iSCSI SAN Configuration Guide

Update 1
ESX 4.0
ESXi 4.0
vCenter Server 4.0

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

EN-000267-03

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

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

Revision	Description
EN-000267-03	■ Updated "EMC CLARiiON Storage Systems," on page 46 to include information on ESX/ESXi support of EMC CLARiiON with ALUA. ■ Updated Appendix A, "iSCSI SAN Configuration Checklist," on page 81 to include information on failover parameters for EMC CLARiiON.
EN-000267-02	Updated the following topics to include information about port binding on EMC CLARiiON: "Networking Configuration for Software iSCSI Storage," on page 32 and "EMC CLARiiON Storage Systems," on page 46.
EN-000267-01	Updated commands in Step 2 in the section "Create a Jumbo Frames-Enabled VMkernel Interface," on page 37.
EN-000267-00	Initial release.

About This Book

The *iSCSI SAN Configuration Guide* explains how to use VMware® ESX™ and VMware ESXi systems with an iSCSI storage area network (SAN). The manual includes conceptual background information and installation requirements for ESX, ESXi, and VMware vCenter™ Server.

Intended Audience

This manual is written for experienced Windows or Linux system administrators who are familiar with virtual machine technology datacenter operations.

Document Feedback

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

VMware vSphere Documentation

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

Technical Support and Education Resources

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

Online and Telephone Support

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

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

Support Offerings

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

VMware Professional Services

VMware Education Services courses offer extensive hands-on labs, case study examples, and course materials designed to be used as on-the-job reference tools. Courses are available onsite, in the classroom, and live online. For onsite pilot programs and implementation best practices, VMware Consulting

Services provides offerings to help you assess, plan, build, and manage your virtual environment. To access information about education classes, certification programs, and consulting services, go to http://www.vmware.com/services.

Using ESX/ESXi with an iSCSI Storage Area Network 1

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

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

This chapter includes the following topics:

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.

Network Virtualization

The virtualization layer guarantees that each virtual machine is isolated from other virtual machines. Virtual machines can talk to 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 so that some virtual machines can connect to the outside, while others are connected only through virtual networks to other virtual machines.

Storage Virtualization

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

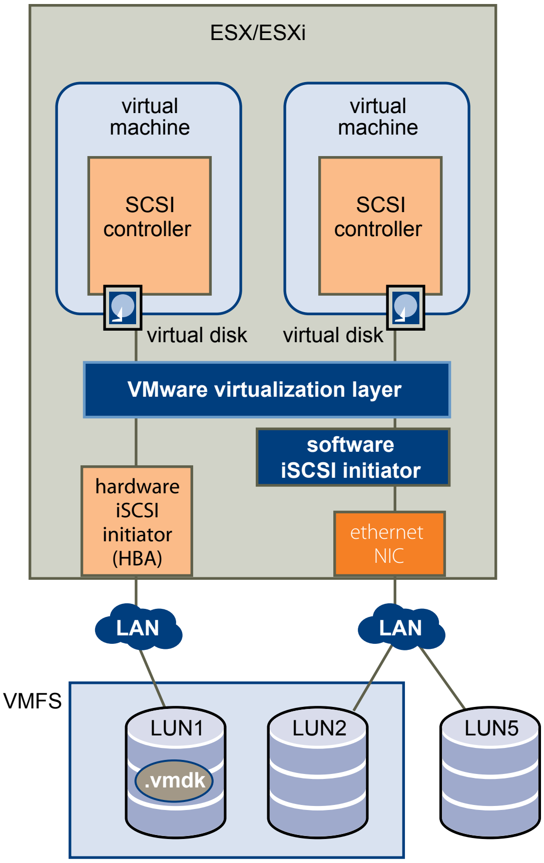
An ESX/ESXi virtual machine uses a virtual hard disk to store its operating system, program files, and other data associated with its activities. A virtual disk is a large physical file, or a set of files, that can be copied, moved, archived, and backed up as easily as any other file. You can configure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

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

Each virtual disk that a virtual machine can access through one of the virtual SCSI controllers resides in the VMware Virtual Machine File System (VMFS) datastore, NFS-based datastore, or on a raw disk. From the standpoint of the virtual machine, each virtual disk appears as if it were a SCSI drive connected to a SCSI controller. Whether the actual physical disk device is being accessed through parallel SCSI, iSCSI, network, or Fibre Channel adapters on the host is transparent to the guest operating system and to applications running on the virtual machine.

Figure 1-1 gives an overview of storage virtualization. The diagram illustrates storage that uses VMFS and storage that uses raw device mapping. The diagram also shows how iSCSI storage is accessed through either iSCSI HBAs or by using a general-purpose NIC that uses iSCSI initiator software.

Figure 1-1. iSCSI SAN Storage Virtualization



iSCSI SAN Concepts

If you are an ESX/ESXi administrator who plans to set up ESX/ESXi hosts to work with SANs, you must have a working knowledge of SAN concepts. You can find information about SAN in print and on the Internet. If you are new to iSCSI SAN technology, read the following sections to familiarize yourself with the basic terminology this document uses.

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

To transfer traffic from host servers to shared storage, the SAN uses the iSCSI protocol that packages SCSI commands into iSCSI packets and transmits them on an Ethernet network.

iSCSI Initiators

To access remote targets, your ESX/ESXi host uses iSCSI initiators. Initiators transport SCSI requests and responses between the ESX/ESXi system and the target storage device on the IP network.

ESX/ESXi supports hardware-based and software-based iSCSI initiators:

Hardware iSCSI Initiator

Uses a specialized iSCSI HBA. The hardware iSCSI initiator is responsible for all iSCSI and network processing and management.

Software iSCSI Initiator

Code built into the VMkernel that allows an ESX/ESXi to connect to the iSCSI storage device through standard network adapters. The software initiator handles iSCSI processing while communicating with the network adapter. With the software initiator, you can use iSCSI technology without purchasing specialized hardware.

Ports in the iSCSI SAN

In the context of this document, a port is an end point of the connection from a device into the iSCSI SAN. Each node in the iSCSI SAN, a host, storage device, and Ethernet switch has one or more ports that connect it to the SAN. Ports are identified in a number of ways.

IP Address

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

iSCSI Name

A worldwide unique name for identifying the port. The iSCSI name starts with either `iqn.` (for iSCSI qualified name) or `eui.` (for extended unique identifier). Multiple iSCSI devices can be present, with multiple iSCSI names, and can be connected through a single physical Ethernet port.

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

iSCSI Alias

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

Multipathing and Path Failover

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

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

Storage System Types

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

ESX/ESXi supports the following types of storage systems:

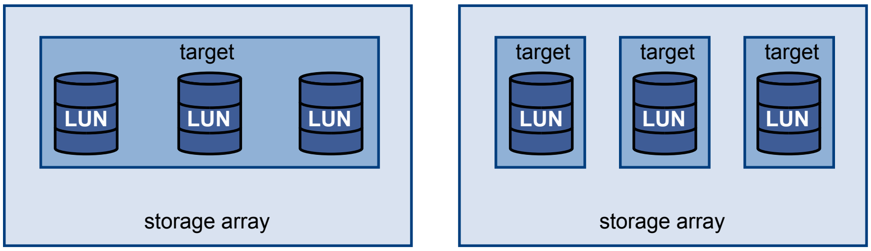
- An active-active storage system, which allows access to the LUNs simultaneously through all the storage ports that are available without significant performance degradation. All the paths are active at all times, unless a path fails.
- An active-passive storage system, in which one port is actively providing access to a given LUN. The other ports act as backup for the LUN and can be actively providing access to other LUN I/O. I/O can be successfully sent only to an active port for a given LUN. If access through the primary storage port fails, one of the secondary ports or storage processors becomes active, either automatically or through administrator intervention.
- A virtual port storage system, which allows access to all available LUNs through a single virtual port. These are active-active storage devices, but hide their multiple connections though a single port. The ESX/ESXi multipathing cannot detect the multiple connections to the storage. These storage systems handle port failover and connection balancing transparently. This is often referred to as transparent failover.

Target Compared to LUN Representations

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

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

Figure 1-2. Target Compared to LUN Representations



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

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

iSCSI Naming Conventions

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

iSCSI names are formatted in two different ways. The first is by an iSCSI qualified name, commonly referred to as an IQN name. The second, much less common method, is through an enterprise unique identifier, also referred to as an EUI name.

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

iSCSI Qualified Names

iSCSI qualified names take the form `iqn.yyyy-mm.naming-authority:unique name`, where:

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

Enterprise Unique Identifiers

Enterprise unique identifiers take the form `eui.`*`<16 hex digits>`*.

For example, `eui.0123456789ABCDEF`.

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

In many cases, the IQN format is chosen over the EUI format for readability and as a more user-friendly method of assigning names.

Discovery, Authentication, and Access Control

You can use several mechanisms to limit which volumes on an iSCSI storage system your ESX/ESXi host can access.

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

Discovery

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

Authentication

iSCSI storage systems authenticate an initiator by a name and key pair. ESX/ESXi supports the CHAP protocol, which VMware recommends for your SAN implementation. The ESX/ESXi host and the iSCSI storage system must have CHAP enabled and have common credentials. In the iSCSI login phrase, the iSCSI storage system exchanges and checks these credentials.

Access Control

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

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

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

Error Correction

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

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

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

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

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

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

Overview of Using ESX/ESXi with a SAN

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

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

- You can store data redundantly and configure multiple paths to your storage, eliminating a single point of failure. ESX/ESXi systems provide multipathing by default for every virtual machine.
- Using a SAN with ESX/ESXi systems extends failure resistance to the server. When you use SAN storage, all applications can instantly be restarted after host failure.
- You can perform live migration of virtual machines using VMware VMotion.
- Use VMware High Availability (HA) in conjunction with a SAN for a cold-standby solution that guarantees an immediate, automatic response.

- Use VMware Distributed Resource Scheduler (DRS) to migrate virtual machines from one host to another for load balancing. Because storage is on a SAN array, applications continue running seamlessly.
- If you use VMware DRS clusters, put an ESX/ESXi host into maintenance mode to have the system migrate all running virtual machines to other ESX/ESXi hosts. You can then perform upgrades or other maintenance operations.

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

ESX/ESXi and SAN Use Cases

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

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

Maintenance with zero downtime

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

Load balancing

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

Storage consolidation and simplification of storage layout

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

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

Disaster recovery

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

Simplified array migrations and storage upgrades

When you purchase new storage systems or arrays, use storage VMotion to perform live automated migration of virtual machine disk files from existing storage to their new destination.

Finding Further Information

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

- Use your storage array vendor's documentation for most setup questions. Your storage array vendor might also offer documentation on using the storage array in an ESX/ESXi environment.
- The VMware Documentation Web site.
- The *Fibre Channel SAN Configuration Guide* discusses the use of ESX/ESXi with Fibre Channel 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.

Specifics of Using SAN Storage with ESX/ESXi

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

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

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

Third-Party Management Applications

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

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

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

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

If you decide to run the SAN management software on a virtual machine, you gain the benefits of running a virtual machine, including failover using VMotion and VMware HA. Because of the additional level of indirection, however, the management software might not be able to detect the SAN. This problem can be resolved by using an RDM.

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

Understanding VMFS Datastores

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

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

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

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

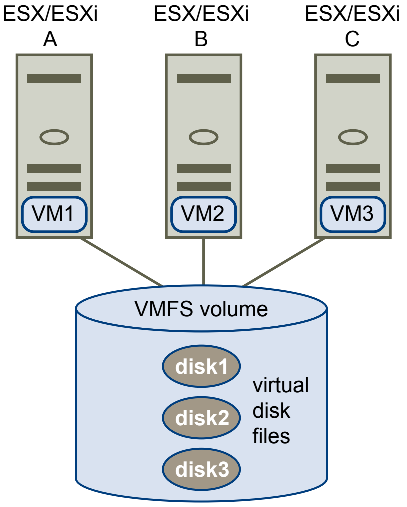
Sharing a VMFS Datastore Across ESX/ESXi Hosts

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

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

Figure 1-3 shows several ESX/ESXi systems sharing the same VMFS volume.

Figure 1-3. Sharing a VMFS Datastore Across ESX/ESXi Hosts



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

Metadata Updates

A VMFS datastore holds virtual machine files, directories, symbolic links, RDMs, and so on. A VMS datastore also maintains a consistent view of all the mapping information for these objects. This mapping information is called metadata.

Metadata is updated each time the attributes of a virtual machine file are accessed or modified when, for example, you perform one of the following operations:

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

Making LUN Decisions

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

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

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

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

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

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

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

When the storage characterization for a virtual machine is not available, there is often no simple answer when you have to decide on the LUN size and number of LUNs to use. You can experiment using either predictive or adaptive scheme.

Use the Predictive Scheme to Make LUN Decisions

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

Procedure

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

   Disk shares are relevant only within a given host. The shares assigned to virtual machines on one host have no effect on virtual machines on other hosts.
5. Run the applications to determine whether virtual machine performance is acceptable.

Use the Adaptive Scheme to Make LUN Decisions

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

Procedure

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

2 Build a VMFS on that LUN.

3 Place four or five virtual disks on the VMFS.

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

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

Use Disk Shares to Prioritize Virtual Machines

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

Procedure

1 Start a vSphere Client and connect to vCenter Server.

2 Select the virtual machine in the inventory panel and click **Edit virtual machine settings** from the menu.

3 Click the **Resources** tab and click **Disk**.

4 Double-click the **Shares** column for the disk to modify and select the required value from the drop-down menu.

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

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

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

How Virtual Machines Access Data on a SAN

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

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

1 When the guest operating system in a virtual machine reads or writes to SCSI disk, it issues SCSI commands to the virtual disk.

2 Device drivers in the virtual machine’s operating system communicate with the virtual SCSI controllers.

3 The virtual SCSI Controller forwards the command to the VMkernel.

4 The VMkernel performs the following tasks.

   - Locates the file in the VMFS volume that corresponds to the guest virtual machine disk.
   - Maps the requests for the blocks on the virtual disk to blocks on the appropriate physical device.
   - Sends the modified I/O request from the device driver in the VMkernel to the iSCSI initiator (hardware or software).

5 If the iSCSI initiator is a hardware iSCSI initiator (iSCSI HBA), the HBA performs the following tasks.

   - Encapsulates I/O requests into iSCSI Protocol Data Units (PDUs).
   - Encapsulates iSCSI PDUs into TCP/IP packets.
   - Sends IP packets over Ethernet to the iSCSI storage system.

6 If the iSCSI initiator is a software iSCSI initiator, the following takes place.

   - The initiator encapsulates I/O requests into iSCSI PDUs.
   - The initiator sends iSCSI PDUs through TCP/IP connections.
   - The VMkernel TCP/IP stack relays TCP/IP packets to a physical NIC.
   - The physical NIC sends IP packets over Ethernet to the iSCSI storage system.

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

   This storage device appears to be a specific disk to the host, but it might be a logical device that corresponds to a physical device on the SAN.

Understanding Multipathing and Failover

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

In case of a failure of any element in the SAN network, such as an adapter, switch, or cable, ESX/ESXi can switch to another physical path, which does not use the failed component. This process of path switching to avoid failed components is known as path failover.

In addition to path failover, multipathing provides load balancing. Load balancing is the process of distributing I/O loads across multiple physical paths. Load balancing reduces or removes potential bottlenecks.

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

Managing Multiple Paths

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

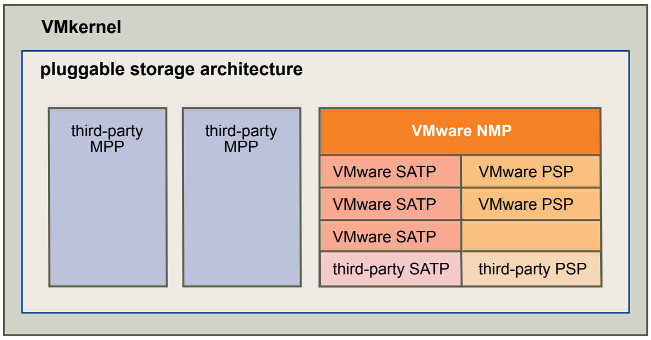
If more multipathing functionality is required, a third party can also provide an MPP to run in addition to, or as a replacement for, the default NMP.

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 1-4 illustrates, multiple third-party MPPs can run in parallel with the VMware NMP. The third-party MPPs can replace the behavior of the NMP and take complete control of the path failover and the load-balancing operations for specified storage devices.

Figure 1-4. Pluggable Storage Architecture



The multipathing modules perform the following operations:

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

VMware Multipathing Module

By default, ESX/ESXi provides an extensible multipathing module called the Native Multipathing 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.

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 an SATP for every type of array that VMware supports. These SATPs include an active/active SATP and active/passive SATP for non-specified storage arrays, and the local SATP for direct-attached storage. Each SATP accommodates special characteristics of a certain class of storage arrays and can perform the array-specific operations required to detect path state and to activate an inactive path. As a result, the NMP module can work with multiple storage arrays without having to be aware of the storage device specifics.

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

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

VMware PSPs

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

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

By default, the VMware NMP supports the following PSPs:

Most Recently Used (MRU)

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

Fixed

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

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

Round Robin (RR)

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

VMware NMP Flow of I/O

When a virtual machine issues an I/O request to a storage device managed by the NMP, the following process takes place.

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

Host-Based Path Failover

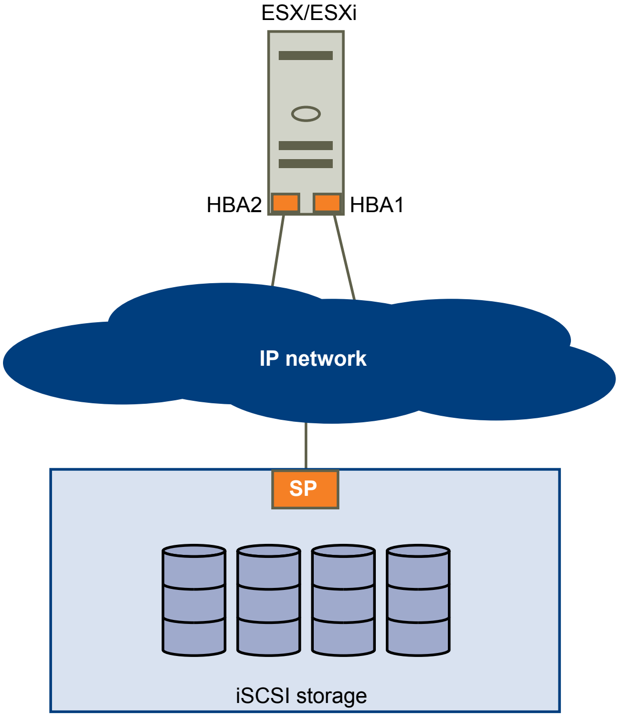
When setting up your ESX/ESXi host for multipathing and failover, you can use multiple iSCSI HBAs with the hardware iSCSI and multiple NICs with the software iSCSI.

Failover with Hardware iSCSI

With the hardware iSCSI, the host typically has two or more hardware iSCSI adapters available, from which the storage system can be reached using one or more switches. Alternatively, the setup might include one adapter and two storage processors so that the adapter can use a different path to reach the storage system.

As Figure 1-5 illustrates, the host has two hardware iSCSI adapters, HBA1 and HBA2, that provide two physical paths to the storage system. Multipathing plug-ins on your host, whether the VMkernel NMP or any third-party MPPs, have access to the paths by default and can monitor health of each physical path. If, for example, HBA1 or the link between HBA1 and the network fails, the multipathing plug-ins can switch the path over to HBA2.

Figure 1-5. Hardware iSCSI and Failover



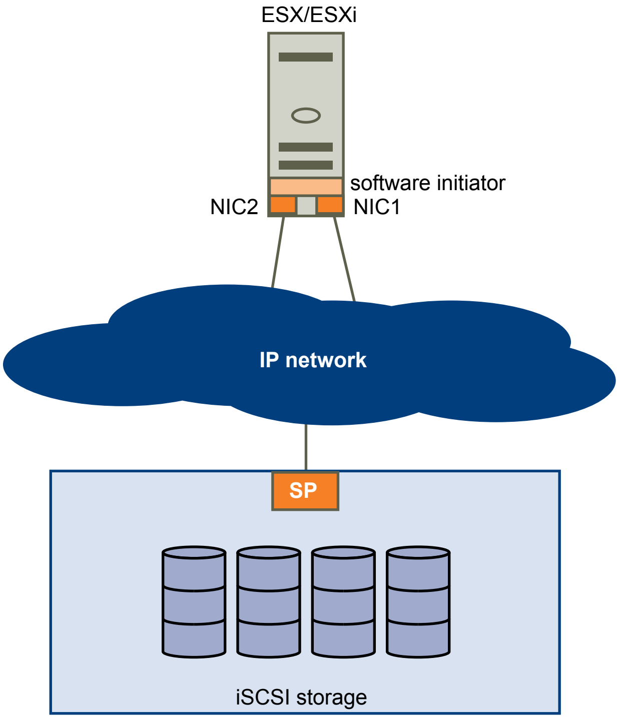
Failover with Software iSCSI

With the software iSCSI, as Figure 1-6 shows, you can use multiple NICs that provide failover and load balancing capabilities for iSCSI connections between your host and storage systems.

For this setup, because multipathing plug-ins do not have direct access to physical NICs on your host, you first need to connect each physical NIC to a separate VMkernel port. You then associate all VMkernel ports with the software iSCSI initiator using a port binding technique. As a result, each VMkernel port connected to a separate NIC becomes a different path that the iSCSI storage stack and its storage-aware multipathing plug-ins can use.

For information on how to configure multipathing for the software iSCSI, see "Networking Configuration for Software iSCSI Storage," on page 32.

Figure 1-6. Software iSCSI and Failover



Array-Based Failover

Some iSCSI storage systems manage path use of their ports automatically (transparently to ESX/ESXi).

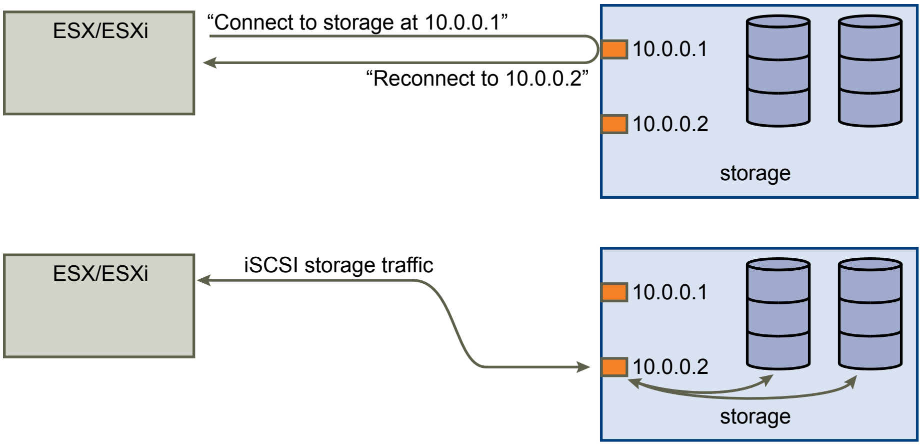
When using one of these storage systems, ESX/ESXi does not see multiple ports on the storage and cannot choose the storage port it connects to. These systems have a single virtual port address that ESX/ESXi uses to initially communicate. During this initial communication, the storage system can redirect ESX/ESXi to communicate with another port on the storage system. The iSCSI initiators in ESX/ESXi obey this reconnection request and connect with a different port on the system. The storage system uses this technique to spread the load across available ports.

If ESX/ESXi loses connection to one of these ports, it automatically attempts to reconnect with the virtual port of the storage system, and should be redirected to an active, usable port. This reconnection and redirection happens quickly and generally does not disrupt running virtual machines. These storage systems can also request that iSCSI initiators reconnect to the system, to change which storage port they are connected to. This allows the most effective use of the multiple ports.

Figure 1-7 shows an example of port redirection. ESX/ESXi attempts to connect to the 10.0.0.1 virtual port. The storage system redirects this request to 10.0.0.2. ESX/ESXi connects with 10.0.0.2 and uses this port for I/O communication.

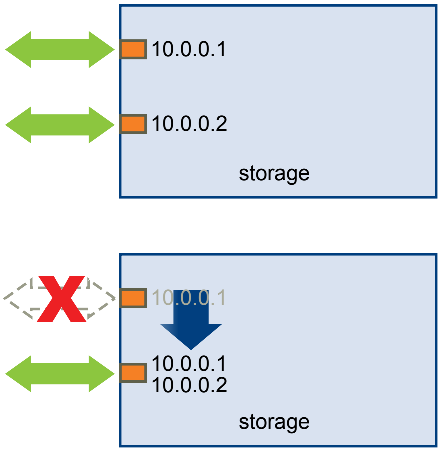
Note The storage system does not always redirect connections. The port at 10.0.0.1 could be used for traffic, also.

Figure 1-7. Port Redirection



If the port on the storage system that is acting as the virtual port becomes unavailable, the storage system reassigns the address of the virtual port to another port on the system. Figure 1-8 shows an example of this type of port reassignment. In this case, the virtual port 10.0.0.1 becomes unavailable and the storage system reassigns the virtual port IP address to a different port. The second port responds to both addresses.

Figure 1-8. Port Reassignment



Choosing Virtual Machine Locations

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

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

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

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

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

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

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

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

Designing for Server Failure

The RAID architecture of SAN storage inherently protects you from failure at the physical disk level. A SAN provides multiple paths between servers and storage, which protects against network or port failures. The final step in making your whole environment failure resistant is to protect against server failure.

Using VMware HA

One of the failover options ESX/ESXi provides is VMware High Availability (HA).

VMware HA allows you to organize virtual machines into failover groups. When a host fails, all its virtual machines are immediately started on different hosts. When a virtual machine is restored on a different host, it loses its memory state, but its disk state is exactly as it was when the host failed (crash-consistent failover). Shared storage (such as a SAN) is required for HA.

NOTE You must be licensed to use VMware HA.

Server Failover and Storage Considerations

When you are configuring your ESX/ESXi host to work in conjunction with SAN, you must make your whole environment failure resistant and protect it against host failures.

For each type of server failover, you must follow these practices:

- 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 system 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 datastores the clustered virtual machines use must be seen by all ESX/ESXi hosts.

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

NOTE As a rule, when you boot from a SAN, each boot volume should be seen only by the host that is booting from that volume. An exception is when you try to recover from a failure by pointing a second host to the same volume. In this case, the SAN volume in question is not really for booting from a SAN. No host is booting from it because it is corrupted. The SAN volume is a regular non-boot volume that is made visible to a host.

LUN Display and Rescan

A SAN is dynamic, and which LUNs are available to a certain host can change based on a number of factors.

The VMkernel discovers LUNs when it boots, and those LUNs are then visible in the vSphere Client. If changes are made to the LUNs, you must rescan to see those changes.

- New LUNs created on the iSCSI storage
- Changes to LUN access control
- Changes in connectivity

Configuring iSCSI Initiators and Storage 2

Before ESX/ESXi can work with a SAN, you must set up your iSCSI initiators and storage. To do this, you must first observe certain basic requirements and then follow best practices for installing and setting up hardware or software iSCSI initiators to access the SAN.

This chapter includes the following topics:

ESX/ESXi iSCSI SAN Requirements

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

- Verify that your SAN storage hardware and firmware combinations are supported in conjunction with ESX/ESXi systems. For an up-to-date list, see the Storage/SAN Compatibility Guide.
- Configure your system to have only one VMFS datastore for each LUN. (In VMFS-3, you do not need to set accessibility.)
- Unless you are using diskless servers (booting from a SAN), 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 for access to any raw disk.
- Set the SCSI controller driver in the guest operating system to 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 allow Windows to better tolerate delayed I/O resulting from path failover.

ESX/ESXi iSCSI SAN Restrictions

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

- ESX/ESXi does not support iSCSI-connected tape devices.
- You cannot use virtual-machine multipathing software to perform I/O load balancing to a single physical LUN.

Setting LUN Allocations

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

Note the following points:

- Storage Provisioning. To ensure that the ESX/ESXi host recognizes LUNs at startup time, make sure to configure all iSCSI storage targets so that your host can access them and use them. Also, configure your host so that it can discover all available iSCSI targets.
- VMotion and VMware DRS. When you use vCenter Server and VMotion or DRS, make sure that the LUNs for the virtual machines are provisioned to all ESX/ESXi hosts. This configuration provides the greatest freedom in moving virtual machines.
- Active-active versus active-passive arrays. When you use VMotion or DRS with an active-passive SAN storage device, make sure that all 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. You must connect the server to the active port on the storage system. This configuration ensures that the LUNs are presented to the ESX/ESXi host.

Network Configuration and Authentication

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

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

Setting Up Hardware iSCSI Initiators

With hardware-based iSCSI storage, you use a specialized third-party adapter capable of accessing iSCSI storage over TCP/IP. This iSCSI initiator handles all iSCSI and network processing and management for your ESX/ESXi system.

You must install and configure the hardware iSCSI adapter for your host to be able to access the iSCSI storage device. For installation information, see vendor documentation.

View Hardware iSCSI Initiators

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

Prerequisites

Before you begin configuring the hardware iSCSI initiator, make sure that the iSCSI HBA is successfully installed and appears on the list of initiators available for configuration. If the initiator is installed, you can view its properties.

Procedure

1. Log in to the vSphere Client, and select a host from the inventory panel.
2. Click the **Configuration** tab and click **Storage Adapters** in the Hardware panel.

   The hardware iSCSI initiator appears in the list of storage adapters.
3. Select the initiator to view.

   The default details for the initiator appear, including the model, iSCSI name, iSCSI alias, IP address, and target and paths information.
4. Click **Properties**.

   The iSCSI Initiator Properties dialog box appears. The **General** tab displays additional characteristics of the initiator.

You can now configure your hardware initiator or change its default characteristics.

Change Name and IP Address for Hardware Initiators

When you configure your hardware iSCSI initiators, make sure that their names and IP addresses are formatted properly.

Procedure

1. Log in to the vSphere Client, and select a host from the inventory panel.
2. Click the **Configuration** tab and click **Storage Adapters** in the Hardware panel.
3. Select the initiator to configure and click **Properties > Configure**.
4. To change the default iSCSI name for your initiator, enter the new name.

   Make sure the name you enter is worldwide unique and properly formatted or some storage devices might not recognize the hardware iSCSI initiator.
5. (Optional) Enter the iSCSI alias.

   The alias is a name that you use to identify the hardware iSCSI initiator.
6. Change the default IP settings.

   You must change the default IP settings so that they are configured properly for the IP SAN. Work with your network administrator to determine the IP setting for the HBA.
7. Click **OK** to save your changes.

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

Setting Up Software iSCSI Initiators

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

Before you configure the software iSCSI initiator, you must perform the following tasks:

1. Create a VMkernel port for physical network adapters.
2. Enable the software iSCSI initiator.
3. If you use multiple network adapters, activate multipathing on your host using the port binding technique.
4. If needed, enable Jumbo Frames. Jumbo Frames must be enabled for each vSwitch through the vSphere CLI. Also, if you use an ESX host, you must create a VMkernel network interface enabled with Jumbo Frames.

Networking Configuration for Software iSCSI Storage

Networking configuration for software iSCSI involves creating an iSCSI VMkernel port and mapping it to a physical NIC that handles iSCSI traffic.

Depending on the number of physical NICs you use for iSCSI traffic, the networking setup can be different:

- If you have one physical NIC, create one VMkernel port on a vSwitch and map the port to the NIC. VMware recommends that you designate a separate network adapter entirely for iSCSI. No additional network configuration steps are required.

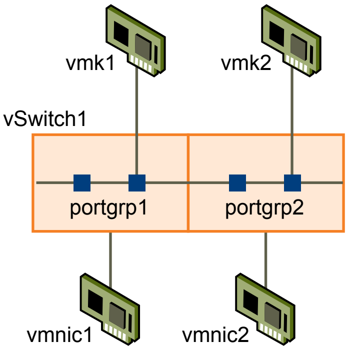
  For information on creating a port, see "Create a VMkernel Port for Software iSCSI," on page 33.

- If you have two or more physical NICs for iSCSI, you can create multiple paths for the software iSCSI by using the port binding technique.

  For background information on multipathing with software iSCSI, see "Host-Based Path Failover," on page 24.

  With port binding, you create a separate VMkernel port for each physical NIC using 1:1 mapping. You can add all network adapter and VMkernel port pairs to a single vSwitch, as Figure 2-1 shows.

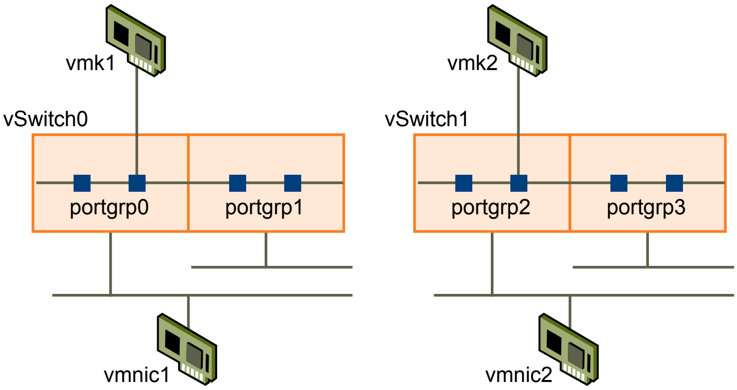
  Figure 2-1. Port Binding on a Single vSwitch



  For information on adding the NIC and VMkernel port pairs to a vSwitch, see "Set Up Multipathing for Software iSCSI," on page 34.

  Another alternative is to create a separate vSwitch for each network adapter and VMkernel port pair, as Figure 2-2 indicates.

Figure 2-2. Port Binding on Separate vSwitches



After you map VMkernel ports to network adapters, use the `esxcli` command to connect the ports with the software iSCSI initiator. For information, see "Activate Multipathing for Software iSCSI Initiator," on page 35.

NOTE Port binding support on EMC CLARiiON storage systems requires initiators in different subnets. See vendor documentation for additional details.

Create a VMkernel Port for Software iSCSI

This procedure lets you connect the VMkernel, which runs services for iSCSI storage, to the physical network adapter.

If you have one physical network adapter to be used for iSCSI traffic, this is the only procedure you must perform to set up your iSCSI networking.

Procedure

1 Log in to the vSphere Client and select the host from the inventory panel.

2 Click the **Configuration** tab and click **Networking**.

3 In the Virtual Switch view, click **Add Networking**.

4 Select **VMkernel** and click **Next**.

5 Select **Create a virtual switch** to create a new vSwitch.

If no adapters appear under **Create a virtual switch**, existing vSwitches are using all of the network adapters in the system. You can use an existing vSwitch for your iSCSI traffic.

6 Select an adapter you want to use for iSCSI traffic.

IMPORTANT Do not use iSCSI on 100Mbps or slower adapters.

7 Click **Next**.

8 Under **Port Group Properties**, enter a network label. Network label is a friendly name that identifies the VMkernel port that you are creating.

9 Click **Next**.

10 Specify the IP settings and click **Next**.

11 Review the information and click **Finish**.

What to do next

If your host uses only one network adapter for iSCSI, no additional network configuration steps are required.

If your host uses more than one physical network adapter for iSCSI, connect additional adapters and associate them with corresponding VMkernel ports using the port binding technique. You have the following options:

- Use a single vSwitch for iSCSI multipathing. You must connect additional network adapters and VMkernel ports to the vSwitch you just created and override the default setup, so that each port maps to only one active adapter. See "Set Up Multipathing for Software iSCSI," on page 34.
- Create separate vSwitches for each additional network adapter.

Set Up Multipathing for Software iSCSI

Use this procedure only if you have two or more NICs you can designate for iSCSI and you want to connect all of your iSCSI NICs to a single vSwitch. In this procedure, you associate VMkernel ports with the network adapters using 1:1 mapping.

You now need to connect additional network adapters to the existing vSwitch and map them to corresponding VMkernel ports.

Prerequisites

You must create one VMkernel port for your network adapter before you can set up multipathing for software iSCSI.

Procedure

1 Log in to the vSphere Client and select the host from the inventory panel.

2 Click the **Configuration** tab and click **Networking**.

3 Select the vSwitch that you use for iSCSI and click **Properties**.

4 Connect additional network adapters to the vSwitch.

  a In the vSwitch Properties dialog box, click the **Network Adapters** tab and click **Add**.

  b Select one or more adapters from the list and click **Next**.

  c Review the information on the Adapter Summary page, and click **Finish**.

  The list of network adapters reappears, showing the network adapters that the vSwitch now claims.

5 Create VMkernel ports for all network adapters that you connected.

  The number of VMkernel ports must correspond to the number of network adapters on the vSwitch.

  a In the vSwitch Properties dialog box, click the **Ports** tab and click **Add**.

  b Select **VMkernel** and click **Next**.

  c Under **Port Group Properties**, enter a network label and click **Next**.

  d Specify the IP settings and click **Next**.

  When you enter subnet mask, make sure that the network adapter is set to the subnet of the storage system it connects to.

  e Review the information and click **Finish**.

CAUTION If the network adapter you add to software iSCSI initiator is not in the same subnet as your iSCSI target, your host is not able to establish sessions from this network adapter to the target.

6 Map each VMkernel port to just one active adapter.

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

a On the **Ports** tab, select a VMkernel port and click **Edit**.

b Click the **NIC Teaming** tab and select **Override vSwitch failover order**.

c Designate only one adapter as active and move all remaining adapters to the **Unused Adapters** category.

7 Repeat the last step for each VMkernel port on the vSwitch.

What to do next

After performing this task, use the `esxcli` command to connect the VMkernel ports to the software iSCSI initiator.

Activate Multipathing for Software iSCSI Initiator

Use this task only if your ESX/ESXi host has two or more physical network adapters that you designate for iSCSI traffic. This task explains how to activate host-based multipathing for your host by connecting the software iSCSI initiator to iSCSI VMkernel ports that you created for the network adapters.

Prerequisites

Before you activate multipathing, complete the following tasks:

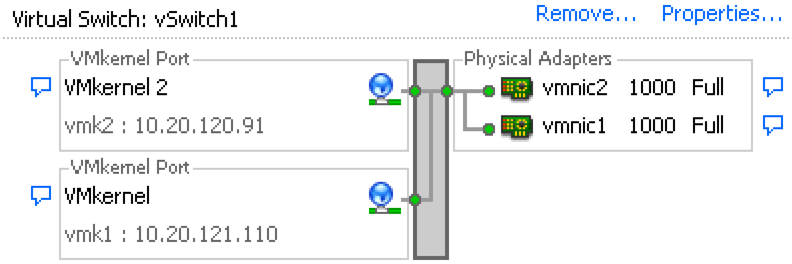
- Create VMkernel ports for the physical network adapters making sure to use 1:1 port to adapter mapping. See "Create a VMkernel Port for Software iSCSI," on page 33 and "Set Up Multipathing for Software iSCSI," on page 34.
- Enable the software iSCSI initiator. See "Enable the Software iSCSI Initiator," on page 36.

Procedure

1 Identify the names of VMkernel iSCSI ports assigned to physical adapters.

The vSphere Client displays the port's name below the network label.

For example, the following graphic shows the ports' names as vmk1 and vmk2.



2 Using the vSphere CLI, connect the software iSCSI initiator to the iSCSI VMkernel ports.

Repeat this command for each port.

```
esxcli swiscsi nic add -n <port_name> -d <vmhba>
```

3 Verify that the ports were added to the software iSCSI initiator by running the following command:

```
esxcli swiscsi nic list -d <vmhba>
```

4 Use the vSphere Client to rescan the software iSCSI initiator.

5 To disconnect the software iSCSI initiator from the ports, run the following command.

If there are active iSCSI sessions between your host and targets, discontinue them before running this command. You can do so by removing static targets that the ports use from the vSphere Client.

```
esxcli swiscsi nic remove -n <port_name> -d <vmhba>
```

Example 2-1. Connecting Software iSCSI Initiator to Two VMkernel Ports

This example shows how to connect the software iSCSI initiator vmhba33 to VMkernel ports vmk1 and vmk2.

1 Connect vmhba33 to vmk1: `esxcli swiscsi nic add -n vmk1 -d vmhba33`

2 Connect vmhba33 to vmk2: `esxcli swiscsi nic add -n vmk2 -d vmhba33`

3 Verify vmhba33 configuration: `esxcli swiscsi nic list -d vmhba33`

Both vmk1 and vmk2 should be listed.

In this example, if you use the vSphere client to display the Paths view for the vmhba33 initiator, you can see that it uses two different paths to access the same target. The runtime names of the paths are vmhba33:C1:T1:L0 and vmhba33:C2:T1:L0. C1 and C2 in this example indicate the two network adapters that are used for multipathing.

Enable the Software iSCSI Initiator

You must enable your software iSCSI initiator so that ESX/ESXi can use it to access iSCSI storage.

Procedure

1 Log in to the vSphere Client, and select a server from the inventory panel.

2 Click the **Configuration** tab and click **Storage Adapters** in the Hardware panel.

The list of available storage adapters appears.

3 Select the iSCSI initiator to configure and click **Properties**.

4 Click **Configure**.

The **General Properties** dialog box displays the initiator's status, default name, and alias.

5 To enable the initiator, select **Enabled**.

6 To change the default iSCSI name for your initiator, enter the new name.

Make sure the name you enter is worldwide unique and properly formatted or some storage devices might not recognize the software iSCSI initiator.

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

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

Enabling Jumbo Frames for Software iSCSI

Jumbo Frames allow ESX/ESXi to send larger frames out onto the physical network. The network must support Jumbo Frames end-to-end for Jumbo Frames to be effective. Jumbo Frames up to 9kB (9000 Bytes) are supported.

Before enabling Jumbo Frames, check with your hardware vendor to ensure your physical network adapter and iSCSI storage support Jumbo Frames.

Jumbo Frames must be enabled for each vSwitch through the vSphere CLI. Also, if you use an ESX host, you must create a VMkernel network interface enabled with Jumbo Frames.

Create a Jumbo Frames-Enabled vSwitch

Configure a vSwitch for Jumbo Frames by changing the MTU size for that vSwitch.

Procedure

1 To set the MTU size for the vSwitch, run the `vicfg-vswitch -m <MTU> <vSwitch>` command from the vSphere CLI.

  This command sets the MTU for all uplinks on that vSwitch. The MTU size should be set to the largest MTU size among all the virtual network adapters connected to the vSwitch.

2 Run the `vicfg-vswitch -l` command to display a list of vSwitches on the host, and check that the configuration of the vSwitch is correct.

Create a Jumbo Frames-Enabled VMkernel Interface

If you are using an ESX host, you must use the command-line interface to create a VMkernel network interface that is enabled with Jumbo Frames.

Procedure

1 Log in directly to the console of the ESX host.

2 Run the following commands to create a VMkernel connection with Jumbo Frame support.

```
esxcfg-vswitch -a <virtual switch>

esxcfg-vswitch -m <MTU> <virtual switch>

esxcfg-vswitch -A <port group name> <virtual switch>

esxcfg-vswitch -L <pnic> <virtual switch>

esxcfg-vmknic -a -I <ip address> -n <netmask> -m <MTU> <port group name>
```

3 Run the `esxcfg-vmknic -l` command to display a list of VMkernel interfaces and check that the configuration of the Jumbo Frame-enabled interface is correct.

4 Check that the VMkernel interface is connected to a vSwitch with Jumbo Frames enabled.

5 Configure all physical switches and any physical or virtual machines to which this VMkernel interface connects to support Jumbo Frames.

Configuring Discovery Addresses for iSCSI Initiators

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

The ESX/ESXi system supports these discovery methods:

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

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

Set Up Dynamic Discovery

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

Procedure

1. Log in to the vSphere Client and select a server from the inventory panel.
2. Click the **Configuration** tab and click **Storage Adapters** in the Hardware panel.

   The list of available storage adapters appears.
3. Select the iSCSI initiator to configure and click **Properties**.
4. In the iSCSI Initiator Properties dialog box, click the **Dynamic Discovery** tab.
5. To add an address for the Send Targets discovery, click **Add**.

   The **Add Send Targets Server** dialog box appears.
6. Enter the IP address or DNS name of the storage system and click **OK**.

   After your host establishes the Send Targets session with this system, any newly discovered targets appear in the Static Discovery list.
7. To delete a specific Send Targets server, select it and click **Remove**.

   After you remove a Send Targets server, it might still appear in the Inheritance field as the parent of static targets. This entry indicates where the static targets were discovered and does not affect the functionality.

Note You cannot change the IP address, DNS name, or port number of an existing Send Targets server. To make changes, delete the existing server and add a new one.

Set Up Static Discovery

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

Procedure

1. Log in to the vSphere Client and select a server from the inventory panel.
2. Click the **Configuration** tab and click **Storage Adapters** in the Hardware panel.

   The list of available storage adapters appears.
3. Select the iSCSI initiator to configure and click **Properties**.
4. In the iSCSI Initiator Properties dialog box, click the **Static Discovery** tab.

   The tab displays all dynamically discovered targets and any static targets already entered.
5. To add a target, click **Add** and enter the target's information.
6. To delete a specific target, select the target and click **Remove**.

Note You cannot change the IP address, DNS name, iSCSI target name, or port number of an existing target. To make changes, remove the existing target and add a new one.

Configuring CHAP Parameters for iSCSI Initiators

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

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

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

Choosing CHAP Authentication Method

ESX/ESXi supports one-way CHAP for both hardware and software iSCSI, and mutual CHAP for software iSCSI only.

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

ESX/ESXi supports the following CHAP authentication methods:

One-way CHAP

In one-way, or unidirectional, CHAP authentication, the target authenticates the initiator, but the initiator does not authenticate the target.

Mutual CHAP (software iSCSI only)

In mutual, or bidirectional, CHAP authentication, an additional level of security enables the initiator to authenticate the target.

For software iSCSI only, you can set one-way CHAP and mutual CHAP for each initiator or at the target level. Hardware iSCSI supports CHAP only at the initiator level.

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

Table 2-1. CHAP Security Level

CHAP Security Level	Description	Supported
Do not use CHAP	The host does not use CHAP authentication. Select this option to disable authentication if it is currently enabled.	Software iSCSI Hardware iSCSI
Do not use CHAP unless required by target	The host prefers a non-CHAP connection, but can use a CHAP connection if required by the target.	Software iSCSI
Use CHAP unless prohibited by target	The host prefers CHAP, but can use non-CHAP connections if the target does not support CHAP.	Software iSCSI Hardware iSCSI
Use CHAP	The host requires successful CHAP authentication. The connection fails if CHAP negotiation fails.	Software iSCSI

Set Up CHAP Credentials for an iSCSI Initiator

For increased security, you can set up all targets to receive the same CHAP name and secret from the iSCSI initiator at the initiator level. By default, all discovery addresses or static targets inherit CHAP parameters that you set up at the initiator level.

Prerequisites

Before setting up CHAP parameters for software iSCSI, determine whether to configure one-way or mutual CHAP. Hardware iSCSI does not support mutual CHAP.

- In one-way CHAP, the target authenticates the initiator.
- In mutual CHAP, both the target and initiator authenticate each other. Make sure to use different secrets for CHAP and mutual CHAP.

When configuring CHAP parameters, make sure that they match the parameters on the storage side.

For software iSCSI, the CHAP name should not exceed 511 and the CHAP secret 255 alphanumeric characters. For hardware iSCSI, the CHAP name should not exceed 255 and the CHAP secret 100 alphanumeric characters.

Procedure

1. Log in to the vSphere Client and select a server from the inventory panel.
2. Click the **Configuration** tab and click **Storage Adapters** in the Hardware panel.

   The list of available storage adapters appears.
3. Select the iSCSI initiator to configure and click **Properties**.
4. On the **General** tab, click **CHAP**.
5. To configure one-way CHAP, under CHAP specify the following.

   a. Select one of the following options:

      - **Do not use CHAP unless required by target** (software iSCSI only)
      - **Use CHAP unless prohibited by target**
      - **Use CHAP** (software iSCSI only). To be able to configure mutual CHAP, you must select this option.

   b. Specify the CHAP name.

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

      - To set the CHAP name to the iSCSI initiator name, select **Use initiator name**.
      - To set the CHAP name to anything other than the iSCSI initiator name, deselect **Use initiator name** and enter a name in the **Name** field.

   c. Enter a one-way CHAP secret to be used as part of authentication. Make sure to use the same secret that you enter on the storage side.
6. To configure mutual CHAP, first configure one-way CHAP by following directions in Step 5.

   Make sure to select **Use CHAP** as an option for one-way CHAP. Then, specify the following under **Mutual CHAP**:

   a. Select **Use CHAP**.

   b. Specify the mutual CHAP name.

   c. Enter the mutual CHAP secret. Make sure to use different secrets for the one-way CHAP and mutual CHAP.

7 Click **OK**.

8 Rescan the initiator.

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

Set Up CHAP Credentials for a Target

For software iSCSI, you can configure different CHAP credentials for each discovery address or static target.

When configuring CHAP parameters, make sure that they match the parameters on the storage side. For software iSCSI, the CHAP name should not exceed 511 and the CHAP secret 255 alphanumeric characters.

Prerequisites

Before setting up CHAP parameters for software iSCSI, determine whether to configure one-way or mutual CHAP.

- In one-way CHAP, the target authenticates the initiator.
- In mutual CHAP, both the target and initiator authenticate each other. Make sure to use different secrets for CHAP and mutual CHAP.

Procedure

1 Log in to the vSphere Client and select a server from the inventory panel.

2 Click the **Configuration** tab and click **Storage Adapters** in the Hardware panel.

The list of available storage adapters appears.

3 Select the software iSCSI initiator to configure and click **Properties**.

4 Select either **Dynamic Discovery** tab or **Static Discovery** tab.

5 From the list of available targets, select a target you want to configure and click **Settings > CHAP**.

6 To configure one-way CHAP, under CHAP specify the following.

a Deselect **Inherit from parent**.

b Select one of the following options:

- **Do not use CHAP unless required by target**
- **Use CHAP unless prohibited by target**
- **Use CHAP**. To be able to configure mutual CHAP, you must select this option.

c Specify the CHAP name.

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

- To set the CHAP name to the iSCSI initiator name, select **Use initiator name**.
- To set the CHAP name to anything other than the iSCSI initiator name, deselect **Use initiator name** and enter a name in the **Name** field.

d Enter a one-way CHAP secret to be used as part of authentication. Make sure to use the same secret that you enter on the storage side.

7. To configure mutual CHAP, first configure one-way CHAP by following directions in Step 6.

   Make sure to select **Use CHAP** as an option for one-way CHAP. Then, specify the following under **Mutual CHAP**:

   a. Deselect **Inherit from parent**.
   b. Select **Use CHAP**.
   c. Specify the mutual CHAP name.
   d. Enter the mutual CHAP secret. Make sure to use different secrets for the one-way CHAP and mutual CHAP.

8. Click **OK**.
9. Rescan the initiator.

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

Disable CHAP

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

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

Procedure

1. Open the CHAP Credentials dialog box.
2. For software iSCSI, to disable just the mutual CHAP, select **Do not use CHAP** under **Mutual CHAP**.
3. To disable one-way CHAP, select **Do not use CHAP** under **CHAP**.

   The mutual CHAP, if set up, automatically turns to **Do not use CHAP** when you disable the one-way CHAP.

4. Click **OK**.

Configuring Additional Parameters for iSCSI

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

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

Table 2-2 lists advanced iSCSI parameters that you can configure using the vSphere Client. In addition, you can use the `vicfg-iscsi` vSphere CLI command to configure some of the advanced parameters. For information, see the *VMware vSphere Command-Line Interface Installation and Reference Guide*.

Table 2-2. Additional Parameters for iSCSI Initiators

Advanced Parameter	Description	Configurable On
Header Digest	Increases data integrity. When header digest is enabled, the system performs a checksum over each iSCSI Protocol Data Unit's (PDU's) header part and verifies using the CRC32C algorithm.	Software iSCSI
Data Digest	Increases data integrity. When data digest is enabled, the system performs a checksum over each PDU's data part and verifies using the CRC32C algorithm. NOTE Systems that use Intel Nehalem processors offload the iSCSI digest calculations for software iSCSI, thus reducing the impact on performance.	Software iSCSI
Maximum Outstanding R2T	Defines the R2T (Ready to Transfer) PDUs that can be in transition before an acknowledge PDU is received.	Software iSCSI
First Burst Length	Specifies the maximum amount of unsolicited data an iSCSI initiator can send to the target during the execution of a single SCSI command, in bytes.	Software iSCSI
Maximum Burst Length	Maximum SCSI data payload in a Data-In or a solicited Data-Out iSCSI sequence, in bytes.	Software iSCSI
Maximum Receive Data Segment Length	Maximum data segment length, in bytes, that can be received in an iSCSI PDU.	Software iSCSI
ARP Redirect	Allows storage systems to move iSCSI traffic dynamically from one port to another. ARP is required by storage systems that do array-based failover.	Hardware iSCSI (Configurable through vSphere CLI)
Delayed ACK	Allows systems to delay acknowledgment of received data packets.	Software iSCSI

Configure Advanced Parameters for iSCSI

The advanced iSCSI settings control such parameters as header and data digest, ARP redirection, delayed ACK, and so on. Generally, you do not need to change these settings because your ESX/ESXi host works with the assigned predefined values.

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

Procedure

1 Log in to the vSphere Client, and select a host from the inventory panel.

2 Click **Configuration** tab and click **Storage Adapters**.

3 Select the iSCSI initiator to configure and click **Properties**.

4 To configure advanced parameters at the initiator level, on the **General** tab, click **Advanced**. Proceed to Step 6.

5 Configure advanced parameters at the target level.

At the target level, advanced parpameters can be configured only for software iSCSI.

a Select either the **Dynamic Discovery** tab or **Static Discovery** tab.

b From the list of available targets, select a target to configure and click **Settings** > **Advanced**.

6 Enter any required values for the advanced parameters you want to modify and click **OK** to save your changes.

Add iSCSI Storage

When you create a datastore on an iSCSI storage device, the Add Storage wizard guides you through the configuration.

Procedure

1 Log in to the vSphere Client and select a server from the inventory panel.

2 Click the **Configuration** tab and click **Storage**.

3 Click **Add Storage**.

4 Select the **Disk/LUN** storage type and click **Next**.

The **Select Disk/LUN** page appears. This can take a few seconds depending on the number of targets.

5 Select the iSCSI device to use for your datastore and click **Next**.

6 Review the current disk layout and click **Next**.

7 Enter a datastore name and click **Next**.

The datastore name appears in the vSphere Client, and the label must be unique within the current VMware vSphere instance.

8 If needed, adjust the file system values and capacity you use for the datastore.

By default, the entire free space available on the storage device is offered to you.

9 Click **Next**.

The **Ready to Complete** page appears.

10 Review the datastore configuration information and click **Finish**.

A datastore is now available on the iSCSI storage device.

Modifying SAN Storage Systems for ESX/ESXi

3

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

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

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

VMware tests ESX/ESXi with storage systems in the following configurations:

Basic Connectivity
Tests whether ESX/ESXi can recognize and operate with the storage system. 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.

Booting from a SAN (with ESX hosts only)
The ESX host boots from a LUN configured on the SAN rather than from the server itself.

General Considerations for iSCSI SAN Storage Systems

When you prepare your iSCSI SAN storage system to work with ESX/ESXi, you need to follow specific general requirements that apply to all storage systems.

For all storage systems, the following general requirements exist:

- LUNs must be presented to each HBA of each host with the same LUN ID number. If different numbers are used, the ESX/ESXi hosts do not recognize different paths to the same LUN. Because instructions on how to configure identical SAN LUN IDs are vendor-specific, consult your storage documentation for more information.
- Unless specified for individual storage systems discussed in this chapter, set the host type for LUNs presented to ESX/ESXi to `Linux or Linux Cluster`, if applicable to your storage system. The method ESX/ESXi uses to access the storage system is most compatible with Linux access, however, this can vary depending on the storage system you are using.
- If you are using VMotion, DRS, or HA, make sure that source and target hosts for virtual machines can see the same LUNs with identical LUN IDs. SAN administrators might find it counterintuitive to have multiple hosts see the same LUNs because they might be concerned about data corruption. However, VMFS prevents multiple virtual machines from writing to the same file at the same time, so provisioning the LUNs to all required ESX/ESXi system is appropriate.
- If you do not have CHAP authentication set up on the LUNs that are being accessed, you must also disable CHAP on the ESX/ESXi host. Otherwise, authentication of the storage system fails, although the LUNs have no CHAP requirement.

EMC CLARiiON Storage Systems

EMC CLARiiON storage systems work with ESX/ESXi hosts in iSCSI SAN configurations. Generally, you use the EMC software to perform configurations.

For more information, see the EMC documentation.

This is an active-passive disk array, so any related issues that apply to all active-passive disk arrays are relevant. In addition, keep in mind the following:

- To avoid the possibility of path thrashing, the default multipathing policy is Most Recently Used, not Fixed. The ESX/ESXi system sets the default policy when it identifies the storage system.
- To boot from a SAN, choose the active storage processor for the boot LUN's target in the HBA BIOS.
- On EMC CLARiiON AX100i and AX150i systems, RDMs are supported only if you use the Navisphere Management Suite for SAN administration. Navisphere Express is not guaranteed to configure them properly.

  To use RDMs successfully, a given LUN must be presented with the same LUN ID to every ESX/ESXi host in the cluster. The AX100i and AX150i do not do this by default.

- When you use an AX100i or AX150i storage system, no host agent periodically checks the host configuration and pushes changes to the storage system. The `axnaviserverutil cli` utility is used to update the changes. This is a manual operation that you should perform as needed.
- Port binding support on EMC CLARiiON storage systems requires initiators in different subnets. See vendor documentation for additional details.
- 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 Symmetrix Storage Systems

To work with ESX/ESXi, EMC Symmetrix storage systems require certain specific settings. Use EMC software to configure the storage system. For information, see your EMC documentation.

The following settings are required for ESX/ESXi operations on the Symmetrix networked storage system:

- Common serial number (C)
- Auto negotiation (EAN) enabled
- SCSI 3 (SC3) set (enabled)
- Unique world wide name (UWN)
- SPC-2 (Decal) (SPC2) SPC-2 flag is required

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

Enable HP StorageWorks MSA1510i to Communicate with ESX/ESXi

This section describes the setup and configuration steps needed to allow an HP StorageWorks MSA1510i storage system to communicate with ESX/ESXi hosts.

Procedure

1 Install, connect, and power up the network devices as detailed in the vendor installation document.

2 Obtain the IP address assigned to the MSA1510i controller management port.

   a Scroll through the messages on the LCD panel until the following message appears: `603 Port MA0 IP` *`<address>`*

   b Record the management port IP address that appears in **Basic MSA1510i information**.

3 From the server or a workstation on the MSA1510i LAN segment, open a Web browser and enter the address obtained in the previous step.

4 When prompted, enter the default access permissions.

   - User name: `root`
   - Password: `root`

5 When prompted, set a unique user name and password.

6 Using the wizard, complete the following actions.

Option	Description
Storage configuration	a Set the Fault Tolerant mode (RAID mode). b Assign a spare disk for appropriate RAID level.
iSCSI configuration (configure an iSCSI portal)	a Select a data port. b Assign an IP address to the data port. c VLANs are set up on the switch and are used as one method of controlling access to the storage. If you are using VLANs, enter the VLAN ID to use (0 = not used). d The wizard suggests a default iSCSI Target Name and iSCSI Target Alias. Accept the default or enter user-defined values. NOTE To configure the remaining data ports, complete the Initial System Configuration Wizard process, and then use tasks available on the **Configure** tab.
Login settings	
Management settings	

7 Click **Finish** to apply the configuration settings.

NOTE Wizards are available for basic configuration tasks only. Use the **Manage** and **Configure** tabs to view and change your configuration.

What to do next

After initial setup, perform the following tasks to complete the configuration:

- Create an array.
- Create a logical drive.
- Create a target.
- Create a portal group.
- Associate or assign the portals created using the wizard with the portal group created.
- Map logical drives to the target.
- Add initiators (initiator IQN name and alias).
- Update the ACLs of the logical drives to provide access to initiators (select the list of initiators to access the logical drive).

HP StorageWorks EVA Storage Systems

The two types of HP StorageWorks EVA systems are EVA_GL, an active-passive system, and EVA_XL, an active-active system. For the systems to work with ESX/ESXi, certain specific settings are required.

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

- For HP EVAgl 3000/5000 (active-passive), use the `000000002200282E` host mode type.
- For HP EVAgl firmware 4.001 (active-active firmware for GL series) and above, use the `VMware` host mode type.
- For EVA4000/6000/8000 active-active arrays with firmware earlier than 5.031, use the `000000202200083E` host mode type.
- For EVA4000/6000/8000 active-active arrays with firmware 5.031 and later, use the `VMware` host mode type.

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

NetApp Storage Systems

For NetApp storage systems to communicate within an ESX/ESXi environment, you must perform specific configuration steps.

For additional documentation on NetApp and VMware best practices and SAN solutions, search the NetApp web page.

Table 3-1. Configuration Steps

Configuration Step	Description
Disable ALUA.	If any of your iSCSI initiators are a part of an initiator group (igroup), disable ALUA on the NetApp filter.
Set up multipathing.	When you set up multipathing between two iSCSI HBAs and multiple ports on a NetApp storage system, give the two HBAs different dynamic or static discovery addresses to connect to the storage. The NetApp storage system only permits one connection for each target and each initiator. Attempts to make additional connections cause the first connection to drop. Therefore, a single HBA should not attempt to connect to multiple IP addresses associated with the same NetApp target.
Set LUN type and initiator group type.	Set the appropriate LUN type and initiator group type for the storage system: ■ LUN type – VMware (if VMware type is not available, use Linux). ■ Initiator group type – VMware (if VMware type is not available, use Linux).
Provision storage.	Use either FilerView or CLI.

Provision Storage by Using FilerView Storage Management

You can use FilerView to provision storage on a NetApp storage system.

Procedure

1 Log in to NetApp storage system management (FilerView).

2 Create a volume.

   a Select **Volumes** and click **Add**.

   b Click **Next**.

   c Select **Flexibility** (Default) or **Traditional**, then click **Next**.

   d Enter a **Volume Name**, select a **Language**, and click **Next**.

   e Enter values for **Containing Aggregate**, **Total Volume Size**, and **Space Guarantee** and click **Next**.

   f Click **Commit** to create the volume.

3 Create LUNs.

   a Select **LUNs** and click **Add**.

   b Enter the following:

   - **Path**: Enter a path, for example, `/vol/vol1/lun1`.
   - **LUN Protocol Type**: VMware.
   - **Description**: A brief description.
   - **Size and Unit**: Enter a size, for example, 10GB and select **Space Reserved**.

4 Create an initiator group.

   a Select **LUNs > Initiator Group** and click **Add**.

   b Enter the following:

      - **Group Name**: Enter a group name
      - **Type**: Choose **iSCSI**.
      - **Operating System**: Enter `VMware`.
      - **Initiators**: Enter fully qualified initiator names. If there is more than one initiator, each initiator has to be separated with a return carriage.

   c Click **Add**.

5 Map the LUN to the initiator group.

   a Select **LUNs** and click **Manage**.

      A LUNs list appears.

   b From this list, click the label on the **Maps** row for the specific LUNs.

   c Click **Add Groups to Map**.

   d Select the initiator group and click **Add**.

   e When prompted, enter the LUN ID (any number from 0 to 255) and click **Apply**.

Provision Storage by Using Command-Line Interface

You can use command-line interface to provision storage on a NetApp storage system.

Procedure

1 Use command-line interface to 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 created.

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

EqualLogic Storage Systems

When setting up your EqualLogic storage systems to work in an ESX/ESXi implementation, you must address certain specific issues.

The following are specific requirements for EqualLogic storage systems to work with ESX/ESXi:

- Multipathing. No special setup is needed because EqualLogic storage systems support storage-processor failover that is transparent to iSCSI. Multiple iSCSI HBAs or NICs can connect to the same target or LUN on the storage side.
- Creating iSCSI LUNs. From the EqualLogic web portal, right-click **Volumes**, and then select **Create Volume**.
- Enable ARP redirection for ESX/ESXi hardware iSCSI HBAs.
- EqualLogic storage systems impose a maximum limit of 512 iSCSI connections per storage pool and 2048 connections per storage group.

For more information about configuring and using EqualLogic storage systems, see the vendor's documentation.

LeftHand Networks SAN/iQ Storage Systems

SAN/iQ SANs support ESX/ESXi iSCSI connections from a software initiator and hardware initiator.

When configuring SAN/iQ, enable automatic volume resignaturing for SAN/iQ storage devices to allow access to SAN/iQ snapshots and remote copies.

For more information on configuring LeftHand Networks SANs for VMware vSphere, see the vendor documentation related to VMware.

Basic configuration steps include several tasks.

1. Install SAN/iQ storage nodes.
2. Create SAN/iQ management groups and clusters.
3. Create volumes.
4. Assign volumes to authentication groups and volume lists.
5. Enable ARP redirection on hardware iSCSI HBAs.

As a best practice, configure virtual IP load balancing in SAN/iQ for all ESX/ESXi authentication groups.

Dell PowerVault MD3000i Storage Systems

When you configure mutual CHAP for the MD3000i iSCSI storage systems, special considerations that apply.

When you configure mutual CHAP for the MD3000i iSCSI array, follow these guidelines:

- On the MD3000i storage system, mutual CHAP configuration requires only a CHAP secret.
- On the ESX/ESXi host, mutual CHAP configuration requires both the name and CHAP secret. When configuring mutual CHAP on the ESX/ESXi host, enter the IQN name of the target as the mutual CHAP name. Make sure the CHAP secret matches the one set on the array.

Booting from an iSCSI SAN with ESX Systems 4

If you use ESX host, you can set up your system to boot from a SAN. The boot image is not stored on the ESX system's local disk, but instead is stored on a SAN LUN. You can boot from a SAN only with hardware iSCSI.

This chapter includes the following topics:

- "Booting from a SAN Overview," on page 53
- "Enable Booting from a SAN," on page 54

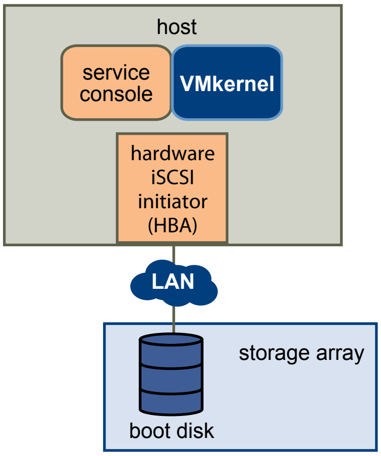
Booting from a SAN Overview

When booting from a SAN, the boot image of the ESX host is installed on one or more LUNs in the SAN storage system. When the host starts, it boots from the LUN on the SAN storage system.

NOTE When you boot from a SAN in conjunction with a VMware ESX system, each server must have its own boot LUN.

Only ESX hosts with hardware iSCSI initiators can boot from SAN.

Figure 4-1. How Booting from a SAN Works



Benefits of Booting from a SAN

Booting your ESX host from a SAN provides numerous benefits.

The benefits of booting from SAN include:

- Cheaper servers – Servers can be more dense and run cooler without internal storage.
- Easier server replacement – You can replace servers and have the new server point to the old boot location.

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

Deciding to Boot from a SAN

Before you consider how to set up your ESX host for booting from a SAN, decide whether it makes sense for your environment.

Boot from a SAN:

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

Do not boot from a SAN if you risk I/O contention between the service console and VMkernel.

Enable Booting from a SAN

You must complete several tasks to enable booting from SAN on an ESX host.

Procedure

1. Review any vendor configuration recommendations that apply to the storage system or the server booting from SAN.
2. Configure the hardware elements of your storage network, including SAN and HBAs.
3. Configure ACLs on your storage system.

   Proper access control on the storage systems is important when an ESX host is booting from iSCSI.

   - Boot LUNs should only be visible to the server using that LUN to boot. No other server or system on the SAN should be permitted to see that boot LUN.
   - Multiple ESX hosts can share a diagnostic partition. ACLs on the storage systems can allow you to do this.
4. Choose the location for the diagnostic partition.

   Diagnostic partitions can be put on the same LUN as the boot partition. Core dumps are stored in diagnostic partitions. If a diagnostic partition is configured in the boot LUN, this LUN cannot be shared between multiple hosts
5. Set up your ESX to boot from CD-ROM first because the VMware installation CD is in the CD-ROM drive.

   To achieve this, change the system boot sequence in your system BIOS setup.

Prepare the SAN

Before you configure the iSCSI HBAs to boot from a SAN, first prepare your storage area network by checking the cabling and switch wiring and configuring the storage system.

⚠ **CAUTION** If you use scripted installation to install ESX when booting from a SAN, you must take special steps to avoid unintended data loss. See VMware knowledge base article 1540.

Procedure

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

2 Ensure IP connectivity between your storage system and server.

  This includes proper configuration of any routers or switches on your storage network. Storage systems must be able to ping the iSCSI HBAs in your ESX hosts.

3 Configure the storage system.

  a Create a volume (or LUN) on the storage system for ESX to boot from.

  b Configure the storage system so that the ESX system has access to the assigned LUN.

    This could involve updating ACLs with the IP addresses, iSCSI names, and the CHAP authentication parameter you use on the ESX system. On some storage systems, in addition to providing access information for the ESX host, you must also explicitly associate the assigned LUN with the host.

  c Ensure that the LUN is presented to the ESX system as LUN 0. The host can also boot from LUN 255.

    On storage systems that present volumes as multiple targets rather than multiple LUNs, the volumes are always presented as LUN 0.

  d Ensure that no other system has access to the configured LUN.

  e Record the iSCSI name and IP addresses of the targets assigned to the ESX host.

    You must have this information to configure your iSCSI HBA.

Configure iSCSI HBAs to Boot from a SAN

This topic discusses how to configure a QLogic iSCSI HBA for booting from a SAN.

On a system set up to boot from a SAN:

- The system BIOS must designate the iSCSI card as the boot controller.
- The BIOS must be enabled on the iSCSI HBA to locate the target boot LUN.

Procedure

1 During server POST, press Crtl+q to enter the QLogic iSCSI HBA configuration menu.

2 Select the I/O port to configure.

  By default, the Adapter Boot mode is set to Disable.

3 Configure the HBA.

  a From the **Fast!UTIL Options** menu, select **Configuration Settings > Host Adapter Settings**.

  b Configure the following settings for your host adapter: initiator IP address, subnet mask, gateway, initiator iSCSI name, and CHAP (if required).

4 Configure iSCSI Boot Settings.

   a From the **Fast!UTIL Options** menu, select **Configuration Settings > iSCSI Boot Settings**.

   b Before you can set SendTargets, set Adapter Boot mode to **Manual**.

   c Select **Primary Boot Device Settings**.

      1 Enter the discovery **Target IP** and **Target Port**.

      2 You can leave the **Boot LUN** and **iSCSI Name** fields blank if only one iSCSI target and one LUN are at the specified address to boot from. Otherwise, you must specify these fields to ensure that you do not boot from a volume for some other system. After the target storage system is reached, these fields will be populated after a rescan.

      3 Save changes.

   d From the **iSCSI Boot Settings** menu, select the primary boot device. An auto rescan of the HBA is made to find new target LUNS.

   e Select the iSCSI target.

   NOTE If more then one LUN exists within the target, you can choose a specific LUN ID by pressing **Enter** after you locate the iSCSI device.

   f Return to the **Primary Boot Device Setting** menu. After the rescan, the **Boot LUN**and **iSCSI Name** fields are populated. Change the value of **Boot LUN** to the desired LUN ID.

5 Save your changes and restart the system.

What to do next

For more information and more up-to-date details about QLogic host adapter configuration settings, see the QLogic host adapter readme file at the QLogic web site.

Managing ESX/ESXi Systems That Use SAN Storage 5

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, iSCSI adapters use iSCSI names.

When you display details for each iSCSI adapter, you see the following information. Certain adapters might need to be configured or enabled before you can view their information.

Table 5-1. Storage Adapter Information

Adapter Information	Description
Model	Model of the adapter.
Targets	Number of targets accessed through the adapter.
iSCSI Name	A unique name formed according to iSCSI standards that identifies the iSCSI adapter.
iSCSI Alias	A friendly name used instead of the iSCSI name.

Table 5-1. Storage Adapter Information (Continued)

Adapter Information	Description
IP Address	An address assigned to the iSCSI 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**.

Copy Storage Adapter Names to Clipboard

You can copy the name of an adapter to a clipboard directly from the UI.

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 In the Details panel, right-click the value in the name field, and select **Copy**.

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 5-2. Storage Device Information

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

Table 5-2. Storage Device Information (Continued)

Device Information	Description
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.

Details for each storage device include the following:

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

Understanding Storage Device Naming

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

Name

A friendly name that the host assigns to a device based on the storage type and manufacturer. You can modify the name using the vSphere Client.

Identifier

A universally unique identifier that the host extracts from the storage. Depending on the type of storage, the host uses different algorithms to extract the identifier. The identifier is persistent across reboots and is the same for all hosts sharing the device.

Runtime Name

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

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.

  Software iSCSI initiators use the channel number to show multiple paths to the same target.
- 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**.

Copy Storage Device Identifiers to Clipboard

A storage device identifier is a universally unique ID that the host assigns to a storage device or LUN. Depending on the type of storage, the host uses different algorithms to create the identifier and it can become quite long and complex. You can copy the storage device identifier directly from the UI.

Procedure

1. Display a list of storage devices.
2. Right-click a device and select **Copy identifier to clipboard**.

Viewing Datastore Information

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

The Datastores pane shows summary information about the datastore.

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

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

- Location of the datastore.
- Total capacity, including the used and available space.

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

Review Datastore Properties

Use the vSphere Client to review datastore properties.

Procedure

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

Resolving Display Issues

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

Perform the following troubleshooting tasks if you have display issues.

Table 5-3. Troubleshooting iSCSI LUN Display

Troubleshooting Task	Description
Check cable connectivity.	If you do not see a port, the problem could be cable connectivity or routing. 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 routing settings.	Controls connectivity between different subnets on your Ethernet configuration. If your ESX/ESXi system and iSCSI storage are not on the same subnet, ensure that appropriate routing exists between the subnets. Also, ensure that the subnet mask and gateway address are set correctly on the iSCSI storage and the iSCSI initiator in the ESX/ESXi host.
Check access control configuration.	If the expected LUNs do not appear after rescan, access control might not be configured correctly on the storage system side: ■ If CHAP is configured, ensure that it is enabled on the ESX/ESXi host and matches the storage system setup. ■ If IP-based filtering is used, ensure that the iSCSI HBA or the VMkernel port group IP address and service console IP address are allowed. ■ If you are using initiator name-based filtering, ensure that the name is a qualified iSCSI name and matches the storage system setup. ■ For booting from a SAN, ensure that each ESX host sees only required LUNs. Do not allow any ESX host to see any boot LUN other than its own. Use storage system software to make sure that the ESX host can see only the LUNs that it is supposed to see. ■ Ensure that the **Disk.MaxLUN** setting allows you to view the LUN you expect to see.
Check storage processor setup.	If a storage system has more than one storage processor, make sure that the SAN switch has a connection to the SP that owns the LUNs you want to access. On some storage systems, only one SP is active and the other SP is passive until a failure occurs. If you are connected to the wrong SP (the one with the passive path) you might not see the expected LUNs, or you might see the LUNs but get errors when trying to access them.

Table 5-3. Troubleshooting iSCSI LUN Display (Continued)

Troubleshooting Task	Description
For software iSCSI, check network configuration.	The software iSCSI initiator in ESX/ESXi requires that a VMkernel network port have access to the iSCSI storage. The software initiator uses the VMkernel for data transfer between the ESX/ESXi system and the iSCSI storage.
Rescan your iSCSI initiator.	Perform a rescan each time you complete the following tasks: ■ Create new LUNs on a SAN. ■ Change the LUN masking on an ESX/ESXi host storage system. ■ Reconnect a cable. ■ Make a change to a host in a cluster. ■ Change CHAP settings or add new discovery addresses.

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 use storage adapter rescan to scan storage devices and datastores.

Typically, your host or the vCenter Server automatically rescans storage adapters and updates storage devices and VMFS datastores after you make configuration changes. In certain cases, you need to manually rescan storage adapters.

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

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

- Create new LUNs on a SAN.
- Change the path masking on a host.
- Reconnect a cable.
- Make a change to a host in a cluster.
- Change CHAP settings or add new discovery addresses.
- 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 Do not rescan when a path is unavailable. If one path fails, another takes over and your system continues to be fully functional. If, however, you rescan at a time when a path is not available, the host removes the path from its list of paths to the device. The path cannot be used by the host until the next time a rescan is performed while the path is active.

Rescan Storage Adapters

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

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

Procedure

1 In the vSphere Client, select a host and click the **Configuration** tab.

2 In the Hardware panel, select **Storage Adapters**, and click **Rescan** above the Storage Adapters panel.

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

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

3 To discover new disks or LUNs, select **Scan for New Storage Devices**.

If new LUNs are discovered, they appear in the device list.

4 To discover new datastores or update a datastore after its configuration has been changed, select **Scan for New VMFS Volumes**.

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

Change the Number of Scanned LUNs

By default, the VMkernel scans for LUN 0 to LUN 255 for every target (a total of 256 LUNs). You can modify the `Disk.MaxLUN` parameter to improve LUN discovery speed.

IMPORTANT You cannot discover LUNs with a LUN ID number that is greater than 255.

Reducing the value can shorten rescan time and boot time. However, the time to rescan LUNs might depend on other factors, including the type of storage system and whether sparse LUN support is enabled.

Procedure

1 In the vSphere Client inventory panel, select the host, click the **Configuration** tab, and click **Advanced Settings**.

2 Select **Disk**.

3 Scroll down to **Disk.MaxLUN**.

4 Change the existing value to the value of your choice, and click **OK**.

The value you enter specifies the LUN after the last one you want to discover.

For example, to discover LUNs from 0 through 31, set **Disk.MaxLUN** to 32.

Disable Sparse LUN Support

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

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

Procedure

1 In the vSphere Client inventory panel, select the host, click the **Configuration** tab, and click **Advanced Settings**.

2 In the Advanced Settings dialog box, select **Disk**.

3 Scroll down to **Disk.SupportSparseLUN**, change the value to 0, and click **OK**.

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

Use the vSphere Client to view which SATP and PSP the host is using for a specific storage device and the status of all available paths for this storage device. If needed, you can change the default VMware PSP using the vSphere Client. To change the default SATP, you need to modify claim rules using the vSphere CLI.

You can find some information about modifying claim rules in Appendix C, "Managing Storage Paths and Multipathing Plug-Ins," on page 85.

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

Viewing the Paths Information

Use the vSphere Client to determine which SATP and PSP the ESX/ESXi host uses for a specific storage device and the status of all available paths for this storage device. You can access the path information from both, the Datastores and Devices views. For datastores, you review the paths that connect to the device the datastore is deployed on.

The path information includes the SATP assigned to manage the device, the path selection policy (PSP), and a list of paths with their physical characteristics, such as an adapter and target each path uses, and the status of each path. The following path status information can appear:

Active Paths available for issuing I/O to a LUN. A single or multiple working paths currently used for transferring data are marked as Active (I/O).

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

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

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

Broken The software cannot connect to the disk through this path.

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

View Datastore Paths

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

Procedure

1 Log in to the vSphere Client and select a server from the inventory panel.

2 Click the **Configuration** tab and click **Storage** in the Hardware panel.

3 Click **Datastores** under View.

4 From the list of configured datastores, select the datastore whose paths you want to view or configure.

The Details panel shows the total number of paths being used to access the device and whether any of them are broken or disabled.

5 Click **Properties > Manage Paths** to open the Manage Paths dialog box.

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

View Storage Device Paths

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

Procedure

1 Log in to the vSphere Client and select a server from the inventory panel.

2 Click the **Configuration** tab and click **Storage** in the Hardware panel.

3 Click **Devices** under View.

4 Click **Manage Paths** to open the Manage Paths dialog box.

Setting a Path Selection Policy

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

By default, VMware supports the following path selection policies. If you have a third-party PSP installed on your host, its policy also appears on the list.

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

Most Recently Used (VMware)
The host uses a path to the disk until the path becomes unavailable. When the path becomes unavailable, the host selects one of the alternative paths. The host does not revert back to the original path when that path becomes available again. There is no preferred path setting with the MRU policy. MRU is the default policy for active-passive storage devices and is required for those devices.

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

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

Table 5-4 summarizes how the behavior of host changes, depending on the type of array and the failover policy.

Table 5-4. Path Policy Effects

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

Change the Path Selection Policy

Generally, you do not have to change the default multipathing settings your host uses for a specific storage device. However, if you want to make any changes, you can use the Manage Paths dialog box to modify a path selection policy and specify the preferred path for the Fixed policy.

Procedure

1 Open the Manage Paths dialog box either from the Datastores or Devices view.

2 Select a path selection policy.

By default, VMware supports the following path selection policies. If you have a third-party PSP installed on your host, its policy also appears on the list.

- **Fixed (VMware)**
- **Most Recently Used (VMware)**
- **Round Robin (VMware)**

3 For the fixed policy, specify the preferred path by right-clicking the path you want to assign as the preferred path, and selecting **Preferred**.

4 Click **OK** to save your settings and exit the dialog box.

Disable Paths

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

Procedure

1 Open the Manage Paths dialog box either from the Datastores or Devices view.

2 In the Paths panel, right-click the path to disable, and select **Disable**.

3 Click **OK** to save your settings and exit the dialog box.

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

Path Management and Manual, or Static, Load Balancing

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

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

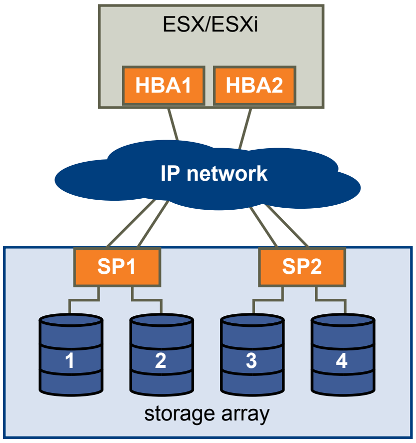
With active/active storage arrays, you can configure your ESX/ESXi host to load balance traffic across multiple adapters by assigning preferred paths to your LUNs. Path policy must be set to Fixed.

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

Assume the following setup, shown in Figure 5-1.

- Active/Active SPs
- An ESX/ESXi system
- Two iSCSI HBAs

Figure 5-1. Manual Load Balancing with iSCSI



For load balancing, set the preferred paths as follows.

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

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

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

Set Guest Operating System Timeout

Increase the standard disk timeout value so that a Windows guest operating system is not extensively disrupted during a path failover.

Path failover occurs 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.

I/O might pause for 30 to 60 seconds until the iSCSI driver determines that the link is unavailable and until failover is complete. 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 multiple breakages, all connections to SAN storage devices might be lost. If none of the connections to the storage device is working, some virtual machines might encounter I/O errors on their virtual SCSI disks.

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

Procedure

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

   After you make this change, Windows waits at least 60 seconds for delayed disk operations to complete before it generates errors.
6. Click **OK** to exit the **Registry Editor**.

Sharing Diagnostic Partitions

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

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

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

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

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

Avoiding and Resolving SAN Problems

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

You should observe these tips for avoiding and resolving problems with your SAN configuration:

- Place only one VMFS datastore on each LUN. Multiple VMFS datastores on one LUN is not recommended.
- Do not change the path policy the system sets for you unless you understand the implications of making such a change. In particular, working with an active-passive array and setting the path policy to **Fixed** can lead to path thrashing.
- Document everything. Include information about configuration, access control, storage, switch, server and iSCSI HBA configuration, software and firmware versions, and storage cable plan.

- Plan for failure:
  - Make several copies of your topology maps. For each element, consider what happens to your SAN if the element fails.
  - Cross off different links, switches, HBAs and other elements to ensure you did not miss a critical failure point in your design.
- Ensure that the iSCSI HBAs are installed in the correct slots in the 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, Ethernet switch statistics, and storage performance statistics.
- Be cautious when changing IDs of the LUNs that have VMFS datastores being used by your ESX/ESXi host. If you change the ID, virtual machines running on the VMFS datastore will fail.

  If there are no running virtual machines on the VMFS datastore, after you change the ID of the LUN, you must use rescan to reset the ID on your host. For information on using rescan, see "Rescan Storage Adapters," on page 62.

Optimizing SAN Storage Performance

Several factors contribute to optimizing a typical SAN environment.

If the network environment is properly configured, the iSCSI components provide adequate throughput and low enough latency for iSCSI initiators and targets. If the network is congested and links, switches or routers are saturated, iSCSI performance suffers and might not be adequate for ESX/ESXi environments.

Storage System Performance

Storage system performance is one of the major factors contributing to the performance of the entire iSCSI environment.

If issues occur with storage system performance, consult your storage system vendor's documentation for any relevant information.

When you assign LUNs, remember that you can access each LUN through 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 that contains the ESX/ESXi LUNs should not include LUNs that other hosts use that are not running ESX/ESXi for I/O intensive applications.

Enable read caching and write caching.

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

SAN storage systems require continual redesign and tuning to ensure that I/O is load balanced across all storage system paths. To meet this requirement, distribute the paths to the LUNs among all the SPs to provide optimal load balancing. Close monitoring indicates when it is necessary to manually rebalance the LUN distribution.

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

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 system. To achieve performance goals, perform the following tasks:

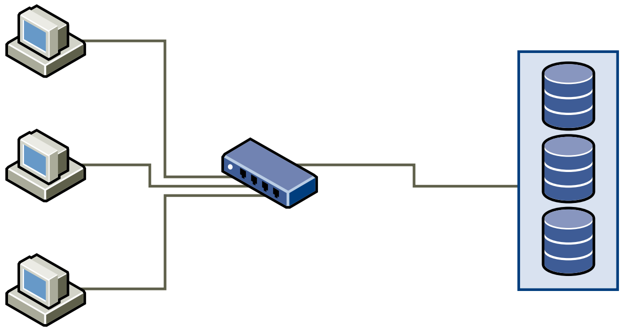
- Place each LUN on a RAID group that provides the necessary performance levels. Pay attention to the activities and resource utilization of other LUNS in the assigned RAID group. A high-performance RAID group that has too many applications doing I/O to it might not meet performance goals required by an application running on the ESX/ESXi host.
- Provide each server with a sufficient number of network adapters or iSCSI hardware adapters to allow maximum throughput for all the applications hosted on the server for the peak period. I/O spread across multiple ports provides higher throughput and less latency for each application.
- To provide redundancy for software iSCSI, make sure the initiator is connected to all network adapters used for iSCSI connectivity.
- When allocating LUNs or RAID groups for 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 SAN can handle.

Network Performance

A typical SAN consists of a collection of computers connected to a collection of storage systems through a network of switches. Several computers often access the same storage.

Figure 5-2 shows several computer systems connected to a storage system through an Ethernet switch. In this configuration, each system is connected through a single Ethernet link to the switch, which is also connected to the storage system through a single Ethernet link. In most configurations, with modern switches and typical traffic, this is not a problem.

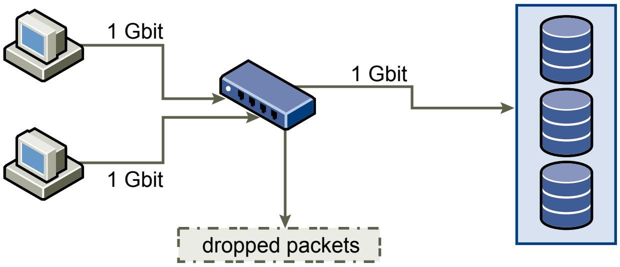
Figure 5-2. Single Ethernet Link Connection to Storage



When systems read data from storage, the maximum response from the storage is to send enough data to fill the link between the storage systems and the Ethernet switch. It is unlikely that any single system or virtual machine gets full use of the network speed, but this situation can be expected when many systems share one storage device.

When writing data to storage, multiple systems or virtual machines might attempt to fill their links. As Figure 5-3 shows, when this happens, the switch between the systems and the storage system has to drop data. This happens because, while it has a single connection to the storage device, it has more traffic to send to the storage system than a single link can carry. In this case, the switch drops network packets because the amount of data it can transmit is limited by the speed of the link between it and the storage system.

Figure 5-3. Dropped Packets



Recovering from dropped network packets results in large performance degradation. In addition to time spent determining that data was dropped, the retransmission uses network bandwidth that could otherwise be used for current transactions.

iSCSI traffic is carried on the network by the Transmission Control Protocol (TCP). TCP is a reliable transmission protocol that ensures that dropped packets are retried and eventually reach their destination. TCP is designed to recover from dropped packets and retransmits them quickly and seamlessly. However, when the switch discards packets with any regularity, network throughput suffers significantly. The network becomes congested with requests to resend data and with the resent packets, and less data is actually transferred than in a network without congestion.

Most Ethernet switches can buffer, or store, data and give every device attempting to send data an equal chance to get to the destination. This ability to buffer some transmissions, combined with many systems limiting the number of outstanding commands, allows small bursts from several systems to be sent to a storage system in turn.

If the transactions are large and multiple servers are trying to send data through a single switch port, a switch's ability to buffer one request while another is transmitted can be exceeded. In this case, the switch drops the data it cannot send, and the storage system must request retransmission of the dropped packet. For example, if an Ethernet switch can buffer 32KB on an input port, but the server connected to it thinks it can send 256KB to the storage device, some of the data is dropped.

Most managed switches provide information on dropped packets, similar to the following:

```
*: interface is up
IHQ: pkts in input hold queue     IQD: pkts dropped from input queue
OHQ: pkts in output hold queue    OQD: pkts dropped from output queue
RXBS: rx rate (bits/sec)          RXPS: rx rate (pkts/sec)
TXBS: tx rate (bits/sec)          TXPS: tx rate (pkts/sec)
TRTL: throttle count
```

Table 5-5. Sample Switch Information

Interface	IHQ	IQD	OHQ	OQD	RXBS	RXPS	TXBS	TXPS	TRTL
*GigabitEthernet 0/1	3	9922	0	0	476303000	62273	477840000	63677	0

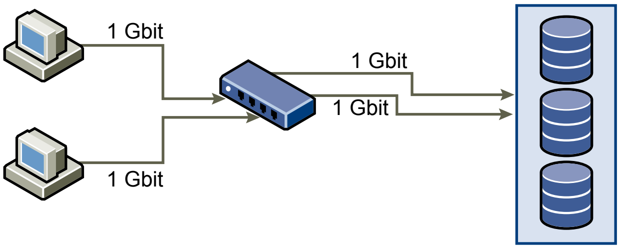
In this example from a Cisco switch, the bandwidth used is 476303000 bits/second, which is less than half of wire speed. In spite of this, the port is buffering incoming packets and has dropped quite a few packets. The final line of this interface summary indicates that this port has already dropped almost 10,000 inbound packets in the IQD column.

Configuration changes to avoid this problem involve making sure several input Ethernet links are not funneled into one output link, resulting in an oversubscribed link. When a number of links transmitting near capacity are switched to a smaller number of links, oversubscription is a possibility.

Generally, applications or systems that write a lot of data to storage, such as data acquisition or transaction logging systems, should not share Ethernet links to a storage device. These types of applications perform best with multiple connections to storage devices.

Figure 5-4 shows multiple connections from the switch to the storage.

Figure 5-4. Multiple Connections from Switch to Storage



Using VLANs or VPNs does not provide a suitable solution to the problem of link oversubscription in shared configurations. VLANs and other virtual partitioning of a network provide a way of logically designing a network, but do not change the physical capabilities of links and trunks between switches. When storage traffic and other network traffic end up sharing physical connections, as they would with a VPN, the possibility for oversubscription and lost packets exists. The same is true of VLANs that share interswitch trunks. Performance design for a SANs must take into account the physical limitations of the network, not logical allocations.

Resolving Performance Issues

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

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

Checking Ethernet Switch Statistics

Many Ethernet switches provide different methods for monitoring switch health.

Switches that have ports operating near maximum throughput much of the time do not provide optimum performance. If you have ports in your iSCSI SAN running near the maximum, reduce the load. If the port is connected to an ESX/ESXi system or iSCSI storage, you can reduce the load by using manual load balancing.

If the port is connected between multiple switches or routers, consider installing additional links between these components to handle more load. Ethernet switches also commonly provide information about transmission errors, queued packets, and dropped Ethernet packets. If the switch regularly reports any of these conditions on ports being used for iSCSI traffic, performance of the iSCSI SAN will be poor.

Resolving Path Thrashing

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

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

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

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

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

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

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

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

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, only one LUN at a time can access all the sectors on the storage that make up a given LUN. The ownership is passed between the storage processors. Storage systems 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, neither SP can process I/O to the LUN.

For active/active arrays, the algorithms allow more fine-grained access to the storage and synchronize caches. Access can happen concurrently through any SP without extra time required.

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.

Resolve Path Thrashing

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

Procedure

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

Equalize Disk Access Between Virtual Machines

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

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

Procedure

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

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

What to do next

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

Reducing SCSI Reservations

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

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

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

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

Setting Maximum Queue Depth for Software iSCSI

If you notice unsatisfactory performance for your software iSCSI LUNs, you can change their maximum queue depth by using the `vicfg-module` command.

On the vSphere CLI, run the following command:

```
vicfg-module -s iscsi_max_lun_queue=value iscsi_mod
```

After you issue this command, reboot your system.

The `iscsi_max_lun_queue` parameter is used to set the maximum outstanding commands, or queue depth, for each LUN accessed through the software iSCSI adapter. The default is 32, and the valid range is 1 to 255.

Caution Setting the queue depth higher than the default can decrease the total number of LUNs supported.

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.

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

Array-Based (Third-Party) Solution

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

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

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

File-Based (VMFS) Solution

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

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

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

Managing Duplicate VMFS Datastores

When a LUN contains a VMFS datastore copy, you can mount the datastore with the existing signature or assign a new signature.

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

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

Mounting VMFS Datastores with Existing Signatures

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

For example, you can maintain synchronized copies of virtual machines at a secondary site as part of a disaster recovery plan. In the event of a disaster at the primary site, you can mount the datastore copy and power on the virtual machines at the secondary site.

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

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

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

Mount a VMFS Datastore with an Existing Signature

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

Prerequisites

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

Procedure

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

   The name present in the VMFS Label column indicates that the LUN is a copy that contains a copy of an existing VMFS datastore.
6. Under Mount Options, select **Keep Existing Signature**.
7. In the Ready to Complete page, review the datastore configuration information and click **Finish**.

What to do next

If you later want to resignature the mounted datastore, you must unmount it first.

Unmount Datastores

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

You can unmount only the following types of datastores:

- NFS datastores
- VMFS datastore copies mounted without resignaturing

Procedure

1. Display the datastores.
2. Right-click the datastore to unmount and select **Unmount**.
3. If the datastore is shared, specify which hosts should no longer access the datastore.
   a. If needed, deselect the hosts where you want to keep the datastore mounted.

      By default, all hosts are selected.
   b. Click **Next**.
   c. Review the list of hosts from which to unmount the datastore, and click **Finish**.
4. Confirm that you want to unmount the datastore.

Resignaturing VMFS Copies

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

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

When you perform datastore resignaturing, consider the following points:

- Datastore resignaturing is irreversible.
- The LUN copy that contains the VMFS datastore that you resignature is no longer treated as a LUN copy.
- A spanned datastore can be resignatured only if all its extents are online.
- The resignaturing process is crash and fault tolerant. If the process is interrupted, you can resume it later.
- You can mount the new VMFS datastore without a risk of its UUID colliding with UUIDs of any other datastore, such as an ancestor or child in a hierarchy of LUN snapshots.

Resignature a VMFS Datastore Copy

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

Prerequisites

To resignature a mounted datastore copy, first unmount it.

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

Procedure

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

   The name present in the VMFS Label column indicates that the LUN is a copy that contains a copy of an existing VMFS datastore.
6. Under Mount Options, select **Assign a New Signature** and click **Next**.
7. In the Ready to Complete page, review the datastore configuration information and click **Finish**.

What to do next

After resignaturing, you might have to do the following:

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

iSCSI SAN Configuration Checklist A

Different storage systems and ESX/ESXi hosts have specific setup requirements.

Table A-1. iSCSI SAN Configuration Requirements

Component	Comments
All storage systems	Write cache must be disabled if not battery backed.
Topology	No single failure should cause HBA and SP failover, especially with active-passive storage arrays.
EMC Symmetrix	Enable the SPC2 and SC3 settings. Contact EMC for the latest settings.
EMC Clariion	Set the EMC Clariion failover mode to 1 or 4. Contact EMC for details.
HP MSA	No specific requirements
HP EVA	For EVA3000/5000 firmware 4.001 and later, and EVA4000/6000/8000 firmware 5.031 and later, set the host type to `VMware`. Otherwise, set the host mode type to `Custom`. The value is: ■ EVA3000/5000 firmware 3.x: 000000002200282E ■ EVA4000/6000/8000: 000000202200083E
NetApp	If any of your iSCSI initiators are a part of an initiator group (igroup), disable ALUA on the NetApp array.
EqualLogic	Make sure ARP Redirect is enabled on hardware iSCSI adapters.
LeftHand	Make sure ARP Redirect is enabled on hardware iSCSI adapters.
ESX/ESXi Configuration	Set the following Advanced Settings for the ESX/ESXi host: ■ Set **Disk.UseLunReset** to **1** ■ Set **Disk.UseDeviceReset** to **0** A multipathing policy of Most Recently Used must be set for all LUNs hosting clustered disks for active-passive arrays. A multipathing policy of Most Recently Used or Fixed may be set for LUNs on active-active arrays. Allow ARP redirection if the storage system supports transparent failover.

VMware vSphere Command-Line Interface

B

In most cases, the vSphere Client is well suited for monitoring an ESX/ESXi host connected to SAN storage. Advanced users might want to use some VMware vSphere command-line interface (vSphere CLI) commands for additional functionality.

For more information, see the *VMware vSphere Command-Line Interface Installation and Reference Guide*.

This appendix includes the following topics:

resxtop Command

The `resxtop` command provides a detailed look at ESX/ESXi resource use in real time.

For detailed information about `resxtop`, see the *Resource Management Guide* and *VMware vSphere Command-Line Interface Installation and Reference Guide*.

vicfg-iscsi Command

The `vicfg-iscsi` command allows you to configure software or hardware iSCSI on ESX/ESXi hosts, set up CHAP parameters, and set up iSCSI networking.

For details, see the *VMware vSphere Command-Line Interface Installation and Reference Guide*.

vicfg-mpath Command

Use the `vicfg-mpath` command to view information about storage devices, paths, and multipathing plug-ins.

For details, see the *VMware vSphere Command-Line Interface Installation and Reference Guide*.

esxcli corestorage claimrule Command

Use the `esxcli corestorage claimrule` command to manage claim rules. Claim rules determine which multipathing module should claim paths to a particular device and manage the device.

For details, see the *VMware vSphere Command-Line Interface Installation and Reference Guide*.

vmkping Command

The vmkping command allows you to verify the VMkernel networking configuration.

Usage example:

```
vmkping [options] [host|IP address]
```

Table B-1. vmkping Command-Line Options

Option	Description
-6	Use IPv6 - ICMPv6 Echo request.
-4	Use IPv4 (default).
-I	Outgoing interface - for IPv6 scope.
-D	VMkernel TCP stack debug mode.
-c <count>	Sets packet count.
-i <interval>	Sets interval.
-s <size>	Sets send size.

Managing Storage Paths and Multipathing Plug-Ins

C

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

You can use the vSphere CLI to 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 additional commands available to manage PSA, see the *vSphere Command-Line Interface Installation and Reference Guide*.

This appendix includes the following topics:

List Claim Rules for the Host

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

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` to list claim rules.

  Example C-1 shows the output of the command.

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

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

This example indicates the following:

- The NMP claims all paths connected to storage devices that use the USB, SATA, IDE, and Block SCSI transportation.
- The MASK_PATH module claims all paths returning SCSI inquiry data with a vendor string of DELL and a model string of Universal Xport. The MASK_PATH module is used to mask paths from your host.
- The MPP_1 module claims all paths connected to any model of the NewVend storage array.
- The MPP_3 module claims the paths to storage devices controlled by the Mega-RAID device driver.
- Any paths not described in the previous rules are claimed by NMP.
- The Class column in the output shows which rules are defined and which are loaded. The `file` parameter in the Class column indicates that the rule is defined. The `runtime` parameter indicates that the rule has been loaded into your system. For a user- defined claim rule to be active, two lines with the same rule number should exist, one line for the rule with the `file` parameter and another line with `runtime`. Several low numbered rules have only one line with the Class of `runtime`. These are system defined claim rules that you cannot modify.

Display Multipathing Modules

Use the vSphere CLI to list all multipathing modules loaded into the system. Multipathing modules manage physical paths that connect your host with storage.

Procedure

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

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

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

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

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

```
MPP_1
MPP_2
MPP_3
MASK_PATH
NMP
```

Display SATPs for the Host

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

Procedure

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

  ```
  esxcli nmp satp list
  ```

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

Keep in mind the following:

- If no SATP is assigned to the device by the claim rules, the default SATP for iSCSI or FC devices is VMW_SATP_DEFAULT_AA. The default PSP is VMW_PSP_FIXED.
- If VMW_SATP_ALUA is assigned to a specific storage device, but the device is not ALUA-aware, there is no claim rule match for this device. In this case, the device is claimed by the default SATP based on the device's transport type.
- The default PSP for all devices claimed by VMW_SATP_ALUA is VMW_PSP_MRU. The VMW_PSP_MRU selects an active/optimized path as reported by the VMW_SATP_ALUA, or an active/unoptimized path if there is no active/optimized path. This path is used until a better path is available (MRU). For example, if the VMW_PSP_MRU is currently using an active/unoptimized path and an active/optimized path becomes available, the VMW_PSP_MRU will switch the current path to the active/optimized one.

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

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

Display NMP Storage Devices

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

Procedure

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

   ```
   esxcli nmp device list
   ```

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

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

Add PSA Claim Rules

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

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

CAUTION When creating new claim rules, be careful to avoid a situation when different physical paths to the same LUN are claimed by different MPPs. Unless one of the MPPs is the MASK_PATH MPP, this configuration will cause performance errors.

Procedure

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

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

For information on the options that the command requires, see "esxcli corestorage Command-Line Options," on page 92.

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

```
esxcli corestorage claimrule load
```

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

Example C-4. Adding a PSA Claim Rule

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

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

2 `# esxcli corestorage claimrule load`

If you now run the **esxcli corestorage claimrule list** command, you can see the new claim rule appearing on the list.

NOTE The two lines for the claim rule, one with the Class of `runtime` another with the Class of `file`, indicate that the new claim rule has been loaded into the system and is active.

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

Delete PSA Claim Rules

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

Procedure

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

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

For information on the options that the command takes, see "esxcli corestorage Command-Line Options," on page 92.

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 -r <claimrule_ID> -t <type> <required_option> -P <MASK_PATH>
```

For information on command-line options, see "esxcli corestorage Command-Line Options," on page 92.

3 Load the MASK_PATH claim rule into your system.

```
esxcli corestorage claimrule load
```

4 Verify that the MASK_PATH claim rule was added correctly.

```
esxcli corestorage claimrule list
```

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

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

6 Run the path claiming rules.

```
esxcli corestorage claimrule run
```

After you assign the MASK_PATH 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 C-5. Masking a LUN

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

1 `#esxcli corestorage claimrule list`

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

3 `#esxcli corestorage claimrule load`

4 `#esxcli corestorage claimrule list`

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

6 `# esxcli corestorage claimrule run`

Unmask Paths

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

Procedure

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

   Run this command for each path to the storage device.

   For example:

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

2. Load path claiming rules into the VMkernel by running the `esxcli corestorage claimrule load` command.

3. Run the path claiming rules by entering the `esxcli corestorage claimrule run`.

Your host can now access the previously masked storage device.

Define NMP SATP Rules

The NMP SATP claim rules specify which SATP should manage a particular storage device. Usually you do not need to modify the NMP SATP rules. If you need to do so, use vSphere CLI to add a rule to the list of claim rules for the specified SATP.

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

Procedure

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

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

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

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

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

Specify the following options for any SATP claim rule:

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

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

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

3 Reboot your host.

Example C-6. 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.

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

esxcli corestorage Command-Line Options

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

Table C-1. esxcli corestorage Command-Line Options

Option	Description	Required Option
`-r <claimrule_ID>`	Use to specify the order number for the claim rule from 0 to 65535.	
`-t <type>`	Use to define the set of paths for the claim rule. Specify one of the following values for the <*type*> variable:	These options change depending on the value you enter for <*type*>.
	`vendor` – Indicate the vendor and model of the storage device used for this path.	`-V` *<vendor>* `-M` *<model>* Use asterisk (*) to specify all vendors or models.
	`location` – Indicate the adapter, channel, target, or LUN used for this path.	Use any of the following: ■ `-A` *<adapter>* ■ `-C` *<channel>* ■ `-T` *<target>* ■ `-L` *<lunID>*
	`driver` – Indicate the driver used for the path.	`-D` *<driver>*
	`transport` – Indicate the transport used for the path.	`-R` *<transport>* Use one of the following for the <transport> variable: ■ `block` – Raid block devices, such as cciss ■ `fc` – Fibre Channel ■ `iscsi` – Default iSCSI ■ `iscsivendor` – iSCSI with vendor supplied IMA ■ `ide` – IDE ■ `sas` – Serial attached SCSI ■ `sata` – Serial ATA ■ `usb` – USB storage devices ■ `parallel` – Parallel SCSI devices ■ `unknown` – Unknown storage device type
`-P <MPP_name>`	Indicate which MPP plug-in should claim the paths defined by the claim rule. Run the `vicfg-mpath --list-plugins` command to see valid values.	

Index