

Fibre Channel SAN Configuration Guide

ESX 4.0

ESXi 4.0

vCenter Server 4.0

This document supports the version of each product listed and supports all subsequent versions until the document is replaced by a new edition. To check for more recent editions of this document, see <http://www.vmware.com/support/pubs>.

EN-000109-05

vmware[®]

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

<http://www.vmware.com/support/>

The VMware Web site also provides the latest product updates.

If you have comments about this documentation, submit your feedback to:

docfeedback@vmware.com

Copyright © 2009, 2010 VMware, Inc. All rights reserved. This product is protected by U.S. and international copyright and intellectual property laws. VMware products are covered by one or more patents listed at <http://www.vmware.com/go/patents>.

VMware is a registered trademark or trademark of VMware, Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.

VMware, Inc.
3401 Hillview Ave.
Palo Alto, CA 94304
www.vmware.com

Contents

Updated Information	5
About This Book	7
1 Overview of VMware ESX/ESXi	9
Introduction to ESX/ESXi	9
Understanding Virtualization	10
Interacting with ESX/ESXi Systems	13
2 Using ESX/ESXi with Fibre Channel SAN	15
Storage Area Network Concepts	15
Overview of Using ESX/ESXi with a SAN	17
Understanding VMFS Datastores	18
Making LUN Decisions	19
Specifics of Using SAN Storage with ESX/ESXi	21
How Virtual Machines Access Data on a SAN	22
Understanding Multipathing and Failover	22
Choosing Virtual Machine Locations	26
Designing for Server Failure	26
Optimizing Resource Use	27
3 Requirements and Installation	29
General ESX/ESXi SAN Requirements	29
ESX Boot from SAN Requirements	31
Installation and Setup Steps	31
4 Setting Up SAN Storage Devices with ESX/ESXi	33
Testing ESX/ESXi SAN Configurations	33
General Setup Considerations for Fibre Channel SAN Arrays	34
EMC CLARiiON Storage Systems	34
EMC Symmetrix Storage Systems	35
IBM System Storage DS4800 Storage Systems	36
IBM Systems Storage 8000 and IBM ESS800	38
HP StorageWorks Storage Systems	39
Hitachi Data Systems Storage	41
Network Appliance Storage	41
5 Using Boot from SAN with ESX Systems	43
Boot from SAN Overview	43
Getting Ready for Boot from SAN	44
Setting Up the QLogic FC HBA for Boot from SAN	46

Setting Up the Emulex FC HBA for Boot from SAN 48

6 Managing ESX/ESXi Systems That Use SAN Storage 51

- Viewing Storage Adapter Information 51
- Viewing Storage Device Information 52
- Viewing Datastore Information 54
- Resolving Display Issues 54
- N-Port ID Virtualization 57
- Path Scanning and Claiming 60
- Path Management and Manual, or Static, Load Balancing 63
- Path Failover 64
- Set Device Driver Options for SCSI Controllers 65
- Sharing Diagnostic Partitions 65
- Disable Automatic Host Registration 66
- Avoiding and Resolving SAN Problems 66
- Optimizing SAN Storage Performance 67
- Resolving Performance Issues 68
- SAN Storage Backup Considerations 71
- Layered Applications 72
- Managing Duplicate VMFS Datastores 73

A Multipathing Checklist 77

B Managing Storage Paths and Multipathing Plugins 79

- List Claim Rules for the Host 79
- Display Multipathing Modules 80
- Display SATPs for the Host 81
- Display NMP Storage Devices 81
- Add PSA Claim Rules 82
- Delete PSA Claim Rules 83
- Mask Paths 83
- Unmask Paths 84
- Define NMP SATP Rules 84
- esxcli corestorage Command-Line Options 86

Index 87

Updated Information

This *Fibre Channel SAN Configuration Guide* is updated with each release of the product or when necessary.

This table provides the update history of the *Fibre Channel SAN Configuration Guide*.

Revision	Description
EN-000109-05	<ul style="list-style-type: none">■ “HP StorageWorks XP,” on page 40 and Appendix A, “Multipathing Checklist,” on page 77 have been changed to include host mode parameters required for HP StorageWorks XP arrays.■ “Boot from SAN Overview,” on page 43 is updated to remove a reference to the restriction on using Microsoft Cluster Service.■ Renamed the topic IBM TotalStorage DS4800 Storage Systems as “IBM System Storage DS4800 Storage Systems,” on page 36 and updated its contents.■ Renamed the topic IBM TotalStorage 8000 as “IBM Systems Storage 8000 and IBM ESS800,” on page 38 and updated its contents.
EN-000109-04	The section “Setting the Timeout for Failover,” on page 30 has been updated to state the recommended HBA timeout value for failover.
EN-000109-03	Appendix A, “Multipathing Checklist,” on page 77 has been updated to include information on the Disk.UseLunReset and Disk.UseDeviceReset parameters for the ESX/ESXi host.
EN-000109-02	<ul style="list-style-type: none">■ “EMC CLARiiON Storage Systems,” on page 34 has been updated to include information on the host automatic registration feature. “Disable Automatic Host Registration,” on page 66 provides instructions about how to turn the feature off.■ “Path Failover,” on page 64 has been updated to remove information on setting HBA timeout parameters for failover. This information is no longer valid for ESX 4.0.
EN-000109-01	<ul style="list-style-type: none">■ The topic “HP StorageWorks EVA,” on page 40 has been updated to remove references to the HP EVA3000/5000 storage devices, which are not supported by ESX 4.0.■ Appendix A, “Multipathing Checklist,” on page 77 has been updated to remove references to the HP EVA3000/5000 storage devices, which are not supported by ESX 4.0.
EN-000109-00	Initial release.

About This Book

The *Fibre Channel SAN Configuration Guide* explains how to use VMware® ESX™ and VMware ESXi systems with a Fibre Channel storage area network (SAN).

The manual discusses conceptual background, installation requirements, and management information in the following main topics:

- Understanding ESX/ESXi – Introduces ESX/ESXi systems for SAN administrators.
- Using ESX/ESXi with a SAN – Discusses requirements, noticeable differences in SAN setup if ESX/ESXi is used, and how to manage and troubleshoot the two systems together.
- Enabling your ESX system to boot from a LUN on a SAN – Discusses requirements, limitations, and management of boot from SAN.

The Fibre Channel SAN Configuration Guide covers ESX, ESXi, and vCenter Server.

Intended Audience

The information presented in this manual is written for experienced Windows or Linux system administrators and who are familiar with virtual machine technology datacenter operations.

Document Feedback

VMware welcomes your suggestions for improving our documentation. If you have comments, send your feedback to docfeedback@vmware.com.

VMware vSphere Documentation

The VMware vSphere documentation consists of the combined VMware vCenter Server and ESX/ESXi documentation set.

Technical Support and Education Resources

The following technical support resources are available to you. To access the current version of this book and other books, go to <http://www.vmware.com/support/pubs>.

Online and Telephone Support

To use online support to submit technical support requests, view your product and contract information, and register your products, go to <http://www.vmware.com/support>.

Customers with appropriate support contracts should use telephone support for the fastest response on priority 1 issues. Go to http://www.vmware.com/support/phone_support.html.

Support Offerings

To find out how VMware support offerings can help meet your business needs, go to <http://www.vmware.com/support/services>.

VMware Professional Services

VMware Education Services courses offer extensive hands-on labs, case study examples, and course materials designed to be used as on-the-job reference tools. Courses are available onsite, in the classroom, and live online. For onsite pilot programs and implementation best practices, VMware Consulting Services provides offerings to help you assess, plan, build, and manage your virtual environment. To access information about education classes, certification programs, and consulting services, go to <http://www.vmware.com/services>.

Overview of VMware ESX/ESXi

You can use ESX/ESXi in conjunction with a Fibre Channel storage area network (SAN), a specialized high-speed network that uses Fibre Channel (FC) protocol to transmit data between your computer systems and high-performance storage subsystems. Using ESX/ESXi with a SAN provides extra storage for consolidation, improves reliability, and helps with disaster recovery.

To use ESX/ESXi effectively with a SAN, you must have a working knowledge of ESX/ESXi systems and SAN concepts.

This chapter includes the following topics:

- [“Introduction to ESX/ESXi,”](#) on page 9
- [“Understanding Virtualization,”](#) on page 10
- [“Interacting with ESX/ESXi Systems,”](#) on page 13

Introduction to ESX/ESXi

The ESX/ESXi architecture allows administrators to allocate hardware resources to multiple workloads in fully isolated environments called virtual machines.

System Components

The main components of ESX/ESXi include a virtualization layer, hardware interface components, and user interface.

An ESX/ESXi system has the following key components.

Virtualization layer This layer provides the idealized hardware environment and virtualization of underlying physical resources to the virtual machines. This layer includes the virtual machine monitor (VMM), which is responsible for virtualization, and VMkernel.

The virtualization layer schedules the virtual machine operating systems and, if you are running an ESX host, the service console. The virtualization layer manages how the operating systems access physical resources. The VMkernel must have its own drivers to provide access to the physical devices. VMkernel drivers are modified Linux drivers, even though the VMkernel is not a Linux variant.

Hardware interface components

The virtual machine communicates with hardware such as CPU or disk by using hardware interface components. These components include device drivers, which enable hardware-specific service delivery while hiding hardware differences from other parts of the system.

User interface

Administrators can view and manage ESX/ESXi hosts and virtual machines in several ways:

- A VMware vSphere Client (vSphere Client) can connect directly to the ESX/ESXi host. This setup is appropriate if your environment has only one host.

A vSphere Client can also connect to vCenter Server and interact with all ESX/ESXi hosts that vCenter Server manages.
- The vSphere Web Access Client allows you to perform many management tasks by using a browser-based interface.
- When you must have command-line access, you can use the VMware vSphere Command-Line Interface (vSphere CLI).

Software and Hardware Compatibility

In the VMware ESX/ESXi architecture, the operating system of the virtual machine (the guest operating system) interacts only with the standard, x86-compatible virtual hardware that the virtualization layer presents. This architecture allows VMware products to support any x86-compatible operating system.

Most applications interact only with the guest operating system, not with the underlying hardware. As a result, you can run applications on the hardware of your choice if you install a virtual machine with the operating system that the application requires.

Understanding Virtualization

The VMware virtualization layer is common across VMware desktop products (such as VMware Workstation) and server products (such as VMware ESX/ESXi). This layer provides a consistent platform for development, testing, delivery, and support of application workloads.

The virtualization layer is organized as follows:

- Each virtual machine runs its own operating system (the guest operating system) and applications.
- The virtualization layer provides the virtual devices that map to shares of specific physical devices. These devices include virtualized CPU, memory, I/O buses, network interfaces, storage adapters and devices, human interface devices, and BIOS.

CPU, Memory, and Network Virtualization

A VMware virtual machine provides complete hardware virtualization. The guest operating system and applications running on a virtual machine can never determine directly which physical resources they are accessing (such as which physical CPU they are running on in a multiprocessor system, or which physical memory is mapped to their pages).

The following virtualization processes occur.

CPU virtualization Each virtual machine appears to run on its own CPU (or a set of CPUs), fully isolated from other virtual machines. Registers, the translation lookaside buffer, and other control structures are maintained separately for each virtual machine.

Most instructions are executed directly on the physical CPU, allowing resource-intensive workloads to run at near-native speed. The virtualization layer safely performs privileged instructions.

Memory virtualization A contiguous memory space is visible to each virtual machine. However, the allocated physical memory might not be contiguous. Instead, noncontiguous physical pages are remapped and presented to each virtual machine. With unusually memory-intensive loads, server memory becomes overcommitted. In that case, some of the physical memory of a virtual machine might be mapped to shared pages or to pages that are unmapped or swapped out.

ESX/ESXi performs this virtual memory management without the information that the guest operating system has and without interfering with the guest operating system's memory management subsystem.

Network virtualization The virtualization layer guarantees that each virtual machine is isolated from other virtual machines. Virtual machines can communicate with each other only through networking mechanisms similar to those used to connect separate physical machines.

The isolation allows administrators to build internal firewalls or other network isolation environments that allow some virtual machines to connect to the outside, while others are connected only through virtual networks to other virtual machines.

Storage Virtualization

ESX/ESXi provides host-level storage virtualization, which logically abstracts the physical storage layer from virtual machines. Virtual machines running on the ESX/ESXi host are not aware of the complexities and specifics of the storage devices to which the host connects.

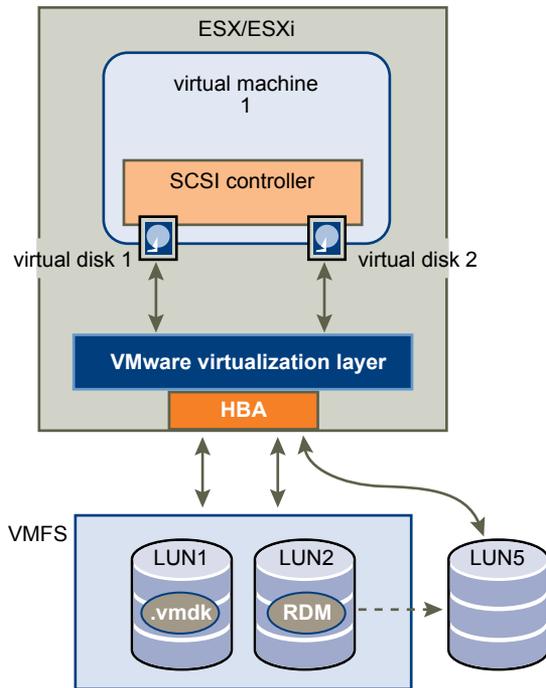
An ESX/ESXi virtual machine uses a virtual hard disk to store its operating system, program files, and other data associated with its activities. A virtual disk is a large physical file, or a set of files, that can be copied, moved, archived, and backed up as easily as any other file. You can configure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

To access virtual disks, a virtual machine uses virtual SCSI controllers. These virtual controllers appear to a virtual machine as different types of controllers, including 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 that a virtual machine can access through one of the virtual SCSI controllers resides in the VMware Virtual Machine File System (VMFS) datastore, NFS-based datastore, or on a raw disk. 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 disk device is being accessed through parallel SCSI, iSCSI, network, or Fibre Channel adapters on the host is transparent to the guest operating system and to applications running on the virtual machine.

Figure 1-1 gives an overview of storage virtualization. The diagram illustrates storage that uses VMFS and storage that uses raw device mapping (RDM).

Figure 1-1. SAN Storage Virtualization



Virtual Machine File System

In a simple configuration, the disks of virtual machines are stored as files within a Virtual Machine File System (VMFS). When guest operating systems issue SCSI commands to their virtual disks, the virtualization layer translates these commands to VMFS file operations.

ESX/ESXi hosts use VMFS to store virtual machine files. With VMFS, multiple virtual machines can run concurrently and have concurrent access to their virtual disk files. Since VMFS is a clustered file system, multiple hosts can have a shared simultaneous access to VMFS datastores on SAN LUNs. VMFS provides the distributed locking to ensure that the multi-host environment is safe.

You can configure a VMFS datastore on either local disks or SAN LUNs. If you use the ESXi host, the local disk is captured and used for the VMFS datastore during the host's first boot.

A VMFS datastore can map to a single SAN LUN or stretch over multiple SAN LUNs. You can expand a datastore while virtual machines are running on it, either by growing the datastore or by adding a new extent. The VMFS datastore can be extended over 32 physical storage extents of the same storage type.

Raw Device Mapping

A raw device mapping (RDM) is a special file in a VMFS volume that acts as a proxy for a raw device, such as a SAN LUN. With the RDM, the SAN LUN can be directly and entirely allocated to a virtual machine. The RDM provides some of the advantages of a virtual disk in the VMFS file system, while keeping some advantages of direct access to physical devices.

An RDM might be required if you use Microsoft Cluster Service (MSCS) or if you run SAN snapshot or other layered applications on the virtual machine. RDMs enable systems to use the hardware features inherent to SAN arrays. However, virtual machines with RDMs do not display performance gains compared to virtual machines with virtual disk files stored on a VMFS datastore.

For more information on the RDM, see the *ESX Configuration Guide* or *ESXi Configuration Guide*.

Interacting with ESX/ESXi Systems

You can interact with ESX/ESXi systems in several different ways. You can use a client or, in special cases, interact programmatically.

Administrators can interact with ESX/ESXi systems in one of the following ways:

- With a GUI client (vSphere Client or vSphere Web Access). You can connect clients directly to the ESX/ESXi host, or you can manage multiple ESX/ESXi hosts simultaneously with vCenter Server.
- Through the command-line interface. vSphere Command-Line Interface (vSphere CLI) commands are scripts that run on top of the vSphere SDK for Perl. The vSphere CLI package includes commands for storage, network, virtual machine, and user management and allows you to perform most management operations. For more information, see *vSphere Command-Line Interface Installation and Reference Guide*.
- ESX administrators can also use the ESX service console, which supports a full Linux environment and includes all vSphere CLI commands. Using the service console is less secure than remotely running the vSphere CLI. The service console is not supported on ESXi.

VMware vCenter Server

vCenter Server is a central administrator for ESX/ESXi hosts. You can access vCenter Server through a vSphere Client or vSphere Web Access.

vCenter Server	vCenter Server acts as a central administrator for your hosts connected on a network. The server directs actions upon the virtual machines and VMware ESX/ESXi.
vSphere Client	The vSphere Client runs on Microsoft Windows. In a multihost environment, administrators use the vSphere Client to make requests to vCenter Server, which in turn affects its virtual machines and hosts. In a single-server environment, the vSphere Client connects directly to an ESX/ESXi host.
vSphere Web Access	vSphere Web Access allows you to connect to vCenter Server by using an HTML browser.

Using ESX/ESXi with Fibre Channel SAN

2

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

This chapter includes the following topics:

- [“Storage Area Network Concepts,”](#) on page 15
- [“Overview of Using ESX/ESXi with a SAN,”](#) on page 17
- [“Understanding VMFS Datastores,”](#) on page 18
- [“Making LUN Decisions,”](#) on page 19
- [“Specifics of Using SAN Storage with ESX/ESXi,”](#) on page 21
- [“How Virtual Machines Access Data on a SAN,”](#) on page 22
- [“Understanding Multipathing and Failover,”](#) on page 22
- [“Choosing Virtual Machine Locations,”](#) on page 26
- [“Designing for Server Failure,”](#) on page 26
- [“Optimizing Resource Use,”](#) on page 27

Storage Area Network Concepts

If you are an ESX/ESXi administrator planning to set up ESX/ESXi 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 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. Usually, LUN masking is performed at the SP or server level.

Ports

In the context of this document, a port is the connection from a device into the SAN. Each node in the SAN, a host, storage device, and 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) In the 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 ESX/ESXi hosts use a SAN, these multiple, unique identifiers allow the assignment of WWNs to individual virtual machines as part of their configuration.

Multipathing and Path Failover

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

If a path or any component along the path, HBA or NIC, cable, switch or switch port, or storage processor, fails, the server selects another of the available paths. The process of detecting a failed path and switching to another is called path failover.

Storage System Types

Storage disk systems can be active-active and active-passive.

ESX/ESXi supports the following types of storage systems:

- An active-active storage system, which 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.
- An active-passive storage system, in which one port is actively providing access to a given LUN. The other ports 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 primary storage port fails, one of the secondary ports or storage processors becomes active, either automatically or through administrator intervention.

Overview of Using ESX/ESXi with a SAN

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

The following are benefits of using ESX/ESXi with a SAN:

- You can store data redundantly and configure multiple paths to your storage, eliminating a single point of failure. ESX/ESXi systems provide multipathing by default for every virtual machine.
- Using a SAN with ESX/ESXi systems extends failure resistance to the server. When you use SAN storage, all applications can instantly be restarted after host failure.
- You can perform live migration of virtual machines using VMware VMotion.
- Use VMware High Availability (HA) in conjunction with a SAN for a cold-standby solution that guarantees an immediate, automatic response.
- Use VMware Distributed Resource Scheduler (DRS) to migrate virtual machines from one host to another for load balancing. Because storage is on a SAN array, applications continue running seamlessly.
- If you use VMware DRS clusters, put an ESX/ESXi host into maintenance mode to have the system migrate all running virtual machines to other ESX/ESXi hosts. You can then perform upgrades or other maintenance operations.

The transportability 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.

ESX/ESXi and SAN Use Cases

You can perform a number of tasks when using ESX/ESXi with SAN.

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

Maintenance with zero downtime

When performing an ESX/ESXi host or infrastructure maintenance, use VMware DRS or VMotion to migrate virtual machines to other servers. If shared storage is on the SAN, you can perform maintenance without interruptions to the user.

Load balancing

Use VMotion or VMware DRS to migrate virtual machines to other hosts for load balancing. If shared storage is on a SAN, you can perform load balancing without interruption to the user.

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. Choosing a SAN for external storage results in a simpler system architecture along with other benefits.

Start by reserving a large LUN and then allocate portions to virtual machines as needed. LUN reservation and creation from the storage device needs to happen only once.

Disaster recovery	Having all data stored on a SAN facilitates the remote storage of data backups. You can restart virtual machines on remote ESX/ESXi hosts for recovery if one site is compromised.
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.

Finding Further Information

In addition to this document, a number of other resources can help you configure your ESX/ESXi system in conjunction with a SAN.

- Use your storage array vendor's documentation for most setup questions. Your storage array vendor might also offer documentation on using the storage array in an ESX/ESXi environment.
- The VMware Documentation Web site.
- The *iSCSI SAN Configuration Guide* discusses the use of ESX/ESXi with iSCSI storage area networks.
- The *VMware I/O Compatibility Guide* lists the currently approved HBAs, HBA drivers, and driver versions.
- The *VMware Storage/SAN Compatibility Guide* lists currently approved storage arrays.
- The *VMware Release Notes* give information about known issues and workarounds.
- The *VMware Knowledge Bases* have information on common issues and workarounds.

Understanding VMFS Datastores

To store virtual disks, ESX/ESXi uses datastores, which are logical containers that hide specifics of storage from virtual machines and provide a uniform model for storing virtual machine files. Datastores that you deploy on storage devices use the VMware Virtual Machine File System (VMFS) format, a special high-performance file system format that is optimized for storing virtual machines.

A VMFS datastore can run multiple virtual machines as one workload. VMFS provides distributed locking for your virtual machine files, so that your virtual machines can operate safely in a SAN environment where multiple ESX/ESXi hosts share a set of LUNs.

Use the vSphere Client to set up a VMFS datastore in advance on any SCSI-based storage device that your ESX/ESXi host discovers. A VMFS datastore can be extended over several physical storage extents, including SAN LUNs and local storage. This feature allows you to pool storage and gives you flexibility in creating the storage volume necessary for your virtual machine.

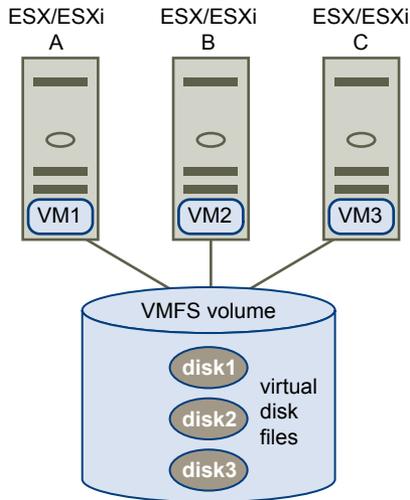
You can increase the capacity of a datastore while virtual machines are running on the datastore. This ability lets you add new space to your VMFS datastores as your virtual machine requires it. ESX/ESXi VMFS is designed for concurrent access from multiple physical machines and enforces the appropriate access controls on virtual machine files.

Sharing a VMFS Datastore Across ESX/ESXi Hosts

As a cluster file system, VMFS lets multiple ESX/ESXi hosts access the same VMFS datastore concurrently.

To ensure that multiple servers do not access the same virtual machine at the same time, VMFS provides on-disk locking. To coordinate access to VMFS internal file system information, ESX/ESXi uses SCSI reservations on the entire LUN.

[Figure 2-1](#) shows several ESX/ESXi systems sharing the same VMFS volume.

Figure 2-1. Sharing a VMFS Datastore Across ESX/ESXi Hosts

Because virtual machines share a common VMFS datastore, it might be difficult to characterize peak-access periods or to optimize performance. You must plan virtual machine storage access for peak periods, but different applications might have different peak-access periods. VMware recommends that you load balance virtual machines over servers, CPU, and storage. Run a mix of virtual machines on each server so that not all experience high demand in the same area at the same time.

Metadata Updates

A VMFS datastore holds virtual machine files, directories, symbolic links, RDMs, and so on. A VMS 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 the attributes of a virtual machine file are accessed or modified when, for example, you perform one of the following operations:

- Creating, growing, or locking a virtual machine file
- Changing a file's attributes
- Powering a virtual machine on or off

Making LUN Decisions

You must plan how to set up storage for your ESX/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 applications in virtual machines that use it.
- One 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 answer when you have to decide on the LUN size and number of LUNs to use. You can experiment using either predictive or adaptive scheme.

Use the Predictive Scheme to Make LUN Decisions

When you plan how to set up your storage for your ESX/ESXi systems before you format LUNs with VMFS datastores, you must decide on the LUN size and number of LUNs to use. You can experiment using the predictive scheme.

Procedure

- 1 Create several LUNs with different storage characteristics.
- 2 Build a VMFS datastore on each LUN, labeling each datastore according to its characteristics.
- 3 Allocate virtual disks to contain the data for virtual machine applications in the VMFS datastores built on LUNs with the appropriate RAID level for the applications' requirements.
- 4 Use disk shares to distinguish high-priority from low-priority virtual machines.
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 you plan how to set up your storage for your ESX/ESXi systems before you format LUNs with VMFS datastores, you must decide on the LUN size and number of LUNs to use. You can experiment using the adaptive scheme.

Procedure

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

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

Use Disk Shares to Prioritize Virtual Machines

If multiple virtual machines access the same VMFS datastore (and therefore the same LUN), use disk shares to prioritize the disk accesses from the virtual machines. Disk shares distinguish high-priority from low-priority virtual machines.

Procedure

- 1 Start a vSphere Client and connect to vCenter Server.
- 2 Select the virtual machine in the inventory panel and click **Edit virtual machine settings** from the menu.
- 3 Click the **Resources** tab and click **Disk**.
- 4 Double-click the **Shares** column for the disk to modify and select the required value from the drop-down menu.

Shares is a value that represents the relative metric for controlling disk bandwidth to all virtual machines. The values Low, Normal, High, and Custom are compared to the sum of all shares of all virtual machines on the server and, on an ESX host, the service console. Share allocation symbolic values can be used to configure their conversion into numeric values.

- 5 Click **OK** to save your selection.

NOTE Disk shares are relevant only within a given ESX/ESXi host. The shares assigned to virtual machines on one host have no effect on virtual machines on other hosts.

Specifics of Using SAN Storage with ESX/ESXi

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

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

- You cannot directly access the virtual machine operating system that uses the storage. With traditional tools, you can monitor only the VMware ESX/ESXi operating system. You use the vSphere Client to monitor virtual machines.
- When you create a virtual machine, it is, by default, configured with one virtual hard disk and one virtual SCSI controller. You can modify the SCSI controller type and SCSI bus sharing characteristics by using the vSphere Client to edit the virtual machine settings. You can also add hard disks to your virtual machine.
- The HBA visible to the SAN administration tools is part of the ESX/ESXi system, not part of the virtual machine.
- Your ESX/ESXi system performs multipathing for you.

Third-Party Management Applications

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

Most SAN hardware is packaged with SAN management software. This software typically runs on the storage array 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 detect the SAN. This problem can be resolved by using an RDM.

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

How Virtual Machines Access Data on a SAN

ESX/ESXi stores a virtual machine's disk files within a VMFS datastore that is deployed on a SAN storage device. When virtual machine guest operating systems issue SCSI commands to their virtual disks, the 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.
 - Locates the file in the VMFS volume that corresponds to the guest virtual machine disk.
 - Maps the requests for the blocks on the virtual disk to blocks on the appropriate physical device.
 - Sends the modified I/O request from the device driver in the VMkernel to the physical HBA (host HBA).
- 5 The host HBA performs the following tasks.
 - Converts the request from its binary data form to the optical form required for transmission on the fiber optic cable.
 - Packages the request according to the rules of the FC protocol.
 - Transmits the request to the SAN.
- 6 Depending on which 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.

This storage device appears to be a specific disk to the host, but it might be a logical device that corresponds to a physical device on the SAN. The switch must determine which physical device is made available to the host for its targeted logical device.

Understanding Multipathing and Failover

To maintain a constant connection between an ESX/ESXi host and its storage, ESX/ESXi supports multipathing. Multipathing is a technique that lets you use more than one physical path that transfers data between the host and external storage device.

In case of a failure of any element in the SAN network, such as an adapter, switch, or cable, ESX/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.

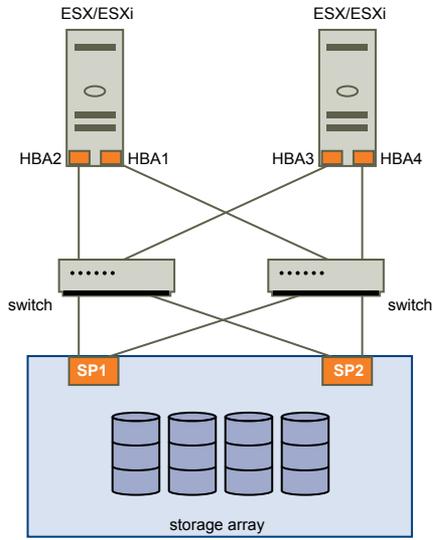
NOTE Virtual machine I/O might be delayed for up to sixty seconds while path failover takes place. These delays allow the SAN to stabilize its configuration after topology changes. In general, the I/O delays might be longer on active-passive arrays and shorter on activate-active arrays.

Host-Based 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 the one or more storage processors on the storage array device itself.

In [Figure 2-2](#), 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 2-2. Multipathing and Failover



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 ESX/ESXi supports HBA and SP failover with its multipathing capability.

Managing Multiple Paths

To manage storage multipathing, ESX/ESXi uses a special VMkernel layer, Pluggable Storage Architecture (PSA). The PSA is an open modular framework that coordinates the simultaneous operation of multiple multipathing plugins (MPPs).

The VMkernel multipathing plugin that ESX/ESXi provides by default is the VMware Native Multipathing Plugin (NMP). The NMP is an extensible module that manages subplugins. There are two types of NMP subplugins, Storage Array Type Plugins (SATPs), and Path Selection Plugins (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.

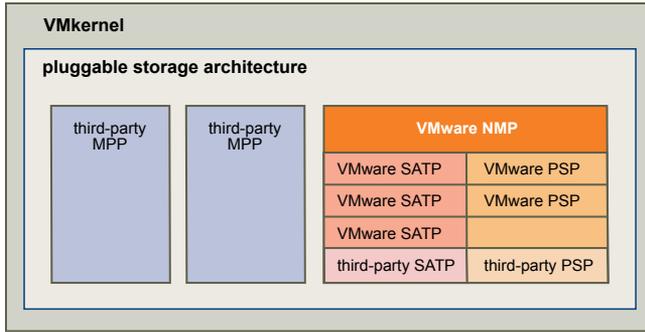
When coordinating the VMware NMP and any installed third-party MPPs, the PSA performs the following tasks:

- Loads and unloads multipathing plugins.
- Hides virtual machine specifics from a particular plugin.
- Routes I/O requests for a specific logical device to the MPP managing that device.
- Handles I/O queuing to the logical devices.
- Implements logical device bandwidth sharing between virtual machines.

- Handles I/O queuing to the physical storage HBAs.
- Handles physical path discovery and removal.
- Provides logical device and physical path I/O statistics.

As [Figure 2-3](#) illustrates, multiple third-party MPPs can run in parallel with the VMware NMP. The third-party MPPs can replace the behavior of the NMP and take complete control of the path failover and the load-balancing operations for specified storage devices.

Figure 2-3. 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.
- 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 abort or reset of logical devices.

VMware Multipathing Module

By default, ESX/ESXi provides an extensible multipathing module called the Native Multipathing Plugin (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 Plugin (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 Plugin (PSP). SATPs and PSPs are sub-plugins within the NMP module.

VMware SATPs

Storage Array Type Plugins (SATPs) run in conjunction with the VMware NMP and are responsible for array-specific operations.

ESX/ESXi offers an SATP for every type of array that VMware supports. These SATPs include an active/active SATP and active/passive SATP for non-specified storage arrays, and the local SATP for direct-attached storage. 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 can work with multiple storage arrays without having to be aware of the storage device specifics.

After the NMP determines which SATP to call 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 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 Plugins (PSPs) run in conjunction with the VMware NMP and are responsible for choosing a physical path for I/O requests.

The VMware NMP assigns a default PSP for every logical device based on the SATP associated with the physical paths for that device. You can override the default PSP.

By default, the VMware NMP supports the following PSPs:

Most Recently Used (MRU)

Selects the path the ESX/ESXi host used most recently to access the given device. If this path becomes unavailable, the host switches to an alternative path and continues to use the new path while it is available.

Fixed

Uses the designated preferred path, if it has been configured. Otherwise, it uses the first working path discovered at system boot time. If the host cannot use the preferred path, it selects a random alternative available path. The host automatically reverts back to the preferred path as soon as that path becomes available.

NOTE With active-passive arrays that have a **Fixed** path policy, path thrashing might be a problem.

Round Robin (RR)

Uses a path selection algorithm that rotates through all available paths enabling load balancing across the paths.

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 If the I/O operation is successful, the NMP reports its completion.
- 4 If the I/O operation reports an error, the NMP calls an appropriate SATP.
- 5 The SATP interprets the I/O command errors and, when appropriate, activates inactive paths.
- 6 The PSP is called to select a new path on which to issue the I/O.

Choosing Virtual Machine Locations

Storage location is an important factor when you want to optimize the performance of your virtual machines. There is always a trade-off 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 SP redundancy, and fibre drives. Uses high-cost spindles.
Mid tier	Offers mid-range performance, lower availability, some SP redundancy, and SCSI drives. Might 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 require the highest performance and most available storage, at least not throughout their entire life cycle.

If you want 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.

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 point-in-time (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.

Designing for Server Failure

The RAID architecture of SAN storage inherently protects you from failure at the physical disk level. A dual fabric, with duplication of all fabric components, protects the SAN from most fabric failures. The final step in making your whole environment failure resistant is to protect against server failure.

Using VMware HA

With VMware HA, you can organize virtual machines into failover groups. When a host fails, all its virtual machines are immediately started on different hosts. HA requires SAN storage.

When a virtual machine is restored on a different host, the virtual machine loses its memory state, but its disk state is exactly as it was when the host failed (crash-consistent failover).

NOTE You must be licensed to use VMware HA.

Using Cluster Services

Server clustering is a method of tying two or more servers together by using a high-speed network connection so that the group of servers functions as a single, logical server. If one of the servers fails, the other servers in the cluster continue operating, picking up the operations that the failed server performs.

VMware tests Microsoft Cluster Service in conjunction with ESX/ESXi systems, but other cluster solutions might also work. Different configuration options are available for achieving failover with clustering:

Cluster in a box	Two virtual machines on one host act as failover servers for each other. When one virtual machine fails, the other takes over. This configuration does not protect against host failures and is most commonly used during testing of the clustered application.
Cluster across boxes	A virtual machine on an ESX/ESXi host has a matching virtual machine on another ESX/ESXi host.
Physical to virtual clustering (N+1 clustering)	A virtual machine on an ESX/ESXi host acts as a failover server for a physical server. Because virtual machines that run on a single host can act as failover servers for numerous physical servers, this clustering method is a cost-effective N+1 solution.

Server Failover and Storage Considerations

For each type of server failover, you must consider storage issues.

- Approaches to server failover work only if each server has access to the same storage. Because multiple servers require a lot of disk space, and because failover for the storage array complements failover for the server, SANs are usually employed in conjunction with server failover.
- When you design a SAN to work in conjunction with server failover, all LUNs that are used by the clustered virtual machines must be detected by all ESX/ESXi hosts. This requirement is counterintuitive for SAN administrators, but is appropriate when using virtual machines.

Although a LUN is accessible to a host, all virtual machines on that host do not necessarily have access to all data on that LUN. A virtual machine can access only the virtual disks for which it has been configured. In case of a configuration error, virtual disks are locked when the virtual machine boots so that no corruption occurs.

NOTE As a rule, when you are using boot from a SAN LUN, only the host that is booting from a LUN should see each boot LUN. An exception is when you are trying to recover from a failure by pointing a second host to the same LUN. In this case, the SAN LUN in question is not really booting from SAN. No host is booting from the SAN LUN because it is corrupted. The SAN LUN is a nonboot LUN that is made visible to a host.

Optimizing Resource Use

VMware vSphere allows you to optimize resource allocation by migrating virtual machines from overused hosts to underused hosts.

You have the following options:

- Migrate virtual machines manually by using VMotion.
- Migrate virtual machines automatically by using VMware DRS.

You can use VMotion or DRS only if the virtual disks are located on shared storage accessible to multiple servers. In most cases, SAN storage is used.

Using VMotion to Migrate Virtual Machines

VMotion allows administrators to manually migrate virtual machines to different hosts. Administrators can migrate a running virtual machine to a different physical server connected to the same SAN without service interruption.

VMotion makes it possible to do the following tasks:

- Perform zero-downtime maintenance by moving virtual machines around so that the underlying hardware and storage can be serviced without disrupting user sessions.
- Continuously balance workloads across the datacenter to most effectively use resources in response to changing business demands.

Using VMware DRS to Migrate Virtual Machines

VMware DRS helps improve resource allocation across all hosts and resource pools.

DRS collects resource usage information for all hosts and virtual machines in a VMware cluster and gives recommendations (or migrates virtual machines) in one of two situations:

Initial placement

When you first power on a virtual machine in the cluster, DRS either places the virtual machine or makes a recommendation.

Load balancing

DRS tries to improve resource use across the cluster by performing automatic migrations of virtual machines (VMotion) or by providing recommendations for virtual machine migrations.

Requirements and Installation

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

This chapter includes the following topics:

- [“General ESX/ESXi SAN Requirements,”](#) on page 29
- [“ESX Boot from SAN Requirements,”](#) on page 31
- [“Installation and Setup Steps,”](#) on page 31

General ESX/ESXi SAN Requirements

In preparation for configuring your SAN and setting up your ESX/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 ESX/ESXi systems.
- Configure your system to have only one VMFS volume per LUN. With VMFS-3, you do not have to set accessibility.
- 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.
- VMware recommends that you use RDMS for access to any raw disk from an ESX Server 2.5 or later machine.
- For multipathing to work properly, each LUN must present the same LUN ID number to all ESX/ESXi hosts.
- Make sure the driver you use in the guest operating system specifies a large enough queue. You can set the queue depth for the physical HBA during system setup.
- 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.

Restrictions for ESX/ESXi with a SAN

This topic lists restrictions that exist when you use ESX/ESXi with a SAN.

- ESX/ESXi does not support FC connected tape devices.
- You cannot use virtual machine multipathing software to perform I/O load balancing to a single physical LUN.
- You cannot use virtual machine logical-volume manager software to mirror virtual disks. Dynamic disks on a Microsoft Windows virtual machine are an exception, but require special configuration.

Setting LUN Allocations

This topic provides some general information on how to allocate LUNs when your ESX/ESXi works in conjunction with SAN.

When you set LUN allocations, note the following points:

- | | |
|--|---|
| Storage provisioning. | To ensure that the ESX/ESXi system recognizes the LUNs at startup time, provision all LUNs to the appropriate HBAs before you connect the SAN to the ESX/ESXi system.

VMware recommends that you provision all LUNs to all ESX/ESXi HBAs at the same time. HBA failover works only if all HBAs see the same LUNs. |
| 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 ESX/ESXi hosts. This provides the greatest freedom in moving 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 ESX/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 the Storage/SAN Compatibility Guide, 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 ESX/ESXi host. |

Setting Fibre Channel HBA

This topic provides general guidelines for setting a FC HBA on your ESX/ESXi host.

During FC HBA setup, consider the following issues.

HBA Default Settings

FC HBAs work correctly with the default configuration settings. Follow the configuration guidelines given by your storage array vendor.

NOTE You should not mix FC HBAs from different vendors in a single server. Having different models of the same HBA is supported, but a single LUN can not be accessed through two different HBA types, only through the same type. Ensure that the firmware level on each HBA is the same.

Static Load Balancing Across HBAs

With both active/active and active/passive storage arrays, you can set up your host to use different paths to different LUNs so that your adapters are being used evenly. See [“Path Management and Manual, or Static, Load Balancing,”](#) on page 63.

Setting the Timeout for Failover

Set the timeout value for detecting when a path fails in the HBA driver. VMware recommends that you set the timeout to 10 seconds to ensure optimal performance.

Dedicated Adapter for Tape Drives

For best results, use a dedicated SCSI adapter for any tape drives that you are connecting to an ESX/ESXi system. FC connected tape drives are not supported. Use the Consolidated Backup proxy, as discussed in the Virtual Machine Backup Guide.

ESX Boot from SAN Requirements

When you have SAN storage configured with your ESX system, you can place the ESX boot image on one of the LUNs on the SAN. This configuration must meet specific criteria.

To enable your ESX system to boot from a SAN, your environment must meet the requirements listed in [Table 3-1](#).

Table 3-1. Boot from SAN Requirements

Requirement	Description
ESX system requirements	ESX 3.x or later is recommended. When you use the ESX 3.x system or later, RDMs are supported in conjunction with boot from SAN. For an ESX Server 2.5.x system, RDMs are not supported in conjunction with boot from SAN.
HBA requirements	<p>The HBA BIOS for your HBA FC card must be enabled and correctly configured to access the boot LUN.</p> <p>The HBA should be plugged into the lowest PCI bus and slot number. This allows the drivers to detect the HBA quickly because the drivers scan the HBAs in ascending PCI bus and slot numbers, regardless of the associated virtual machine HBA number.</p>
Boot LUN considerations	<ul style="list-style-type: none"> ■ When you boot from an active/passive storage array, the SP whose WWN is specified in the BIOS configuration of the HBA must be active. If that SP is passive, the HBA cannot support the boot process. ■ To facilitate BIOS configuration, mask each boot LUN so that only its own ESX system can see it. Each ESX system should see its own boot LUN, but not the boot LUN of any other ESX system.
SAN considerations	<ul style="list-style-type: none"> ■ 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. ■ Redundant and nonredundant configurations are supported. In the redundant case, ESX collapses the redundant paths so that only a single path to a LUN is presented to the user.
Hardware- specific considerations	If you are running an IBM eServer BladeCenter and use boot from SAN, you must disable IDE drives on the blades.

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 ESX/ESXi.

Follow these steps to configure your ESX/ESXi SAN environment.

- 1 Design your SAN if it is not already configured. Most existing SANs require only minor modification to work with ESX/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 ESX/ESXi.
- 4 Set up the HBAs for the hosts you have connected to the SAN.
- 5 Install ESX/ESXi on the hosts.
- 6 Create virtual machines.
- 7 (Optional) Set up your system for VMware HA failover or for using Microsoft Clustering Services.
- 8 Upgrade or modify your environment as needed.

Setting Up SAN Storage Devices with ESX/ESXi

4

This section discusses many of the storage devices supported in conjunction with VMware ESX/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 Guide, check with your storage array vendor, and explore the VMware knowledge base articles.

This chapter includes the following topics:

- [“Testing ESX/ESXi SAN Configurations,”](#) on page 33
- [“General Setup Considerations for Fibre Channel SAN Arrays,”](#) on page 34
- [“EMC CLARiiON Storage Systems,”](#) on page 34
- [“EMC Symmetrix Storage Systems,”](#) on page 35
- [“IBM System Storage DS4800 Storage Systems,”](#) on page 36
- [“IBM Systems Storage 8000 and IBM ESS800,”](#) on page 38
- [“HP StorageWorks Storage Systems,”](#) on page 39
- [“Hitachi Data Systems Storage,”](#) on page 41
- [“Network Appliance Storage,”](#) on page 41

Testing ESX/ESXi SAN Configurations

ESX/ESXi supports a variety of SAN storage systems in different configurations. Generally, VMware tests ESX/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 ESX/ESXi, and vendors might have specific positions of support with regard to ESX/ESXi.

Basic connectivity	Tests whether ESX/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 (with ESX only)	The ESX 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 and with only an FC cable. 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 ESX/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 ESX/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 ESX/ESXi system is appropriate.

EMC CLARiiON Storage Systems

EMC CLARiiON storage systems work with ESX/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.

ESX/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. If you prefer to use storage management software, such as EMC Navisphere, to perform manual registration, turn off the ESX/ESXi auto-registration feature. Turning it off helps you avoid overwriting the manual user registration. For information, see [“Disable Automatic Host Registration,”](#) on page 66.

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

- To avoid the possibility of path thrashing, the default multipathing policy is **Most Recently Used**, not **Fixed**. The ESX/ESXi system sets the default policy when it identifies the array.
- Automatic volume resignaturing is not supported for AX100 storage devices.
- To use boot from SAN, make sure that the active SP is chosen for the boot LUN's target in the HBA BIOS.

IMPORTANT If the EMC CLARiiON CX storage systems use the ALUA protocol, your host cannot boot from the systems or display VMFS datastores deployed on them. With ALUA enabled, these storage systems do not support SCSI-2 reservations that ESX/ESXi requires for its operations.

EMC CLARiiON AX100 and RDM

On EMC CLARiiON AX100 systems, RDMs are supported only if you use the Navisphere Management Suite for SAN administration. Navilight is not guaranteed to work properly.

To use RDMs successfully, a given LUN must be presented with the same LUN ID to every ESX/ESXi host in the cluster. By default, the AX100 does not support this configuration.

AX100 Display Problems with Inactive Connections

When you use an AX100 FC storage device directly connected to an ESX/ESXi system, you must verify that all connections are operational and unregister any connections that are no longer in use. If you do not, ESX/ESXi cannot discover new LUNs or paths.

Consider the following scenario:

An ESX/ESXi system is directly connected to an AX100 storage device. The ESX/ESXi has two FC HBAs. One of the HBAs was previously registered with the storage array and its LUNs were configured, but the connections are now inactive.

When you connect the second HBA on the ESX/ESXi host to the AX100 and register it, the ESX/ESXi host correctly shows the array as having an active connection. However, none of the LUNs that were previously configured to the ESX/ESXi host are visible, even after repeated rescans.

To resolve this issue, remove the inactive HBA, unregister the connection to the inactive HBA, or make all inactive connections active. This causes only active HBAs to be in the storage group. After this change, rescan to add the configured LUNs.

Pushing Host Configuration Changes to the Array

When you use an AX100 storage array, no host agent periodically checks the host configuration and pushes changes to the array. The `axnavi serverutil cli` utility is used to update the changes. This is a manual operation and should be performed as needed.

The utility runs only on the service console and is not available with ESXi.

EMC Symmetrix Storage Systems

EMC Symmetrix storage systems work with ESX/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 ESX/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 System Storage DS4800 Storage Systems

IBM System Storage DS4800 systems used to be called IBM FAStT. A number of storage array vendors (including LSI and StorageTek) make SAN storage arrays that are compatible with the DS4800.

For your host to work with IBM DS4800 and other compatible systems, make sure to set the multipathing policy on your host to Most Recently Used or VMW_PSP_FIXED_AP.

Configuring the Hardware for SAN Failover with DS4800 Storage Servers

This topic provides information on how to set up a highly available SAN failover configuration with an ESX/ESXi host and DS4800 storage.

You must have the following hardware components:

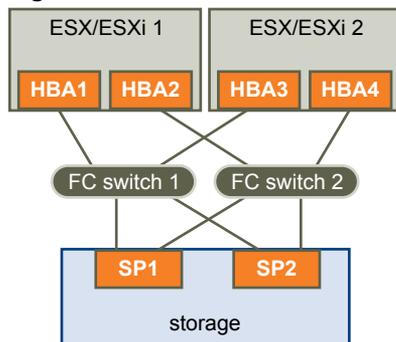
- Two FC HBAs, such as QLogic or Emulex, on each ESX/ESXi machine.
- Two FC switches connecting the HBAs to the SAN (for example, FC switch 1 and FC switch 2).
- Two SPs (for example, SP1 and SP2).

Each SP must have at least two ports connected to the SAN.

Use the following connection settings for the ESX/ESXi host, as shown in [Figure 4-1](#):

- Connect each HBA on each ESX/ESXi machine to a separate switch. For example, connect HBA1 to FC switch 1 and HBA2 to FC switch 2.
- On FC switch 1, connect SP1 to a lower switch port number than SP2, to ensure that SP1 is listed first. For example, connect SP1 to FC switch 1 port 1 and SP2 to FC switch 1 port 2.
- On FC switch 2, connect SP1 to a lower switch port number than SP2, to ensure that SP1 is listed first. For example, connect SP1 to port 1 on FC switch 2 and SP2 to port 2 on FC switch 2.

Figure 4-1. SAN Failover



This configuration provides two paths from each HBA, so that each element of the connection can fail over to a redundant path. The order of the paths in this configuration provides HBA and switch failover without the need to trigger SP failover. The storage processor that the preferred paths are connected to must own the LUNs. In the preceding example configuration, SP1 owns them.

NOTE The preceding example assumes that the switches are not connected through an Inter-Switch Link (ISL) in one fabric.

Verify the Storage Processor Port Configuration

You can verify the SP port configuration by comparing the vSphere Client information with the information in the DS4800 subsystem profile.

Procedure

- 1 Connect to the ESX/ESXi host by using the vSphere Client.
- 2 Select the host and choose the **Configuration** tab.
- 3 Click **Storage Adapters** in the Hardware panel.
- 4 Select each storage adapter to see its WWPN.
- 5 Select **Storage** to see the available datastores.

Compare the WWPN information to the information listed in the DS4800 storage subsystem profile.

Disabling Auto Volume Transfer

To avoid the possibility of path thrashing, disable Auto Volume Transfer (AVT) on the SAN storage processors. If AVT is enabled, the two storage processors can alternately take ownership of the LUN in certain situations, resulting in performance degradation. AVT is also known as ADT (Auto Disk Transfer).

To disable AVT, in the DS 4800 Storage Manager, for each port defined in each host group that contains HBAs for one or more ESX/ESXi machines, set the host type to LNXCL or, in later versions, to VMware.

You must reboot the ESX/ESXi host after you change the AVT configuration.

Configure Storage Processor Sense Data

A DS4800 SP that runs Windows as a guest operating system should return `Not Ready` sense data when it is quiescent. Returning `Unit Attention` might cause the Windows guest to fail during a failover.

Procedure

- 1 Determine the index for the LNXCL host type by using the following commands in a shell window.

Press Enter after each command.

```
SMcli.exe <ip-addr-for-SPA> show hosttopology; <Enter>
SMcli.exe <ip-addr-for-SPB> show hosttopology; <Enter>
```

The following commands assume that 13 is the index corresponding to LNXCL in the NVSRAM host type definitions. If your storage processors have LNXCL at a different index, substitute that index for 13 in the following commands.

- 2 Execute these commands for SPA to have it return `Not Ready` sense data.

Press Enter only after you enter all commands.

```
SMcli.exe <ip-addr-for-SPA>
set controller [a] HostNVSRAMBYTE [13,0x12]=0x01;
set controller [a] HostNVSRAMBYTE [13,0x13]=0x00;
reset Controller [a];
<Enter>
```

- 3 Execute these commands for SPB to have it return `Not Ready` sense data.

Press Enter only after you enter all commands.

```
SMcli.exe <ip-addr-for-SPB>
set controller [b] HostNVSRAMBYTE [13,0x12]=0x01;
set controller [b] HostNVSRAMBYTE [13,0x13]=0x00;
reset Controller [b];
<Enter>
```

NOTE If you use the DS4800 Storage Manager GUI, paste the configuration commands for both storage processors into a single script and configure both storage processors at the same time. If you use `SMcli.exe`, make individual connections to each SP.

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 ESX/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 ESX/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 ESX 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 ESX/ESXi at the HP web site.

HP StorageWorks MSA

This section lists issues of interest if you are using the active/passive version of the HP StorageWorks MSA.

Set the Profile Name to Linux

To use HP StorageWorks MSA 1000 and MSA 1500 with ESX/ESXi systems, configure the FC connections between the SAN array and the ESX/ESXi host with the Profile Name set to Linux.

Procedure

- 1 Create a static connection on the MSA 1000 by using the MSA 1000 command-line interface.

For information on installing and configuring the command-line interface, see the HP StorageWorks MSA 1000 documentation.

NOTE You cannot create connection settings by using the HP Array Configuration utility.

- 2 Connect the MSA 1000 command-line interface to the MSA 1000.
- 3 Verify that the FC network between the MSA 1000 and the ESX/ESXi host is working.
- 4 Start the command-line interface and enter the following at the prompt:

SHOW CONNECTIONS

The output displays a connection specification for each FC WWNN and WWPNN attached to the MSA 1000.

```
Connection Name: <unknown>
Host WWNN = 20:02:00:a0:b8:0c:d5:56
Host WWPNN = 20:03:00:a0:b8:0c:d5:57
Profile Name = Default
Unit Offset 0
Controller 1 Port 1 Status = Online
Controller 2 Port 1 Status = Online
```

- 5 Make sure the host's WWNN and WWPNN show the correct connection for each FC adapter on the ESX/ESXi machine.
- 6 Create a static connection.

```
ADD CONNECTION ESX_CONN_1 WWNN=20:02:00:a0:b8:0c:d5:56 WWPNN=20:03:00:a0:b8:0c:d5:57
PROFILE=LINUX
```

- 7 Verify the connection by entering the following:

SHOW CONNECTIONS

The output displays a single connection with the WWNN and WWPN pair 20:02:00:a0:b8:0c:d5:56 and 20:03:00:a0:b8:0c:d5:57 and the Profile Name set to Linux:

```
Connection Name: ESX_CONN_1
Host WWNN = 20:02:00:a0:b8:0c:d5:56
Host WWPN = 20:03:00:a0:b8:0c:d5:57
Profile Name = Linux
Unit Offset = 0
Controller 1 Port 1 Status = Online
Controller 2 Port 1 Status = Online
```

NOTE Make sure WWNN = 20:02:00:a0:b8:0c:d5:56 and WWPN = 20:03:00:a0:b8:0c:d5:57 display a single connection.

There should be no connection with the Connection Name unknown for WWNN=20:02:00:a0:b8:0c:d5:56 and WWPN =20:03:00:a0:b8:0c:d5:57.

- 8 Add static connections with different connection name values for each WWNN and WWPN on the ESX/ESXi host.

Hub Controller Issues

The ESX/ESXi system might not function correctly with the MSA hub controller. Use the 2/8 internal switch or the single port controller instead.

HP StorageWorks EVA

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

Set the connection type to Custom when you present a LUN to an ESX/ESXi 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 ESX/ESXi system.

See the VMware Infrastructure, HP StorageWorks Best Practices on 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 ESX/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 ESX/ESXi. If your microcode is not supported, interaction with ESX/ESXi is usually not possible.
Modes	<p>The modes you set depend on the model you are using, for example:</p> <ul style="list-style-type: none"> ■ 9900 and 9900v uses Netware host mode. ■ 9500v series uses Hostmode1: standard and Hostmode2: SUN Cluster. <p>Check with your HDS representative for host mode settings for the models not listed here.</p>

Network Appliance Storage

When configuring a Network Appliance 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.

Provision Storage from a Network Appliance Storage Device

You can use CLI or the FilerView GUI to provision storage on a Network Appliance storage system.

For additional information on how to use Network Appliance Storage with VMware technology, see the Network Appliance documents.

Procedure

- 1 Using CLI or the FilerView GUI, create an Aggregate if required.

```
aggr create <vmware-aggr> <number of disks>
```

- 2 Create a Flexible Volume.

```
vol create <aggregate name> <volume size>
```

- 3 Create a Qtree to store each LUN.

```
qtree create <path>
```

- 4 Create a LUN.

```
lun create -s <size> -t vmware <path>
```

- 5 Create an initiator group.

```
igroup create -f -t vmware <igroup name>
```

- 6 Map the LUN to the initiator group you just created.

```
lun map (<path>) <igroup name> <LUN ID>
```


Using Boot from SAN with ESX Systems

5

This section discusses the benefits of boot from SAN and describes the tasks you need to perform to have the ESX boot image stored on a SAN LUN.

NOTE Skip this information if you do not plan to have your ESX host boot from a SAN.

This chapter includes the following topics:

- [“Boot from SAN Overview,”](#) on page 43
- [“Getting Ready for Boot from SAN,”](#) on page 44
- [“Setting Up the QLogic FC HBA for Boot from SAN,”](#) on page 46
- [“Setting Up the Emulex FC HBA for Boot from SAN,”](#) on page 48

Boot from SAN Overview

Before you consider how to set up your system for boot from SAN, decide whether it makes sense for your environment.

Use boot from SAN in the following circumstances:

- If you do not want to handle maintenance of local storage.
- If you need easy cloning of service consoles.
- In diskless hardware configurations, such as on some blade systems.

You should not use boot from SAN if you expect I/O contention to occur between the service console and VMkernel.

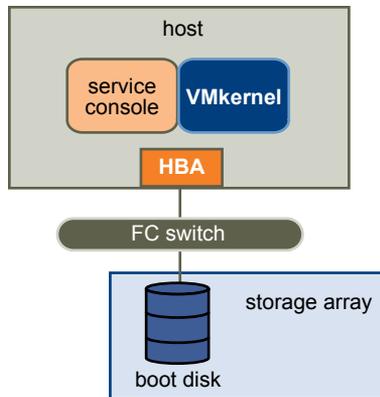
NOTE With ESX Server 2.5, you could not use boot from SAN together with RDM. With ESX 3.x or later, this restriction is removed.

How Boot from a SAN Works

When you set up your host to boot from a SAN, the boot image is not stored on the ESX host’s local disk, but instead is stored on a SAN LUN. The host is informed about the boot image location. When the host is started, it boots from the LUNs on the SAN array.

On a system set up to boot from a SAN:

- The HBA BIOS must designate the FC card as the boot controller.
- The FC card must be configured to initiate a primitive connection to the target boot LUN.

Figure 5-1. How Boot from a SAN Works

NOTE When you use boot from SAN in conjunction with ESX hosts, each host must have its own boot LUN.

Benefits of Boot from SAN

Booting your ESX host from a SAN provides numerous benefits.

The benefits include:

- 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.
- Easier backup processes – The system boot images in the SAN can be backed up as part of the overall SAN backup procedures.
- Improved management – Creating and managing the operating system image is easier and more efficient.

Getting Ready for Boot from SAN

In addition to the general ESX with SAN configuration tasks, complete the following tasks to enable your ESX host to boot from SAN.

- 1 Ensure that the configuration settings meet the basic boot from SAN requirements.
- 2 Prepare the hardware elements.

This includes your HBA, network devices, and storage system. Refer to the product documentation for each device.

- 3 Configure LUN masking on your SAN.

This ensures that each ESX host has a dedicated LUN for the boot partitions. The boot LUN must be dedicated to a single server.

- 4 Choose the location for the diagnostic partition.

Diagnostic partitions can be put on the same LUN as the boot partition. Core dumps are stored in diagnostic partitions.

IMPORTANT Your host cannot boot from the EMC CLARiiON CX storage systems that use the ALUA protocol. With ALUA enabled, these storage systems do not support SCSI-2 reservations, which ESX requires to boot from a SAN LUN.

Before You Begin

When preparing your ESX host and storage array for the boot from SAN setup, review any available information, including specific recommendations and requirements, vendor's documentation, and so on.

Review the following information:

- The recommendations or sample setups for the type of configuration you want:
 - Single or redundant paths to the boot LUN.
 - FC switch fabric.
 - Any specific recommendations that apply to the type of storage array you have.
- Restrictions and requirements, including:
 - Boot-from-SAN restrictions.
 - The vendor's recommendation for the storage array to be used for booting from a SAN.
 - The vendor's recommendation for the server booting from a SAN.
- Find the WWN for the boot path HBA by using one of the following methods:
 - Go into the FC HBA BIOS upon boot.
 - Find the WWN on the physical card. It is similar to a MAC address.

LUN Masking in Boot from SAN Mode

Proper LUN masking is critical in boot from SAN mode.

- Each server can see only its own boot LUN, not the boot LUNs of other servers.
- Multiple servers can share a diagnostic partition. You can use LUN masking to achieve this.

For information on how to mask paths to specific LUNs on your host, see [“Mask Paths,”](#) on page 83.

Prepare the SAN

This section lists the steps for preparing the SAN storage array for boot from SAN.

Procedure

- 1 Connect the FC and Ethernet cables, referring to any cabling guide that applies to your setup.
Check the FC switch wiring, if there is any.
- 2 Configure the storage array.
 - a From the SAN storage array, make the ESX host visible to the SAN. (This is often referred to as creating an object.)
 - b From the SAN storage array, set up the ESX host to have the WWPNs of the host's FC adapters as port names or node names.
 - c Create LUNs.
 - d Assign LUNs.
 - e Record the IP addresses of the FC switches and storage arrays.
 - f Record the WWPN for each SP and host adapter involved.



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

- 3 Configure the HBA BIOS for boot from SAN.
- 4 Boot your ESX system from the ESX installation CD.

The QLogic BIOS uses a search list of paths (`wwpn:lun`) to locate a boot image. If one of the `wwpn:lun` paths is associated with a passive path, for example, when you use CLARiiON or IBM TotalStorage DS 4000 systems, the BIOS stays with the passive path and does not locate an active path. If you are booting your ESX system from a SAN LUN, the boot fails while the host tries to access the passive path.

Minimizing the Number of Initiators

Be sure the zone contains the minimum number of host and storage ports possible. The Emulex and QLogic BIOS can become unresponsive if several other initiators are in the same zone and you try to select a boot LUN.

For example, if fifteen initiators and four Symmetrix ports are in one zone, you might not be able to select a boot device from either the Emulex or QLogic BIOS because it becomes unresponsive. If you zone the two host ports to see only the four storage ports, you can select a boot LUN.

Set Up Your System to Boot from CD-ROM First

Because the VMware installation CD is in the CD-ROM drive, set up your system to boot from CD-ROM first. To achieve this, change the system boot sequence in your system BIOS setup.

For example, you need to complete the following on the IBM X-Series 345 server.

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]**.

You can now install the ESX system.

Setting Up the QLogic FC HBA for Boot from SAN

Configuring the QLogic HBA BIOS to boot ESX from a SAN includes enabling the QLogic HBA BIOS, enabling the selectable Boot, and selecting the Boot LUN.

NOTE If you are using an IBM BladeCenter, disconnect all your local disk drives from the server.

Enable the QLogic HBA BIOS

When configuring the QLogic HBA BIOS to boot ESX from SAN, start with enabling the QLogic HBA BIOS.

Procedure

- 1 Enter the BIOS Fast!UTIL configuration utility.
 - a Boot the server.
 - b While booting the server, press Ctrl+Q.
- 2 Perform the appropriate action depending on the number of HBAs.

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

- 3 In the Fast!UTIL Options page, select **Configuration Settings** and press Enter.
- 4 In the Configuration Settings page, select **Host 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.

Enable the Selectable Boot

You need to enable the selectable boot.

Procedure

- 1 Choose **Selectable Boot Settings** and press Enter.
- 2 In the Selectable Boot Settings page, choose **Selectable Boot**.
- 3 Press Enter to toggle the value to **Enabled**.

Select the Boot LUN

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.

Procedure

- 1 Use the cursor keys to select the first entry in the list of storage processors.
- 2 Press Enter to open the Select Fibre Channel Device page.

- 3 Use the cursor keys to select the chosen SP and press Enter.
 - If the SP has only one LUN attached, it is selected as the boot LUN, and you can skip to [Step 4](#).
 - If the SP has more than one LUN attached, the Select LUN page opens. Use the arrow keys to position to the selected LUN and press Enter.

If any remaining storage processors show in the list, position to those entries and press C to clear the data.
- 4 Press Esc twice to exit.
- 5 Press Enter to save the setting.

Setting Up the Emulex FC HBA for Boot from SAN

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

Enable the BootBIOS Prompt

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

Procedure

- 1 From the ESX service console or a Linux command prompt, run `lputil`.

NOTE Consider booting the ESX host from a Linux Administration CD that loads the Emulex driver, then run `lputil` from there.

- 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 ESX from SAN, you need to enable BIOS.

Procedure

- 1 Reboot the ESX machine.
- 2 Press ALT+E at the Emulex prompt.
 - a Select an adapter (with BIOS support).
 - b Select <2> **Configure Adapter's Parameters**.
 - c Select <1> **Enable or Disable BIOS**.
 - d Select <1> to enable BIOS.
 - e Select <x> to exit and <N> to return to the main menu.

- 3 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.

Managing ESX/ESXi Systems That Use SAN Storage

6

This section helps you manage your ESX/ESXi system, use SAN storage effectively, and perform troubleshooting. It also explains how to find information about storage devices, adapters, multipathing, and so on.

This chapter includes the following topics:

- [“Viewing Storage Adapter Information,”](#) on page 51
- [“Viewing Storage Device Information,”](#) on page 52
- [“Viewing Datastore Information,”](#) on page 54
- [“Resolving Display Issues,”](#) on page 54
- [“N-Port ID Virtualization,”](#) on page 57
- [“Path Scanning and Claiming,”](#) on page 60
- [“Path Management and Manual, or Static, Load Balancing,”](#) on page 63
- [“Path Failover,”](#) on page 64
- [“Set Device Driver Options for SCSI Controllers,”](#) on page 65
- [“Sharing Diagnostic Partitions,”](#) on page 65
- [“Disable Automatic Host Registration,”](#) on page 66
- [“Avoiding and Resolving SAN Problems,”](#) on page 66
- [“Optimizing SAN Storage Performance,”](#) on page 67
- [“Resolving Performance Issues,”](#) on page 68
- [“SAN Storage Backup Considerations,”](#) on page 71
- [“Layered Applications,”](#) on page 72
- [“Managing Duplicate VMFS Datastores,”](#) on page 73

Viewing Storage Adapter Information

In the vSphere Client, you can display storage adapters that your host uses and review their information.

When you list all available adapters, you can see their models, types, such as Fibre Channel, Parallel SCSI, or iSCSI, and, if available, their unique identifiers.

As unique identifiers, Fibre Channel HBAs use World Wide Names (WWNs).

When you display details for each Fibre Channel HBA, you see the following information.

Table 6-1. Storage Adapter Information

Adapter Information	Description
Model	Model of the adapter.
Targets	Number of targets accessed through the adapter.
WWN	A World Wide Name formed according to Fibre Channel standards that uniquely identifies the FC adapter.
Devices	All storage devices or LUNs the adapter can access.
Paths	All paths the adapter uses to access storage devices.

View Storage Adapter Information

Use the vSphere Client to display storage adapters and review their information.

Procedure

- 1 In Inventory, select **Hosts and Clusters**.
- 2 Select a host and click the **Configuration** tab.
- 3 In Hardware, select **Storage Adapters**.
- 4 To view details for a specific adapter, select the adapter from the Storage Adapters list.
- 5 To list all storage devices the adapter can access, click **Devices**.
- 6 To list all paths the adapter uses, click **Paths**.

Viewing Storage Device Information

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

For each storage adapter, you can display a separate list of storage devices accessible just through this adapter. When you review a list of storage devices, you typically see the following information.

Table 6-2. Storage Device Information

Device Information	Description
Name	A friendly name that the host assigns to the device based on the storage type and manufacturer.
Identifier	A universally unique identifier that is intrinsic to the storage device.
Runtime Name	The name of the first path to the device.
LUN	The LUN number that shows the position of the LUN within the target.
Type	Type of device, for example, disk or CD-ROM.
Transport	Transportation protocol your host uses to access the device.
Capacity	Total capacity of the storage device.
Owner	The plugin, such as the NMP or a third-party plugin, the host uses to manage the storage device.

Details for each storage device include the following:

- A path to the storage device in the `/vmfs/devices/` directory.
- Primary and logical partitions, including a VMFS datastore, if configured.

Understanding Storage Device Naming

In the vSphere Client, each storage device, or LUN, is identified by several names.

Name	A friendly name that the host assigns to a device based on the storage type and manufacturer. You can modify the name using the vSphere Client.
Identifier	A universally unique identifier that the host extracts from the storage. Depending on the type of storage, the host uses different algorithms to extract the identifier. The identifier is persistent across reboots and is the same for all hosts sharing the device.
Runtime Name	<p>The name of the first path to the device. The runtime name is created by the host. The name is not a reliable identifier for the device, and is not persistent.</p> <p>The runtime name has the following format:</p> <p>vmhba#:C#:T#:L#, where</p> <ul style="list-style-type: none"> ■ vmhba# 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. ■ C# is the storage channel number. ■ T# is the target number. Target numbering is decided by the host and might change if there is a change in the mappings of targets visible to the host. Targets that are shared by different hosts might not have the same target number. ■ L# 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.

Display Storage Devices for a Host

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

Procedure

- 1 In Inventory, select **Hosts and Clusters**.
- 2 Select a host and click the **Configuration** tab.
- 3 In Hardware, select **Storage**.
- 4 Click **Devices**.
- 5 To view additional details about a specific device, select the device from the list.

Display Storage Devices for an Adapter

For each storage adapter on your host, you can display a list of storage devices accessible just through this adapter.

Procedure

- 1 In Inventory, select **Hosts and Clusters**.
- 2 Select a host and click the **Configuration** tab.
- 3 In Hardware, select **Storage Adapters**.
- 4 Select the adapter from the Storage Adapters list.
- 5 Click **Devices**.

Viewing Datastore Information

You can view a list of available datastores and analyze their properties.

The Datastores pane shows summary information about the datastore.

- Target storage device where the datastore is located.
- File system the datastore uses.
- Total capacity and available space.

For each datastore, you can also review the following details:

- Location of the datastore.
- Total capacity, including the used and available space.
- Individual extents that the datastore spans and their capacity. To view extent details, click **Properties** and select the **Extents** panel.
- Paths used to access the storage device.

Review Datastore Properties

Use the vSphere Client to review datastore properties.

Procedure

- 1 Select a host in the inventory and click the **Configuration** tab.
- 2 In Hardware, select **Storage**.
- 3 Click the **Datastores** view.
- 4 To display details for a particular datastore, select the datastore from the list.

Resolving Display Issues

When you use the vSphere Client to view storage devices available to your ESX/ESXi host and the output differs from what you expect, perform troubleshooting tasks.

Perform the following troubleshooting tasks if you have display issues.

Table 6-3. Troubleshooting Fibre Channel LUN Display

Troubleshooting Task	Description
Check cable connectivity.	If you do not see a port, the problem could be cable connectivity. Check the cables first. Ensure that cables are connected to the ports and a link light indicates that the connection is good. If each end of the cable does not show a good link light, replace the cable.
Check zoning.	Zoning limits access to specific storage devices, increases security, and decreases traffic over the network. Some storage vendors allow only single-initiator zones. In that case, an HBA can be in multiple zones to only one target. Other vendors allow multiple-initiator zones. See your storage vendor's documentation for zoning requirements. Use the SAN switch software to configure and manage zoning.
Check access control configuration.	<ul style="list-style-type: none"> ■ The MASK_PATH plugin allows you to prevent your host from seeing a specific storage array or specific LUNs on a storage array. If your host does not see the expected LUNs on the array, path masking could have been set up incorrectly. ■ For booting from a SAN, ensure that each ESX host sees only required LUNs. Do not allow any ESX host to see any boot LUN other than its own. Use storage system software to make sure that the ESX host can see only the LUNs that it is supposed to see. ■ Ensure that the Disk.MaxLUN setting allows you to view the LUN you expect to see.
Check storage processor setup.	If a disk array has more than one storage processor (SP), make sure that the SAN switch has a connection to the SP that owns the LUNs you want to access. On some disk arrays, only one SP is active and the other SP is passive until there is a failure. If you are connected to the wrong SP (the one with the passive path), you might see the LUNs but get errors when trying to access them.
Rescan your HBA.	Perform a rescan each time you complete the following tasks: <ul style="list-style-type: none"> ■ Create new LUNs on a SAN. ■ Change the path masking configuration on an ESX/ESXi host storage system. ■ Reconnect a cable. ■ Make a change to a host in a cluster.

Storage Refresh and Rescan Operations

The refresh operation updates the datastore lists and storage information, such as the datastore capacity, displayed in the vSphere Client. When you make changes in your ESX/ESXi host or SAN configuration, you need to use the rescan operation.

You can rescan all adapters on your host. If the changes you make are isolated to a specific adapter, rescan only this adapter. If your vSphere Client is connected to a vCenter Server system, you can rescan adapters on all hosts managed by the vCenter Server system.

Perform a rescan each time you make one of the following changes.

- Zone a new disk array on the SAN to an ESX/ESXi host.
- Create new LUNs on a SAN.
- Change the path masking on a host.
- Reconnect a cable.
- Make a change to a host in a cluster.

IMPORTANT Do not rescan when a path is unavailable. If one path fails, another takes over and your system continues to be fully functional. If, however, you rescan at a time when a path is not available, the host removes the path from its list of paths to the device. The path cannot be used by the host until the next time a rescan is performed while the path is active.

Rescan Storage Adapters

When you make changes in your ESX/ESXi host or SAN configuration, you might need to rescan your storage adapters. You can rescan all adapters on your host. If the changes you make are isolated to a specific adapter, rescan only this adapter.

Use this procedure if you want to limit the rescan to a particular host or an adapter on the host. If you want to rescan adapters on all hosts managed by your vCenter Server system, you can do so by right-clicking a datacenter, cluster, or folder that contains the hosts and selecting **Rescan for Datastores**.

Procedure

- 1 In the vSphere Client, select a host and click the **Configuration** tab.
- 2 In the Hardware panel, select **Storage Adapters**, and click **Rescan** above the Storage Adapters panel.

You can also right-click an individual adapter and click **Rescan** to rescan just that adapter.

IMPORTANT On ESXi, it is not possible to rescan a single storage adapter. If you rescan a single adapter, all adapters are rescanned.

- 3 To discover new disks or LUNs, select **Scan for New Storage Devices**.
If new LUNs are discovered, they appear in the device list.
- 4 To discover new datastores or update a datastore after its configuration has been changed, select **Scan for New VMFS Volumes**.

If new datastores or VMFS volumes are discovered, they appear in the datastore list.

Change the Number of Scanned LUNs

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 In the vSphere Client inventory panel, select the host, click the **Configuration** tab, and click **Advanced Settings**.
- 2 Select **Disk**.
- 3 Scroll down to **Disk.MaxLUN**.
- 4 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.

Disable Sparse LUN Support

You can disable the default sparse LUN support to decrease the time ESX/ESXi needs to scan for LUNs.

The VMkernel provides sparse LUN support by default. The sparse LUN support enables the VMkernel to perform uninterrupted LUN scanning when a storage system presents LUNs with nonsequential LUN numbering, for example 0, 6, and 23. If all LUNs that your storage system presents are sequential, you can disable the sparse LUN support.

Procedure

- 1 In the vSphere Client inventory panel, select the host, click the **Configuration** tab, and click **Advanced Settings**.
- 2 In the Advanced Settings dialog box, select **Disk**.
- 3 Scroll down to **Disk.SupportSparseLUN**, change the value to 0, and click **OK**.

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 ESX/ESXi to another, the VPORT is closed on the first host and opened on the destination host.

If NPIV is enabled, four 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 up to the maximum of four. 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.

NOTE If a host has four physical HBAs as paths to the storage, all physical paths must be zoned to the virtual machine by the SAN administrator. This is required to support multipathing even though only one path at a time will be active.

Requirements for Using NPIV

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

The following requirements and limitations exist:

- NPIV can only be used for virtual machines with RDM disks. Virtual machines with regular virtual disks use the WWNs of the host's physical HBAs.
- For this implementation of NPIV, the physical HBAs on an ESX/ESXi host, using their own WWNs, must have access to all LUNs that are to be accessed by virtual machines running on that host.
- The ESX/ESXi host's physical HBAs must support NPIV. Currently, the following vendors and types of HBA provide this support:
 - QLogic – any 4GB HBA.
 - Emulex – 4GB HBAs that have NPIV-compatible firmware.
- Only four WWN pairs are generated per virtual machine.
- When a virtual machine or template with a WWN assigned to it is cloned, the clones do not retain the WWN.
- The switches used must be NPIV-aware.
- When configuring an NPIV LUN for access at the storage level, make sure that the NPIV LUN number and NPIV target ID match the physical LUN and Target ID.
- Always use the vSphere Client to manipulate virtual machines with WWNs.



CAUTION 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

You can assign a WWN to a new virtual machine with an RDM disk when you create this virtual machine, or to an existing virtual machine you can temporarily power off.

Procedure

- 1 From the vSphere Client, click **Inventory** in the navigation bar, and expand the inventory as needed.
- 2 In the inventory list, select the managed host to which you want to add a new virtual machine.
- 3 Select **File > New > Virtual Machine**.

The New Virtual Machine wizard appears.

- 4 Select **Custom**, and click **Next**.
- 5 Type a virtual machine name, and click **Next**.
- 6 Select a folder or the root of a datacenter, and click **Next**.
- 7 If the resource pool option is available, expand the tree until you locate the resource pool in which you want to run the virtual machine, highlight it, and click **Next**.
- 8 Select a datastore in which to store the virtual machine files, and click **Next**.
- 9 Under **Guest operating system**, select the operating system family (Microsoft Windows, Linux, Novell NetWare, Solaris, or Other).
- 10 Select the version from the pull-down menu, and click **Next**.

- 11 Select the number of virtual processors in the virtual machine from the pull-down list, and click **Next**.
- 12 Configure the virtual machine's memory size by selecting the number of megabytes, and click **Next**.
- 13 Configure network connections, and click **Next**.
- 14 Choose the type of SCSI adapter you want to use with the virtual machine.
- 15 Select **Raw Device Mapping**, and click **Next**.
- 16 From a list of SAN disks or LUNs, select a raw LUN you want your virtual machine to access directly.
- 17 Select a datastore for the RDM mapping file.

You can place the RDM file on the same datastore where your virtual machine files reside, or select a different datastore.

NOTE If you want to use VMotion for a virtual machine with enabled NPIV, make sure that the RDM file is located on the same datastore where the virtual machine configuration file resides. You cannot perform Storage VMotion, or VMotion between datastores, when NPIV is enabled.

- 18 Select a compatibility mode, either physical or virtual.
Depending on your choice, subsequent screens offer different options.
 - Physical compatibility mode allows the guest operating system to access the hardware directly. Physical compatibility is useful if you are using SAN-aware applications in the virtual machine. However, a virtual machine with the physical compatibility RDM cannot be cloned, made into a template, or migrated if the migration involves copying the disk.
 - Virtual compatibility allows the RDM to behave as if it were a virtual disk, so you can use such features as snapshotting, cloning, and so on.
- 19 On the Specify Advanced Options page, you can change the virtual device node and click **Next**.
- 20 Assign WWNs to the virtual machine.
- 21 On the Ready to Complete New Virtual Machine page, select the **Edit the virtual machine settings before completion** check box and click **Next**.

Assign or Modify WWNs

After you create a virtual machine with an RDM, you can assign virtual WWNs to it. You can also modify WWN assignments for an existing virtual machine with an RDM.

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 ESX/ESXi host to access it.

Procedure

- 1 Open the Virtual Machine Properties dialog box.

Option	Action
New virtual machine	For a new virtual machine, after creating the virtual machine, on the Ready to Complete New Virtual Machine page select the Edit the virtual machine settings before submitting the creation task checkbox, and click Continue .
Existing virtual machine	For an existing virtual machine, select the virtual machine from the inventory panel, and click the Edit Settings link.

- 2 Select the **Options** tab.
- 3 Select **Fibre Channel NPIV**.
- 4 In the dialog box that opens, select one of the following options:

Option	Description
Leave unchanged	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.
Generate new WWNs	New WWNs are generated and assigned to the virtual machine, overwriting any existing WWNs (those of the HBA itself are unaffected).
Remove WWN assignment	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. CAUTION Removing or changing a virtual machine's existing WWN assignments causes it to lose connectivity to the storage LUNs

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

Path Scanning and Claiming

When you start your ESX/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 defined in the `/etc/vmware/esx.conf` file, the host determines which multipathing plugin (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 plugins or the native multipathing plugin (NMP).

For the paths managed by the NMP module, a second set of claim rules is applied. These rules determine which SATP should be used to manage the paths from a specific array type, and which PSP is to be used for each storage device. For example, for a storage device that belongs to the EMC CLARiiON CX storage family, the default SATP is `VMW_SATP_CX` and the default PSP is Most Recently Used.

Use the vSphere 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 vSphere 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 [Appendix B, “Managing Storage Paths and Multipathing Plugins,”](#) on page 79.

For detailed descriptions of the commands available to manage PSA, see the *vSphere Command-Line Interface Installation and Reference Guide*.

Viewing the Paths Information

Use the vSphere Client to determine which SATP and PSP the ESX/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 path selection policy (PSP), and a list of paths with their physical characteristics, such as an adapter and target each path uses, 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).

NOTE For hosts that run ESX/ESXi 3.5 or earlier, the term active means the only path that the host is using to issue I/O to a LUN.

Standby The path is operational and can be used for I/O if active paths fail.

Disabled The path is disabled and no data can be transferred.

Broken 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.

View Datastore Paths

Use the vSphere Client to review the paths that connect to storage devices the datastores are deployed on.

Procedure

- 1 Log in to the vSphere Client and select a server from the inventory panel.
- 2 Click the **Configuration** tab and click **Storage** in the Hardware panel.
- 3 Click **Datastores** under View.
- 4 From the list of configured datastores, select the datastore whose paths you want to view or configure. The Details panel shows the total number of paths being used to access the device and whether any of them are broken or disabled.
- 5 Click **Properties > Manage Paths** to open the Manage Paths dialog box.

You can use the Manage Paths dialog box to enable or disable your paths, set multipathing policy, and specify the preferred path.

View Storage Device Paths

Use the vSphere Client to view which SATP and PSP the host uses for a specific storage device and the status of all available paths for this storage device.

Procedure

- 1 Log in to the vSphere Client and select a server from the inventory panel.
- 2 Click the **Configuration** tab and click **Storage** in the Hardware panel.
- 3 Click **Devices** under View.
- 4 Click **Manage Paths** to open the Manage Paths dialog box.

Setting a Path Selection Policy

For each storage device, the ESX/ESXi host sets the path selection policy based on the claim rules defined in the `/etc/vmware/esx.conf` file.

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 always uses the preferred path to the disk when that path is available. If the host cannot access the disk through the preferred path, it tries the alternative paths. The default policy for active-active storage devices is Fixed.

Most Recently Used (VMware) The host uses a path to the disk until the path becomes unavailable. When the path becomes unavailable, the host selects one of the alternative paths. 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 active-passive storage devices and is required for those devices.

Round Robin (VMware) The host uses an automatic path selection algorithm rotating through all available paths. This implements load balancing across all the available physical paths.

Load balancing is the process of spreading server I/O requests across all available host paths. The goal is to optimize performance in terms of throughput (I/O per second, megabytes per second, or response times).

[Table 6-4](#) summarizes how the behavior of host changes, depending on the type of array and the failover policy.

Table 6-4. Path Policy Effects

Policy/Controller	Active-Active	Active-Passive
Most Recently Used	Administrator action is required to fail back after path failure.	Administrator action is required to fail back after path failure.
Fixed	VMkernel resumes using the preferred path when connectivity is restored.	VMkernel attempts to resume using the preferred path. This can cause path thrashing or failure when another SP now owns the LUN.
Round Robin	No fail back.	Next path in round robin scheduling is selected.

Change the Path Selection Policy

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 Open the Manage Paths dialog box either from the Datastores or Devices view.
- 2 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)**

- 3 For the fixed policy, specify the preferred path by right-clicking the path you want to assign as the preferred path, and selecting **Preferred**.
- 4 Click **OK** to save your settings and exit the dialog box.

Disable Paths

You can temporarily disable paths for maintenance or other reasons. You can do so using the vSphere Client.

Procedure

- 1 Open the Manage Paths dialog box either from the Datastores or Devices view.
- 2 In the Paths panel, right-click the path to disable, and select **Disable**.
- 3 Click **OK** to save your settings and exit the dialog box.

You can also disable a path from the adapter's Paths view by right-clicking the path in the list and selecting **Disable**.

Path Management and Manual, or Static, Load Balancing

Balancing loads among available paths improves performance. With both active/active and active/passive storage arrays, you can set up your host to use different paths to different LUNs so that your adapters are being used evenly.

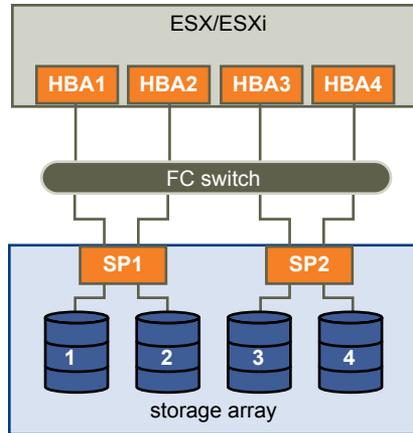
If a path fails, the surviving paths carry all the traffic. Path failover might take a minute or more, because the SAN might converge with a new topology to try to restore service. This delay is necessary to allow the SAN to stabilize its configuration after topology changes.

With active/active storage arrays, you can configure your ESX/ESXi host to load balance traffic across multiple adapters by assigning preferred paths to your LUNs. Path policy must be set to Fixed.

The following example demonstrates how manual load balancing is performed with an active/active array.

Assume the following setup, shown in [Figure 6-1](#).

- Active/Active SPs
- An ESX/ESXi system
- Four Fibre Channel HBAs in each server
- Director class software

Figure 6-1. Manual Load Balancing with Fibre Channel

For load balancing, set the preferred paths as follows. Load balancing can be performed with as few as two HBAs, although this example uses four.

- For LUN 1: HBA1-SP1-LUN1
- For LUN 2: HBA2-SP1-LUN2
- For LUN 3: HBA3-SP2-LUN3
- For LUN 4: HBA4-SP2-LUN4

With active/passive arrays, you can perform load balancing if the array supports two active paths and the HBA ports can access both SPs in an array.

NOTE Active/passive arrays use the MRU path policy which does not have a preferred path. If a path failure occurs, there is no failback. As a result, static load balancing can become out of balance over time.

Path Failover

Path failover refers to situations when the active path to a LUN is changed from one path to another, usually because of some SAN component failure along the current path. A server usually has one or two HBAs and each HBA is connected to one or two storage processors on a given SAN array. You can determine the active path, the path currently used by the server, by looking at the LUN's properties.

When an FC cable is pulled, I/O might pause for 30-60 seconds until the FC driver determines that the link is unavailable and failover has occurred. As a result, the virtual machines, with their virtual disks installed on SAN storage, can appear unresponsive. If you attempt to display the host, its storage devices, or its adapter, the operation might appear to stall. After failover is complete, I/O resumes normally.

In case of disastrous events that include multiple breakages, all connections to SAN storage devices might be lost. If none of the connections to the storage device is working, some virtual machines might encounter I/O errors on their virtual SCSI disks.

Set Operating System Timeout

You might want to increase the standard disk timeout value so that a Windows guest operating system is not extensively disrupted during failover.

For Windows 2000 and Windows Server 2003 guest operating systems, you can set operating system timeout using the registry.

Prerequisites

Before you begin, back up your Windows registry.

Procedure

- 1 Select **Start > Run**.
- 2 In the command window, type **regedit.exe**, and click **OK**.
- 3 In the left panel hierarchy view, double-click first **HKEY_LOCAL_MACHINE**, then **System**, then **CurrentControlSet**, then **Services**, and then **Disk**.
- 4 Select the **TimeOutValue** and set the data value to **x03c** (hexadecimal) or **60** (decimal).

After you've made this change, Windows waits at least 60 seconds for delayed disk operations to complete before it generates errors.

- 5 Click **OK** to exit the Registry Editor.

Set Device Driver Options for SCSI Controllers

This section describes how to set device driver options for QLogic, Emulex, or other SCSI card drivers.

Procedure

- 1 Back up the file `/etc/vmware/esx.conf`, and open it for editing.

The file includes a section for each SCSI device, as in the following example.

```
/device/002:02.0/class = "0c0400"
/device/002:02.0/devID = "2312"
/device/002:02.0/irq = "19"
/device/002:02.0/name = "QLogic Corp QLA231x/2340 (rev 02)"
/device/002:02.0/options = ""
/device/002:02.0/owner = "vmkernel"
/device/002:02.0/subsysDevID = "027d"
/device/002:02.0/subsysVendor = "1014"
/device/002:02.0/vendor = "1077"
/device/002:02.0/vmname = "vmhba0"
```

- 2 Find the options line right under the name line and modify it as appropriate.
- 3 Repeat for every SCSI adapter that is controlled by the same driver if needed.

Sharing Diagnostic Partitions

Generally, you use the local disc of your ESX/ESXi host as a diagnostic partition. If you have diskless ESX servers that boot from a SAN, multiple hosts can share one diagnostic partition on the same SAN LUN.

If more than one ESX/ESXi system uses the same LUN as the diagnostic partition, that LUN must be zoned so that all the servers can access it.

Each server needs 100MB of space, so the size of the LUN determines how many servers can share it. Each ESX/ESXi system is mapped to a diagnostic slot. VMware recommends at least 16 slots (1600MB) of disk space if servers share a diagnostic partition.

If there is only one diagnostic slot on the device, all ESX/ESXi systems sharing that device map to the same slot. This setup can easily create problems. If two ESX/ESXi systems perform a core dump at the same time, the core dumps are overwritten on the last slot on the diagnostic partition.

If you allocate enough disk space for 16 slots, it is unlikely that core dumps are mapped to the same location on the diagnostic partition, even if two ESX/ESXi systems perform a core dump at the same time.

Disable Automatic Host Registration

When you use EMC CLARiiON or Invista arrays for storage, the hosts must register with the arrays. ESX/ESXi performs automatic host registration by sending the host's name and IP address to the array. If you prefer to perform manual registration by using storage management software, turn off the ESX/ESXi auto-registration feature.

Procedure

- 1 In the vSphere Client, select the host in the inventory panel.
- 2 Click the **Configuration** tab and select **Advanced Settings** under Software.
- 3 Select **Disk** in the left panel and scroll down to `Disk.EnableNaviReg` on the right.
- 4 Change the default value to 0.
- 5 Reboot the host for the changes to take effect.

This action disables the automatic host registration enabled by default.

Avoiding and Resolving SAN Problems

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

You should observe these tips for avoiding and resolving 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. In particular, working with an active-passive array and setting the path policy to **Fixed** can lead to path thrashing.
- 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 ESX/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 ESX/ESXi 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 ESX/ESXi 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 [“Rescan Storage Adapters,”](#) on page 56.

Optimizing 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.

When assigning LUNs, remember that each LUN is accessed by a number of ESX/ESXi hosts, and that a number of virtual machines can run on each host. One LUN used by an ESX/ESXi host can service I/O from many different applications running on different operating systems. Because of this diverse workload, the RAID group containing the ESX/ESXi LUNs should not include LUNs used by other hosts that are not running ESX/ESXi for I/O intensive applications.

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 manually 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 ESX/ESXi.

Server Performance

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 ESX/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 ESX/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 ESX/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 ESX/ESXi LUNs.
- When using multiple ESX/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 ESX/ESXi system should match the number of I/Os the HBA and storage array can handle.

Resolving Performance Issues

The vSphere Client offers extensive facilities for collecting performance information. The information is graphically displayed in the vSphere Client. The vSphere Client updates its display periodically.

You can also use the `resxtop` vSphere CLI command that allows you to examine how ESX/ESXi hosts use resources. For information about `resxtop`, see the *Resource Management Guide* or *vSphere Command-Line Interface Installation and Reference Guide*.

Resolving Path Thrashing

If your server is unable to access a LUN, or access is very slow, you might have a problem with path thrashing (also called LUN thrashing). Path thrashing might occur when two hosts access the LUN through different SPs and, as a result, the LUN is never actually available.

Only specific SAN configurations in conjunction with the following conditions can cause the path thrashing:

- You are working with an active-passive array. Path thrashing only occurs on active-passive arrays. For active-active arrays or arrays that provide transparent failover, path thrashing does not occur.
- Two hosts access the same LUN using different storage processors (SPs). For example, the LUN is configured to use the Fixed PSP. On Host A, the preferred path to the LUN is set to use a path through SP A. On Host B, the preferred path to the LUN is configured to use a path through SP B.

Path thrashing can also occur if the LUN is configured to use either the Fixed PSP or the MRU PSP and Host A can access the LUN only with paths through SP A, while Host B can access the LUN only with paths through SP B.

This problem can also occur on a direct connect array (such as AX100) with HBA failover on one or more nodes.

Path thrashing is a problem that you typically do not experience with other operating systems:

- No other common operating system uses shared LUNs for more than two servers. That setup is typically reserved for clustering.
- If only one server is issuing I/Os to the LUN at a time, path thrashing does not become a problem.

In contrast, multiple ESX/ESXi systems might issue I/O to the same LUN concurrently.

Resolve Path Thrashing

Use this procedure to resolve path thrashing. Path thrashing occurs on active-passive arrays when two hosts access the LUN through different SPs and, as a result, the LUN is never actually available.

Procedure

- 1 Ensure that all hosts sharing the same set of LUNs on the active-passive arrays use the same storage processor.
- 2 Correct any cabling inconsistencies between different ESX/ESXi hosts and SAN targets so that all HBAs see the same targets in the same order.
- 3 Configure the path to use the Most Recently Used PSP (the default).

Understanding Path Thrashing

The SPs in a storage array are like independent computers that have access to some shared storage. Algorithms determine how concurrent access is handled.

For active/passive arrays, all the sectors on the storage that make up a given LUN can be accessed by only one SP at a time. The LUN ownership is passed around between the storage processors. The reason is that storage arrays use caches and SP A must not write anything to disk that invalidates the SP B cache. Because the SP has to flush the cache when it finishes the operation, it takes a little time to move the ownership. During that time, no I/O to the LUN can be processed by either SP.

Some active/passive arrays attempt to look like active/active arrays by passing the ownership of the LUN to the various SPs as I/O arrives. This approach works in a clustering setup, but if many ESX/ESXi systems access the same LUN concurrently through different SPs, the result is path thrashing.

Consider how path selection works:

- On an active/active array the ESX/ESXi system starts sending I/O down the new path.
- On an active/passive arrays, the ESX/ESXi system checks all standby paths. The SP of the path that is currently under consideration sends information to the system on whether it currently owns the LUN.
 - If the ESX/ESXi system finds an SP that owns the LUN, that path is selected and I/O is sent down that path.
 - If the ESX/ESXi host cannot find such a path, the ESX/ESXi host picks one of the standby paths and sends the SP of that path a command to move the LUN ownership to the SP.

Path thrashing can occur as a result of the following path choice: If server A can reach a LUN only through one SP, and server B can reach the same LUN only through a different SP, they both continually cause the ownership of the LUN to move between the two SPs, effectively ping-ponging the ownership of the LUN. Because the system moves the ownership quickly, the storage array cannot process any I/O (or can process only very little). As a result, any servers that depend on the LUN will experience low throughput due to the long time it takes to complete each I/O request.

Equalize Disk Access Between Virtual Machines

You can adjust the maximum number of outstanding disk requests with the `Disk.SchedNumReqOutstanding` parameter in the vSphere Client. When two or more virtual machines are accessing the same LUN, this parameter controls the number of outstanding requests that each virtual machine can issue to the LUN. Adjusting the limit can help equalize disk access between virtual machines.

This limit does not apply when only one virtual machine is active on a LUN. In that case, the bandwidth is limited by the queue depth of the storage adapter.

Procedure

- 1 In the vSphere Client, select the host in the inventory panel.
- 2 Click the **Configuration** tab and click **Advanced Settings** under Software.
- 3 Click **Disk** in the left panel and scroll down to **Disk.SchedNumReqOutstanding**.
- 4 Change the parameter value to the number of your choice and click **OK**.

This change can impact disk bandwidth scheduling, but experiments have shown improvements for disk-intensive workloads.

What to do next

If you adjust this value in the VMkernel, you might also want to adjust the queue depth in your storage adapter.

Reducing SCSI Reservations

Operations that require getting a file lock or a metadata lock in VMFS result in short-lived SCSI reservations. SCSI reservations lock an entire LUN. Excessive SCSI reservations by a server can cause performance degradation on other servers accessing the same VMFS.

Examples of operations that require getting file locks or metadata locks include:

- Virtual machine power on.
- VMotion.
- Virtual machines running with virtual disk snapshots.
- File operations that require opening files or doing metadata updates.

Performance degradation can occur if such operations occur frequently on multiple servers accessing the same VMFS. For instance, VMware recommends that you do not run many virtual machines from multiple servers that are using virtual disk snapshots on the same VMFS. Limit the number of VMFS file operations when many virtual machines run on the VMFS.

Adjust Queue Depth for a QLogic HBA

If you are not satisfied with the performance of your QLogic adapter, you can change its maximum queue depth.

You can adjust the maximum queue depth for a QLogic qla2xxx series adapter by using the vSphere CLI.

Procedure

- 1 Verify which QLogic HBA module is currently loaded by entering the following command:
`vmkload_mod -l | grep qla2xxx.`

- 2 Run the following commands.

The example shows the qla2300_707 module. Use the appropriate module based on the outcome of the previous step.

```
vicfg-module -s ql2xmaxqdepth=64 qla2300_707
```

In this case, the HBA represented by ql2x will have its LUN queue depth set to 64.

- 3 Reboot your host.

Adjust Queue Depth for an Emulex HBA

If you are not satisfied with the performance of your Emulex adapter, you can change its maximum queue depth.

Procedure

- 1 Verify which Emulex HBA module is currently loaded by entering the `vmkload_mod -l | grep lpfcdd` command.
- 2 Run the following command.

The example shows the `lpfcdd_7xx` module. Use the appropriate module based on the outcome of [Step 1](#).

```
vicfg-module -s lpfc0_lun_queue_depth=16 lpfcdd_7xx
```

In this case, the HBA represented by `lpfc0` will have its LUN queue depth set to 16.

- 3 Reboot the host.

SAN Storage Backup Considerations

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 reprovision the data. 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.

Snapshot Software

Snapshot software allows an administrator to make an instantaneous copy of any single virtual disk defined within the disk subsystem.

Snapshot software is available at different levels:

- ESX/ESXi hosts allow you to create snapshots of virtual machines. This software is included in the basic ESX/ESXi package.
- Third-party backup software might allow for more comprehensive backup procedures and might contain more sophisticated configuration options.

Administrators make snapshots for a variety of reasons:

- Backup
- Disaster recovery
- Availability of multiple configurations, versions, or both
- Forensics (looking at a snapshot to find the cause of problems while your system is running)
- Data mining (looking at a copy of your data to reduce load on production systems)

Using a Third-Party Backup Package

Using third-party software has the advantage of a uniform environment. However, the additional cost of the third-party snapshotting software can become higher as your SAN grows.

If you are using third-party backup software, make sure that the software is supported with ESX/ESXi hosts.

If you use snapshots to back up your data, consider the following points:

- Some vendors support snapshots for both VMFS and RDMs. If both are supported, you can make either a snapshot of the whole virtual machine file system for a host, or snapshots for the individual virtual machines (one per disk).
- Some vendors support snapshots only for a setup using RDM. If only RDM is supported, you can make snapshots of individual virtual machines.

See your storage vendor's documentation.

NOTE ESX/ESXi systems also include a Consolidated Backup component.

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 ESX/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 ESX/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 RDM, data always takes the same path, which results in easier performance management.
- Security is more transparent to the storage administrator when you use RDM and an array-based solution because with RDM, 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 RDM, 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 ESX/ESXi system in conjunction with a SAN, you must decide whether host-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 RDM, a new LUN is required for each virtual machine.
- Snapshotting is included with your ESX/ESXi host at no extra cost. The file-based solution is therefore more cost-effective than the array-based solution.
- Using VMFS is easier for ESX/ESXi administrators.
- ESX/ESXi administrators who use the file-based solution are more independent from the SAN administrator.

Managing Duplicate VMFS Datastores

When a LUN 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 LUN has a unique UUID that is stored in the file system superblock. When the LUN is replicated or snapshotted, the resulting LUN copy is identical, byte-for-byte, with the original LUN. As a result, if the original LUN contains a VMFS datastore with UUID X, the LUN copy appears to contain an identical VMFS datastore, or a VMFS datastore copy, with exactly the same UUID X.

ESX/ESXi can determine whether a LUN contains the VMFS datastore copy, and either mount the datastore copy with its original UUID or change the UUID, thus resignaturing the datastore.

Mounting VMFS Datastores with Existing Signatures

You might not have to resignature a VMFS datastore copy. You can mount a VMFS datastore copy without changing its signature.

For example, you can 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 can mount the datastore copy and power on the virtual machines at the secondary site.

IMPORTANT You can mount a VMFS datastore only if it does not collide with an already mounted VMFS datastore that has the same UUID.

When you mount the VMFS datastore, ESX/ESXi allows both reads and writes to the datastore residing on the LUN copy. The LUN copy must be writable. The datastore mounts are persistent and valid across system reboots.

Because ESX/ESXi does not allow you to resignature the mounted datastore, unmount the datastore before resignaturing.

Mount a VMFS Datastore with an Existing Signature

If you do not need to resignature a VMFS datastore copy, you can mount it without changing its signature.

Prerequisites

Before you mount a VMFS datastore, perform a storage rescan on your host so that it updates its view of LUNs presented to it.

Procedure

- 1 Log in to the vSphere Client and select the server from the inventory panel.
- 2 Click the **Configuration** tab and click **Storage** in the Hardware panel.
- 3 Click **Add Storage**.
- 4 Select the **Disk/LUN** storage type and click **Next**.
- 5 From the list of LUNs, select the LUN that has a datastore name displayed in the VMFS Label column and click **Next**.

The name present in the VMFS Label column indicates that the LUN is a copy that contains a copy of an existing VMFS datastore.

- 6 Under Mount Options, select **Keep Existing Signature**.
- 7 In the Ready to Complete page, 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.

Unmount Datastores

When you unmount a datastore, it remains intact, but can no longer be seen from the hosts that you specify. It continues to appear on other hosts, where it remains mounted.

You can unmount only the following types of datastores:

- NFS datastores
- VMFS datastore copies mounted without resignaturing

Procedure

- 1 Display the datastores.
- 2 Right-click the datastore to unmount and select **Unmount**.
- 3 If the datastore is shared, specify which hosts should no longer access the datastore.
 - a If needed, deselect the hosts where you want to keep the datastore mounted.
By default, all hosts are selected.
 - b Click **Next**.
 - c Review the list of hosts from which to unmount the datastore, and click **Finish**.
- 4 Confirm that you want to unmount the datastore.

Resignaturing VMFS Copies

Use datastore resignaturing to retain the data stored on the VMFS datastore copy. When resignaturing a VMFS copy, ESX/ESXi assigns a new UUID and a new label to the copy, and mounts the copy as a datastore distinct from the original.

The default format of the new label assigned to the datastore is `snap-<snapID>-<oldLabel>`, where *<snapID>* is an integer and *<oldLabel>* is the label of the original datastore.

When you perform datastore resignaturing, consider the following points:

- Datastore resignaturing is irreversible.
- The LUN copy that contains the VMFS datastore that you resignature is no longer treated as a LUN copy.
- 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 colliding with UUIDs of any other datastore, such as an ancestor or child in a hierarchy of LUN snapshots.

Resignature a VMFS Datastore Copy

Use datastore resignaturing if you want to retain the data stored on the VMFS datastore copy.

Prerequisites

To resignature a mounted datastore copy, first unmount it.

Before you resignature a VMFS datastore, perform a storage rescan on your host so that the host updates its view of LUNs presented to it and discovers any LUN copies.

Procedure

- 1 Log in to the vSphere Client and select the server from the inventory panel.
- 2 Click the **Configuration** tab and click **Storage** in the Hardware panel.
- 3 Click **Add Storage**.
- 4 Select the **Disk/LUN** storage type and click **Next**.
- 5 From the list of LUNs, select the LUN that has a datastore name displayed in the VMFS Label column and click **Next**.

The name present in the VMFS Label column indicates that the LUN is a copy that contains a copy of an existing VMFS datastore.

- 6 Under Mount Options, select **Assign a New Signature** and click **Next**.
- 7 In the Ready to Complete page, review the datastore configuration information and click **Finish**.

What to do next

After resignaturing, you might have to do the following:

- If the resignatured datastore contains virtual machines, update references to the original VMFS datastore in the virtual machine files, including `.vmtx`, `.vmdk`, `.vmsd`, and `.vmsn`.
- To power on virtual machines, register them with vCenter Server.



Multipathing Checklist

Storage arrays have different multipathing setup requirements.

Table A-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">■ Host Mode 1: Standard.■ Host Mode 2: Sun Cluster 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	All Initiator records must have: <ul style="list-style-type: none">■ Failover Mode = 1■ Initiator Type = "Clariion Open"■ Array CommPath = "Enabled" or 1
HP MSA	Host type must be Linux. Set the connection type for each HBA port to Linux.
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">■ On XP128/1024/10000/12000, set the host mode to Windows (0x0C).■ On XP24000/20000, set the host mode to 0x01.
NetApp	No specific requirements
ESX/ESXi Configuration	Set the following Software Advanced Settings for the ESX/ESXi host: <ul style="list-style-type: none">■ Set Disk.UseLunReset to 1■ Set Disk.UseDeviceReset to 0 A PSP of Most Recently Used must be used for all LUNs hosting clustered disks for active-passive arrays. A PSP of Most Recently Used or Fixed may be used for LUNs on active-active arrays. All FC HBAs must be of the same model.

Managing Storage Paths and Multipathing Plugins

B

Use the vSphere CLI to manage the Pluggable Storage Architecture (PSA) multipathing plugins and storage paths assigned to them.

You can use the vSphere CLI to display all multipathing plugins 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 plugin should claim the paths.

For more information about additional commands available to manage PSA, see the *vSphere Command-Line Interface Installation and Reference Guide*.

This appendix includes the following topics:

- [“List Claim Rules for the Host,”](#) on page 79
- [“Display Multipathing Modules,”](#) on page 80
- [“Display SATPs for the Host,”](#) on page 81
- [“Display NMP Storage Devices,”](#) on page 81
- [“Add PSA Claim Rules,”](#) on page 82
- [“Delete PSA Claim Rules,”](#) on page 83
- [“Mask Paths,”](#) on page 83
- [“Unmask Paths,”](#) on page 84
- [“Define NMP SATP Rules,”](#) on page 84
- [“esxcli corestorage Command-Line Options,”](#) on page 86

List Claim Rules for the Host

Use the vSphere CLI to list all claim rules from 0 to 65535.

Claim rules indicate which multipathing plugin, 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

Procedure

- ◆ Use the `esxcli corestorage claimrule list` to list claim rules.

[Example B-1](#) shows the output of the command.

Example B-1. Sample Output of the `esxcli corestorage claimrule list` Command

Rule	Class	Type	Plugin	Matches
0	runtime	transport	NMP	transport=usb
1	runtime	transport	NMP	transport=sata
2	runtime	transport	NMP	transport=ide
3	runtime	transport	NMP	transport=block
101	runtime	vendor	MASK_PATH	vendor=DELL model=Universal Xport
101	file	vendor	MASK_PATH	vendor=DELL model=Universal Xport
200	runtime	vendor	MPP_1	vendor=NewVend model=*
200	file	vendor	MPP_1	vendor=NewVend model=*
201	runtime	location	MPP_2	adapter=vmhba41 channel=* target=* lun=*
201	file	location	MPP_2	adapter=vmhba41 channel=* target=* lun=*
202	runtime	driver	MPP_3	driver=megaraid
202	file	driver	MPP_3	driver=megaraid
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.
- The MASK_PATH module claims all paths returning SCSI inquiry data with a vendor string of DELL and a model string of Universal Xport. The MASK_PATH module is used to mask paths from your host.
- 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 Class column in the output 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 vSphere CLI to list all multipathing modules loaded into the system. Multipathing modules manage physical paths that connect your host with storage.

Procedure

- ◆ To list all multipathing modules, run the following command:

```
vicfg-mpath --server <server> --list-plugins,
```

where `<server>` is your vSphere CLI administration server. You might be prompted for a user name and password.

At a minimum, this command returns the NMP module. If any third-party MPPs have been loaded, they are listed as well.

Example B-2. Sample Output of the vicfg-mpath Command

```
MPP_1
MPP_2
MPP_3
MASK_PATH
NMP
```

Display SATPs for the Host

Use the vSphere CLI to list all VMware NMP SATPs loaded into the system.

Procedure

- ◆ To list all VMware SATPs, run the following command.

```
esxcli nmp satp list
```

For each SATP, the command displays information that shows the type of storage array or system this SATP supports and the default PSP for any LUNs using this SATP.

Keep in mind the following:

- 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.
- If VMW_SATP_ALUA is assigned to a specific storage device, but the device is not ALUA-aware, there is no claim rule match for this device. In this case, 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.

Example B-3. Sample Output of the esxcli nmp satp list Command

Name	Default PSP	Description
VMW_SATP_ALUA_CX	VMW_PSP_FIXED	Supports EMC CX that use the ALUA protocol
VMW_SATP_SVC	VMW_PSP_FIXED	Supports IBM SVC
VMW_SATP_MSA	VMW_PSP_MRU	Supports HP MSA
VMW_SATP_EQL	VMW_PSP_FIXED	Supports EqualLogic arrays
VMW_SATP_INV	VMW_PSP_FIXED	Supports EMC Invista
VMW_SATP_SYMM	VMW_PSP_FIXED	Supports EMC Symmetrix

Display NMP Storage Devices

Use vSphere CLI to list all storage devices controlled by the VMware NMP and display SATP and PSP information associated with each device.

Procedure

- 1 To list all storage devices, run the following command:

```
esxcli nmp device list
```

- 2 To show information for a specific device, run the following:

```
esxcli nmp device list -d <device_ID>
```

Add PSA Claim Rules

Use the vSphere CLI to add a new 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 plugin (MPP) and need to define which paths this module should claim. You may need to create a new 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 when 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 errors.

Procedure

- 1 To define a new claim rule, on the vSphere CLI, run the following command:

```
esxcli corestorage claimrule add -r <claimrule_ID> -t <type> <required_option (based on type)>
-P <MPP_name>
```

For information on the options that the command requires, see “[esxcli corestorage Command-Line Options](#),” on page 86.

- 2 To load the new claim rule into your system, run the following command:

```
esxcli corestorage claimrule load
```

This command has no options. It loads all newly created claim rules from your system's configuration file.

Example B-4. Adding a PSA Claim Rule

In the following example, you define the claim rule # 500, which specifies that the NMP module claims all paths to the NewMod model of the NewVend storage array. You then load this claim rule into your system.

```
1 # esxcli corestorage claimrule add -r 500 -t vendor -V NewVend -M NewMod -P NMP
2 # esxcli corestorage claimrule load
```

If you now run the `esxcli corestorage 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 another with the Class of file, indicate that the new claim rule has been loaded into the system and is active.

Rule	Class	Type	Plugin	Matches
0	runtime	transport	NMP	transport=usb
1	runtime	transport	NMP	transport=sata
2	runtime	transport	NMP	transport=ide
3	runtime	transport	NMP	transport=block
101	runtime	vendor	MASK_PATH	vendor=DELL model=Universal Xport
101	file	vendor	MASK_PATH	vendor=DELL model=Universal Xport
500	runtime	vendor	NMP	vendor=NewVend model=NewMod
500	file	vendor	NMP	vendor=NewVend model=NewMod

Delete PSA Claim Rules

Use the vSphere CLI to remove a PSA claim rule from the set of claim rules on the system.

Procedure

- 1 Delete a claim rule from the set of claim rules.

```
esxcli corestorage claimrule delete -r <claimrule_ID>
```

For information on the options that the command takes, see “[esxcli corestorage Command-Line Options](#),” on page 86.

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.

- 2 Remove the claim rule from the ESX/ESXi system.

```
esxcli corestorage claimrule load
```

Mask Paths

You can prevent the ESX/ESXi host from accessing storage devices or LUNs or from using individual paths to a LUN. Use the vSphere CLI commands to mask the paths.

When you mask paths, you create claim rules that assign the MASK_PATH plugin to the specified paths.

Procedure

- 1 Check what the next available rule ID is.

```
esxcli corestorage 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 plugin to a path by creating a new claim rule for the plugin.

```
esxcli corestorage claimrule add -r <claimrule_ID> -t <type> <required_option> -P <MASK_PATH>
```

For information on command-line options, see “[esxcli corestorage Command-Line Options](#),” on page 86.

- 3 Load the MASK_PATH claim rule into your system.

```
esxcli corestorage claimrule load
```

- 4 Verify that the MASK_PATH claim rule was added correctly.

```
esxcli corestorage claimrule list
```

- 5 If a claim rule for the masked path exists, remove the rule.

```
esxcli corestorage claiming unclaim <type> <required_option>
```

- 6 Run the path claiming rules.

```
esxcli corestorage claimrule run
```

After you assign the MASK_PATH plugin 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 B-5. 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 corestorage claimrule list

2  #esxcli corestorage claimrule add -P MASK_PATH -r 109 -t location -A vmhba2 -C 0 -T 1 -L 20
   #esxcli corestorage claimrule add -P MASK_PATH -r 110 -t location -A vmhba3 -C 0 -T 1 -L 20
   #esxcli corestorage claimrule add -P MASK_PATH -r 111 -t location -A vmhba2 -C 0 -T 2 -L 20
   #esxcli corestorage claimrule add -P MASK_PATH -r 112 -t location -A vmhba3 -C 0 -T 2 -L 20

3  #esxcli corestorage claimrule load

4  #esxcli corestorage claimrule list

5  #esxcli corestorage claiming unclaim -t location -A vmhba2
   #esxcli corestorage claiming unclaim -t location -A vmhba3

6  # esxcli corestorage claimrule run

```

Unmask Paths

When you need the host to access the masked storage device, unmask the paths to the device.

Procedure

- 1 Unmask a path to the storage device by running the `esxcli corestorage claiming unclaim` command.

Run this command for each path to the storage device.

For example:

```
esxcli corestorage claiming unclaim -t location -A vmhba0 -C 0 -T 0 -L 149
```

- 2 Load path claiming rules into the VMkernel by running the `esxcli corestorage claimrule load` command.
- 3 Run the path claiming rules by entering the `esxcli corestorage claimrule run`.

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 vSphere CLI to add a rule to the list of claim rules for the specified SATP.

You might need to create a new SATP rule when you install a third-party SATP for a specific storage array.

Procedure

- 1 To add a claim rule for a specific SATP, run the following command.

```
esxcli nmp satp addrule <rule_parameter> -e <description> -o <option> -s <SATP_name>
```

Use the following options for *<rule_parameter>*. The *-V* and *-M* options can be used at the same time. They cannot be used in conjunction with the *-R* or *-D* options.

NOTE When searching the SATP rules to locate an 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.

- *-D <driver>* -- Driver string to set when adding the SATP claim rule.
- *-V <vendor>* -- Vendor string to set when adding the SATP claim rule.
- *-M <model>* -- Model string to set when adding the SATP claim rule.
- *-R <transport>* -- Transport type string to set when adding the SATP claim rule.

Specify the following options for any SATP claim rule:

- *-e <description>* -- Description string to set when adding the SATP claim rule.
- *-o <option>* -- Claim option string to set when adding the SATP claim rule. This string is passed to the SATP when the SATP claims a path. The contents of this string, and how the SATP behaves as a result, are unique to each SATP. For example, some SATPs support the claim option strings *tpgs_on* and *tpgs_off*. If *tpgs_on* is specified, the SATP will claim the path only if the ALUA Target Port Group support is enabled on the storage device.

- 2 To delete a rule from the list of claim rules for the specified SATP, run the following command. You can run this command with the same options you used for *addrule*.

```
esxcli nmp satp deleterule <rule_parameter> -s <SATP_name>
```

- 3 Reboot your host.

Example B-6. Defining an NMP SATP Rule

The following sample command assigns the *VMW_SATP_INV* plugin to manage storage arrays with vendor string *NewVend* and model string *NewMod*.

```
# esxcli nmp satp addrule -V NewVend -M NewMod -s VMW_SATP_INV
```

If you run the **esxcli nmp satp listrules -s VMW_SATP_INV** command, you can see the new rule added to the list of *VMW_SATP_INV* rules.

Name	Vendor	Model	Driver	Transport	Options	Claim Options	Description
VMW_SATP_INV	EMC	Invista					
VMW_SATP_INV	EMC	LUNZ					Invista LUNZ
VMW_SATP_INV	NewVend	NewMod					

esxcli corestorage Command-Line Options

Some `esxcli corestorage` commands, for example the commands that you run to add new claim rules, remove the rules, or mask paths, require that you specify certain options.

Table B-1. `esxcli corestorage` Command-Line Options

Option	Description	Required Option
<code>-r <claimrule_ID></code>	Use to specify the order number for the claim rule from 0 to 65535.	
<code>-t <type></code>	Use to define the set of paths for the claim rule. Specify one of the following values for the <code><type></code> variable:	These options change depending on the value you enter for <code><type></code> .
	<code>vendor</code> – Indicate the vendor and model of the storage device used for this path.	<code>-V <vendor></code> <code>-M <model></code> Use asterisk (*) to specify all vendors or models.
	<code>location</code> – Indicate the adapter, channel, target, or LUN used for this path.	Use any of the following: <ul style="list-style-type: none"> ■ <code>-A <adapter></code> ■ <code>-C <channel></code> ■ <code>-T <target></code> ■ <code>-L <lunID></code>
	<code>driver</code> – Indicate the driver used for the path.	<code>-D <driver></code>
	<code>transport</code> – Indicate the transport used for the path.	<code>-R <transport></code> Use one of the following for the <code><transport></code> variable: <ul style="list-style-type: none"> ■ <code>block</code> – Raid block devices, such as <code>cciss</code> ■ <code>fc</code> – Fibre Channel ■ <code>iscsi</code> – Default iSCSI ■ <code>iscsivendor</code> – iSCSI with vendor supplied IMA ■ <code>ide</code> – IDE ■ <code>sas</code> – Serial attached SCSI ■ <code>sata</code> – Serial ATA ■ <code>usb</code> – USB storage devices ■ <code>parallel</code> – Parallel SCSI devices ■ <code>unknown</code> – Unknown storage device type
<code>-P <MPP_name></code>	Indicate which MPP plug-in should claim the paths defined by the claim rule. Run the <code>vicfg-mpath --list-plugins</code> command to see valid values.	

Index

Symbols

* next to path 61

A

access, equalizing disk access 69
active-active disk arrays 16, 38, 62
active-passive disk arrays, path policy reset 66
active/active disk arrays, managing paths 63
active/passive disk arrays
 boot from SAN 31
 HP StorageWorks MSA 39
 managing paths 63
 path thrashing 69
adaptive scheme 20
allocations 30
applications, layered 72
array-based (third-party) solution 73
asterisk next to path 61
auto volume transfer 37
automatic host registration, disabling 66
avoiding problems 66
AVT 37
AX100
 display problems 35
 inactive connections 35
axnaviserverutil cli utility 35

B

backups
 considerations 71
 third-party backup package 72
basic connectivity 33
benefits 17
BIOS
 enabling for BFS 48
 enabling Qlogic HBA for BFS 47
boot BIOS prompt, enabling for BFS 48
boot from CD-ROM 46
boot from SAN
 benefits 44
 boot LUN considerations 31
 conceptual overview 43
 diagnostic partitions 44
 Emulex FC HBA 48
 enabling Qlogic HBA BIOS 47

ESX requirements 31
HBA requirements 31
introduction 43
LUN masking 45
preparing installation 44
Qlogic FC HBA 46
recommendations 45
requirements 31
boot LUN 47
boot LUN, selecting 47
BusLogic, queue depth 29

C

can't see LUN 54
CD-ROM, booting from 46
claim rules
 adding 82
 deleting 83
cluster across boxes 27
cluster in a box 27
cluster services 27
clustering 33
commands
 SDK 13
 vSphere CLI 13
configuration, storage processor sense data 38
configuring hardware for SAN failover,
 DS4800 36
CPU virtualization 11
current multipathing state 61

D

datastore copies, mounting 74
datastores
 managing duplicate 73
 mounting 74
 paths 61
 refresh 55
 reviewing properties 54
 unmounting 74
 viewing information 54
design, for server failure 26
device driver options 65
device drivers 9

- diagnostic partitions
 - boot from SAN **44**
 - sharing **65**
- direct connect **33**
- disabling auto volume transfer **37**
- disabling paths **63**
- disaster recovery **17**
- disk access, equalizing **69**
- disk arrays
 - active-active **62**
 - active-passive **62**
 - active/active **30**
 - active/passive **30, 47, 69**
 - zoning disk array **55, 56**
- disk shares **21**
- Disk.MaxLUN **56**
- Disk.SchedNumReqOutstanding parameter **69**
- Disk.SupportSparseLUN **57**
- display problems, AX100 **35**
- distributed locking **12**
- drivers, device drivers **65**
- DRS **28**
- DS4800, configuring hardware for SAN
 - failover **36**
- dump partitions, sharing **65**

E

- educational support **7**
- EMC CLARiiON **34**
- EMC Symmetrix, pseudo LUNs **35**
- Emulex FC HBA
 - boot from SAN **48**
 - NPIV support **58**
- Emulex HBA, adjusting queue depth **71**
- enabling BIOS for BFS **48**
- enabling boot BIOS prompt for BFS **48**
- equalizing disk access **69**
- esx, host type **34**
- ESX/ESXi
 - introduction **9**
 - sharing VMFS **18**
- ESX/ESXi and SAN, requirements **29**
- esxcli corestorage command options **86**
- EVA (HP StorageWorks) **40**
- extents **12**

F

- failover
 - FAST storage **36**
 - transparent **16**
- failover paths, status **61**
- failure **26**
- FC HBA setup **30**

- Fibre Channel, concepts **15**
- file-based (VMFS) solution **73**
- finding information **18**
- Fixed path policy, path thrashing **68**

H

- HA **26**
- hardware compatibility **10**
- HBA
 - Emulex **48, 58**
 - enabling Qlogic HBA BIOS for BFS **47**
 - Qlogic **46, 58**
 - queue depth **70**
 - setup **30**
 - static load balancing **30**
- high-tier storage **26**
- Hitachi Data Systems storage, microcode **41**
- host registration, disabling **66**
- host type **34**
- host-based failover **23**
- HP StorageWorks
 - EVA **40**
 - MSA **39**
 - XP **40**
- hub controller issues **40**

I

- IBM ESS800 **38**
- IBM FASTT **36**
- IBM System Storage DS4800 **36**
- IBM Systems Storage 8000 **38**
- initiators, minimizing the number **46**
- installation
 - preparing for boot from SAN **44**
 - steps **31**
- Inter-Switch Link **36**
- interacting with ESX/ESXi systems **13**
- ISL **36**
- issues
 - hub controller **40**
 - performance **68**
 - visibility **54**

L

- layered applications **72**
- Linux
 - host type **34**
 - profile name **39**
 - VMkernel **9**
- Linux Cluster, host type **34**
- load balancing, manual **63**
- locations of virtual machines **26**
- locking **12**

- lower-tier storage **26**
- LSILogic queue depth **29**
- LUN decisions
 - adaptive scheme **20**
 - predictive scheme **20**
- LUN masking, boot from SAN **45**
- LUN not visible, SP visibility **54**
- LUNs
 - 1 VMFS volume **29**
 - allocations **30**
 - boot LUN **47**
 - can't see **54**
 - changing number scanned **56**
 - creating, and rescan **54–56**
 - decisions **19**
 - making changes and rescan **55**
 - masking **83**
 - masking changes and rescan **54, 56**
 - multipathing policy **62**
 - NPIV-based access **57**
 - number scanned **56**
 - selecting boot LUN **47**
 - setting multipathing policy **62**
 - sparse **57**
- M**
- maintenance **17**
- manual load balancing **63**
- mapping file **13**
- masking LUNs **83**
- maximum HBA queue depth **70**
- memory virtualization **11**
- metadata updates **19**
- microcode, Hitachi Data Systems storage **41**
- Microsoft Cluster Service **13, 33**
- mid-tier storage **26**
- Most Recently Used path policy, path thrashing **68**
- mounting VMFS datastores **74**
- MPPs
 - displaying **80**
 - See also* multipathing plugins
- MRU path policy **62**
- MSA (HP StorageWorks) **39**
- MSCS **33**
- multipathing
 - active paths **61**
 - broken paths **61**
 - disabled paths **61**
 - standby paths **61**
 - viewing the current state of **61**
- multipathing plugins, path claiming **60**
- multipathing policy **62**
- multipathing setup requirements **77**
- multipathing state **61**
- N**
- N-Port ID Virtualization **57**
- N-Port ID Virtualization (NPIV), requirements **58**
- N+1 clustering **27**
- Native Multipathing Plugin **23, 24**
- Netware host mode **41**
- Network Appliance storage, provisioning storage **41**
- network virtualization **11**
- NFS datastores, unmounting **74**
- NMP, path claiming **60**
- NPIV **57**
- number of extents **12**
- number of outstanding disk requests **69**
- O**
- operating system timeout **64**
- optimizing resource utilization **27**
- outstanding disk requests **69**
- P**
- passive disk arrays, path thrashing **69**
- path claiming **60**
- path failover **22, 23**
- path failure rescan **55, 56**
- path management **22, 63**
- path policies
 - changing defaults **63**
 - Fixed **25, 62**
 - Most Recently Used **25, 62**
 - MRU **62**
 - Round Robin **25, 62**
- path policy reset, active-passive disk array **66**
- Path Selection Plugins **25**
- path thrashing, resolving **69**
- paths
 - disabling **63**
 - masking **83**
 - preferred **61**
 - unmasking **84**
- performance
 - issues **68**
 - optimizing **67**
 - SCSI reservations **18**
- physical to virtual clustering **27**
- Pluggable Storage Architecture **23**
- Port_ID **16**
- ports, configuration **37**
- predictive scheme **20**

- preferred path **61**
- prioritizing virtual machines **21**
- problems
 - avoiding **66**
 - hub controller **40**
 - performance **68**
 - visibility **54**
- profile name, Linux **39**
- PSA, *See* Pluggable Storage Architecture
- PSA claim rules, deleting **83**
- PSPs, *See* Path Selection Plugins

Q

- Qlogic FC HBA
 - boot from SAN **46**
 - NPIV support **58**
- Qlogic HBA BIOS, enabling for BFS **47**
- queue depth **70, 71**

R

- raw device mapping, mapping file **13**
- RDM
 - mapping file **13**
 - Microsoft Cluster Service **13**
- refresh **55**
- requirements, boot from SAN **31**
- rescan
 - adding disk array **55, 56**
 - LUN creation **54–56**
 - LUN masking **54**
 - path masking **55, 56**
 - when path is down **55, 56**
- reservations, reducing SCSI reservations **70**
- resolving problems **66**
- resource utilization, optimizing **27**
- restrictions **29**
- Round Robin path policy **25, 62**

S

- SAN
 - backup considerations **71**
 - hardware failover **36**
 - preparing **45**
 - requirements **29**
 - server failover **27**
 - specifics **21**
- SAN fabric **15**
- SAN management software **21**
- SAN storage performance, optimizing **67**
- SAN storage, benefits **17**
- SANs, accessing **22**
- SATP rules, adding **84**
- scanning, changing number **56**

- SCSI controllers **11**
- SCSI controllers, device driver options **65**
- SCSI reservations, reducing **70**
- SDK **13**
- selectable boot, enabling **47**
- server failover **27**
- server failure **26**
- server performance **67**
- service console **43**
- setup steps **31**
- sharing diagnostic partitions **65**
- sharing VMFS across servers **18**
- snapshot software **72**
- software compatibility **10**
- SP visibility, LUN not visible **54**
- sparse LUN support **57**
- storage adapter, displaying in vSphere Client **52**
- storage adapters, viewing in vSphere Client **51**
- SATPs, displaying **81**
- Storage Array Type Plugins **24**
- storage arrays
 - configuring **33**
 - multipathing requirements **77**
 - performance **67**
- storage devices
 - accessible through adapters **54**
 - available to hosts **53**
 - displaying **81**
 - naming **53**
 - paths **62**
 - viewing information **52**
- storage processors
 - configuring sense data **38**
 - port configuration **37**
 - sense data **38**
- storage systems
 - EMC CLARiiON **34**
 - EMC Symmetrix **35**
 - Hitachi **41**
 - HP StorageWorks **39**
 - Network Appliance **41**
 - types **16**
- storage virtualization **11**
- supported devices **34**

T

- tape devices **30**
- technical support **7**
- third-party backup package **72**
- third-party management applications **21**
- timeout **64**
- TimeoutValue parameter **29**

troubleshooting **66**

U

use cases **17**

V

vCenter Server, accessing **13**

Virtual Machine File System **12**

Virtual Machine Monitor **9**

virtual machines

accessing SANs **22**

assigning WWNs to **58**

equalizing disk access **69**

locations **26**

prioritizing **21**

virtual ports (VPORTs) **57**

virtualization **10**

visibility issues **54**

VMFS

1 volume per LUN **29**

creating new volume **12**

locking **12**

minimum size **12**

number of extents **12**

sharing across ESX/ESXi hosts **18**

volume resignaturing **73**

VMFS datastores

changing signatures **75**

resignaturing copies **75**

unmounting **74**

VMFS volume resignaturing **73**

VMkernel **9**

VMM **9**

VMotion **17, 28, 30**

vmware, host type **34**

VMware DRS **17, 28**

VMware HA **17, 26**

VMware NMP

I/O flow **25**

See also Native Multipathing Plugin

VMware vSphere Client **9**

volume resignaturing **73, 75**

vSphere CLI, See vSphere Command-Line Interface

vSphere Client **9, 13**

vSphere Command-Line Interface **13**

vSphere SDK **13**

vSphere Web Access **9, 13**

W

World Wide Names (WWNs)

assigning to virtual machines **58**

World Wide Port Names (WWPNs) **37, 57**

WWNs

assigning **59**

changing **59**

WWPN **16**

X

XP (HP StorageWorks) **40**

Z

zoning **15**

