Fibre Channel SAN Configuration Guide

ESX 4.1
ESXi 4.1
vCenter Server 4.1

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-000290-02

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

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

Contents

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-000290-02	Removed reference to the IBM System Storage DS4800 Storage Systems. These devices are not supported with ESX/ESXi 4.1.
EN-000290-01	■ "HP StorageWorks XP," on page 36 and Appendix A, "Multipathing Checklist," on page 75 have been changed to include host mode parameters required for HP StorageWorks XP arrays. ■ "Boot from SAN Restrictions and Benefits," on page 39 is updated to remove a reference to the restriction on using Microsoft Cluster Service.
EN-000290-00	Initial release.

About This Book

This manual, 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:

- Overview of VMware ESX/ESXi – Introduces ESX/ESXi systems for SAN administrators.
- Using ESX/ESXi with a Fibre Channel SAN – Discusses requirements, noticeable differences in SAN setup if ESX/ESXi is used, and how to manage and troubleshoot the two systems together.
- Using Boot from SAN with ESX/ESXi Systems – Discusses requirements, limitations, and management of boot from SAN.

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

Intended Audience

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

VMware Technical Publications Glossary

VMware Technical Publications provides a glossary of terms that might be unfamiliar to you. For definitions of terms as they are used in VMware technical documentation, go to http://www.vmware.com/support/pubs.

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 1

You can use ESX/ESXi in conjunction with the Fibre Channel storage area network (SAN), a specialized high-speed network that uses the Fibre Channel (FC) protocol to transmit data between your computer systems and high-performance storage subsystems. SANs allow hosts to share storage, provide extra storage for consolidation, improve reliability, and help with disaster recovery.

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

This chapter includes the following topics:

Introduction to ESX/ESXi

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

ESX/ESXi 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 the VMkernel. The VMkernel manages most of the physical resources on the hardware, including memory, physical processors, storage, and networking controllers.

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.

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 a number of 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.

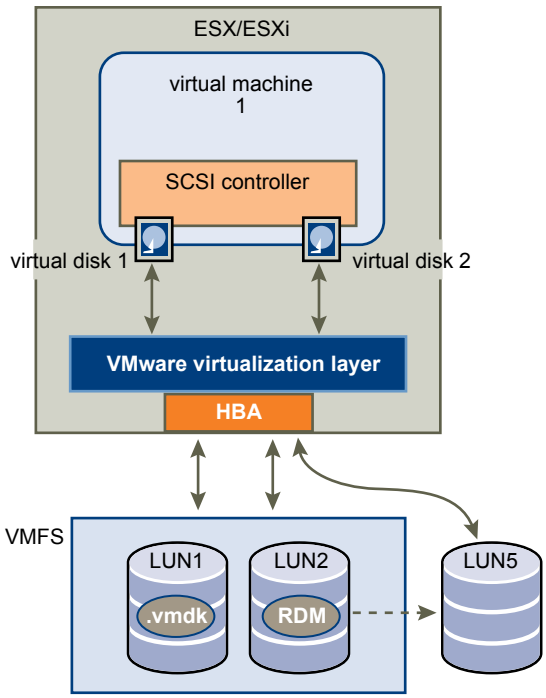
An ESX/ESXi virtual machine uses a virtual disk to store its operating system, program files, and other data associated with its activities. A virtual disk is a large physical file, or a set of files, that can be copied, moved, archived, and backed up as easily as any other file. You can configure virtual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

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

Each virtual disk that a virtual machine can access through one of the virtual SCSI controllers resides on a VMware Virtual Machine File System (VMFS) datastore, an 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 on 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 detected and used to create the VMFS datastore during the host's first boot.

A VMFS datastore can map to a single SAN LUN or local disk or stretch over multiple SAN LUNs or local disks. You can expand a datastore while virtual machines are running on it, either by growing the datastore or by adding a new physical extent. The VMFS datastore can be extended to span 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, an entire SAN LUN can be directly allocated to a virtual machine. The RDM provides some of the advantages of a virtual disk in a VMFS datastore, 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 a particular SAN device. 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 the *vSphere Command-Line Interface Installation and Scripting Guide* and the *vSphere Command-Line Interface Reference*.
- 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 storage arrays, 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

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 the Fibre Channel (FC) protocol that packages SCSI commands into Fibre Channel frames.

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

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

Ports

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

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

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

When N-Port ID Virtualization (NPIV) is used, a single FC HBA port (N-port) can register with the fabric by using several WWPNs. This method allows an N-port to claim multiple fabric addresses, each of which appears as a unique entity. When 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 technique known as multipathing. Multipathing allows you to have more than one physical path from the ESX/ESXi host to a LUN on a storage system.

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

Storage System Types

ESX/ESXi supports different storage systems and arrays.

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

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

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

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

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, failover, and load balancing technologies.

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

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

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

ESX/ESXi and SAN Use Cases

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

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

Maintenance with zero downtime

When performing 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 users of the virtual machines.

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 users of the virtual machines.

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 without interruptions to the users of the virtual machines.

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 typically 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. 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 the same VMFS datastore.

Use the vSphere Client to set up a VMFS datastore in advance on a block-based storage device that your ESX/ESXi host discovers. A VMFS datastore can be extended to span several physical storage extents, including SAN LUNs and local storage. This feature allows you to pool storage and gives you flexibility in creating the datastore necessary for your virtual 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. VMFS is designed for concurrent access from multiple physical machines and enforces the appropriate access controls on virtual machine files.

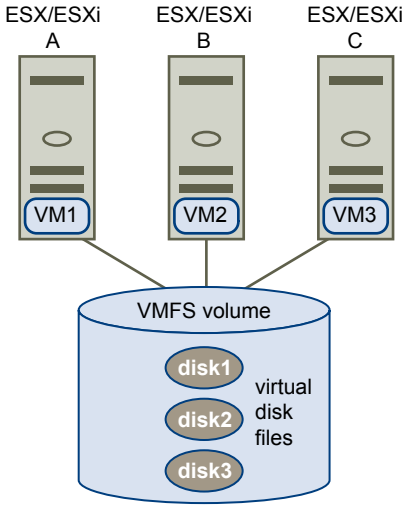
Sharing a VMFS Datastore Across ESX/ESXi Hosts

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.

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, RDM descriptor files, and so on. The datastore also maintains a consistent view of all the mapping information for these objects. This mapping information is called metadata.

Metadata is updated each time 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 the applications running in virtual machines that use the LUN.
- 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 method to determine the number and size of LUNs to provision. You can experiment using either a predictive or adaptive scheme.

Use the Predictive Scheme to Make LUN Decisions

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

Procedure

1 Provision several LUNs with different storage characteristics.

2 Create a VMFS datastore on each LUN, labeling each datastore according to its characteristics.

3 Create virtual disks to contain the data for virtual machine applications in the VMFS datastores created on LUNs with the appropriate RAID level for the applications' requirements.

4 Use disk shares to distinguish high-priority from low-priority virtual machines.

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

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

Use the Adaptive Scheme to Make LUN Decisions

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

Procedure

1 Provision a large LUN (RAID 1+0 or RAID 5), with write caching enabled.

2 Create a VMFS on that LUN.

3 Create four or five virtual disks on the VMFS.

4 Run the applications to determine whether disk performance is acceptable.

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

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.
- 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.

Using Zoning

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

Zoning has the following effects:

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

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

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

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 see the SAN. In this case, you can use an RDM.

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

How Virtual Machines Access Data on a SAN

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

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

1. When the guest operating system in a virtual machine reads or writes to SCSI disk, it issues SCSI commands to the virtual disk.
2. Device drivers in the virtual machine's operating system communicate with the virtual SCSI controllers.
3. The virtual SCSI Controller forwards the command to the VMkernel.
4. The VMkernel performs the following tasks.
   - 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.
5. The physical HBA performs the following tasks.
   - Packages the I/O 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.

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 an 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.

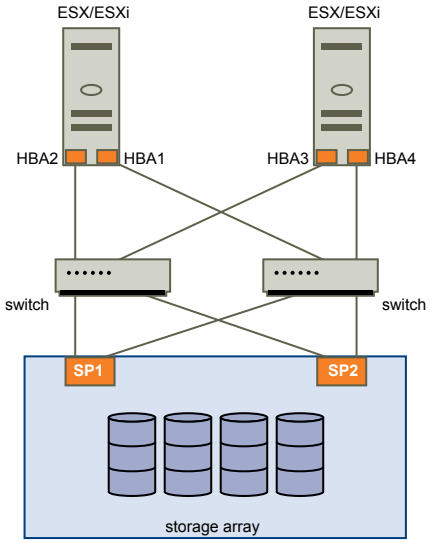
NOTE Virtual machine I/O might be delayed for up to sixty seconds while path failover takes place. These delays allow the SAN to stabilize its configuration after topology changes. In general, the I/O delays might be longer on active-passive arrays and shorter on activate-active arrays.

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 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 both HBA and SP failovers with its multipathing capability.

Managing Multiple Paths

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

The VMkernel multipathing plug-in that ESX/ESXi provides by default is the VMware Native Multipathing Plug-In (NMP). The NMP is an extensible module that manages sub plug-ins. There are two types of NMP sub plug-ins, Storage Array Type Plug-Ins (SATPs), and Path Selection Plug-Ins (PSPs). SATPs and PSPs can be built-in and provided by VMware, or can be provided by a third party.

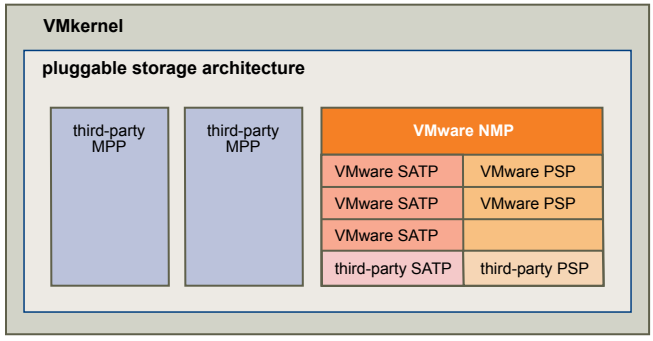
If more multipathing functionality is required, a third party can also provide an MPP to run in addition to, or as a replacement for, the default NMP.

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

- Loads and unloads multipathing plug-ins.
- Hides virtual machine specifics from a particular plug-in.
- Routes I/O requests for a specific logical device to the MPP managing that device.
- Handles I/O queuing to the logical devices.
- Implements logical device bandwidth sharing between virtual machines.
- Handles I/O queueing to the physical storage HBAs.
- Handles physical path discovery and removal.
- Provides logical device and physical path I/O statistics.

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

Figure 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.
- Support path failure detection and remediation.

- Process I/O requests to logical devices:
  - Select an optimal physical path for the request.
  - Depending on a storage device, perform specific actions necessary to handle path failures and I/O command retries.
- Support management tasks, such as abort or reset of logical devices.

VMware Multipathing Module

By default, ESX/ESXi provides an extensible multipathing module called the Native Multipathing Plug-In (NMP).

Generally, the VMware NMP supports all storage arrays listed on the VMware storage HCL and provides a default path selection algorithm based on the array type. The NMP associates a set of physical paths with a specific storage device, or LUN. The specific details of handling path failover for a given storage array are delegated to a Storage Array Type Plug-In (SATP). The specific details for determining which physical path is used to issue an I/O request to a storage device are handled by a Path Selection Plug-In (PSP). SATPs and PSPs are sub plug-ins within the NMP module.

Upon installation of ESX/ESXi, the appropriate SATP for an array you use will be installed automatically. You do not need to obtain or download any SATPs.

VMware SATPs

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

ESX/ESXi offers a SATP for every type of array that VMware supports. It also provides default SATPs that support non-specific active-active and ALUA storage arrays, and the local SATP for direct-attached devices. Each SATP accommodates special characteristics of a certain class of storage arrays and can perform the array-specific operations required to detect path state and to activate an inactive path. As a result, the NMP module itself can work with multiple storage arrays without having to be aware of the storage device specifics.

After the NMP determines which SATP to use for a specific storage device and associates the SATP with the physical paths for that storage device, the SATP implements the tasks that include the following:

- Monitors the health of each physical path.
- Reports changes in the state of each physical path.
- Performs array-specific actions necessary for storage fail-over. For example, for active-passive devices, it can activate passive paths.

VMware PSPs

Path Selection Plug-Ins (PSPs) run with the VMware NMP and are responsible for choosing a physical path for I/O requests.

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

By default, the VMware NMP supports the following PSPs:

Most Recently Used (VMW_PSP_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. MRU is the default path policy for active-passive arrays.

Fixed (VMW_PSP_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 reverts back to the preferred path as soon as that path becomes available. Fixed is the default path policy for active-active arrays.

CAUTION If used with active-passive arrays, the **Fixed** path policy might cause path thrashing.

VMW_PSP_FIXED_AP Extends the Fixed functionality to active-passive and ALUA mode arrays.

Round Robin (VMW_PSP_RR) Uses a path selection algorithm that rotates through all available active 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 The NMP issues the I/O request on the path selected by the PSP.

4 If the I/O operation is successful, the NMP reports its completion.

5 If the I/O operation reports an error, the NMP calls the appropriate SATP.

6 The SATP interprets the I/O command errors and, when appropriate, activates the inactive paths.

7 The PSP is called to select a new path on which to issue the I/O.

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 (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-tier characteristics in software.

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

- How critical is the virtual machine?
- What are the virtual machine and the applications' I/O requirements?
- What are the virtual machine point-in-time (PiT) restoration and availability requirements?
- What are its backup requirements?
- What are its replication requirements?

A virtual machine might change tiers during 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 a shared 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 linking 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 performed.

VMware supports 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 multiple 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.

NOTE As a rule, when you are booting from a SAN LUN, only the host that is booting from that LUN should see the LUN.

Optimizing Resource Use

VMware vSphere allows you to optimize resource allocation by migrating virtual machines from overloaded hosts to less busy 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 perform live migration of running virtual machines from one host to another without service interruption. The hosts should be connected to the same SAN.

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 automatically 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 CPU and memory resource use across the cluster by performing automatic migrations of virtual machines using vMotion, or by providing recommendations for virtual machine migrations.

Requirements and Installation 3

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

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.
- Use RDMs to access raw disks, or LUNs, from an ESX/ESXi host.
- For multipathing to work properly, each LUN must present the same LUN ID number to all ESX/ESXi hosts.
- Make sure the storage device driver specifies a large enough queue. You can set the queue depth for the physical HBA during system setup.
- 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

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

- 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 is an exception, but requires special configuration.

Setting LUN Allocations

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

When you set LUN allocations, be aware of 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.

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

vMotion and VMware DRS

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

Active/active compared to active-passive arrays

When you use vMotion or DRS with an active-passive SAN storage device, make sure that all 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 HBAs

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 cannot 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 59.

Setting the Timeout for Failover

Set the timeout value for detecting a failover. The default timeout is 10 seconds. To ensure optimal performance, do not change the default value.

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*.

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 and install guest operating systems.

7 (Optional) Set up your system for VMware HA failover or for using Microsoft Clustering Services.

8 Upgrade or modify your environment as needed.

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

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

Direct connect	The server connects to the array without using switches. For all other tests, a fabric connection is used. FC Arbitrated Loop (AL) is not supported.
Clustering	The system is tested with Microsoft Cluster Service running in the virtual machine.

General Setup Considerations for Fibre Channel SAN Arrays

When you prepare your FC SAN storage to work with 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. However, 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 61.

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

- The default multipathing policy for CLARiiON arrays that do not support ALUA is Most Recently Used. For CLARiiON arrays that support ALUA, the default multipathing policy is VMW_PSP_FIXED_AP. 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 For ESX/ESXi to support EMC CLARiiON with ALUA, check the HCLs to make sure that you use the correct firmware version on the storage array. For additional information, contact your storage vendor.

EMC 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.

EMC CLARiiON 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 `axnaviserverutil 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 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 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 at the HP Web site.

HP StorageWorks XP

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

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

Hitachi Data Systems Storage

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

LUN masking To mask LUNs on an 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 The modes you set depend on the model you are using, for example:

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

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

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
```

LSI-Based Storage Systems

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

Otherwise, ESX installation might fail.

Using Boot from SAN with ESX/ESXi Systems 5

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

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

This chapter includes the following topics:

Boot from SAN Restrictions and Benefits

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

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

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

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

- Cheaper servers. Servers can be more dense and run cooler without internal storage.
- Easier server replacement. You can replace servers and have the new server point to the old boot location.
- Less wasted space. Servers without local disks often take up less space.
- Easier backup processes. You can backup the system boot images in the SAN as part of the overall SAN backup procedures. Also, you can use advanced array features such as snapshots on the boot image.

- Improved management. Creating and managing the operating system image is easier and more efficient.
- Better reliability. You can access the boot disk through multiple paths, which protects the disk from being a single point of failure.

Boot from SAN Requirements and Considerations

Your ESX/ESXi boot configuration must meet specific requirements.

Table 5-1 specifies the criteria your ESX/ESXi environment must meet.

Table 5-1. Boot from SAN Requirements

Requirement	Description
ESX/ESXi system requirements	Follow vendor recommendation for the server booting from a SAN.
Adapter requirements	Enable and correctly configure the adapter, so it can access the boot LUN. See your vendor documentation.
Access control	■ Each host must have access to its own boot LUN only, not the boot LUNs of other hosts. Use storage system software to make sure that the host accesses only the designated LUNs. ■ Multiple servers can share a diagnostic partition. You can use array specific LUN masking to achieve this.
Multipathing support	Multipathing to a boot LUN on active-passive arrays is not supported because the BIOS does not support multipathing and is unable to activate a standby path.
SAN considerations	SAN connections must be through a switched topology if the array is not certified for direct connect topology. If the array is certified for direct connect topology, the SAN connections can be made directly to the array. Boot from SAN is supported for both switched topology and direct connect topology if these topologies for the specific array are certified.
Hardware- specific considerations	If you are running an IBM eServer BladeCenter and use boot from SAN, you must disable IDE drives on the blades.

Getting Ready for Boot from SAN

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

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

1 Configure SAN Components and Storage System on page 40

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

2 Configure Storage Adapter to Boot from SAN on page 41

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

3 Set Up Your System to Boot from Installation Media on page 41

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

Configure SAN Components and Storage System

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

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

Procedure

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

Check the switch wiring, if there is any.

2 Configure the storage array.

a From the SAN storage array, make the ESX/ESXi host visible to the SAN. This process is often referred to as creating an object.

b From the SAN storage array, set up the host to have the WWPNs of the host's adapters as port names or node names.

c Create LUNs.

d Assign LUNs.

e Record the IP addresses of the switches and storage arrays.

f Record the WWPN for each SP.

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

Configure Storage Adapter to Boot from SAN

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

Prerequisites

Determine the WWPN for the storage adapter.

Procedure

- Configure the storage adapter to boot from SAN.

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

Set Up Your System to Boot from Installation Media

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

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

Procedure

1 During your system power up, enter the system BIOS Configuration/Setup Utility.

2 Select **Startup Options** and press Enter.

3 Select **Startup Sequence Options** and press Enter.

4 Change the **First Startup Device** to **[CD-ROM]**.

You can now install ESX/ESXi.

Configure Emulex HBA to Boot from SAN

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

Procedure

1 Enable the BootBIOS Prompt on page 42

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

2 Enable the BIOS on page 42

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

Enable the BootBIOS Prompt

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

Procedure

1 Run `lputil`.

2 Select **3. Firmware Maintenance**.

3 Select an adapter.

4 Select **6. Boot BIOS Maintenance**.

5 Select **1. Enable Boot BIOS**.

Enable the BIOS

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

Procedure

1 Reboot the ESX/ESXi host.

2 To configure the adapter parameters, press ALT+E at the Emulex prompt and follow these steps.

a Select an adapter (with BIOS support).

b Select **2. Configure This Adapter's Parameters.**

c Select **1. Enable or Disable BIOS.**

d Select **1** to enable BIOS.

e Select **x** to exit and **Esc** to return to the previous menu.

3 To configure the boot device, follow these steps from the Emulex main menu.

   a Select the same adapter.

   b Select **1. Configure Boot Devices.**

   c Select the location for the Boot Entry.

   d Enter the two-digit boot device.

   e Enter the two-digit (HEX) starting LUN (for example, **08**).

   f Select the boot LUN.

   g Select **1. WWPN.** (Boot this device using WWPN, not DID).

   h Select **x** to exit and **Y** to reboot.

4 Boot into the system BIOS and move Emulex first in the boot controller sequence.

5 Reboot and install on a SAN LUN.

Configure QLogic HBA to Boot from SAN

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

Procedure

1 While booting the server, press **Ctrl+Q** to enter the Fast!UTIL configuration utility.

2 Perform the appropriate action depending on the number of HBAs.

Option	Description
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. 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 **Adapter Settings** and press **Enter**.

5 Set the BIOS to search for SCSI devices.

   a In the Host Adapter Settings page, select **Host Adapter BIOS**.

   b Press Enter to toggle the value to **Enabled**.

   c Press Esc to exit.

6 Enable the selectable boot.

   a Select **Selectable Boot Settings** and press **Enter**.

   b In the Selectable Boot Settings page, select **Selectable Boot**.

   c Press **Enter** to toggle the value to **Enabled**.

7 Use the cursor keys to select the Boot Port Name entry in the list of storage processors (SPs) and press **Enter** to open the Select Fibre Channel Device screen.

8 Use the cursor keys to select the specific SP and press **Enter**.

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

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

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

10 If any remaining storage processors show in the list, press **C** to clear the data.

11 Press **Esc** twice to exit and press **Enter** to save the setting.

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

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.
WWN	A World Wide Name formed according to Fibre Channel standards that uniquely identifies the FC adapter.
Targets	Number of targets accessed through the 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 plug-ins, storage devices available through the plug-ins also appear on the list.

For each storage adapter, you can display a separate list of storage devices 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. You can change this name to a name of your choice.
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 plug-in, such as the NMP or a third-party plug-in, the host uses to manage the storage device.
Hardware Acceleration	Information on whether the storage device assists the host with various virtual machine management operations. The status can be Supported, Not Supported, or Unknown.
Location	A path to the storage device in the `/vmfs/devices/` directory.
Partitions	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.

Runtime Name The name of the first path to the device. The runtime name is created by the host. The runtime name is not a reliable identifier for the device, and is not persistent.

The runtime name has the following format:

vmhba#:C#:T#:L#, where

- 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 plug-ins, storage devices available through the plug-ins 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 display all datastores available to your hosts and analyze their properties.

Datastores are added to the vSphere Client in the following ways:

- Created on an available storage device.
- Discovered when a host is added to the inventory. When you add a host to the inventory, the vSphere Client displays any datastores available to the host.

If your vSphere Client is connected to a vCenter Server system, you can see datastore information in the Datastores view. This view displays all datastores in the inventory, arranged by a datacenter. Through this view, you can organize datastores into folder hierarchies, create new datastores, edit their properties, or remove existing datastores.

This view is comprehensive and shows all information for your datastores including hosts and virtual machines using the datastores, storage reporting information, permissions, alarms, tasks and events, storage topology, and storage reports. Configuration details for each datastore on all hosts connected to this datastore are provided on the Configuration tab of the Datastores view.

NOTE The Datastores view is not available when the vSphere client connects directly to your host. In this case, review datastore information through the host storage configuration tab.

The following table describes the datastore details that you can see when you review datastores.

Table 6-3. Datastore Information

Datastore Information	Description
Identification	Editable friendly name that you assign to the datastore.
Device	Storage device, on which the datastore is deployed. If the datastore spans over multiple storage devices, only the first storage device is shown.
Capacity	Total formatted capacity of the datastore.
Free	Available space.
Type	File system that the datastore uses, either VMFS or NFS.
Storage I/O Control	Allows cluster-wide storage I/O prioritization. See the *vSphere Resource Management Guide*.
Hardware Acceleration	Information on whether the datastore assists the host with various virtual machine management operations. The status can be Supported, Not Supported, or Unknown.
Location	A path to the datastore in the `/vmfs/volumes/` directory.
Extents	Individual extents that the datastore spans and their capacity (VMFS datastores only).

Table 6-3. Datastore Information (Continued)

Datastore Information	Description
Path Selection	Path selection policy the host uses to access storage (VMFS datastores only).
Paths	Number of paths used to access storage and their status (VMFS datastores only).

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 Storage 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 6-4. 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.	■ The MASK_PATH plug-in allows you to prevent your host from accessing a specific storage array or specific LUNs on a storage array. If your host is detecting devices and paths that you do not want the host to access, path masking could have been set up incorrectly. ■ For booting from a SAN, ensure that each host sees only required LUNs. Do not allow any host to see any boot LUN other than its own. Use storage system software to make sure that the 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: ■ 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.

Datastore Refresh and Storage Rescan Operations

The datastore refresh operation updates the datastore lists and storage information, such as the datastore capacity, displayed in the vSphere Client. When you perform datastore management tasks or make changes in the SAN configuration, you might need to rescan your storage.

When you perform VMFS datastore management operations, such as creating a VMFS datastore or RDM, adding an extent, and increasing or deleting a VMFS datastore, your host or the vCenter Server automatically rescans and updates your storage. You can disable the automatic rescan feature by turning off the Host Rescan Filter. See "Turn off vCenter Server Storage Filters," on page 51.

In certain cases, you need to perform a manual rescan. You can rescan all storage available to your host, or, if you are using the vCenter Server, to all hosts in a folder, cluster, and datacenter.

If the changes you make are isolated to storage connected through a specific adapter, perform a rescan for this adapter.

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

- Zone a new disk array on a SAN.
- Create new LUNs on a SAN.
- Change the path masking on a host.
- Reconnect a cable.
- Add a single host to the vCenter Server after you have edited or removed from the vCenter Server a datastore shared by the vCenter Server hosts and the single host.

IMPORTANT If you rescan when a path is unavailable, the host removes the path from the list of paths to the device. The path reappears on the list as soon as it becomes available and starts working again.

Perform Storage Rescan

When you make changes in your SAN configuration, you might need to rescan your storage. You can rescan all storage available to your host. If the changes you make are isolated to storage accessed through a specific adapter, perform rescan for only this adapter.

Use this procedure if you want to limit the rescan to storage available to a particular host or accessed through a particular adapter on the host. If you want to rescan storage available to 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.

Turn off vCenter Server Storage Filters

When you perform VMFS datastore management operations, vCenter Server uses default storage filters. The filters help you to avoid storage corruption by retrieving only the storage devices, or LUNs, that can be used for a particular operation. Unsuitable LUNs are not displayed for selection. You can turn off the filters to view all LUNs.

Before making any changes to the LUN filters, consult with the VMware support team. You can turn off the filters only if you have other methods to prevent LUN corruption.

Procedure

1 In the vSphere Client, select **Administration > vCenter Server Settings**.

2 In the settings list, select **Advanced Settings**.

3 In the **Key** text box, type a key.

Key	Filter Name
config.vpxd.filter.vmfsFilter	VMFS Filter
config.vpxd.filter.rdmFilter	RDM Filter
config.vpxd.filter.SameHostAndTransportsFilter	Same Host and Transports Filter
config.vpxd.filter.hostRescanFilter	Host Rescan Filter NOTE If you turn off the Host Rescan Filter, your hosts continue to perform a rescan each time you present a new LUN to a host or a cluster.

4 In the **Value** text box, type **`False`** for the specified key.

5 Click **Add**.

6 Click **OK**.

You are not required to restart the vCenter Server system.

vCenter Server Storage Filtering

vCenter Server provides storage filters to help you avoid storage device corruption or performance degradation that can be caused by an unsupported use of LUNs. These filters are available by default.

Table 6-5. Storage Filters

Filter Name	Description	Key
VMFS Filter	Filters out storage devices, or LUNs, that are already used by a VMFS datastore on any host managed by vCenter Server. The LUNs do not show up as candidates to be formatted with another VMFS datastore or to be used as an RDM.	config.vpxd.filter.vmfsFilter
RDM Filter	Filters out LUNs that are already referenced by an RDM on any host managed by vCenter Server. The LUNs do not show up as candidates to be formatted with VMFS or to be used by a different RDM. If you need virtual machines to access the same LUN, the virtual machines must share the same RDM mapping file. For information about this type of configuration, see *Setup for Failover Clustering and Microsoft Cluster Service*.	config.vpxd.filter.rdmFilter

Table 6-5. Storage Filters (Continued)

Filter Name	Description	Key
Same Host and Transports Filter	Filters out LUNs ineligible for use as VMFS datastore extents because of host or storage type incompatibility. Prevents you from adding the following LUNs as extents: ■ LUNs not exposed to all hosts that share the original VMFS datastore. ■ LUNs that use a storage type different from the one the original VMFS datastore uses. For example, you cannot add a Fibre Channel extent to a VMFS datastore on a local storage device.	config.vpxd.filter.SameHostAndTransportsFilter
Host Rescan Filter	Automatically rescans and updates VMFS datastores after you perform datastore management operations. The filter helps provide a consistent view of all VMFS datastores on all hosts managed by vCenter Server. NOTE If you present a new LUN to a host or a cluster, the hosts automatically perform a rescan no matter whether you have the Host Rescan Filter on or off.	config.vpxd.filter.hostRescanFilter

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** under Software.

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** under Software.

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

Requirements for Using NPIV

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

The following requirements exist:

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

  For information, see the *VMware Compatibility Guide* and refer to you vendor documentation.

  - Use HBAs of the same type, either all QLogic or all Emulex. VMware does not support heterogeneous HBAs on the same host accessing the same LUNs.
  - If a host uses multiple physical HBAs as paths to the storage, zone all physical paths to the virtual machine. This is required to support multipathing even though only one path at a time will be active.
  - Make sure that physical HBAs on the host have access to all LUNs that are to be accessed by NPIV-enabled virtual machines running on that host.
- The switches in the fabric must be NPIV-aware.
- When configuring a LUN for NPIV access at the storage level, make sure that the NPIV LUN number and NPIV target ID match the physical LUN and Target ID.
- Use the vSphere Client to manipulate virtual machines with WWNs.

NPIV Capabilities and Limitations

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

ESX/ESXi with NPIV supports the following items:

- NPIV supports vMotion. When you use vMotion to migrate a virtual machine it retains the assigned WWN.

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

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

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

Assign WWNs to Virtual Machines

You can assign a WWN to a new virtual machine with an RDM disk when you create this virtual machine.

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

Procedure

1 Open the New Virtual Machine wizard.

2 Select **Custom**, and click **Next**.

3 Follow all steps required to create a custom virtual machine.

4 On the Select a Disk page, select **Raw Device Mapping**, and click **Next**.

5 From a list of SAN disks or LUNs, select a raw LUN you want your virtual machine to access directly.

6 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.

---

7 Follow the steps required to create a virtual machine with the RDM.

8 On the Ready to Complete page, select the **Edit the virtual machine settings before completion** check box and click **Continue**.

The Virtual Machine Properties dialog box opens.

9 Assign WWNs to the virtual machine.

a Click the **Options** tab, and select **Fibre Channel NPIV**.

b Select **Generate new WWNs**.

c Specify the number of WWNNs and WWPNs.

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

10 Click **Finish**.

The host creates WWN assignments for the virtual machine.

What to do next

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

Modify WWN Assignments

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

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

Prerequisites

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

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

Procedure

1 Open the Virtual Machine Properties dialog box by clicking the **Edit Settings** link for the selected virtual machine.

2 Click the **Options** tab and select **Fibre Channel NPIV**.

The Virtual Machine Properties dialog box opens.

3 Edit the WWN assignments by selecting one of the following options:

Option	Description
Temporarily disable NPIV for this virtual machine	Disable the WWN assignments for the virtual machine.
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.

4 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 plug-in (MPP) should claim the paths to a particular device and become responsible for managing the multipathing support for the device.

By default, the host performs a periodic path evaluation every 5 minutes causing any unclaimed paths to be claimed by the appropriate MPP.

The claim rules are numbered. For each physical path, the host runs through the claim rules starting with the lowest number first. The attributes of the physical path are compared to the path specification in the claim rule. If there is a match, the host assigns the MPP specified in the claim rule to manage the physical path. This continues until all physical paths are claimed by corresponding MPPs, either third-party multipathing plug-ins or the native multipathing plug-in (NMP).

For general information on multipathing plug-ins, see "Managing Multiple Paths," on page 24.

For the paths managed by the NMP module, a second set of claim rules is applied. These rules determine which Storage Array Type Plug-In (SATP) should be used to manage the paths for a specific array type, and which Path Selection Plug-In (PSP) is to be used for each storage device. For example, for a storage device that belongs to the EMC CLARiiON CX storage family and is not configured as ALUA device, 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 "Managing Storage Paths and Multipathing Plug-Ins," on page 77.

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

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.

Dead — The software cannot connect to the disk through this path.

If you are using the **Fixed**path policy, you can see which path is the preferred path. The preferred path is marked with an asterisk (*) in the Preferred column.

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 (VMW_PSP_FIXED)
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.

Fixed AP (VMW_PSP_FIXED_AP)
Extends the Fixed functionality to active-passive and ALUA mode arrays.

Most Recently Used (VMW_PSP_MRU)
The host selects the path that it used recently. When the path becomes unavailable, the host selects an alternative path. The host does not revert back to the original path when that path becomes available again. There is no preferred path setting with the MRU policy. MRU is the default policy for active-passive storage devices.

Round Robin (VMW_PSP_RR)
The host uses an automatic path selection algorithm rotating through all active paths when connecting to active-passive arrays, or through all available paths when connecting to active-active arrays. This implements load balancing across the physical paths available to your host.

Load balancing is the process of spreading I/O requests across the paths. The goal is to optimize performance in terms of throughput, such as I/O per second, megabytes per second, or response times.

Table 6-6 summarizes how the behavior of host changes, depending on the type of array and the failover policy.

Table 6-6. 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.
Fixed AP	For ALUA arrays, VMkernel picks the path set to be the preferred path. For both A/A and A/P and ALUA arrays, VMkernel resumes using the preferred path, but only if the path-thrashing avoidance algorithm allows the fail-back.	

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 (VMW_PSP_FIXED)**
- **Fixed AP (VMW_PSP_FIXED_AP)**
- **Most Recently Used (VMW_PSP_MRU)**
- **Round Robin (VMW_PSP_RR)**

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. If your array supports the ALUA protocol, your host uses that protocol to perform load balancing. If your array does not support the ALUA protocol and you want your host to do automatic load balancing, configure your devices to use the Round Robin PSP. If you want to do the load balancing yourself, you can perform manual load balancing.

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.

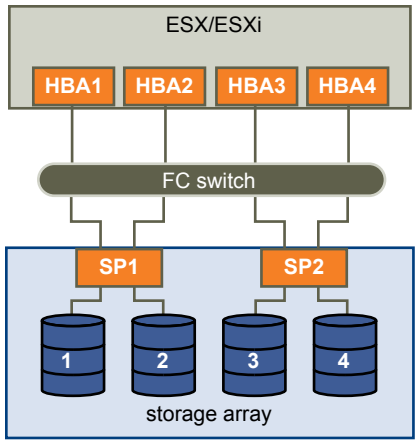
You can configure your ESX/ESXi host to load balance traffic across multiple adapters by assigning preferred paths to your LUNs. Path policy must be set to Fixed.

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.

You can use the VMW_PSP_FIXED_AP path selection policies to do static path load balancing on active-passive arrays.

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 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.

Sharing Diagnostic Partitions

Generally, you use the local disk 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 diagnostic partition.

Disable Automatic Host Registration

When you use EMC CLARiiON or Invista arrays for storage, it is required that the hosts 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 using storage management software, disable 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 click **Advanced Settings** under Software.

3 Click **Disk** in the left panel and scroll down to Disk.EnableNaviReg on the right.

4 Change the default value to 0.

This 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.
- 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 "Perform Storage Rescan," on page 50.

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 the *vSphere Command-Line Interface Installation and Scripting 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). This can happen in two ways.
  - 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 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 or masking inconsistencies between different ESX/ESXi hosts and SAN targets so that all HBAs see the same targets.
3. Configure the path to use the Most Recently Used PSP (the default) or VMW_PSP_FIXED_AP.

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.

NOTE ESX/ESXi uses the SCSI reservations mechanism only when a LUN is not VAAI capable. If a LUN is VAAI capable and supports Hardware Acceleration, ESX/ESXi uses atomic test and set (ATS) algorithm to lock the LUN.

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:
`vicfg-module --server` ***`server`*** **`-l |grep qla`**.

2 Run the following command.

Use the appropriate module based on the outcome of the previous step.

```
vicfg-module --server server -s ql2xmaxqdepth=64 qla2xxx
```

In this example, the queue depth is 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.

You can adjust the maximum queue depth for an Emulex HBA using vSphere CLI.

Procedure

1 Verify which Emulex HBA module is currently loaded by entering the
`vicfg-module --server` ***`server`*** **`-l | grep lpfc`** command.

2 Run the following command.

The example shows the `lpfc820` module. Use the appropriate module based on the outcome of Step 1.

```
vicfg-module --server server -s lpfc0_lun_queue_depth=16 lpfc820
```

In this example, the HBA will have its LUN queue depth set to 16.

3 Reboot your 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 copy only if it does not collide with the original VMFS datastore that has the same UUID. To mount the copy, the original VMFS datastore has to be offline.

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

You cannot unmount an active mounted datastore.

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 `.vmx`, `.vmdk`, `.vmsd`, and `.vmsn`.
- To power on virtual machines, register them with vCenter Server.

Storage Hardware Acceleration

The hardware acceleration functionality enables your host to offload specific virtual machine and storage management operations to compliant storage hardware. With the storage hardware assistance, your host performs these operations faster and consumes less CPU, memory, and storage fabric bandwidth.

To implement the hardware acceleration functionality, the Pluggable Storage Architecture (PSA) uses a combination of special array integration plug-ins, called VAAI plug-ins, and an array integration filter, called VAAI filter. The PSA automatically attaches the VAAI filter and vendor-specific VAAI plug-ins to those storage devices that support the hardware acceleration.

To view and manage the VAAI filter and VAAI plug-ins available on your host, use the vSphere CLI commands.

You can find some information about managing the VAAI filter and VAAI plug-ins in "Managing Hardware Acceleration Filter and Plug-Ins," on page 84.

For descriptions of the commands, see the *vSphere Command-Line Interface Installation and Scripting Guide* and the *vSphere Command-Line Interface Reference*.

Hardware Acceleration Requirements and Benefits

The hardware acceleration functionality works only if you use an appropriate host and storage array combination.

Use the following hosts and storage arrays:

- ESX/ESXi version 4.1 or later.
- Storage arrays that support storage-based hardware acceleration. ESX/ESXi version 4.1 does not support hardware acceleration with NAS storage devices.

On your host, the hardware acceleration is enabled by default. To enable the hardware acceleration on the storage side, check with your storage vendor. Certain storage arrays require that you explicitly activate the hardware acceleration support on the storage side.

When the hardware acceleration functionality is supported, the host can get hardware assistance and perform the following operations faster and more efficiently:

- Migration of virtual machines with Storage vMotion
- Deployment of virtual machines from templates
- Cloning of virtual machines or templates
- VMFS clustered locking and metadata operations for virtual machine files
- Writes to thin provisioned and thick virtual disks
- Creation of fault-tolerant virtual machines

Hardware Acceleration Support Status

For each storage device and datastore, the vSphere Client displays the hardware acceleration support status in the Hardware Acceleration column of the Devices view and the Datastores view.

The status values are Unknown, Supported, and Not Supported. The initial value is Unknown. The status changes to Supported after the host successfully performs the offload operation. If the offload operation fails, the status changes to Not Supported.

When storage devices do not support or provide only partial support for the host operations, your host reverts to its native methods to perform unsupported operations.

Turn off Hardware Acceleration

If your storage devices do not support the hardware acceleration functionality, you can turn it off by using the vSphere Client advanced settings.

As with any advanced settings, before disabling the hardware acceleration, consult with the VMware support team.

Procedure

1 In the vSphere Client inventory panel, select the host.

2 Click the **Configuration** tab, and click **Advanced Settings** under **Software.**

3 Click VMFS3 and change the value in the **VMFS3.HardwareAcceleratedLocking** field to zero.

4 Click **DataMover** and change the values in each of the following fields to zero:

- **DataMover.HardwareAcceleratedMove**
- **DataMover.HardwareAcceleratedInit**

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

Multipathing Checklist

A

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: ■ 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	Set the EMC Clariion failover mode to 1 or 4. Contact EMC for details.
HP MSA	Host type must be `Linux`. Set the connection type for each HBA port to `Linux`.
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	■ 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	■ For all LUNs hosting clustered disks on active-passive arrays, use the Most Recently Used or VMW_PSP_FIXED_AP PSP.
	■ For LUNs on active-active arrays, you can use the Most Recently Used or Fixed PSP.
	■ With either active-passive or active-active arrays, you can use the Round Robin PSP.
	■ All FC HBAs must be of the same model.
	■ Set the following Software Advanced Settings for the ESX/ESXi host:
	  ■ Set Disk.UseLunReset to 1
	  ■ Set Disk.UseDeviceReset to 0

Managing Multipathing Modules and Hardware Acceleration Plug-Ins

B

Use the vSphere CLI to manage the Pluggable Storage Architecture (PSA) multipathing plug-ins and Hardware Acceleration plug-ins.

This appendix includes the following topics:

Managing Storage Paths and Multipathing Plug-Ins

Using the vSphere CLI you can manage the PSA multipathing plug-ins and storage paths assigned to them.

You can display all multipathing plug-ins available on your host. You can list any third-party MPPs, as well as your host's NMP and SATPs and review the paths they claim. You can also define new paths and specify which multipathing plug-in should claim the paths.

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

List Multipathing Claim Rules for the Host

Use the vSphere CLI to list available multipathing claim rules.

Claim rules indicate which multipathing plug-in, the NMP or any third-party MPP, manages a given physical path. Each claim rule identifies a set of paths based on the following parameters:

- Vendor/model strings
- Transportation, such as SATA, IDE, Fibre Channel, and so on
- Adapter, target, or LUN location
- Device driver, for example, Mega-RAID

Procedure

- Use the `esxcli corestorage claimrule list --claimrule-class=MP` to list the multipathing claim rules.

  "Example: Sample Output of the esxcli corestorage claimrule list Command," on page 78 shows the output of the command.

Example: Sample Output of the esxcli corestorage claimrule list Command

```
Rule Class  Rule   Class    Type       Plugin     Matches
MP          0      runtime  transport  NMP        transport=usb
MP          1      runtime  transport  NMP        transport=sata
MP          2      runtime  transport  NMP        transport=ide
MP          3      runtime  transport  NMP        transport=block
MP          4      runtime  transport  NMP        transport=unknown
MP          101    runtime  vendor     MASK_PATH  vendor=DELL model=Universal Xport
MP          101    file     vendor     MASK_PATH  vendor=DELL model=Universal Xport
MP          200    runtime  vendor     MPP_1      vendor=NewVend model=*
MP          200    file     vendor     MPP_1      vendor=NewVend model=*
MP          201    runtime  location   MPP_2      adapter=vmhba41 channel=* target=* lun=*
MP          201    file     location   MPP_2      adapter=vmhba41 channel=* target=* lun=*
MP          202    runtime  driver     MPP_3      driver=megaraid
MP          202    file     driver     MPP_3      driver=megaraid
MP          65535  runtime  vendor     NMP        vendor=* model=*
```

This example indicates the following:

- The NMP claims all paths connected to storage devices that use the USB, SATA, IDE, and Block SCSI transportation.
- The MASK_PATH module by default 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 Rule Class column in the output describes the category of a claim rule. It can be MP (multipathing plug-in), Filter, or VAAI.
- The Class column shows which rules are defined and which are loaded. The `file` parameter in the Class column indicates that the rule is defined. The `runtime` parameter indicates that the rule has been loaded into your system. For a user-defined claim rule to be active, two lines with the same rule number should exist, one line for the rule with the `file` parameter and another line with `runtime`. Several low numbered rules have only one line with the Class of `runtime`. These are system-defined claim rules that you cannot modify.

Display Multipathing Modules

Use the 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:

  ```
  esxcli corestorage plugin list --plugin-class=MP
  ```

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

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.

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 Multipathing Claim Rules

Use the vSphere CLI to add a new multipathing PSA claim rule to the set of claim rules on the system. For the new claim rule to be active, you first define the rule and then load it into your system.

You add a new PSA claim rule when, for example, you load a new multipathing plug-in (MPP) and need to define which paths this module should claim. You may need to create a claim rule if you add new paths and want an existing MPP to claim them.

CAUTION When creating new claim rules, be careful to avoid a situation 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 problems.

Procedure

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

```
esxcli corestorage claimrule add
```

For information on the options that the command requires, see "esxcli corestorage claimrule Options," on page 87.

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

```
esxcli corestorage claimrule load
```

This command loads all newly created multipathing claim rules from your system's configuration file.

Example: Defining Multipathing Claim Rules

- Add rule # 500 to claim all paths with the NewMod model string and the NewVend vendor string for the NMP plug-in.

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

After you load the claim rule and 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` and another with the Class of `file`, indicate that the new claim rule has been loaded into the system and is active.

```
Rule Class  Rule   Class    Type       Plugin     Matches
MP          0      runtime  transport  NMP        transport=usb
MP          1      runtime  transport  NMP        transport=sata
MP          2      runtime  transport  NMP        transport=ide
MP          3      runtime  transport  NMP        transport=block
MP          4      runtime  transport  NMP        transport=unknown
MP          101    runtime  vendor     MASK_PATH  vendor=DELL model=Universal Xport
MP          101    file     vendor     MASK_PATH  vendor=DELL model=Universal Xport
MP          500    runtime  vendor     NMP        vendor=NewVend model=NewMod
MP          500    file     vendor     NMP        vendor=NewVend model=NewMod
```

- Add rule # 321 to claim the path on adapter vmhba0, channel 0, target 0, LUN 0 for the NMP plug-in.

```
# esxcli corestorage claimrule add -r 321 -t location -A vmhba0 -C 0 -T 0 -L 0 -P NMP
```

- Add rule # 1015 to claim all paths provided by Fibre Channel adapters for the NMP plug-in.

```
# esxcli corestorage claimrule add -r 1015 -t transport -R fc -P NMP
```

- Add a rule with a system assigned rule id to claim all paths provided by Fibre Channel type adapters for the NMP plug-in.

```
# esxcli corestorage claimrule add --autoassign -t transport -R fc -P NMP
```

Delete Multipathing Claim Rules

Use the vSphere CLI to remove a multipathing 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 claimrule Options," on page 87.

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 plug-in 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 plug-in to a path by creating a new claim rule for the plug-in.

```
esxcli corestorage claimrule add –P MASK_PATH
```

For information on command-line options, see "esxcli corestorage claimrule Options," on page 87.

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
```

6 Run the path claiming rules.

```
esxcli corestorage claimrule run
```

After you assign the MASK_PATH plug-in to a path, the path state becomes irrelevant and is no longer maintained by the host. As a result, commands that display the masked path's information might show the path state as dead.

Example: Masking a LUN

In this example, you mask the LUN 20 on targets T1 and T2 accessed through storage adapters vmhba2 and vmhba3.

1 `#esxcli 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 Delete the MASK_PATH claim rule.

esxcli *conn_options* corestorage claimrule delete –r *rule#*

2 Verify that the claim rule was deleted correctly.

esxcli *conn_options* corestorage claimrule list

3 Reload the path claiming rules from the configuration file into the VMkernel.

esxcli *conn_options* corestorage claimrule load

4 Run the `esxcli corestorage claiming unclaim` command for each path to the masked storage device.

For example:

esxcli *conn_options* corestorage claiming unclaim –t location –A vmhba0 –C 0 –T 0 –L 149

5 Run the path claiming rules.

esxcli *conn_options* 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 an 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 `esxcli nmp satp addrule` command. The command takes the following options.

Option	Description
-c|--claim-option	Set the claim option string when adding a 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.
-e|--description	Set the claim rule description when adding a SATP claim rule.
-d|--device	Set the device when adding SATP claim rules. Device rules are mutually exclusive with vendor/model and driver rules.
-D|--driver	Set the driver string when adding a SATP claim rule. Driver rules are mutually exclusive with vendor/model rules.
-f|--force	Force claim rules to ignore validity checks and install the rule anyway.
-h|--help	Show the help message.
-M|--model	Set the model string when adding SATP a claim rule. Vendor/Model rules are mutually exclusive with driver rules.
-o|--option	Set the option string when adding a SATP claim rule.
-P|--psp	Set the default PSP for the SATP claim rule.
-O|--psp-option	Set the PSP options for the SATP claim rule.
-s|--satp	The SATP for which a new rule will be added.
-R|--transport	Set the claim transport type string when adding a SATP claim rule.
-V|--vendor	Set the vendor string when adding SATP claim rules. Vendor/Model rules are mutually exclusive with driver rules.

NOTE When searching the SATP rules to locate a SATP for a given device, the NMP searches the driver rules first. If there is no match, the vendor/model rules are searched, and finally the transport rules. If there is still no match, NMP selects a default SATP for the device.

2 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
```

3 Reboot your host.

Example: Defining an NMP SATP Rule

The following sample command assigns the VMW_SATP_INV plug-in to manage storage arrays with vendor string NewVend and model string NewMod.

```
# esxcli 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.

Managing Hardware Acceleration Filter and Plug-Ins

The hardware acceleration, or VAAI, filter in combination with vendor-specific VAAI plug-ins are attached to storage devices that support the hardware acceleration. Using the vSphere CLI you can display and manipulate the VAAI filter and VAAI plug-ins.

Display Hardware Acceleration Filter

Use the vSphere CLI to view the hardware acceleration, or VAAI, filter currently loaded into your system.

Procedure

- Run the `esxcli corestorage plugin list --plugin-class=Filter` command.

The output of this command is similar to the following:

```
Plugin name  Plugin class
VAAI_FILTER  Filter
```

Display Hardware Acceleration Plug-Ins

Use the vSphere CLI to view hardware acceleration plug-ins, also called VAAI plug-ins, currently loaded into your system.

Procedure

- Run the `esxcli corestorage plugin list --plugin-class=VAAI` command.

The output of this command is similar to the following:

```
Plugin name         Plugin class
VMW_VAAIP_EQL       VAAI
VMW_VAAIP_NETAPP    VAAI
VMW_VAAIP_CX        VAAI
```

Verify Hardware Acceleration Status of a Storage Device

Use the vSphere CLI to verify the hardware acceleration support status of a particular storage device. This command also helps to determine which VAAI filter is attached to the device.

Procedure

- Run the **`esxcli corestorage device list --d`** ***`device_ID`*** command.

The output shows the hardware acceleration, or VAAI, status that can be unknown, supported, or unsupported. If the device supports the hardware acceleration, the output also lists the VAAI filter attached to the device.

```
# esxcli corestorage device list --d naa.60a98000572d43595a4a52644473374c
naa.60a98000572d43595a4a52644473374c
 Display Name: NETAPP Fibre Channel Disk(naa.60a98000572d43595a4a52644473374c)
 Size: 20480
 Device Type: Direct-Access
 Multipath Plugin: NMP
 Devfs Path: /vmfs/devices/disks/naa.60a98000572d43595a4a52644473374c
 Vendor: NETAPP
 Model: LUN
 Revision: 8000
 SCSI Level: 4
 Is Pseudo: false
 Status: on
```

```
Is RDM Capable: true
Is Local: false
Is Removable: false
Attached Filters: VAAI_FILTER
VAAI Status: supported
Other UIDs: vml.020003000060a98000572d43595a4a52644473374c4c554e202020
```

View Hardware Acceleration Plug-In for a Device

Use the vSphere CLI to view the hardware acceleration, or VAAI, plug-in attached to a storage device that supports the hardware acceleration.

Procedure

- Run the **esxcli vaai device list --d *device_ID*** command.

  For example:

  ```
  # esxcli vaai device list –d naa.6090a028d00086b5d0a4c44ac672a233
  naa.6090a028d00086b5d0a4c44ac672a233
   Device Display Name: EQLOGIC iSCSI Disk (naa.6090a028d00086b5d0a4c44ac672a233)
   VAAI Plugin Name: VMW_VAAIP_EQL
  ```

List Hardware Acceleration Claim Rules

For each storage device that supports the hardware acceleration functionality, the claim rules specify the hardware acceleration filter and the hardware acceleration plug-in to manage this storage device. You can use the vSphere CLI to list the hardware acceleration filter and plug-in claim rules.

Procedure

1 To list the filter claim rules, run the `esxcli corestorage claimrule list --claimrule-class=Filter` command.

  In this example, the filter claim rules specify devices that should be claimed by the VAAI_FILTER filter.

  ```
  # esxcli corestorage claimrule list --claimrule-class=Filter
  Rule Class   Rule   Class   Type   Plugin      Matches
  Filter       65430  runtime vendor VAAI_FILTER vendor=EMC model=SYMMETRIX
  Filter       65430  file    vendor VAAI_FILTER vendor=EMC model=SYMMETRIX
  Filter       65431  runtime vendor VAAI_FILTER vendor=DGC model=*
  Filter       65431  file    vendor VAAI_FILTER vendor=DGC model=*
  ```

2 To list the VAAI plug-in claim rules, run the `esxcli corestorage claimrule list --claimrule-class=VAAI` command.

  In this example, the VAAI claim rules specify devices that should be claimed by a particular VAAI plug-in.

  ```
  esxcli corestorage claimrule list --claimrule-class=VAAI
  Rule Class   Rule   Class   Type    Plugin         Matches
  VAAI         65430  runtime vendor  VMW_VAAIP_SYMM vendor=EMC model=SYMMETRIX
  VAAI         65430  file    vendor  VMW_VAAIP_SYMM vendor=EMC model=SYMMETRIX
  VAAI         65431  runtime vendor  VMW_VAAIP_CX   vendor=DGC model=*
  VAAI         65431  file    vendor  VMW_VAAIP_CX   vendor=DGC model=*
  ```

Add Hardware Acceleration Claim Rules

To configure hardware acceleration for a new array, you need to add two claim rules, one for the VAAI filter and another for the VAAI plug-in. For the new claim rules to be active, you first define the rules and then load them into your system.

Procedure

1 Define a new claim rule for the VAAI filter by using the `esxcli corestorage claimrule add --claimrule-class=Filter --plugin=VAAI_FILTER` command.

For information about the options that the command requires, see "esxcli corestorage claimrule Options," on page 87.

2 Define a new claim rule for the VAAI plug-in by using the `esxcli corestorage claimrule add --claimrule-class=VAAI` command.

3 Load both claim rules by using the following commands:

```
esxcli corestorage claimrule load --claimrule-class=Filter
esxcli corestorage claimrule load --claimrule-class=VAAI
```

4 Run the VAAI filter claim rule by using the `esxcli corestorage claimrule run --claimrule-class=Filter` command.

NOTE Only the Filter-class rules need to be run. When the VAAI filter claims a device, it automatically finds the proper VAAI plug-in to attach.

Example: Defining Hardware Acceleration Claim Rules

To configure Hardware Acceleration for IBM arrays using the VMW_VAAI_T10 plug-in, use the following sequence of commands:

```
# esxcli corestorage claimrule add --claimrule-class=Filter --plugin=VAAI_FILTER --type=vendor --vendor=IBM --autoassign
# esxcli corestorage claimrule add --claimrule-class=VAAI --plugin=VMW_VAAI_T10 --type=vendor --vendor=IBM --autoassign
# esxcli corestorage claimrule load --claimrule-class=Filter
# esxcli corestorage claimrule load --claimrule-class=VAAI
# esxcli corestorage claimrule run --claimrule-class=Filter
```

Delete Hardware Acceleration Claim Rules

Use the vSphere CLI to delete existing hardware acceleration claim rules.

Procedure

- Use the following commands:

```
esxcli corestorage claimrule delete -r claimrule_ID --claimrule-class=Filter
esxcli corestorage claimrule delete -r claimrule_ID --claimrule-class=VAAI
```

esxcli corestorage claimrule Options

Certain `esxcli corestorage claimrule` commands, for example the commands that you run to add new claim rules, remove the rules, or mask paths, require that you specify a number of options.

Table B-1. esxcli corestorage claimrule Options

Option	Description
-A|--adapter	Indicate the adapter of the paths to use in this operation.
-u|--autoassign	The system will auto assign a rule ID.
-C|--channel	Indicate the channel of the paths to use in this operation.
-c|--claimrule-class	Indicate the claim rule class to use in this operation. Valid values are: MP, Filter, VAAI
-d|--device	Indicate the device Uid to use for this operation.
-D|--driver	Indicate the driver of the paths to use in this operation.
-f|--force	Force claim rules to ignore validity checks and install the rule anyway.
-h|--help	Show the help message.
-L|--lun	Indicate the LUN of the paths to use in this operation.
-M|--model	Indicate the model of the paths to use in this operation.
-P|--plugin	Indicate which PSA plug-in to use for this operation.
-r|--rule	Indicate the claim rule ID to use for this operation.
-T|--target	Indicate the target of the paths to use in this operation.
-R|--transport	Indicate the transport of the paths to use in this operation. Valid values are: block, fc, iscsi, iscsivendor, ide, sas, sata, usb, parallel, unknown.
-t|--type	Indicate which type of matching is used for claim/unclaim or claimrule. Valid values are: vendor, location, driver, transport, device.
-V|--vendor	Indicate the vendor of the paths to use in this operation.

Index

N

O

P

Q

R

W

X

Z