volume column.  
  
The value present in the Snapshot Volume column indicates that the device is a copy that contains a copy of an existing VMFS datastore.
- 5 Under Mount Options, select **Assign a New Signature** and click **Next**.
- 6 Review the datastore configuration information and click **Finish**.

## Upgrading VMFS Datastores

If your datastores were formatted with VMFS2 or VMFS3, you can upgrade the datastores to VMFS5.

When you perform datastore upgrades, consider the following items:

- To upgrade a VMFS2 datastore, you use a two-step process that involves upgrading VMFS2 to VMFS3 first. Because ESXi 5.0 and later hosts cannot access VMFS2 datastores, use a legacy host, ESX/ESXi 4.x or earlier, to access the VMFS2 datastore and perform the VMFS2 to VMFS3 upgrade.

After you upgrade your VMFS2 datastore to VMFS3, the datastore becomes available on the ESXi 5.x host, where you complete the process of upgrading to VMFS5.

- You can perform a VMFS3 to VMFS5 upgrade while the datastore is in use with virtual machines powered on.
- While performing an upgrade, your host preserves all files on the datastore.
- The datastore upgrade is a one-way process. After upgrading your datastore, you cannot revert it back to its previous VMFS format.

An upgraded VMFS5 datastore differs from a newly formatted VMFS5.

**Table 16-3. Comparing Upgraded and Newly Formatted VMFS5 Datastores**

Characteristics	Upgraded VMFS5	Formatted VMFS5
File block size	1, 2, 4, and 8MB	1MB
Subblock size	64KB	8KB
Partition format	MBR. Conversion to GPT happens only after you expand the datastore to a size larger than 2TB.	GPT
Datastore limits	Retains limits of VMFS3 datastore.	

## Upgrade a Datastore to VMFS5 in the vSphere Web Client

You can upgrade your VMFS3 datastore to VMFS5. VMFS5 is a new version of the VMware cluster file system that provides performance and scalability improvements.

You can perform an upgrade while the datastore is in use with virtual machines powered on.

### Prerequisites

- If you have a VMFS2 datastore, you must first upgrade it to VMFS3 using an ESX/ESXi 3.x or ESX/ESXi 4.x host. Use the vSphere Host Client to access the host.
- All hosts accessing the datastore must support VMFS5.
- Verify that the volume to be upgraded has at least 2MB of free blocks available and 1 free file descriptor.
- Browse to Datastores in the vSphere Web Client navigator. See [Display Datastore Information in the vSphere Web Client](#).

### Procedure

- 1 Click the datastore to upgrade.
- 2 Click the **Manage** tab, and click **Settings**.
- 3 Click **Upgrade to VMFS5**.
- 4 Verify that the hosts accessing the datastore support VMFS5.
- 5 Click **OK** to start the upgrade.
- 6 Perform a rescan on all hosts that are associated with the datastore.

## Results

The datastore is upgraded to VMFS5 and is available to all hosts that are associated with the datastore.

## Increasing VMFS Datastore Capacity

If your VMFS datastore requires more space, increase the datastore capacity. You can dynamically increase the capacity by growing a datastore extent or by adding a new extent.

Use one of the following methods to increase a datastore capacity:

- Dynamically grow any expandable datastore extent, so that it fills the available adjacent capacity. The extent is considered expandable when the underlying storage device has free space immediately after the extent.
- Dynamically add a new extent. The datastore can span over up to 32 extents with the size of each extent of more than 2TB, yet appear as a single volume. The spanned VMFS datastore can use any or all of its extents at any time. It does not need to fill up a particular extent before using the next one.

---

**Note** Datastores that only support hardware assisted locking, also called atomic test and set (ATS) mechanism cannot span over non-ATS devices. For more information, see [VMFS Locking Mechanisms](#).

---

## Increase VMFS Datastore Capacity in the vSphere Web Client

When you need to add virtual machines to a datastore, or when the virtual machines running on a datastore require more space, you can dynamically increase the capacity of a VMFS datastore.

If a shared datastore has powered on virtual machines and becomes 100% full, you can increase the datastore's capacity only from the host with which the powered on virtual machines are registered.

### Prerequisites

Browse to Datastores in the vSphere Web Client navigator. See [Display Datastore Information in the vSphere Web Client](#).

### Procedure

- 1 Select the datastore to grow and click the Increase Datastore Capacity icon.

## 2 Select a device from the list of storage devices.

Your selection depends on whether an expandable storage device is available.

Option	Description
<b>To expand an existing extent</b>	Select the device for which the Expandable column reads YES. A storage device is reported as expandable when it has free space immediately after the extent.
<b>To add a new extent</b>	Select the device for which the Expandable column reads NO.

## 3 Review the **Current Disk Layout** to see the available configurations and click **Next**.

## 4 Select a configuration option from the bottom panel.

Depending on the current layout of the disk and on your previous selections, the options you see might vary.

Option	Description
<b>Use free space to add new extent</b>	Adds the free space on this disk as a new extent.
<b>Use free space to expand existing extent</b>	Expands an existing extent to a required capacity.
<b>Use free space</b>	Deploys an extent in the remaining free space of the disk. This option is available only when you are adding an extent.
<b>Use all available partitions</b>	Dedicates the entire disk to a single extent. This option is available only when you are adding an extent and when the disk you are formatting is not blank. The disk is reformatted, and the datastores and any data that it contains are erased.

## 5 Set the capacity for the extent.

The minimum extent size is 1.3GB. By default, the entire free space on the storage device is available.

## 6 Click **Next**.

## 7 Review the proposed layout and the new configuration of your datastore, and click **Finish**.

# Unmount VMFS or NFS Datastores

When you unmount a datastore, it remains intact, but can no longer be seen from the hosts that you specify. The datastore continues to appear on other hosts, where it remains mounted. If you unmount an NFS datastore from all hosts, it disappears from the inventory.

Do not perform any configuration operations that might result in I/O to the datastore while the unmount is in progress.

---

**Note** For NFS or VMFS datastore, make sure that the datastore is not used by vSphere HA heartbeating. vSphere HA heartbeating does not prevent you from unmounting the datastore. However, if the datastore is used for heartbeating, unmounting it might cause the host to fail and restart any active virtual machine.

---

### Prerequisites

Before unmounting a VMFS datastores, make sure that the following prerequisites are met:

- No virtual machines reside on the datastore.
- The datastore is not managed by Storage DRS.
- Storage I/O control is disabled for this datastore.

Browse to Datastores in the vSphere Web Client navigator. See [Display Datastore Information in the vSphere Web Client](#).

### Procedure

- 1 Right-click the datastore to unmount and select **Unmount Datastore**.
- 2 If the datastore is shared, specify which hosts should no longer access the datastore.
- 3 Confirm that you want to unmount the datastore.

### Results

After you unmount a VMFS datastore, the datastore is marked as inactive.

### What to do next

If you unmounted the VMFS datastore as a part of an orderly storage removal procedure, you can now detach the storage device that is backing the datastore. See [Detach Storage Devices](#).

## Mount VMFS Datastores

You can mount an unmounted VMFS datastore.

### Prerequisites

Browse to Datastores in the vSphere Web Client navigator. See [Display Datastore Information in the vSphere Web Client](#).

### Procedure

- 1 Right-click the datastore to mount and select **Mount Datastore**.

---

**Note** The datastore that is unmounted from some hosts while being mounted on others, is shown as active in the inventory.

---



- 2 Select the hosts that should access the datastore.

## Remove VMFS Datastores in the vSphere Web Client

You can delete any type of VMFS datastore, including copies that you have mounted without resignaturing. When you delete a datastore, it is destroyed and disappears from all hosts that have access to the datastore.

---

**Note** The datastore delete operation permanently deletes all files associated with virtual machines on the datastore. Although you can delete the datastore without unmounting, it is preferable that you unmount the datastore first.

---

### Prerequisites

- Remove or migrate all virtual machines from the datastore.
- Make sure that no other host is accessing the datastore.
- Disable Storage DRS for the datastore.
- Disable Storage I/O control for the datastore.
- Make sure that the datastore is not used for vSphere HA heartbeating.
- Browse to Datastores in the vSphere Web Client navigator. See [Display Datastore Information in the vSphere Web Client](#).

### Procedure

- 1 Right-click the datastore to remove.
- 2 Select **All vCenter Actions > Delete Datastore**.
- 3 Confirm that you want to remove the datastore.

## Storage Filtering

vCenter Server provides storage filters to help you avoid storage device corruption or performance degradation that might be caused by an unsupported use of storage devices. These filters are available by default.

Table 16-4. Storage Filters

Filter Name	Description
config.vpxd.filter.vmfsFilter (VMFS Filter)	Filters out storage devices, or LUNs, that are already used by a VMFS datastore on any host managed by vCenter Server. The LUNs do not show up as candidates to be formatted with another VMFS datastore or to be used as an RDM.
config.vpxd.filter.rdmFilter (RDM Filter)	<p>Filters out LUNs that are already referenced by an RDM on any host managed by vCenter Server. The LUNs do not show up as candidates to be formatted with VMFS or to be used by a different RDM.</p> <p>For your virtual machines to access the same LUN, the virtual machines must share the same RDM mapping file. For information about this type of configuration, see the <i>vSphere Resource Management</i> documentation.</p>
config.vpxd.filter.sameHostsAndTransportsFilter (Same Hosts and Transports Filter)	<p>Filters out LUNs ineligible for use as VMFS datastore extents because of host or storage type incompatibility. Prevents you from adding the following LUNs as extents:</p> <ul style="list-style-type: none"> <li>■ LUNs not exposed to all hosts that share the original VMFS datastore.</li> <li>■ LUNs that use a storage type different from the one the original VMFS datastore uses. For example, you cannot add a Fibre Channel extent to a VMFS datastore on a local storage device.</li> </ul>
config.vpxd.filter.hostRescanFilter (Host Rescan Filter)	<p>Automatically rescans and updates VMFS datastores after you perform datastore management operations. The filter helps provide a consistent view of all VMFS datastores on all hosts managed by vCenter Server.</p> <p><b>Note</b> If you present a new LUN to a host or a cluster, the hosts automatically perform a rescan no matter whether you have the Host Rescan Filter on or off.</p>

## Turn Off Storage Filters

When you perform VMFS datastore management operations, vCenter Server uses default storage protection filters. The filters help you to avoid storage corruption by retrieving only the storage devices that can be used for a particular operation. Unsuitable devices are not displayed for selection. You can turn off the filters to view all devices.

### Prerequisites

Before you change the device filters, consult with the VMware support team. You can turn off the filters only if you have other methods to prevent device corruption.

### Procedure

- 1 Browse to the vCenter Server in the vSphere Web Client object navigator.
- 2 Click the **Manage** tab, and click **Settings**.
- 3 Under **Settings**, click **Advanced Settings**, and click **Edit**.

#### 4 Specify the filter to turn off.

- a In the **Name** text box at the bottom of the page, enter an appropriate filter name.

Name	Description
<code>config.vpxd.filter.vmfsFilter</code>	VMFS Filter
<code>config.vpxd.filter.rdmFilter</code>	RDM Filter
<code>config.vpxd.filter.sameHostsAndTransportsFilter</code>	Same Hosts and Transports Filter
<code>config.vpxd.filter.hostRescanFilter</code>	Host Rescan Filter
<b>Note</b> If you turn off this filter, your hosts continue to perform a rescan each time you present a new LUN to a host or a cluster.	

- b In the **Value** text box, enter **False** for the specified key.

#### 5 Click **Add**, and click **OK** to save your changes.

You are not required to restart the vCenter Server system.

## Set Up Dynamic Disk Mirroring

Typically, you cannot use logical-volume manager software on virtual machines to mirror virtual disks. However, if your Microsoft Windows virtual machines support dynamic disks, you can protect the virtual machines from an unplanned storage device loss by mirroring virtual disks across two SAN LUNs.

### Prerequisites

- Use a Windows virtual machine that supports dynamic disks.
- Required privilege: **Advanced**

### Procedure

- 1 Create a virtual machine with two virtual disks.  
Make sure to place the disks on different datastores.
- 2 Log in to your virtual machine and configure the disks as dynamic mirrored disks.  
See Microsoft documentation.
- 3 After the disks synchronise, power off the virtual machine.
- 4 Change virtual machine settings to allow the use of dynamic disk mirroring.
  - a Right-click the virtual machine and select **Edit Settings**.
  - b Click the **VM Options** tab and expand the **Advanced** menu.
  - c Click **Edit Configuration** next to Configuration Parameters.

- d Click **Add Row** and add the following parameters:

Name	Value
scsi#.returnNoConnectDuringAPD	True
scsi#.returnBusyOnNoConnectStatus	False

- e Click **OK**.

## Collecting Diagnostic Information for ESXi Hosts on a Storage Device

During a host failure, ESXi must be able to save diagnostic information to a preconfigured location for diagnostic and technical support purposes.

Typically, a partition to collect diagnostic information, also called VMkernel core dump, is created on a local storage device during ESXi installation. You can override this default behavior if, for example, you use shared storage devices instead of local storage. To prevent automatic formatting of local devices, detach the devices from the host before you install ESXi and power on the host for the first time. You can later set up a location for collecting diagnostic information on a local or remote storage device.

When you use storage devices, you can select between two options of setting up core dump collection. You can use a preconfigured diagnostic partition on a storage device or use a file on a VMFS datastore.

- [Set Up a Device Partition as Core Dump Location](#)

Create a diagnostic partition for your ESXi host.

- [Set Up a File as Core Dump Location](#)

If the size of your available core dump partition size is insufficient, you can configure ESXi to generate core dump as a file.

## Set Up a Device Partition as Core Dump Location

Create a diagnostic partition for your ESXi host.

When you create a diagnostic partition, the following considerations apply:

- You cannot create a diagnostic partition on an iSCSI LUN accessed through the software iSCSI or dependent hardware iSCSI adapter. For more information about diagnostic partitions with iSCSI, see [General Boot from iSCSI SAN Recommendations](#).
- You cannot create a diagnostic partition on a software FCoE LUN.
- Unless you are using diskless servers, set up a diagnostic partition on a local storage.
- Each host must have a diagnostic partition of 2.5 GB. If multiple hosts share a diagnostic partition on a SAN LUN, the partition should be large enough to accommodate core dumps of all hosts.

- If a host that uses a shared diagnostic partition fails, reboot the host and extract log files immediately after the failure. Otherwise, the second host that fails before you collect the diagnostic data of the first host might not be able to save the core dump.

#### Procedure

1 Browse to the host in the vSphere Web Client navigator.

2 Right-click the host, and select **Add Diagnostic Partition**.

If you do not see this option, the host already has a diagnostic partition.

3 Specify the type of diagnostic partition.

Option	Description
<b>Private local</b>	Creates the diagnostic partition on a local disk. This partition stores fault information only for your host.
<b>Private SAN storage</b>	Creates the diagnostic partition on a non-shared SAN LUN. This partition stores fault information only for your host.
<b>Shared SAN storage</b>	Creates the diagnostic partition on a shared SAN LUN. This partition is accessed by multiple hosts and can store fault information for more than one host.

4 Click **Next**.

5 Select the device to use for the diagnostic partition and click **Next**.

6 Review the partition configuration information and click **Finish**.

### Verify a Diagnostic Partition

Use the `esxcli` command to verify whether a diagnostic partition is set.

In the procedure, `--server=server_name` specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

#### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

#### Procedure

- ◆ List partitions to verify that a diagnostic partition is set.

```
esxcli --server=server_name system coredump partition list
```

## Results

If a diagnostic partition is set, the command displays information about it. Otherwise, the command shows that no partition is activated and configured.

## What to do next

To manage the host's diagnostic partition, use the vCLI commands. See *vSphere Command-Line Interface Concepts and Examples*.

## Set Up a File as Core Dump Location

If the size of your available core dump partition size is insufficient, you can configure ESXi to generate core dump as a file.

Typically, the core dump partition for new installations of ESXi 5.5 and later is 2.5GB. For upgrades from previous releases to ESXi 5.5, the core dump partition is limited to 100MB for ESXi crashes. For many large systems, this limit is not enough space and the core dump log files are truncated. For upgrades with limits on 100MB partitions, during the boot process the system might create a core dump file on a VMFS datastore. If it does not create a core dump file, you can manually create the file.

---

**Note** Software iSCSI and software FCoE are not supported for core dump file locations.

---

## Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

## Procedure

- 1 Create a VMFS datastore core dump file by running the following command:

```
esxcli system coredump file add -d DATASTORE_UUID -f FILENAME
```

- 2 Verify that the file has been created:

```
esxcli system coredump file list
```

You can see the output similar to the following:

Path	Active Configured Size		
-----	-----	-----	-----
/vmfs/volumes/52b021c3-.../vmkdump/test.dumpfile	false	false	104857600

- 3 Set the core dump file for the host:

```
esxcli system coredump file set -p /vmfs/volumes/DATASTORE_UUID/vmkdump/FILENAME
```

For example:

```
esxcli system coredump file set -p /vmfs/volumes/52b021c3-  
f6b3da50-4c76-001d0904c5a5/vmkdump/test.dumpfile
```

- 4 Verify that the core dump file is active and configured:

#### **esxcli system coredump file list**

The output similar to the following indicates that the core dump file is active and configured:

Path	Active	Configured	Size
/vmfs/volumes/52b021c3-.../vmkdump/test.dumpfile	True	True	104857600

#### **What to do next**

For information about other commands you can use to manage the core dump files, see the *vSphere Command-Line Interface Reference* documentation.

## Checking Metadata Consistency with VOMA

Use VMware Ondisk Metadata Analyser (VOMA) when you experience problems with your VMFS datastore and need to check metadata consistency of VMFS or logical volume backing the VMFS volume.

#### **Problem**

The following examples show circumstances in which you might need to perform a metadata check:

- You experience SAN outages.
- After you rebuild RAID or perform a disk replacement.
- You see metadata errors in the `vmkernel.log` file.
- You are unable to access files on the VMFS datastore that are not in use by any other host.

#### **Solution**

To check metadata consistency, run VOMA from the CLI of an ESXi host version 5.1 or later. VOMA can check both the logical volume and the VMFS for metadata inconsistencies. You can use VOMA on VMFS3 and VMFS5 datastores. VOMA runs in a read-only mode and serves only to identify problems. VOMA does not fix errors that it detects. Consult VMware Support to resolve errors reported by VOMA.

Follow these guidelines when you use the VOMA tool:

- Make sure that the VMFS datastore you analyze does not span multiple extents. You can run VOMA only against a single-extent datastore.
- Power off any virtual machines that are running or migrate them to a different datastore.

Follow these steps when you use the VOMA tool to check VMFS metadata consistency.

- 1 Obtain the name and partition number of the device that backs the VMFS datastore that you need to check.

## #esxcli storage vmfs extent list

The Device Name and Partition columns in the output identify the device. For example:

Volume Name	XXXXXXX	Device Name	Partition
1TB_VMFSS	XXXXXXX	naa.600508e000000000b367477b3be3d703	3

### 2 Run VOMA to check for VMFS errors.

Provide the absolute path to the device partition that backs the VMFS datastore, and provide a partition number with the device name. For example:

```
# voma -m vmfs -f check -d /vmfs/devices/disks/
naa.600508e000000000b367477b3be3d703:3
```

The output lists possible errors. For example, the following output indicates that the heartbeat address is invalid.

```
XXXXXXXXXXXXXXXXXXXXX
Phase 2: Checking VMFS heartbeat region
ON-DISK ERROR: Invalid HB address
Phase 3: Checking all file descriptors.
Phase 4: Checking pathname and connectivity.
Phase 5: Checking resource reference counts.

Total Errors Found:          1
```

The VOMA tool uses the following options.

**Table 16-5. VOMA Command Options**

Command Option	Description
-m   --module	The module to run: vmfs or lvm. If you specify vmfs, minimal checks are performed for LVM as well. The default module is vmfs.
-f   --func	Functions to be performed: query - list functions supported by module. check - check for errors.
-d   --device	Device or disk to be inspected. Make sure to provide the absolute path to the device partition backing the VMFS datastore. For example, /vmfs/devices/disks/naa.00000000000000000000000000000000:1.
-s   --logfile	Specify the log file to output the results.
-v   --version	Display the version of VOMA.
-h   --help	Display the help message for the VOMA command.

## Configuring VMFS Pointer Block Cache

You can use advanced VMFS parameters to configure pointer block cache.



As the size of virtual machine files on VMFS datastores increases, the number of pointer blocks used by those files also increases. Pointer blocks are used to address file blocks in large virtual machine files and virtual disks on a VMFS datastore.

You can configure the minimum and maximum sizes of the pointer block cache on each ESXi host. When the size of the pointer block cache approaches the configured maximum size, an eviction mechanism removes some pointer block entries from the cache to allow for new pointer blocks to be read in.

Base the maximum size of the pointer block cache on the working size of all open virtual disk files that reside on VMFS datastores. All VMFS datastores on the host use a single pointer block cache.

The minimum value is based on the minimum guaranteed memory that the system can allocate to the cache. 1TB of open file space requires approximately 4MB of memory.

Use the following parameters to configure the minimum and maximum values for pointer block cache.

**Table 16-6. Advanced Parameters to Regulate Pointer Block Cache**

Parameter	Values	Description
VMFS3.MaxAddressableSpaceTB	Default value is 32 (in TB).	Maximum size of all open files that VMFS cache will support before eviction starts.
VMFS3.MinAddressableSpaceTB	Default value is 10 (in TB).	Minimum size of all open files that VMFS cache guarantees to support.

For information about how to configure these advanced parameters, see [Set Advanced Host Attributes in the vSphere Web Client](#).

## Set Advanced Host Attributes in the vSphere Web Client

You can set advanced attributes for a host.

**Caution** Changing advanced options is considered unsupported unless VMware technical support or a KB article instruct you to do so. In all other cases, changing these options is considered unsupported. In most cases, the default settings produce the optimum result.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab and click **Settings**.
- 3 Under **System**, click **Advanced System Settings**.
- 4 In Advanced Settings, select the appropriate item.
- 5 Click the **Edit** button to edit the value.
- 6 Click **OK**.

# Raw Device Mapping

# 17

Raw device mapping (RDM) provides a mechanism for a virtual machine to have direct access to a LUN on the physical storage subsystem (Fibre Channel or iSCSI only).

The following topics contain information about RDMs and provide instructions on how to create and manage RDMs.

This chapter includes the following topics:

- [About Raw Device Mapping](#)
- [Raw Device Mapping Characteristics](#)
- [Create Virtual Machines with RDMs in the vSphere Web Client](#)
- [Manage Paths for a Mapped LUN in the vSphere Web Client](#)

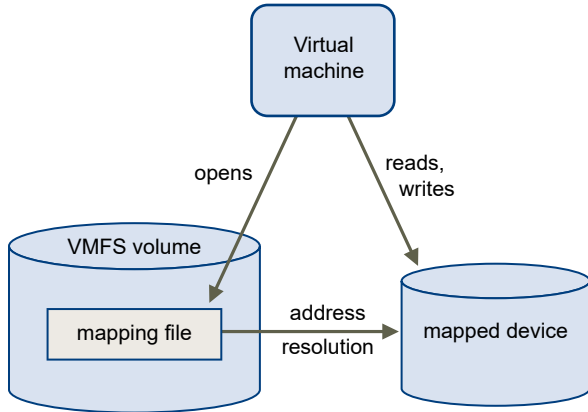
## About Raw Device Mapping

An RDM is a mapping file in a separate VMFS volume that acts as a proxy for a raw physical storage device. The RDM allows a virtual machine to directly access and use the storage device. The RDM contains metadata for managing and redirecting disk access to the physical device.

The file gives you some of the advantages of direct access to a physical device while keeping some advantages of a virtual disk in VMFS. As a result, it merges VMFS manageability with raw device access.

RDMs can be described in terms such as mapping a raw device into a datastore, mapping a system LUN, or mapping a disk file to a physical disk volume. All these terms refer to RDMs.

Figure 17-1. Raw Device Mapping



Although VMware recommends that you use VMFS datastores for most virtual disk storage, on certain occasions, you might need to use raw LUNs or logical disks located in a SAN.

For example, you need to use raw LUNs with RDMs in the following situations:

- When SAN snapshot or other layered applications run in the virtual machine. The RDM better enables scalable backup offloading systems by using features inherent to the SAN.
- In any MSCS clustering scenario that spans physical hosts — virtual-to-virtual clusters as well as physical-to-virtual clusters. In this case, cluster data and quorum disks should be configured as RDMs rather than as virtual disks on a shared VMFS.

Think of an RDM as a symbolic link from a VMFS volume to a raw LUN. The mapping makes LUNs appear as files in a VMFS volume. The RDM, not the raw LUN, is referenced in the virtual machine configuration. The RDM contains a reference to the raw LUN.

Using RDMs, you can:

- Use vMotion to migrate virtual machines using raw LUNs.
- Add raw LUNs to virtual machines using the vSphere Web Client.
- Use file system features such as distributed file locking, permissions, and naming.

Two compatibility modes are available for RDMs:

- Virtual compatibility mode allows an RDM to act exactly like a virtual disk file, including the use of snapshots.
- Physical compatibility mode allows direct access of the SCSI device for those applications that need lower level control.

## Benefits of Raw Device Mapping

An RDM provides a number of benefits, but it should not be used in every situation. In general, virtual disk files are preferable to RDMs for manageability. However, when you need raw devices, you must use the RDM.

RDM offers several benefits.

### **User-Friendly Persistent Names**

Provides a user-friendly name for a mapped device. When you use an RDM, you do not need to refer to the device by its device name. You refer to it by the name of the mapping file, for example:

```
/vmfs/volumes/myVolume/myVMDirectory/myRawDisk.vmdk
```

### **Dynamic Name Resolution**

Stores unique identification information for each mapped device. VMFS associates each RDM with its current SCSI device, regardless of changes in the physical configuration of the server because of adapter hardware changes, path changes, device relocation, and so on.

### **Distributed File Locking**

Makes it possible to use VMFS distributed locking for raw SCSI devices. Distributed locking on an RDM makes it safe to use a shared raw LUN without losing data when two virtual machines on different servers try to access the same LUN.

### **File Permissions**

Makes file permissions possible. The permissions of the mapping file are enforced at file-open time to protect the mapped volume.

### **File System Operations**

Makes it possible to use file system utilities to work with a mapped volume, using the mapping file as a proxy. Most operations that are valid for an ordinary file can be applied to the mapping file and are redirected to operate on the mapped device.

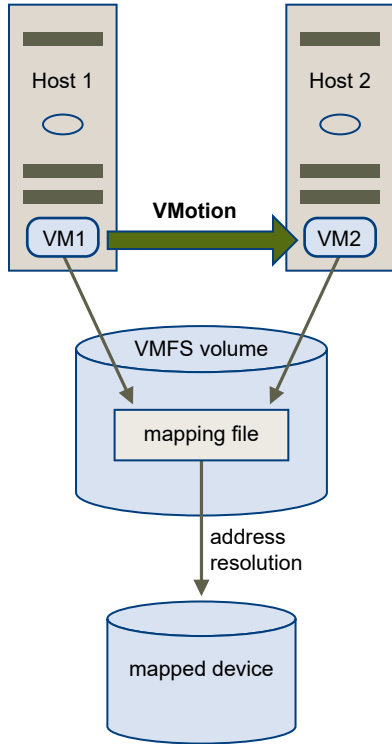
### **Snapshots**

Makes it possible to use virtual machine snapshots on a mapped volume. Snapshots are not available when the RDM is used in physical compatibility mode.

### **vMotion**

Lets you migrate a virtual machine with vMotion. The mapping file acts as a proxy to allow vCenter Server to migrate the virtual machine by using the same mechanism that exists for migrating virtual disk files.

Figure 17-2. vMotion of a Virtual Machine Using Raw Device Mapping



### SAN Management Agents

Makes it possible to run some SAN management agents inside a virtual machine. Similarly, any software that needs to access a device by using hardware-specific SCSI commands can be run in a virtual machine. This kind of software is called SCSI target-based software. When you use SAN management agents, select a physical compatibility mode for the RDM.

### N-Port ID Virtualization (NPIV)

Makes it possible to use the NPIV technology that allows a single Fibre Channel HBA port to register with the Fibre Channel fabric using several worldwide port names (WWPNs). This ability makes the HBA port appear as multiple virtual ports, each having its own ID and virtual port name. Virtual machines can then claim each of these virtual ports and use them for all RDM traffic.

---

**Note** You can use NPIV only for virtual machines with RDM disks.

---

VMware works with vendors of storage management software to ensure that their software functions correctly in environments that include ESXi. Some applications of this kind are:

- SAN management software
- Storage resource management (SRM) software
- Snapshot software
- Replication software

Such software uses a physical compatibility mode for RDMs so that the software can access SCSI devices directly.

Various management products are best run centrally (not on the ESXi machine), while others run well on the virtual machines. VMware does not certify these applications or provide a compatibility matrix. To find out whether a SAN management application is supported in an ESXi environment, contact the SAN management software provider.

## RDM Considerations and Limitations

Certain considerations and limitations exist when you use RDMs.

- The RDM is not available for direct-attached block devices or certain RAID devices. The RDM uses a SCSI serial number to identify the mapped device. Because block devices and some direct-attach RAID devices do not export serial numbers, they cannot be used with RDMs.
- If you are using the RDM in physical compatibility mode, you cannot use a snapshot with the disk. Physical compatibility mode allows the virtual machine to manage its own, storage-based, snapshot or mirroring operations.

Virtual machine snapshots are available for RDMs with virtual compatibility mode.

- You cannot map to a disk partition. RDMs require the mapped device to be a whole LUN.
- If you use vMotion to migrate virtual machines with RDMs, make sure to maintain consistent LUN IDs for RDMs across all participating ESXi hosts.
- Flash Read Cache does not support RDMs in physical compatibility. Virtual compatibility RDMs are supported with Flash Read Cache.

## Raw Device Mapping Characteristics

An RDM is a special mapping file in a VMFS volume that manages metadata for its mapped device. The mapping file is presented to the management software as an ordinary disk file, available for the usual file-system operations. To the virtual machine, the storage virtualization layer presents the mapped device as a virtual SCSI device.

Key contents of the metadata in the mapping file include the location of the mapped device (name resolution), the locking state of the mapped device, permissions, and so on.

## RDM Virtual and Physical Compatibility Modes

You can use RDMs in virtual compatibility or physical compatibility modes. Virtual mode specifies full virtualization of the mapped device. Physical mode specifies minimal SCSI virtualization of the mapped device, allowing the greatest flexibility for SAN management software.

In virtual mode, the VMkernel sends only READ and WRITE to the mapped device. The mapped device appears to the guest operating system exactly the same as a virtual disk file in a VMFS volume. The real hardware characteristics are hidden. If you are using a raw disk in virtual mode, you can realize the benefits of VMFS such as advanced file locking for data protection and snapshots for streamlining development processes. Virtual mode is also more portable across storage hardware than physical mode, presenting the same behavior as a virtual disk file.

In physical mode, the VMkernel passes all SCSI commands to the device, with one exception: the REPORT LUNs command is virtualized so that the VMkernel can isolate the LUN to the owning virtual machine. Otherwise, all physical characteristics of the underlying hardware are exposed. Physical mode is useful to run SAN management agents or other SCSI target-based software in the virtual machine. Physical mode also allows virtual-to-physical clustering for cost-effective high availability.

VMFS5 supports greater than 2TB disk size for RDMs in virtual and physical modes. You cannot relocate larger than 2TB RDMs to datastores other than VMFS5.

## Dynamic Name Resolution

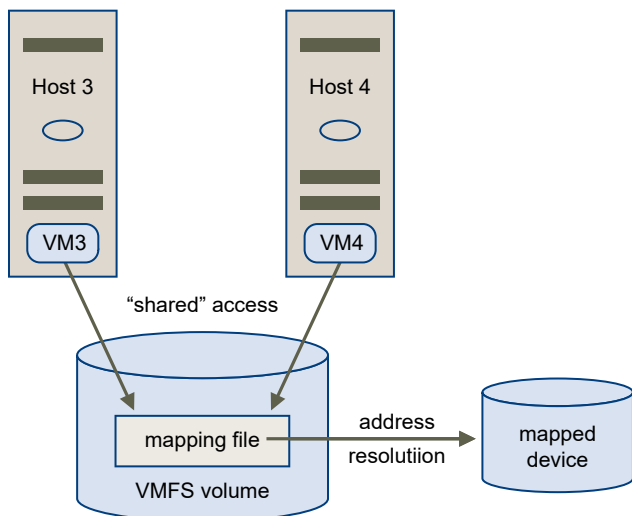
The RDM file supports dynamic name resolution when a path to a raw device changes.

VMFS uniquely identifies all mapped storage devices, and the identification is stored in its internal data structures. Any change in the path to a raw device, such as a Fibre Channel switch failure or the addition of a new HBA, can change the device name. Dynamic name resolution resolves these changes and automatically associates the original device with its new name.

## Raw Device Mapping with Virtual Machine Clusters

Use an RDM with virtual machine clusters that need to access the same raw LUN for failover scenarios. The setup is similar to that of a virtual machine cluster that accesses the same virtual disk file, but an RDM replaces the virtual disk file.

Figure 17-3. Access from Clustered Virtual Machines



## Comparing Available SCSI Device Access Modes

The ways of accessing a SCSI-based storage device include a virtual disk file on a VMFS datastore, virtual mode RDM, and physical mode RDM.

To help you choose among the available access modes for SCSI devices, the following table provides a quick comparison of features available with the different modes.

**Table 17-1. Features Available with Virtual Disks and Raw Device Mappings**

ESXi Features	Virtual Disk File	Virtual Mode RDM	Physical Mode RDM
SCSI Commands Passed Through	No	No	Yes REPORT LUNs is not passed through
vCenter Server Support	Yes	Yes	Yes
Snapshots	Yes	Yes	No
Distributed Locking	Yes	Yes	Yes
Clustering	Cluster-in-a-box only	Cluster-in-a-box cluster-across-boxes	Physical-to-virtual clustering cluster-across-boxes
SCSI Target-Based Software	No	No	Yes

VMware recommends that you use virtual disk files for the cluster-in-a-box type of clustering. If you plan to reconfigure your cluster-in-a-box clusters as cluster-across-boxes clusters, use virtual mode RDMs for the cluster-in-a-box clusters.

## Create Virtual Machines with RDMs in the vSphere Web Client

When you give your virtual machine direct access to a raw SAN LUN, you create an RDM disk that resides on a VMFS datastore and points to the LUN. You can create the RDM as an initial disk for a new virtual machine or add it to an existing virtual machine. When creating the RDM, you specify the LUN to be mapped and the datastore on which to put the RDM.

Although the RDM disk file has the same .vmdk extension as a regular virtual disk file, the RDM contains only mapping information. The actual virtual disk data is stored directly on the LUN.

This procedure assumes that you are creating a new virtual machine. For information, see the *vSphere Virtual Machine Administration* documentation.

### Procedure

- 1 Right-click any inventory object that is a valid parent object of a virtual machine, such as a datacenter, folder, cluster, resource pool, or host, and select **New Virtual Machine**.
- 2 Select **Create a new virtual machine** and click **Next**.
- 3 Follow the steps required to create a virtual machine.
- 4 On the Customize Hardware page, click the **Virtual Hardware** tab.



- 5 (Optional) To delete the default virtual hard disk that the system created for your virtual machine, move your cursor over the disk and click the **Remove** icon.
- 6 From the **New** drop-down menu at the bottom of the page, select **RDM Disk** and click **Add**.
- 7 From the list of SAN devices or LUNs, select a raw LUN for your virtual machine to access directly and click **OK**.

The system creates an RDM disk that maps your virtual machine to the target LUN. The RDM disk is shown on the list of virtual devices as a new hard disk.

- 8 Click the **New Hard Disk** triangle to expand the properties for the RDM disk.
- 9 Select a location for the RDM disk.

You can place the RDM on the same datastore where your virtual machine configuration files reside, or select a different datastore.

**Note** To use vMotion for virtual machines with enabled NPIV, make sure that the RDM files and the virtual machine files are located on the same datastore. You cannot perform Storage vMotion when NPIV is enabled.

- 10 Select a compatibility mode.

Option	Description
<b>Physical</b>	Allows the guest operating system to access the hardware directly. Physical compatibility is useful if you are using SAN-aware applications on the virtual machine. However, a virtual machine with a physical compatibility RDM cannot be cloned, made into a template, or migrated if the migration involves copying the disk.
<b>Virtual</b>	Allows the RDM to behave as if it were a virtual disk, so you can use such features as taking snapshots, cloning, and so on. When you clone the disk or make a template out of it, the contents of the LUN are copied into a .vmdk virtual disk file. When you migrate a virtual compatibility mode RDM, you can migrate the mapping file or copy the contents of the LUN into a virtual disk.

- 11 If you selected virtual compatibility mode, select a disk mode.

Disk modes are not available for RDM disks using physical compatibility mode.

Option	Description
<b>Dependent</b>	Dependent disks are included in snapshots.
<b>Independent - Persistent</b>	Disks in persistent mode behave like conventional disks on your physical computer. All data written to a disk in persistent mode are written permanently to the disk.
<b>Independent - Nonpersistent</b>	Changes to disks in nonpersistent mode are discarded when you power off or reset the virtual machine. With nonpersistent mode, you can restart the virtual machine with a virtual disk in the same state every time. Changes to the disk are written to and read from a redo log file that is deleted when you power off or reset.

12 Click **OK**.

## Manage Paths for a Mapped LUN in the vSphere Web Client

When you use virtual machines with RDMs, you can manage paths for mapped raw LUNs.

### Procedure

- 1 In the vSphere Web Client, browse to the virtual machine.
- 2 Right-click the virtual machine and select **Edit Settings**.
- 3 Click the **Virtual Hardware** tab and click **Hard Disk** to expand the disk options menu.
- 4 Click **Manage Paths**.
- 5 Use the Edit Multipathing Policies dialog box to enable or disable paths, set multipathing policy, and specify the preferred path.

For information on managing paths, see [Chapter 23 Understanding Multipathing and Failover](#).

# Working with Solid State Disks

# 18

In addition to regular storage hard disk drives, ESXi supports Solid State Disks (SSDs).

Unlike the regular hard disks that are electromechanical devices containing moving parts, SSDs use semiconductors as their storage medium and have no moving parts. Typically, SSDs are very resilient and provide faster access to data.

Support of SSDs with ESXi provides the following benefits:

- You can use VMFS datastores that are created on SSD storage devices to allocate space for ESXi host cache.
- You can use SSDs for such features as Flash Read Cache and Virtual SAN. Local SSDs are required for these features.
- Use of SSDs increases virtual machine consolidation ratio as SSDs can provide very high I/O throughput.
- Guest operating systems might be able to identify virtual SSD devices.

To detect SSD devices, ESXi uses an inquiry mechanism based on T10 standards. ESXi can automatically detect SSD devices on a number of storage arrays. Check with your vendor whether your storage array supports ESXi SSD device detection.

This chapter includes the following topics:

- [Benefits of SSD Enablement](#)
- [Auto-Detection of SSD Devices](#)
- [Tag Devices as SSD](#)
- [Untag an SSD Device](#)
- [Untag an Automatically Detected SSD Device](#)
- [Tag Devices as Local](#)
- [Monitor SSD Devices](#)
- [Identify SSD Devices](#)
- [Identifying a Virtual SSD Device](#)

- [Best Practices for SSD Devices](#)
- [Swapping to Host Cache](#)
- [Local SSDs Are Undetectable](#)

## Benefits of SSD Enablement

SSDs are very resilient and provide faster access to data.

SSD enablement results in several benefits:

- Use VMFS datastores that are created on SSD storage devices to allocate space for ESXi host cache.
- Use local SSDs for such features as Flash Read Cache and Virtual SAN.
- Use of SSDs increases virtual machine consolidation ratio as SSDs can provide very high I/O throughput.
- It supports identification of virtual SSD device by the guest operating system.

## Auto-Detection of SSD Devices

ESXi enables automatic detection of SSD devices by using an inquiry mechanism based on T10 standards.

ESXi enables detection of the SSD devices on a number of storage arrays. Check with your vendor whether your storage array supports ESXi SSD device detection.

You can use PSA SATP claim rules to tag devices that cannot be auto-detected.

## Tag Devices as SSD

You can use PSA SATP claim rules to tag SSD devices that are not detected automatically.

Only devices that are consumed by the PSA Native Multipathing (NMP) plugin can be tagged.

If you need more information about the commands listed in this topic, see the *Getting Started with vSphere Command-Line Interfaces* and *vSphere Command-Line Interface Concepts and Examples* documentation.

### Procedure

- 1 Identify the device to be tagged and its SATP.

```
esxcli storage nmp device list
```

The command results in the following information.

```
naa.6006016015301d00167ce6e2ddb3de11
Device Display Name: DGC Fibre Channel Disk (naa.6006016015301d00167ce6e2ddb3de11)
Storage Array Type: VMW_SATP_CX
Storage Array Type Device Config: {navireg ipfilter}
Path Selection Policy: VMW_PSP_MRU
Path Selection Policy Device Config: Current Path=vmhba4:C0:T0:L25
Working Paths: vmhba4:C0:T0:L25
```

**2** Note down the SATP associated with the device.

**3** Add a PSA claim rule to mark the device as SSD.

- ◆ You can add a claim rule by specifying the device name.

```
esxcli storage nmp satp rule add -s SATP --device device_name --option="enable_ssd"
```

- ◆ You can add a claim rule by specifying the vendor name and the model name.

```
esxcli storage nmp satp rule add -s SATP -V vendor_name -M model_name --option="enable_ssd"
```

- ◆ You can add a claim rule based on the transport protocol.

```
esxcli storage nmp satp rule add -s SATP --transport transport_protocol --option="enable_ssd"
```

- ◆ You can add a claim rule based on the driver name.

```
esxcli storage nmp satp rule add -s SATP --driver driver_name --option="enable_ssd"
```

**4** Reclaim the device.

```
esxcli storage core claiming reclaim --device device_name
```

**5** Verify if devices are tagged as SSD.

```
esxcli storage core device list -d device_name
```

The command output indicates if a listed device is tagged as SSD.

```
Is SSD: true
```

### What to do next

If the SSD device that you want to tag is shared among multiple hosts, make sure that you tag the device from all the hosts that share the device.

## Untag an SSD Device

You can untag a device tagged as SSD whenever required.

This topic lists commands that you need to use to untag an SSD device. For more information about these commands, see the *Getting Started with vSphere Command-Line Interfaces* and *vSphere Command-Line Interface Concepts and Examples* documentation.

### Procedure

- 1 Identify the device that you want to untag.

```
esxcli storage nmp device list
```

You might see the output similar to the following:

```
naa.6006016015301d00167ce6e2ddb3de11
Device Display Name: DGC Fibre Channel Disk (naa.6006016015301d00167ce6e2ddb3de11)
Storage Array Type: VMW_SATP_CX
Storage Array Type Device Config: {navireg ipfilter}
Path Selection Policy: VMW_PSP_MRU
Path Selection Policy Device Config: Current Path=vmhba4:C0:T0:L25
Working Paths: vmhba4:C0:T0:L25
```

- 2 Note down the SATP associated with the device.
- 3 Remove the claim rule that was used to tag the device.

```
esxcli storage nmp satp rule remove -s SATP --device device_name
```

For example,

```
esxcli storage nmp satp rule remove -s VMW_SATP_CX --device naa.6006016042fa19010a12d9b16d6ade11
```

- 4 Reclaim the device.

```
esxcli storage core claiming reclaim --device naa.6006016042fa19010a12d9b16d6ade11
```

- 5 Check the device status by running the following command.

```
esxcli storage core device list -d device_name
```

The command output indicates whether the disk is untagged.

```
Is SSD: false
```

## Untag an Automatically Detected SSD Device

You can tag an automatically detected SSD device as a non-SSD device.

### Procedure

- 1 Identify the device that you want to untag.

```
esxcli storage nmp device list
```

You might see the output similar to the following:

```
naa.6006016015301d00167ce6e2ddb3de11
Device Display Name: DGC Fibre Channel Disk (naa.6006016015301d00167ce6e2ddb3de11)
Storage Array Type: VMW_SATP_CX
Storage Array Type Device Config: {navireg ipfilter}
Path Selection Policy: VMW_PSP_MRU
Path Selection Policy Device Config: Current Path=vmhba4:C0:T0:L25
Working Paths: vmhba4:C0:T0:L25
```

**2** Note down the SATP associated with the device.

**3** Add a claim rule with the option *disable\_ssd*.

```
esxcli storage nmp satp rule add -s SATP --device device_name --option disable_ssd
```

For example,

```
esxcli storage nmp satp rule add -s VMW_SATP_CX --device naa.6006016042fa19010a12d9b16d6ade11 --
option disable_ssd
```

**4** Reclaim the device.

```
esxcli storage core claiming reclaim --device naa.6006016042fa19010a12d9b16d6ade11
```

**5** Check the device status by running the following command.

```
esxcli storage core device list -d device_name
```

The command output indicates whether the disk is untagged.

```
Is SSD: false
```

## Tag Devices as Local

ESXi enables you to tag devices as local. This is useful in cases when ESXi is unable to determine whether certain SAS devices are local or remote.

For more information about the commands listed in this topic, see the *Getting Started with vSphere Command-Line Interfaces* and *vSphere Command-Line Interface Concepts and Examples* documentation.

### Prerequisites

- Make sure that the device is not shared.
- Power off virtual machines that reside on the device and unmount an associated datastore.

## Procedure

- 1 Identify the device to be tagged and its SATP:

```
esxcli storage nmp device list
```

You might see the output similar to the following:

```
naa.000000000000000000000000000000001234
Device Display Name: DGC Fibre Channel Disk (naa.000000000000000000000000000000001234)
Storage Array Type: VMW_SATP_CX
Storage Array Type Device Config: {navireg ipfilter}
Path Selection Policy: VMW_PSP_MRU
Path Selection Policy Device Config: Current Path=vmhba4:C0:T0:L25
Working Paths: vmhba4:C0:T0:L25
```

- 2 Note down the SATP associated with the device.
- 3 Run this command to add a PSA claim rule that marks the devices as local. Use the SATP associated with the device from the output in [Step 1](#).

```
esxcli storage nmp satp rule add -s SATP_name --device device_name --option="enable_local"
```

For example,

```
esxcli storage nmp satp rule add -s VMW_SATP_CX --device naa.000000000000000000000000000000001234 --
option="enable_local"
```

- 4 Reclaim the device. For example,

```
esxcli storage core claiming reclaim --device naa.000000000000000000000000000000001234
```

- 5 Check the status by running the following command:

```
esxcli storage core device list -d device_name
```

The command output indicates that the disk is local.

```
Is Local: true
```

## Monitor SSD Devices

You can monitor certain critical SSD parameters, including Media Wearout Indicator, Temperature, and Reallocated Sector Count, from an ESXi host.

Use the `esxcli` command to monitor SSD devices.

In the procedure, **--server=*server\_name*** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.



### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

### Procedure

- ◆ Run the following command to display the SSD statistics:

```
esxcli server=server_name storage core device smart get -d=SSD device_ID
```

## Identify SSD Devices

You can identify the SSD devices in your storage network.

### Prerequisites

Before you identify an SSD device, ensure that the device is tagged as SSD.

### Procedure

- 1 List the devices.

```
esxcli storage core device list
```

The command output includes the following information about the listed device.

```
Is SSD: true
```

- 2 Verify whether the value of the flag `Is SSD` is true.

## Identifying a Virtual SSD Device

ESXi allows operating systems to auto-detect VMDKs residing on SSD datastores as SSD devices.

To verify if this feature is enabled, guest operating systems can use standard inquiry commands such as SCSI VPD Page (B1h) for SCSI devices and ATA IDENTIFY DEVICE (Word 217) for IDE devices.

For linked clones, native snapshots, and delta-disks, the inquiry commands report the virtual SSD status of the base disk.

Operating systems can auto-detect a VMDK as SSD under the following conditions:

- Detection of virtual SSDs is supported on ESXi 5.x hosts and Virtual Hardware version 8 or later.
- Detection of virtual SSDs is supported only on VMFS5 or later.
- If VMDKs are located on shared VMFS datastores with SSD device extents, the device must be marked as SSD on all hosts.

- For a VMDK to be detected as virtual SSD, all underlying physical extents should be SSD-backed.

## Best Practices for SSD Devices

Follow these best practices when you use SSD devices in vSphere environment.

- Make sure to use the latest firmware with SSD devices. Frequently check with your storage vendors for any updates.
- Carefully monitor how intensively you use the SSD device and calculate its estimated lifetime. The lifetime expectancy depends on how actively you continue to use the SSD device.

## Estimate Lifetime of Flash Devices

When working with flash devices, monitor how actively you use them and calculate their estimated lifetime.

Typically, storage vendors provide reliable lifetime estimates for a flash device under ideal conditions. For example, a vendor might guarantee a lifetime of 5 years under the condition of 20 GB writes per day. However, the more realistic life expectancy of the device depends on how many writes per day your ESXi host actually generates. Follow these steps to calculate the lifetime of the flash device.

### Prerequisites

Note the number of days passed since the last reboot of your ESXi host. For example, ten days.

### Procedure

- 1 Obtain the total number of blocks written to the flash device since the last reboot.

Run the **esxcli storage core device stats get -d=device\_ID** command. For example:

```
~ # esxcli storage core device stats get -d t10.xxxxxxxxxxxxxx
Device: t10.xxxxxxxxxxxxxx
Successful Commands: xxxxxxxx
Blocks Read: xxxxxxxx
Blocks Written: 629145600
Read Operations: xxxxxxxx
```

The Blocks Written item in the output shows the number of blocks written to the device since the last reboot. In this example, the value is 629,145,600. After each reboot, it resets to 0.

- 2 Calculate the total number of writes and convert to GB.

One block is 512 bytes. To calculate the total number of writes, multiply the Blocks Written value by 512, and convert the resulting value to GB.

In this example, the total number of writes since the last reboot is approximately 322 GB.

### 3 Estimate the average number of writes per day in GB.

Divide the total number of writes by the number of days since the last reboot.

If the last reboot was ten days ago, you get 32 GB of writes per day. You can average this number over the time period.

### 4 Estimate lifetime of your device by using the following formula:

*vendor provided number of writes per day times vendor provided life span divided by actual average number of writes per day*

For example, if your vendor guarantees a lifetime of 5 years under the condition of 20 GB writes per day, and the actual number of writes per day is 30 GB, the life span of your flash device will be approximately 3.3 years.

## Swapping to Host Cache

You can create VMFS datastores on Solid State Disks (SSD), and then use these datastores to allocate space for host cache. The host reserves a certain amount of space for swapping to host cache.

The host cache is made up of files on a low-latency disk that ESXi uses as a write back cache for virtual machine swap files. The cache is shared by all virtual machines running on the host. Host-level swapping of virtual machine pages makes the best use of potentially limited SSD space.

Using swap to host cache is not the same as placing regular swap files on SSD-backed datastores. Even if you enable swap to host cache, the host still needs to create regular swap files. However, when you use swap to host cache, the speed of the storage where the host places regular swap files is less important.

The Host Cache Configuration page allows you to view the amount of space on a datastore that a host can use to swap to host cache. Only SSD-backed datastores appear in the list of datastores on the Host Cache Configuration page.

If you have a virtual flash license, you can use virtual flash resource to configure cache on the host. For information, see [Chapter 21 About VMware vSphere Flash Read Cache](#).

## Configure the Host Cache in the vSphere Web Client

You can change the percentage of space allocated for host cache or disable the host's ability to swap to host cache.

Use this procedure if your ESXi host is version 5.1 or earlier, or if you do not have a virtual flash licence. To configure host cache on an ESXi 5.5 or later host that has a virtual flash licence, use the virtual flash resource. For information, see [Chapter 21 About VMware vSphere Flash Read Cache](#).

### Prerequisites

You must have an SSD-backed datastore in your inventory.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab.
- 3 Under the Storage tab, click **Host Cache Configuration**.
- 4 Select the datastore in the list and click the **Allocate space for host cache** icon.
- 5 Select a size for the host cache allocation on the drive.
- 6 To disable the ability for the host to swap to host cache on a per-datastore basis, deselect the **Allocate space for host cache** check box.
- 7 Click **OK**.

## Local SSDs Are Undetectable

If you query for local SSDs during creation of a Flash Read Cache resource, the ESXi host might not return a complete list of the local SSDs.

### Problem

ESXi might not be able to automatically detect SSDs, or recognize them as local.

### Cause

ESXi does not recognize certain devices as SSDs when their vendors do not support automatic SSD detection. In other cases, some non-SATA SAS SSDs might not be detected as local. When disks are not recognized as local SSDs, they are excluded from the list of SSDs available for such features as Flash Read Cache because the feature requires only local SSDs.

### Solution

- ◆ If ESXi does not automatically recognize its disks as SSDs, tag them as SSD disks. See [Tag Devices as SSD](#).
- ◆ If ESXi does not detect SSD disks as local, manually set them as local. See [Tag Devices as Local](#).

# Working with Virtual SAN

# 19

Virtual SAN is a distributed layer of software that runs natively as a part of the ESXi hypervisor. Virtual SAN aggregates local or direct-attached storage disks of a host cluster and creates a single storage pool shared across all hosts of the cluster.

While supporting VMware features that require shared storage, such as HA, vMotion, and DRS, Virtual SAN eliminates the need for an external shared storage and simplifies storage configuration and virtual machine provisioning activities.

For information about troubleshooting Virtual SAN, see the *vSphere Troubleshooting* documentation.

This chapter includes the following topics:

- [About Virtual SAN](#)
- [Enabling Virtual SAN](#)
- [Managing Disk Groups](#)
- [Monitoring Virtual SAN](#)
- [Managing Virtual SAN](#)
- [Virtual SAN and Storage Policy-Based Management](#)

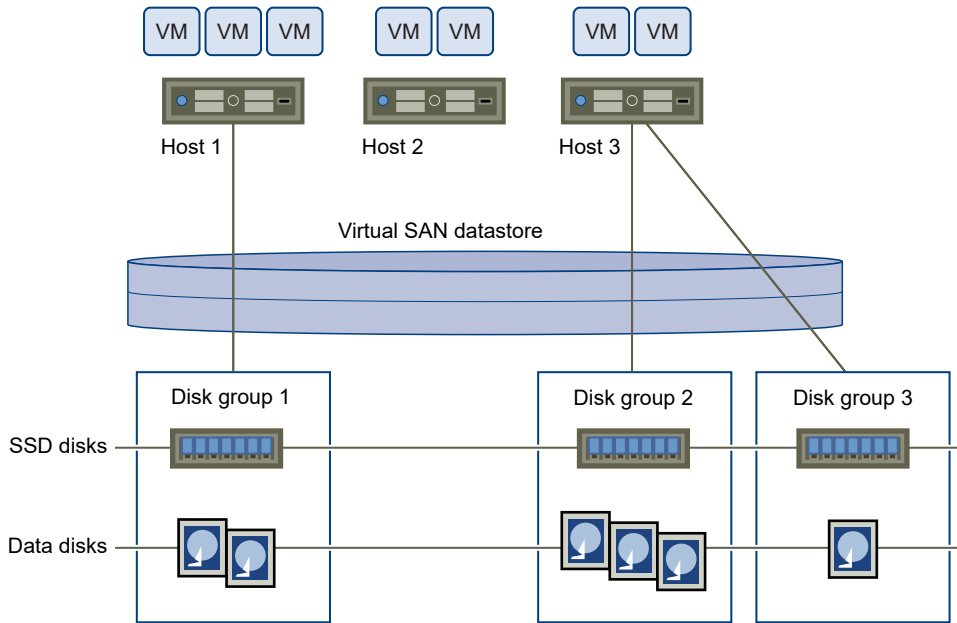
## About Virtual SAN

Virtual SAN virtualizes local physical storage resources of ESXi hosts and turns them into pools of storage that can be carved up and assigned to virtual machines and applications according to their quality of service requirements.

You can activate Virtual SAN when you create host clusters or enable Virtual SAN on existing clusters. When enabled, Virtual SAN aggregates all local storage disks available on the hosts into a single datastore shared by all hosts. You can later expand the datastore by adding storage devices or hosts to the cluster.

The hosts in the Virtual SAN cluster do not need to be identical. Even the hosts that have no local disks can participate and run their virtual machines on the Virtual SAN datastore, as, for example, host 2 in the illustration.

If a host contributes its local storage to the Virtual SAN datastore, the host must provide one SSD and at least one HDD, also called data disk. The disks on the contributing host form a disk group. Each disk group must include one SSD and at least one or multiple data disks. The disk group uses the SSD disk for read caching and write buffering, while the data disks are used for persistent storage. You can have multiple disk groups per host. For example, in the illustration, host 1 contains one disk group, while host 3 has two disk groups.



## Virtual SAN and Virtual Machine Storage Policies

When enabled, Virtual SAN works together with virtual machine storage policies. You use storage policies to define virtual machine storage requirements such as performance and availability in the form of a profile. The policy requirements are pushed to the Virtual SAN layer when the virtual disk is created. Virtual SAN lays out the virtual disk across the Virtual SAN datastore to meet the specified requirements.

Virtual SAN monitors and reports on the policy compliance during the lifecycle of the virtual machine. If the policy becomes noncompliant because of a host, disk, or network failure, Virtual SAN takes remedial actions. It reconfigures the data of the affected virtual machines and optimizes the use of resources across the cluster. Any reconfiguration processes occur with minimal impact on the regular workload.

For information on how Virtual SAN and virtual machine storage policies interact, see [Virtual SAN and Storage Policy-Based Management](#).

For general information about storage policies, see [Chapter 20 About Virtual Machine Storage Policies](#).

## Virtual SAN and Object-Based Storage

Virtual SAN stores and manages data in the form of flexible data containers called objects. An object is a logical volume that has its data and metadata distributed and accessed across the entire cluster.

When you provision a virtual machine on a Virtual SAN datastore, Virtual SAN creates an object for each virtual disk. It also creates a container object that stores all metadata files of the virtual machine. Virtual SAN can store and manage multiple objects in a single cluster.

Virtual SAN provisions and manages each object individually creating a RAID configuration over the network for every object. When creating an object for a virtual disk and determining how to distribute the object in the cluster, Virtual SAN considers the following parameters:

- Virtual disk requirements specified by the administrator in a form of a virtual machine storage policy.
- Cluster resources and their utilization at the time of provisioning.

For example, based on the availability policy, Virtual SAN determines how many replicas to create. While the performance policy determines the amount of SSD allocated for caching for each replica and how many stripes to create for each replica.

As the Virtual SAN cluster and the workloads evolve over time, Virtual SAN monitors the compliance of the virtual disk with its policies and makes necessary adjustments to bring the object back to compliance or to optimize the utilization of the cluster resources. Virtual SAN actively throttles the storage and network throughput used for reconfiguration to minimize the impact of reconfiguration to normal workload execution.

## Characteristics of Virtual SAN

This topic summarizes characteristics that apply to Virtual SAN, as well as its clusters and datastores.

When you work with Virtual SAN, consider the following:

- Virtual SAN is implemented directly in the ESXi hypervisor.
- Virtual SAN supports VMware features that require shared storage, such as HA, vMotion, and DRS. For example, if a host becomes overloaded, DRS can migrate virtual machines to other hosts in the cluster.
- Virtual SAN integrates with vSphere data management features traditionally used with VMFS and NFS storage. These features include snapshots, linked clones, vSphere Replication, and vSphere APIs for Data Protection.
- Virtual SAN works with Virtual Machine Storage Policies to support a virtual machine-centric storage approach.

## Limitations of Virtual SAN

- Virtual SAN does not support multiple Virtual SAN clusters for each host.

- Virtual SAN does not support virtual machines with large capacity virtual disks, or disks greater than 2TB.
- Virtual SAN supports only SATA, SAS HDD, and PCIe storage. You cannot use storage attached through USB, Fibre Channel, or iSCSI.
- Virtual SAN does not support such features as Fault Tolerance, vSphere DPM, and Storage I/O Control.
- Virtual SAN does not support SE Sparse disks.
- Virtual SAN does not support SCSI reservations.
- Virtual SAN does not support RDM, VMFS, diagnostic partition, and other disk access features.

## Characteristics of a Virtual SAN Cluster

- Virtual SAN clusters can include hosts with storage disks, or without. The minimum requirement is three hosts with storage disks.
- If a host contributes storage devices, it must have one SSD and at least one HDD device.
- Only local or direct-attached storage devices can participate in a Virtual SAN cluster. Virtual SAN does not use any other storage, such as SAN or NAS, attached to cluster.
- Virtual SAN takes up entire disks, including SSD and data disks, and does not share disks with other features.
- You can have multiple Virtual SAN clusters for each vCenter Server instance.

## Characteristics of a Virtual SAN Datastore

- Virtual SAN creates a single Virtual SAN datastore accessible to all hosts in the cluster, whether or not they have disks. All hosts can also mount any other datastores, VMFS or NFS.
- If you have multiple Virtual SAN clusters for each vCenter Server, each cluster provides a Virtual SAN datastore. You can use Storage vMotion to move virtual machines between the Virtual SAN datastores.
- Only HDD disks contribute to the capacity of the datastore. The capacity of SSDs is not counted as part of the datastore.
- In automatic mode, a Virtual SAN datastore dynamically grows when you add hosts to a Virtual SAN cluster, or disks to any cluster member.

## About Virtual SAN Datastores

After you enable Virtual SAN on a cluster, a single Virtual SAN datastore is created. It appears as another type of datastore on the list of datastores in addition to VMFS and NFS.

A Virtual SAN datastore uses multiple storage components, or disk groups, distributed across the ESXi hosts in the cluster.



A single Virtual SAN datastore provides different Quality of Service (QoS) levels for each virtual machine or each virtual disk.

In the vCenter Server, storage characteristics of the Virtual SAN datastore appear as a set of capabilities. You can reference these capabilities when creating a storage policy for virtual machines. When you later deploy virtual machines, Virtual SAN uses this policy to place virtual machines in the optimal manner based on the virtual machine requirements.

See [Chapter 20 About Virtual Machine Storage Policies](#).

## Requirements for Virtual SAN

Before you activate Virtual SAN, verify that your environment meets all requirements.

Your environment must meet specific host and storage requirements. For networking requirements, see [Virtual SAN Networking Requirements and Best Practices](#).

### Host Requirements

A minimum of three hosts contributing local disks is required. Each host must contribute one SSD and at least one HDD.

Install ESXi 5.5 or later on each host. You can add the hosts to the Virtual SAN cluster during initial setup of the cluster, or you can add hosts later.

The hosts must be managed by vCenter Server 5.5. Not all hosts need to have local storage. ESXi memory requirements: at least 6GB of RAM on each host.

### Storage Requirements

Virtual SAN supports a limited number of storage devices, which are listed on the *vSphere Compatibility Guide*. Your environment must meet the following requirements.

- One SAS or SATA or PCIe SSD. Make sure that Flash Read Cache does not use SSDs. Do not format SSDs with VMFS or any other file system.
- One SAS or SATA Hard Disk (HDD). On each host, SSDs and HDDs are combined into disk groups. The capacity ratio of an SSD to HDDs typically depends on use cases and workloads, but the best practice is to use SSD capacity of at least 10 percent of the total consumed HDD capacity in each disk group, without counting the protection copies. For more information about disk groups, see [Managing Disk Groups](#).
- An SAS or SATA HBA, or RAID controller that is set up in RAID0 mode or passthrough.

### Memory Requirements

The minimum memory requirement for one disk group is 6GB. For hosts with seven or more disks per a disk group, the memory requirement is 10GB or more. For more information about Virtual SAN memory requirement, see the *Configuration Maximums* documentation.

## USB or SD Device Requirements for Persistent Logging

The minimum requirement for installing ESXi on USB or SD devices is a 4GB or larger USB or SD device. VMware recommends using a high quality USB flash drive of 16GB or larger. The extra space will be used for an expanded coredump partition on the USB or SD device.

If you allocate local storage to Virtual SAN, you might not have enough local disk or datastore left for persistent logging. To avoid this, configure the ESXi Dump Collector and Syslog Collector to redirect ESXi memory dumps and system logs onto a network server.

Due to the I/O sensitivity of USB and SD devices the installer does not create a scratch partition on these devices. When installing on USB or SD devices, the installer attempts to allocate a scratch region on an available local disk or datastore. If no local disk or datastore is found, /scratch is placed on the ramdisk. After the installation, you should reconfigure /scratch to use a persistent datastore. For more information about reconfiguring /scratch, see the *vSphere Installation and Setup* documentation.

## Virtual SAN Networking Requirements and Best Practices

Virtual SAN requires correctly configured network interfaces.

The hosts in your Virtual SAN cluster must be part of a Virtual SAN network and must be on the same subnet. On each host, configure at least one Virtual SAN interface. You must configure this interface on all host in the cluster, no matter whether the hosts contribute storage or not.

---

**Note** Virtual SAN does not support IPv6.

---

Follow these guidelines:

- Virtual SAN requires a private 1Gb network. As a best practice, use 10Gb network.
- On each host, dedicate at minimum a single physical 1Gb Ethernet NIC to Virtual SAN. You can also provision one additional physical NIC as a failover NIC.
- You can use vSphere standard switches on each host, or you can configure your environment with a vSphere Distributed Switch.
- For each network that you use for Virtual SAN, configure a VMkernel port group with the Virtual SAN port property activated.
- Use the same Virtual SAN Network label for each port group and ensure that the labels are consistent across all hosts.
- Use Jumbo Frames for best performance.
- Virtual SAN supports IP-hash load balancing, but cannot guarantee improvement in performance for all configurations. You can benefit from IP-hash when Virtual SAN is among its many consumers. In this case, IP-hash performs the load balancing. However, if Virtual SAN is the only consumer, you might not notice changes. This specifically applies to 1G

environments. For example, if you use four 1G physical adapters with IP-hash for Virtual SAN, you might not be able to use more than 1G. This also applies to all NIC teaming policies that we currently support. For more information on NIC teaming, see the *Networking Policies* section of the *vSphere Networking Guide*.

- Virtual SAN does not support multiple VMkernel adapters on the same subnet for load balancing. Multiple VMkernel adapters on different networks, such as VLAN or separate physical fabric, are supported.
- You should connect all hosts participating in Virtual SAN to a single L2 network, which has multicast (IGMP snooping) enabled. If the hosts participating in Virtual SAN span across multiple switches or even across L3 boundaries, you must ensure that your network is configured correctly to enable multicast connectivity. You can change multicast addresses from the defaults if your network environment requires, or if you are running multiple Virtual SAN clusters on the same L2 network.

For more information, see the *vSphere Networking* documentation.

## Set Up Networking for Virtual SAN

To be able to use Virtual SAN, you must correctly configure network interfaces on the ESXi host that are included in the Virtual SAN cluster. Hosts use Virtual SAN network to communicate with each other.

On each host that participates in the Virtual SAN cluster, create a VMkernel network adapter to use as a Virtual SAN interface.

### Procedure

- 1 In the vSphere Web Client, navigate to the host.
- 2 Click the **Manage** tab and click **Networking**.
- 3 Select **VMkernel adapters** and click the **Add host networking** icon.
- 4 On the **Select connection type** page, select **VMkernel Network Adapter** and click **Next**.
- 5 Configure a target device.
- 6 On the **Port properties** page, select **Virtual SAN traffic**.
- 7 Complete the VMkernel adapter configuration.
- 8 Verify that the Virtual SAN column shows Enabled as the status for the VMkernel adapter.

### Results

Virtual SAN is enabled for networking.

### What to do next

You can now activate Virtual SAN on the host cluster.

## Using Virtual SAN and vSphere HA

You can enable vSphere HA and Virtual SAN on the same cluster. As with traditional datastores, vSphere HA provides the same level of protection for virtual machines that reside on Virtual SAN datastores. However, specific considerations exist when vSphere HA and Virtual SAN interact.

### ESXi Host Requirements

You can use Virtual SAN with a vSphere HA cluster only if the following conditions are met:

- The cluster's ESXi hosts all must be version 5.5 or later.
- The cluster must have a minimum of three ESXi hosts.

### Networking Differences

Virtual SAN has its own network. When Virtual SAN and vSphere HA are enabled for the same cluster, the HA interagent traffic flows over this storage network rather than the management network. The management network is used by vSphere HA only when Virtual SAN is disabled. vCenter Server chooses the appropriate network when vSphere HA is configured on a host.

---

**Note** Virtual SAN can only be enabled when vSphere HA is disabled.

---

The following table shows the differences in vSphere HA networking when Virtual SAN is used or not.

**Table 19-1. vSphere HA networking differences**

	<b>Virtual SAN Enabled</b>	<b>Virtual SAN Disabled</b>
Network used by vSphere HA	Virtual SAN storage network	Management network
Heartbeat datastores	Any datastore mounted to > 1 host, but not Virtual SAN datastores	Any datastore mounted to > 1 host
Host declared isolated	Isolation addresses not pingable and Virtual SAN storage network inaccessible	Isolation addresses not pingable and management network inaccessible

---

If you change the Virtual SAN network configuration, the vSphere HA agents do not automatically pick up the new network settings. So to make changes to the Virtual SAN network, you must take the following steps in the vSphere Web Client:

- 1 Disable Host Monitoring for the vSphere HA cluster.
- 2 Make the Virtual SAN network changes.
- 3 Right-click all hosts in the cluster and select **Reconfigure HA**.
- 4 Re-enable Host Monitoring for the vSphere HA cluster.

## Capacity Reservation Settings

When you reserve capacity for your vSphere HA cluster with an admission control policy, this setting must be coordinated with the corresponding Virtual SAN setting that ensures data accessibility on failures. Specifically, the Number of Failures Tolerated setting in the Virtual SAN rule set must not be lower than the capacity reserved by the vSphere HA admission control setting.

For example, if the Virtual SAN rule set allows for only two failures, the vSphere HA admission control policy must reserve capacity that is equivalent to only one or two host failures. If you are using the Percentage of Cluster Resources Reserved policy for a cluster that has eight hosts, you must not reserve more than 25% of the cluster resources. In the same cluster, with the Host Failures Cluster Tolerates policy, the setting must not be higher than two hosts. If less capacity is reserved by vSphere HA, failover activity might be unpredictable, while reserving too much capacity overly constrains the powering on of virtual machines and inter-cluster vMotion migrations.

## Virtual SAN and vSphere HA Behavior in a Multiple Host Failure Situation

After a Virtual SAN cluster fails with a loss of failover quorum for a virtual machine object, vSphere HA might not be able to restart the virtual machine even when the cluster quorum has been restored. vSphere HA guarantees the restart only when it has a cluster quorum and can access the most recent copy of the virtual machine object. The most recent copy is the last copy to be written.

Consider an example where a Virtual SAN virtual machine is provisioned to tolerate one host failure. The virtual machine runs on a Virtual SAN cluster that includes three hosts, H1, H2, and H3. All three hosts fail in a sequence with H3 being the last host to fail.

After H1 and H2 recover, the cluster has a quorum (one host failure tolerated). Despite this, vSphere HA is unable to restart the virtual machine because the last host that failed (H3) contains the most recent copy of the virtual machine object and is still inaccessible.

In this example, either all three hosts must recover at the same time, or the two-host quorum must include H3. If neither condition is met, HA attempts to restart the virtual machine when host H3 comes back on line.

## Enabling Virtual SAN

To use Virtual SAN, you must create a host cluster and enable Virtual SAN on the cluster.

A Virtual SAN cluster can include hosts with storage disks and hosts without storage disks. Follow these guidelines when you create a Virtual SAN cluster.

- A Virtual SAN cluster must include the minimum of three ESXi hosts.
- For the Virtual SAN cluster to tolerate host and disk failures, at least three hosts that join the Virtual SAN cluster must have storage disks.
- Only ESXi 5.5 hosts can join the Virtual SAN cluster.

- Before you move a host from a Virtual SAN cluster to another cluster, make sure the destination cluster is enabled with Virtual SAN.
- To be able to access the Virtual SAN datastore, an ESXi host must be a member of the Virtual SAN cluster.

After you enable Virtual SAN, the Virtual SAN storage provider is automatically registered with vCenter Server and the Virtual SAN datastore is created. For information about storage providers, see [Chapter 26 Using Storage Providers](#).

## Create a Virtual SAN Cluster

You can enable Virtual SAN when you create a cluster.

### Procedure

- 1 Right-click a datacenter in the vSphere Web Client and select **New Cluster**.
- 2 Type a name for the cluster in the **Name** text box.  
This name appears in the vSphere Web Client navigator.
- 3 Select the Virtual SAN **Turn ON** check box.
- 4 Select the mode for storage disks to be claimed and click **OK**.

Option	Description
<b>Automatic</b>	Claims all empty disks on the included hosts for Virtual SAN. Virtual SAN in automatic mode claims only local disks on the ESXi hosts in the cluster. You can add any remote non-shared disks manually.
<b>Manual</b>	Requires manual claiming of the disks on the included hosts. New disks on the host are not added to Virtual SAN. With the manual mode, two methods of organizing disks into disk groups exist, semi-automatic and manual.  <b>Note</b> When you use this mode, a Virtual SAN datastore is created, but its initial size is 0 bytes. The datastore remains unusable until you manually claim disks.

- 5 Add hosts to the Virtual SAN cluster.

Virtual SAN clusters can include hosts with disks and hosts without disks.

### Results

The cluster appears in the inventory.

### What to do next

Verify that the Virtual SAN datastore has been created. See [View Virtual SAN Datastore](#).

Verify that the Virtual SAN storage provider is registered. See [View Virtual SAN Storage Providers](#).

If you use manual mode for Virtual SAN, claim disks or create disk groups. See [Managing Disk Groups](#).

## Enable Virtual SAN on Existing Clusters

You can edit cluster properties to turn on Virtual SAN for an existing cluster.

### Prerequisites

Verify that your environment meets all requirements. See [Requirements for Virtual SAN](#).

### Procedure

- 1 Browse to the host cluster in the vSphere Web Client navigator.
- 2 Click the **Manage** tab and click **Settings**.
- 3 Under Virtual SAN, select **General** and click **Edit** to edit Virtual SAN settings.
- 4 Select the mode for storage disks to be claimed.

Option	Description
<b>Automatic</b>	Claims all empty disks on the included hosts for Virtual SAN. Virtual SAN in automatic mode claims only local disks on the ESXi hosts in the cluster. You can add any remote non-shared disks manually.
<b>Manual</b>	Requires manual claiming of the disks on the included hosts. New disks on the host are not added to Virtual SAN. With the manual mode, two methods of organizing disks into disk groups exist, semi-automatic and manual.  <b>Note</b> When you use this mode, a Virtual SAN datastore is created, but its initial size is 0 bytes. The datastore remains unusable until you manually claim disks.

- 5 Click **OK**.

### Results

Enabling Virtual SAN creates a Virtual SAN datastore and registers the Virtual SAN storage provider. The Virtual SAN storage provider communicates to vCenter Server storage capabilities of the datastore.

### What to do next

Verify that the Virtual SAN datastore has been created. See [View Virtual SAN Datastore](#).

Verify that the Virtual SAN storage provider is registered. See [View Virtual SAN Storage Providers](#).

If you use manual mode for Virtual SAN, claim disks or create disk groups. See [Managing Disk Groups](#).

If you want to disable Virtual SAN on a cluster, see [Disable Virtual SAN](#).

## Assign a License Key to a Virtual SAN Cluster

After you obtain an appropriate license key, assign it to your Virtual SAN cluster. Alternatively, you can use an existing Virtual SAN license key if one is available in the inventory of the vCenter Server system.

When you enable Virtual SAN for a cluster, you can use Virtual SAN in evaluation mode to explore its features. The evaluation period starts when Virtual SAN is enabled and expires after 60 days. To use Virtual SAN, you must license the cluster before the evaluation period expires. For more information, see the License Management and Reporting section of the *vCenter Server and Host Management Guide*.

### Prerequisites

- Required privilege: **Global.Licenses**
- If you plan to assign an existing license, verify that the license key is available in the vCenter Server system.

### Procedure

- 1 In the vSphere Web Client, navigate to a cluster with enabled Virtual SAN.
- 2 On the **Manage** tab, click **Settings**.
- 3 Under Configuration, select **Virtual SAN Licensing** and click **Assign License Key**.
- 4 Select a licensing option.

Option	Description
<b>Assign an existing license key</b>	Select the license key to assign and click <b>OK</b>
<b>Assign a new license key</b>	<ol style="list-style-type: none"> <li>a Type or copy and paste the license key, and type an optional label for the key.</li> <li>b To verify that the license key is in the correct format, and that it is for Virtual SAN, click <b>Decode</b>.</li> <li>c Click <b>OK</b>.</li> </ol>

## View Virtual SAN Datastore

After you enable Virtual SAN, a single datastore is created. You can review the capacity of the Virtual SAN datastore.

### Prerequisites

Activate Virtual SAN and configure disk groups.

### Procedure

- 1 Browse to Datastores in the vSphere Web Client navigator.
- 2 Select the Virtual SAN datastore.
- 3 Click the **Manage** tab, and click **Settings**.



#### 4 Review the Virtual SAN datastore capacity.

The size of the Virtual SAN datastore depends on the number of HDDs per ESXi host and the number of ESXi hosts in the cluster. For example, if a host has seven 2TB HDDs, and the cluster includes eight hosts, the approximate storage capacity would be  $7 \times 2 \text{ TB} \times 8 = 112\text{TB}$ . Some capacity is allocated for metadata.

#### What to do next

Use the storage capabilities of the Virtual SAN datastore to create a storage policy for virtual machines. For information, see [Chapter 20 About Virtual Machine Storage Policies](#).

## Disable Virtual SAN

You can turn off Virtual SAN for a host cluster.

When you disable the Virtual SAN cluster, all virtual machines located on the shared Virtual SAN datastore become inaccessible. If you intend to use virtual machine while Virtual SAN is disabled, make sure you migrate virtual machines from Virtual SAN datastore to another datastore before disabling the Virtual SAN cluster.

#### Prerequisites

Verify that the hosts are in maintenance mode.

#### Procedure

- 1 Browse to the host cluster in the vSphere Web Client navigator.
- 2 Click the **Manage** tab and click **Settings**.
- 3 Under Virtual SAN, select **General** and click **Edit** to edit Virtual SAN settings.
- 4 Deselect the **Turn On Virtual SAN** check box.

## Managing Disk Groups

On each ESXi host that contributes its local disks to a Virtual SAN cluster, disks are organized into disk groups. A disk group is a main unit of storage on a host. Each disk group includes one SSD and one or multiple HDDs (magnetic disks).

Virtual SAN uses an aggregation of disk groups to back up a single datastore that is created when you enable Virtual SAN.

In the disk group, the SSD primarily serves as a read cache write buffer, while the HDDs are used for permanent storage. Typically, a higher SSD to HDD ratio, both in size and quantity, improves performance.

Depending on the mode you select when enabling Virtual SAN on a cluster, you can use different ways to organize disks into groups.

#### Automatic Mode

Virtual SAN claims all available and usable disks and organizes them into default groups with one SSD and one or multiple HDDs. If you add more disks to hosts or add new hosts to the Virtual SAN cluster, all applicable disks are claimed by Virtual SAN. Virtual SAN in automatic mode claims only local disks on the ESXi hosts in the cluster. You can add any remote nonshared disks manually.

## Manual Mode

You must specify hosts and disks on the hosts to be used for the Virtual SAN datastore. You have two methods of organizing disks into disk groups, semi-automatic and manual.

When you use the semi-automatic method, Virtual SAN organizes the disks that you specify into default disk groups.

Another option is to manually create user-defined disk groups and select disks for each group. When you create a disk group manually, your main consideration should be the ratio of SSD to Raw HDD capacity. Although the ratios depend on use cases and workloads, the best practice is to use SSD capacity of at least 10 percent of the total consumed HDD capacity in each disk group, without counting the protection copies. For example, if the size of your Raw HDD capacity on the disk group is 4TB, the recommended SSD capacity is 400GB.

Virtual SAN disks cannot be used by other features, such as RDM, VMFS, diagnostic partition, and so on.

## Use Semi-Automatic Method to Claim Disks for Virtual SAN

When Virtual SAN is enabled in the manual mode for the host cluster, you must select SSD and HDD disks to be claimed by Virtual SAN. After you select the disks, Virtual SAN organizes the disks into default disk groups. This is a semi-automatic method of creating disk groups.

### Procedure

- 1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
- 2 Click the **Manage** tab and click **Settings**.
- 3 Under Virtual SAN, click **Disk Management**.
- 4 Click the **Claim Disks** icon.
- 5 Select disks to be used for Virtual SAN.

On each host, you must select at least one SSD and one or multiple HDDs. If you select several SSDs, the number of HDDs must be equal to the number of SSDs, or greater.

### Results

Virtual SAN claims the disks that you select and organizes them into default disk groups that back the Virtual SAN datastore.

## Use Manual Method to Claim Disks for Virtual SAN

When Virtual SAN is enabled in the manual mode for the host cluster, you can use a manual method of combining specific SSDs with specific HDDs and creating user-defined disk groups on a particular host.

This method contrasts with the semi-automatic claiming of disks for Virtual SAN, where you can select multiple disks to be consumed by Virtual SAN, but do not organize them into groups. Instead, after you select the disks, Virtual SAN creates default disk groups for you. In contrast, when you manually create a disk group, you can be more specific and organize selected disks into disk groups. You can select an SSD and HDD disks of your choice to form a disk group. The disk groups that you create belong to the particular host where the disks reside, and you cannot reassign the groups to another host.

When you create a disk group, your main consideration should be the ratio of SSD to HDD capacity. Although the ratios depend on use cases and workloads, the best practice is to use 1:10 ratio of the SSD to HDD capacity.

### Procedure

- 1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
- 2 Click the **Manage** tab and click **Settings**.
- 3 Under Virtual SAN, click **Disk Management**.
- 4 Click the **Create a new disk group** icon.
- 5 Select an SSD disk, and then select one or multiple non-SSD disks.  
For each disk group, you can select only one SSD and up to six non-SSD disks.
- 6 Click **OK**.

### Results

The new disk group appears on the list.

## Add a Disk to the Disk Group

When you use Virtual SAN in manual mode, you can add local disks to existing disk groups.

### Procedure

- 1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
- 2 Click the **Manage** tab and click **Settings**.
- 3 Under Virtual SAN, click **Disk Management**.
- 4 Select the disk group, and click the **Add a disk to the selected disk group** icon.
- 5 Select a disk to add and click **OK**.

## Results

The disk appears on the list of disks that belong to the disk group.

## Remove Disks from Virtual SAN

You can remove an entire disk group or selected disks from the disk group.

You can remove disks or disk groups only when Virtual SAN is in manual mode. For the automatic disk claim mode, the remove action is not supported.

---

**Note** Removing a single SSD disk or all non-SSDs from the disk group, removes the entire disk group.

---

### Prerequisites

Because removing disks might be disruptive for the Virtual SAN datastore and virtual machines on the datastore, make sure that no virtual machines use the disks or the disk group that you remove. To achieve this, place the Virtual SAN host in maintenance mode.

### Procedure

- 1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
- 2 Click the **Manage** tab and click **Settings**.
- 3 Under Virtual SAN, click **Disk Management**.
- 4 Remove a disk group or selected disks.

Option	Description
<b>Remove the Disk Group</b>	<p>Under Disk Groups, select the disk group to remove, and click the <b>Remove the Disk Group</b> icon.</p> <hr/> <p><b>Note</b> The <b>Remove the Disk Group</b> icon does not appear when the Virtual SAN cluster is set up in automatic mode. The icon appears only when the cluster is in manual mode.</p> <hr/>
<b>Remove the Selected Disk</b>	<ol style="list-style-type: none"> <li>a Under Disk Groups, select the disk group that contains the disk to remove.</li> <li>b Under Disks, select the disk to remove, and click the <b>Remove the Selected Disks</b> icon.</li> </ol> <hr/>

## Results

Any SSDs that you remove from Virtual SAN become available for such features as Flash Read Cache.

## Monitoring Virtual SAN

Virtual SAN supports extensive monitoring facilities. You can monitor your Virtual SAN environment from a vSphere Web Client.

You can monitor different objects in your Virtual SAN setup. You can monitor hosts that participate in a Virtual SAN cluster just like any other host that is managed by vCenter Server. In addition, you can monitor the Virtual SAN datastore. For more information on monitoring objects and storage resources in your Virtual SAN environment, see *vSphere Monitoring and Performance* documentation.

## Monitor the Virtual SAN Cluster

You can monitor the Virtual SAN cluster and all the objects that are related to the cluster.

### Procedure

- 1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
- 2 Select the **Monitor** tab and click **Virtual SAN**.
- 3 Select **Physical Disks** to review all hosts and storage disks that belong to the cluster.  
Information about storage disks, such as capacity, used capacity, and so on appears.
- 4 Select a non-SSD disk and click **Virtual Disks** to review virtual machines that use the disk.  
You can review all virtual machine objects and whether they are in compliance with the storage policies assigned to them.

## Monitor Disks that Participate in Virtual SAN Datastores

Verify the status of the disks that back up the Virtual SAN datastore. You can check whether the disks experience any problems.

### Procedure

- 1 Browse to Datastores in the vSphere Web Client navigator.
- 2 Select the Virtual SAN datastore.
- 3 Click the **Manage** tab, and click **Settings**.
- 4 Click **Device Backing**.
- 5 Select the disk group and review the local disks listed in the **Disks** table.
- 6 If the Issue column for the disks is not visible, do the following:
  - a Right-click the column heading and select **Show/Hide Columns**.
  - b Select **Issue** to display the column.

The Issue column alerts you if the disks become ineligible. If the disks experience no problems, the column does not display any values.

## Managing Virtual SAN

After you have set up your Virtual SAN cluster, you can use the vSphere Web Client to add hosts and disks, remove hosts and disks, and manage failure scenarios.

## Expanding a Virtual SAN Cluster

You can expand an existing Virtual SAN cluster by adding hosts or disks to the hosts without disrupting any ongoing operations.

Use one of the following methods to expand your Virtual SAN cluster.

- Add new ESXi hosts directly to the cluster. See [Add a Host to the Virtual SAN Cluster](#).
- Move existing hosts to the cluster by using Host Profile. See [Add a Host to the Virtual SAN Cluster using Host Profile](#).

New cluster members add storage and compute capacity. If you use Virtual SAN in automatic mode, local storage disks on the new cluster member will be automatically aggregated into a disk group and claimed by the Virtual SAN datastore. If Virtual SAN is set to manual, you must manually create a subset of disk groups from the local storage disks on the newly added host. See [Use Manual Method to Claim Disks for Virtual SAN](#).

- Add new storage disks to ESXi hosts that are cluster members. If you use Virtual SAN in automatic mode, the new disk that you add joins an existing disk group. If Virtual SAN is set to manual, you must manually add the disk to the disk group on the host. See [Add a Disk to the Disk Group](#).

### Add a Host to the Virtual SAN Cluster

You can add an ESXi host to a running Virtual SAN cluster without disrupting any ongoing operations. The host's resources become associated with the cluster.

#### Procedure

- 1 Browse to the Virtual SAN cluster in the vSphere Web Client navigator.
- 2 Right-click the cluster and select **Add Host**.
- 3 Enter the host name, user name, and password, and click **Next**.
- 4 View the summary information and click **Next**.
- 5 Assign an existing or a new license key and click **Next**.
- 6 (Optional) Enable lockdown mode to prevent remote users from logging directly into the host.

You can configure this option later by editing the Security Profile in host settings.

- 7 Select what to do with the host's virtual machines and resource pools.

- **Put this host's virtual machines in the cluster's root resource pool**

vCenter Server removes all existing resource pools of the host. The virtual machines in the host's hierarchy are all attached to the root. Because share allocations are relative to a resource pool, you might have to manually change a virtual machine's shares, which destroys the resource pool hierarchy.

- **Create a resource pool for this host's virtual machines and resource pools**

vCenter Server creates a top-level resource pool that becomes a direct child of the cluster and adds all children of the host to that new resource pool. You can type a name for that new top-level resource pool. The default is **Grafted from <host\_name>**.

**8** Review the settings and click **Finish**.

## Results

The host is added to the cluster.

## Add a Host to the Virtual SAN Cluster using Host Profile


You can add an ESXi host to the Virtual SAN cluster by creating a new Host Profile. A new profile can be created by extracting and reusing the designated reference host's configuration.

## Prerequisites

Verify that the host is in maintenance mode.


## Procedure

**1** Create a new Host Profile.

- a Navigate to the Host profiles view.
  - b Click the **Extract Profile from a Host** icon (  ).
  - c Select the host that you intend to use as the reference host and click **Next**.
- The selected host must be an active host.
- d Type a name and description of the new profile and click **Next**.
  - e Review the summary information of the new host profile and click **Finish**.

The new profile appears in the profile list.

**2** Attach the host to the intended host profile.

- a From the Profile List in the Host Profiles main view, select the host profile to be applied to the Virtual SAN host.
  - b Click the **Attach/Detach Hosts and clusters to a host profile** icon (  ).
  - c Select the host from the expanded list and click **Attach** to attach the host to the profile.
- The host is added to the Attached Entities list.
- d Click **Next**.
  - e Click **Finish** to complete attaching the host to the profile.

### 3 Detach the referenced Virtual SAN host from the Host Profile.

When a Host Profile is attached to a cluster, the host or hosts within that cluster are also attached to the Host Profile. However, when the Host Profile is detached from the cluster, the association between the host or hosts within the cluster and that of the Host Profile remains intact.

- a From the Profile List in the Host Profiles main view, select the Host Profile to be detached from a host or cluster.

- b Click the **Attach/Detach Hosts and clusters to a host profile** icon (  ).

- c Select the host or cluster from the expanded list and click **Detach**.

The selected host or cluster is added to the Attached Entities list.

- d Click **Detach All** to detach all the listed hosts and clusters from the profile.

- e Click **Next**.

- f Click **Finish** to complete attaching the host to the profile.

### 4 Verify that the compliance of the Virtual SAN host to its attached Host Profile and determine if any, configuration parameters on the host are different from those specified in the Host Profile.

- a Navigate to a Host Profile.

The **Objects** tab lists all Host Profiles, the number of hosts attached to that Host Profile, and summarized results of the last compliance check.

- b Click the **Check Host Profile Compliance** icon (  ).

In order to see specific detail on which parameters differ between the host that failed compliance and the Host Profile, click on the **Monitor** tab and select the Compliance view. Then, expand the object hierarchy and select the failing host. The differing parameters are displayed in the Compliance window, below the hierarchy.

In the event of a compliance failure, use the Remediate action to apply the Host Profile settings onto the host. The action changes all Host Profile managed parameters to the values contained in the Host Profile attached to the host.

### 5 Remediate the host to fix any compliance error on the host.

- a Select the **Monitor** tab, then click **Compliance**.

- b Right-click the host or hosts that you want to remediate and select **All vCenter Actions > Host Profiles > Remediate**.

You can update or change the user input parameters for the Host Profiles policies by customizing the host.

- c Click **Next**.

- d Review the tasks that are necessary to remediate the Host Profile and click **Finish**.



The host is now part of the Virtual SAN cluster and its resources are accessible to the Virtual SAN cluster. The host can also access all existing Virtual SAN Storage I/O Policies in the VSAN cluster.

## Place a Member of Virtual SAN Cluster in Maintenance Mode

Before you shut down, reboot, or disconnect a host that is a member of a Virtual SAN cluster, you must put it in maintenance mode. When you place a host in maintenance mode, you can select a specific evacuation mechanism. Note that when any member node of a Virtual SAN cluster enters maintenance mode, the cluster capacity is automatically reduced as the member node no longer contributes storage to the cluster.

### Prerequisites

Check that your environment has the capabilities required for the option you select.

### Procedure

- 1 Right-click the host and select **Enter Maintenance Mode**.
- 2 Select one of the evacuation modes and click **OK**.

Option	Description
<b>Ensure accessibility</b>	<p>This is the default option. When you power off or remove the host from the cluster, Virtual SAN ensures that all accessible virtual machines on this host remain accessible. Typically, only partial data evacuation is required.</p> <p>Select this option if you take the host out of the cluster temporarily, for example, to install upgrades, and plan to have the host back in the cluster. This option is not appropriate when you permanently remove the host from the cluster.</p>
<b>Full data migration</b>	<p>Virtual SAN evacuates all data to other hosts in the cluster and maintains or fixes availability compliance for the affected components in the cluster. This option results in the largest amount of data transfer and consumes the most time and resources. Select this option if you plan to migrate the host permanently. When evacuating data from the last host in the cluster, make sure you migrate the virtual machines to another datastore and then place the host in maintenance mode.</p> <p><b>Note</b> The host will not enter maintenance mode if a virtual machine object that has data on the host is not accessible and cannot be fully evacuated.</p>
<b>No data Migration</b>	<p>Virtual SAN does not evacuate any data from this host. If you power off or remove the host from the cluster, some virtual machines might become inaccessible.</p>

## Virtual SAN and Storage Policy-Based Management

When you use Virtual SAN, you can define virtual machine storage requirements, such as performance and availability, in the form of a policy. The policy requirements are then pushed

down to the Virtual SAN layer when a virtual machine is being created. The virtual disk is distributed across the Virtual SAN datastore to meet the requirements.

When you enable Virtual SAN on a host cluster, a single Virtual SAN datastore is created. In addition, enabling Virtual SAN configures and registers Virtual SAN storage providers. Virtual SAN storage providers are built-in software components that communicate datastore capabilities to vCenter Server.

A storage capability is typically represented by a key-value pair, where the key is a specific property that the datastore can offer and the value is a metric, or a range, that the datastore guarantees for a provisioned object, such as a virtual machine metadata object or a virtual disk. You can also use tags to create user-defined storage capabilities and reference them when defining a storage policy for a virtual machine. For information on how to use and apply tags to datastores, see [Understanding Storage Capabilities](#).

When you know storage requirements of your virtual machines, you can create a storage policy referencing capabilities that the datastore advertises. You can create several policies to capture different types or classes of requirements.

You can apply these storage policies when you create or edit virtual machines.

---

**Note** If you do not apply a storage policy to a virtual machine, it will use a default Virtual SAN policy with one number of failures to tolerate and a single disk stripe per object.

---

**Table 19-2. Storage Policy Attributes**

Capability	Description
Number of failures to tolerate	<p>Defines the number of host, disk, or network failures a virtual machine object can tolerate. For <math>n</math> failures tolerated, <math>n+1</math> copies of the virtual machine object are created and <math>2n+1</math> hosts with storage are required. Default value is 1. Maximum value is 3.</p> <hr/> <p><b>Note</b> If the host is a member of a Virtual SAN cluster, and if any virtual machine object on the host uses the <b>Number of failures to tolerate=0</b> setting in its storage policy, the host might experience unusual delays when entering maintenance mode. The delay occurs because Virtual SAN has to evacuate this object from the host for the maintenance operation to complete successfully.</p> <hr/>
Number of disk stripes per object	<p>The number of HDDs across which each replica of a virtual machine object is striped. A value higher than 1 might result in better performance, but also results in higher use of system resources. Default value is 1. Maximum value is 12.</p> <hr/>
Object space reservation	<p>Percentage of the logical size of the object that should be reserved, or thick provisioned, during virtual machine creation. The rest of the storage object is thin provisioned. This value applies only when you chose thin provisioning for a virtual disk. If your disk is thick provisioned, 100% is reserved automatically. Default value is 0%. Maximum value is 100%.</p> <hr/>

Table 19-2. Storage Policy Attributes (continued)

Capability	Description
Flash read cache reservation	<p>SSD capacity reserved as read cache for the virtual machine object. Specified as a percentage of the logical size of the object. Should be used only to address read performance issues. Reserved flash capacity cannot be used by other objects. Unreserved flash is shared fairly among all objects.</p> <p>Default value is 0%. Maximum value is 100%.</p> <hr/> <p><b>Note</b> By default, Virtual SAN dynamically allocates read cache to storage objects based on demand. This represents the most flexible and the most optimal use of resources. As a result, typically, you do not need to change the default 0 value for this parameter.</p> <p>To increase the value when solving a performance problem, exercise caution. Over-provisioned cache reservation across several virtual machines can cause SSD space to be wasted on over-reservations and not being available to service the workloads that need the space at a given time. This might lead to performance degradation.</p> <hr/>
Force provisioning	<p>If the option is set to <b>Yes</b>, the object will be provisioned even if the policy specified in the storage policy is not satisfiable by the datastore. Use this parameter in bootstrapping scenarios and during an outage when standard provisioning is not longer possible.</p> <p>The default <b>No</b> is acceptable for most production environments. Virtual SAN attempts to provision a virtual machine even when all requirements cannot be met.</p> <hr/>

## View Virtual SAN Storage Providers

Virtual SAN automatically configures and registers a storage provider for each host in the Virtual SAN cluster.

The Virtual SAN storage providers report a set of underlying storage capabilities to vCenter Server. They also communicate with the Virtual SAN layer to report the storage requirements of the virtual machines. For more information about storage providers, see [Chapter 26 Using Storage Providers](#).

Virtual SAN registers a separate storage provider for each host in the Virtual SAN cluster, using the following URL:

```
http://host_ip:8080/version.xml
```

where *host\_ip* is the actual IP of the host.

Verify that the storage providers are registered.

### Procedure

- 1 Browse to the vCenter Server in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage Providers**.

## Results

The storage providers for Virtual SAN appear on the list. All hosts have a storage provider, but only one is active. Storage provider that belong to other hosts are in standby. If the host that currently has the online storage provider fails, another host will bring its provider online.

## What to do next

If you delete or unregister the Virtual SAN storage provider for troubleshooting purposes, you can register it manually. See [Register Storage Providers](#).

## Define a Virtual Machine Storage Policy for Virtual SAN

When you need to define storage requirements and a type of storage service for a virtual machine and its virtual disks, you create a storage policy. In this policy, you reference storage capabilities that the Virtual SAN datastore reports.

### Prerequisites

- Verify that the Virtual SAN storage provider is available. See [View Virtual SAN Storage Providers](#).
- Ensure that Storage Policies is enabled. See [Enable Virtual Machine Storage Policies](#).

### Procedure

- 1 From the vSphere Web Client Home, click **Rules and Profiles > VM Storage Policies**.
- 2 Click the Create a New VM Storage Policy icon.
- 3 Select the vCenter Server instance.
- 4 Type a name and a description for the storage policy.
- 5 On the Rule-Set 1 window, define the first rule set.
  - a Select **vsan** from the Rules Based on Vendor Specific Capabilities drop-box.  
The page expands to show capabilities reported by the Virtual SAN datastore.
  - b Add a capability and specify its value.  
Make sure that the values you provide are within the range of values advertised by the capability profile of the Virtual SAN datastore.
  - c (Optional) Add tag-based capabilities. See [Add or Edit Tag-Based Rules](#).
- 6 (Optional) Add another rule set.
- 7 Review the list of datastores that match this policy and click **Finish**.  
To be eligible, a datastore does not need to satisfy all rule sets within the policy. The datastore must satisfy at least one rule set and all rules within this set. Make sure that the Virtual SAN datastore meets the requirements set in the storage policy and appears on the list of compatible datastores.

## Results

The new policy is added to the list.

## What to do next

Apply this policy to a virtual machine and its virtual disks. Virtual SAN will place the virtual machine objects in accordance with the requirements specified in the policy. For information about using and editing storage policies, see [Chapter 20 About Virtual Machine Storage Policies](#).

# About Virtual Machine Storage Policies

# 20

Use virtual machine storage policies, formerly called virtual machine storage profiles, to ensure that virtual machines are placed to storage that guarantees a specific level of capacity, performance, availability, redundancy, and so on.

When you define a storage policy, you specify storage requirements for applications that would run on virtual machines. After you apply this storage policy to a virtual machine, the virtual machine is placed to a specific datastore that can satisfy the storage requirements.

You manage storage placement by using storage policies.

- 1 If you use storage policies in conjunction with storage providers, verify that an appropriate storage provider is registered.  
See [Chapter 26 Using Storage Providers](#).
- 2 Apply storage capability tags to datastores. See [Assign Tags to Datastores](#).
- 3 Enable storage policies for a host or cluster. See [Enable Virtual Machine Storage Policies](#).
- 4 Create storage policies by defining requirements for applications running on a virtual machine. See [Define a Storage Policy for a Virtual Machine](#).
- 5 Associate a storage policy with the virtual machine files or virtual disks. See [Apply a Storage Policy to a Virtual Machine](#).
- 6 Verify that virtual machines and virtual disks use datastores that are compliant with the assigned storage policy. See [Check Storage Compliance for a Virtual Machine](#).

This chapter includes the following topics:

- [Understanding Storage Capabilities](#)
- [Understanding Virtual Machine Storage Policies](#)

## Understanding Storage Capabilities

When you define storage requirements for virtual machines and virtual disks, you reference vendor specific or user-defined storage capabilities in a storage policy.

## Vendor Specific Storage Capabilities

Storage systems that use vStorage APIs for Storage Awareness, also called VASA, are represented by storage providers. Storage providers inform vCenter Server about specific storage devices, and present characteristics of the devices and datastores deployed on the devices as storage capabilities. Such storage capabilities are system-defined and vendor specific.

A storage system can advertise multiple capabilities. The capabilities are grouped into one or more capability profile. Capabilities outline the quality of service that the storage system can deliver. They guarantee that the storage system can provide a specific set of characteristics for capacity, performance, availability, redundancy, and so on.

Vendor specific capabilities appear in the Storage Policy-Based Management system. When you create a storage policy for your virtual machine, you reference these vendor specific storage capabilities, so that your virtual machine is placed on the datastore with these capabilities.

For more information about storage providers, see [Chapter 26 Using Storage Providers](#).

## Tag-Based Storage Capabilities

You can use tags to create user-defined storage capabilities and apply those tags to datastores. You can apply more than one tag to a datastore.

Typically, tag-based capabilities serve the following purposes:

- Encode policy-relevant information that is not advertised through vStorage APIs for Storage Awareness, such as geographical location or administrative group.
- Attach a broad storage level definition to a datastore that is not represented by any storage provider.

All tags applied to datastores appear in the Storage Policy-Based Management system. As with vendor specific capabilities, you can use the tags when you create storage policies.

## Assign Tags to Datastores

When your datastore is not represented by a storage provider and does not display its capabilities in the Storage Profile-Based Management system, use tags to encode information about the datastore. You can then reference these tags when defining a storage policy for a virtual machine.

You can apply a new tag that contains storage information to a datastore. For information about tags, their categories, and how to manage tags, see the *vCenter Server and Host Management* documentation.

### Procedure

- 1 Browse to a datastore in the vSphere Web Client inventory.
- 2 Click the **Manage** tab and click **Tags**.
- 3 Click the **New Tag** icon.

- 4 In the **vCenter Server** drop-down menu, select the vCenter Server instance on which to create the tag.
- 5 Type a name and a description for the tag.

For example, you can specify a broad storage level definition, such as Gold Storage, or indicate a property that is not communicated through a storage provider, such as geographical location or administrative group.

Tag Property	Example
Name	Fault Tolerance
Description	Storage that has a capacity over 2TB and is fault-tolerant

- 6 In the **Category** drop-down menu, select an existing category or create a new category.
- 7 (Optional) To create a new category, follow these steps:
  - a Select **New Category** to expand the dialog box.
  - b Specify the category options.

Category Property	Example
Category Name	Storage Category
Description	Category for tags related to storage
Cardinality	<b>Many tags per object</b>
Associable Object Types	<b>Datastore</b> and <b>Datastore Cluster</b>

- 8 Click **OK**.

## Results

The new tag is assigned to the datastore.

## What to do next

You can reference the tag when adding tag-based rules to the storage policy. See [Add or Edit Tag-Based Rules](#).

The datastore appears on a list of compatible storage resources for virtual machines that use the policy.

# Understanding Virtual Machine Storage Policies

Storage policies that you define for virtual machines, capture storage characteristics that virtual machine home files and virtual disks require to run applications within the virtual machine.

When you create a storage policy, you can reference storage capabilities advertised by a storage system. You can also reference user-defined datastore tags.



Although storage policies and storage capabilities have similar semantics, policies describe what users require for their virtual machines, while storage capabilities refer to what the system can offer.

You can create several storage policies to define different types and classes of storage requirements.

Each storage policy is not only a set of constraints that apply simultaneously. A single policy can include alternative sets of subpolicies, or rule-sets, that represent equally acceptable storage requirements.

The virtual machine home files (.vmx, .vmsd, .nvram, .log, and so on) and the virtual disks (.vmdk) can have separate storage policies as shown in the following table.

**Table 20-1. Example Storage Policy for a Virtual Machine**

Example Virtual Machine Files	Example for a Storage Policy	Example for a Datastore Compliant with the Storage Policy
<i>windows_2008r2_test.vmx</i>	Storage Policy 2	datastore02, datastore05, datastore10
<i>windows_2008r2_test.vmx</i>		
<i>windows_2008r2_test.log</i>		
<i>windows_2008r2_test.nvram</i>		
<i>windows_2008r2_test.vmem</i>		
<i>windows_2008r2_test.vmsd</i>		
<i>windows_2008r2_test.vmdk</i>	Storage Policy 3	datastore05
<i>windows_2008r2_test_1.vmdk</i>	Storage Policy 5	datastore10

When you create, clone, or migrate a virtual machine, you can apply the storage policy to the virtual machine. You can then place the virtual machine to one of the datastores that has capabilities compatible with the policy requirements.

If you select a datastore that does not match the policy, the Policy shows that the virtual machine is using noncompliant storage.

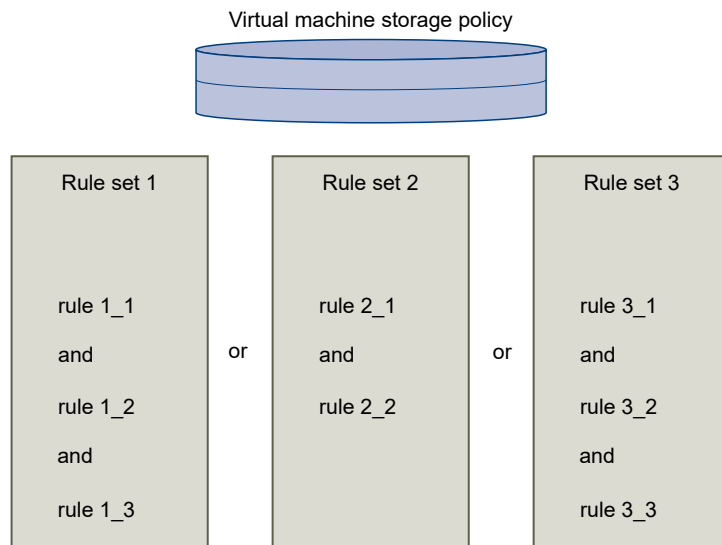
## About Rule Sets

A storage policy includes one or several rule sets that describe requirements for virtual machine storage resources.

When you define a storage policy for your virtual machines, you can add one or multiple rule sets, also called subprofiles. A single rule set contains one or several rules. Each rule describes a specific quality or quantity that needs to be provided by a storage resource. Each rule can either be an underlying storage capability or a user-defined storage tag. One rule set can include rules from only a single storage provider.

To create a policy, one rule set is required, while additional rule sets are optional. Multiple rule sets allow a single policy to define alternative selection parameters, often from several storage providers.

The relationship between all rule sets within a policy is defined by the Boolean operator OR, while the relationship between all rules within a single rule set is defined by AND. Meeting all the rules of any one rule set is sufficient to satisfy the entire policy. Each rule set represents an equally acceptable set of constraints.



## Enable Virtual Machine Storage Policies

To use storage policies for virtual machines, you must enable them for a host or a cluster.

### Prerequisites

Verify that a single host or all the hosts in a cluster, where you want to enable the policies, include the Storage Policies license.

### Procedure

- 1 From the vSphere Web Client Home, click **Rules and Profiles > VM Storage Policies**.
- 2 Click the **Enable VM Storage Policies** icon (📁✅).
- 3 Select the vCenter Server instance.  
All clusters and hosts available for the selected server appear. You can see the server licensing status and whether storage policies are enabled.
- 4 To enable storage policies, select a host or a cluster whose status is Disabled or Unknown and click **Enable**.

### Results

The status of the host or cluster changes to Enabled.

### What to do next

You can apply storage policies to virtual machines that run on the enabled host or cluster.

## Define a Storage Policy for a Virtual Machine

When you define storage policies for virtual machines, you specify storage requirements for virtual machine objects, such as home files and virtual disks.

Each storage policy can reference storage capabilities advertised by a storage system, or user-defined datastore tags.

### Prerequisites

If you plan to include system-defined storage capabilities in the storage policy, make sure to register an appropriate storage provider. See [Chapter 26 Using Storage Providers](#).

### Procedure

- 1 From the vSphere Web Client Home, click **Rules and Profiles > VM Storage Policies**.
- 2 Click the **Create a New VM Storage Policy** icon.
- 3 Select the vCenter Server instance.
- 4 Type a name and a description for the storage policy.
- 5 On the Rule-Set 1 screen, define the first rule set.
  - a Select a storage provider from the Rules Based on Vendor Specific Capabilities drop-box.  
The page expands to show system-defined capabilities.
  - b Add a capability and specify its value.  
Make sure that the values you provide are within the range of values advertised by the capability profile of a storage system.
  - c (Optional) Add user-defined capabilities. See [Add or Edit Tag-Based Rules](#).
- 6 (Optional) Add another rule set.
- 7 Review the list of datastores that match this policy.  
To be eligible, a datastore does not need to satisfy all rule sets within the policy. The datastore must satisfy at least one rule set and all rules within this set.
- 8 Confirm that the policy information is correct and click **Finish**.

### Results

The new storage policy appears on the list of available policies.

### What to do next

Apply the policy to a virtual machine home file and its virtual disks.

## Add or Edit Tag-Based Rules

When you define or edit a storage policy for virtual machines, you can create or modify a rule that references tags that you used for particular datastores. The datastores become compatible with that type of storage policy.

When you use tags in the policies, follow these guidelines:

- If the rule set contains other rules, the datastore with the assigned tag must satisfy all of the rules in the rule set.
- If you add several tags from the same category within the same rule, the tags are treated as alternative constraints. Either of the tags can be satisfied.
- If you add the tags in separate rules within the same rule set, all tags must be satisfied.

### Prerequisites

Create storage tags and apply them to datastores. See [Assign Tags to Datastores](#).

### Procedure

- 1 On the Rule Set page, perform one of the following:
  - To add a rule, click the **Add tag-based rule** button.
  - To modify an existing rule, select the rule and click the **Modify rule** icon.
- 2 Specify a category.
- 3 Make a tag selection or modify an existing selection.

### Results

Datastores that use the tags that you selected become compatible with the storage policy.

## Edit a Storage Policy

You can change the storage requirements for virtual machines and virtual disks by modifying the storage policy.

### Procedure

- 1 From the vSphere Web Client Home, click **Rules and Profiles > VM Storage Policies**.
- 2 Select a storage policy, and click the **Edit a VM Storage Policy** icon (✎).
- 3 Modify the policy and click **OK**.

- 4 If the storage policy is used by a virtual machine, reapply the policy to the virtual machine.

Option	Description
<b>Manually later</b>	If you select this option, the compliance status for all virtual disks and virtual machine home objects associated with the storage policy changes to Out of Date. To update configuration and compliance, manually reapply the storage policy to all associated entities. See <a href="#">Reapply Virtual Machine Storage Policy</a> .
<b>Now</b>	Update virtual machine and compliance status immediately after editing the storage policy.

## Results

The storage policy is changed.

## Delete a Virtual Machine Storage Policy

You can delete a storage policy if you are not using it for any virtual machine or virtual disk.

### Procedure

- 1 From the vSphere Web Client Home, click **Rules and Profiles > VM Storage Policies**.
- 2 In the VM Storage Policies view, select a policy that you want to delete and click the **Delete a VM Storage Policy** icon (✖).
- 3 Click **Yes**.

## Results

The policy is removed from the inventory.

## Apply a Storage Policy to a Virtual Machine

You can apply a storage policy to a virtual machine to specify storage requirements for applications that run on the virtual machine.

You can apply a policy to a powered-off or powered-on virtual machine.

### Procedure

- 1 In the vSphere Web Client, browse to the virtual machine.
- 2 Click the **Manage** tab, and click **VM Storage Policies**.
- 3 Click **Manage VM Storage Policies**.
- 4 Apply the storage policy to the virtual machine configuration files by selecting the policy from the **Home VM Storage Policy** drop-down menu.

---

**Note** The virtual machine configuration files include `.vmx`, `.vmsd`, `.nvram`, and so on.

---

- 5 Apply the storage policy to virtual disks.
  - Click **Apply to disks** to apply the same home policy to all virtual disks.

- Under VM Storage Policies for Virtual Disks, select a virtual disk and apply a policy from the **VM Storage Policy** drop-down menu. This option associates each virtual disk with a specified policy. Use this option if you store your virtual disks separately from the virtual machine configuration files.

6 Click **OK**.

## Results

The storage policy is assigned to the virtual machine and its disks.

## Monitor Storage Compliance for Virtual Machines

When you associate a policy with virtual machine objects, and select the datastores on which virtual machines and virtual disks run, you can check whether virtual machines and virtual disks use datastores that are compliant with the policy.

If you check the compliance of a virtual machine whose host or cluster has storage policies disabled, the result of the check is Not Compliant because the feature is disabled.

## Prerequisites

To perform a compliance check for a storage policy, associate the policy with at least one virtual machine or virtual disk.

## Procedure

- 1 From the vSphere Web Client Home, click **Rules and Profiles > VM Storage Policies**.
- 2 Double-click a storage policy.
- 3 Click the **Monitor** tab, and click **VMs and Virtual Disks**.
- 4 Click **Trigger VM Storage Policy Compliance Check**.

The Compliance Status column shows the storage compliancy status for virtual machines and their policies.

Compliance Status	Description
Compliant	The datastore that the virtual machine or virtual disk uses has the storage capabilities that are required by the policy.
Not Compliant	The datastore supports specified storage requirements, but cannot currently satisfy the virtual machine storage policy. For example, the status might become Not Compliant when physical resources that back up the datastore are unavailable or exhausted. You can bring the datastore into compliance by making changes in the physical configuration, for example, by adding hosts or disks to the cluster. If additional resources satisfy the virtual machine storage policy, the status changes to Compliant.

Compliance Status	Description
Out of Date	The status indicates that the policy has been edited, but the new requirements have not been communicated to the datastore where the virtual machine objects reside. To communicate the changes, reapply the policy to the objects that are out of date.
Not Applicable	This storage policy references datastore capabilities that are not supported by the datastore where virtual machine resides.

### What to do next

When you cannot bring the non-compliant datastore into compliance, migrate the files or virtual disks to a compatible datastore. See [Find a Compatible Storage Resource for a Noncompliant Virtual Machine](#).

If the status is Out of Date, reapply the policy to the objects. See [Reapply Virtual Machine Storage Policy](#).

## Check Storage Compliance for a Virtual Machine

You can check whether a virtual machine uses a compatible datastore.

### Prerequisites

Ensure that the virtual machine has a storage policy associated with it.

### Procedure

- 1 In the vSphere Web Client, browse to the virtual machine.
- 2 From the right-click menu, select **All vCenter Actions > VM Storage Policies > Check VM Storage Policy Compliance**.

The system verifies the compliance.

- 3 Click the **Summary** tab for the virtual machine.
- 4 View the compliance status in the VM Storage Policies pane.

Compliance Status	Description
Compliant	The datastore that the virtual machine or virtual disk uses has the storage capabilities that are required by the policy.
Not Compliant	The datastore supports specified storage requirements, but cannot currently satisfy the virtual machine storage policy. For example, the status might become Not Compliant when physical resources that back up the datastore are unavailable or exhausted. You can bring the datastore into compliance by making changes in the physical configuration, for example, by adding hosts or disks to the cluster. If additional resources satisfy the virtual machine storage policy, the status changes to Compliant.

Compliance Status	Description
Out of Date	The status indicates that the policy has been edited, but the new requirements have not been communicated to the datastore where the virtual machine objects reside. To communicate the changes, reapply the policy to the objects that are out of date.
Not Applicable	This storage policy references datastore capabilities that are not supported by the datastore where virtual machine resides.

### What to do next

When you cannot bring the non-compliant datastore into compliance, migrate the files or virtual disks to a compatible datastore. See [Find a Compatible Storage Resource for a Noncompliant Virtual Machine](#).

If the status is Out of Date, reapply the policy to the objects. See [Reapply Virtual Machine Storage Policy](#).

## Find a Compatible Storage Resource for a Noncompliant Virtual Machine

Determine which datastore is compatible with the storage policy associated with your virtual machine.

Occasionally, a storage policy assigned to a virtual machine can be in the noncompliant status. This status indicates that the virtual machine or its disks use datastores incompatible with the policy. You can migrate the virtual machine files and virtual disks to compatible datastores.

Use this task to determine which datastores satisfy the requirements of the policy.

### Prerequisites

Verify that the **VM Storage Policy Compliance** field on the virtual machine **Summary** tab displays the Not Compliant status.

### Procedure

**1** In the vSphere Web Client, browse to the virtual machine.

**2** Click the **Summary** tab.

The VM Storage Policy Compliance field on the VM Storage Policies pane shows the Not Compliant status.

**3** Click the policy link in the **VM Storage Policies** field.

**4** Click the **Monitor** tab, and click **VMs and Virtual Disks** to determine which virtual machine files are noncompliant.

**5** Click **Matching Resources**.

The list of datastores that match the requirements of the policy appears.



### What to do next

You can migrate the virtual machine or its disks to one of the datastores on the list.

## Reapply Virtual Machine Storage Policy

After you edit a storage policy that is already associated with a virtual machine object, you need to reapply the policy. By reapplying the policy you communicate new storage requirements to the datastore where the virtual machine object resides.

### Prerequisites

The compliance status for a virtual machine is Out of Date. The status indicates that the policy has been edited, but the new requirements have not been communicated to the datastore.

### Procedure

- 1 In the vSphere Web Client, browse to the virtual machine.
- 2 Click the **Manage** tab, and click **VM Storage Policies**.
- 3 Verify that the compliance status shows Out of Date.
- 4 Click the **Reapply VM Storage Policy** icon.

### Results

The compliance status changes to Compliant.

# About VMware vSphere Flash Read Cache

# 21

Flash Read Cache™ lets you accelerate virtual machine performance through the use of host resident flash devices as a cache.

You can reserve a Flash Read Cache for any individual virtual disk. The Flash Read Cache is created only when a virtual machine is powered on, and it is discarded when a virtual machine is suspended or powered off. When you migrate a virtual machine you have the option to migrate the cache. By default the cache is migrated if the virtual flash module on the source and destination hosts are compatible. If you do not migrate the cache, the cache is rewarmed on the destination host. You can change the size of the cache while a virtual machine is powered on. In this instance, the existing cache is discarded and a new write-through cache is created, which results in a cache warm up period. The advantage of creating a new cache is that the cache size can better match the application's active data.

Flash Read Cache supports write-through or read caching. Write-back or write caching are not supported. Data reads are satisfied from the cache, if present. Data writes are dispatched to the backing storage, such as a SAN or NAS. All data that is read from or written to the backing storage is unconditionally stored in the cache.

---

**Note** Not all workloads benefit with a Flash Read Cache. The performance boost depends on your workload pattern and working set size. Read-intensive workloads with working sets that fit into the cache can benefit from a Flash Read Cache configuration. By configuring Flash Read Cache for your read-intensive workloads additional I/O resources become available on your shared storage, which can result in a performance increase for other workloads even though they are not configured to use Flash Read Cache.

---

This chapter includes the following topics:

- [DRS Support for Flash Read Cache](#)
- [vSphere High Availability Support for Flash Read Cache](#)
- [Using Virtual Flash Resource to Configure a Swap Cache](#)
- [Virtual Flash Volume Limits](#)
- [Set Up Virtual Flash Resource](#)
- [Set Up Multiple Virtual Flash Resources](#)

- [Configure Host Swap Cache with Virtual Flash Resource](#)
- [Configure Flash Read Cache for a Virtual Machine](#)
- [Remove Virtual Flash Resource](#)
- [Migrate Virtual Machines with Flash Read Cache](#)
- [Using esxcli Commands with Virtual Flash](#)

## DRS Support for Flash Read Cache

DRS supports virtual flash as a resource.

DRS manages virtual machines with Flash Read Cache reservations. Every time DRS runs, it displays the available virtual flash capacity reported by the ESXi host. Each host supports one virtual flash resource. DRS selects a host that has sufficient available virtual flash capacity to start a virtual machine. DRS treats powered-on virtual machines with a Flash Read Cache as soft affinity to their current host and moves them only for mandatory reasons or if necessary to correct host over-utilization.

## vSphere High Availability Support for Flash Read Cache

Flash Read Cache is supported by High Availability (HA).

When vSphere HA restarts a virtual machine configured with Flash Read Cache, the virtual machine is restarted on a host in the cluster where the Flash Read Cache, CPU, Memory and overhead reservations are met. vSphere HA will not restart a virtual machine if unreserved flash is insufficient to meet the virtual flash reservation. You must manually reconfigure a virtual machine to reduce or drop the Flash Read Cache, if the target host does not have sufficient virtual flash resource available.

## Using Virtual Flash Resource to Configure a Swap Cache

You can use virtual flash resource on an ESXi host to configure a swap cache. The host swap cache is allocated from a portion of the virtual flash resource.

You can virtualize flash devices on a host to use as a virtual flash resource on that host, regardless of the type of flash device. Virtual flash does not require every host in the cluster to have flash, but only hosts with flash devices can have cache devices configured. You can use virtual flash with existing VMware solutions such as vMotion, DRS and HA.

You can configure vSphere to enable ESXi host swapping to a solid-state disk (SSD). In high memory usage, host swapping is used to reclaim memory from virtual machines. vSphere randomly selects the virtual machine physical pages to be swapped out, compresses memory of the virtual machine pages that are targeted by ESXi to be swapped out, and lets you configure a swap cache on the SSD.

## Virtual Flash Volume Limits

The following virtual flash volume limitations are supported.

Table 21-1. Virtual flash volume limits

Resource	Limit
virtual flash volumes per host	1 (local only)
SSDs per virtual flash volume	8 or fewer
SSD size	4TB or less
virtual flash size	32TB or less

## Set Up Virtual Flash Resource

You can set up a virtual flash resource or add capacity to existing virtual flash resource.

To set up a virtual flash resource you use a solid-state disk (SSD) connected to your host. To increase the capacity of your virtual flash resource, you can add more SSDs, up to a total of eight. An individual SSD must be exclusively allocated to a virtual flash resource and cannot be shared with any other vSphere service, such as Virtual SAN.

### Procedure

- 1 In the vSphere Web Client, navigate to the host.
- 2 Click the **Manage** tab and click **Settings**.
- 3 Under Virtual Flash, select **Virtual Flash Resource Management** and click **Add Capacity**.
- 4 From the list of available SSDs, select one or more SSDs to use for virtual flash and click **OK**.

### Results

The virtual flash resource is created. The Device Backing area lists all SSDs that you use for the virtual flash resource.

### What to do next

You can use the virtual flash resource for cache configuration on the host and Flash Read Cache configuration on virtual disks. You can increase the capacity by adding more SSDs to the virtual flash resource.

## Set Up Multiple Virtual Flash Resources

You can set up multiple virtual flash resources using a batch configuration.

Set up multiple virtual flash resources at the same time or increase virtual flash resource capacity on hosts already configured with virtual flash resource.

### Prerequisites

Verify that hosts are configured with ESXi 5.5 or higher and have eligible SSDs so the hosts can appear in the **Add Virtual Flash Capacity** list.

### Procedure

- 1 In the vSphere Web Client, navigate to the host.
- 2 Right-click the host, select **All vCenter Actions > Add Virtual Flash Capacity**.
- 3 Select one or more SSD.

The selected SSDs are formatted so that the virtual flash resource can use them. All data on the disks is erased.

- 4 Click **OK**.

### Results

Multiple virtual flash resources are created.

## Configure Host Swap Cache with Virtual Flash Resource

You can reserve a certain amount of virtual flash resource for swapping to host cache.

For more information, see [Swapping to Host Cache](#).

### Prerequisites

Set up a virtual flash resource.

### Procedure

- 1 In the vSphere Web Client, navigate to the host.
- 2 Click the **Manage** tab and click **Settings**.
- 3 Under Virtual Flash, select **Virtual Flash Host Swap Cache Configuration** and click **Edit**.
- 4 Select the **Enable virtual flash host swap cache** check box.
- 5 Specify the amount of virtual flash resource to reserve for swapping to host cache.
- 6 Click **OK**.

## Configure Flash Read Cache for a Virtual Machine

You can configure Flash Read Cache for a virtual machine compatible with ESXi 5.5 or later.

Enabling Flash Read Cache lets you specify block size and cache size reservation.

Block size is the minimum number of contiguous bytes that can be stored in the cache. This block size can be larger than the nominal disk block size of 512 bytes, between 4KB and 1024KB. If a guest operating system writes a single 512 byte disk block, the surrounding cache block size bytes will be cached. Do not confuse cache block size with disk block size.

Reservation is a reservation size for cache blocks. There is a minimum number of 256 cache blocks. If the cache block size is 1MB, then the minimum cache size is 256MB. If the cache block size is 4K, then the minimum cache size is 1MB.

For more information about sizing guidelines, search for the *Performance of vSphere Flash Read Cache in VMware vSphere 5.5* white paper on the VMware web site.

### Prerequisites

Verify that the virtual machine is compatible with ESXi 5.5 or later.

### Procedure

- 1 To locate a virtual machine, select a datacenter, folder, cluster, resource pool, host, or vApp.
- 2 Click the **Related Objects** tab and click **Virtual Machines**.
- 3 Right-click the virtual machine and select **Edit Settings**.
- 4 On the **Virtual Hardware** tab, expand **Hard disk** to view the disk options.
- 5 To enable Flash Read Cache for the virtual machine, enter a value in the **Virtual Flash Read Cache** text box.
- 6 Click **Advanced** to specify the following parameters.

Option	Description
<b>Reservation</b>	Select a cache size reservation.
<b>Block Size</b>	Select a block size.

- 7 Click **OK**.

## Remove Virtual Flash Resource

You might need to remove a virtual flash resource on a local SSD to use the device for another resource, such as Virtual SAN.

You cannot remove a virtual flash resource if it is configured with host swap cache or if the host has virtual machines configured with Flash Read Cache that are powered on.

### Procedure

- 1 In the vSphere Web Client, navigate to the host that has virtual flash configured.
- 2 Click the **Manage** tab and click **Settings**.

- 3 Under Virtual Flash, select **Virtual Flash Resource Management** and click **Remove All**.

After you remove the virtual flash resource and erase the SSD, the disk is available for other operations.

## Migrate Virtual Machines with Flash Read Cache

When you migrate a powered on virtual machine from one host to another, you can specify whether or not to migrate Flash Read Cache contents with the virtual disks.

### Prerequisites

If you plan to migrate Flash Read Cache contents, configure a sufficient virtual flash resource on the destination host.

### Procedure

- 1 Right-click the running virtual machine and select **Migrate**.
- 2 Specify the migration type.

Option	Description
<b>Change host</b>	Migrate the virtual machines to another host or cluster.
<b>Change both host and datastore</b>	Migrate the virtual machines to a specific host or cluster and their storage to a specific datastore or datastore cluster.

- 3 Select the destination host and click **Next**.
- 4 Specify a migration setting for all virtual disks configured with virtual Flash Read Cache. This migration parameter does not appear when you do not change the host, but only change the datastore.

Flash Read Cache Migration Settings	Description
<b>Always migrate the cache contents</b>	Virtual machine migration proceeds only if all of the cache contents can be migrated to the destination host. This option is useful when the cache is small or the cache size closely matches the application's active data.
<b>Do not migrate the cache contents</b>	Flash Read Cache is not migrated, the cache is dropped. Flash Read Cache will be recreated on the destination host. This option is useful when the cache size is large or the cache size is larger than the application's active data.

- 5 If you have multiple virtual disks with Flash Read Cache, you can adjust the migration setting for each individual disk.
  - a Click **Advanced**.
  - b Select a virtual disk for which you want to modify the migration setting.
  - c From the drop-down menu in the **Virtual Flash Read Cache Migration Setting** column, select an appropriate option.

6 Complete your migration configuration and click **Finish**.

### What to do next

Verify the successful migration by looking at the **Summary** tab of the virtual machine:

- Make sure that the tab displays the correct IP address of the destination host.
- Make sure that the VM Hardware panel displays correct Virtual Flash Read Cache information for each virtual disk.

## Using esxcli Commands with Virtual Flash

Use the esxcli commands to obtain information about virtual flash and to troubleshoot your environment.

The following commands are available.

Command	Options	Description
esxcli storage vflash		Virtual flash Management Operations on the system.
	cache	Flash Read Cache Management Operations.
	module	Virtual flash module Management Operations.
	device	Virtual flash SSD Device Management Operations.
esxcli storage vflash module		
	get	Get virtual flash module info.
	list	List virtual flash modules.
esxcli storage vflash cache		
	get	Get individual Flash Read Cache info.
	list	List individual Flash Read Caches.
esxcli storage vflash cache stats		
	get	Get Flash Read Cache statistics.
	reset	Reset Flash Read Cache statistics.
esxcli storage vflash device list		List all SSD devices for virtual flash.
	-e --eligible [true   false]	List the eligible or ineligible SSD devices for virtual flash.
	-u --used [true   false]	List the used or unused SSD devices for virtual flash.

The following is a sample output that you get when you use the **vflash device list** command.

```
~ # esxcli storage vflash device list
Name                Size  Is Local  Is Used in vFlash  Eligibility
-----
naa.5001e82002600000 190782    true      true               N/A
naa.5001e82002611111 190782    false     false              Detected as a remote SAS SSD
naa.5001e82002622222 190782    false     false              Detected as a remote SAS SSD
naa.5001e82002633333 190782    true      false              Yes, this is a blank disk.
```



# VMkernel and Storage

# 22

The VMkernel is a high-performance operating system that runs directly on the ESXi host. The VMkernel manages most of the physical resources on the hardware, including memory, physical processors, storage, and networking controllers.

To manage storage, VMkernel has a storage subsystem that supports several Host Bus Adapters (HBAs) including parallel SCSI, SAS, Fibre Channel, FCoE, and iSCSI. These HBAs connect a wide variety of active-active, active-passive, and ALUA storage arrays that are certified for use with the VMkernel. See the *vSphere Compatibility Guide* for a list of the supported HBAs and storage arrays.

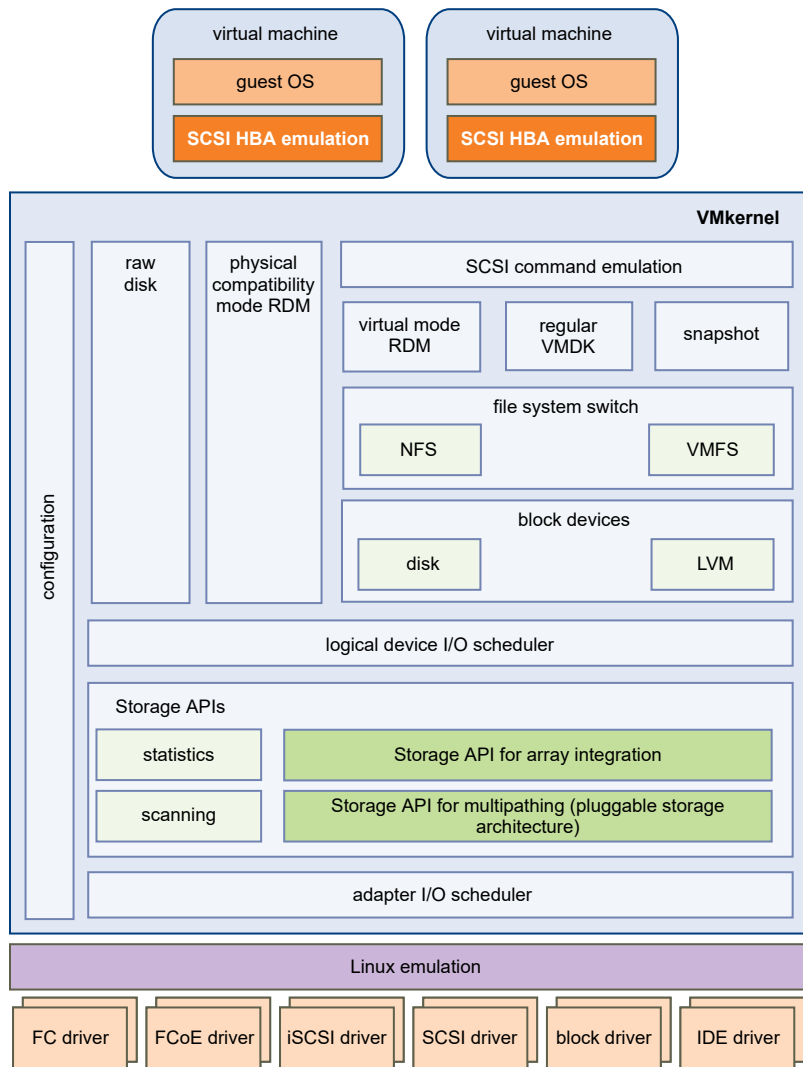
The primary file system that the VMkernel uses is the VMware Virtual Machine File System (VMFS). VMFS is a cluster file system designed and optimized to support large files such as virtual disks and swap files. The VMkernel also supports the storage of virtual disks on NFS file systems.

The storage I/O path provides virtual machines with access to storage devices through device emulation. This device emulation allows a virtual machine to access files on a VMFS or NFS file system as if they were SCSI devices. The VMkernel provides storage virtualization functions such as the scheduling of I/O requests from multiple virtual machines and multipathing.

In addition, VMkernel offers several Storage APIs that enable storage partners to integrate and optimize their products for vSphere.

The following graphic illustrates the basics of the VMkernel core, with special attention to the storage stack. Storage-related modules reside between the logical device I/O scheduler and the adapter I/O scheduler layers.

Figure 22-1. VMkernel and Storage



This chapter includes the following topics:

- [Storage APIs](#)

## Storage APIs

Storage APIs is a family of APIs used by third-party hardware, software, and storage providers to develop components that enhance several vSphere features and solutions.

This publication describes the following sets of Storage APIs and explains how they contribute to your storage environment. For information about other APIs from this family, including Storage API - Data Protection and Storage API - Site Recovery Manager, see the VMware Web site.

- Storage APIs - Multipathing, also known as the Pluggable Storage Architecture (PSA). PSA is a collection of VMkernel APIs that allows storage partners to enable and certify their arrays asynchronous to ESXi release schedules, as well as deliver performance-enhancing, multipathing and load-balancing behaviors that are optimized for each array. For more information, see [Managing Multiple Paths](#).
- Storage APIs - Array Integration, formerly known as VAAI, include the following APIs:
  - Hardware Acceleration APIs. Allows arrays to integrate with vSphere to transparently offload certain storage operations to the array. This integration significantly reduces CPU overhead on the host. See [Chapter 24 Storage Hardware Acceleration](#).
  - Array Thin Provisioning APIs. Help to monitor space use on thin-provisioned storage arrays to prevent out-of-space conditions, and to perform space reclamation. See [Array Thin Provisioning and VMFS Datastores](#).
- Storage APIs - Storage Awareness. These vCenter Server-based APIs enable storage arrays to inform the vCenter Server about their configurations, capabilities, and storage health and events. See [Chapter 26 Using Storage Providers](#).

# Understanding Multipathing and Failover

# 23

To maintain a constant connection between a host and its storage, ESXi supports multipathing. Multipathing is a technique that lets you use more than one physical path that transfers data between the host and an external storage device.

In case of a failure of any element in the SAN network, such as an adapter, switch, or cable, ESXi can switch to another physical path, which does not use the failed component. This process of path switching to avoid failed components is known as path failover.

In addition to path failover, multipathing provides load balancing. Load balancing is the process of distributing I/O loads across multiple physical paths. Load balancing reduces or removes potential bottlenecks.

---

**Note** Virtual machine I/O might be delayed for up to sixty seconds while path failover takes place. These delays allow the SAN to stabilize its configuration after topology changes. In general, the I/O delays might be longer on active-passive arrays and shorter on active-active arrays.

---

This chapter includes the following topics:

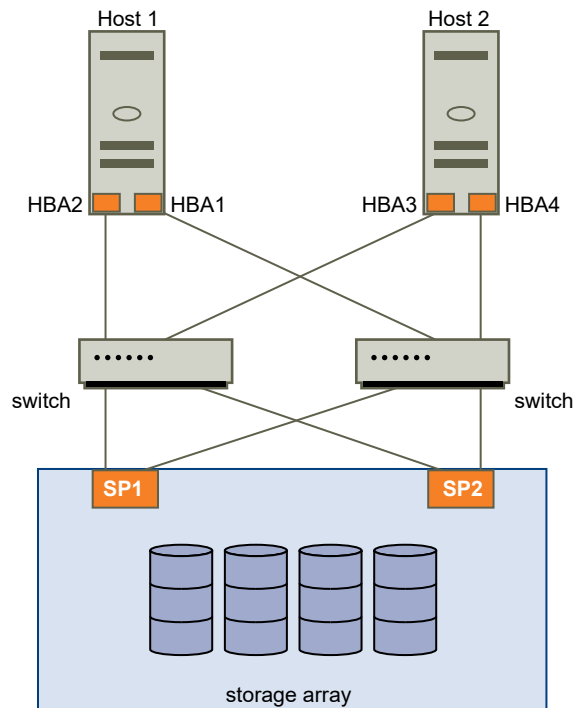
- [Failover with Fibre Channel](#)
- [Host-Based Failover with iSCSI](#)
- [Array-Based Failover with iSCSI](#)
- [Path Failover and Virtual Machines](#)
- [Managing Multiple Paths](#)
- [VMware Multipathing Module](#)
- [Path Scanning and Claiming](#)
- [Managing Storage Paths and Multipathing Plug-Ins](#)

## Failover with Fibre Channel

To support multipathing, your host typically has two or more HBAs available. This configuration supplements the SAN multipathing configuration that generally provides one or more switches in the SAN fabric and one or more storage processors on the storage array device itself.

In the following illustration, multiple physical paths connect each server with the storage device. For example, if HBA1 or the link between HBA1 and the FC switch fails, HBA2 takes over and provides the connection between the server and the switch. The process of one HBA taking over for another is called HBA failover.

**Figure 23-1. Multipathing and Failover with Fibre Channel**



Similarly, if SP1 fails or the links between SP1 and the switches breaks, SP2 takes over and provides the connection between the switch and the storage device. This process is called SP failover. VMware ESXi supports both HBA and SP failovers with its multipathing capability.

## Host-Based Failover with iSCSI

When setting up your ESXi host for multipathing and failover, you can use multiple iSCSI HBAs or multiple NICs depending on the type of iSCSI adapters on your host.

For information on different types of iSCSI adapters, see [iSCSI Initiators](#).

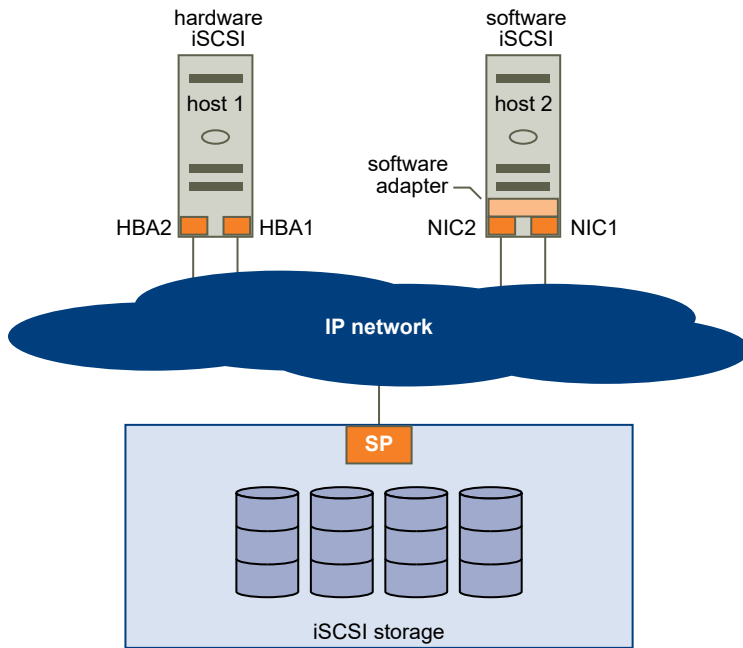
When you use multipathing, specific considerations apply.

- ESXi does not support multipathing when you combine an independent hardware adapter with software iSCSI or dependent iSCSI adapters in the same host.

- Multipathing between software and dependent adapters within the same host is supported.
- On different hosts, you can mix both dependent and independent adapters.

The following illustration shows multipathing setups possible with different types of iSCSI initiators.

**Figure 23-2. Host-Based Path Failover**



## Failover with Hardware iSCSI

With hardware iSCSI, the host typically has two or more hardware iSCSI adapters available, from which the storage system can be reached using one or more switches. Alternatively, the setup might include one adapter and two storage processors so that the adapter can use a different path to reach the storage system.

On the Host-Based Path Failover illustration, Host1 has two hardware iSCSI adapters, HBA1 and HBA2, that provide two physical paths to the storage system. Multipathing plug-ins on your host, whether the VMkernel NMP or any third-party MPPs, have access to the paths by default and can monitor health of each physical path. If, for example, HBA1 or the link between HBA1 and the network fails, the multipathing plug-ins can switch the path over to HBA2.

## Failover with Software iSCSI

With software iSCSI, as shown on Host 2 of the Host-Based Path Failover illustration, you can use multiple NICs that provide failover and load balancing capabilities for iSCSI connections between your host and storage systems.

For this setup, because multipathing plug-ins do not have direct access to physical NICs on your host, you first need to connect each physical NIC to a separate VMkernel port. You then associate all VMkernel ports with the software iSCSI initiator using a port binding technique. As a result, each VMkernel port connected to a separate NIC becomes a different path that the iSCSI storage stack and its storage-aware multipathing plug-ins can use.

For information on how to configure multipathing for software iSCSI, see [Setting Up iSCSI Network](#).

## Array-Based Failover with iSCSI

Some iSCSI storage systems manage path use of their ports automatically and transparently to ESXi.

When using one of these storage systems, your host does not see multiple ports on the storage and cannot choose the storage port it connects to. These systems have a single virtual port address that your host uses to initially communicate. During this initial communication, the storage system can redirect the host to communicate with another port on the storage system. The iSCSI initiators in the host obey this reconnection request and connect with a different port on the system. The storage system uses this technique to spread the load across available ports.

If the ESXi host loses connection to one of these ports, it automatically attempts to reconnect with the virtual port of the storage system, and should be redirected to an active, usable port. This reconnection and redirection happens quickly and generally does not disrupt running virtual machines. These storage systems can also request that iSCSI initiators reconnect to the system, to change which storage port they are connected to. This allows the most effective use of the multiple ports.

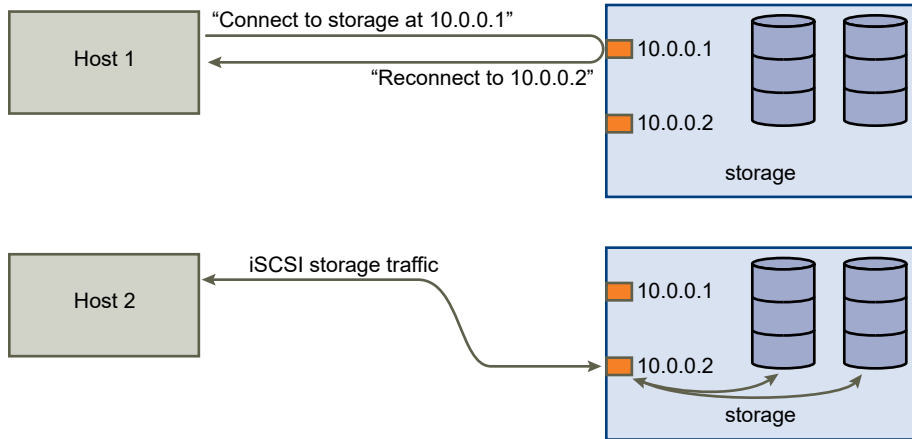
The Port Redirection illustration shows an example of port redirection. The host attempts to connect to the 10.0.0.1 virtual port. The storage system redirects this request to 10.0.0.2. The host connects with 10.0.0.2 and uses this port for I/O communication.

---

**Note** The storage system does not always redirect connections. The port at 10.0.0.1 could be used for traffic, also.

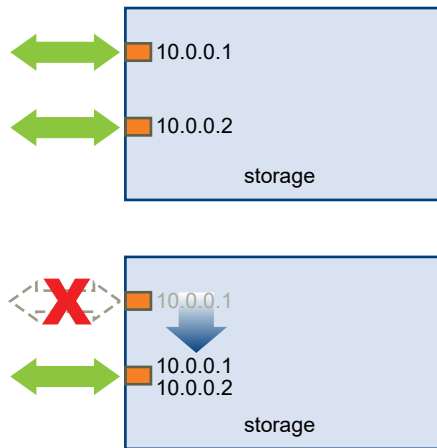
---

Figure 23-3. Port Redirection



If the port on the storage system that is acting as the virtual port becomes unavailable, the storage system reassigns the address of the virtual port to another port on the system. Port Reassignment shows an example of this type of port reassignment. In this case, the virtual port 10.0.0.1 becomes unavailable and the storage system reassigns the virtual port IP address to a different port. The second port responds to both addresses.

Figure 23-4. Port Reassignment



With this form of array-based failover, you can have multiple paths to the storage only if you use multiple ports on the ESXi host. These paths are active-active. For additional information, see [iSCSI Session Management](#).

## Path Failover and Virtual Machines

Path failover occurs when the active path to a LUN is changed from one path to another, usually because of a SAN component failure along the current path.



When a path fails, storage I/O might pause for 30 to 60 seconds until your host determines that the link is unavailable and completes failover. If you attempt to display the host, its storage devices, or its adapters, the operation might appear to stall. Virtual machines with their disks installed on the SAN can appear unresponsive. After failover is complete, I/O resumes normally and the virtual machines continue to run.

However, when failovers take a long time to complete, a Windows virtual machine might interrupt the I/O and eventually fail. To avoid the failure, set the disk timeout value for the Windows virtual machine to at least 60 seconds.

## Set Timeout on Windows Guest OS

Increase the standard disk timeout value on a Windows guest operating system to avoid disruptions during a path failover.

This procedure explains how to change the timeout value by using the Windows registry.

### Prerequisites

Back up the Windows registry.

### Procedure

- 1 Select **Start > Run**.
- 2 Type **regedit.exe**, and click **OK**.
- 3 In the left-panel hierarchy view, double-click **HKEY\_LOCAL\_MACHINE > System > CurrentControlSet > Services > Disk**.
- 4 Double-click **TimeOutValue**.
- 5 Set the value data to 0x3c (hexadecimal) or 60 (decimal) and click **OK**.

After you make this change, Windows waits at least 60 seconds for delayed disk operations to complete before it generates errors.

- 6 Reboot guest OS for the change to take effect.

## Managing Multiple Paths

To manage storage multipathing, ESXi uses a collection of Storage APIs, also called the Pluggable Storage Architecture (PSA). The PSA is an open, modular framework that coordinates the simultaneous operation of multiple multipathing plug-ins (MPPs). The PSA allows 3rd party software developers to design their own load balancing techniques and failover mechanisms for particular storage array, and insert their code directly into the ESXi storage I/O path.

Topics discussing path management use the following acronyms.

Table 23-1. Multipathing Acronyms

Acronym	Definition
PSA	Pluggable Storage Architecture
NMP	Native Multipathing Plug-In. Generic VMware multipathing module.
PSP	Path Selection Plug-In, also called Path Selection Policy. Handles path selection for a given device.
SATP	Storage Array Type Plug-In, also called Storage Array Type Policy. Handles path failover for a given storage array.

The VMkernel multipathing plug-in that ESXi provides by default is the VMware Native Multipathing Plug-In (NMP). The NMP is an extensible module that manages sub plug-ins. There are two types of NMP sub plug-ins, Storage Array Type Plug-Ins (SATPs), and Path Selection Plug-Ins (PSPs). SATPs and PSPs can be built-in and provided by VMware, or can be provided by a third party.

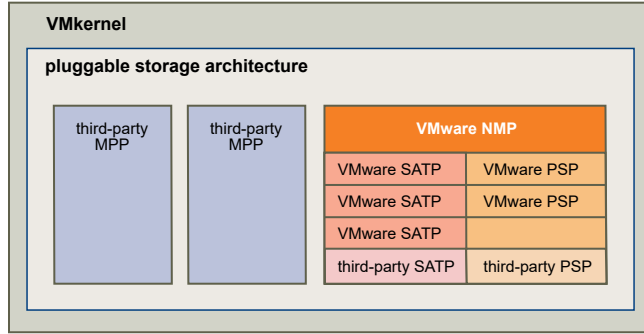
If more multipathing functionality is required, a third party can also provide an MPP to run in addition to, or as a replacement for, the default NMP.

When coordinating the VMware NMP and any installed third-party MPPs, the PSA performs the following tasks:

- Loads and unloads multipathing plug-ins.
- Hides virtual machine specifics from a particular plug-in.
- Routes I/O requests for a specific logical device to the MPP managing that device.
- Handles I/O queueing to the logical devices.
- Implements logical device bandwidth sharing between virtual machines.
- Handles I/O queueing to the physical storage HBAs.
- Handles physical path discovery and removal.
- Provides logical device and physical path I/O statistics.

As the Pluggable Storage Architecture illustration shows, multiple third-party MPPs can run in parallel with the VMware NMP. When installed, the third-party MPPs replace the behavior of the NMP and take complete control of the path failover and the load-balancing operations for specified storage devices.

Figure 23-5. Pluggable Storage Architecture



The multipathing modules perform the following operations:

- Manage physical path claiming and unclaiming.
- Manage creation, registration, and deregistration of logical devices.
- Associate physical paths with logical devices.
- Support path failure detection and remediation.
- Process I/O requests to logical devices:
  - Select an optimal physical path for the request.
  - Depending on a storage device, perform specific actions necessary to handle path failures and I/O command retries.
- Support management tasks, such as reset of logical devices.

## VMware Multipathing Module

By default, ESXi provides an extensible multipathing module called the Native Multipathing Plug-In (NMP).

Generally, the VMware NMP supports all storage arrays listed on the VMware storage HCL and provides a default path selection algorithm based on the array type. The NMP associates a set of physical paths with a specific storage device, or LUN. The specific details of handling path failover for a given storage array are delegated to a Storage Array Type Plug-In (SATP). The specific details for determining which physical path is used to issue an I/O request to a storage device are handled by a Path Selection Plug-In (PSP). SATPs and PSPs are sub plug-ins within the NMP module.

With ESXi, the appropriate SATP for an array you use will be installed automatically. You do not need to obtain or download any SATPs.

## VMware SATPs

Storage Array Type Plug-Ins (SATPs) run in conjunction with the VMware NMP and are responsible for array-specific operations.

ESXi offers a SATP for every type of array that VMware supports. It also provides default SATPs that support non-specific active-active and ALUA storage arrays, and the local SATP for direct-attached devices. Each SATP accommodates special characteristics of a certain class of storage arrays and can perform the array-specific operations required to detect path state and to activate an inactive path. As a result, the NMP module itself can work with multiple storage arrays without having to be aware of the storage device specifics.

After the NMP determines which SATP to use for a specific storage device and associates the SATP with the physical paths for that storage device, the SATP implements the tasks that include the following:

- Monitors the health of each physical path.
- Reports changes in the state of each physical path.
- Performs array-specific actions necessary for storage fail-over. For example, for active-passive devices, it can activate passive paths.

## VMware PSPs

Path Selection Plug-Ins (PSPs) are sub plug-ins of the VMware NMP and are responsible for choosing a physical path for I/O requests.

The VMware NMP assigns a default PSP for each logical device based on the SATP associated with the physical paths for that device. You can override the default PSP. For information, see [Path Scanning and Claiming](#).

By default, the VMware NMP supports the following PSPs:

### VMW\_PSP\_MRU

The host selects the path that it used most recently. When the path becomes unavailable, the host selects an alternative path. The host does not revert back to the original path when that path becomes available again. There is no preferred path setting with the MRU policy. MRU is the default policy for most active-passive storage devices.

The VMW\_PSP\_MRU ranking capability allows you to assign ranks to individual paths. To set ranks to individual paths, use the `esxcli storage nmp psp generic pathconfig set` command. For details, see the VMware knowledge base article at <http://kb.vmware.com/kb/2003468>.

The policy is displayed in the client as the Most Recently Used (VMware) path selection policy.

### VMW\_PSP\_FIXED

The host uses the designated preferred path, if it has been configured. Otherwise, it selects the first working path discovered at system boot time. If you want the host to use a particular

preferred path, specify it manually. Fixed is the default policy for most active-active storage devices.

---

**Note** If the host uses a default preferred path and the path's status turns to Dead, a new path is selected as preferred. However, if you explicitly designate the preferred path, it will remain preferred even when it becomes inaccessible.

---

Displayed in the client as the Fixed (VMware) path selection policy.

### **VMW\_PSP\_RR**

The host uses an automatic path selection algorithm rotating through all active paths when connecting to active-passive arrays, or through all available paths when connecting to active-active arrays. RR is the default for a number of arrays and can be used with both active-active and active-passive arrays to implement load balancing across paths for different LUNs.

Displayed in the client as the Round Robin (VMware) path selection policy.

## **VMware NMP Flow of I/O**

When a virtual machine issues an I/O request to a storage device managed by the NMP, the following process takes place.

- 1 The NMP calls the PSP assigned to this storage device.
- 2 The PSP selects an appropriate physical path on which to issue the I/O.
- 3 The NMP issues the I/O request on the path selected by the PSP.
- 4 If the I/O operation is successful, the NMP reports its completion.
- 5 If the I/O operation reports an error, the NMP calls the appropriate SATP.
- 6 The SATP interprets the I/O command errors and, when appropriate, activates the inactive paths.
- 7 The PSP is called to select a new path on which to issue the I/O.

## **Path Scanning and Claiming**

When you start your ESXi host or rescan your storage adapter, the host discovers all physical paths to storage devices available to the host. Based on a set of claim rules, the host determines which multipathing plug-in (MPP) should claim the paths to a particular device and become responsible for managing the multipathing support for the device.

By default, the host performs a periodic path evaluation every 5 minutes causing any unclaimed paths to be claimed by the appropriate MPP.

The claim rules are numbered. For each physical path, the host runs through the claim rules starting with the lowest number first. The attributes of the physical path are compared to the path specification in the claim rule. If there is a match, the host assigns the MPP specified in the claim rule to manage the physical path. This continues until all physical paths are claimed by corresponding MPPs, either third-party multipathing plug-ins or the native multipathing plug-in (NMP).

For the paths managed by the NMP module, a second set of claim rules is applied. These rules determine which Storage Array Type Plug-In (SATP) should be used to manage the paths for a specific array type, and which Path Selection Plug-In (PSP) is to be used for each storage device.

Use the vSphere Web Client to view which SATP and PSP the host is using for a specific storage device and the status of all available paths for this storage device. If needed, you can change the default VMware PSP using the client. To change the default SATP, you need to modify claim rules using the vSphere CLI.

You can find some information about modifying claim rules in [Managing Storage Paths and Multipathing Plug-Ins](#).

For more information about the commands available to manage PSA, see *Getting Started with vSphere Command-Line Interfaces*.

For a complete list of storage arrays and corresponding SATPs and PSPs, see the SAN Array Model Reference section of the *vSphere Compatibility Guide*.

## Viewing the Paths Information

You can review the storage array type policy (SATP) and path selection policy (PSP) that the ESXi host uses for a specific storage device and the status of all available paths for this storage device. You can access the path information from both the Datastores and Devices views. For datastores, you review the paths that connect to the device the datastore is deployed on.

The path information includes the SATP assigned to manage the device, the PSP, a list of paths, and the status of each path. The following path status information can appear:

### Active

Paths available for issuing I/O to a LUN. A single or multiple working paths currently used for transferring data are marked as Active (I/O).

### Standby

If active paths fail, the path can quickly become operational and can be used for I/O.

### Disabled

The path is disabled and no data can be transferred.

### Dead

The software cannot connect to the disk through this path.

If you are using the **Fixed** path policy, you can see which path is the preferred path. The preferred path is marked with an asterisk (\*) in the Preferred column.

For each path you can also display the path's name. The name includes parameters that describe the path: adapter ID, target ID, and device ID. Usually, the path's name has the format similar to the following:

```
fc.adapterID-fc.targetID-naa.deviceID
```

---

**Note** When you use the host profiles editor to edit paths, you must specify all three parameters that describe a path, adapter ID, target ID, and device ID.

---

## View Datastore Paths in the vSphere Web Client

Review the paths that connect to storage devices backing your datastores.

### Prerequisites

Browse to Datastores in the vSphere Web Client navigator. See [Display Datastore Information in the vSphere Web Client](#).

### Procedure

- 1 Select the datastore.
- 2 Click the **Manage** tab, and click **Settings**.
- 3 Click **Connectivity and Multipathing**.
- 4 If the datastore is shared, select a host to view multipathing details for its devices.
- 5 Under Multipathing Details, review the multipathing policies and paths for the storage device that backs your datastore.

## View Storage Device Paths in the vSphere Web Client

View which multipathing policies the host uses for a specific storage device and the status of all available paths for this storage device.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Devices**.
- 4 Select the storage device whose paths you want to view.
- 5 Click the **Properties** tab and review details under Multipathing Policies.
- 6 Click the **Paths** tab to review all paths available for the storage device.

## Setting a Path Selection Policy

For each storage device, the ESXi host sets the path selection policy based on the claim rules.

By default, VMware supports the following path selection policies. If you have a third-party PSP installed on your host, its policy also appears on the list.

### Fixed (VMware)

The host uses the designated preferred path, if it has been configured. Otherwise, it selects the first working path discovered at system boot time. If you want the host to use a particular preferred path, specify it manually. Fixed is the default policy for most active-active storage devices.

---

**Note** If the host uses a default preferred path and the path's status turns to Dead, a new path is selected as preferred. However, if you explicitly designate the preferred path, it will remain preferred even when it becomes inaccessible.

---

### Most Recently Used (VMware)

The host selects the path that it used most recently. When the path becomes unavailable, the host selects an alternative path. The host does not revert back to the original path when that path becomes available again. There is no preferred path setting with the MRU policy. MRU is the default policy for most active-passive storage devices.

### Round Robin (VMware)

The host uses an automatic path selection algorithm rotating through all active paths when connecting to active-passive arrays, or through all available paths when connecting to active-active arrays. RR is the default for a number of arrays and can be used with both active-active and active-passive arrays to implement load balancing across paths for different LUNs.

## Change the Path Selection Policy in the vSphere Web Client

Generally, you do not have to change the default multipathing settings your host uses for a specific storage device. However, if you want to make any changes, you can use the Manage Paths dialog box to modify a path selection policy and specify the preferred path for the Fixed policy.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Devices**.
- 4 Select the storage device whose paths you want to change and click the **Properties** tab.
- 5 Under Multipathing Properties, click **Edit**.



## 6 Select a path selection policy.

By default, VMware supports the following path selection policies. If you have a third-party PSP installed on your host, its policy also appears on the list.

- Fixed (VMware)
- Most Recently Used (VMware)
- Round Robin (VMware)

## 7 For the fixed policy, specify the preferred path.

## 8 Click **OK** to save your settings and exit the dialog box.

# Disable Paths in the vSphere Web Client

You can temporarily disable paths for maintenance or other reasons.

You disable a path using the Paths panel. You have several ways to access the Paths panel, from a datastore, a storage device, or an adapter view. This task explains how to disable a path using a storage device view.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage**.
- 3 Click **Storage Devices**.
- 4 Select the storage device whose paths you want to disable and click the **Paths** tab.
- 5 Select the path to disable and click **Disable**.

# Managing Storage Paths and Multipathing Plug-Ins

Use the `esxcli` commands to manage the PSA multipathing plug-ins and storage paths assigned to them.

You can display all multipathing plug-ins available on your host. You can list any third-party MPPs, as well as your host's NMP and SATPs and review the paths they claim. You can also define new paths and specify which multipathing plug-in should claim the paths.

For more information about commands available to manage PSA, see the *Getting Started with vSphere Command-Line Interfaces*.

## Multipathing Considerations

Specific considerations apply when you manage storage multipathing plug-ins and claim rules.

The following considerations help you with multipathing:

- If no SATP is assigned to the device by the claim rules, the default SATP for iSCSI or FC devices is `VMW_SATP_DEFAULT_AA`. The default PSP is `VMW_PSP_FIXED`.

- When the system searches the SATP rules to locate a SATP for a given device, it searches the driver rules first. If there is no match, the vendor/model rules are searched, and finally the transport rules are searched. If no match occurs, NMP selects a default SATP for the device.
- If VMW\_SATP\_ALUA is assigned to a specific storage device, but the device is not ALUA-aware, no claim rule match occurs for this device. The device is claimed by the default SATP based on the device's transport type.
- The default PSP for all devices claimed by VMW\_SATP\_ALUA is VMW\_PSP\_MRU. The VMW\_PSP\_MRU selects an active/optimized path as reported by the VMW\_SATP\_ALUA, or an active/unoptimized path if there is no active/optimized path. This path is used until a better path is available (MRU). For example, if the VMW\_PSP\_MRU is currently using an active/unoptimized path and an active/optimized path becomes available, the VMW\_PSP\_MRU will switch the current path to the active/optimized one.
- While VMW\_PSP\_MRU is typically selected for ALUA arrays by default, certain ALUA storage arrays need to use VMW\_PSP\_FIXED. To check whether your storage array requires VMW\_PSP\_FIXED, see the *VMware Compatibility Guide* or contact your storage vendor. When using VMW\_PSP\_FIXED with ALUA arrays, unless you explicitly specify a preferred path, the ESXi host selects the most optimal working path and designates it as the default preferred path. If the host selected path becomes unavailable, the host selects an alternative available path. However, if you explicitly designate the preferred path, it will remain preferred no matter what its status is.
- By default, the PSA claim rule 101 masks Dell array pseudo devices. Do not delete this rule, unless you want to unmask these devices.

## List Multipathing Claim Rules for the Host

Use the `esxccli` command to list available multipathing claim rules.

Claim rules indicate which multipathing plug-in, the NMP or any third-party MPP, manages a given physical path. Each claim rule identifies a set of paths based on the following parameters:

- Vendor/model strings
- Transportation, such as SATA, IDE, Fibre Channel, and so on
- Adapter, target, or LUN location
- Device driver, for example, Mega-RAID

In the procedure, **—server=server\_name** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

## Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

## Procedure

- ◆ Run the **`esxcli --server=server_name storage core claimrule list --claimrule-class=MP`** command to list the multipathing claim rules.

## Example: Sample Output of the `esxcli storage core claimrule list` Command

Rule Class	Rule	Class	Type	Plugin	Matches
MP	0	runtime	transport	NMP	transport=usb
MP	1	runtime	transport	NMP	transport=sata
MP	2	runtime	transport	NMP	transport=ide
MP	3	runtime	transport	NMP	transport=block
MP	4	runtime	transport	NMP	transport=unknown
MP	101	runtime	vendor	MASK_PATH	vendor=DELL model=Universal Xport
MP	101	file	vendor	MASK_PATH	vendor=DELL model=Universal Xport
MP	200	runtime	vendor	MPP_1	vendor=NewVend model=*
MP	200	file	vendor	MPP_1	vendor=NewVend model=*
MP	201	runtime	location	MPP_2	adapter=vmhba41 channel=* target=* lun=*
MP	201	file	location	MPP_2	adapter=vmhba41 channel=* target=* lun=*
MP	202	runtime	driver	MPP_3	driver=megaraid
MP	202	file	driver	MPP_3	driver=megaraid
MP	65535	runtime	vendor	NMP	vendor=* model=*

This example indicates the following:

- The NMP claims all paths connected to storage devices that use the USB, SATA, IDE, and Block SCSI transportation.
- You can use the MASK\_PATH module to hide unused devices from your host. By default, the PSA claim rule 101 masks Dell array pseudo devices with a vendor string of DELL and a model string of Universal Xport.
- The MPP\_1 module claims all paths connected to any model of the NewVend storage array.
- The MPP\_3 module claims the paths to storage devices controlled by the Mega-RAID device driver.
- Any paths not described in the previous rules are claimed by NMP.
- The Rule Class column in the output describes the category of a claim rule. It can be MP (multipathing plug-in), Filter, or VAAI.

- The Class column shows which rules are defined and which are loaded. The `file` parameter in the Class column indicates that the rule is defined. The `runtime` parameter indicates that the rule has been loaded into your system. For a user-defined claim rule to be active, two lines with the same rule number should exist, one line for the rule with the `file` parameter and another line with `runtime`. Several low numbered rules, have only one line with the Class of `runtime`. These are system-defined claim rules that you cannot modify.

## Display Multipathing Modules

Use the `esxcli` command to list all multipathing modules loaded into the system. Multipathing modules manage physical paths that connect your host with storage.

In the procedure, **`--server=server_name`** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

### Procedure

- ◆ To list multipathing modules, run the following command:

```
esxcli --server=server_name storage core plugin list --plugin-class=MP
```

### Results

This command typically shows the NMP and, if loaded, the MASK\_PATH module. If any third-party MPPs have been loaded, they are listed as well.

## Display SATPs for the Host

Use the `esxcli` command to list VMware NMP SATPs loaded into the system. Display information about the SATPs.

In the procedure, **`--server=server_name`** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

**Procedure**

- ◆ To list VMware SATPs, run the following command:

```
esxcli --server=server_name storage nmp satp list
```

**Results**

For each SATP, the output displays information that shows the type of storage array or system this SATP supports and the default PSP for any LUNs using this SATP. Placeholder (plugin not loaded) in the Description column indicates that the SATP is not loaded.

## Display NMP Storage Devices

Use the `esxcli` command to list all storage devices controlled by the VMware NMP and display SATP and PSP information associated with each device.

In the procedure, `--server=server_name` specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

**Prerequisites**

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

**Procedure**

- ◆ To list all storage devices, run the following command:

```
esxcli --server=server_name storage nmp device list
```

Use the `--device | -d=device_ID` option to filter the output of this command to show a single device.

## Add Multipathing Claim Rules

Use the `esxcli` commands to add a new multipathing PSA claim rule to the set of claim rules on the system. For the new claim rule to be active, you first define the rule and then load it into your system.

You add a new PSA claim rule when, for example, you load a new multipathing plug-in (MPP) and need to define which paths this module should claim. You may need to create a claim rule if you add new paths and want an existing MPP to claim them.

---

**Caution** When creating new claim rules, be careful to avoid a situation where different physical paths to the same LUN are claimed by different MPPs. Unless one of the MPPs is the MASK\_PATH MPP, this configuration will cause performance problems.

---

In the procedure, **---server=server\_name** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

### Procedure

- 1 To define a new claim rule, run the following command:

```
esxcli ---server=server_name storage core claimrule add
```

The command takes the following options:

Option	Description
<b>-A --adapter=&lt;str&gt;</b>	Indicate the adapter of the paths to use in this operation.
<b>-u --autoassign</b>	The system will auto assign a rule ID.
<b>-C --channel=&lt;long&gt;</b>	Indicate the channel of the paths to use in this operation.
<b>-c --claimrule-class=&lt;str&gt;</b>	Indicate the claim rule class to use in this operation. Valid values are: MP, Filter, VAAI.
<b>-d --device=&lt;str&gt;</b>	Indicate the device Uid to use for this operation.
<b>-D --driver=&lt;str&gt;</b>	Indicate the driver of the paths to use in this operation.
<b>-f --force</b>	Force claim rules to ignore validity checks and install the rule anyway.
<b>--if-unset=&lt;str&gt;</b>	Execute this command if this advanced user variable is not set to 1.
<b>-i --iqn=&lt;str&gt;</b>	Indicate the iSCSI Qualified Name for the target to use in this operation.
<b>-L --lun=&lt;long&gt;</b>	Indicate the LUN of the paths to use in this operation.
<b>-M --model=&lt;str&gt;</b>	Indicate the model of the paths to use in this operation.
<b>-P --plugin=&lt;str&gt;</b>	Indicate which PSA plugin to use for this operation. (required)
<b>-r --rule=&lt;long&gt;</b>	Indicate the rule ID to use for this operation.
<b>-T --target=&lt;long&gt;</b>	Indicate the target of the paths to use in this operation.
<b>-R --transport=&lt;str&gt;</b>	Indicate the transport of the paths to use in this operation. Valid values are: block, fc, iscsi, iscsivendor, ide, sas, sata, usb, parallel, unknown.
<b>-t --type=&lt;str&gt;</b>	Indicate which type of matching is used for claim/unclaim or claimrule. Valid values are: vendor, location, driver, transport, device, target. (required)
<b>-V --vendor=&lt;str&gt;</b>	Indicate the vendor of the paths to user in this operation.

Option	Description
<b>--wwnn=&lt;str&gt;</b>	Indicate the World-Wide Node Number for the target to use in this operation.
<b>--wwpn=&lt;str&gt;</b>	Indicate the World-Wide Port Number for the target to use in this operation.

- 2 To load the new claim rule into your system, run the following command:

```
esxcli --server=server_name storage core claimrule load
```

This command loads all newly created multipathing claim rules from your system's configuration file.

## Example: Defining Multipathing Claim Rules

In the following example, you add and load rule # 500 to claim all paths with the NewMod model string and the NewVend vendor string for the NMP plug-in.

```
# esxcli --server=server_name storage core claimrule add -r 500 -t vendor -V NewVend -M NewMod -P NMP
```

```
# esxcli --server=server_name storage core claimrule load
```

After you run the **esxcli --server=server\_name storage core claimrule list** command, you can see the new claim rule appearing on the list.

**Note** The two lines for the claim rule, one with the Class of runtime and another with the Class of file, indicate that the new claim rule has been loaded into the system and is active.

Rule	Class	Rule	Class	Type	Plugin	Matches
MP		0	runtime	transport	NMP	transport=usb
MP		1	runtime	transport	NMP	transport=sata
MP		2	runtime	transport	NMP	transport=ide
MP		3	runtime	transport	NMP	transport=block
MP		4	runtime	transport	NMP	transport=unknown
MP		101	runtime	vendor	MASK_PATH	vendor=DELL model=Universal Xport
MP		101	file	vendor	MASK_PATH	vendor=DELL model=Universal Xport
MP		500	runtime	vendor	NMP	vendor=NewVend model=NewMod
MP		500	file	vendor	NMP	vendor=NewVend model=NewMod

## Delete Multipathing Claim Rules

Use the **esxcli** commands to remove a multipathing PSA claim rule from the set of claim rules on the system.

In the procedure, **--server=server\_name** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

## Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

## Procedure

- 1 Delete a claim rule from the set of claim rules.

```
esxcli --server=server_name storage core claimrule remove
```

---

**Note** By default, the PSA claim rule 101 masks Dell array pseudo devices. Do not delete this rule, unless you want to unmask these devices.

---

The command takes the following options:

Option	Description
<b>-c --claimrule-class=&lt;str&gt;</b>	Indicate the claim rule class to use in this operation (MP, Filter, VAAI).
<b>-P --plugin=&lt;str&gt;</b>	Indicate the plugin to use for this operation.
<b>-r --rule=&lt;long&gt;</b>	Indicate the rule ID to use for this operation.

This step removes the claim rule from the File class.

- 2 Remove the claim rule from the system.

```
esxcli --server=server_name storage core claimrule load
```

This step removes the claim rule from the Runtime class.

## Mask Paths

You can prevent the host from accessing storage devices or LUNs or from using individual paths to a LUN. Use the `esxcli` commands to mask the paths. When you mask paths, you create claim rules that assign the MASK\_PATH plug-in to the specified paths.

In the procedure, **--server=*server\_name*** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

## Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

## Procedure

- 1 Check what the next available rule ID is.

```
esxcli --server=server_name storage core claimrule list
```



The claim rules that you use to mask paths should have rule IDs in the range of 101 – 200. If this command shows that rule 101 and 102 already exist, you can specify 103 for the rule to add.

- 2 Assign the MASK\_PATH plug-in to a path by creating a new claim rule for the plug-in.

```
esxcli --server=server_name storage core claimrule add -P MASK_PATH
```

- 3 Load the MASK\_PATH claim rule into your system.

```
esxcli --server=server_name storage core claimrule load
```

- 4 Verify that the MASK\_PATH claim rule was added correctly.

```
esxcli --server=server_name storage core claimrule list
```

- 5 If a claim rule for the masked path exists, remove the rule.

```
esxcli --server=server_name storage core claiming unclaim
```

- 6 Run the path claiming rules.

```
esxcli --server=server_name storage core claimrule run
```

## Results

After you assign the MASK\_PATH plug-in to a path, the path state becomes irrelevant and is no longer maintained by the host. As a result, commands that display the masked path's information might show the path state as dead.

## Example: Masking a LUN

In this example, you mask the LUN 20 on targets T1 and T2 accessed through storage adapters vmhba2 and vmhba3.

```
1 #esxcli --server=server_name storage core claimrule list
```

```
2 #esxcli --server=server_name storage core claimrule add -P MASK_PATH -r 109 -t location -A vmhba2
  -C 0 -T 1 -L 20
  #esxcli --server=server_name storage core claimrule add -P MASK_PATH -r 110 -t location -A vmhba3
  -C 0 -T 1 -L 20
  #esxcli --server=server_name storage core claimrule add -P MASK_PATH -r 111 -t location -A vmhba2
  -C 0 -T 2 -L 20
  #esxcli --server=server_name storage core claimrule add -P MASK_PATH -r 112 -t location -A vmhba3
  -C 0 -T 2 -L 20
```

```
3 #esxcli --server=server_name storage core claimrule load
```

```
4 #esxcli --server=server_name storage core claimrule list
```

```
5 #esxcli --server=server_name storage core claiming unclaim -t location -A vmhba2
  #esxcli --server=server_name storage core claiming unclaim -t location -A vmhba3
```

```
6 #esxcli --server=server_name storage core claimrule run
```

## Unmask Paths

When you need the host to access the masked storage device, unmask the paths to the device.

In the procedure, **--server=*server\_name*** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

### Procedure

- 1 Delete the MASK\_PATH claim rule.

```
esxcli --server=server_name storage core claimrule remove -r rule#
```

- 2 Verify that the claim rule was deleted correctly.

```
esxcli --server=server_name storage core claimrule list
```

- 3 Reload the path claiming rules from the configuration file into the VMkernel.

```
esxcli --server=server_name storage core claimrule load
```

- 4 Run the **esxcli --server=*server\_name* storage core claiming unclaim** command for each path to the masked storage device.

For example:

```
esxcli --server=server_name storage core claiming unclaim -t location -A vmhba0 -C 0 -T 0 -L 149
```

- 5 Run the path claiming rules.

```
esxcli --server=server_name storage core claimrule run
```

### Results

Your host can now access the previously masked storage device.

## Define NMP SATP Rules

The NMP SATP claim rules specify which SATP should manage a particular storage device. Usually you do not need to modify the NMP SATP rules. If you need to do so, use the `esxcli` commands to add a rule to the list of claim rules for the specified SATP.

You might need to create a SATP rule when you install a third-party SATP for a specific storage array.

In the procedure, **--server=server\_name** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

### Procedure

- 1 To add a claim rule for a specific SATP, run the **esxcli --server=server\_name storage nmp satp rule add** command. The command takes the following options.

Option	Description
<b>-b --boot</b>	This is a system default rule added at boot time. Do not modify <code>esx.conf</code> or add to host profile.
<b>-c --claim-option=string</b>	Set the claim option string when adding a SATP claim rule.
<b>-e --description=string</b>	Set the claim rule description when adding a SATP claim rule.
<b>-d --device=string</b>	Set the device when adding SATP claim rules. Device rules are mutually exclusive with vendor/model and driver rules.
<b>-D --driver=string</b>	Set the driver string when adding a SATP claim rule. Driver rules are mutually exclusive with vendor/model rules.
<b>-f --force</b>	Force claim rules to ignore validity checks and install the rule anyway.
<b>-h --help</b>	Show the help message.
<b>-M --model=string</b>	Set the model string when adding SATP a claim rule. Vendor/Model rules are mutually exclusive with driver rules.
<b>-o --option=string</b>	Set the option string when adding a SATP claim rule.
<b>-P --psp=string</b>	Set the default PSP for the SATP claim rule.
<b>-O --psp-option=string</b>	Set the PSP options for the SATP claim rule.
<b>-s --satp=string</b>	The SATP for which a new rule will be added.
<b>-R --transport=string</b>	Set the claim transport type string when adding a SATP claim rule.

Option	Description
<b>-t --type=string</b>	Set the claim type when adding a SATP claim rule.
<b>-V --vendor=string</b>	Set the vendor string when adding SATP claim rules. Vendor/Model rules are mutually exclusive with driver rules.

**Note** When searching the SATP rules to locate a SATP for a given device, the NMP searches the driver rules first. If there is no match, the vendor/model rules are searched, and finally the transport rules. If there is still no match, NMP selects a default SATP for the device.

## 2 Reboot your host.

### Example: Defining an NMP SATP Rule

The following sample command assigns the VMW\_SATP\_INV plug-in to manage storage arrays with vendor string NewVend and model string NewMod.

```
# esxcli --server=server_name storage nmp satp rule add -V NewVend -M NewMod -s VMW_SATP_INV
```

If you run the **esxcli --server=server\_name storage nmp satp list -s VMW\_SATP\_INV** command, you can see the new rule added to the list of VMW\_SATP\_INV rules.

# Storage Hardware Acceleration

# 24

The hardware acceleration functionality enables the ESXi host to integrate with compliant storage arrays and offload specific virtual machine and storage management operations to storage hardware. With the storage hardware assistance, your host performs these operations faster and consumes less CPU, memory, and storage fabric bandwidth.

The hardware acceleration is supported by block storage devices, Fibre Channel and iSCSI, and NAS devices.

For additional details, see the VMware knowledge base article at <http://kb.vmware.com/kb/1021976>.

This chapter includes the following topics:

- [Hardware Acceleration Benefits](#)
- [Hardware Acceleration Requirements](#)
- [Hardware Acceleration Support Status](#)
- [Hardware Acceleration for Block Storage Devices](#)
- [Hardware Acceleration on NAS Devices](#)
- [Hardware Acceleration Considerations](#)

## Hardware Acceleration Benefits

When the hardware acceleration functionality is supported, the host can get hardware assistance and perform several tasks faster and more efficiently.

The host can get assistance with the following activities:

- Migrating virtual machines with Storage vMotion
- Deploying virtual machines from templates
- Cloning virtual machines or templates
- VMFS clustered locking and metadata operations for virtual machine files
- Writes to thin provisioned and thick virtual disks

- Creating fault-tolerant virtual machines
- Creating and cloning thick disks on NFS datastores

## Hardware Acceleration Requirements

The hardware acceleration functionality works only if you use an appropriate host and storage array combination.

**Table 24-1. Hardware Acceleration Storage Requirements**

ESXi	Block Storage Devices	NAS Devices
ESX/ESXi version 4.1	Support block storage plug-ins for array integration (VAAI)	Not supported
ESXi version 5.x	Support T10 SCSI standard or block storage plug-ins for array integration (VAAI)	Support NAS plug-ins for array integration

**Note** If your SAN or NAS storage fabric uses an intermediate appliance in front of a storage system that supports hardware acceleration, the intermediate appliance must also support hardware acceleration and be properly certified. The intermediate appliance might be a storage virtualization appliance, I/O acceleration appliance, encryption appliance, and so on.

## Hardware Acceleration Support Status

For each storage device and datastore, the vSphere Web Client display the hardware acceleration support status.

The status values are Unknown, Supported, and Not Supported. The initial value is Unknown.

For block devices, the status changes to Supported after the host successfully performs the offload operation. If the offload operation fails, the status changes to Not Supported. The status remains Unknown if the device provides partial hardware acceleration support.

With NAS, the status becomes Supported when the storage can perform at least one hardware offload operation.

When storage devices do not support or provide partial support for the host operations, your host reverts to its native methods to perform unsupported operations.

## Hardware Acceleration for Block Storage Devices

With hardware acceleration, your host can integrate with block storage devices, Fibre Channel or iSCSI, and use certain storage array operations.

ESXi hardware acceleration supports the following array operations:

- Full copy, also called clone blocks or copy offload. Enables the storage arrays to make full copies of data within the array without having the host read and write the data. This operation reduces the time and network load when cloning virtual machines, provisioning from a template, or migrating with vMotion.
- Block zeroing, also called write same. Enables storage arrays to zero out a large number of blocks to provide newly allocated storage, free of previously written data. This operation reduces the time and network load when creating virtual machines and formatting virtual disks.
- Hardware assisted locking, also called atomic test and set (ATS). Supports discrete virtual machine locking without use of SCSI reservations. This operation allows disk locking per sector, instead of the entire LUN as with SCSI reservations.

Check with your vendor for the hardware acceleration support. Certain storage arrays require that you activate the support on the storage side.

On your host, the hardware acceleration is enabled by default. If your storage does not support the hardware acceleration, you can disable it.

In addition to hardware acceleration support, ESXi includes support for array thin provisioning. For information, see [Array Thin Provisioning and VMFS Datastores](#).

## Disable Hardware Acceleration for Block Storage Devices in the vSphere Web Client

On your host, the hardware acceleration for block storage devices is enabled by default. You can use the vSphere Web Client advanced settings to disable the hardware acceleration operations.

As with any advanced settings, before you disable the hardware acceleration, consult with the VMware support team.

### Procedure

- 1 Browse to the host in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Settings**.
- 3 Under System, click **Advanced System Settings**.
- 4 Change the value for any of the options to 0 (disabled):
  - VMFS3.HardwareAcceleratedLocking
  - DataMover.HardwareAcceleratedMove
  - DataMover.HardwareAcceleratedInit

## Managing Hardware Acceleration on Block Storage Devices

To integrate with the block storage arrays and to benefit from the array hardware operations, vSphere uses the ESXi extensions referred to as Storage APIs - Array Integration, formerly called VAAI.

In the vSphere 5.x release, these extensions are implemented as the T10 SCSI based commands. As a result, with the devices that support the T10 SCSI standard, your ESXi host can communicate directly and does not require the VAAI plug-ins.

If the device does not support T10 SCSI or provides partial support, ESXi reverts to using the VAAI plug-ins, installed on your host, or uses a combination of the T10 SCSI commands and plug-ins. The VAAI plug-ins are vendor-specific and can be either VMware or partner developed. To manage the VAAI capable device, your host attaches the VAAI filter and vendor-specific VAAI plug-in to the device.

For information about whether your storage requires VAAI plug-ins or supports hardware acceleration through T10 SCSI commands, see the *vSphere Compatibility Guide* or check with your storage vendor.

You can use several `esxcli` commands to query storage devices for the hardware acceleration support information. For the devices that require the VAAI plug-ins, the claim rule commands are also available. For information about `esxcli` commands, see *Getting Started with vSphere Command-Line Interfaces*.

### Display Hardware Acceleration Plug-Ins and Filter

To communicate with the devices that do not support the T10 SCSI standard, your host uses a combination of a single VAAI filter and a vendor-specific VAAI plug-in. Use the `esxcli` command to view the hardware acceleration filter and plug-ins currently loaded into your system.

In the procedure, **`--server=server_name`** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

#### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

#### Procedure

- ◆ Run the **`esxcli --server=server_name storage core plugin list --plugin-class=value`** command.

For *value*, enter one of the following options:

- Type VAAI to display plug-ins.



The output of this command is similar to the following example:

```
#esxcli --server=server_name storage core plugin list --plugin-class=VAAI
Plugin name      Plugin class
VMW_VAAIP_EQL    VAAI
VMW_VAAIP_NETAPP VAAI
VMW_VAAIP_CX     VAAI
```

- Type Filter to display the Filter.

The output of this command is similar to the following example:

```
esxcli --server=server_name storage core plugin list --plugin-class=Filter
Plugin name  Plugin class
VAAI_FILTER Filter
```

## Verify Hardware Acceleration Support Status

Use the `esxcli` command to verify the hardware acceleration support status of a particular storage device.

In the procedure, `--server=server_name` specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

### Procedure

- ◆ Run the `esxcli --server=server_name storage core device list -d=device_ID` command.

The output shows the hardware acceleration, or VAAI, status that can be unknown, supported, or unsupported.

```
# esxcli --server=server_name storage core device list -d naa.XXXXXXXXXXX4c
naa.XXXXXXXXXXX4c
Display Name: XXXX Fibre Channel Disk(naa.XXXXXXXXXXX4c)
Size: 20480
Device Type: Direct-Access
Multipath Plugin: NMP
XXXXXXXXXXXXXXX
Attached Filters: VAAI_FILTER
VAAI Status: supported
XXXXXXXXXXXXXXX
```

## Verify Hardware Acceleration Support Details

Use the `esxcli` command to query the block storage device about the hardware acceleration support the device provides.

In the procedure, `--server=server_name` specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

### Procedure

- ◆ Run the `esxcli --server=server_name storage core device vaaip status get -d=device_ID` command.

If the device is managed by a VAAI plug-in, the output shows the name of the plug-in attached to the device. The output also shows the support status for each T10 SCSI based primitive, if available. Output appears in the following example:

```
# esxcli --server=server_name storage core device vaaip status get -d naa.XXXXXXXXXXX4c
naa.XXXXXXXXXXX4c
VAAI Plugin Name: VMW_VAAIP_SYMM
ATS Status: supported
Clone Status: supported
Zero Status: supported
Delete Status: unsupported
```

## List Hardware Acceleration Claim Rules

Each block storage device managed by a VAAI plug-in needs two claim rules, one that specifies the hardware acceleration filter and another that specifies the hardware acceleration plug-in for the device. You can use the `esxcli` commands to list the hardware acceleration filter and plug-in claim rules.

### Procedure

- 1 To list the filter claim rules, run the `esxcli --server=server_name storage core claimrule list --claimrule-class=Filter` command.

In this example, the filter claim rules specify devices that should be claimed by the VAAI\_FILTER filter.

```
# esxcli --server=server_name storage core claimrule list --claimrule-class=Filter
Rule Class Rule Class Type Plugin Matches
```

Filter	65430	runtime	vendor	VAAI_FILTER	vendor=EMC model=SYMMETRIX
Filter	65430	file	vendor	VAAI_FILTER	vendor=EMC model=SYMMETRIX
Filter	65431	runtime	vendor	VAAI_FILTER	vendor=DGC model=*
Filter	65431	file	vendor	VAAI_FILTER	vendor=DGC model=*

- 2 To list the VAAI plug-in claim rules, run the

**esxcli --server=*server\_name* storage core claimrule list --claimrule-class=VAAI** command.

In this example, the VAAI claim rules specify devices that should be claimed by a particular VAAI plug-in.

```
esxcli --server=server_name storage core claimrule list --claimrule-class=VAAI
```

Rule Class	Rule	Class	Type	Plugin	Matches
VAAI	65430	runtime	vendor	VMW_VAAIP_SYMM	vendor=EMC model=SYMMETRIX
VAAI	65430	file	vendor	VMW_VAAIP_SYMM	vendor=EMC model=SYMMETRIX
VAAI	65431	runtime	vendor	VMW_VAAIP_CX	vendor=DGC model=*
VAAI	65431	file	vendor	VMW_VAAIP_CX	vendor=DGC model=*

## Add Hardware Acceleration Claim Rules

To configure hardware acceleration for a new array, you need to add two claim rules, one for the VAAI filter and another for the VAAI plug-in. For the new claim rules to be active, you first define the rules and then load them into your system.

This procedure is for those block storage devices that do not support T10 SCSI commands and instead use the VAAI plug-ins.

In the procedure, **--server=*server\_name*** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run **esxcli** commands in the ESXi Shell.

### Procedure

- 1 Define a new claim rule for the VAAI filter by running the **esxcli --server=*server\_name* storage core claimrule add --claimrule-class=Filter --plugin=VAAI\_FILTER** command.
- 2 Define a new claim rule for the VAAI plug-in by running the **esxcli --server=*server\_name* storage core claimrule add --claimrule-class=VAAI** command.
- 3 Load both claim rules by running the following commands:
 

```
esxcli --server=server_name storage core claimrule load --claimrule-class=Filter
```

```
esxcli --server=server_name storage core claimrule load --claimrule-class=VAAI
```

- 4 Run the VAAI filter claim rule by running the

```
esxcli --server=server_name storage core claimrule run --claimrule-class=Filter
```

command.

---

**Note** Only the Filter-class rules need to be run. When the VAAI filter claims a device, it automatically finds the proper VAAI plug-in to attach.

---

### Example: Defining Hardware Acceleration Claim Rules

This example shows how to configure hardware acceleration for IBM arrays using the VMW\_VAAIP\_T10 plug-in. Use the following sequence of commands. For information about the options that the command takes, see [Add Multipathing Claim Rules](#).

```
# esxcli --server=server_name storage core claimrule add --claimrule-class=Filter --
plugin=VAAI_FILTER --type=vendor --vendor=IBM --autoassign
```

```
# esxcli --server=server_name storage core claimrule add --claimrule-class=VAAI --
plugin=VMW_VAAIP_T10 --type=vendor --vendor=IBM --autoassign
```

```
# esxcli --server=server_name storage core claimrule load --claimrule-class=Filter
```

```
# esxcli --server=server_name storage core claimrule load --claimrule-class=VAAI
```

```
# esxcli --server=server_name storage core claimrule run --claimrule-class=Filter
```

### Delete Hardware Acceleration Claim Rules

Use the `esxcli` command to delete existing hardware acceleration claim rules.

In the procedure, `--server=server_name` specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

#### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

#### Procedure

- ◆ Run the following commands:

```
esxcli --server=server_name storage core claimrule remove -r claimrule_ID --
claimrule-class=Filter
```

```
esxcli --server=server_name storage core claimrule remove -r claimrule_ID --
claimrule-class=VAAI
```

## Hardware Acceleration on NAS Devices

Hardware acceleration allows your host to integrate with NAS devices and use several hardware operations that NAS storage provides.

The following list shows the supported NAS operations:

- Full file clone. This operation is similar to the VMFS block cloning except that NAS devices clone entire files instead of file segments.
- Reserve space. Enables storage arrays to allocate space for a virtual disk file in thick format.

Typically, when you create a virtual disk on an NFS datastore, the NAS server determines the allocation policy. The default allocation policy on most NAS servers is thin and does not guarantee backing storage to the file. However, the reserve space operation can instruct the NAS device to use vendor-specific mechanisms to reserve space for a virtual disk. As a result, you can create thick virtual disks on the NFS datastore.

- Lazy file clone. Allows VMware View to offload creation of linked clones to a NAS array.
- Extended file statistics. Enables storage arrays to accurately report space utilization.

With NAS storage devices, the hardware acceleration integration is implemented through vendor-specific NAS plug-ins. These plug-ins are typically created by vendors and are distributed as VIB packages through a web page. No claim rules are required for the NAS plug-ins to function.

There are several tools available for installing and upgrading VIB packages. They include the `esxcli` commands and vSphere Update Manager. For more information, see the *vSphere Upgrade* and *Installing and Administering VMware vSphere Update Manager* documentation.

## Install NAS Plug-In

Install vendor-distributed hardware acceleration NAS plug-ins on your host.

This topic provides an example for a VIB package installation using the `esxcli` command. For more details, see the *vSphere Upgrade* documentation.

In the procedure, **`--server=server_name`** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

### Procedure

- 1 Place your host into the maintenance mode.

**2** Set the host acceptance level:

```
esxcli --server=server_name software acceptance set --level=value
```

The command controls which VIB package is allowed on the host. The *value* can be one of the following:

- VMwareCertified
- VMwareAccepted
- PartnerSupported
- CommunitySupported

**3** Install the VIB package:

```
esxcli --server=server_name software vib install -v|--viburl=URL
```

The *URL* specifies the URL to the VIB package to install. http:, https:, ftp:, and file: are supported.

**4** Verify that the plug-in is installed:

```
esxcli --server=server_name software vib list
```

**5** Reboot your host for the installation to take effect.

## Uninstall NAS Plug-Ins

To uninstall a NAS plug-in, remove the VIB package from your host.

This topic discusses how to uninstall a VIB package using the `esxcli` command. For more details, see the *vSphere Upgrade* documentation.

In the procedure, `--server=server_name` specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

### Procedure

**1** Uninstall the plug-in:

```
esxcli --server=server_name software vib remove -n|--vibName=name
```

The *name* is the name of the VIB package to remove.

**2** Verify that the plug-in is removed:

```
esxcli --server=server_name software vib list
```

- 3 Reboot your host for the change to take effect.

## Update NAS Plug-Ins

Upgrade hardware acceleration NAS plug-ins on your host when a storage vendor releases a new plug-in version.

In the procedure, **—server=server\_name** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

This topic discusses how to update a VIB package using the `esxcli` command. For more details, see the *vSphere Upgrade* documentation.

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

### Procedure

- 1 Upgrade to a new plug-in version:

```
esxcli --server=server_name software vib update -v|--viburl=URL
```

The *URL* specifies the URL to the VIB package to update. `http:`, `https:`, `ftp:`, and `file:` are supported.

- 2 Verify that the correct version is installed:

```
esxcli --server=server_name software vib list
```

- 3 Reboot the host.

## Verify Hardware Acceleration Status for NAS

In addition to the client, you can use the `esxcli` command to verify the hardware acceleration status of the NAS device.

In the procedure, **—server=server\_name** specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

## Procedure

- ◆ Run the **esxcli --server=*server\_name* storage nfs list** command.

The Hardware Acceleration column in the output shows the status.

## Hardware Acceleration Considerations

When you use the hardware acceleration functionality, certain considerations apply.

Several reasons might cause a hardware-accelerated operation to fail.

For any primitive that the array does not implement, the array returns an error. The error triggers the ESXi host to attempt the operation using its native methods.

The VMFS data mover does not leverage hardware offloads and instead uses software data movement when one of the following occurs:

- The source and destination VMFS datastores have different block sizes.
- The source file type is RDM and the destination file type is non-RDM (regular file).
- The source VMDK type is eagerzeroedthick and the destination VMDK type is thin.
- The source or destination VMDK is in sparse or hosted format.
- The source virtual machine has a snapshot.
- The logical address and transfer length in the requested operation are not aligned to the minimum alignment required by the storage device. All datastores created with the vSphere Web Client are aligned automatically.
- The VMFS has multiple LUNs or extents, and they are on different arrays.

Hardware cloning between arrays, even within the same VMFS datastore, does not work.



# Storage Thin Provisioning

# 25

With ESXi, you can use two models of thin provisioning, array-level and virtual disk-level.

Thin provisioning is a method that optimizes storage utilization by allocating storage space in a flexible on-demand manner. Thin provisioning contrasts with the traditional model, called thick provisioning. With thick provisioning, large amount of storage space is provided in advance in anticipation of future storage needs. However, the space might remain unused causing underutilization of storage capacity.

The VMware thin provisioning features help you eliminate storage underutilization problems at the datastore and storage array level.

This chapter includes the following topics:

- [Storage Over-Subscription](#)
- [Virtual Disk Thin Provisioning](#)
- [Array Thin Provisioning and VMFS Datastores](#)

## Storage Over-Subscription

Thin provisioning allows you to report more virtual storage space than there is real physical capacity. This discrepancy can lead to storage over-subscription, also called over-provisioning.

When you use thin provisioning, you should monitor actual storage usage to avoid conditions when you run out of physical storage space.

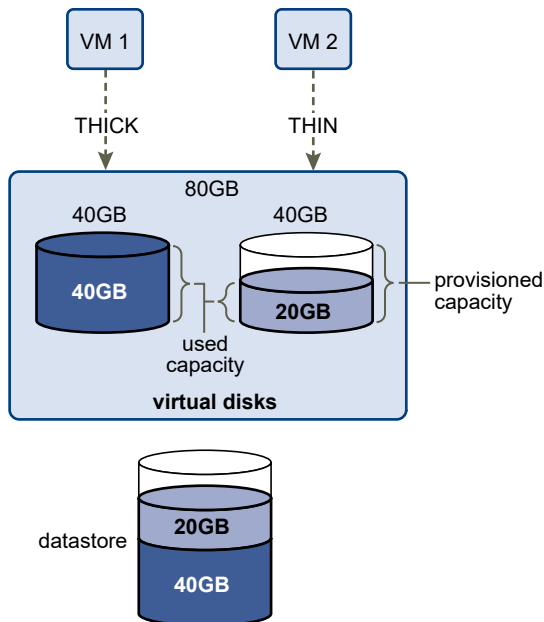
## Virtual Disk Thin Provisioning

When you create a virtual machine, a certain amount of storage space on a datastore is provisioned to virtual disk files.

By default, ESXi offers a traditional storage provisioning method for virtual machines. With this method, you first estimate how much storage the virtual machine will need for its entire life cycle. You then provision a fixed amount of storage space to its virtual disk in advance, for example, 40GB, and have the entire provisioned space committed to the virtual disk. A virtual disk that immediately occupies the entire provisioned space is a thick disk.

ESXi supports thin provisioning for virtual disks. With the disk-level thin provisioning feature, you can create virtual disks in a thin format. For a thin virtual disk, ESXi provisions the entire space required for the disk's current and future activities, for example 40GB. However, the thin disk uses only as much storage space as the disk needs for its initial operations. In this example, the thin-provisioned disk occupies only 20GB of storage. As the disk requires more space, it can grow into its entire 40GB provisioned space.

**Figure 25-1. Thick and thin virtual disks**



## About Virtual Disk Provisioning Policies

When you perform certain virtual machine management operations, such as creating a virtual disk, cloning a virtual machine to a template, or migrating a virtual machine, you can specify a provisioning policy for the virtual disk file.

NFS datastores with Hardware Acceleration and VMFS datastores support the following disk provisioning policies. On NFS datastores that do not support Hardware Acceleration, only thin format is available.

You can use Storage vMotion or cross-host Storage vMotion to transform virtual disks from one format to another.

### Thick Provision Lazy Zeroed

Creates a virtual disk in a default thick format. Space required for the virtual disk is allocated when the disk is created. Data remaining on the physical device is not erased during creation, but is zeroed out on demand at a later time on first write from the virtual machine. Virtual machines do not read stale data from the physical device.

### Thick Provision Eager Zeroed

A type of thick virtual disk that supports clustering features such as Fault Tolerance. Space required for the virtual disk is allocated at creation time. In contrast to the thick provision lazy zeroed format, the data remaining on the physical device is zeroed out when the virtual disk is created. It might take longer to create virtual disks in this format than to create other types of disks.

### Thin Provision

Use this format to save storage space. For the thin disk, you provision as much datastore space as the disk would require based on the value that you enter for the virtual disk size. However, the thin disk starts small and at first, uses only as much datastore space as the disk needs for its initial operations. If the thin disk needs more space later, it can grow to its maximum capacity and occupy the entire datastore space provisioned to it.

Thin provisioning is the fastest method to create a virtual disk because it creates a disk with just the header information. It does not allocate or zero out storage blocks. Storage blocks are allocated and zeroed out when they are first accessed.

---

**Note** If a virtual disk supports clustering solutions such as Fault Tolerance, do not make the disk thin.

---

You can manually inflate the thin disk, so that it occupies the entire provisioned space. If physical storage space is exhausted and the thin provisioned disk cannot grow, the virtual machine becomes unusable.

## Create Thin Provisioned Virtual Disks in the vSphere Web Client

To save storage space, you can create a virtual disk in thin provisioned format. The thin provisioned virtual disk starts small and grows as more disk space is required. You can create thin disks only on the datastores that support disk-level thin provisioning.

This procedure assumes that you are creating a new virtual machine. For information, see the *vSphere Virtual Machine Administration* documentation.

### Procedure

- 1 Right-click any inventory object that is a valid parent object of a virtual machine, such as a datacenter, folder, cluster, resource pool, or host, and select **New Virtual Machine**.
- 2 Select **Create a new virtual machine** and click **Next**.
- 3 Follow the steps required to create a virtual machine.
- 4 On the Customize Hardware page, click the **Virtual Hardware** tab.
- 5 Click the **New Hard Disk** triangle to expand the hard disk options.
- 6 (Optional) Adjust the default disk size.

With a thin virtual disk, the disk size value shows how much space is provisioned and guaranteed to the disk. At the beginning, the virtual disk might not use the entire provisioned space and the actual storage usage value could be less than the size of the virtual disk.

**7** Select **Thin Provision** for Disk Provisioning.

**8** Finish virtual machine creation.

### Results

You created a virtual machine with a disk in thin format.

### What to do next

If you created a virtual disk in the thin format, you can later inflate it to its full size.

## View Virtual Machine Storage Resources in the vSphere Web Client

You can view how datastore storage space is allocated for your virtual machines.

Storage Usage shows how much datastore space is occupied by virtual machine files, including configuration and log files, snapshots, virtual disks, and so on. When the virtual machine is running, the used storage space also includes swap files.

For virtual machines with thin disks, the actual storage usage value might be less than the size of the virtual disk.

### Procedure

- 1** In the vSphere Web Client, browse to the virtual machine.
- 2** Double-click the virtual machine and click the **Summary** tab.
- 3** Review the storage usage information in the upper right area of the **Summary** tab.

## Determine the Disk Format of a Virtual Machine in the vSphere Web Client

You can determine whether your virtual disk is in thick or thin format.

### Procedure

- 1** In the vSphere Web Client, browse to the virtual machine.
- 2** Right-click the virtual machine and select **Edit Settings**.
- 3** Click the **Virtual Hardware** tab.
- 4** Click the **Hard Disk** triangle to expand the hard disk options.

The **Type** text box shows the format of your virtual disk.


### What to do next

If your virtual disk is in the thin format, you can inflate it to its full size.

## Inflate Thin Virtual Disks in the vSphere Web Client

If you created a virtual disk in the thin format, you can convert the thin disk to a virtual disk in thick provision format.

### Procedure

- 1 In the vSphere Web Client, browse to the virtual machine.
- 2 Double-click the virtual machine.
- 3 Click the **Related Objects** tab and click **Datastores**.  
The datastore that stores the virtual machine files is listed.
- 4 Click the datastore link to open the datastore management panel.
- 5 Click the **Manage** tab and click **Files**.
- 6 Open the virtual machine folder and browse to the virtual disk file that you want to convert.  
The file has the .vmdk extension and is marked with the virtual disk  icon.
- 7 Right-click the virtual disk file and select **Inflate**.

### Results

The inflated virtual disk occupies the entire datastore space originally provisioned to it.

## Handling Datastore Over-Subscription

Because the provisioned space for thin disks can be greater than the committed space, a datastore over-subscription can occur, which results in the total provisioned space for the virtual machine disks on the datastore being greater than the actual capacity.

Over-subscription can be possible because usually not all virtual machines with thin disks need the entire provisioned datastore space simultaneously. However, if you want to avoid over-subscribing the datastore, you can set up an alarm that notifies you when the provisioned space reaches a certain threshold.

For information on setting alarms, see the *vCenter Server and Host Management* documentation.

If your virtual machines require more space, the datastore space is allocated on a first come first served basis. When the datastore runs out of space, you can add more physical storage and increase the datastore.

See [Increase VMFS Datastore Capacity in the vSphere Web Client](#).

## Array Thin Provisioning and VMFS Datastores

You can use thin provisioned storage arrays with ESXi.

Traditional LUNs that arrays present to the ESXi host, are thick-provisioned. The entire physical space needed to back each LUN is allocated in advance.

ESXi also supports thin-provisioned LUNs. When a LUN is thin-provisioned, the storage array reports the LUN's logical size, which might be larger than the real physical capacity backing that LUN.

A VMFS datastore that you deploy on the thin-provisioned LUN can detect only the logical size of the LUN. For example, if the array reports 2TB of storage while in reality the array provides only 1TB, the datastore considers 2TB to be the LUN's size. As the datastore grows, it cannot determine whether the actual amount of physical space is still sufficient for its needs.

However, when you use the Storage APIs - Array Integration, the host can integrate with physical storage and become aware of underlying thin-provisioned LUNs and their space usage.

Using thin provision integration, your host can perform these tasks:

- Monitor the use of space on thin-provisioned LUNs to avoid running out of physical space. As your datastore grows or if you use Storage vMotion to migrate virtual machines to a thin-provisioned LUN, the host communicates with the LUN and warns you about breaches in physical space and about out-of-space conditions.
- Inform the array about the datastore space that is freed when files are deleted or removed from the datastore by Storage vMotion. The array can then reclaim the freed blocks of space.

---

**Note** ESXi does not support enabling and disabling of thin provisioning on a storage device.

---

## Requirements

To use the thin provision reporting feature, your host and storage array must meet the following requirements:

- ESXi version 5.0 or later.
- Storage array has appropriate firmware that supports T10-based Storage APIs - Array Integration (Thin Provisioning). For information, contact your storage provider and check the HCL.

## Space Usage Monitoring

The thin provision integration functionality helps you to monitor the space usage on thin-provisioned LUNs and to avoid running out of space.

The following sample flow demonstrates how the ESXi host and the storage array interact to generate breach of space and out-of-space warnings for a datastore with underlying thin-provisioned LUN. The same mechanism applies when you use Storage vMotion to migrate virtual machines to the thin-provisioned LUN.

- 1 Using storage-specific tools, your storage administrator provisions a thin LUN and sets a soft threshold limit that, when reached, triggers an alert. This step is vendor-specific.
- 2 Using the vSphere Web Client, you create a VMFS datastore on the thin-provisioned LUN. The datastore spans the entire logical size that the LUN reports.

- 3 As the space used by the datastore increases and reaches the specified soft threshold, the following actions take place:

- a The storage array reports the breach to your host.
- b Your host triggers a warning alarm for the datastore.

You can contact the storage administrator to request more physical space or use Storage vMotion to evacuate your virtual machines before the LUN runs out of capacity.

- 4 If no space is left to allocate to the thin-provisioned LUN, the following actions take place:

- a The storage array reports out-of-space condition to your host.

---

**Caution** In certain cases, when a LUN becomes full, it might go offline or get unmapped from the host.

---

- b The host pauses virtual machines and generates an out-of-space alarm.

You can resolve the permanent out-of-space condition by requesting more physical space from the storage administrator.

## Identify Thin-Provisioned Storage Devices

Use the `esxcli` command to verify whether a particular storage device is thin-provisioned.

In the procedure, `--server=server_name` specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

### Procedure

- ◆ Run the `esxcli --server=server_name storage core device list -d=device_ID` command.

### Results

The following thin provisioning status indicates that the storage device is thin-provisioned.

```
# esxcli --server=server_name storage core device list -d naa.XXXXXXXXXXX4c
naa.XXXXXXXXXXX4c
Display Name: XXXX Fibre Channel Disk(naa.XXXXXXXXXXX4c)
Size: 20480
Device Type: Direct-Access
Multipath Plugin: NMP
-----
Thin Provisioning Status: yes
```

```
Attached Filters: VAAI_FILTER
VAAI Status: supported
-----
```

An unknown status indicates that a storage device is thick.

**Note** Some storage systems present all devices as thin-provisioned no matter whether the devices are thin or thick. Their thin provisioning status is always yes. For details, check with your storage vendor.

## Reclaim Accumulated Storage Space

When VMFS datastores reside on thin-provisioned LUNs, you can use the `esxcli` command to reclaim unused storage blocks.

In the procedure, `--server=server_name` specifies the target server. The specified target server prompts you for a user name and password. Other connection options, such as a configuration file or session file, are supported. For a list of connection options, see *Getting Started with vSphere Command-Line Interfaces*.

### Prerequisites

Install vCLI or deploy the vSphere Management Assistant (vMA) virtual machine. See *Getting Started with vSphere Command-Line Interfaces*. For troubleshooting, run `esxcli` commands in the ESXi Shell.

### Procedure

- ◆ Reclaim unused storage blocks on the VMFS5 datastore for the thin-provisioned device by running the following command:

```
esxcli --server=server_name storage vmfs unmap --volume-label=volume_label|--volume-uuid=volume_uuid --reclaim-unit=number
```

The command takes these options:

Option	Description
<code>-l --volume-label=volume_label</code>	The label of the VMFS volume to unmap. This is a mandatory argument. If you specify this argument, do not use <code>-u --volume-uuid=volume_uuid</code> .
<code>-u --volume-uuid=volume_uuid</code>	The UUID of the VMFS volume to unmap. This is a mandatory argument. If you specify this argument, do not use <code>-l --volume-label=volume_label</code> .
<code>-n --reclaim-unit=number</code>	Number of VMFS blocks to unmap per iteration. This is an optional argument. If it is not specified, the command uses the default value of 200.

### What to do next

**Important** For additional details, see the VMware knowledge base article at <http://kb.vmware.com/kb/2014849>.



# Using Storage Providers

# 26

vSphere uses storage providers to integrate with external physical storage and storage abstractions. When using storage providers, vSphere can gain a better insight into resources and obtain comprehensive and meaningful storage data. vSphere can also communicate to storage any requirements that are based on storage policies you define for virtual machines.

A storage provider is a software component that is either provided by vSphere or developed by a third party through the vSphere APIs for Storage Awareness (VASA). The third-party storage provider is typically installed on the storage side and acts as a storage awareness service in the vSphere environment. ESXi hosts and vCenter Server connect to the storage provider and obtain information about available storage topology, capabilities, and status. Subsequently vCenter Server provides this information to the vSphere Web Client. This information helps storage policy based management (SPBM) make decisions about virtual disk placement and compliance.

To find out whether your storage supports the storage provider plug-ins, contact your storage vendor.

If your storage supports third-party storage providers, use the vSphere Web Client to register and manage each storage provider component.

Built-in storage providers typically do not require registration. For example, the storage provider that supports Virtual SAN becomes registered automatically when you enable Virtual SAN. For more information, see [Chapter 19 Working with Virtual SAN](#).

This chapter includes the following topics:

- [Storage Providers and Storage Data Representation](#)
- [Storage Provider Requirements and Considerations](#)
- [Storage Status Reporting](#)
- [Register Storage Providers](#)
- [Securing Communication with Storage Providers](#)
- [View Storage Provider Information in the vSphere Web Client](#)
- [Unregister Storage Providers in the vSphere Web Client](#)
- [Update Storage Providers in the vSphere Web Client](#)

## Storage Providers and Storage Data Representation

The vCenter Server communicates with the storage provider to obtain information that the storage provider collects from underlying physical and software-defined storage. The vCenter Server can then display the storage data in the vSphere Web Client.

Information that the storage provider supplies can be divided into the following categories:

- **Storage topology.** Information about storage elements appears in storage reports. It includes such data as storage arrays, array IDs, and so on.

This type of information can be helpful when you need to track virtual machine-to-storage relationships and configuration, or to identify any changes in physical storage configuration.

For more information, see the *vSphere Monitoring and Performance* documentation.

- **Vendor-specific storage capabilities.** The storage provider collects and communicates information about characteristics and services that underlying storage offers.

You reference these capabilities when you define storage requirements for virtual machines and virtual disks in a storage policy. Virtual machines that use this storage policy are placed to a compatible storage that guarantees a specified level of capacity, performance, availability, redundancy, and so on. For details, see [Understanding Storage Capabilities](#).

- **Storage status.** This category includes reporting about status of various storage entities. It also includes alarms and events for notifying about configuration changes.

This type of information can help you troubleshoot storage connectivity and performance problems. It can also help you to correlate array-generated events and alarms to corresponding performance and load changes on the array.

## Storage Provider Requirements and Considerations

When you use the storage provider functionality, certain requirements and considerations apply.

The storage provider functionality is implemented as an extension to the VMware vCenter Storage Monitoring Service (SMS). Because the SMS is a part of the vCenter Server, the storage provider functionality does not need special installation or enablement on the vCenter Server side.

To use storage providers, follow these requirements:

- vCenter Server version 5.0 or later.
- ESX/ESXi hosts version 4.0 or later.
- Storage arrays that support Storage APIs - Storage Awareness plug-ins. The storage provider component must be installed on the storage side. See the *vSphere Compatibility Guide* or check with your storage vendor.

---

**Note** Fibre Channel over Ethernet (FCoE) does not support storage providers.

---

The following considerations exist when you use the storage providers:

- Both block storage and file system storage devices can use storage providers.
- Storage providers can run anywhere, except the vCenter Server.
- Multiple vCenter Servers can simultaneously connect to a single instance of a storage provider.
- A single vCenter Server can simultaneously connect to multiple different storage providers. It is possible to have a different storage provider for each type of physical storage device available to your host.

## Storage Status Reporting

If you use storage providers, the vCenter Server can collect status characteristics for physical storage devices and display this information in the vSphere Web Client.

The status information includes events and alarms.

- Events indicate important changes in the storage configuration. Such changes might include creation and deletion of a LUN, or a LUN becoming inaccessible due to LUN masking.
- Alarms indicate a change in storage system availability. For example, when you use profile-based storage management, you can specify virtual machine storage requirements. When changes to underlying storage occur that might violate the storage requirements of the virtual machine, an alarm gets triggered.

For more information about events and alarms, see the *vSphere Monitoring and Performance* documentation.

Thin-provisioned LUNs have special reporting requirements. For information about space monitoring on thin-provisioned LUNs, see [Array Thin Provisioning and VMFS Datastores](#).

## Register Storage Providers

To establish a connection between the vCenter Server and a storage provider, you must register the storage provider. Make sure to register a separate storage provider for each host in the cluster.

---

**Note** If you use Virtual SAN, the storage providers for Virtual SAN are registered and appear on the list of storage providers automatically. See [View Virtual SAN Storage Providers](#). If for troubleshooting purposes you need to remove or unregister the Virtual SAN storage providers, you should remove the hosts for the storage providers and add the hosts back to the Virtual SAN cluster. Make sure at least one storage provider is in active state. Virtual SAN does not support manual registration of storage providers.

---

## Prerequisites

Verify that the storage provider component is installed on the storage side and obtain its credentials from your storage administrator.

## Procedure

- 1 Browse to the vCenter Server in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage Providers**.
- 3 Click the **Register a new storage provider** icon.
- 4 Type connection information for the storage provider, including the name, URL, and credentials.
- 5 (Optional) To direct the vCenter Server to the storage provider certificate, select the **Use Storage Provider Certificate** option and specify the certificate's location.

If you do not select this option, a thumbprint of the certificate is displayed. You can check the thumbprint and approve it.

This step is not valid for Virtual SAN and is optional for all other types of storage providers.

- 6 Click **OK** to complete the registration.

## Results

The vCenter Server has registered the storage provider and established a secure SSL connection with it.

## What to do next

If your storage provider fails to register, see the VMware Knowledge Base article <http://kb.vmware.com/kb/2079087>. For problems with Virtual SAN storage provider registration, see <http://kb.vmware.com/kb/2105018>.

# Securing Communication with Storage Providers

To communicate with a storage provider, the vCenter Server uses a secure SSL connection. The SSL authentication mechanism requires that both parties, the vCenter Server and the storage provider, exchange SSL certificates and add them to their truststores.

The vCenter Server can add the storage provider certificate to its truststore as part of the storage provider installation. If the certificate is not added during the installation, use one of the following methods to add it when registering the storage provider:

- Direct the vCenter Server to the storage provider certificate. In the **New Storage Provider** dialog box, select the **Use storage provider certificate** option and specify the certificate's location.

- Use a thumbprint of the storage provider certificate. If you do not direct the vCenter Server to use the provider certificate, the certificate thumbprint is displayed. You can check the thumbprint and approve it. The vCenter Server adds the certificate to the truststore and proceeds with the connection.

The storage provider adds the vCenter Server certificate to its truststore when the vCenter Server first connects to the provider.

## View Storage Provider Information in the vSphere Web Client

After you register a storage provider component with the vCenter Server, the Storage provider appears on the storage providers list.

View general Storage provider information and details for each storage component.

### Procedure

- 1 Browse to the vCenter Server in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage Providers**.
- 3 In the Storage Providers list, view the storage provider components registered with the vCenter Server.

The list shows general vendor information including the name, URL, and the time of the last view refresh.

- 4 To display additional details, select a specific storage provider from the list.

The details include storage array vendors and array models that the provider supports.

---

**Note** A single Storage provider can support storage arrays from multiple different vendors.

---

## Unregister Storage Providers in the vSphere Web Client

Unregister storage providers that you do not need.

### Procedure

- 1 Browse to the vCenter Server in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage Providers**.
- 3 From the list of storage providers, select the one you want to unregister and click the **Unregister the storage provider** icon.

### Results

The vCenter Server terminates the connection and removes the storage provider from its configuration.

## Update Storage Providers in the vSphere Web Client

The vCenter Server periodically updates storage data in its database. The updates are partial and reflect only those storage changes that storage providers communicate to the vCenter Server. When needed, you can perform a full database synchronisation for the selected storage provider.

### Procedure

- 1 Browse to the vCenter Server in the vSphere Web Client navigator.
- 2 Click the **Manage** tab, and click **Storage Providers**.
- 3 From the list, select the storage provider that you want to synchronise with and click the **Rescan the storage provider** icon.

### Results

The vSphere Web Client updates the storage data for the provider.

# Using vmkfstools

# 27

`vmkfstools` is one of the ESXi Shell commands for managing VMFS volumes and virtual disks. You can perform many storage operations using the `vmkfstools` command. For example, you can create and manage VMFS datastores on a physical partition, or manipulate virtual disk files, stored on VMFS or NFS datastores.

---

**Note** After you make a change using the `vmkfstools`, the vSphere Web Client might not be updated immediately. You need to use a refresh or rescan operation from the client.

---

For more information on the ESXi Shell, see *Getting Started with vSphere Command-Line Interfaces*.

This chapter includes the following topics:

- [vmkfstools Command Syntax](#)
- [vmkfstools Options](#)

## vmkfstools Command Syntax

Generally, you do not need to log in as the root user to run the `vmkfstools` commands. However, some commands, such as the file system commands, might require the root user login.

The `vmkfstools` command supports the following command syntax:

`vmkfstools conn_options options target.`

Target specifies a partition, device, or path to apply the command option to.

Table 27-1. vmkfstools command arguments

Argument	Description
options	<p>One or more command-line options and associated arguments that you use to specify the activity for vmkfstools to perform, for example, choosing the disk format when creating a new virtual disk.</p> <p>After entering the option, specify a target on which to perform the operation. Target can indicate a partition, device, or path.</p>
partition	<p>Specifies disk partitions. This argument uses a <i>disk_ID:P</i> format, where <i>disk_ID</i> is the device ID returned by the storage array and <i>P</i> is an integer that represents the partition number. The partition digit must be greater than zero (0) and should correspond to a valid VMFS partition.</p>
device	<p>Specifies devices or logical volumes. This argument uses a path name in the ESXi device file system. The path name begins with /vmfs/devices, which is the mount point of the device file system.</p> <p>Use the following formats when you specify different types of devices:</p> <ul style="list-style-type: none"> <li>■ /vmfs/devices/disks for local or SAN-based disks.</li> <li>■ /vmfs/devices/lvm for ESXi logical volumes.</li> <li>■ /vmfs/devices/generic for generic SCSI devices.</li> </ul>
path	<p>Specifies a VMFS file system or file. This argument is an absolute or relative path that names a directory symbolic link, a raw device mapping, or a file under /vmfs.</p> <ul style="list-style-type: none"> <li>■ To specify a VMFS file system, use this format: <div data-bbox="644 1010 1015 1039" data-label="Text"> <pre>/vmfs/volumes/<i>file_system_UUID</i></pre> </div> <p>or</p> <div data-bbox="644 1134 1024 1161" data-label="Text"> <pre>/vmfs/volumes/<i>file_system_label</i></pre> </div> </li> <li>■ To specify a file on a VMFS datastore, use this format: <div data-bbox="644 1245 1313 1297" data-label="Text"> <pre>/vmfs/volumes/<i>file_system_label</i>/<i>file_system_UUID</i>/[dir]/ myDisk.vmdk</pre> </div> </li> </ul> <p>You do not need to enter the entire path if the current working directory is the parent directory of myDisk.vmdk.</p>

## vmkfstools Options

The vmkfstools command has several options. Some of the options are suggested for advanced users only.

The long and single-letter forms of the options are equivalent. For example, the following commands are identical.

```
vmkfstools --createfs vmfs5 --blocksize 1m disk_ID:P
vmkfstools -C vmfs5 -b 1m disk_ID:P
```



## -v Suboption

The `-v` suboption indicates the verbosity level of the command output.

The format for this suboption is as follows:

```
-v --verbose number
```

You specify the *number* value as an integer from 1 through 10.

You can specify the `-v` suboption with any `vmkfstools` option. If the output of the option is not suitable for use with the `-v` suboption, `vmkfstools` ignores `-v`.

---

**Note** Because you can include the `-v` suboption in any `vmkfstools` command line, `-v` is not included as a suboption in the option descriptions.

---

## File System Options

File system options allow you to create and manage VMFS datastores. These options do not apply to NFS. You can perform many of these tasks through the vSphere Web Client.

### Listing Attributes of a VMFS Volume

Use the `vmkfstools` command to list attributes of a VMFS volume.

```
-P --queryfs
    -h --humanreadable
```

When you use this option on any file or directory that resides on a VMFS volume, the option lists the attributes of the specified volume. The listed attributes include the file system label, if any, the number of extents comprising the specified VMFS volume, the UUID, and a listing of the device names where each extent resides.

---

**Note** If any device backing VMFS file system goes offline, the number of extents and available space change accordingly.

---

You can specify the `-h` suboption with the `-P` option. If you do so, `vmkfstools` lists the capacity of the volume in a more readable form, for example, 5k, 12.1M, or 2.1G.

### Creating a VMFS Datastore

Use the `vmkfstools` command to create a VMFS datastore.

```
-C --createfs [vmfs3|vmfs5]
    -b --blocksize block_size kK|mM
    -S --setfsname datastore
```

This option creates a VMFS3 or VMFS5 datastore on the specified SCSI partition, such as `disk_ID:P`. The partition becomes the file system's head partition.

---

**Note** Use the VMFS3 option when you need legacy hosts to access the datastore.

---

You can specify the following suboptions with the `-C` option:

- `-b --blocksize` – Define the block size for the VMFS datastore.

For VMFS5, the only available block size is 1MB. For VMFS3, the default block size is 1MB. Depending on your needs, the block size can be 1MB, 2MB, 4MB, and 8MB. When you enter the size, indicate the unit type by adding a suffix, such as `m` or `M`. The unit type is not case sensitive.

- `-S --setfsname` – Define the volume label of the VMFS datastore you are creating. Use this suboption only in conjunction with the `-C` option. The label you specify can be up to 128 characters long and cannot contain any leading or trailing blank spaces.

---

**Note** vCenter Server supports the 80 character limit for all its entities. If a datastore name exceeds this limit, the name gets shortened when you add this datastore to vCenter Server.

---

After you define a volume label, you can use it whenever you specify the VMFS datastore for the `vmkfstools` command. The volume label appears in listings generated for the `ls -l` command and as a symbolic link to the VMFS volume under the `/vmfs/volumes` directory.

To change the VMFS volume label, use the `ln -sf` command. Use the following as an example:

```
ln -sf /vmfs/volumes/UUID /vmfs/volumes/datastore
```

*datastore* is the new volume label to use for the *UUID* VMFS.

---

**Note** If your host is registered with the vCenter Server, any changes you make to the VMFS volume label get overwritten by the vCenter Server. This guarantees that the VMFS label is consistent across all vCenter Server hosts.

---

## Example for Creating a VMFS File System

This example illustrates creating a new VMFS datastore named `my_vmfs` on the `naa.ID:1` partition. The file block size is 1MB.

```
vmkfstools -C vmfs5 -b 1m -S my_vmfs /vmfs/devices/disks/naa.ID:1
```

## Extending an Existing VMFS Volume

Use the `vmkfstools` command to add an extent to a VMFS volume.

```
-Z --spanfs span_partition head_partition
```

This option extends the VMFS file system with the specified head partition by spanning it across the partition specified by *span\_partition*. You must specify the full path name, for example `/vmfs/devices/disks/disk_ID:1`. Each time you use this option, you extend a VMFS volume with a new extent so that the volume spans multiple partitions.

---

**Caution** When you run this option, you lose all data that previously existed on the SCSI device you specified in *span\_partition*.

---

## Example for Extending a VMFS Volume

In this example, you extend the logical file system by allowing it to span to a new partition.

```
vmkfstools -Z /vmfs/devices/disks/naa.disk_ID_2:1 /vmfs/devices/disks/naa.disk_ID_1:1
```

The extended file system spans two partitions—*naa.disk\_ID\_1:1* and *naa.disk\_ID\_2:1*. In this example, *naa.disk\_ID\_1:1* is the name of the head partition.

## Growing an Existing Extent

Instead of adding a new extent to a VMFS datastore, you can grow an existing extent using the `vmkfstools -G` command.

Use the following option to increase the size of a VMFS datastore after the underlying storage had its capacity increased.

```
-G --growfs device device
```

This option grows an existing VMFS datastore or its extent. For example,

```
vmkfstools --growfs /vmfs/devices/disks/disk_ID:1 /vmfs/devices/disks/disk_ID:1
```

## Upgrading a VMFS Datastore

You can upgrade a VMFS3 to VMFS5 datastore.

---

**Caution** The upgrade is a one-way process. After you have converted a VMFS3 datastore to VMFS5, you cannot revert it back.

---

When upgrading the datastore, use the following command: `vmkfstools -T|-- upgradevmfs /vmfs/volumes/UUID`

---

**Note** All hosts accessing the datastore must support VMFS5. If any ESX/ESXi host version 4.x or earlier is using the VMFS3 datastore, the upgrade fails and the host's mac address is displayed. with the Mac address details of the Host which is actively using the Datastore

---

## Virtual Disk Options

Virtual disk options allow you to set up, migrate, and manage virtual disks stored in VMFS and NFS file systems. You can also perform most of these tasks through the vSphere Web Client.

## Supported Disk Formats

When you create or clone a virtual disk, you can use the `-d --diskformat` suboption to specify the format for the disk.

Choose from the following formats:

- `zeroedthick` (default) – Space required for the virtual disk is allocated during creation. Any data remaining on the physical device is not erased during creation, but is zeroed out on demand at a later time on first write from the virtual machine. The virtual machine does not read stale data from disk.
- `eagerzeroedthick` – Space required for the virtual disk is allocated at creation time. In contrast to `zeroedthick` format, the data remaining on the physical device is zeroed out during creation. It might take much longer to create disks in this format than to create other types of disks.
- `thin` – Thin-provisioned virtual disk. Unlike with the thick format, space required for the virtual disk is not allocated during creation, but is supplied, zeroed out, on demand at a later time.
- `rdm:device` – Virtual compatibility mode raw disk mapping.
- `rdmp:device` – Physical compatibility mode (pass-through) raw disk mapping.
- `2gbsparse` – A sparse disk with 2GB maximum extent size. You can use disks in this format with hosted VMware products, such as VMware Fusion, Player, Server, or Workstation. However, you cannot power on sparse disk on an ESXi host unless you first re-import the disk with `vmkfstools` in a compatible format, such as `thick` or `thin`.

See [Migrate Virtual Machines Between Different VMware Products](#).

## NFS Disk Formats

The only disk formats you can use for NFS are `thin`, `thick`, `zeroedthick` and `2gbsparse`.

Thick, `zeroedthick` and `thin` formats usually behave the same because the NFS server and not the ESXi host determines the allocation policy. The default allocation policy on most NFS servers is `thin`. However, on NFS servers that support Storage APIs - Array Integration, you can create virtual disks in `zeroedthick` format. The reserve space operation enables NFS servers to allocate and guarantee space.

For more information on array integration APIs, see [Chapter 24 Storage Hardware Acceleration](#).

## Creating a Virtual Disk

Use the `vmkfstools` command to create a virtual disk.

```
-c --createvirtualdisk size[kK|mM|gG]
  -a --adaptype [buslogic|lsilogic|ide|lsisas|pvscsi] srcfile
  -d --diskformat [thin|zeroedthick|eagerzeroedthick]
```

This option creates a virtual disk at the specified path on a datastore. Specify the size of the virtual disk. When you enter the value for *size*, you can indicate the unit type by adding a suffix of *k* (kilobytes), *m* (megabytes), or *g* (gigabytes). The unit type is not case sensitive. `vmkfstools` interprets either *k* or *K* to mean kilobytes. If you don't specify a unit type, `vmkfstools` defaults to bytes.

You can specify the following suboptions with the `-c` option.

- `-a` specifies the controller that a virtual machine uses to communicate with the virtual disks. You can choose between BusLogic, LSI Logic, IDE, LSI Logic SAS, and VMware Paravirtual SCSI.
- `-d` specifies disk formats.

## Example for Creating a Virtual Disk

This example illustrates creating a two-gigabyte virtual disk file named `rh6.2.vmdk` on the VMFS file system named `myVMFS`. This file represents an empty virtual disk that virtual machines can access.

```
vmkfstools -c 2048m /vmfs/volumes/myVMFS/rh6.2.vmdk
```

## Initializing a Virtual Disk

Use the `vmkfstools` command to initialize a virtual disk.

```
-w --writezeros
```

This option cleans the virtual disk by writing zeros over all its data. Depending on the size of your virtual disk and the I/O bandwidth to the device hosting the virtual disk, completing this command might take a long time.

---

**Caution** When you use this command, you lose any existing data on the virtual disk.

---

## Inflating a Thin Virtual Disk

Use the `vmkfstools` command to inflate a thin virtual disk.

```
-j --inflatedisk
```

This option converts a thin virtual disk to `eagerzeroedthick`, preserving all existing data. The option allocates and zeroes out any blocks that are not already allocated.

## Removing Zeroed Blocks

Use the `vmkfstools` command to convert any thin, zeroedthick, or eagerzeroedthick virtual disk to a thin disk with zeroed blocks removed.

```
-K --punchzero
```

This option deallocates all zeroed out blocks and leaves only those blocks that were allocated previously and contain valid data. The resulting virtual disk is in thin format.

## Converting a Zeroedthick Virtual Disk to an Eagerzeroedthick Disk

Use the `vmkfstools` command to convert any zeroedthick virtual disk to an eagerzeroedthick disk.

```
-k --eagerzero
```

While performing the conversion, this option preserves any data on the virtual disk.

## Deleting a Virtual Disk

This option deletes a virtual disk file at the specified path on the VMFS volume.

```
-U --deletevirtualdisk
```

## Renaming a Virtual Disk

This option renames a virtual disk file at the specified path on the VMFS volume.

You must specify the original file name or file path *oldName* and the new file name or file path *newName*.

```
-E --renamevirtualdisk oldName newName
```

## Cloning or Converting a Virtual Disk or RDM

Use the `vmkfstools` command to create a copy of a virtual disk or raw disk you specify.

A non-root user is not allowed to clone a virtual disk or an RDM. You must specify the original file name or file path *oldName* and the new file name or file path *newName*.

```
-i|--clonevirtualdisk oldName newName
  -d|--diskformat [thin|zeroedthick|eagerzeroedthick|rdm:device|rdmp:device]
  -W|--objecttype [file|vsan|vvol]
  --policyFile fileName
  -N|--avoidnativeclone
```

Use the following suboptions to change corresponding parameters for the copy you create.

- `-d|--diskformat` specifies disk formats.
- `-W|--objecttype` specifies whether the virtual disk is a file on a VMFS or NFS datastore, or an object on a Virtual SAN or Virtual Volumes datastore.
- `--policyFile fileName` specifies VM storage policy for the disk.

By default, ESXi uses its native methods to perform the cloning operations. If your array supports the cloning technologies, you can off-load the operations to the array. Specify the `-N|--avoidnativeclone` option to avoid the ESXi native cloning.

### Example: Example for Cloning or Converting a Virtual Disk

This example illustrates cloning the contents of a gold virtual disk from the templates repository to a virtual disk file named `myOS.vmdk` on the `myVMFS` file system.

```
vmkfstools -i /vmfs/volumes/myVMFS/templates/gold.vmdk /vmfs/volumes/myVMFS/myOS.vmdk
```

You can configure a virtual machine to use this virtual disk by adding lines to the virtual machine configuration file, as in the following example:

```
scsi0:0.present = TRUE
scsi0:0.fileName = /vmfs/volumes/myVMFS/myOS.vmdk
```

If you want to convert the format of the disk, use the `-d|--diskformat` suboption.

This suboption is useful when you import virtual disks in a format not compatible with ESXi, for example `2gbsparse` format. After you convert the disk, you can attach this disk to a new virtual machine you create in ESXi.

For example:

```
vmkfstools -i /vmfs/volumes/myVMFS/templates/gold.vmdk /vmfs/volumes/myVMFS/myOS.vmdk -d thin
```

## Migrate Virtual Machines Between Different VMware Products

Typically, you use VMware Converter to migrate virtual machines from other VMware products into your ESXi system. However, you can use the `vmkfstools -i` command to import virtual disks in `2gbsparse` format into ESXi and then attach this disk to a new virtual machine you create in ESXi.

You must import the virtual disk first because you cannot power on disks in `2gbsparse` format on the ESXi host.

### Procedure

- 1 Import a disk in `2gbsparse` format into the ESXi host by running the following command. Make sure to select the disk format compatible with ESXi.

```
vmkfstools -i <input> <output> -d <format>
```

- 2 Use the vSphere Web Client to attach the disk you imported to a virtual machine in ESXi.

For information, see the *vSphere Virtual Machine Administration*.

## Extending a Virtual Disk

This option extends the size of a disk allocated to a virtual machine after the virtual machine has been created.

```
-X --extendvirtualdisk newSize [kK|mM|gG]
```

You must power off the virtual machine that uses this disk file before you enter this command. You might have to update the file system on the disk so the guest operating system can recognize and use the new size of the disk and take advantage of the extra space.

---

**Note** You cannot hot-extend virtual SATA disks, or any virtual disk if the capacity after extension is equal to or greater than 2TB.

---

You specify the `newSize` parameter in kilobytes, megabytes, or gigabytes by adding a `k` (kilobytes), `m` (megabytes), or `g` (gigabytes) suffix. The unit type is not case sensitive. `vmkfstools` interprets either `k` or `K` to mean kilobytes. If you don't specify a unit type, `vmkfstools` defaults to kilobytes.

The `newSize` parameter defines the entire new size, not just the increment you add to the disk.

For example, to extend a 4g virtual disk by 1g, enter: `vmkfstools -X 5g disk name`.

You can extend the virtual disk to the `eagerzeroedthick` format by using the `-d eagerzeroedthick` option.

---

**Note** Do not extend the base disk of a virtual machine that has snapshots associated with it. If you do, you can no longer commit the snapshot or revert the base disk to its original size.

---

## Upgrading Virtual Disks

This option converts the specified virtual disk file from ESX Server 2 format to the ESXi format.

```
-M --migratevirtualdisk
```

## Creating a Virtual Compatibility Mode Raw Device Mapping

This option creates a Raw Device Mapping (RDM) file on a VMFS volume and maps a raw LUN to this file. After this mapping is established, you can access the LUN as you would a normal VMFS virtual disk. The file length of the mapping is the same as the size of the raw LUN it points to.

```
-r --createrdm device
```

When specifying the `device` parameter, use the following format:

```
/vmfs/devices/disks/disk_ID:P
```

## Example for Creating a Virtual Compatibility Mode RDM

In this example, you create an RDM file named `my_rdm.vmdk` and map the `disk_ID` raw disk to that file.

```
vmkfstools -r /vmfs/devices/disks/disk_ID my_rdm.vmdk
```



You can configure a virtual machine to use the `my_rdm.vmdk` mapping file by adding the following lines to the virtual machine configuration file:

```
scsi0:0.present = TRUE
scsi0:0.fileName = /vmfs/volumes/myVMFS/my_rdm.vmdk
```

## Creating a Physical Compatibility Mode Raw Device Mapping

This option lets you map a pass-through raw device to a file on a VMFS volume. This mapping lets a virtual machine bypass ESXi SCSI command filtering when accessing its virtual disk. This type of mapping is useful when the virtual machine needs to send proprietary SCSI commands, for example, when SAN-aware software runs on the virtual machine.

```
-z --createrdmpassthru device
```

After you establish this type of mapping, you can use it to access the raw disk just as you would any other VMFS virtual disk.

When specifying the *device* parameter, use the following format:

```
/vmfs/devices/disks/disk_ID
```

## Listing Attributes of an RDM

This option lets you list the attributes of a raw disk mapping.

```
-q --queryrdm
```

This option prints the name of the raw disk RDM. The option also prints other identification information, like the disk ID, for the raw disk.

## Displaying Virtual Disk Geometry

This option gets information about the geometry of a virtual disk.

```
-g --geometry
```

The output is in the form: Geometry information C/H/S, where C represents the number of cylinders, H represents the number of heads, and S represents the number of sectors.

---

**Note** When you import virtual disks from hosted VMware products to the ESXi host, you might see a disk geometry mismatch error message. A disk geometry mismatch might also be the cause of problems loading a guest operating system or running a newly-created virtual machine.

---

## Checking and Repairing Virtual Disks

Use this option to check or repair a virtual disk in case of an unclean shutdown.

```
-x , --fix [check|repair]
```

## Checking Disk Chain for Consistency

With this option, you can check the entire disk chain. You can determine if any of the links in the chain are corrupted or any invalid parent-child relationships exist.

```
-e --chainConsistent
```

## Storage Device Options

Device options allows you to perform administrative task for physical storage devices.

### Managing SCSI Reservations of LUNs

The `-L` option lets you reserve a SCSI LUN for exclusive use by the ESXi host, release a reservation so that other hosts can access the LUN, and reset a reservation, forcing all reservations from the target to be released.

```
-L --lock [reserve|release|lunreset|targetreset|busreset] device
```

**Caution** Using the `-L` option can interrupt the operations of other servers on a SAN. Use the `-L` option only when troubleshooting clustering setups.

Unless specifically advised by VMware, never use this option on a LUN hosting a VMFS volume.

You can specify the `-L` option in several ways:

- `-L reserve` – Reserves the specified LUN. After the reservation, only the server that reserved that LUN can access it. If other servers attempt to access that LUN, a reservation error results.
- `-L release` – Releases the reservation on the specified LUN. Other servers can access the LUN again.
- `-L lunreset` – Resets the specified LUN by clearing any reservation on the LUN and making the LUN available to all servers again. The reset does not affect any of the other LUNs on the device. If another LUN on the device is reserved, it remains reserved.
- `-L targetreset` – Resets the entire target. The reset clears any reservations on all the LUNs associated with that target and makes the LUNs available to all servers again.
- `-L busreset` – Resets all accessible targets on the bus. The reset clears any reservation on all the LUNs accessible through the bus and makes them available to all servers again.

When entering the *device* parameter, use the following format:

```
/vmfs/devices/disks/disk_ID:P
```

### Breaking Device Locks

The `-B` option allows you to forcibly break the device lock on a particular partition.

```
-B --breaklock device
```

When entering the *device* parameter, use the following format:

```
/vmfs/devices/disks/disk_ID:P
```

You can use this command when a host fails in the middle of a datastore operation, such as grow extent, add extent, or resignaturing. When you issue this command, make sure that no other host is holding the lock.