

EMC CLARiiON Integration with VMware ESX Server

Applied Technology

Abstract

This white paper provides a quick overview of how to connect an EMC® CLARiiON® storage system to a VMware ESX Server. It highlights the VMotion, VMware HA, and Distributed Resource Scheduling capabilities of the VMware ESX Server, as well as clustering of virtual machines when connected to a CLARiiON storage system. It provides a general overview of how the VMware ESX Server integrates with the CLARiiON storage system.

May 2008

Copyright © 2006, 2007, 2008 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

All other trademarks used herein are the property of their respective owners.

Part number H1416.6

Table of Contents

Executive summary	5
Introduction	5
Audience	5
CLARiiON terminology	5
VMware terminology	5
VMware ESX Server overview	6
VMware ESX Server features	7
VMware VMotion	7
Distributed Resource Scheduling and VMware High Availability	8
VMware Clustering	8
VMware ESX 3.x/ESX 3i Consolidated Backup	8
VMware N_Port ID Virtualization	9
VMware Site Recovery Manager (SRM)	9
EMC CLARiiON overview	9
Why use CLARiiON with VMware ESX Server?	11
CLARiiON integration with VMware	11
Basic connectivity	11
Booting from a CLARiiON storage system	13
Booting VMware ESX Server from CLARiiON LUNs with ESX 3.x and ESX 2.5.x	13
Booting guest operating systems on CLARiiON LUNs	14
Navisphere management	14
VMware native failover with CLARiiON	15
LUN partitioning	18
Raw disks	19
VMFS volumes	19
Raw device mapping	20
LUN layout recommendations	20
Using CLARiiON virtual LUN technology with ESX 3.x/ESX 3i and 2.x	21
Expanding and migrating LUNs used as raw device mapping	21
Expanding and migrating LUNs used as VMFS volumes	21
Using CLARiiON replication software with ESX 3.x/ESX 3i and 2.5.x	22
CLARiiON replication considerations with VMware ESX Server	22
CLARiiON replication software considerations when using VMFS volumes	22
CLARiiON replication software considerations when using RDM volumes	23
CLARiiON and VMotion	23
VMotion with VMFS volumes	24
VMotion with RDM volumes	24
CLARiiON with VMware Distributed Resource Scheduling and High Availability	27
CLARiiON and virtual machine clustering	27
In-the-box cluster	27
Out-of-the-box cluster	28
Virtual-to-virtual clustering	29
Physical-to-virtual clustering	30

CLARiiON and VMware Consolidated Backup	31
CLARiiON and VMware NPIV support	32
CLARiiON and VMware Site Recovery Manager (SRM)	34
SRM Protection Groups	38
SRM recovery plan	38
Testing the SRM recovery plan	38
Executing an SRM recovery plan	39
Failback scenarios	40
Conclusion	40
References	40
Appendix A: Copying data from a VMFS to RDM volume	41
Appendix B: Using vm-support on VMware ESX Server	42

Executive summary

EMC is aggressively expanding product sets from high-end to midtier markets. Through VMware—the industry leader of x86 server-virtualization software—and CLARiiON®, which offers the best performance in midtier storage, EMC is integrating cutting-edge virtualization technology into its core storage business.

Introduction

This white paper outlines the benefits of using VMware virtualization products with the CLARiiON storage system and how to combine features to complement one another. This paper also discusses the connectivity aspect of attaching a VMware ESX Server to the CLARiiON storage system.

Audience

This paper is intended for customers, partners, and EMC field personnel requiring information about the features, parameters, and configuration of the VMware ESX Server. It includes information about how these features integrate with the CLARiiON storage system. It is assumed that the audience is familiar with CLARiiON hardware and software products, and has a general idea of how VMware ESX Server works.

CLARiiON terminology

CLARiiON LUN — Logical subdivisions of RAID groups in a CLARiiON storage system.

MirrorView™ — Software designed for disaster recovery solutions by mirroring local production data to a remote disaster recovery site. It offers two complementary remote mirroring products: MirrorView/Synchronous and MirrorView/Asynchronous.

RAID groups — One or more disks grouped together under a unique identifier in a CLARiiON storage system.

SAN Copy™ — Data mobility software that runs on the CLARiiON.

SnapView™ — Software used to create replicas of the source LUN. These point-in-time replicas can be pointer-based snapshots or full binary copies called *clones* or *BCVs*.

VMware terminology

Bridge — A connection that links a virtual network adapter in your virtual machine to the physical Ethernet adapter in your host computer.

Cluster — A cluster within VirtualCenter 2.0 is a collection of ESX Server hosts and associated virtual machines that share resources and a management interface.

ESX Server — VMware's high-end server product that installs directly on the physical hardware and therefore offers the best performance. ESX Server supports more virtual machines per physical CPU than its other virtualization products such as VMware Server (previously called GSX server).

Farm or Data Center— The primary organizational structure used in VirtualCenter, which contains hosts and virtual machines. The term *Farm* is used with VirtualCenter 1.x while the term *Data Center* is used with VirtualCenter 2.x.

Guest operating system — An operating system that runs on a virtual machine.

ISO image — A CD image that can be downloaded and burnt on a CD-ROM or mounted as a loopback device.

Management User Interface (MUI) — A web-based graphical interface for VMware that manages a VMware ESX 2.5.x server.

Mapping file — A VMFS file containing metadata used to map and manage a raw device.

Network label — A unique name given to a virtual switch on the ESX Server.

Raw device mapping (RDM) – Raw device mapping volumes consist of a pointer in a .vmdk file and a physical raw device. The pointer in the .vmdk points to the physical raw device. The .vmdk file resides on a VMFS volume, which must reside on shared storage.

Service console (COS) — The modified Linux kernel that serves as the management interface to the ESX Server. Not to be confused with VMkernel.

Templates —A means to import virtual machines and store them as templates that can be deployed at a later time to create new virtual machines.

VirtualCenter — A virtual infrastructure management product that manages and provide valuable services for virtual machines and underlying virtualization platforms from a central, secure location.

Virtual machine — A virtualized x86 PC environment on which a guest operating system and associated application software can run. Multiple virtual machines can operate on the same physical machine concurrently.

Virtual machine configuration file — A file containing a virtual machine configuration that is created by the Configuration Wizard or the Configuration Editor. The VMware ESX Server uses this file to identify and run a specific virtual machine. It usually has a .vmx extension.

Virtual switch — A switch that allows virtual network interface cards (NICs) to communicate with one another. Additionally, you can connect one or more physical network adapters to a virtual switch, so that virtual machines can communicate with the outside world.

VMFS — A clustered file system that stores virtual disks and other files that are used by virtual machines.

VMkernel — A kernel that controls the server hardware and schedules virtual machine computations and I/O operations.

VMware ESX Server overview

VMware ESX Server consists of virtualization software that provides server consolidation by allowing several instances of similar and dissimilar operating systems to run as virtual machines on one physical machine. This cost-effective, highly scalable virtual machine platform offers advanced resource management capabilities. VMware ESX Server minimizes the total cost of ownership (TCO) of computing infrastructure by:

- Increasing resource utilization.
- Decreasing the number of servers and all associated costs.
- Maximizing server manageability.

Figure 1 shows the architecture of two VMware ESX Servers (ESX 1 and ESX 2) with virtual machines containing guest operating systems that sit on top of the virtualization layer.

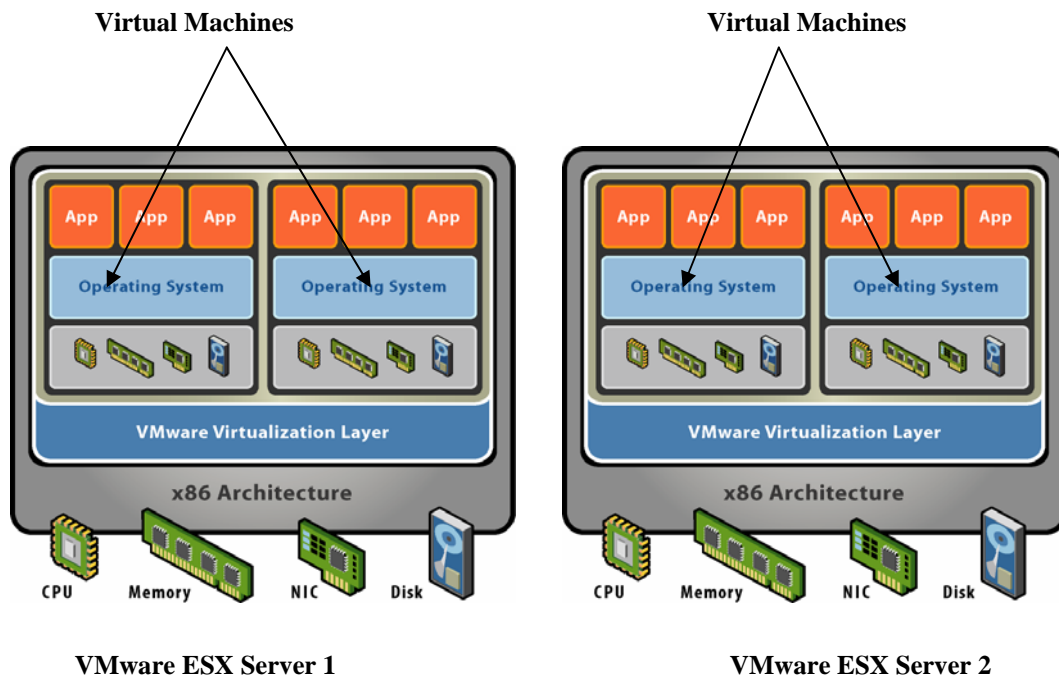


Figure 1. Architecture of VMware ESX Server

ESX Server runs directly on the hardware, and allows virtual machines to run on top of the virtualization layer provided by VMware. The combination of an operating system and applications is referred to as a virtual machine. ESX 2.x servers are managed using the Management User Interface (MUI). VirtualCenter, a VMware MUI, allows you to manage a number of ESX Servers, as well as to perform operations such as VMotion. VirtualCenter 1.x is used to manage one or more ESX 2.x servers. No MUI is available for ESX 3.x management. VirtualCenter 2.0 is used to manage a single ESX 3.x server or multiple ESX (3.x and 2.x) servers.

VMware ESX 3.5 and ESX 3i were introduced in December 2007. VMware ESX 3.5 is similar to the previous versions of ESX and includes a service console (COS) to boot ESX Server. VMware ESX 3i is a “thin” version that does not include a service console. The management of the ESX 3i version is done using a VirtualCenter server or a client whose operating system is embedded in the hardware or can be installed on a hard disk. On Windows and Linux platforms, you can also use Remote CLI to issue commands directly to the ESX 3i server. For more information on ESX 3.5/ESX 3i, please see www.vmware.com.

VMware ESX Server features

The ESX Server has several features that work with the CLARiiON storage system. The four features discussed in this paper are VMotion, Distributed Resource Scheduling and VMware HA, VMware Clustering, and Consolidated Backup technology. Distributed Resource Scheduling, VMware HA, and Consolidated Backup are features introduced in ESX 3.x/ESX 3i.

VMware VMotion

ESX Server version 2.0.1 was the first platform to support VMotion. System administrators can use VMotion, a systems management and provisioning product that works through VMware VirtualCenter, to quickly provision and reprovision servers with any number of virtual machines.

VMotion technology provides the ability to migrate a running virtual machine from one physical ESX Server to another—without application service interruption—allowing for fast reconfiguration and

optimization of resources without impacting users. With VMotion, VMware allows administrators to move virtual machine partitions from machine to machine on the fly, as real workloads run in the partitions. This allows administrators to do hardware maintenance without interrupting applications and users. It also allows the administrator to do dynamic load balancing to maintain high utilization and performance.

Storage VMotion

This feature, which was also introduced in ESX 3.5/3i, allows you to migrate a virtual machine from one storage system to another while the virtual machine is up and running. For example, using the CLI interface, virtual machine files can be moved from one FC LUN to another FC LUN without taking the virtual machine offline. Storage VMotion is supported for VMFS volumes and qualified for FC-to-FC storage. For more information on VMFS, see the “LUN partitioning” section.

Distributed Resource Scheduling and VMware High Availability

The VMware Distributed Resource Scheduling (DRS) feature improves resource allocation across all hosts by collecting resource (such as CPU and memory) usage information for all hosts and virtual machines in the cluster and generating recommendations for virtual machine placement. These recommendations can be applied automatically or manually. Depending on the configured DRS automation level, DRS can display or automatically implement recommendations. The result is a self-managing, highly optimized, highly efficient computer cluster with built-in resource and load balancing. In VMware 3.5/3i, VMware’s Distributed Power Management reduces power consumption by intelligently balancing a data center’s workload. Distributed Power Management, which is part of VMware DRS, automatically powers off servers whose resources are not immediately required and returns power to these servers when they are needed. Support for DRS Distributed Power Management is experimental only.

VMware High Availability (HA) detects ESX hardware failures and automatically restarts virtual machines and their resident applications and services on alternate ESX hardware, enabling servers to recover more rapidly and deliver a higher level of availability. Using VMware HA and DRS together combines automatic failover with load balancing. This combination results in a fast rebalancing of virtual machines after HA has moved virtual machines to different hosts. In VMware Infrastructure 3.5/3i, enhanced HA provides experimental support for monitoring individual virtual machine failures. VMware HA can now be configured to restart the failed virtual machine or send a notification to the administrator.

VMware Clustering

The ESX Server can be clustered at a virtual machine level within a single ESX Server (referred to as an *in-the-box-cluster*) or between two or more ESX Servers (referred to as an *outside-the-box-cluster*). The cluster setup within a box is useful for providing high availability when software or administrative errors are the likely causes of failure. Users who want a higher level of protection in the event of hardware failures, as well as software/logical failures, benefit from clustering outside the box.

VMware ESX 3.x/ESX 3i Consolidated Backup

ESX Server version 3.x/ESX 3i offers LAN-free backup (without running backup agents on virtual machines) with higher performance and greater manageability using VMware Consolidated Backup technology. Backup of a virtual machine begins with a built-in, automatic quiescing of the virtual disk; no backup agents are required in the virtual machines. Next, ESX Server creates a *virtual disk snapshot*: a hot, online disk snapshot of the virtual machine’s disk through the ESX Server’s VMkernel. Finally, a separate dedicated proxy server (physical server), which is optimized for I/O, reads the virtual disk snapshot. With this procedure, the customer’s standard backup software agent on that proxy server backs the data up with optimal performance, fully offloaded from ESX Servers, and with just one agent to manage. Backing up to a physical server also allows the use of Fibre Channel tape devices instead of transferring everything over the network. Consolidated Backup can perform file-level backups for virtual machines with Windows guest operating systems. It can also perform image-level backups for any virtual machine regardless of guest operating system.

VMware N_Port ID Virtualization

N_Port ID Virtualization (NPIV) within the Fibre Channel protocol allows multiple virtual N_Port IDs to share a single physical N_Port. In other words, you can define multiple virtual initiators through a single initiator. This feature, which was introduced with ESX 3.5/3i, enables SAN tools that provide QoS at the storage system level to guarantee service levels for virtual machine applications.

Within VMware ESX, NPIV is enabled for each virtual machine, so that physical HBAs on ESX Server can assign virtual initiators to each virtual machine. As a result, within ESX Server, a virtual machine has virtual initiators (WWNs) available for each HBA. These initiators can log in to the storage like any other host. VMware NPIV support is limited to RDM volumes. For more details about this configuration, please see the “CLARiiON and VMware NPIV support” section.

VMware Site Recovery Manager (SRM)

VMware Site Recovery Manager (SRM) integrates various EMC replication software products (such as MirrorView/S) to automate the failover process for virtual machines. SRM centralizes the creation and management of the disaster recovery strategies that are implemented at the secondary site. SRM uses EMC’s array-based snapshot technologies to test the failover process, and to ensure that the recovery image is consistent.

SRM requires that the protected (primary) site and the recovery (secondary) site each has two independent virtual infrastructure servers to facilitate the failover process. Array-based Site Recovery Adapters (SRAs) are also installed at both sites; these SRAs communicate with the storage systems (arrays). For more information about using SRM with CLARiiON storage systems, please see the “CLARiiON and VMware Site Recovery Manager (SRM)” section

EMC CLARiiON overview

The EMC CLARiiON family of networked storage systems brings best-in-class performance to the midtier with a wide range of storage solutions—all based on the powerful, proven, eight generations of CLARiiON architecture. They provide multiple tiers of storage (both Fibre Channel and SATA) in a single storage system, which significantly reduces acquisition costs and management costs by allowing multiple tiers to be managed with a single management interface. The next-generation CLARiiON systems, called the CX3 UltraScale™ series, provide an end-to-end 4 Gb/s design that is optimized to deliver native 4 Gb/s performance. Products with multiple back ends such as the CX3-40 and CX3-80 can support disks operating at both 2 Gb/s and 4 Gb/s simultaneously.

CLARiiON storage systems address a wide range of storage requirements by providing flexible levels of capacity, functionality, and performance. The AX4 is an entry-level system that consists of single-controller and dual-controller models. It supports both Serial Attached SCSI (SAS) and SATA drives and connectivity for up to 64 high availability (HA) connected hosts. The CX3-20 supports up to 120 drives and connectivity for up to 128 HA hosts. The CX3-40 storage system expands the family, supporting up to 128 HA hosts and up to 240 drives. The high-end CX3-80 adds even more capability, supporting up to 256 HA hosts and up to 480 drives. For midsize environments that require the economy and familiarity of IP networking through iSCSI host connectivity, the EMC CLARiiON CX3-40C, CX3-20C, and AX4 networked storage systems are available. Table 1 and Table 2 on page 10 summarize the basic features for the CLARiiON CX3, CX, AX, and iSCSI series storage systems.

Table 1. CLARiiON CX3 and AX4 (Fibre Channel) storage systems feature summary

Feature	CX3-80	CX3-40	CX3-20	AX4
Maximum disks	480	240	120	60
Storage processors (SP)	2	2	2	1 or 2
Front-end FC ports/SP	4 @ 4 Gb/s	2 @ 4 Gb/s	6 @ 4 Gb/s	2 @ 4 Gb/s
Back-end FC ports/SP	4 @ 4 Gb/s	2 @ 4 Gb/s	1 @ 4 Gb/s	1 @ 4 Gb/s
Cache	16 GB	8 GB	4 GB	2 GB
High availability hosts	256	128	128	10/64*
Minimum physical size	9U	5U	5U	2U
Maximum standard LUNs	2048	1024	512	512
SnapView snapshots	Yes	Yes	Yes	Yes
SnapView clones	Yes	Yes	Yes	Yes**
SAN Copy	Yes	Yes	Yes	Yes**
MirrorView/S	Yes	Yes	Yes	Yes**
MirrorView/A	Yes	Yes	Yes	Yes**

* Support for 10 hosts with the base pack and 64 hosts with the expansion enabler

** Available in Q1 2008 with the AX4 Service Pack 1

Table 2. CLARiiON CX3 and AX4 series (iSCSI) storage systems feature summary

Feature	CX3-40C	CX3-20C	CX3-10C	AX4i
Maximum disks	240	120	60	60
Storage processors (SP)	2	2	2	1 or 2
Front-end FC ports/SP	2 @ 4 Gb/s	2 @ 4 Gb/s	2 @ 4 Gb/s	N/A
Front-end iSCSI ports/SP	4 @ 1 Gb/s	4 @ 1 Gb/s	2 @ 1 Gb/s	2 @ 1 Gb/s
Back-end FC ports/SP	2 @ 4 Gb/s	1 @ 4 Gb/s	1 @ 4 Gb/s	1 @ 2 Gb/s
Cache	8 GB	4 GB	2 GB	2 GB
High availability hosts	128	128	64	10/64*
Minimum physical size	5U	5U	5U	2U
Maximum standard LUNs	1024	512	256	512
SnapView snapshots	Yes	Yes	Yes	Yes
SnapView clones	Yes	Yes	No	Yes**
SAN Copy	Yes	Yes	N/A	N/A
MirrorView/S	Yes	Yes	N/A	N/A

MirrorView/A	Yes	Yes	Yes	N/A
--------------	-----	-----	-----	-----

*Support for 10 hosts with the base pack and 64 hosts with the expansion enabler

**Available in Q1 2008 with the AX4 Service Pack 1

Why use CLARiiON with VMware ESX Server?

CLARiiON and VMware complement each other with the features that they provide. Some of the reasons CLARiiON is an ideal fit for VMware in the midrange storage market include:

- CLARiiON storage systems provide flexible levels of models including Fibre Channel and iSCSI systems. This allows the user to make the optimal choice of a storage system based on capacity, performance, and cost.
- CLARiiON storage systems can scale quickly to manage anticipated data growth, especially as the storage need for virtual machines increases on the VMware ESX Server.
- CLARiiON storage can be shared across multiple ESX Servers allowing storage consolidation to provide efficient use of storage resources, which is valuable for clustering and VMotion.
- Virtual machine applications running on CLARiiON storage systems enhance performance and therefore maximize functionality, reliability, and efficiency of the VMware ESX Server.
- Navisphere® Manager suite provides web-based centralized control of global disk space, availability, security, and quality-of-service for virtual machines provisioned by the CLARiiON storage system.
- The redundant architecture of the CLARiiON storage system provides no single point of failure, thereby reducing application downtime and minimizing business impact for storage upgrades.
- The CLARiiON storage system's modular architecture allows a mixture of FC and SATA drives. FC drives are used for I/O intensive applications, while SATA drives are used for backup and offloading old data, among other things.

CLARiiON integration with VMware

This section discusses how CLARiiON hardware and software technologies integrate with VMware ESX Server. It includes topics such as booting from SAN, CLARiiON array-based software implementation, multipathing, and failover software from VMware.

Basic connectivity

Connecting a VMware ESX Server to the CLARiiON storage system requires LUN masking (Access Logix™) to be enabled on the CLARiiON storage system. CLARiiON assigns storage (Fibre Channel or iSCSI) to the VMware ESX Server and not to individual virtual machines. LUNs presented to the virtual machines are called virtual disks; the underlying CLARiiON LUNs are transparent to the guest operating system. Virtual disks that are assigned to the VMware ESX Server are not automatically assigned to the virtual machines; VirtualCenter or the Management User Interface of the VMware ESX Server assigns virtual disks to the individual virtual machines.

ESX 3.x/ESX 3i allows the user to add virtual disks to a virtual machine without powering the virtual machine down. This functionality is provided in the Add dialog box in VirtualCenter 2.0, which is shown in Figure 2. In ESX 2.5.x, the user must power the virtual machine down before adding virtual disks, as shown in Figure 3.

When connecting a CLARiiON (CX3 or AX4) Fibre Channel storage system to VMware ESX both direct and FC-SW connections are supported. For specific versions of VMware ESX that support direct and FC-SW connect, consult the E-Lab™ Navigator on Powerlink®.

When connecting CLARiiON iSCSI storage to a VMware ESX Server (ESX 3.x), both the software and hardware initiators are supported. The drivers for the software and hardware initiator are installed by

default during the installation of the ESX operating system. The software initiators for network cards and HBA hardware initiators are configured through the VirtualCenter interface. Please see VMware ESX 3.x/ESX 3i documentation for more details.

ESX 3.x/ESX 3i support both Fibre Channel and iSCSI storage. However, VMware and EMC do not support connecting an ESX 3.x/ESX 3i server to CLARiiON Fibre Channel and iSCSI devices simultaneously.

Both 2 Gb/s and 4 Gb/s Fibre Channel connections are supported with the VMware ESX Server when connected to the CLARiiON CX3 UltraScale series storage systems.

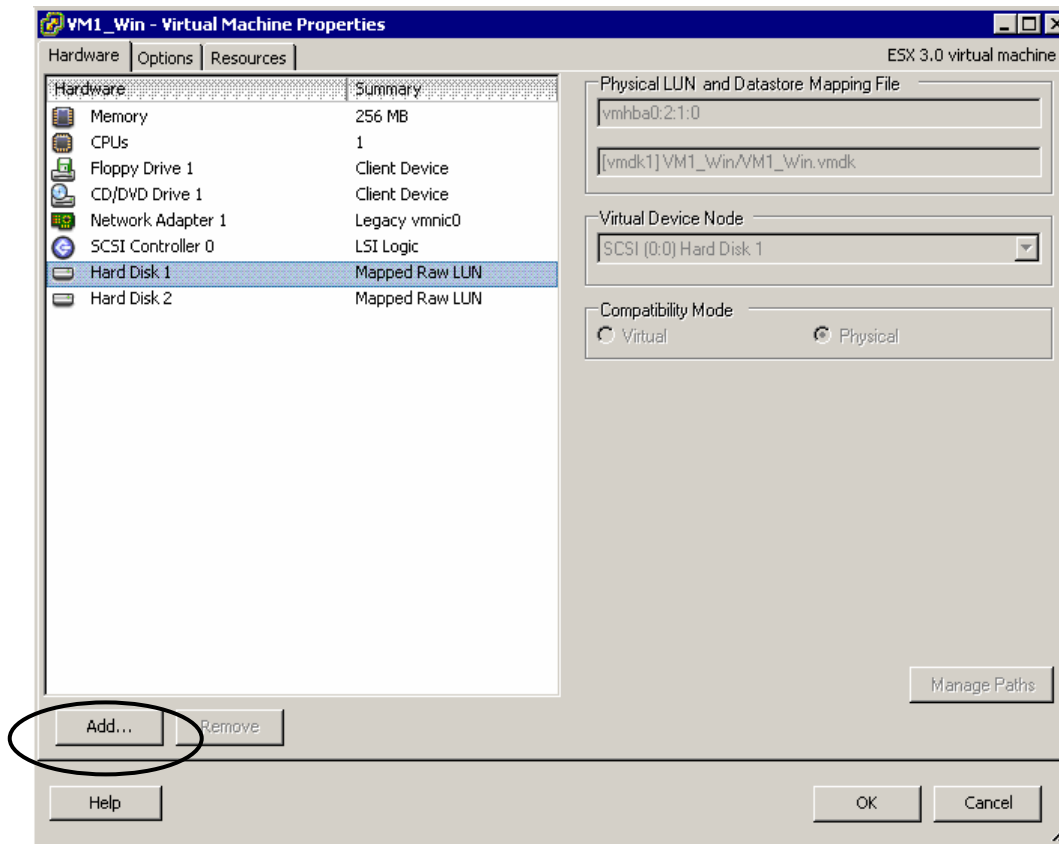


Figure 2. Adding a CLARiiON LUN to a virtual machine (guest operating system) using VirtualCenter 2.0 for the ESX 3.x/3i server

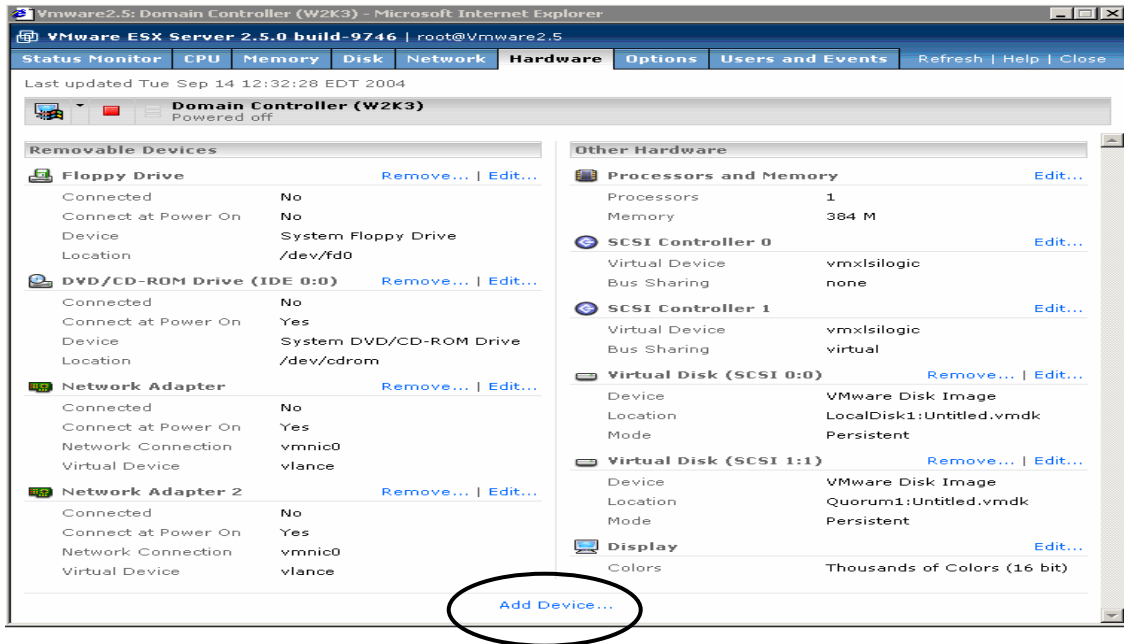


Figure 3. Adding a CLARiiON LUN to a virtual machine (guest operating system) using the Management User Interface (MUI)

Booting from a CLARiiON storage system

This section discusses the procedure for booting the VMware ESX Server and guest operating systems—Windows, Linux, NetWare, and Solaris—from CLARiiON LUNs.

Booting VMware ESX Server from CLARiiON LUNs with ESX 3.x and ESX 2.5.x

In ESX Server 2.5 and 3.x, the service console can boot from a CLARiiON LUN through the HBAs if the CLARiiON is a Fibre Channel storage system with Fibre Channel HBAs. To boot an ESX 2.5.x server from a LUN, the HBAs must be shared between the service console and the virtual machines. In ESX version 3.x, this is not necessary because there is no concept of shared or dedicated HBAs.

There is no boot-from-SAN support for the VMware ESX 3i operating system image.

VMware ESX Server accesses LUNs through the Fibre Channel HBA. If the ESX Server machine has more than one HBA, all its HBAs must be the same model.

To boot through a LUN you must:

- Configure the BIOS settings for the Fibre Channel HBA to select the CLARiiON LUN as the boot device.
- With ESX version 2.5, make sure that the boot LUN is **/dev/sda** and **ID 0**—the lowest-numbered LUN visible to the ESX Server. This is not necessary for ESX 3.x, since it uses the device UUID for LUN identification.
- The internal SCSI device (controller) must be disabled for the CLARiiON LUN to map as **/dev/sda**. Since the CLARiiON storage system consists of active/passive path configuration, the lowest-numbered path to the boot LUN must be the active path.
- For ESX 2.5, install the VMware ESX Server software on the CLARiiON LUN using the **boot-from-SAN** option. For ESX 3.x, when installing the server software select the CLARiiON LUN from which the operating system will boot.

Note: VMware ESX Server version 2.5 or later is supported for booting ESX over the SAN.

Booting guest operating systems on CLARiiON LUNs

Virtual machines can run on CLARiiON LUNs, as well as on internal disks. Virtual machines can boot using both Fibre Channel and iSCSI CLARiiON storage. Booting virtual machines from shared storage is also a requirement for VMware VMotion.

When booting virtual machines from a CLARiiON storage system, the LUNs are first presented to the VMware ESX Server.

LUNs presented to virtual machines are called virtual disks; to the virtual machine it appears that it is accessing a physical disk. These disks are created when the virtual machine is created using VirtualCenter or the Management User Interface (MUI). The operating system can be installed on these virtual disks using a CD or ISO image. When the virtual machine is created, a configuration file with a .vmx extension is also generated. This file contains the location of virtual disks, memory size, and some basic hardware setup information (CD-ROM drive, floppy drive, network connections) for the virtual machine.

Once the virtual machine is up and running, it is highly recommended that VMware Tools be installed on each virtual machine. VMware Tools will optimize the use of the VMware ESX Server resources. VMware Tools also provide a VGA device driver and a heartbeat mechanism for the virtual machine to communicate with the VMkernel.

The procedure for connecting the VMware ESX Server to the EMC CLARiiON storage system is found in the *Host Connectivity Guide for VMware ESX Server 2.x*.

VMware does not support booting ESX Server over iSCSI storage using the software initiator; however, it does support VMs residing on iSCSI LUNs, which is a requirement for VMotion.

Navisphere management

Navisphere Agent (for CX series only) or the Server Utility must be installed on the ESX service console to register the ESX 3.x servers with the CLARiiON storage system. ESX 3i does not have a service console to install or run the host agent or server utility; instead, the CLARiiON storage system is automatically registered when you rescan the EXC 3i server host bus adapters or when the ESX 3i server reboots. Manual registration is not necessary. However, note that the ESX 3i server that appears in Navisphere Manager will not have an OS name, OS revision information, or device mapping information.

Navisphere CLI and array initialization software for the CX and AX4 series storage systems can run on the ESX Server console, as well as the individual virtual machines.

Navisphere CLI and Navisphere Host Agent come bundled as an .rpm package and are supported with ESX version 2.5 and later. The VMware Navisphere Agent and CLI package are the same as the Linux kit that provides device mappings of the **vmhba** device names and allows path registration with the storage system. It does not provide the device mappings information from the virtual machines since the agent is installed on the ESX service console.

For Navisphere Agent/CLI to work with ESX 3.x when connected to a CLARiiON storage system, the ports for agent and CLI need to be opened. This can be done by executing the following command on the ESX service console:

```
# esxcfg-firewall -o --openPort <port,tcp|udp,in|out,name>
```

For example:

```
esxcfg-firewall -o 6389,tcp,in,naviagent
```

For detailed information on which ports to open, see the *CLARiiON Server Support Products for Linux and VMware ESX Server Installation Guide* available on Powerlink[®], EMC's password-protected extranet for customers and partners.

If Navisphere Agent is installed before the preceding command is executed, restart the agent so that it communicates with the storage system and sends updated information.

When Navisphere CLI is installed on virtual machines, some commands (for example, **lunmapinfo**) that require Navisphere Agent must be directed to the ESX service console and not to the virtual machines. Check the Navisphere Agent/CLI release notes on Linux and VMware for more details. Figure 4 shows the device mapping information that is listed when the **lunmapinfo** command is issued from Navisphere CLI on a Windows virtual machine. This command is directed to the agent residing on the ESX service console.

```
C:\Program Files\EMC\Navisphere CLI>navicli -h 10.14.17.73 lunmapinfo
Logical Drives:      vmhba0:1:0, /boot, /
Physical Device:     sda

Logical Drives:      vmhba1:0:0
Physical Device:     sdc

Logical Drives:      vmhba1:0:1
Physical Device:     sdd

Logical Drives:      vmhba1:0:2
Physical Device:     sde

No storage systems were found. Certain fields could not be displayed.
```

Figure 4. Executing the lunmapinfo command issued on a Windows virtual machine and sent to the Navisphere Agent on the ESX service console

You can use the **symm inq** utility to get device mapping information from the virtual machine to the CLARiiON LUN level. The virtual disks assigned to the virtual machine must be configured as raw device mappings for this to work correctly. Figure 5 shows output for the **inq -clar_wwn** command.

```
C:\>
C:\>inq -clar_wwn
Inquiry utility, Version U7.3-623 (Rev 0.0)      (SIL Version U6.0.0.0 (Edit Level 623)
Copyright (C) by EMC Corporation, all rights reserved.
For help type inq -h.
...
-----
CLARiiON Device      Array Serial #      SP IP Address      LUN      WWN (all 32 hex
digits required)
-----
\\.\PHYSICALDRIVE0  WRE00022100934     A  10.14.17.78      0x0006    60060160b7a5080
03ec2197bffc2d911
\\.\PHYSICALDRIVE1  WRE00022100934     B  10.14.17.79      0x000a    60060160b7a5080
07df535167eadd911
\\.\PHYSICALDRIVE2  WRE00022100934     A  10.14.17.78      0x0003    60060160b7a5080
03bc2197bffc2d911
```

Figure 5. Output for the inq -clar_wwn command that provides device mapping information from the virtual machine level to the CLARiiON LUN level

VMware native failover with CLARiiON

The CLARiiON storage system supports VMware ESX Server’s built-in failover mechanism; this mechanism provides failover, but not active I/O load balancing. The CLARiiON storage system also supports the nondisruptive upgrade (NDU) operation while the VMware ESX Server is online. The E-Lab Navigator has a list of ESX Server versions for which NDU operations are supported.

The native failover software provides a listing of the paths—whether active or passive—from the VMware ESX Server to the CLARiiON storage system. The **esxcfg-mpath** command in ESX 3.x provides details on all devices (Fibre Channel, iSCSI, and local) and the number of paths attached to that device. With VMware ESX 3i, you can use VirtualCenter or the remote CLI package to see the number of paths to a given LUN. If you are using an ESX version with a service console, type:

```
# esxcfg-mpath -l
```

```
[root@esx3 /]# esxcfg-mpath -l
Disk vmhba0:0:0 /dev/sda (2048MB) has 2 paths and policy of Most Recently Used
FC 1:8.0 10000000c92930e7<->500601601060174b vmhba0:0:0 On active preferred
FC 1:8.0 10000000c92930e7<->500601681060174b vmhba0:1:0 Standby

Disk vmhba0:0:1 /dev/sdb (5120MB) has 2 paths and policy of Most Recently Used
FC 1:8.0 10000000c92930e7<->500601601060174b vmhba0:0:1 Standby preferred
FC 1:8.0 10000000c92930e7<->500601681060174b vmhba0:1:1 On active

Disk vmhba0:0:3 /dev/sdc (5120MB) has 2 paths and policy of Most Recently Used
FC 1:8.0 10000000c92930e7<->500601601060174b vmhba0:0:3 Standby preferred
FC 1:8.0 10000000c92930e7<->500601681060174b vmhba0:1:3 On active

Disk vmhba0:0:5 /dev/sdd (1024MB) has 2 paths and policy of Most Recently Used
FC 1:8.0 10000000c92930e7<->500601601060174b vmhba0:0:5 On active preferred
FC 1:8.0 10000000c92930e7<->500601681060174b vmhba0:1:5 Standby

Disk vmhba0:2:0 (0MB) has 2 paths and policy of Most Recently Used
FC 1:8.0 10000000c92930e7<->5006016010208b46 vmhba0:2:0 On active preferred
FC 1:8.0 10000000c92930e7<->5006016810208b46 vmhba0:3:0 On

Disk vmhba2:0:0 /dev/sdq (17366MB) has 1 paths and policy of Fixed
Local 5:6.0 vmhba2:0:0 On active preferred

Processor Device vmhba2:6:0 (0MB) has 1 paths and policy of Fixed
Local 5:6.0 vmhba2:6:0 On active preferred
```

Figure 6. VMware ESX Server 3.0 path information for Fibre Channel devices

Figure 6 shows the seven devices attached to the CLARiiON storage system. The **vmhba0:x:x** devices are Fibre Channel devices. All Fibre Channel devices have paths to both SP A and SP B. The **active** mode for each path shows the path the ESX Server uses to access the disk. The **preferred** mode, although it is displayed, is not honored (it is ignored) since the policy is set to **Most Recently Used** (MRU). Device **vmhba2:0:0** is the internal boot device and has a single path.

Figure 7 shows the three devices attached to the CLARiiON storage system. The **vmhba40:0:x** devices are iSCSI devices. All iSCSI devices have paths going to both SP A and SP B. If using VMware NIC teaming, the NICs must be on the same subnet and use the same IP address for failover to work between multiple iSCSI NICs (uplink adapters). As a best practice, create dedicated virtual switches for iSCSI traffic.

With ESX 3.5/3i, VMware supports the configuration of two virtual switches on separate subnets that go to different network switches if you use the iSCSI software initiator that is built in to ESX Server.

```

Disk vmhba2:0:0 /dev/sdq (17366MB) has 1 paths and policy of Fixed
Local 5:6.0 vmhba2:0:0 On active preferred

Processor Device vmhba2:6:0 (0MB) has 1 paths and policy of Fixed
Local 5:6.0 vmhba2:6:0 On active preferred

Disk vmhba40:0:0 /dev/sda (10240MB) has 2 paths and policy of Most Recently Used
iScsi sw iqn.1998-01.com.vmware:esx3-1eedf183<->iqn.1992-04.com.emc:cx.apm00042102262.a0 vmhba40:0:0 Standby preferred
iScsi sw iqn.1998-01.com.vmware:esx3-1eedf183<->iqn.1992-04.com.emc:cx.apm00042102262.b0 vmhba40:1:0 On active

Disk vmhba40:0:1 /dev/sdb (10240MB) has 2 paths and policy of Most Recently Used
iScsi sw iqn.1998-01.com.vmware:esx3-1eedf183<->iqn.1992-04.com.emc:cx.apm00042102262.a0 vmhba40:0:1 Standby preferred
iScsi sw iqn.1998-01.com.vmware:esx3-1eedf183<->iqn.1992-04.com.emc:cx.apm00042102262.b0 vmhba40:1:1 On active

Disk vmhba40:0:2 /dev/sdc (10240MB) has 2 paths and policy of Most Recently Used
iScsi sw iqn.1998-01.com.vmware:esx3-1eedf183<->iqn.1992-04.com.emc:cx.apm00042102262.a0 vmhba40:0:2 Standby preferred
iScsi sw iqn.1998-01.com.vmware:esx3-1eedf183<->iqn.1992-04.com.emc:cx.apm00042102262.b0 vmhba40:1:2 On active

Disk vmhba40:0:3 /dev/sdd (14336MB) has 2 paths and policy of Most Recently Used
iScsi sw iqn.1998-01.com.vmware:esx3-1eedf183<->iqn.1992-04.com.emc:cx.apm00042102262.a0 vmhba40:0:3 Standby preferred
iScsi sw iqn.1998-01.com.vmware:esx3-1eedf183<->iqn.1992-04.com.emc:cx.apm00042102262.b0 vmhba40:1:3 On active

```

Figure 7. VMware ESX Server 3.0 path information for iSCSI devices

The `vmkmultipath` command, when issued on an ESX 2.x server, provides details about the devices and the number of paths attached to each device. At the service console of the VMware ESX Server, type:

```
# vmkmultipath -q
```

```

[root@Vmware2 root]# vmkmultipath -q
Disk and multipath information follows:

Disk vmhba0:0:0 (10,236 MB) has 2 paths. Policy is mru.
  vmhba0:0:0      on (active, preferred)
  vmhba1:0:0      on

Disk vmhba0:1:1 (40,954 MB) has 2 paths. Policy is mru.
  vmhba0:1:1     on (active, preferred)
  vmhba1:1:1     on

Disk vmhba2:0:0 (34,726 MB) has only 1 path.

```

Figure 8. VMware ESX Server 2.x path information through the native failover software

The most recently used MRU policy is the default policy for active/passive storage devices in ESX 2.x and 3.0. The policy for the path should be set to MRU for CLARiiON storage systems to avoid path thrashing. When using the MRU policy, there is no concept of preferred path; in this case, the preferred path can be disregarded. The MRU policy uses the most recent path to the disk until this path becomes unavailable. As a result, ESX Server does not automatically revert to the original path until a manual restore is executed.

If you connect two ESX Servers with path one from HBA1 to SPA, and path two from HBA0 to SPB, a single LUN configured as a VMFS volume can be accessed by multiple ESX Servers; in this example a LUN can be accessed by both ESX Servers.

If the HBA1-SPA path on ESX1 fails, it issues a trespass command to the array, and SPB takes ownership of the LUN. If the path from HBA1-SPB on ESX2 then fails, the LUN will trespass back and forth between the SPs, which could result in performance degradation.

When the CLARiiON LUN policy is set to MRU and an ESX Server with two HBAs is configured so that each HBA has a path to both storage processors, ESX Server accesses all LUNs through one HBA and does not use the second HBA. You can edit the path configuration settings so the other HBA is the active path for some LUNs; however, this configuration is not persistent across reboots. After a reboot, the LUNs will be on a single HBA. The advantage of this configuration is it prevents unnecessary trespasses of LUNs in the case of failure.

The failover time can be adjusted at the HBA, ESX, and virtual machine levels. The *VMware ESX Server 3.0.1 and VirtualCenter 2.0.1 SAN Configuration Guide* at the following link provides recommendations for setting the failover time at the HBA and virtual machine level:

http://www.vmware.com/pdf/vi3_301_201_san_cfg.pdf

The ESX Server periodically evaluates the state of each path. The default evaluation period is 300 seconds. This can be changed by modifying the `/proc/vmware/config/disk/PathEvalTime vmkernel config` value. This can also be done thru the MUI for ESX 2.x by going to **Advanced Setting** and changing the **Disk.PathEvalTime** parameter. The evaluation period can be set to any value between 30 and 1500 seconds. Note that reducing the **PathEvalTime** causes path evaluation to run more frequently. This puts a slightly higher CPU load on the system. Reducing this value (to 90 for example) will help improve failover time (keeping in mind the preceding caveat).

PowerPath[®] multipathing and failover software is not supported on the VMware ESX Server or on the virtual machines.

LUN partitioning

LUNs presented to the ESX Server are ultimately presented to the virtual machines. Any storage device presented to any virtual machine is represented as a virtual disk. To the virtual machine, the virtual disk appears to be a physical disk. A virtual machine can have multiple virtual disks of different/multiple virtual disk types. A CLARiiON LUN presented to the ESX Server can be partitioned using one of the three methods:

- Raw disks
- VMFS volumes
- Raw device mappings

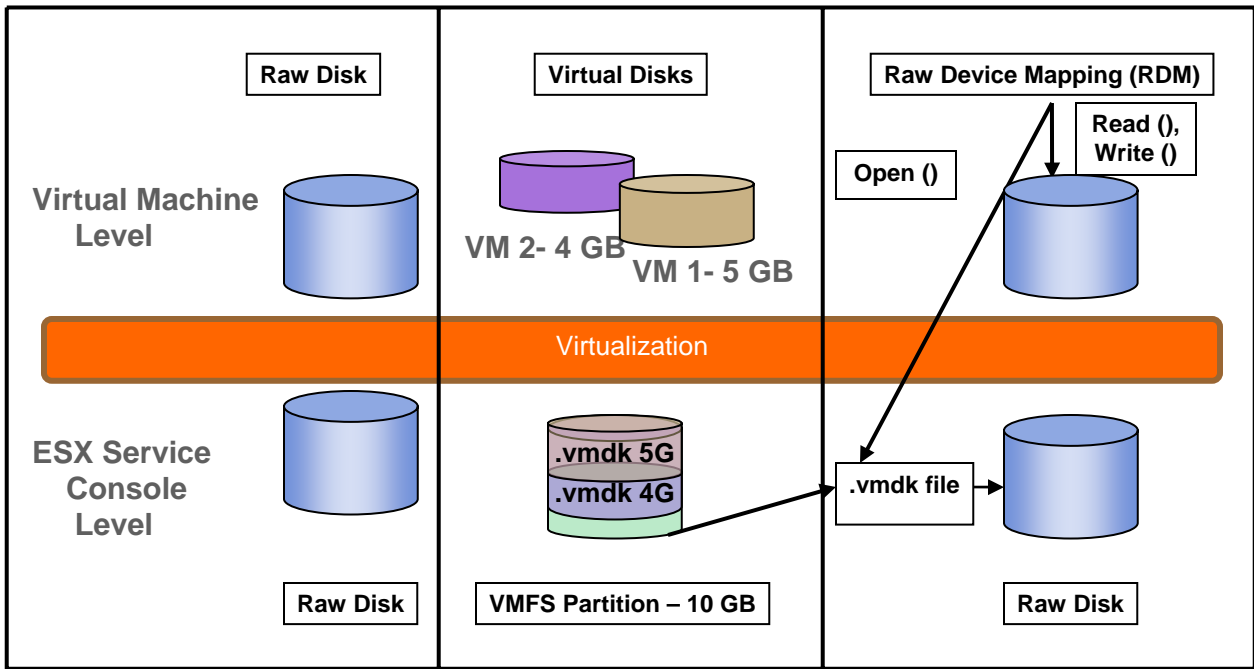


Figure 9. Partitioning a CLARiiON LUN

Raw disks

For raw disks, an entire CLARiiON LUN is presented to a single virtual machine without being partitioned at the ESX service console level. When a virtual machine is configured to use a raw disk, VMware directly accesses the local disk/partition as a raw device. Raw devices are available in ESX 2.5, but it is recommended that the LUN be configured as a raw device mapping device (RDM) instead. RDMs are very similar to raw disks except that RDMs are compatible with VMotion.

VMFS volumes

When a CLARiiON LUN is configured as a VMFS volume, this volume can be partitioned and presented to a number of virtual machines. For example, if you present a 10 GB CLARiiON LUN to your ESX Server, a VMFS file system can be created on that LUN. New VMFS3 volumes created with 3.5/3i must be 1200 MB or larger. For previous versions of ESX Server, the VMFS3 requirement was 600 MB. The user has the option of presenting this entire VMFS volume to an individual virtual machine or presenting portions of this volume to a number of virtual machines. In Figure 9, the VMFS volume is used to create two virtual disks (.vmdk files)—one is 5 GB and the other is 4 GB. Each of these virtual disks is presented to a different virtual machine. It is also possible to create a virtual disk on an entire VMFS volume and assign this virtual disk to a single virtual machine.

In ESX 3.x/ESX 3i, the swap files, NVRAM files, and configuration (.vmx) files for a virtual machine reside on a VMFS-3 volume. On ESX 2.0, these files reside on an ext3 file system on the service console.

ESX 2.x supports undoable disk mode that allows you to keep or discard changes to a virtual disk using snapshot technology. Snapshot technology on the ESX Server is supported for VMFS-3 and VMFS-2 volumes. In ESX 3.x/ESX 3i, the snapshot technology allows all virtual disks within a VM configured as VMFS-3 volumes to be snapshot together along with VM memory, processor, and other states using the consolidated backup solution.

Raw device mapping

VMware ESX 2.5 introduced a new technology called raw device mapping (RDM). This technology has a **SCSI pass-through** mode that allows virtual machines to pass SCSI commands directly to the physical hardware. Utilities like `admsnap` and `admhost`, when installed on virtual machines, can directly access the virtual disk when the virtual disk is in physical compatibility mode. In virtual compatibility mode, a raw device mapping volume looks like a virtual disk in a VHMS volume. This streamlines the development process by providing advance file locking data protection and VMware snapshots.

Using a raw CLARiiON LUN, a user can create a raw device mapping volume by creating a mapping file on a VMFS volume. This mapping file, which contains a `.vmdk` extension, points to the raw device, as shown in Figure 9. The mapping file is created when the raw device is ready to be assigned to a virtual machine. The entire CLARiiON LUN is presented to an individual virtual machine. The virtual machine opens the mapping file information from the VMFS volume and can directly access the raw device mappings volume for reading and writing.

For more information on configuring VMFS and raw device mapping volumes, refer to the *VMware ESX Server 3.0.1 and VirtualCenter 2.0.1 Server Configuration Guide*.

LUN layout recommendations

This section discusses some of the best practices for optimal capacity when designing and implementing the LUN layout for VMware ESX Servers connected to CLARiiON storage systems.

OS images and application data images of virtual machines can reside on CLARiiON LUNs. Since VMFS is a clustered file system, when LUNs are configured as VMFS volumes, many ESX Servers can share different virtual disks on the same LUN (VMFS) volume. Hence, the number of virtual machines images installed on that particular LUN, and the workload on those virtual machines and the ESX Servers that are accessing the LUN, will dictate the number of spindles that need to be assigned to that particular LUN (VMFS volume). Use of striped metaLUNs to distribute the load across different RAID groups when booting a number of OS images on a given LUN (VMFS volume) should also be considered. When installing a guest operating on a CLARiiON LUN, configure the LUN to use RAID 1/0 or RAID 5. Choose RAID 1/0 instead of RAID 5 to reduce rebuild times in the event of disk failures.

For I/O-intensive application data volumes, it is best to separate OS images from application data. In this case, EMC recommends that you use RDM volumes; since they are dedicated to only one virtual machine (that is, the entire LUN is presented to the virtual machine), replication and backup of applications are similar to that of a physical server. RDMs also allow users to perform online expansion of the underlying LUN using CLARiiON metaLUNs and file systems (OS dependent) at the virtual machine level. A mix of VMFS and raw device mapping volumes are allowed on an ESX Server. However, note that ESX Server 2.x has a limit of 128 SCSI disks. This limit includes both local devices and SAN LUNs. With ESX 3.x/ESX 3i, the limit is increased to 256 SCSI disks.

Also, because of the use of VMware redo logs, EMC recommends that you use separate disks for test and development applications, virtual machine templates (because of sequential I/O intensity), and production LUNs.

ESX 3.x/ESX 3i provides performance and reliability improvements where a single swap file is available for each virtual machine. These swap files and NVRAM files for a given VM can reside on a VMFS-3 volume. Ensure that the VMFS-3 volume has enough space to accommodate the swap files. With ESX 2.x, a single swap file is used for all virtual machines.

Application data disks residing on virtual machines should be aligned with the CLARiiON disk stripe, just as they are on physical servers. When aligning RDMs, align them at the virtual machine level. For Windows virtual machines, use `diskpart` to perform the alignment.

For VMFS-2 volumes, the alignment can be done at the ESX Server level and virtual machine level using `fdisk`. VMFS-3 volumes are already aligned to 64KB during creation; however, for Intel-based systems the virtual disks from a VMFS-3 volume need to be aligned at the virtual machine level. OS disks do not need alignment, since it is almost impossible to align them at the virtual machine level. For best

performance, use VI Client or Virtual Infrastructure Web Access to set up your VMFS3 partitions instead of using the ESX Server 3.x service console. Using VI Client or VI Web Access ensures that the starting sectors of partitions are 64K-aligned, which improves storage performance.

Using CLARiiON virtual LUN technology with ESX 3.x/ESX 3i and 2.x

CLARiiON virtual LUN technology provides an additional layer of abstraction between the host and back-end disks. This technology consists of two features: CLARiiON metaLUNs and the CLARiiON LUN migration that is available on the CLARiiON storage system. This section explains how CLARiiON metaLUNs and CLARiiON LUN migration work with VMware ESX Server.

CLARiiON metaLUNs are a collection of individual LUNs. They are presented to a host or application as a single storage entity. MetaLUNs allow users to expand existing volumes on the fly using the stripe or concatenation method.

CLARiiON LUN migration allows users to change performance and other characteristics of existing LUNs without disrupting host applications. It moves data—with the change characteristics that the user selects—from a source LUN to a destination LUN of the same or larger size. LUN migration can also be used on a metaLUN.

Expanding and migrating LUNs used as raw device mapping

A LUN presented to a VMware ESX Server (ESX 3.x or ESX 2.x) can be expanded with metaLUNs using the striping or concatenation method. After the CLARiiON completes the expansion, rescan the HBAs using either VirtualCenter for ESX 3.x/ESX 3i or the Management User Interface for 2.x to ensure the ESX service console and VMkernel see the additional space. Since the LUN is presented to the virtual machine, expansion must take place at the virtual machine level. Use the native tools available on the virtual machine to perform the file system expansion at the virtual machine level.

LUN migration can be conducted on VMFS and RDM volumes and is transparent to the guest OS. For RDM volumes, if the destination LUN is larger than the source LUN after the migration process completes, use the procedure previously outlined to rescan the HBAs, and then expand the disk at the virtual machine level. Note that RDM volumes must use the **physical compatibility** mode for expansion when using the CLARiiON metaLUN technology.

Expanding and migrating LUNs used as VMFS volumes

VMware ESX supports the volume management functions where VMFS volumes can be concatenated together as a single volume. This procedure is also called VMFS spanning in ESX 2.x and is done at the ESX Server level. ESX 3.x/ESX 3i also provides volume management functionality for VMFS volumes through a process called Adding Extents within VirtualCenter. The difference between VMFS-2 spanning within ESX 2.x and volume management using VMFS-3 within ESX 3.x/ESX 3i is:

- Unlike VMFS-2, a VMFS-3 volume can be extended while in use.
- With VMFS-2, loss of any partition renders the whole volume inaccessible. For VMFS-3, except for the head partition, loss of a partition renders only the data on that partition inaccessible.

The first option to expand a CLARiiON LUN configured as a VMFS-2 or VMFS-3 volume is to add a new CLARiiON LUN and concatenate the two LUNs using the VMFS spanning process in ESX 2.x, or by adding extents in ESX 3.x/ESX 3i. To expand the virtual disk presented to the virtual machine, use the **vmkfstools** utility available on ESX Server.

The other option is to expand a VMFS-3 volume using CLARiiON metaLUNs and span or add an extent using the additional space with the original VMFS volume available before expansion. Refer to Knowledgebase case emc128545 for additional details.

As a best practice, always have a backup copy in place before performing any of these procedures.

Using CLARiiON replication software with ESX 3.x/ESX 3i and 2.5.x

CLARiiON replication software products including SnapView, MirrorView, and SAN Copy are supported with the VMware ESX Server using both VMFS and RDM volumes. The OS image and the application/data can be replicated using CLARiiON replication software. The following considerations apply to iSCSI and FC storage systems. Please note that remote replication software (MirrorView and SAN Copy) is supported on CLARiiON iSCSI storage systems:

CLARiiON replication considerations with VMware ESX Server

Please note that:

- Use of RDM volumes for replication is not supported when an ESX 2.5.x server is booted from a SAN LUN. In other words, when the Fibre Channel HBAs are shared between the service console and the virtual machines, RDM cannot be configured on an ESX 2.5.x server. There is no such restriction with ESX 3.x/ESX 3i.
- `admsnap` and `admhost` must be installed on the virtual machines and not the ESX Server service console.
- With ESX 2.5.x, ensure that a CLARiiON snapshot, clone, or mirror is not in a **device not ready** state (snapshot not activated, session not started, clone not fractured, or secondary mirror not promoted) when it is assigned to an ESX Server. The ESX service console does not create a device file for a LUN in this state. This restriction only applies at the ESX service console level and not the virtual machine level. Users can execute activate and deactivate operations at the virtual machine level using `admsnap` and `admhost` utilities after the ESX Server sees the device the first time. For an AX100 system running Navisphere Express and connected to an ESX 2.5.x server, the replica must be presented to a secondary physical server (and not an ESX Server) since the replica cannot be activated through the Navisphere Express GUI. There is no such restriction with VMware ESX 3.x/ESX 3i; it sees the snapshot, clone, or mirror in spite of the device's not ready condition.

CLARiiON replication software considerations when using VMFS volumes

Please note that:

- The virtual disks in a VMFS-2 volume must be in **persistent** mode during the replication process.
- When using VMFS-2 volumes, do not present two copies of the same VMFS volume to the same ESX Server. For example, an ESX Server participating in VMotion with the primary ESX Server has access to the original source LUNs. Hence, this ESX Server should not be a target when a replica is presented.

With VMFS-3 volumes on ESX 3.x/ESX 3i, the replica can be presented to the same ESX Server or a standby ESX Server. This can be done using VirtualCenter by enabling the **LVM.EnableResignature** parameter in the ESX Server. After a rescan, the replica is resignatured and can be assigned to a virtual machine.

See the “Automatic Volume Resignaturing” section in the *VMware ESX Server 3.0.1 and VirtualCenter 2.0.1 SAN Configuration Guide* on www.vmware.com.

- When replicating an entire VMFS (VMFS-2 and VMFS-3) volume that contains a number of virtual disks on a single CLARiiON LUN, the granularity of replication is the entire LUN with all its virtual disks.
- CLARiiON VSS Provider is not supported on VMFS volumes.
- When making copies of VMFS volumes that span multiple CLARiiON LUNs, use the array-based consistency technology.

-
- ESX Server-based VM snapshot copies should not be used in conjunction with CLARiiON replication software copies on the same VMFS volume for ESX 2.5.x since VM snapshot copies require the virtual disk to be in nonpersistent mode.
 - Most of the **admsnap** and **admhost** commands when issued on VMFS volumes will fail since VMFS volumes do not support SCSI pass-through commands to communicate with the CLARiiON storage system. Use Navisphere Manager or Navisphere CLI instead. The only commands that will work are **admsnap flush** and **admhost flush**.
 - VMFS volumes are not supported when replicating application data images from a physical (native) server to an ESX Server.
 - For ESX 2.5.x, ensure that a CLARiiON snapshot, clone, or mirror of a VMFS volume is not in a **device not ready** state (snapshot not activated, session not started, clone not fractured, or secondary mirror not promoted) when it is assigned to an ESX Server.

Since VMFS-3 volumes may contain VM configuration files, swap files, and NVRAM files, these files can be replicated using CLARiiON replication software.

CLARiiON replication software considerations when using RDM volumes

Please note that:

- You should configure the LUNs to use the **physical compatibility mode** option when replicating RDM volumes using CLARiiON replication software. Otherwise, **admsnap** and **admhost** will not work on the virtual machine.
- VMware ESX Servers do not write a *signature* on RDM volumes. Hence, replicas can be presented back to the same VMware ESX Server for use. The copies cannot be used on the source virtual machines unless the guest OS supports this feature. However, they can be assigned as raw devices to another virtual machine.
- EMC Replication Manager is supported with RDM volumes.
- RDM volumes are supported on the ESX Server when replicating application data images from a physical (native) server to an ESX Server.

When replicating OS disks while the virtual machine is powered on, the replica or copy will be in a crash-consistent state since there is no mechanism available to quiesce a boot image. For application data disks, native tools available with the application can be deployed to quiesce the application to get a consistent replica. The array-based consistency technology can be used when applications span multiple LUNs or for write-order dependent applications such as databases. For automation, scripts may need to be developed to integrate CLARiiON replication software with the different applications running on the virtual machines.

CLARiiON and VMotion

Migration with VMotion allows you to move a virtual machine between two ESX Servers while the virtual machine is powered on and performing transactions. When a migration with VMotion is performed, the operations of the virtual machine can continue uninterrupted. The virtual machine must reside on a SAN LUN accessible to both source and destination hosts. VMotion only moves the virtual machine configuration file and memory contents to the alternate host. Any disks assigned to the VM are moved by transferring their ownership.

The conditions for VMotion to work are:

- The VMware VirtualCenter server and client must be installed on a Windows system.
- A Gigabit Ethernet connection is required between the two ESX Server hosts participating in VMotion migration.
- The guest operating system must boot from a CLARiiON LUN. The virtual machine boot LUN and its associated data disks must be shared between the source and destination hosts.

-
- VMware VMotion is supported with both Fibre Channel and iSCSI CLARiiON storage systems.
 - VMFS and RDM volumes are supported with VMotion.
 - Both hosts must have identical CPUs (both CPUs are P4, for instance). The CPUs must also be manufactured by the same vendor.

Migration with VMotion cannot occur when virtual machines are in a raw, clustered, or nonpersistent mode. Figure 10 shows two ESX Servers with three NICs each. The recommended number of NICs is three; two is the minimum requirement. A crossover or gigabit LAN connection should exist between the two gigabit NICs on the two VMware ESX Servers. For ESX 2.x, a virtual switch must be created with the same network label on both servers for the two network cards to communicate. In ESX 3.x/ESX 3i, a VMotion VMkernel port group on a vSwitch must be created on both ESX Servers for VMotion migration.

VMotion with VMFS volumes

In order to perform VMotion with VMFS volumes, the following prerequisites must be met:

- All VMFS volumes assigned to the virtual machine that is to be migrated must be shared by both ESX Servers.
- VMFS volumes must be in public access mode.

The VMware ESX Server identifies VMFS-2 volumes by their label information. For VMFS-3 volumes, they are identified by the Host LUN number. Hence, as a best practice for VMotion with VMFS volumes create a single storage group for all ESX Servers in a farm or cluster since LUNs may be assigned different Host LUN numbers if separate storage groups are created. For additional information, see EMC Knowledgebase cases emc151686 and emc153719.

VMotion with RDM volumes

In order to perform VMotion with RDM volumes, the following prerequisites must be met:

- Both the RDM LUN and the VMFS volume that contains the mapping file must be shared by both ESX Servers.
- VMFS volumes must be in public access mode. This mode allows the volume to be accessed by multiple ESX Servers.
- The RDM LUN must have the same host LUN number on the source and destination ESX Servers. This can be set when adding LUNs to the storage group in Navisphere Manager.
- RDM volumes in both physical and virtual compatibility mode are supported with VMotion.

When using RDMs volumes, the recommendation is to create a single storage group since LUNs may be assigned different Host LUN numbers if separate storage groups are created.

Initially, swap files for virtual machines had to be stored on shared storage for VMotion. VMware Infrastructure 3 (ESX Server 3.5 hosts and VirtualCenter 2.5 or later) now allows swap files to be stored on local storage during VMotion migrations for virtual machines. For virtual machines with swap files on local storage, when local storage is the destination during a VMotion migration or a failover, the virtual machine swap file is re-created. The creation time for the virtual machine swap file depends on either local disk I/O, or on how many concurrent virtual machines are starting due to an ESX Server host failover with VMware HA.

If you have an AX4 system running Navisphere Express, you can use Navisphere CLI to implement VMotion with ESX 3.x/ESX 3i servers, and ensure that the Host LUN number is consistent across all ESX Servers using the storagegroup command. Your other option is to set the LVM.Disallowsnapshot and LVM.EnableResignature parameter to 0 on all the ESX Servers participating in VMotion/DRS/VMware HA.

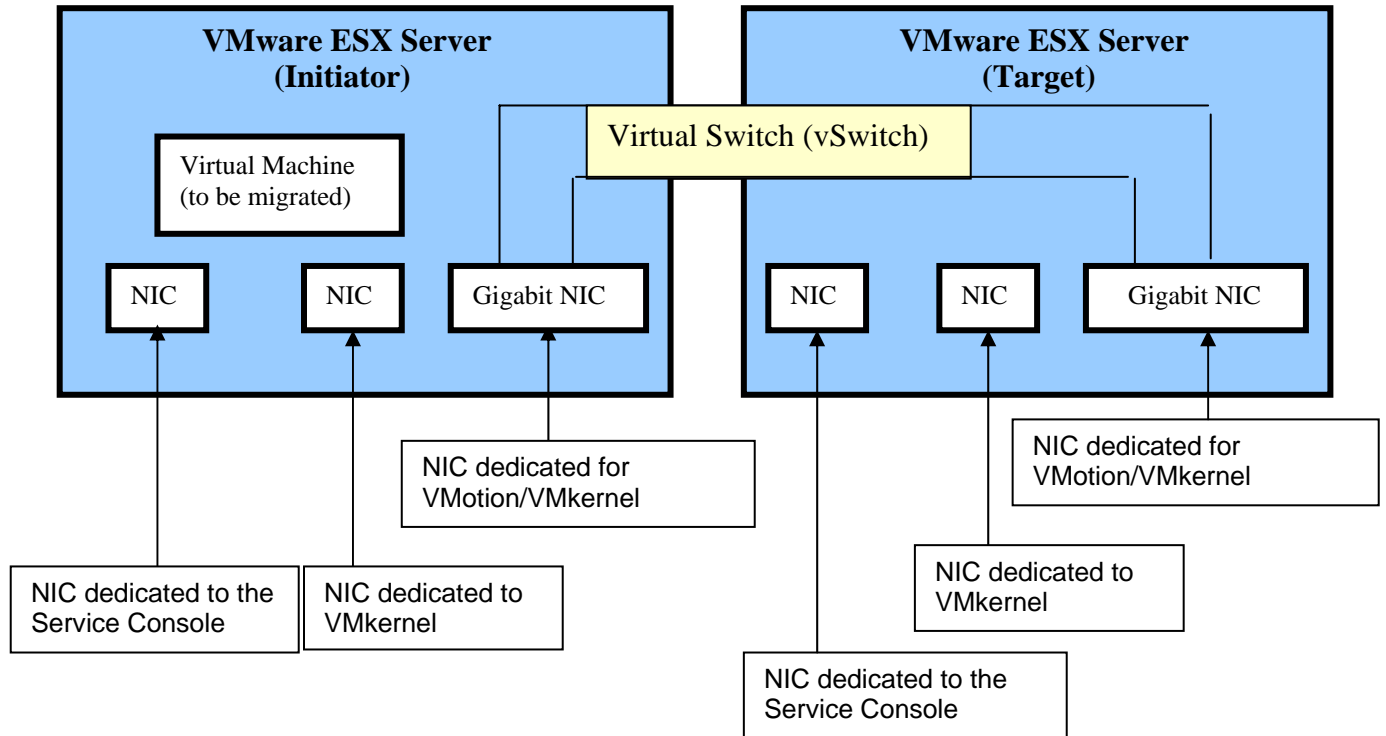


Figure 10. Two VMware ESX Servers ready for VMotion migration

Using the VirtualCenter console, you can initiate the migration process by right-clicking the virtual machine for the initiator host, as shown in Figure 11. After the migration request has been initiated, a wizard opens requesting initiator and target information. VMotion migration takes place through the gigabit network connection setup between the two ESX Servers. After the VMotion migration process completes, the virtual machine is automatically resumed on the target VMware ESX Server with the same name and characteristics as the initiator virtual machine.

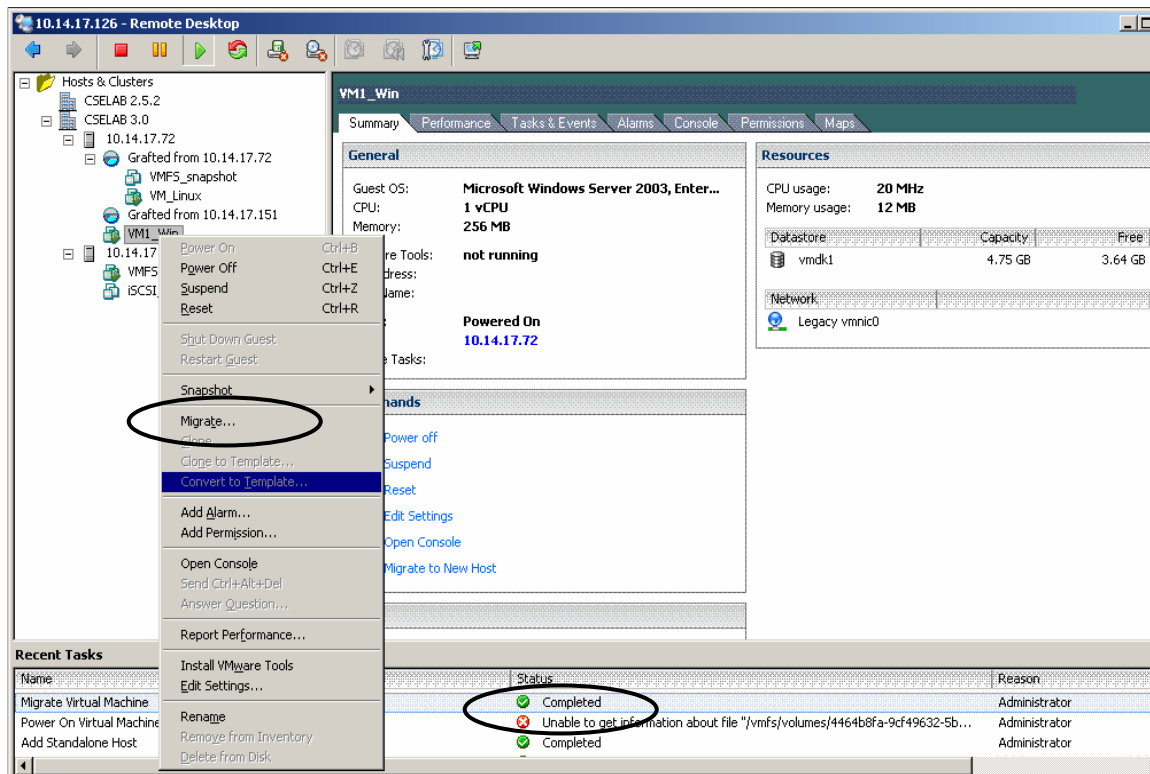


Figure 11. VirtualCenter 2.0 management screen showing how VMotion migration is initiated and completed

CLARiiON with VMware Distributed Resource Scheduling and High Availability

Both VMware Distributed Resource Scheduling (DRS) and VMware High Availability (HA), when used with VMotion technology, provide load balancing and automatic failover for virtual machines with ESX 3.x/ESX 3i. To use VMware DRS and HA, a cluster definition must be created using VirtualCenter 2.0. The ESX hosts in a cluster share resources including CPU, memory, and disks. All virtual machines and their configuration files on ESX Servers in a cluster must reside on CLARiiON storage, so that you can power on the virtual machines from any host in the cluster. Furthermore, the hosts must be configured to have access to the same virtual machine network so VMware HA can monitor heartbeats between hosts on the console network for failure detection.

The conditions for VMware DRS and HA to work are as follows:

- The guest operating system must boot from a CLARiiON LUN. The virtual machine boot LUN, its associated data disks, and configuration files must be shared among all ESX Servers in a cluster.
- VMware HA and DRS are supported with both Fibre Channel and iSCSI CLARiiON storage systems.
- Both VMFS and RDM volumes are supported and are configured exactly as they would be for VMotion. See the “CLARiiON and VMotion” section on page 23.
- DRS is based on the VMotion technology, therefore, ESX Servers configured for a DRS cluster must pass the CPU type check (identical CPUs, same manufacturer). VMware HA does not depend on CPU type check.

CLARiiON and virtual machine clustering

Clustering refers to providing services through a group of servers to achieve high availability and/or scalability. Clustering with the VMware ESX Server is supported at the virtual machine level. Refer to the E-Lab Navigator for the cluster software products that are supported on virtual machines. To implement clustering at the virtual machine level, the virtual machines must boot from local disks and not CLARiiON disks. There are two types of cluster configuration at the virtual machine level:

- In-the-box cluster
- Out-of-the-box cluster
 - Virtual to virtual
 - Virtual to physical

In-the-box cluster

This section outlines clustering virtual machines on a CLARiiON storage system between virtual machines on the same VMware ESX Server. This provides simple clustering to protect against software crashes or administrative errors. The cluster consists of multiple virtual machines on a single ESX Server.

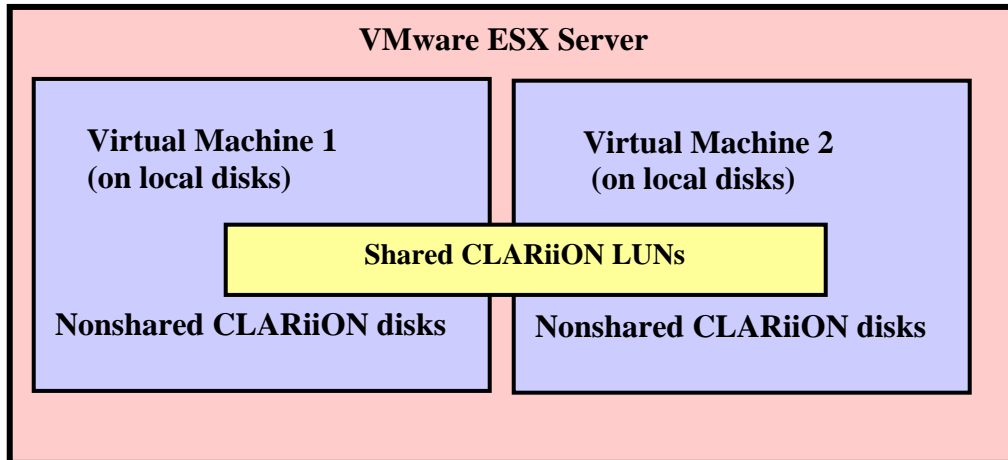


Figure 12. In-the-box cluster configuration

In Figure 12, a single VMware ESX Server consists of two virtual machines. The quorum device and/or clustered applications are shared between the two virtual machines. Each virtual machine can have virtual disks that are local to the virtual machine, that is, virtual disks not shared between virtual machines. The device containing the clustered applications is added to the storage group of the VMware ESX Server and is assigned to both virtual machines using the VMware MUI for ESX Server 2.5.x or VirtualCenter for ESX 3.x/ESX 3i. Additional CLARiiON disks can be assigned to each virtual machine for running other non-clustered applications. Only virtual disks on VMFS volumes are supported with this configuration for ESX 2.5.x. With ESX Server 3.0, RDM volumes are also supported with an in-the-box cluster.

Out-of-the-box cluster

An out-of-the-box cluster consists of virtual machines on multiple physical machines. The virtual disks are stored on shared physical disks, so all virtual machines can access them. Using this type of cluster, you can protect against the crash of a physical machine.

Virtual-to-virtual clustering

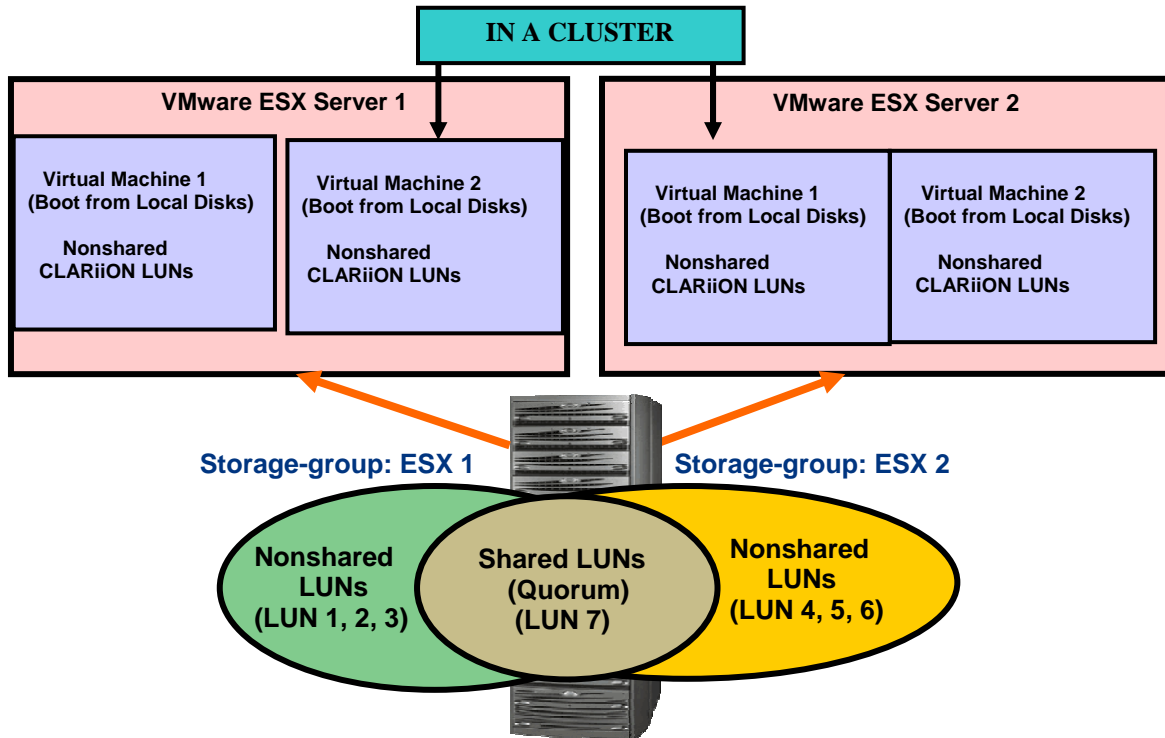


Figure 13. Out-of-the-box cluster configuration

Figure 13 shows two VMware ESX Servers configured with two virtual machines each. The virtual machine boot image must reside on local disks to cluster virtual machines. The bus sharing must be set to **physical**. The quorum device and/or clustered applications residing on CLARiiON disks are shared at the CLARiiON level by assigning the respective LUNs to both storage groups. In this case, LUN 7 is assigned to both storage groups and is a shared LUN. This device is then assigned to **Virtual Machine 2** for ESX Server 1 and **Virtual Machine 1** for ESX Server 2, using VirtualCenter for ESX 3.x/ESX 3i or the MUI for ESX 2.5.x. Additional CLARiiON disks can be assigned to each virtual machine for running non-clustered applications.

The shared resource can be configured as raw disks, VMFS, and/or RDM volumes for a virtual-to-virtual cluster configuration with ESX 2.5.x. If using VMFS volumes for this configuration, the access mode for VMFS volumes must be **shared**. For ESX 3.x/ESX 3i, only RDM volumes (set to physical and virtual compatibility mode) are supported for a virtual-to-virtual cluster configuration. For RDM volumes, both the RDM volume and the VMFS volume that contain the mapping file must be shared by both ESX Servers. The VMFS volume that contains the mapping file must be in **public** access mode.

Physical-to-virtual clustering

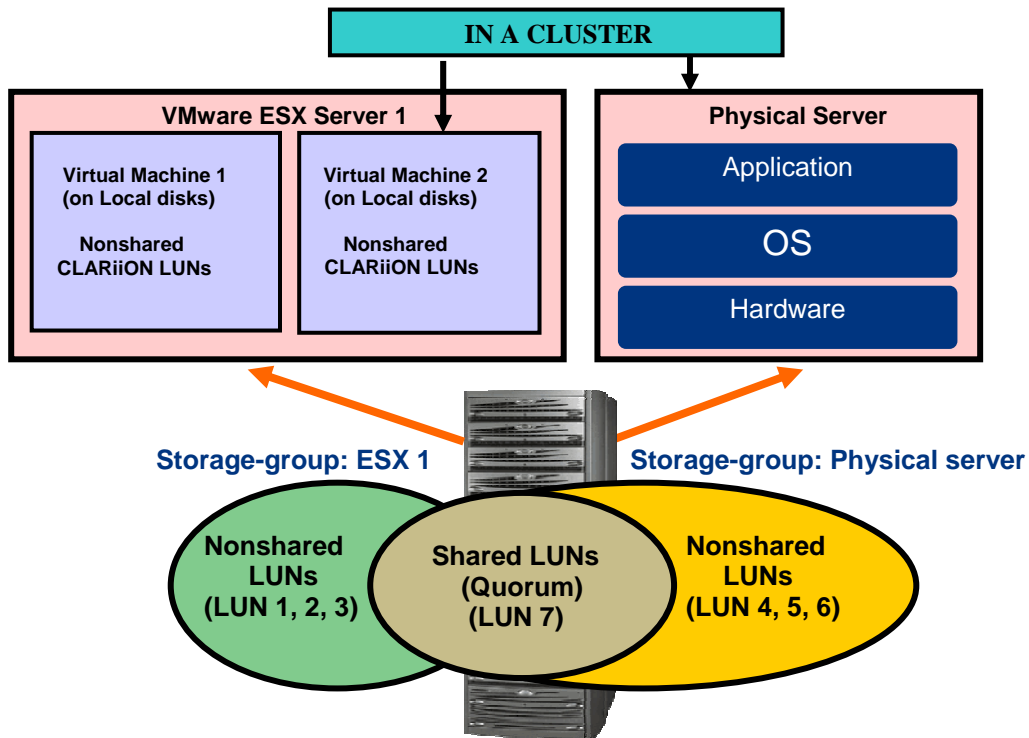


Figure 14. Out-of-the-box cluster configuration (physical to virtual)

Figure 14 shows a VMware ESX Server configured with two virtual machines and a physical server. The virtual machine boot image for the VMware ESX Server must reside on local disks to cluster virtual machines. The quorum device and/or clustered applications residing on CLARiiON disks are shared at the CLARiiON level by assigning the respective LUNs to both storage groups. In this case, LUN 7 is assigned to both storage groups and hence is a shared LUN. This device is then assigned to virtual machine 2 using the VirtualCenter for ESX 3.x/ESX 3i or MUI for ESX 2.5.x. Additional CLARiiON disks can be assigned to each virtual machine for running other non-clustered applications.

The shared resource can be configured as raw disks (ESX 2.x) and/or RDM (physical compatibility mode) volumes, and are supported for a virtual-to-physical cluster configuration. For RDM volumes, only the RDM volume needs to be assigned to both servers. The VMFS volume that contains the mapping file can be in **public** access mode.

CLARiiON and VMware Consolidated Backup

VMware Consolidated Backup, introduced in ESX 3.x/ESX 3i, provides backup of virtual machines residing on a SAN to a physical (proxy) server. Backing up to a proxy server takes the load off the ESX Server and eliminates the need to run backup agents on each virtual machine. The VMware snapshot technology available within the ESX Server provides a mechanism where a redo log for a virtual disk is created; writes are redirected to the redo log, and reads that are not available in the redo log are fetched from the virtual disk. The virtual disk is then accessed by the proxy server. The backup performed by the proxy server can be a file-level or full virtual-machine-level backup.

For a file-level backup, volumes (virtual disks) are mounted on the proxy server as junction points that correspond to a drive letter assigned to each partition in a virtual machine. Third-party software installed on the proxy server performs a file-level backup of these drive letters and exports them to a tape or backup device. File-level backups are only supported for Windows virtual machines.

For full virtual-machine-level backup, the virtual-machine disk images and the configuration files for the virtual machine are exported to a local drive on the proxy server. The backup software accesses these files and moves them to the backup medium. Full virtual-machine-level backup is supported on all guest operating systems.

For restores, the data is transferred to a proxy server or a virtual machine that is running the backup agent. The data is then copied to the respective virtual machine using the CIFS protocol. The other option is to run backup agents on every virtual machine and restore from the backup medium accordingly.

Conditions for VMware Consolidated Backup to work are:

- The physical (Proxy) server must run Windows Server 2003 and the third-party backup software must be installed and configured properly on the proxy server.
- The guest operating system must boot from a CLARiiON LUN.
- EMC recommends that the virtual machine boot LUN and its associated data disks be shared between the ESX and the proxy server. If that is not feasible, the backup can be done over the network.
- VMFS and RDM (virtual compatibility mode) volumes are supported with VMware Consolidated Backup.
- VMware Tools must be installed on the virtual machine for quiescing the virtual machine for backup.
- VMware Consolidated Backup is supported with CLARiiON iSCSI storage.

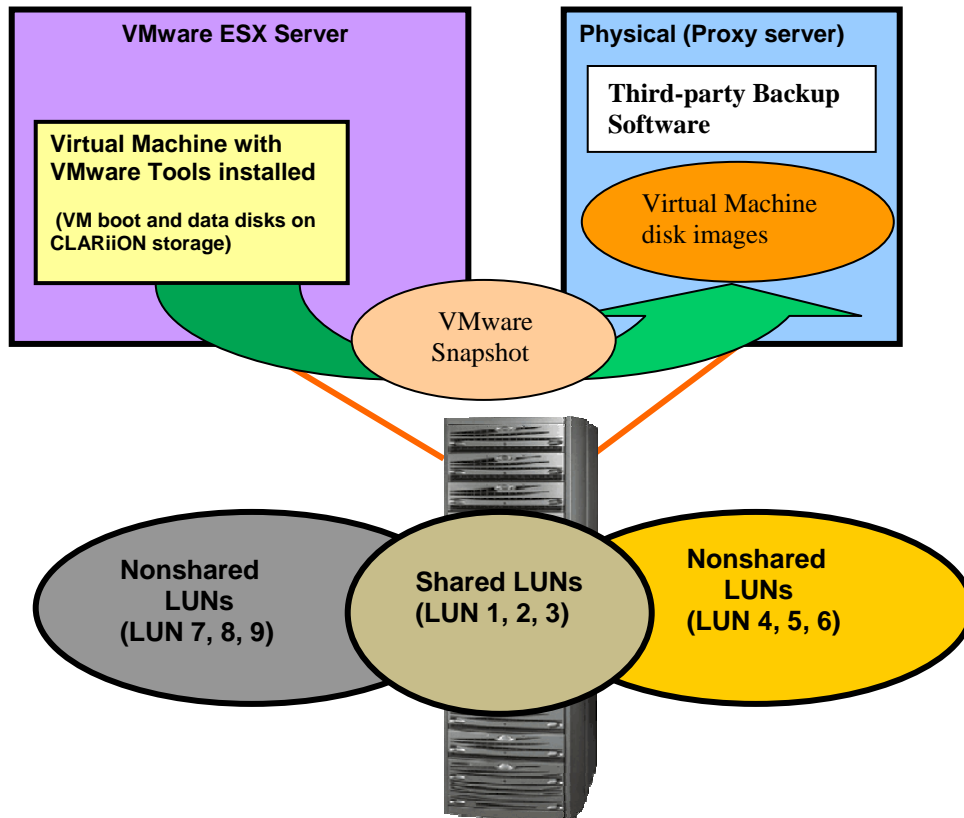


Figure 15. VMware Consolidated Backup with CLARiiON storage

Figure 15 shows a virtual machine installed on the ESX Server and connected to CLARiiON storage. The virtual machine boot and data disks reside on CLARiiON storage. Third-party software is installed on the Windows Server 2003 proxy server and is connected to the CLARiiON system. LUNs 1, 2, and 3 are assigned to the virtual machine and are shared between the storage groups of the ESX and the proxy server. VMware Tools are used to quiesce the virtual machine. Disk images of the drive (at file level) or the entire virtual machines are created using VMware snapshot technology. These images are then made accessible to the proxy server where they can be backed up for archival purposes.

CLARiiON and VMware NPIV support

VMware ESX 3.5/3i added support for NPIV that allows individual virtual machines to have their own virtual WWNs. This allows SAN vendors to implement their QoS tools to distinguish I/O from an ESX Server and a virtual machine. The VMware NPIV is still primitive with a lot of restrictions; however, this section gives the user an idea on how to configure VMware NPIV with CLARiiON storage systems,

To configure VMware NPIV, the HBAs (within the ESX Server) and the FC switches must support NPIV, and NPIV must be enabled on each virtual machine. In order to enable NPIV on a virtual machine, at least one RDM volume must be assigned to the virtual machine. In addition, to use the NPIV feature within VMware, LUNs must be masked to both the VMware ESX Server and the virtual machine that is NPIV enabled.

Figure 16 shows how to enable NPIV for a virtual machine. As mentioned, for the NPIV feature to be enabled, an RDM volume must be presented through the ESX Server to the virtual machine. Note that once NPIV is enabled virtual WWNs are assigned to that virtual machine.

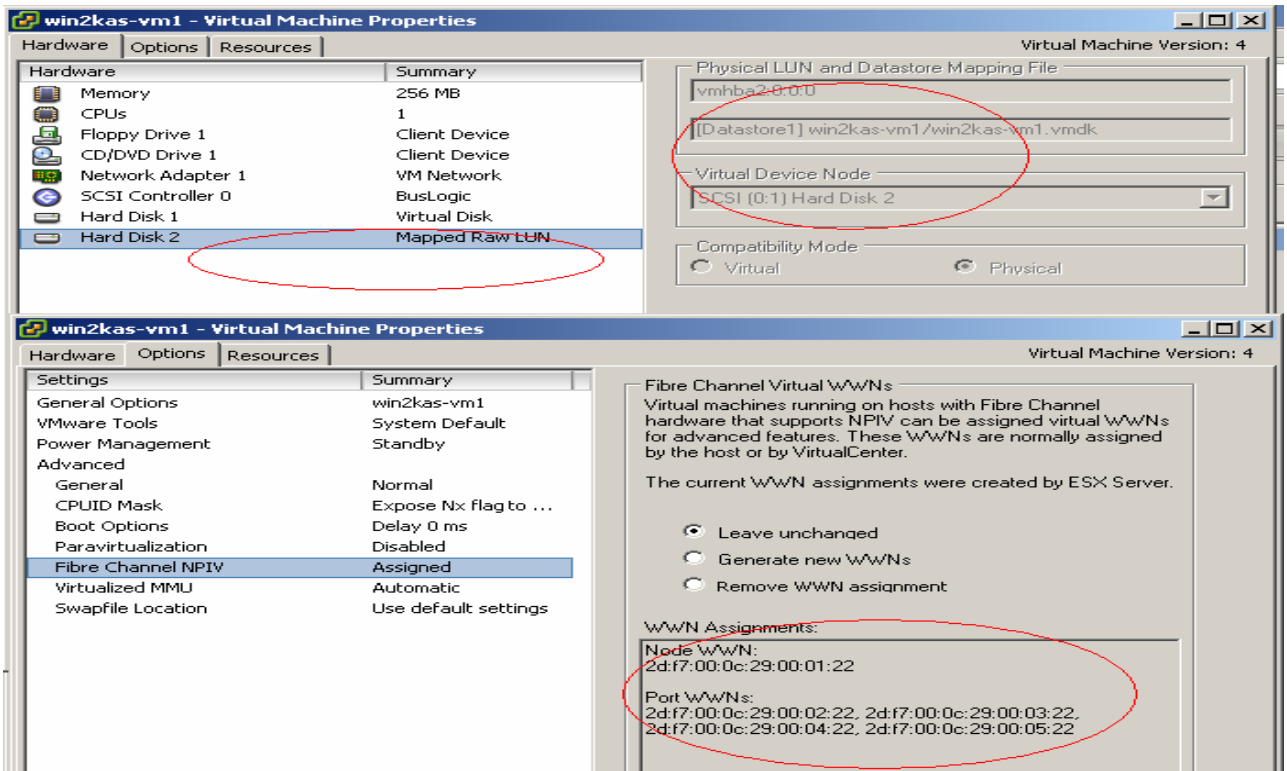


Figure 16. Enable NPIV for a virtual machine after adding a RDM volume

For the switch to see the virtual WWNs, the virtual WWNs' names must be entered manually within the switch interface and then zoned to the storage system. The CLARiON storage system can then see the initiator records for the virtual machine (Virtual WWNs). These initiator records need to be manually registered as shown in Figure 17. A separate storage group can be created for each virtual machine that is NPIV enabled; however, additional LUNs to that virtual machine must be masked to both the ESX Server and the virtual machine that is NPIV enabled.

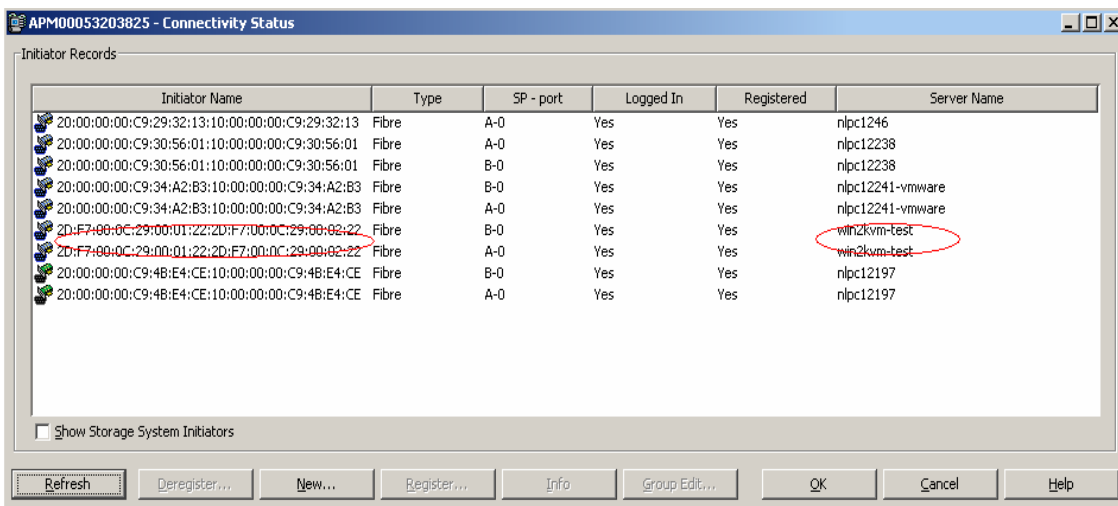


Figure 17. Manually register virtual machine (virtual WWN) initiator records

The following points summarize the steps required to configure NPIV:

1. Ensure the HBA, switch, and ESX version support NPIV.
2. Assign a RDM volume to the ESX Server and then to the virtual machine.
3. Enable NPIV for that virtual machine to create virtual WWNs.
4. Manually enter the virtual WWNs within the switch interface.
5. Zone the virtual WWNs to the CLARiiON storage systems using the switch interface.
6. Manually register the initiator records for that virtual machine using Navisphere.
7. Add the virtual machine (host) to the same storage group as the ESX Server or assign them to a different storage group.
8. To add LUNs to the virtual machine ensure that:
 - a. LUNs are masked to the ESX Server and the virtual machine storage group.
 - b. LUNs have the same host LUN number (HLU) as the ESX Server.
 - c. VMs are defined in different storage groups.

CLARiiON and VMware Site Recovery Manager (SRM)

VMware Site Recovery Manager (SRM) provides a standardized framework to automate site failover in conjunction with Storage Replication Adapters (SRAs) provided by storage vendors. CLARiiON has an SRA for MirrorView that works within the SRM framework to automate most of the steps required for a site failover operation. The initial version of the EMC CLARiiON SRA supports MirrorView/S.

MirrorView/S can be used to synchronously replicate production data changes between data centers. Given the synchronous nature of MirrorView/S, the distances between data centers that communicate over FC are usually less than 200 km. The distances between data centers that communicate over iSCSI are usually shorter, due to the higher latency of IP connectivity. MirrorView/S replicates writes from the source CLARiiON LUN to the target CLARiiON LUN. The write operation from the virtual machine is not acknowledged until both CLARiiON arrays (storage systems) have a copy of the data in their cache. Please consult the white paper *MirrorView Knowledgebook* on Powerlink for further information about MirrorView/S.

SRM requires that the protected (primary) site and the recovery (secondary) site each has two independent virtual infrastructure clients. To use the MirrorView SRA, mirrors need to be created, and secondary LUNs need to be added and placed in a MirrorView/S consistency group. To leverage the test functionality within SRM, SnapView snapshots of the mirrors must exist at the recovery site within the proper CLARiiON Storage Group. (We also recommend that you create snapshots for the mirrors at the protected site, in case a failback is necessary). For installation and configuration information please see the *EMC MirrorView Adapter for VMware Site Recovery Manager Version 1.0 Release Notes*.

The following steps outline the process for initializing an SRM environment using Navisphere Manager and/or Navisphere SecureCLI. The commands must be issued from a management host that is network connected to the production CLARiiON storage array. Note that all of these commands can be performed in the Navisphere Manager GUI or in CLI.

Using Navisphere Manager to configure MirrorView/S

To configure sync mirrors, open the wizard and follow the instructions in the wizard.

NOTE: The MirrorView SRA only supports the MirrorView/S mirror type; also, we recommend setting the sync rate to **High**.

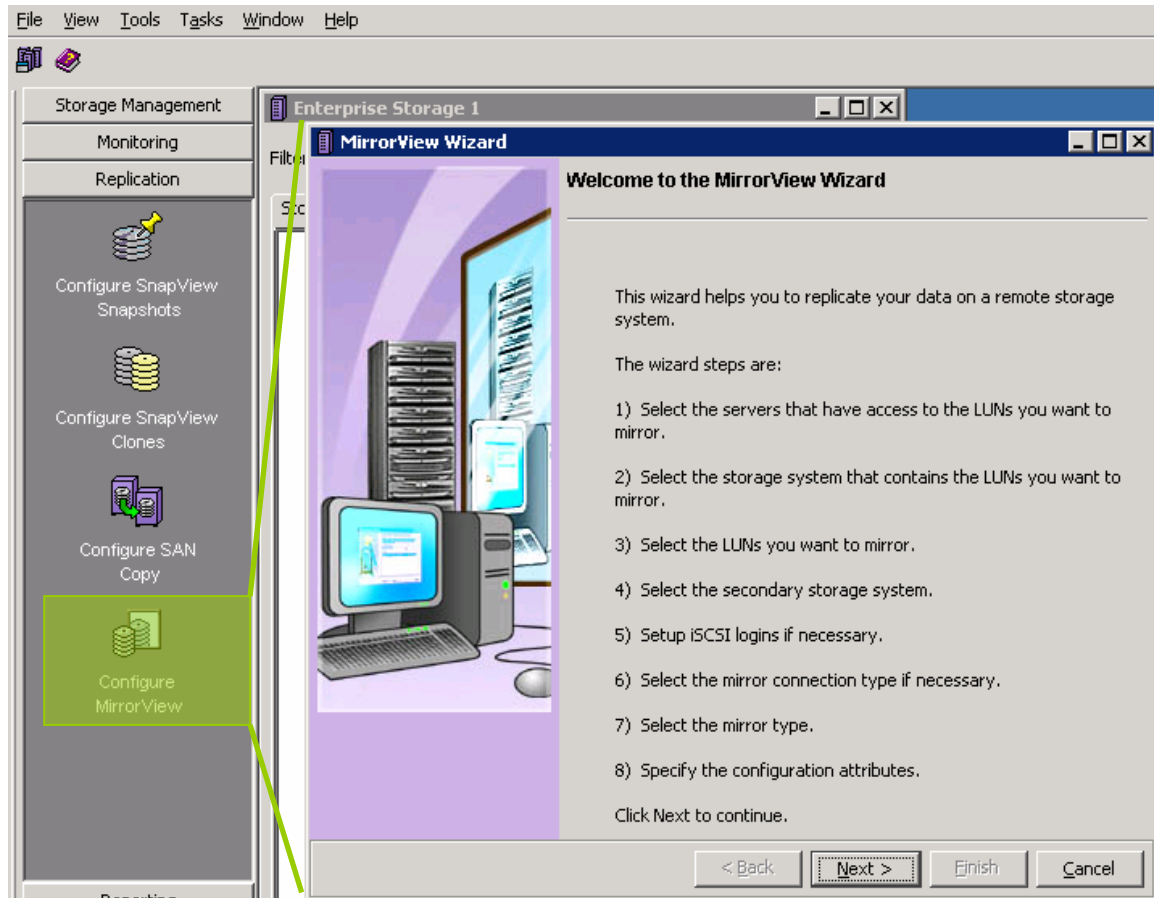


Figure 18. MirrorView Wizard

Configuring sync mirrors via NaviSecCLI

1. If not already established, create a path or paths for remote mirroring between the primary and secondary CLARiiON with this command:

```
naviseccli -h SP ipaddress mirror -sync -enablepath SPhostname  
[-connection type fibre|iscsi]
```

2. Once you have created mirror paths, create a remote mirror of the LUN(s) that you wish to protect with SRM. The LUN(s) on which the mirror is created becomes the primary image.

```
naviseccli -h SP ipaddress mirror -sync -create -lun <LUN_Number>
```

3. The secondary image on the remote CLARiiON can then be added to the primary image. After the secondary image is added, the initial synchronization between the primary and the secondary images is started. The following command assumes that the LUN(s) are already created on the remote CLARiiON storage system.

```
naviseccli -h SP ipaddress mirror -sync -addimage -name <name>  
-arrayhost <sp-hostname| sp ipaddress> -lun <lunnumber| lun uid>
```

4. Even if there is only a single LUN being replicated to the secondary site, you still need to create a consistency group for SRM. The following commands show how to create a consistency group and add existing mirrors to the consistency group.

```
naviseccli -h SP ipaddress mirror -sync -creategroup -name <name>
```

```
naviseccli -h SP ipaddress mirror -sync -addgroup -name <name>  
-mirrorname <mirrorname>
```

5. If for some reason the mirrors are fractured, the **syncimage** option (shown below), can be used to resynchronize the primary and secondary images:

```
naviseccli -h SP ipaddress mirror -sync -syncgroup -name <name>
```

6. While the mirrors are synchronizing or a consistent state, you can add all the LUNs (if you have not already done so) to the ESX Server CLARiiON Storage Group at the protected and recovery site using the following command:

```
naviseccli -h SP ipaddress storagegroup -addhlu -gname <ESX CLARiiON Storage Group  
Name> -hlu <Host Device ID> -alu <Array LUN ID>
```

Using SnapView to configure SnapView snapshots for SRM testing purposes

For SRM testing purposes, you need to create snapshots on the array at the SRM recovery site. Use the wizard to create and configure these snapshots. This wizard will create LUNs automatically to be placed within the Reserved LUN Pool. The default is to allocate 30% storage capacity to the LUN where the snapshot is created. If you have determined that this is not enough for your environment, override the value and select the appropriate percentage. Use the wizard to add the snapshots to the proper CLARiiON Storage Group at the SRM recovery site.

You can also use the **Configure SnapView Snapshot** wizard to create snapshots on the array at the SRM protection site, so that if a failback is necessary, this step has already been performed.

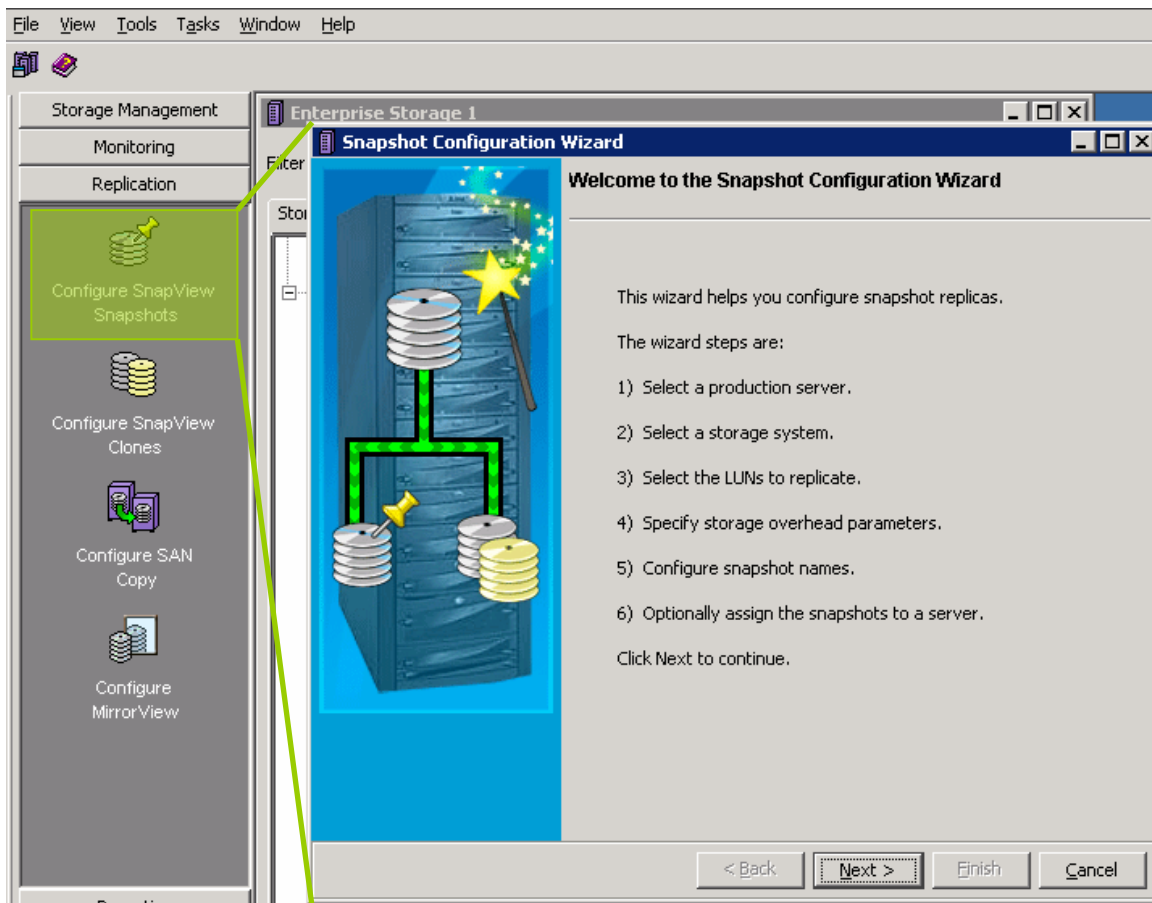


Figure 19. SnapView SnapShot Configuration Wizard

Configuring SnapView snapshots for SRM testing purposes via NaviSecCli

1. Add the LUNs bound for SnapView Sessions into the Reserved LUN Pool.
naviseccli -h SP ipaddress reserved -lunpool -addlun <LUN IDS separated by spaces>
2. Create a snapshot for each LUN at the recovery site, and add the Snapshot to the ESX Server's CLARiiON Storage Group at the recovery site.

(NOTE: This snapshot will not be activated until a user tests the SRM failover operation, in which SRM will create a session and activate it with the corresponding snapshot.)

**naviseccli -h SP ipaddress snapview -createsnapshot <LUN ID>
 -snapshotname VMWARE_SRM_SNAP¹_LUNID**

naviseccli -h SP ipaddress storagegroup -addsnapshot -gname <ESX CLARiiON Storage Group name> -snapshotname <name of snapshot>

¹ The text **VMWARE_SRM_SNAP** must be somewhere in this name for the SRA adapter to function properly.

For more information about using Navisphere CLI with MirrorView, please see the *MirrorView/Synchronous Command Line Interface Reference* available on Powerlink.

After completing the steps listed previously, you need to install SRM and CLARiiON MirrorView Adapter within the Virtual Infrastructure client on the protected and recovery sites. Refer to the *VMware SRM Administration Guide* along with the *EMC MirrorView Adapter for VMware Site Recovery Manager Version 1.0 Release Notes* for installation and configuration instructions.

SRM Protection Groups

A *Protection Group* specifies the items you want to transition to the recovery site in the event of a disaster. A Protection Group may specify things such as virtual machines (VMs), resource pools, datastores, and networks. Protection Groups are created at the primary site. Depending on what the SRM will be protecting, you can define the Protection Group using VMs or based on the application being protected (for example, distributed application across multi-VMs). Usually there is a 1-to-1 mapping between a SRM Protection Group and a CLARiiON consistency group. However, if your CLARiiON model does not support the number of devices being protected within a Protection Group, you can create multiple CLARiiON consistency groups for each Protection Group. Table 3 shows the maximum number of devices allowed per consistency group.

Table 3. Maximum number of mirrors and consistency groups

Parameter	CX3-10c	CX3-20c, CX3-20f	CX3-40c, CX3-40f, CX3-80
Total mirrors per storage system	50	100	200
Total mirrors with write intent log per storage system	25	50	100
Total mirrors per consistency group	8	8	16
Total consistency groups per storage system	8	8	16

SRM recovery plan

The SRM recovery plan is the list of steps required to switch operation of the data center from the protected site to the recovery site. Recovery plans are created at the recovery site, and are associated with a Protection Group created at the protected site. More than one recovery plan may be defined for a Protection Group if different recovery priorities are needed during failover. The purpose of a recovery plan is to ensure priority of a failover. For example, if a database management server needs to be powered on before an application server, the recovery plan can start the database management server, and then start the application server. Once the priorities are established, the recovery plan should be tested to ensure the ordering of activities has been properly aligned for the business to continue running at the recovery site.

Testing the SRM recovery plan

Once the SRM recovery plan is created, it is important to test that the plan performs the operations expected. A recovery plan is shown in Figure 20. To test the plan, click the **Test** button on the menu bar.

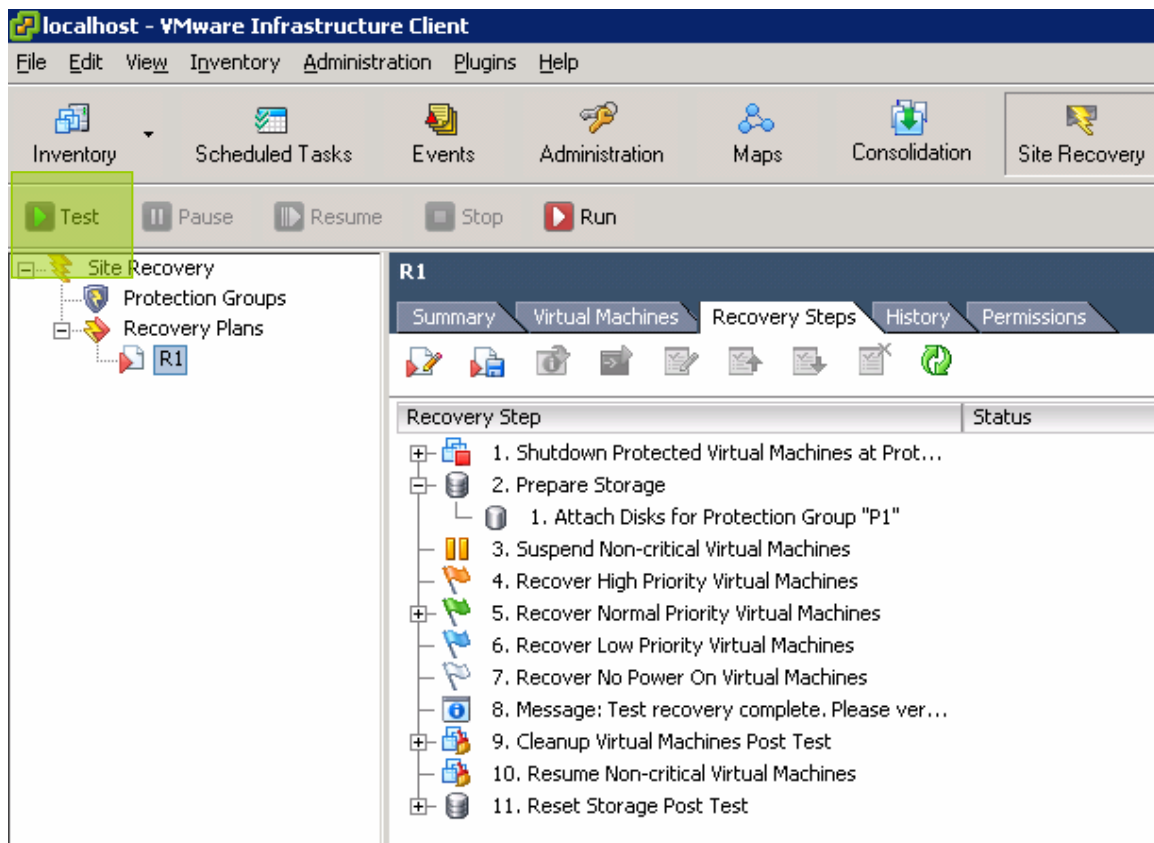


Figure 20. SRM recovery plan

During this test, you would see the following events occur:

1. Production VMs shut down
2. CLARiiON SnapView Sessions are created and activated against the snapshots created above
3. All resources created within the SRM Protection Group carry over to the recovery site
4. VMs power on in the order defined within the recovery plan

Once all the VMs are powered on according to the recovery plan, SRM will wait for the user to verify that the test works correctly. You verify this by opening a console for the VM started at the recovery site and checking the data. After checking your data, click the **Continue** button, and the environment will revert back to its original production state. For more information concerning SRM recovery plans and protection groups, please see the *VMware SRM Administration Guide*.

Executing an SRM recovery plan

Executing an SRM recovery plan is similar to testing the environment with the following differences:

- Execution of the SRM recovery plan is a one-time activity, while running an SRM Test can be done multiple times without user intervention.
- SnapView snapshots are not involved during an executed SRM recovery plan.
- The MirrorView/S secondary copies are promoted as the new primary LUNs to be used for production operation.
- After executing a recovery plan manual steps are needed to resume operation at the original production site.

You should execute a SRM recovery plan only in the event of a declared disaster, to resume operation at the recovery site.

Failback scenarios

The nature of the disaster, and which components of the data center infrastructure are affected, will dictate what steps are necessary to restore the original production data center. For details on how to address different failback scenarios for MirrorView/S, please see the white paper *MirrorView Knowledgebook* on Powerlink. For details on how to address these failback scenarios with the MirrorView SRA, please see the *EMC MirrorView Adapter for VMware Site Recovery Manager Version 1.0 Release Notes*.

Conclusion

EMC CLARiiON and VMware technologies provide the complete Information Lifecycle Management solutions that customers need to consolidate their storage and servers at a low cost. Tools like VMotion, when used with CLARiiON storage, provide online migration of server application workloads without any downtime. VMware HA and Distributed Resource Scheduling coupled with CLARiiON high availability and performance provide reliable and cost-effective solutions. Clustering of virtual machines within the box provides protection against software errors within the cluster.

VMware provides virtualization at the server level while CLARiiON provides protection, performance, and backup at the disk level. Both technologies complement each other, with the high level of functionality and features they provide, to satisfy customer needs.

References

The following documents and resources can be found on Powerlink, EMC's password-protected extranet for partners and customers:

- *EMC Navisphere Manager Administrator's Guide*
- *EMC SnapView for Navisphere Administrator's Guide*
- *EMC SAN Copy for Navisphere Administrator's Guide*
- *EMC MirrorView/Synchronous for Navisphere Administrator's Guide*
- *EMC MirrorView/Asynchronous for Navisphere Administrator's Guide*
- *VMware ESX Server using EMC CLARiiON Storage Systems Solutions Guide*
- *Host Connectivity Guide for VMware ESX Server Version 2.x*
- VMware ESX with CLARiiON Best Practices presentation
- E-Lab Navigator

The following documents and resources can be found on VMware.com:

- VMware resource documents
http://www.vmware.com/support/resources/esx_resources.html
- *VMware ESX Server 3.0.1 and VirtualCenter 2.0.1 Server Configuration Guide*
http://www.vmware.com/pdf/vi3_server_config.pdf
- *VMware ESX Server 3.0.1 and VirtualCenter 2.0.1 SAN Configuration Guide*
http://www.vmware.com/pdf/vi3_esx_san_cfg.pdf
- *VMware ESX Server 3.0.1 and VirtualCenter 2.0.1 Virtual Machine Backup Guide*
http://www.vmware.com/pdf/vi3_vm_backup.pdf
- *VMware ESX Server 3.0.1 and VirtualCenter 2.0.1 Setup for Microsoft Cluster Service*
http://www.vmware.com/pdf/vi3_vm_and_mscs.pdf

Appendix A: Copying data from a VMFS to RDM volume

The export/import CLI command can be used at the ESX Server level to copy data from a VMFS volume to a raw device mapping volume. Consider an example;

The virtual machine resides on a VMFS volume and has the virtual disk name `test.vmdk`. The data on this virtual machine needs to be copied to a RDM volume. Following is an outline of the steps required to move the data.

Store the data (`test.vmdk`) that resides on a VMFS volume (for example, `vmfsprod`) to a temporary location, say `vmimages`, using the `vmkfstools` export function:

For ESX 2.5.x, execute the following command:

```
vmkfstools -e /vmimages/test1.vmdk /vmfs/vmfsprod/test.vmdk
```

For ESX 3.x, execute the following command:

```
vmkfstools -e /vmimages/test1.vmdk  
/vmfs/volumes/vmfsprod/test.vmdk
```

Create a raw device mapping on `vmhba0:0:1:0`, which is a CLARiiON LUN, and the mapping file called `rvm.vmdk` that resides on a VMFS volume (for example, `VMFS1`):

For ESX 2.5.x, execute the following command:

```
vmkfstools -r vmhba0:0:1:0 /vmfs/vmfs1/rvm.vmdk
```

Import the data from `vmimages` to the mapping file residing on a VMFS (for example, `vmfs1`), which points to the RDM volume:

```
vmkfstools -i /vmimages/test1.vmdk /vmfs/vmfs1/rvm.vmdk
```

You may see geometry errors after this command is executed. Ignore these errors.

For ESX 3.x, execute the following command:

```
vmkfstools -i /vmimages/test1.vmdk -d  
rdm:/vmfs/devices/disks/vmhba0:0:1:0 /vmfs/volumes/vmfs1/rvm.vmdk
```

Note: The import command for ESX 3.x creates a raw device mapping volume and imports the data from VMFS volume. Assign the RDM volume to the virtual machine. Power on the virtual machine to ensure the data on the virtual machine is intact.

Appendix B: Using vm-support on VMware ESX Server

VM support is the command tool used to aid in diagnostics and/or troubleshooting of the ESX Server. This service tool is supported on ESX 3.x and 2.x. For ESX 3i, use VI Client's Export Diagnostics Data option to get vm-support files.

The following procedure outlines the steps executed on the ESX 3.x service console. Enter the **vm-support** command on the ESX service console. This script generates a .tgz file in the current directory. Extract this file using the following command

```
tar -zxvf "Vm-support file name"
```

Note that WinZip cannot be used to extract vm-support script output. You have to have a Linux machine to extract these files. Once these files get extracted, a folder with the version of vm-support is created.

Important files to look at within this folder from the storage point of view are as follows:

- /tmp/vmware_XXX.txt – ESX version and patch information
- /var/log/messages – For hardware BIOS versions
- /tmp/chkconfig.*.txt – Confirm naviagent is installed
- /proc/scsi/lpfc or /proc/scsi/qla_2xx – HBA driver versions
- /tmp/esxcfg-swiscsi – Software iSCSI initiator information for ESX 3.x only
- /proc/scsi/vmkiscsi – iSCSI target IP address and inq information for ESX 3.x only
- /tmp/esxcfg-mpath – ESX path information for ESX 3.x
- /tmp/vmkmultipath – ESX path information for ESX 2.x

Additional information related to vmkernel configuration can be found at:

- vmkpcidivv – Device configuration information available only on ESX 2.x systems
- /proc/vmware/config – For additional SCSI, disk, and file system configuration information
- /home/vmware/ – Contains VM configuration files information
- var/log/vmkernel and /var/log/vmkernel.1 – For additional troubleshooting information