EMC® Host Connectivity Guide for VMware ESX Server

P/N 300-002-304
REV A17

EMC Corporation
Corporate Headquarters:
Hopkinton, MA 01748-9103
1-508-435-1000
www.EMC.com

Published December, 2008

Contents

Figures

Tables

Preface

As part of an effort to improve and enhance the performance and capabilities of its product line, EMC from time to time releases revisions of its hardware and software. Therefore, some functions described in this document may not be supported by all revisions of the software or hardware currently in use. For the most up-to-date information on product features, refer to your product release notes.

If a product does not function properly or does not function as described in this document, please contact your EMC representative.

This guide describes the features and setup procedures for VMware ESX Server host interfaces to EMC Symmetrix and CLARiiON Storage Array systems.

Audience

This guide is intended for use by storage administrators, system programmers, or operators who are involved in acquiring, managing, or operating Symmetrix, CLARiiON, and host devices.

Readers of this guide are expected to be familiar with the following topics:

- Symmetrix or CLARiiON system operation
- VMware ESX Server operating environment

Related documentation

For the most up-to-date information for supported server and HBA combinations, always consult the *EMC Support Matrix* (ESM), available through E-Lab Interoperability Navigator (ELN) at: http://elabnavigator.EMC.com, under the **PDFs and Guides** tab.

For documentation on Navisphere, refer to the Navisphere documentation in the Documentation Library on EMC Powerlink:

http://Powerlink.EMC.com

For VMware specific documentation, such as the *VMware ESX Server Release Notes*, *ESX Server Administration Guide*, and the *ESX Server Installation Guide*, go to:

http://www.vmware.com/support

For a list of supported guest operating systems, refer to the VMware *Guest Operating System Installation Guide*, located at:

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

Conventions used in this guide

EMC uses the following conventions for notes and cautions.

Note: A note presents information that is important, but not hazard-related.

CAUTION

A caution contains information essential to avoid damage to the system or equipment. The caution may apply to hardware or software.

Typographical conventions

EMC uses the following type style conventions in this guide:

Normal	Used in running (nonprocedural) text for: • Names of interface elements (such as names of windows, dialog boxes, buttons, fields, and menus) • Names of resources, attributes, pools, Boolean expressions, buttons, DQL statements, keywords, clauses, environment variables, filenames, functions, utilities • URLs, pathnames, filenames, directory names, computer names, links, groups, service keys, file systems, notifications
Bold	Used in running (nonprocedural) text for: • Names of commands, daemons, options, programs, processes, services, applications, utilities, kernels, notifications, system call, man pages Used in procedures for: • Names of interface elements (such as names of windows, dialog boxes, buttons, fields, and menus) • What user specifically selects, clicks, presses, or types

Italic:	Used in all text (including procedures) for: • Full titles of publications referenced in text • Emphasis (for example a new term) • Variables
`Courier`	Used for: • System output, such as an error message or script • URLs, complete paths, filenames, prompts, and syntax when shown outside of running text
`Courier bold`	Used for: • Specific user input (such as commands)
`Courier italic`	Used in procedures for: • Variables on command line • User input variables
< >	Angle brackets enclose parameter or variable values supplied by the user
[ ]	Square brackets enclose optional values
|	Vertical bar indicates alternate selections - the bar means “or”
{ }	Braces indicate content that you must specify (that is, x or y or z)
...	Ellipses indicate nonessential information omitted from the example

Where to get help

EMC support, product, and licensing information can be obtained as follows.

Product information — For documentation, release notes, software updates, or for information about EMC products, licensing, and service, go to the EMC Powerlink website (registration required) at:

`http://Powerlink.EMC.com`

Technical support — For technical support, go to EMC Customer Service on Powerlink. To open a service request through Powerlink, you must have a valid support agreement. Please contact your EMC sales representative for details about obtaining a valid support agreement or to answer any questions about your account.

Your comments

Your suggestions will help us continue to improve the accuracy, organization, and overall quality of the user publications. Please send your opinion of this guide to:

`techpub_comments@EMC.com`

1

Introduction

This chapter provides information in preparation for the installation and configuration of VMware ESX Server attached to EMC CLARiiON and Symmetrix storage systems.

Overview

This document is meant to assist in the installation and configuration of VMware ESX Server attached to EMC® CLARiiON® and Symmetrix® storage systems.

Related documentation

Refer to "Related documentation" on page 11.

HBA settings for use with VMware ESX Server

This section discusses the EMC recommended settings to be used with Emulex and QLogic HBAs in a VMware ESX Server environment.

Recommended Emulex settings

EMC recommends using the firmware settings in provided by Emulex. Additionally, no modifications need to be made to the driver to enhance its performance in VMware ESX Server.

Recommended QLogic NVRAM settings

EMC recommends that no modifications be made to the QLogic driver to enhance its performance in VMware ESX Server. However, EMC does recommend making some modifications to the NVRAM to improve the HBAs behavior.

Table 1 on page 17 lists the QLogic NVRAM parameters and their values. The QLogic default values are those that ship with a standard HBA that wasn't included in an EMC package. The EMC-recommended settings depend upon whether or not failover functionality is used. The settings listed under the **EMC Recommended Settings for VMware ESX** heading are those that have been tested and determined to be applicable in a Linux environment.

The settings are configurable in NVRAM using the **Host Adapter Settings**, **Advanced Settings**, and **Extended Firmware Settings** menus.

Table 1 Recommended QLogic NVRAM settings

Category	Parameter	QLogic Default	Recommended Setting
Host Adapter Settings	Host Adapter BIOS	Disabled	Disabled
	Frame Size	2048	2048
	Loop Reset Delay	5	5
	Adapter Hard Loop IP	Disabled	Disabled
	Hard Loop ID	0	125
	Spin-up Delay	Disabled	Disabled
	Connection Options (topology)	2	1 [a]
	Fibre Channel Tape Support	Enabled	Disabled
	Data Rate	2	2
Advanced Adapter Settings	Execution Throttle	16	256
	LUNs per Target	8	256
	Enable LIP Reset	No	No
	Enable LIP Full Login	Yes	Yes
	Enable Target Reset	Yes	Yes
	Login Retry Count	8	8
	Port Down Retry Count	8	30
	Link Down Timeout	15	30
	Extended Error Logging	Disabled	Disabled
	Operation Mode	0	0
	Interrupt Delay Timer	0	0

Table 1 Recommended QLogic NVRAM settings (continued)

Category	Parameter	QLogic Default	Recommended Setting
Selectable Boot Settings	Selectable Boot	Disabled	Disabled
	(Primary) Boot Port Name, LUN	WWNN [b]	WWNN [b]
	Boot Port Name, LUN	0	0
	Boot Port Name, LUN	0	0
	Boot Port Name, LUN	0	0

a. The values for Connection Options are: 0 for Loop Only, 1 for Point-to-Point, or 2 for Loop preferred, otherwise Point-to-Point. For VMware ESX Server-attach, EMC recommends setting Connection Options to 1 when attached to a fabric and to 0 when attached directly to an EMC storage array.

b. The WWNN of the server's boot LUN must be listed in this field.

VMware ESX Server features

VMware ESX Server is similar to Linux, but is its own operating system. VMware GSX Server is an application that runs on top of a standard Linux or Windows operating system.

VMware ESX Server consists of virtualization software that provides server consolidation and allows several instances of operating systems running as virtual machines on one physical machine. Currently, VMware ESX Server installs and operates as a 32-bit operating system on Intel-based and Opteron-based servers.

VMotion

VMotion is part of VMware Virtual Center, a systems management and provisioning tool used by system administrators. VMotion technology allows the migration of virtual machines from one physical server to another physical server. For instance, VMotion could be used to migrate virtual machines from one server to another so that maintenance may be performed on the original server. This allows the operating systems running in the virtual machine partitions to continue running their workloads.

Raw disk usage and caveats

Raw disk (non-RDM) is used only in ESX Server versions previous to 2.5. For ESX Server 2.5 and greater, please use RDMs for direct access to SAN LUNs. Please skip to the "Raw Device Mappings (RDM)" on page 20.

There are occasions where the Virtual Machine configuration file will need to be manually edited when using directly accessed raw devices (and not RDMs). If an event occurs that forces a change in the paths to the devices, then it is possible that access to a raw disk device may be lost due to a dead path. In such a case as this, the Virtual Machine configuration file will need to be modified to reflect a usable path for the raw device.

For instance, if all the LUNs referenced by the Virtual Machine are still accessible by the VMware ESX Server, but the original paths have changed due to a problem such as a lost cable, then the Virtual Machine configuration file will need to be edited.

Example Assume a LUN originally has four paths with the following notation:

```
vmhba3:3:5
vmhba3:4:5
vmhba4:3:5
vmhba4:4:5
```

The Virtual Machine configuration file contains a line referencing a path that is now dead:

```
scsi0:0.present = "TRUE"
scsi0:0.name = "vmhba3:3:5:0"
```

A path fault occurs and raw disk device `"vmhba3:3:5"` is no longer. The user will need to locate the new canonical name for the LUN. If all paths from vmhba3 are no longer accessible, and assuming that target numbers have not changed, then the canonical name will be `"vmhba4:3:5"`. The user will have to edit the Virtual Machine configuration file and change the scsi0:0.name value:

```
scsi0:0.present = "TRUE"
scsi0:0.name = "vmhba4:3:5:0"
```

If Raw Device Mappings (RDMs) are used instead of raw disks, and all the LUNs referenced by the RDMs are still accessible by the ESX Server using the original LUN numbers, no user action is required. ESX will access the LUNs via the currently available paths.

Raw Device Mappings (RDM)

VMware ESX Server v2.5.0 introduced a new technology called Raw Device Mapping (RDM). Essentially, raw device mapping is a SCSI pass-through technology that allows Virtual Machines to pass SCSI commands directly to the physical hardware. This is an improvement over using raw devices.

Raw Device Mappings use a mapping file located on a VMFS volume that hold the metadata pertaining to a raw device. This mapping file becomes a symbolic link from the VMFS volume to the raw LUN. Instead of allocating the raw LUN to a Virtual Machine, the mapping file is now used. The metadata contained in the mapping file is used to manage and redirect disk access to the physical device.

Note: Using RDMs in shared mode with Fibre Channel HBA is *not* supported. The HBAs must be allocated solely to the Virtual Machines in order to use the raw disks.

2

Installing VMware ESX Server v2.x with EMC Arrays

This chapter covers installation information for the VMware ESX Server v2.x with Symmetrix and CLARiiON arrays.

Overview

The Fibre Channel HBA driver functions as a device driver layer below the standard VMware SCSI adapter driver. The Fibre Channel interface is therefore transparent to the VMware disk administration system.

Note: VMware ESX Server is NOT Linux and, therefore, requires its own EMC-qualified drivers, not the standard EMC-qualified Linux drivers.

Install VMware ESX Server from the CD. The VMware ESX Server installation is based on a modified Red Hat v7.2 v2.4.9 kernel. However, VMware ESX Server is NOT Linux.

VMware recommends partitioning the internal hard drive into four partitions. For example:

/boot	50 MB	primary
swap	256 MB to 1 GB	primary
/	2.5 GB	primary
/vmimages	4 GB	

Note: A swap partition is required by both the service console and the VMkernel. The reference to the swap partition in this example is for the service console. The swap partition and core dump partitions for the VMkernel can be created using the GUI after the initial installation of the service console OS.

Note: EMC supports fabric boot with VMWare ESX Server v2.5.x and later using the QLogic or Emulex HBAs. Virtual Machines are recommended to boot over the fabric.

- Refer to the *EMC Support Matrix* for specifics on which operating systems are supported as Virtual Machines.

You may customize the installation according to your server and the amount of memory and hard disk space you have available.

Recommendations

The following are a few recommendations for the installation:

- Use static IP addresses.

- Set the hardware clock when prompted.
- Create at least one user account other than root.

Note: emacs, samba, and NFS are not enabled by default in the Console OS.

- Reboot the system after completing the installation.
- For VMware ESX Server v2.1.x installations, when the system reboots, you will be prompted with two options in the LILO boot prompt:

  `linux-up,linux`

  Assuming the system on which you are installing has multiple CPUs, the default boot image will be `linux`. This is expected and acceptable to use.

- For VMware ESX Server v2.5.x installations, when the system reboots, you are prompted with three options in the LILO boot prompt:

  `esx, linux-up,linux`

  The default boot image for VMware ESX Server v2.5.x is `esx`.

  You can now log in to the VMware server using IE or Netscape with the host name or IP address. This allows you access to the VMware host using the Management User Interface (MUI).

  After you log into the host, the first dialog is the **VMware ESX Server Status Monitor** dialog.

  The **Status Monitor** dialog provides a summary reporting of the system CPU and processor usage by the Virtual Machines as well as a visual status report of the Virtual Machines themselves.

VMware ESX Server options

Select the **Options** tab. The **Options** dialog provides you with a list of actions that can be performed to configure or modify the VMware ESX Server, as shown in Figure 1.

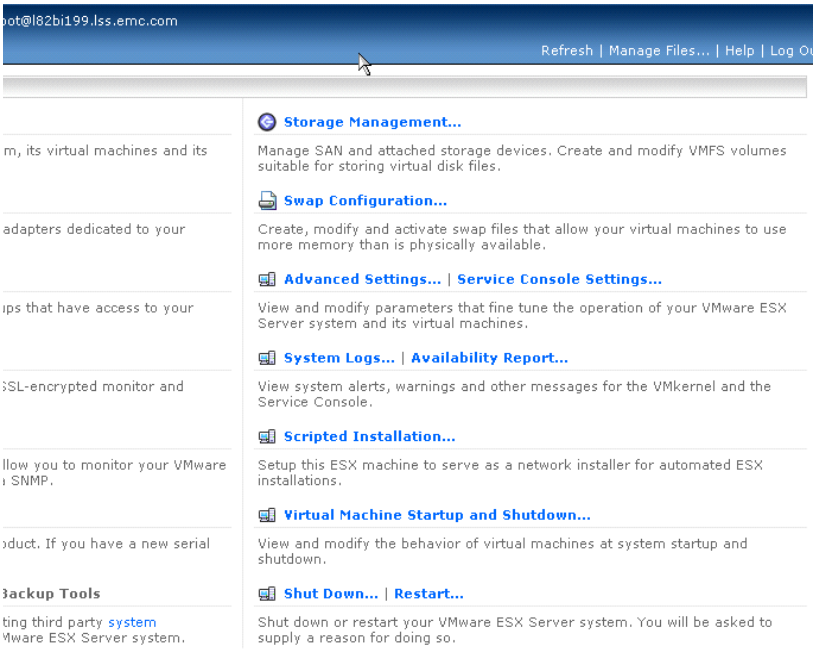

Figure 1 **VMware ESX Server options**

From the **Options** tab, select the **Startup Profile** menu. The Startup Profile creates the VMware ESX Server boot configuration and allows you to specify whether the server resources should be allocated to the Service Console, the Virtual Machines, or shared between the two.

Fibre Channel and SCSI HBAs can be allocated for use with the Virtual Machines only or can be shared between the Virtual Machines and Service Console. In the case of booting from the fabric, the FC HBAs must be shared so that the Service Console and the Virtual Machines can access the boot device.

Networking cards cannot be shared between the Service Console and the Virtual Machines so they must be allocated either to the Service

Console or to the Virtual Machines. Therefore, a minimum of two NICs must be installed in the system.

For an example of the **Startup Profile** dialog box, see Figure 2.

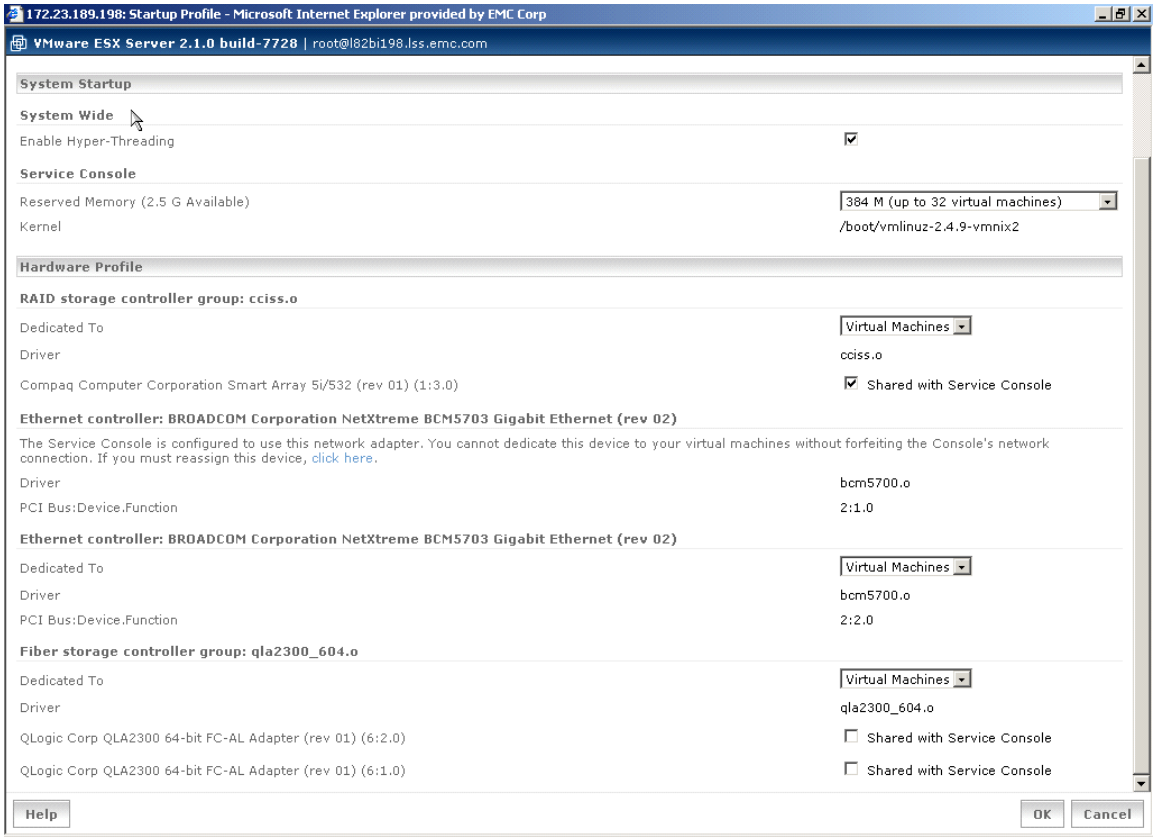

Figure 2 **Startup Profile dialog box**

After the changes appropriate to your system are made, save the configuration by clicking **OK**.

After saving the configuration changes, reboot the host so that the changes can take effect.

After the VMware host reboots, proceed with the VMkernel configuration.

- The **Network Connections** must be set up according to the network interface cards (NICs) being used and the customer environment. Note that the server must have at least two NICs. Because NICs may not be shared, one must be allocated to the Console and at least one must be allocated to the Virtual Machines.

  Note: It is not necessary to reboot the VMware ESX Server into the `esx` kernel in order for network configuration changes to take effect.

- The **Security Settings** selected are entirely dependent upon the customer's desires. VMware recommends that customers use the High Security setting so all connections into ESX will be encrypted. If FTP and Telnet are required, the Medium Security setting will need to be selected. The **Security Settings** dialog box is shown in Figure 3 on page 27.

Figure 3 **Security Settings dialog box**

- **Swap Configuration** creates, modifies, and activates swap files that allow your virtual machines to use more memory than is physically available. Whether or not you want to enable a swap file depends on the amount of memory in the server. It is recommended that a swap file be created so that the virtual machines will be able to use more memory than is actually available in the server.

Note: When booting from an internal hard drive, it is recommended that the swap file resides on the internal device, rather than a SAN-attached device.

- If required, the SNMP agent is enabled in the **SNMP Configuration** section.
- The **Advanced Settings** section allows you to modify the VMkernel parameters. Currently, EMC tests with and supports the majority of the default VMkernel parameters. The table below shows the parameters, their descriptions, and the VMware and EMC default settings. It is highly recommended that the EMC default parameters are used when attached to EMC Storage Arrays.

VMkernel parameters

A number of parameters define the behavior of the VMkernel. Currently, EMC recommends that the majority of the parameter values remain the same as the VMware defaults. The exception at this time is the `Disk.MaxLun` parameter.

These values can be viewed and modified using the MUI and in the subdirectories within the `/proc/vmware/config` directory.

To view the listing of directories in the `/proc/vmware/config` directory, change to the directory and perform a listing:

```
cd /proc/vmware/config
ls -la
dr-xr-xr-x 2 root root 0 Jan 18 06:07 Cpu
dr-xr-xr-x 2 root root 0 Jan 18 06:07 Disk
dr-xr-xr-x 2 root root 0 Jan 18 06:07 FileSystem
dr-xr-xr-x 2 root root 0 Jan 18 06:07 Irq
dr-xr-xr-x 2 root root 0 Jan 18 06:07 Mem
dr-xr-xr-x 2 root root 0 Jan 18 06:07 Migrate
dr-xr-xr-x 2 root root 0 Jan 18 06:07 Misc
dr-xr-xr-x 2 root root 0 Jan 18 06:07 Net
dr-xr-xr-x 2 root root 0 Jan 18 06:07 Numa
dr-xr-xr-x 2 root root 0 Jan 18 06:07 Scsi
```

These directories provide the types of parameters that are listed in each. For instance, the `/proc/vmware/config/Disk` directory contains the parameters pertaining to the Disk subsystem. Any changes made to the parameters in this directory will be specific to Disks. For example, `Disk.MaxLUN` is part of the `/proc/vmware/config` file located in the Disk directory and the parameter itself is the `MaxLUN` setting.

If desired, changes to the values of these parameters made by performed using the MUI. Log into the server via the MUI and select the **Options** tab. In the line for Advanced Settings | Service Console

Settings, select **Advanced Settings**. You will then be presented with a window that contains the parameters and their values and descriptions of the parameters.

The parameters are listed in alphabetical order. An example of the MUI dialog listing the beginning of the parameter list is shown in Figure 4.

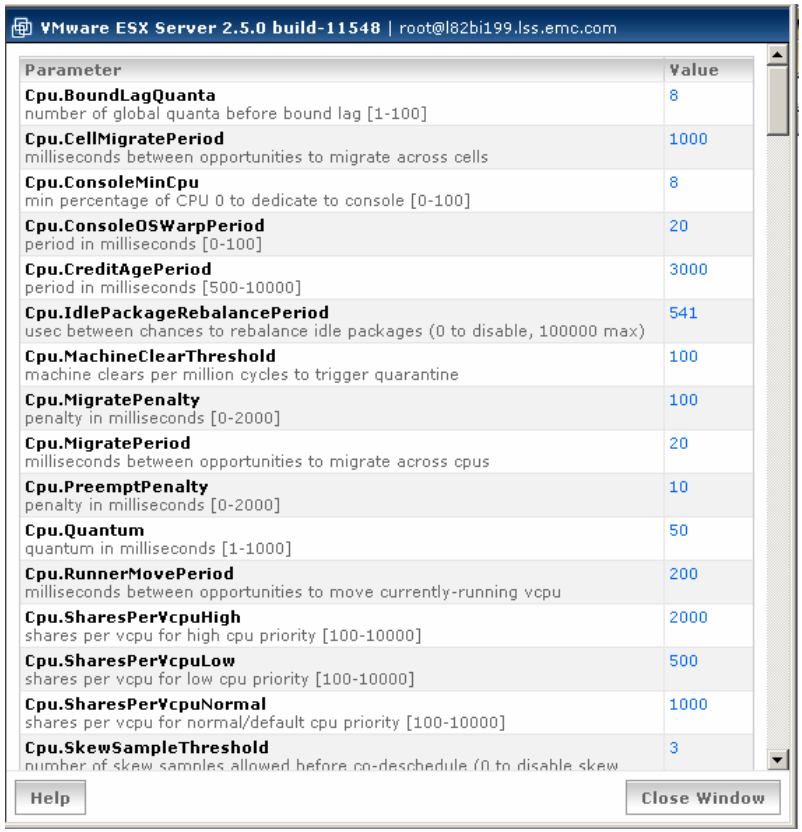

Figure 4 **VMkernel parameter listing**

The values of the VMkernel parameters also can be set at the command line. This is done by echoing the `/proc` filesystem. As mentioned previously, the parameters are ultimately listed in the `/proc/vmware/config` directory. The following example demonstrates the steps to view and change the value of a particular setting.

In this particular example, the LogMultiPath parameter is being enabled.

```
[root@l82bi199 /]# cd /proc/vmware/config/Scsi
[root@l82bi199 Scsi]# ls
ConflictRetries LogAborts LogMultiPath PassthroughLocking
[root@l82bi199 Scsi]# cat LogMultiPath
LogMultiPath (Log path state changes) [default = 0]: 0
[root@l82bi199 Scsi]# echo 1 >> LogMultiPath
[root@l82bi199 Scsi]# cat LogMultiPath
LogMultiPath (Log path state changes) [default = 0]: 1
```

- If the parameter is changed by echoing the /proc at the command line, this change will not persist through a reboot.
- Although it is not recommended, if you prefer to use the command line, you may modify the /etc/rc.local file so that the changes you have via echoing /proc at the command line will be run upon a reboot.
- The /etc/rc.local file is that script that contains your initialization parameters and is run after all of the standard system init scripts. It is recommended that you append your modifications to the end of the file.

  For example, see the end of the /etc/rc.local file for the addition of the /proc modifications:

```
[root@l82bi199 root]# more /etc/rc.local
#!/bin/sh
#
# This script will be executed *after* all the other init scripts.
# You can put your own initialization stuff in here if you don't
# want to do the full Sys V style init stuff.

# BEGINNING_OF_VMWARE_RC_DOT_LOCAL
if ( uname -a | grep -q vmnix ); then
 R="VMware ESX Server 2.5.0"
else
 R="Linux"
fi
arch=$(uname -m)
a="a"
case "_$arch" in
 _a*) a="an";;
 _i*) a="an";;
esac
```

```
NUMPROC=`egrep -c "^cpu[0-9]+" /proc/stat`
if [ "$NUMPROC" -gt "1" ]; then
 SMP="$NUMPROC-processor "
 if [ "$NUMPROC" = "8" -o "$NUMPROC" = "11" ]; then
 a="an"
 else
 a="a"
 fi
fi

# This will overwrite /etc/issue at every boot. So, make any changes
   you
# want to make to /etc/issue here or you will lose them when you reboot.
echo "" > /etc/issue
echo "$R" >> /etc/issue
echo "Kernel $(uname -r) on $a $SMP$(uname -m)" >> /etc/issue

cp -f /etc/issue /etc/issue.net
echo >> /etc/issue

#Echo the /proc filesystem to enable the
#multipath logging. This will ensure that
#the change persists through reboots.

echo 1 >> /proc/vmware/config/Scsi/LogMultiPath

# END_OF_VMWARE_RC_DOT_LOCAL
```

By default, the maximum number of LUNs scanned per disk array by ESX v2.x is 8. Depending upon the EMC storage array configuration and the devices that are being allocated to the VMware host, it is recommended that the following parameter be changed in the VMkernel configuration to accommodate the LUNs allocated to the host.

For example:

```
Disk.MaxLun == 128 (The default is 8.)
```

If the Symmetrix or CLARiiON Storage Group has LUNs allocated to the VMkernel addressable as greater than eight, then this parameter must be increased. The `DiskMaxLun` parameter can be as large as 255. However, the maximum number of LUNs that can be managed by a single VMware ESX Server is 128. The VMware ESX Server will simply stop scanning the SCSI bus once it has seen 128 actual LUNs.

Note: Please note that increasing the `Disk.MaxLun` parameter may slow down the SCSI scanning time.

In VMware ESX Server v2.1.x, the VMkernel is configured with support for sparse LUNs by default. The `Disk.SupportSparseLun` parameter allows the VMkernel to scan past non-existent LUNs, that is, allows LUN skipping. When attaching to EMC storage arrays, it is strongly recommended that the `Disk.SupportSparseLun` parameter be set at the default of 1.

Driver configuration in VMware ESX Server v2.x

VMware ESX Server is NOT Linux and, therefore, requires its own EMC-qualified drivers, not the standard EMC-qualified Linux drivers.

Note: Refer to the *EMC Support Matrix* for the latest supported drivers.

VMware ESX Server v2.x contains QLogic and Emulex drivers native to the esx kernel. At the time this document was written, the drivers qualified for VMware ESX Server v2.1x are the QLogic v6.04.02 driver and the Emulex v4.02q driver. The drivers qualified for VMware ESX Server v2.5.x are the QLogic v6.07.02 driver and the Emulex v2.01g driver.

These drivers are the defaults for their respective HBA families and are contained within the VMware ESX Server v2.x kernel.

No modifications need to be made to the VMkernel in order to use these drivers. There is no need to follow the documentation to download, compile, and install the EMC-qualified Linux drivers. The source code for VMware ESX Server v2.x is not included in the VMware distribution so the Linux drivers may not be compiled. Additionally, the Linux Emulex and QLogic drivers need to be ported from Linux to VMware ESX Server in order for the Linux-specific drivers to be usable in VMware ESX Server.

The drivers should be enabled in the **Startup Profile** menu. The Startup Profile may be modified if the HBAs are added to the system after the installation and initial configuration of VMware ESX Server.

The command **vmkpcidivy** can also be used to create a boot configuration for the VMkernel. When using **vmkpcidivy** interactively, you may specify whether your storage controllers are allocated to the Service Console or to the Virtual Machines or shared. The boot configuration created with `vmkpcidivy` will automatically include the default QLogic and Emulex drivers included in the kernel.

An example of the output produced when running `vmkpcidivy` in interactive mode is as follows:

```
[root@l82bi199 /]# vmkpcidivy -i

Checking for existing VMnix Boot Configurations.

The following VMnix kernel images are defined on your system:

Boot image configuration: vmnix
 Image file: /boot/vmlinuz-2.4.9-vmnix2
 Memory: 384M
 Service Console devices:
 Ethernet controller: BROADCOM Corporation NetXtreme BCM5703 Gigabit Ethernet (rev
   02)
 RAID storage controller: Compaq Computer Corporation Smart Array 5i/532 (rev 01)
   (shared)
 VM devices:
 Ethernet controller: BROADCOM Corporation NetXtreme BCM5703 Gigabit Ethernet (rev
   02)
 Fiber storage controller: QLogic Corp QLA231x/2340 (rev 02)
 Ethernet controller: Intel Corporation 82557 [Ethernet Pro 100] (rev 08)
 RAID storage controller: Compaq Computer Corporation Smart Array 5i/532 (rev 01)
   (shared)

Type in the name of the boot image configuration you wish to configure
or type "new" to create a new image [vmnix]:

Selected configuration was 'vmnix'.

Configuring your setup. Enter in the options that you want. Defaults have
      been provided.

Name of this configuration [vmnix]:

Now allocating memory to the Service Console. The memory that is not allocated
to the Service Console will be allocated to the VMs. Be sure to indicate a
number smaller than you actually have on your system or you may not be able to
run any VMs.

Your system is reporting approximately 2,559 MB of memory.
Note: you cannot allocate more than 800 MB to the Service Console.

If you plan to run 4 virtual machines or fewer, 128 MB should suffice.
Set this to 192 MB for 8 virtual machines, 272 MB for 16 virtual
machines, 384 MB for 32 virtual machines, or 512 MB for more than
32 virtual machines.

Amount of memory (in megs) to allocate to the Service Console [384]:
Enable hyperthreading? (Only works with 8 or less physical processors) [y]:
```

```
Now to divide the SCSI controllers and network adapters. In this process,
you will select whether to give the previous devices to the Service Console or
the virtual machines that you will be running. In order to ensure optimal
performance for your virtual machines, we suggest you allocate as many devices
as possible to the virtual machines, leaving only the ones necessary to
configure and maintain the Service Console.

Type 'c' to allocate to the Service Console or 'v' to allocate to the virtual
machines, or 's' to share between both.

RAID storage controller Compaq Computer Corporation Smart Array 5i/532 (rev 01)
   running on bus 1 device
 3 function 0 (cciss.o) [vmhba4] [s]:
Ethernet controller BROADCOM Corporation NetXtreme BCM5703 Gigabit Ethernet (rev
   02) running on bus 2
device 1 function 0 (bcm5700.o) [] [c]:
Ethernet controller BROADCOM Corporation NetXtreme BCM5703 Gigabit Ethernet (rev
   02) running on bus 2
device 2 function 0 (bcm5700.o) [vmnic0] [v]:
Fiber storage controller QLogic Corp QLA231x/2340 (rev 02) running on bus 6 device
   1 (qla2300_604.o) [v
mhba2] [v]:
Ethernet controller Intel Corporation 82557 [Ethernet Pro 100] (rev 08) running
   on bus 6 device 2 function
 0 (e100.o) [vmnic1] [v]:

Configuration complete.
Commit changes (y/n)? [y]:
```

Rescanning the SCSI bus

For use in the VMkernel itself, rather than using the standard Linux `/proc` file system to obtain information on the HBAs, refer to the `/proc/vmware/scsi` directory. This directory contains entries for each HBA in the system allocated to the Virtual Machines. Each HBA is denoted by `vmhba`*N* (where *N* indicates the file for each adapter owned by the Virtual Machines).

You may confirm that the QLogic or Emulex driver is loaded in the VMkernel with the **vmkload_mod** command. For example, the following command may be issued on this host to list the QLogic driver:

```
root@l82bi199 /]# vmkload_mod -l
Name R/O Addr Length R/W Addr Length ID Loaded
vmklinux 0x4dd000 0xf000 0x12414a0 0x53000 1 Yes
nfshaper 0x4ec000 0x1000 0x1297f28 0x1000 2 Yes
qla2300_604 0x4ed000 0x19000 0x1298f30 0x22000 3 Yes
bond 0x506000 0x2000 0x150e1d8 0x2000 4 Yes [
```

Assuming that the devices have been allocated to the VMware host and that the VMkernel has been configured appropriately, the **Storage Management** dialog in the MUI reports the devices that are allocated to the host.

You may also view the device listing at the command line. This may be done with the following commands:

- **more /proc/scsi/qla2300/N**

  where *N* is the sequential value of each QLogic HBA installed in the system, beginning with the number after the last host adapter number entry in the file. (The first host adapter number entry begins with zero.) This is the same as it would be for standard Linux.

- **ls –la /proc/vmware/scsi/vmhbaN**

  where *N* indicates the file for each adapter owned by the VMkernel and, subsequently, the Virtual Machines

After the host is added to the Storage Group, the SCSI bus needs to be rescanned.

Rescanning the SCSI bus may be accomplished in one of the following ways.

For VMware ESX Server v2.1.x:

1. It is recommended that the MUI be used to rescan the SCSI bus. Simply go to the **Storage Management** dialog and click on the **Rescan SAN** option to rescan the SCSI bus.
2. You may reboot the host, but since this is disruptive, it is not recommended.
3. At the command line, rescanning the SCSI bus is a two-step process for QLogic HBAs. First, you will need to force the QLogic HBAs to rescan the SCSI bus first and secondly; the VMkernel will need to perform the rescan.

   Note: Please note that this step is required both when using the command line and when using the MUI.

   a. Issue the following command first:

      echo "scsi-qlascan" > /proc/scsi/qla2300/<hba_id>

      For example:

      ```
      echo "scsi-qlascan" > /proc/scsi/qla2300/2
      ```

   b. Next, you will need to have the VMkernel rescan the SCSI bus.

      – If you are using the command line, issue the following command:

        vmkfstools -s *<vmware_hba_id>*

        For example:

        ```
        vmkfstools -s vmhba0
        ```

      – If you are using the MUI, then go to the **Storage Management** dialog and click on the **Rescan SAN** option to rescan the SCSI bus.

4. At the command line, you only need to rescan the VMkernel using Emulex HBAs.

   • If you are using the command line, issue the following command:

     vmkfstools -s *<vmware_hba_id>*

For example:

```
vmkfstools -s vmhba0
```

- If you are using the MUI, then go to the **Storage Management** dialog and click on the **Rescan SAN** option to rescan the SCSI bus.

5. For either QLogic or Emulex HBAs, you may also remove and reinsert the driver from the VMkernel to initiate a rescan of the SCSI bus.

   - For QLogic:

```
vmkload_mod -l
vmkload_mod -u qla2300_604
vmkload_mod /usr/lib/vmware/vmkmod/qla2300_604.o vmhba
```

   - For Emulex:

```
vmkload_mod -l
vmkload_mod -u lpfcdd
vmkload_mod /usr/lib/vmware/vmkmod/lpfcdd.o vmhba
```

Note: It is recommended in the case of using either the QLogic or Emulex HBAs that the **Rescan SAN** option in the MUI be used to rescan the SCSI bus.

For VMware ESX Server v2.5.x:

1. It is recommended that the MUI be used to rescan the SCSI bus. Simply go to the **Storage Management** dialog, and click on the **Rescan SAN** option to rescan the SCSI bus. This process calls the `cos-rescan.sh`.

2. If the command line is preferred, you may use the `cos-rescan.sh` script to initiate the rescan.

   For example:

```
[root@l82bi199 /]# cos-rescan.sh vmhba2
Rescanning vmhba2...done.
On scsi1, removing:
On scsi1, adding:
```

3. You may reboot the host, but since this is disruptive, it is not recommended.

4. If booting from an internal hard drive rather than from the SAN, you may also remove and reinsert the driver from the VMkernel to initiate a rescan of the SCSI bus. This is pertinent for both QLogic and Emulex HBAs.

   - For QLogic:

```
vmkload_mod -l
vmkload_mod -u qla2300_604
vmkload_mod /usr/lib/vmware/vmkmod/qla2300_604.o vmhba
```

   - For Emulex:

```
vmkload_mod -l
vmkload_mod -u lpfcdd
vmkload_mod /usr/lib/vmware/vmkmod/lpfcdd.o vmhba
```

In VMware ESX Server v2.5.x, the **vmkfstools** command should not be used to rescan the SCSI bus; otherwise, the Service Console will not see or report the changes made.

Persistent binding

EMC recommends, but does not require, that persistent binding be enabled in VMware ESX Server v2.x. When persistent binding is enabled, the target IDs are assigned to specific storage array devices. This assignment persists through reboots so the association remains the same. This is particularly helpful when using raw devices and RDMS (Raw Device Mappings).

Persistent binding may be enabled or disabled using the MUI or the command line.

If using the MUI, select the **Options** tab, and then select the **Storage Management** tab. Under Storage Management, there are three options: **Disks and LUNs**, **Failover Paths**, and **Adapter Bindings**. To view the current information for your system, select the **Adapter Bindings** option.

An example of the output from the **Adapter Bindings** dialog is shown in Figure 5:

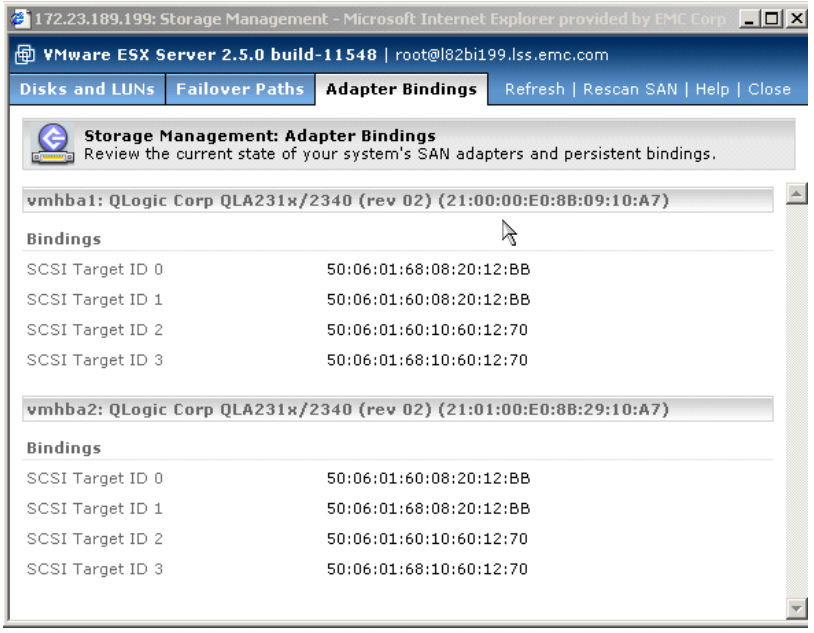

Figure 5 Adapter Bindings dialog example

To view and reset the persistent binding from the command line, follow the steps below.

1. Log in as `root`.
2. To view the current persistent binding entries for the HBAs in the system, look at the `pbindings` file.

   An example of what the `pbindings` file may look like is as follows:

```
[root@l82bi199 /]# more /etc/vmware/pbindings
6.1.1.qla2300_607 = "scsi-qla1-tgt-0-di-0-port=50060160082012bb;scsi-qla1-tgt-1-
di-0-port=50060168082012bb;scsi-qla1-tgt-2-di-0-port=5006016010601270;scsi-qla1-
tgt-3-di-0-port=5006016810601270;"
6.1.0.qla2300_607 = "scsi-qla0-tgt-0-di-0-port=50060168082012bb;scsi-qla0-tgt-1-
di-0-port=50060160082012bb;scsi-qla0-tgt-2-di-0-port=5006016010601270;scsi-qla0-
tgt-3-di-0-port=5006016810601270;"
```

3. To remove the existing persistent binding settings, either the `pbindings` file may be deleted or the `pbind.pl` script may be run. The `pbind.pl` is located in the `/usr/sbin` directory and may be run as root from any directory location.

   To delete the file, use the following command:

   ```
   rm –f /etc/vmware/pbindings
   ```

   To remove the settings using the `pbind.pl` script, use the –D switch. This switch deletes the persistent binding information for all of the adapters within the system.

   ```
   pbind.pl -D
   ```

   For example:

   ```
   [root@l82bi199 /]# pbind.pl -D
   Deleting ALL bindings
   ```

4. Reboot the host.
5. After the system reboots successfully, log in and again run the `pbind.pl` script. This time use the –A switch which will add the persistent binding information for all of the adapters within the system.

   ```
   pbind.pl –A
   ```

For example:

```
[root@l82bi199 /]# pbind.pl -A
+----------------------------------------+
| Setting up the following bindings... |
+--------+---------+----------------------+
| Host | Target | World Wide Name |
| 01 | 000 | 50060160082012bb |
| 01 | 001 | 50060168082012bb |
| 01 | 002 | 5006016010601270 |
| 01 | 003 | 5006016810601270 |
| 00 | 000 | 50060168082012bb |
| 00 | 001 | 50060160082012bb |
| 00 | 002 | 5006016010601270 |
| 00 | 003 | 5006016810601270 |
+----------------------------------------+
```

6. Edit the `vmware` initialization script. This script is run when the system is booting up and starts and manages the services needed by the VMkernel.

   ```
   vi /etc/init.d/vmware
   ```

7. Search for the line containing the text `save_san_persistent_bindings`. The code should be located approximately around line 896 in the file.

```
896 # Save the SAN persistent bindings for the next boot
897 save_san_persistent_bindings
```

8. Comment out the line `save_san_persistent_bindings`.

   This is only a temporary change. When this line is commented out, the adapter to target bindings will not be saved and therefore, will not persist through reboots.

```
896 # Save the SAN persistent bindings for the next boot
897 #save_san_persistent_bindings
```

9. Save the file and reboot the host again.
10. Once the system is rebooted, the persistent binding line will need to be uncommented.

    Change to the `/etc/init.d/vmware` directory again:

    ```
    vi /etc/init.d/vmware
    ```

    Uncomment the line for `save_san_persistent_bindings` and save the file.

11. Now run `wwpn.pl` to associate the targets, HBA WWPN, and `vmhba`.

```
[root@l82bi199 /]# wwpn.pl -v
WWPN 1.02 Copyright VMware 2003
Display WW port names and VMHBA information for fibre channel adapters
For each vmhba here are the corresponding Qlogic and Emulex WW Port Names
 Adapter WWPN PCI (decimal)
vmhba1: 210000e08b0910a7 (Qlogic) 6:1:0 /proc/scsi/qla2300/0
vmhba1:0: 50060168082012bb scsi-qla0-port-0=0000000000000000:50060168082012bb;
vmhba1:1: 50060160082012bb scsi-qla0-port-1=0000000000000000:50060160082012bb;
vmhba1:2: 5006016010601270 scsi-qla0-port-2=0000000000000000:5006016010601270;
vmhba1:3: 5006016810601270 scsi-qla0-port-3=0000000000000000:5006016810601270;
vmhba2: 210100e08b2910a7 (Qlogic) 6:1:1 /proc/scsi/qla2300/1
vmhba2:0: 50060160082012bb scsi-qla1-port-0=0000000000000000:50060160082012bb;
vmhba2:1: 50060168082012bb scsi-qla1-port-1=0000000000000000:50060168082012bb;
vmhba2:2: 5006016010601270 scsi-qla1-port-2=0000000000000000:5006016010601270;
vmhba2:3: 5006016810601270 scsi-qla1-port-3=0000000000000000:5006016810601270;
```

12. The persistent binding is now reset to the most recent information.

LUN spanning

When a volume is spanned, the volume is extended to include multiple VMFS-2 disks or partitions. A physical extent is defined by the VMFS-2 disks or partitions that compose a logical spanned VMFS-2 volume.

There are some restrictions to using spanned volumes. They are:

- Spanning volumes is supported only with VMFS-2 volumes, not with VMFS-1 volumes.
- The maximum number of files may not be changed.
- A spanned volume may not be removed.
- Data on the physical extents are lost when they are added to the VMFS-2 spanned volume. Therefore, it is strongly recommended that only newly created partitions or newly added devices are used for spanning volumes.
- A VMFS-2 volume may span up to a maximum of 32 LUNs.

To span a volume using the MUI, select the **Options** tab then select Storage Management. Select the **Disks and LUNs** category. In the **Disks and LUNs** category, you are provided with a list of the devices that are assigned to the host. Select the one that you want to expand. The configuration window for the device you chose will be shown on your screen and will provide you with the specifics on the selected device plus the devices that may be added to the selected device to create a spanned VMFS-2 volume.

See Figure 6 on page 45 for an example using the device `vmhba1:2:13`.

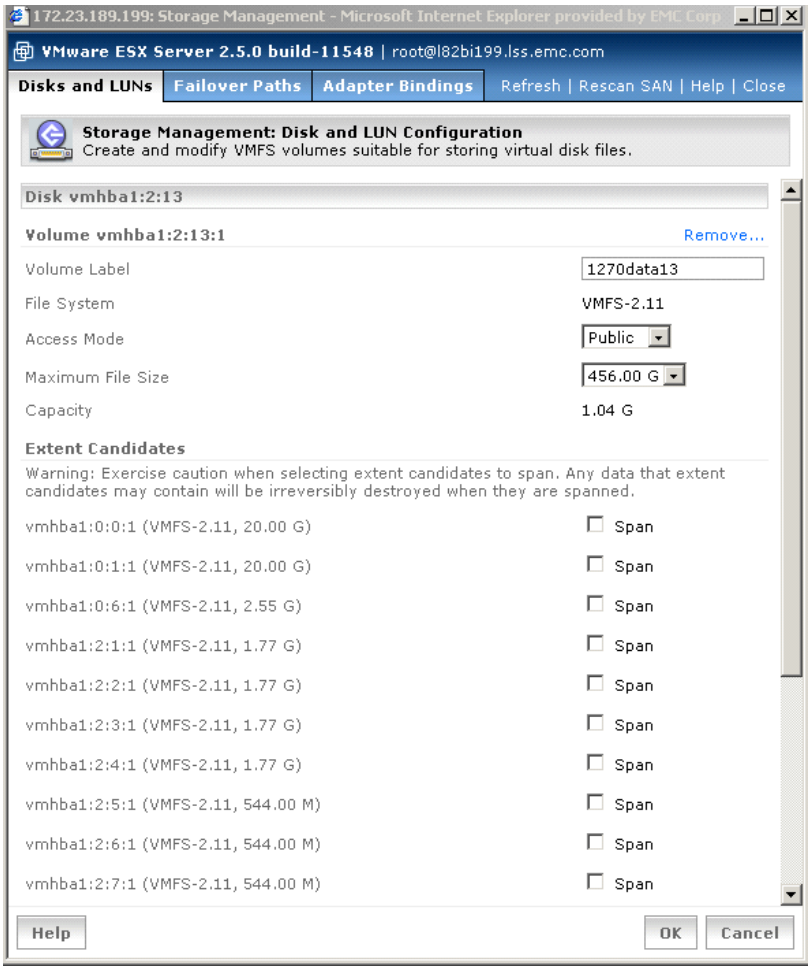

Figure 6 Example using the device vmhba1:2:13

After you have spanned a VMFS-2 volume, you may view the changes immediately because the `/vmfs` directory is updated dynamically. The command `vmkfstools` should be used to list the information on the particular device and to provide the spanning and physical extent information.

For example:

```
[root@l82bi199 /]# cd /vmfs
[root@l82bi199 /]# vmhba1:0:0:1 vmhba1:2:11:1 vmhba1:2:13:1 vmhba1:2:3:1
```

```
[root@l82bi199 /]# vmkfstools --list vmhba1:2:13:1
Name: 1270data13 (public)
Capacity 2826683392 (2679 file blocks * 1048576), 1701838848 (1623 blocks) avail
Permission Uid Gid Attr Bytes (Blocks) Last Modified Filename
rw------- 0 0 disk 1107296256 ( 1056) Aug 26 06:28 w2k3_data4.dsk

[root@l82bi199 /]# vmkfstools -P vmhba1:2:13:1
vmhba1:2:13:1 is a VMFS-2.11 volume spanning 4 physical extents.
Volume label (if any): 1270data13
UUID (if any): 409bb858-be8577b8-ccf1-0002b3248a69
Physical Extents:
 vmhba1:2:13:1
 vmhba1:2:5:1
 vmhba1:2:6:1
 vmhba1:2:7:1
```

To view the same volume information from the MUI, go to the **Options** tab and select **Storage Management**. Select the **Disks and LUNs** category and scroll to view the VMFS-2 volume that was spanned. Please see the reference to Disk `vmhba1:2:13` in Figure 7 for an example.

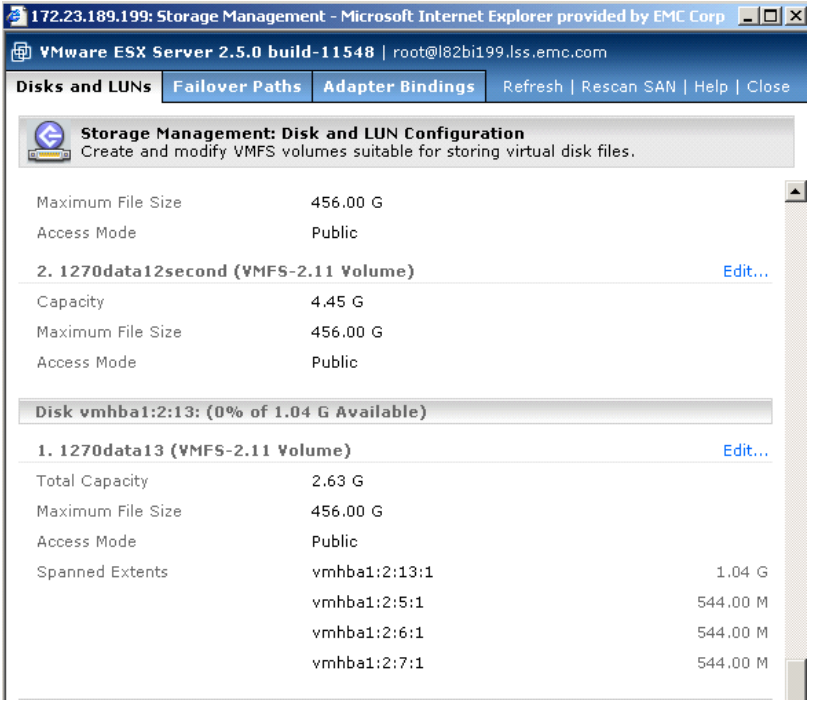

Figure 7 Disks and LUNs example

Additional notes

- VMware ESX Server is supported on both CLARiiON and Symmetrix Storage Arrays.
- VMware ESX Server hosts may be attached to CLARiiON and Symmetrix Storage Arrays simultaneously. The failover policy should be correctly selected to be *Most Recently Used* (MRU) for the CLARiiON devices and *Fixed* for the Symmetrix devices
- EMC software will function on neither the Service Console nor the Virtual Machines for VMware ESX Server versions v2.0.x and v2.1.x as the currently-release versions of the CLARiiON-based software applications do not include support for these VMware ESX Server versions.

  This has changed with the introduction of VMware ESX Server v2.5.x.
- Versions of VMware ESX Server v2.1.x and prior do not support accessing SnapView™, MirrorView™, or SAN Copy™ LUNs. However, with VMware ESX Server v2.5.x, support for accessing SnapView, MirrorView, or SAN Copy LUNs exists.
- With VMware ESX Server versions prior to v2.5.x, the host needs to be registered manually on the CLARiiON Storage Array as the Naviagent/CLI is not supported.
- With the introduction of VMware ESX Server v2.5.x, the Naviagent/CLI is now supported and should be installed on the VMware ESX Server itself, rather than on the Virtual Machines.
- The restriction stating that VMware ESX Server v2.1.x hosts must be offline for CLARiiON-licensed (FLARE®) upgrades and Storage Processor replacements has been removed. The native failover functionality within VMware ESX Server itself has been successfully qualified and provides the necessary path failover. NDUs and SP replacements may be performed while the VMware ESX Server is online.
- PowerPath® is not available for the VMkernel. The failover mechanism supported with VMware ESX Server v2.x and EMC Storage Arrays is native to the VMware ESX Server v2.x kernel itself. The native failover functionality performs failover only, not load balancing.

- For the Virtual Machines, it is recommended, but not required, that the boot LUN(s) be in a CLARiiON RAID group that is separate from the RAID group for the data volumes.
- Prior to VMware ESX Server v2.5.x, the VMware Server itself was not support booting from the array. VMware ESX Server v2.5.x adds support for booting from the EMC Storage Arrays.
- When booting VMware ESX Server v2.5.x from an EMC Storage Array, it is recommended that the boot LUN be in a RAID group separate from the data volumes.
- When booting the VMware ESX Server from an internal drive, the VMware Core Dump and Swap Files should be created on internal drives, rather than SAN-attached devices.
- It is highly recommended that the vmware-tools be installed on each Virtual Machine. The use of vmware-tools will save system resources and will provide better memory management. The `vmware-tools` will also provide a VGS driver and a heartbeat mechanism for the Virtual Machines to communicate with the VMk.
- Booting the virtual machines from SAN is recommended, irrespective of where the ESX Server itself is booting from. The virtual machines have the ESX Server vmkernel virtualization layer between their virtual disks and the SAN LUNs, and the BusLogic/LSI Logic emulation layer. These two layers protect the virtual machines from SAN LUN failures. The only scenario where the virtual machine (which is booting from SAN) will not be protected is during an “all paths down” condition. This means that the ESX Server has lost all paths to a SAN LUN simultaneously. In this particular case, the virtual machine will hang for a certain period of time and may need to be rebooted when the SAN LUNs become available again. Please note that the VMware ESX Server native failover policy must be configured correctly depending on the storage array being used when ESX Server and/or the virtual machines are booting from and/or accessing data from the SAN.

3

VMware ESX Server v2.x and Symmetrix Fibre Channel Environment

This chapter lists Symmetrix Fibre Channel support information specific to the VMware ESX Server environment.

VMware/Symmetrix environment

This section provides Symmetrix Fibre Channel support information specific to the VMware ESX Server environment.

Host and hardware connectivity

Refer to the *EMC Support Matrix* or contact your EMC representative for the latest information on qualified hosts, host bus adapters, and connectivity equipment.

Note: EMC does not support mixing HBAs in the same system, including HBA models from the same vendor.

Symmetrix connectivity

The Symmetrix system is configured by an EMC Customer Engineer via the Symmetrix service processor.

The EMC Customer Engineer (CE) should contact the EMC Configuration Specialist for updated online information. This information is necessary to configure the Symmetrix system to support the customer's host environment.

When attaching a VMware ESX Server 2.x host to a Symmetrix storage array, the following FA settings must be enabled: C, VCM, SC3, and UWN.

ESX Server 2.x does not require the SPC-2 bit to be set but will not behave differently if the SPC-2 bit is set.

After the EMC CE has assigned target IDs and LUNs and brought the Symmetrix channel and disk directors online, reboot the network operating systems, and go into the configuration program.

Note: All qualified HBAs are listed in the *EMC Support Matrix*.

Note that the VMware ESX Server installer will recognize LUNs 25MB or less as management LUNs. This includes any gatekeepers assigned to the VMware host via Solutions Enabler.

A couple of possible configuration scenarios are described in "Example 1" on page 51 and "Example 2" on page 51.

Example 1

In this example as shown in Figure 8, one host with two HBAs is attached to one Symmetrix array using two separate switches. Two FA ports on each of two FAs within the array are being used. HBA1 is zoned to 15D0 and to 16C1. HBA2 is zoned to 15D1 and to 16C0. The zones should be composed of a single initiator and a single target so they would be created with one HBA and on FA port. Two HBAs with two paths each to the array totals four paths and if using single initiator / single target zoning, there are 4 zones.

In this particular example, two switches are used. Using only one switch is supported, but such a configuration would lack redundancy. Preferably, a minimum of two switches should be used to add another level of redundancy. Alternatively, for additional redundancy, two separate fabrics could be utilized.

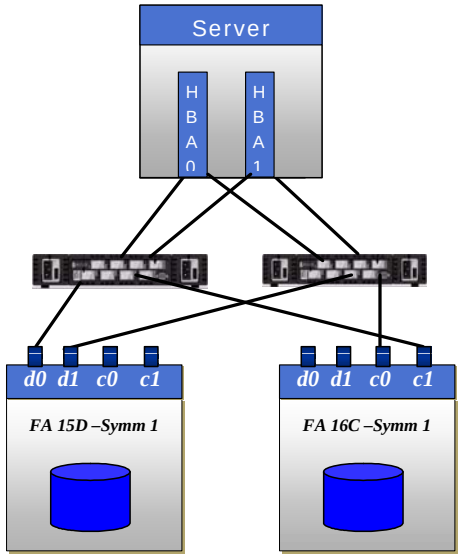

Figure 8 **One host, two switches, and one Symmetrix array**

Example 2

In this example, as shown in Figure 9 on page 52, one host with two HBAs is attached using a two-switch fabric to four Symmetrx arrays FA ports. In this configuration, the zones are created with one HBA and one FA port. That is,

- HBA1 is zoned to one 15D0 port on each of the four arrays.
- HBA2 is zoned to one 16C1 port on each of the four arrays.

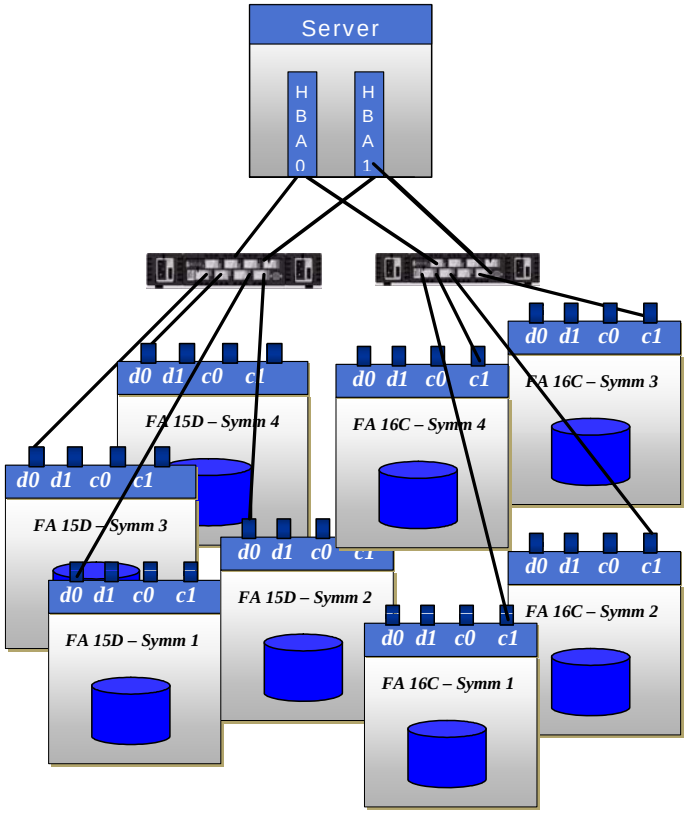

Figure 9 **One host, two switches, and four Symmetrix arrays**

Note: All qualified HBAs are listed in the *EMC Support Matrix*.

When assigning Symmetrix LUNs to a VMware ESX Server host, the LUNs should be assigned to the host across both FAs since the Symmetrix is an active / active array.

Partition alignment

Modern hard disk systems use the logical block address (LBA) to position the head. This is true for both SCSI and IDE disks. However, older disks systems used a different addressing scheme called CHS (cylinder, head, and sectors) to describe the geometry of the drive. Hard disks using this addressing scheme expect three numbers to position the disk head accurately. Various specifications for IDE and BIOS have evolved over the years to accommodate larger disk storage capacities. These standards provide various combinations for the maximum value for CHS. These ranges from 1024-65536 for cylinders, 16-255 for heads and 1-255 sectors per track.

The BIOS of all x86 based computers still provide support for the CHS addressing. In addition, the BIOS also provide mechanism that maps LBA addresses to CHS addresses using the geometry information provided by the disks. Modern operating systems such as Linux and VMware ESX Server do not normally use the mapping information provided by the BIOS to access the disk. However, these operating systems need the geometry information when communicating with the BIOS or with other operating systems that use CHS mapping information, such as DOS or Microsoft Windows.

The first cylinder of all hard disks contains a reserved area called the master boot record (MBR). When an IBM compatible system is booted, the BIOS reads the master boot record (MBR) from the first available disk. The bootstrap loader code found at this location is used to load the operating system. The MBR also contains critical partition table information for four entries that describes the location of the primary data partitions on the disk. The partition table structure looks like:

```
struct partition {
char active;    /* 0x80: bootable, 0: not bootable */
char begin[3];  /* CHS for first sector */
char type;
char end[3];    /* CHS for last sector */
int start;      /* 32 bit sector number (counting from 0) */
int length;     /* 32 bit number of sectors */
};
```

The information in the structure is redundant – the location of a partition is given both by the 24-bit begin and end fields, and by the 32-bit start and length fields. Only one of the two sets of fields is needed to describe the location of the partition. VMware ESX Server uses the start and length fields of the partition table structure. By

default, the VMware ESX Server will create first data partition starting at the first available LBA after the area reserved for MBR.

EMC Symmetrix storage arrays present disk geometry that has 32 sectors per track. Therefore, when EMC Symmetrix storage arrays are used with VMware ESX Server, the starting address for the first data partition will be located at LBA address 32 (i.e., 16,384 bytes). The EMC Symmetrix storage arrays utilize either 32 KB or 64 KB track size. Therefore, utilizing default configuration for disk partitions will result in inefficient utilization of storage components.

Prior experience with misaligned Windows partitions and file systems has shown as much as 20% to 30% degradation in performance. Similar performance impact can be expected for a VMware ESX Server environment. Aligning the data partitions on 64 KB boundary will result in positive improvements in overall IO response time experienced by all hosts connected to the shared storage array.

Partition alignment for Virtual Machines using VMFS volumes

EMC Symmetrix storage arrays manage the data on physical disk using a logical construct known as track. A track can be either 32 KB or 64 KB in size. A hypervolume can thus be viewed as a collection of 32 KB tracks. A data partition created by VMware ESX Server consists of all or subset of the tracks that represent the hypervolume. This is pictorially represented in Figure 10 on page 55.

The figure also shows the layout of the data partition that is created by default by VMware ESX Servers.

Note: The hypervolume appears as a physical disk to the VMware ESX Server.

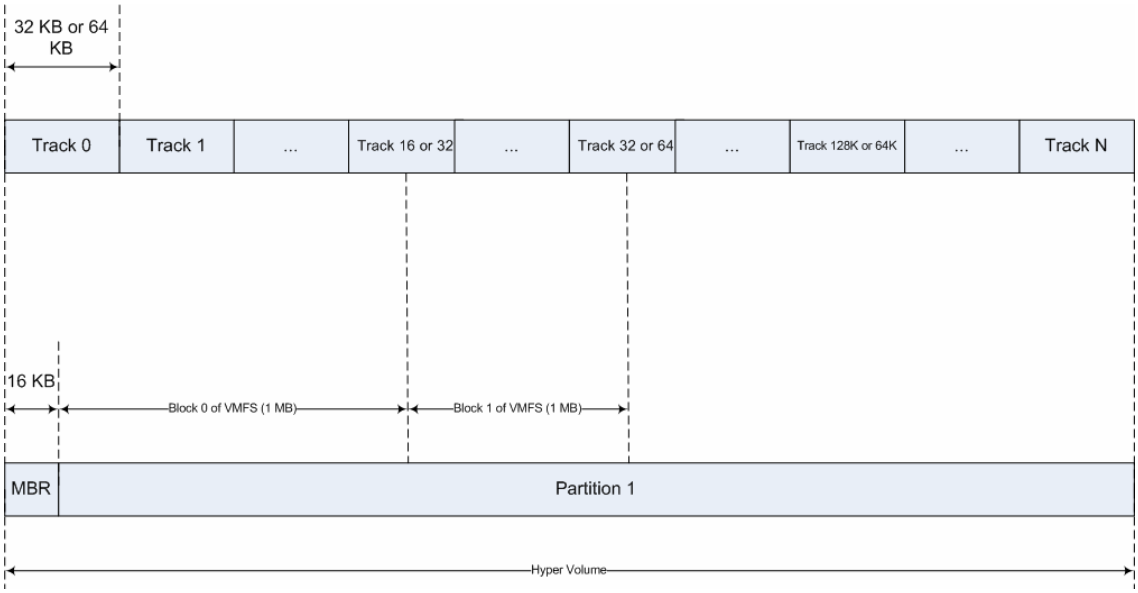

Figure 10 **Pictorial representation: Hypervolume to Symmetrix track mapping**

It is evident from Figure 10 that the data partition if created using the default settings of VMware ESX Server will not be aligned on track boundary. However, unlike the behavior exhibited by Microsoft Windows hosts, the partition is not sector misaligned — a 4K IO at block 0 of the data partition will not span multiple tracks.

Note: Please refer to the white paper *Windows File-System Alignment*, available on Powerlink®, for further details on sector misalignment problem in Microsoft Windows environment. It is important to note that track misalignment does not cause as severe a performance impact as sector misalignment.

VMware ESX Server creates VMware filesystem on the data partition using one MB block size. Since the block size is a multiple of track size, file allocations will be in even multiple of the track size. This is also shown in Figure 10. Thus, virtual disks created on the partitions normally created by VMware ESX Server will always be track misaligned.

A virtual disk created on VMware filesystem will be presented to the guest operating system with geometry of 63 sectors. The Phoenix BIOS used by the virtual machine will reserve one track of the virtual

disk for storing master boot record. The data partition created on the virtual disk will start at sector 64. The layout of the virtual disk and Figure 11. It can be seen from the figure that the guest operating system layout will exacerbate the track misalignment problem created by the VMware ESX Server — the IOs generated by the guest operating system will be sector misaligned.

Important EMC recommends a two step process to address the sector misalignment issue. Aligning both the VMware filesystem and the virtual disk on track boundary will ensure the optimal performance from the storage subsystem. Furthermore, EMC recommends aligning the partitions on 64 KB boundaries. This will ensure optimum performance on all EMC storage platforms.

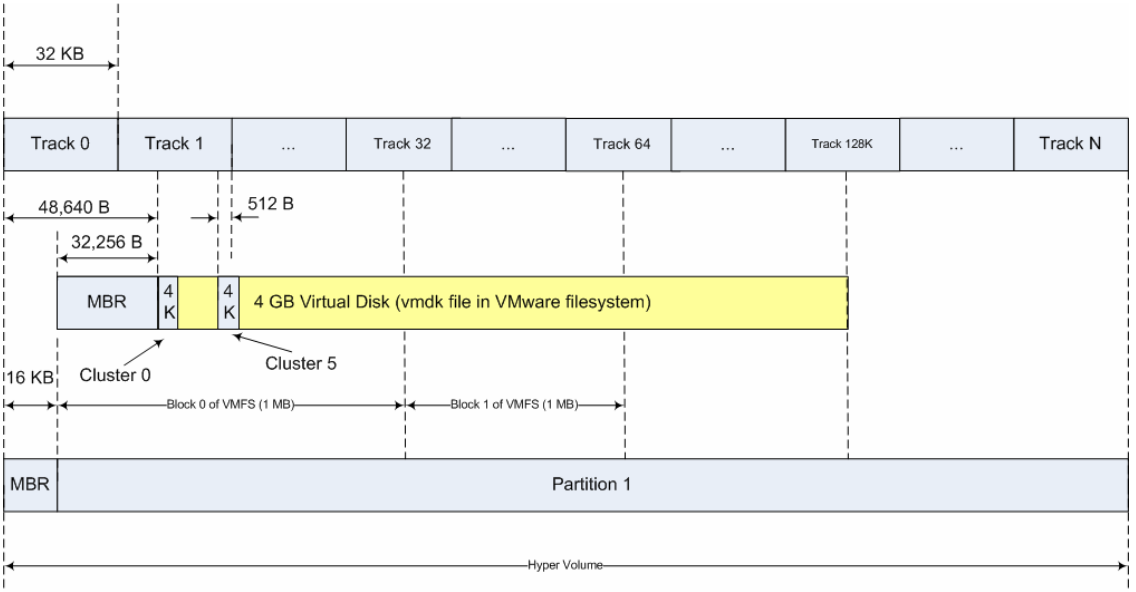

Figure 11 Track misalignment: VMware ESX Servers using VMware file system

Creating track aligned VMware filesystem

To align the VMware filesystem on 64KB track boundary the following steps need to be implemented:

1. Map the canonical name of the device on which the VMware filesystem will be created to the device name assigned by the service console. This can be achieved by executing the command, **vmkpcidivy -q vmhba_devs**. A sample output from this command is shown in Figure 12.

   In the following steps, the device name **/dev/sd<x>** will be used to represent the device name obtained from the aforementioned mapping process.

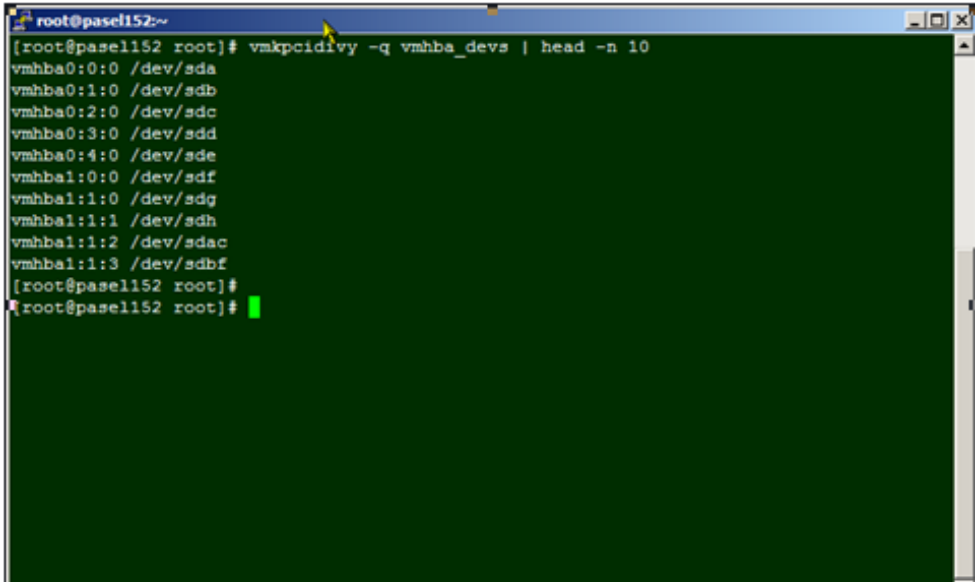

Figure 12 **Sample output from vmkpcidivy**

2. On the service console, execute **fdisk -l /dev/sd<x>**. There should be no existing partitions on this disk. If there are any existing partitions, **ensure they are not in use** and delete them.

CAUTION

Extreme care should be exercised when deleting existing partitions. Deleting partitions can result in potential loss of data.

3. Once there are no partitions on the device under consideration, execute **fdisk /dev/sd<x>**.
4. Type **n** to create a new partition.
5. Type **p** to create a primary partition.
6. Type **1** to create partition 1.
7. Select the default values offered by fdisk to create the partition. This will result in a partition that spans the whole disk.
8. Type **x** to enter expert mode. The alignment of partition can be performed in this mode. A screenshot displaying the steps up to this point is shown in Figure 13.

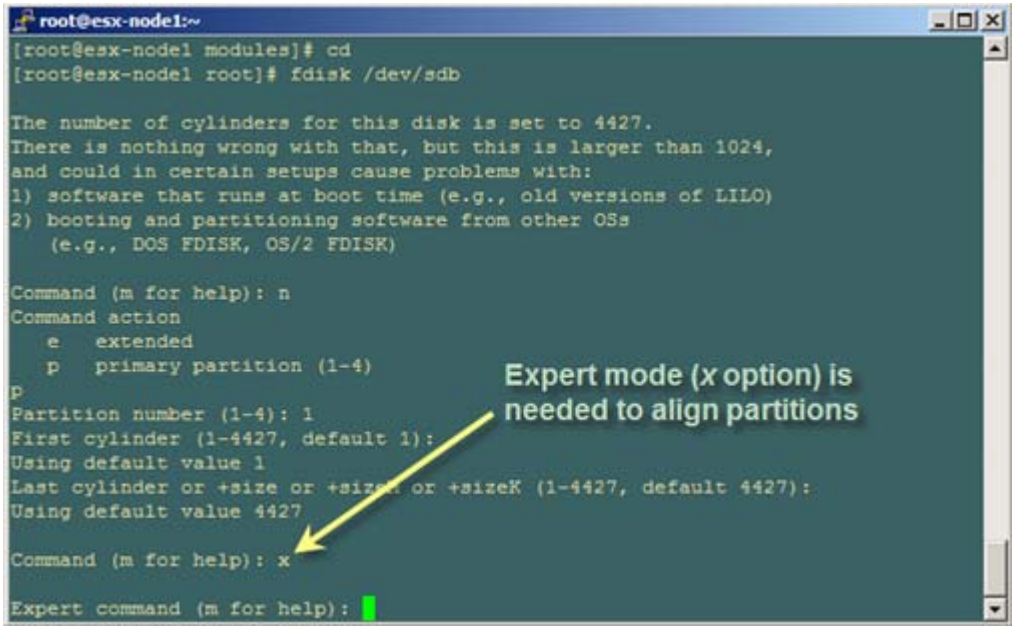

Figure 13 Using fdisk to create track-aligned partition

9. Type **b** to specify the starting block number for partition.
10. fdisk will ask the partition number for which the starting block number will be specified. Type **1** to select partition number 1.

11. The starting block number for partition number 1 can now be specified. Type **128** to specify a starting block number of 128. Since each block is 512 bytes in size, this corresponds to a starting offset of 64 KB for partition number 1. This is depicted in Figure 14.

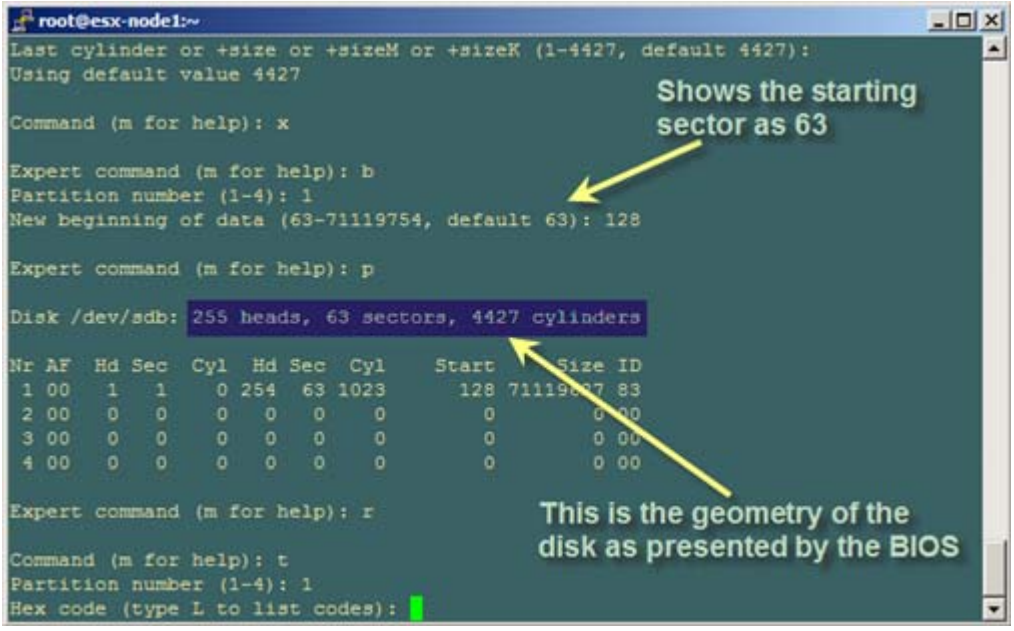

Figure 14 **Aligning partition on VMware ESX Server**

12. Type **r** to return to the main menu of fdisk.
13. The VMware ESX Server creates VMware filesystem on VMFS volume. The VMware ESX Server recognizes a VMFS volume if the partition type is set to **fb**. Type **t** to change partition type.

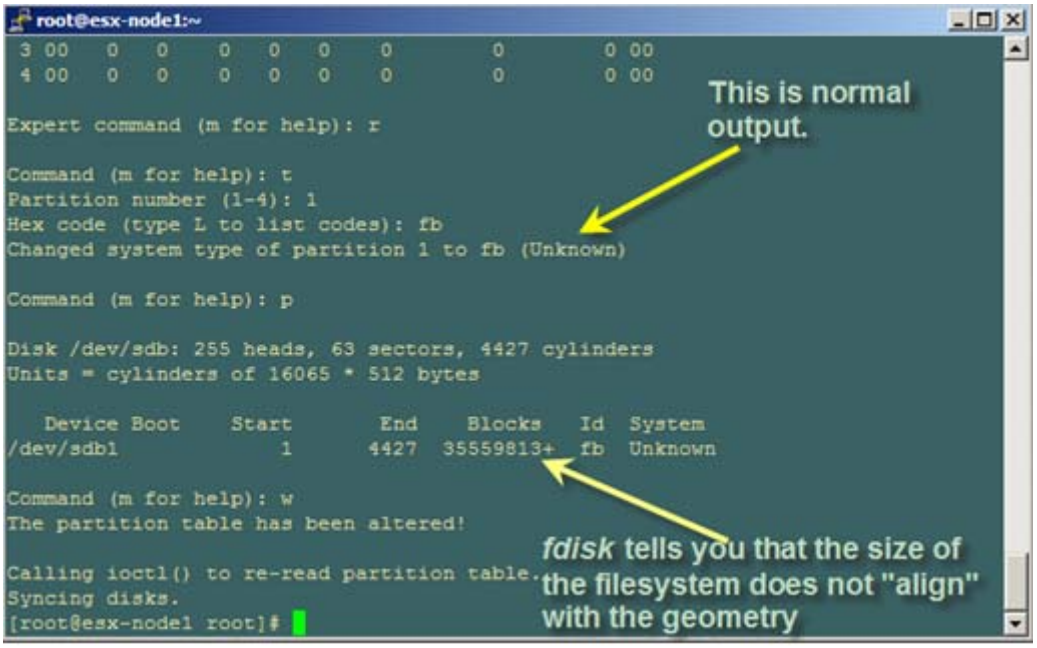

Figure 15 **Setting partition type for the aligned primary partition**

14. Type **fb** to set the partition type to VMFS volume. fdisk will state that the partition is being set to an unknown type. This is normal behavior since the service console is unaware of the partition type to specify VMFS volumes.

15. Type **w** to write the label and partition information to disk, and exit fdisk. Figure 15 shows the output from commands listed in steps 12-15.

A VMware filesystem can now be created on the aligned VMware volume using either command line utility (**vmkfstools**) or the MUI. The storage management screen of the MUI will show the aligned partition as an "Unformatted VMFS volume" on doing a rescan. This is shown in Figure 16 on page 61.

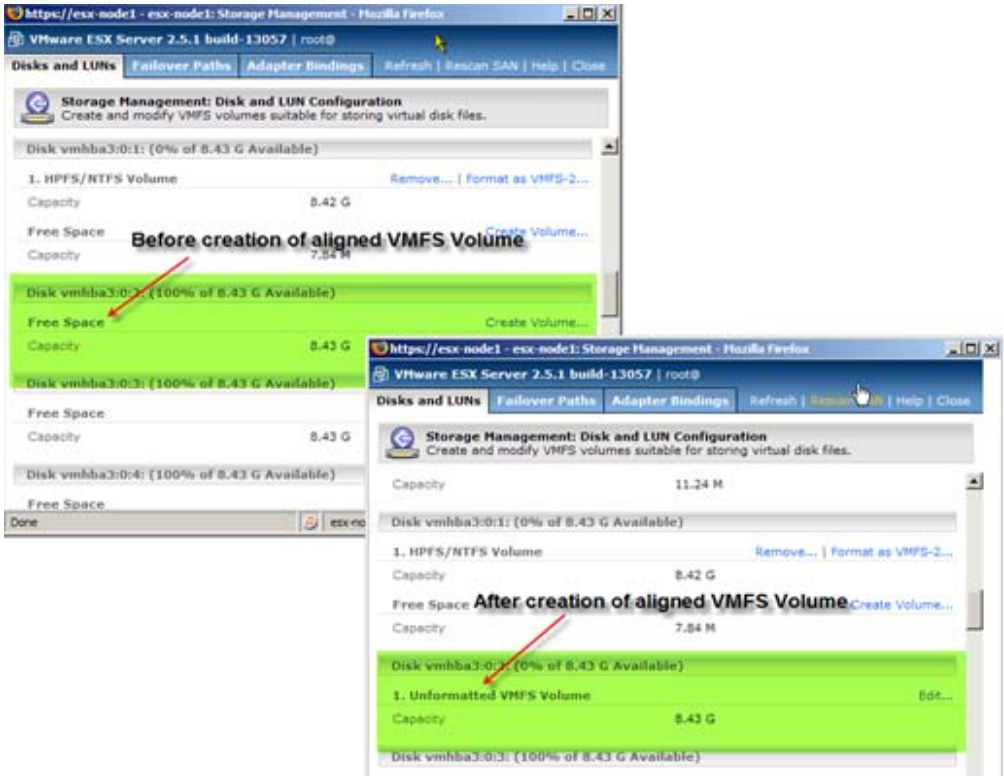

Figure 16 View of aligned partitions in the MUI

Creating track-aligned virtual disks

The virtual disks created on VMware filesystem are presented to the guest OS with a geometry of 63 sectors. It is critical to align the virtual disks on track boundary in addition to aligning the VMware filesystem. The process of aligning virtual disks should be performed in the virtual machine, and is not different from the one used for physical servers.

The alignment process for Microsoft Windows servers has been addressed extensively before. Readers should follow the process described in the white paper, *Windows File-System Alignment* to align disks presented to virtual machines running Microsoft Windows operating system.

The track alignment procedure for creating aligned VMFS volume should be used to align disks in virtual machines with Linux as the

guest operating system. Steps 13-15 should not be executed when aligning virtual disk in the Linux guest operating system.

An aligned VMware filesystem and virtual disks will ensure optimum use of the EMC Symmetrix storage system. Figure 17 shows the layout of the track aligned virtual disk and VMware filesystem and its relation to the EMC Symmetrix track table. Comparison of Figure 17 and Figure 11 on page 56 clearly shows that the alignment process described in the previous section removes the sector misalignment problem described earlier.

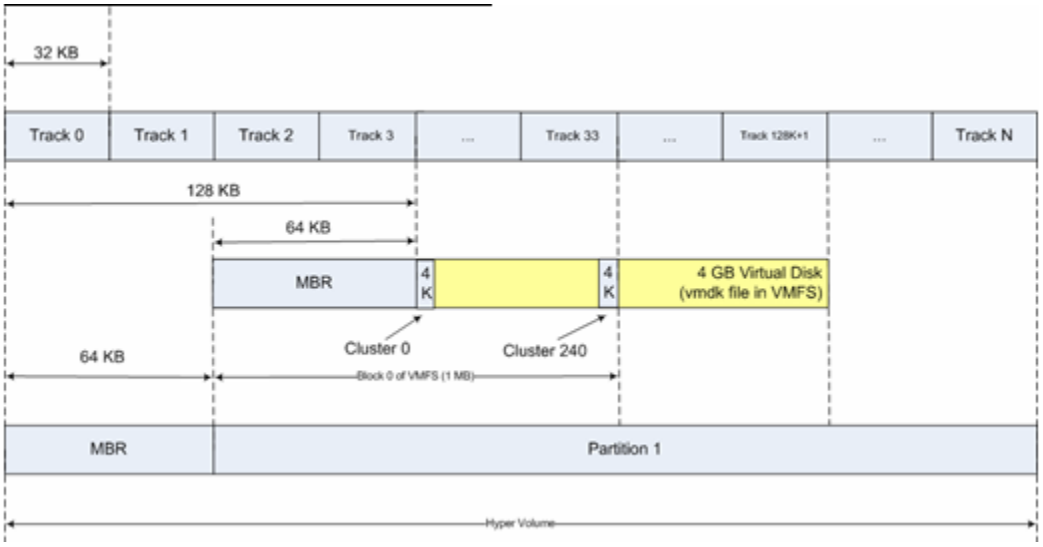

Figure 17 Track aligned VMware file system and virtual disks

Partition alignment for Virtual Machines using raw disk mapping (RDM)

EMC Symmetrix devices that are accessed by virtual machines using raw disk mapping (RDM) do not contain the VMware filesystem. In this configuration, the alignment problem is the same as that seen on physical servers. The same process deployed for aligning partitions on physical servers needs to be used in the virtual machines.

Booting the VMware ESX Server v2.5.x from Symmetrix storage arrays

VMware ESX Server v2.5.x hosts have been qualified for booting from Symmetrix devices interfaced through Fibre Channel as specified in the *EMC Support Matrix*.

EMC does not recommend booting VMware ESX Server from the Symmetrix system unless the host is using VMware's native failover functionality.

Note: If VMware's native failover functionality is not used and it is necessary to use a Symmetrix device as a boot disk, you should shut down the Virtual Machines and the ESX Server during any maintenance procedures that might cause the boot disk to become unavailable to the host.

The Symmetrix device that is to contain the Master Boot Record (MBR) for the host must have a logical unit number (LUN) between the range of 0-127 and must have a LUN number different than any other device visible to the host.

Note that this includes the case of the Symmetrix Volume Logix database device. The Volume Logix device is write-protected so the installer will fail to write the MBR to this device.

- To force the installer to avoid an attempt to write to this device, EMC recommends masking the LUN. The administrative host must first initialize the Volume Logix database. The LUN may then be masked by modifying the active configuration file to enable the `fba flag2` to restrict access to the Volume Logix database.
- An alternative to masking the VCM DB (Volume Logix Database), is to map the Volume Logix database so that it is the highest LUN presented to the host However, the LUN number should not be higher than 254 (FE) if it is to be used by Solutions Enabler or EMC ControlCenter® running on a Microsoft host.

  Note: The EMC-recommended method is to use LUN masking.

When attaching the VMware host to the Symmetrix storage array, use the adapter in the lowest-numbered PCI slot in the server.

To keep the configuration and installation simple, it is recommended that only that HBA be cabled to the array. Ensure that the boot BIOS

or firmware has been applied to your HBA and that the boot BIOS has been enabled on the HBA to be used for boot.

Prior to the installation on a Symmetrix LUN, the Linux host HBA must have successfully logged into the array. Using Solutions Enabler from another host, at least one LUN must be assigned to the host.

During the installation procedure, it is recommended, but not required, that only one LUN be allocated to the host for ease of use. After the installation has completed, additional LUNs can be assigned to the host.

When performing the installation, you will need to manually type **bootfromsan or bootfromsan-text** at the installation prompt.

By selecting this option, the HBAs will automatically be set to be shared between the Service Console and the Virtual Machines. Remember that this parameter should not be changed when creating your VMware ESX Server configuration.

Cautions

CAUTION

If VMware loses all paths to the array for a long enough period, the disks disappear from the system. A hard reboot is required to bring the system back to a usable state.

Any of these events could crash a system booting from a Symmetrix storage array:

- Lost connection to the Symmetrix system (pulled or damaged cable connection).
- Symmetrix service and upgrade procedures, such as on-line Symmetrix microcode upgrades and / or configuration changes.
- Symmetrix director failures, including failed lasers.
- Symmetrix system power failure.
- Storage area network service / upgrade procedures such as firmware upgrades or hardware replacements.

Restrictions

- Raw disk mapping is not supported when booting from the fabric.
- Only FC-SW environments are supported when booting from the SAN.

Fabric zoning

When using VMware ESX Server hosts in a fabric environment, the zoning must be set up as single initiator and single target zoning. A single initiator / single target zone is composed of one HBA and one Symmetrix storage array port.

Note: Multi-initiator zones are not supported in a VMware ESX Server fabric environment.

Required storage system configuration

Symmetrix system configuration is performed by an EMC Customer Engineer (CE) through the Symmetrix Service Processor.

The CE will configure the Symmetrix Storage Arrays settings for each Fibre Channel port. The procedures in this document assume that any switches and storage systems to be used in this configuration have been installed, and that the Symmetrix Fibre Channel Adapter ports have been connected to the switch ports.

Note: EMC highly recommends using Volume Logix to mask LUNs.

To verify that the VMware ESX Server host can see all of the Symmetrix target devices, configure the host as described in the remainder of this chapter.

Useful VMware ESX Server utilities and functions

Table 2 describes some useful utilities and functions.

Table 2 Utility/function descriptions

Utility/Function	Description
fdisk	Command used to create and manipulate partition tables.
grep	Command used to search through a file or files to find specific text.
vmkpcidivy	Command used to create a boot configuration for the VMkernel.
vmkfstools	Command used to create and manipulate files on LUNs owned by the VMware ESX Server host.
vmkload_mod	Command used to view, load, remove driver modules in the VMkernel.
vmkmultipath	Command used to display and set the configuration on a particular device and the paths attached to that device.
vm-support	Command used to gather information about the VMware ESX Server itself and virtual machines to assist in debugging issues or to obtain performance information for the virtual machines.
vmkdump	Command used to manage the VMkernel's dump partition.

Addressing Symmetrix devices

This section discusses the following:

Fabric addressing

Each port on a device attached to a fabric is assigned a unique 64-bit identifier called a World Wide Port Name (WWPN). These names are factory-set on the HBAs in the hosts, and are generated on the Fibre Channel directors in the Symmetrix system.

Note: For comparison to Ethernet terminology, an HBA is analogous to a NIC card, and a WWPN to a MAC address.

Note: The ANSI standard also defines a World Wide Node Name (WWNN), but this name has not been consistently defined by the industry.

When an N_Port (host server or storage device) connects to the fabric, a login process occurs between the N_Port and the F_Port on the fabric switch. During this process, the devices agree on such operating parameters as class of service, flow control rules, and fabric addressing. The N_Port's fabric address is assigned by the switch and sent to the N_Port. This value becomes the source ID (SID) on the N_Port's outbound frames and the destination ID (DID) on the N_Port's inbound frames.

The physical address is a pair of numbers that identify the switch and port, in the format **s,p**, where **s** is a domain ID and **p** is a value associated to a physical port in the domain. The physical address of the N_Port can change when a link is moved from one switch port to another switch port. The WWPN of the N_Port, however, does not change. A Name Server in the switch maintains a table of all logged-in devices, so an N_Port may adjust automatically to changes in the fabric address by keying off the WWPN.

The highest level of login that occurs is the process login. This is used to establish connectivity between the upper-level protocols on the nodes. An example is the login process that occurs at the SCSI FCP level between the HBA and the Symmetrix system.

SCSI-3 FCP addressing

The Symmetrix director extracts the SCSI Command Descriptor Blocks (CDB) from the frames received through the Fibre Channel link. Standard SCSI-3 protocol is used to determine the addressing mode and to address specific devices.

The Symmetrix supports three addressing methods based on a single-layer hierarchy as defined by the SCSI-3 Controller Commands (SCC):

- Peripheral Device Addressing
- Logical Unit Addressing
- Volume Set Addressing

All three methods use the first two bytes (0 and 1) of the 8-byte LUN addressing structure. The remaining six bytes are set to 0s.

For Logical Unit and Volume Set addressing, the Symmetrix port identifies itself as an Array Controller in response to a host's Inquiry command sent to LUN 00. This identification is done by returning the byte 0x0C in the **Peripheral Device Type** field of the returned data for Inquiry. If the Symmetrix system returns the byte 0x00 in the first byte of the returned data for Inquiry, the Symmetrix system is identified as a *direct access* device.

Upon identifying the Symmetrix system as an array controller device, the host should issue a SCSI-3 Report LUNS command (0xA0), in order to discover the LUNs.

The three addressing modes, contrasted in Table 3, differ in the addressing schema (Target ID, LUN, and Virtual Bus) and number of addressable devices.

Table 3 Symmetrix SCSI-3 addressing modes

Addressing mMode	Code[a]	"A" Bit	"V" Bit	Response to "Inquiry"	LUN discovery method	Possible addresses	Maximum logical devices[b]
Peripheral Device	00	0	X	0x00 Direct Access	Access LUNs directly	16,384	256
Logical Unit	10	1	0	0x0C Array Controller	Host issues "Report LUNS" command	2,048	128
Volume Set	01	1	1	0x0C Array Controller	Host issues "Report LUNS" command	16,384	512

a. Bits 7-6 of byte 0 of the address.

b. The actual number of supported devices may be limited by the type host or host bus adapter used.

Note: The addressing modes are provided to allow flexibility in interfacing with various hosts. In all three cases the received address is converted to the internal Symmetrix addressing structure. Volume Set addressing is the default for Symmetrix systems. Select the addressing mode that is appropriate to your host.

4

VMware ESX Server v2.x and CLARiiON Fibre Channel Environment

This chapter lists CLARiiON Fibre Channel support information specific to the VMware ESX Server environment.

Refer to the *E-Lab Navigator* or contact your EMC representative for the latest information on qualified hosts.

VMware/CLARiiON environment

This section provides CLARiiON Fibre Channel support information specific to the VMware ESX Server environment.

Host and hardware connectivity

Refer to the *EMC Support Matrix* or contact your EMC representative for the latest information on qualified hosts, host bus adapters, and connectivity equipment.

Note: EMC does not support mixing HBAs in the same system, including HBA models from the same vendor.

CLARiiON connectivity

Access Logix™ must be installed on the CLARiiON storage array to which the VMware ESX Server is being attached.

When using versions of VMware ESX Server v2.1.x and prior, the VMware ESX Server host must be manually registered on the CLARiiON array. Support for the Naviagent/CLI has been introduced with the release of VMware ESX Server v2.5.0.

VMware ESX Server uses the Linux version of the Navisphere Agent/CLI. The Naviagent must be loaded on the Service Console of the ESX Server while the Naviagent/CLI is supported on both the Service Console and the Virtual Machines.

VMware ESX Server owns the HBAs, not the operating systems running in the virtual machines. As a result, the VMware ESX Server's HBAs will be registered on the CLARiiON and assigned to a Storage Group.

The virtual machines will be assigned LUNs through the VMware ESX Server Service Console.

When attaching a VMware ESX Server host to a CLARiiON Storage Array, follow the appropriate guidelines, depending upon the environment.

- If the VMware ESX Server is not using the native failover functionality within the kernel, then the failover mode on the CLARiiON must be set to 0.

- If the VMware ESX Server is using the native failover functionality within the kernel, then the failover mode must be set to 1. This is the default setting.

Example 1

In this example as shown in Figure 18 on page 73, one host with two HBAs is attached to one CLARiiON array using two separate switches. Two SP ports on each SP within the array are being used. HBA1 is zoned to SPA0 and to SPB1. HBA2 is zoned to SPA1 and to SPB0. The zones should be composed of a single initiator and a single target so they would be created with one HBA and on SP port. Two HBAs with two paths each to the array totals four paths and if using single initiator / single target zoning, there are 4 zones.

In this particular example, two switches are used. Using only one switch is supported, but such a configuration would lack redundancy. Preferably, a minimum of two switches should be used as this adds another level of redundancy. Alternatively, for additional redundancy, two separate fabrics can be utilized.

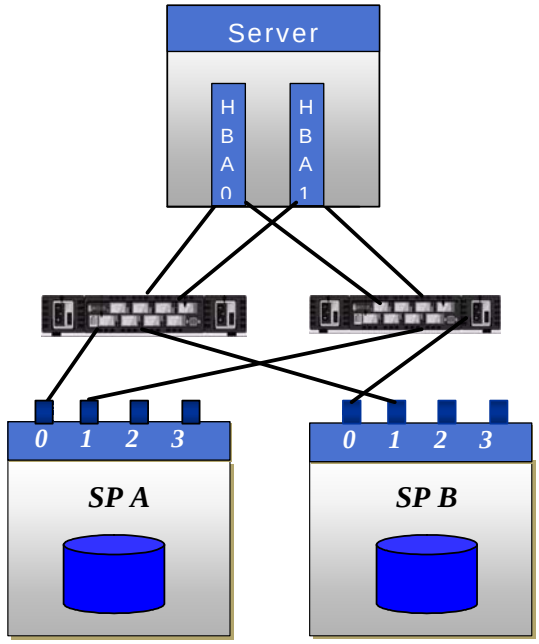

Figure 18 One host, two switches, and one CLARiiON array

Example 2 In this example as shown in Figure 19, one host with two HBAs is attached using a two-switch fabric to four CLARiiON array SPs. In this configuration, the zones are created with one HBA and one SP port. For instance:

- HBA1 is zoned to one SPA port on each of the four arrays.
- HBA2 is zoned to one SPB port on each of the four arrays.

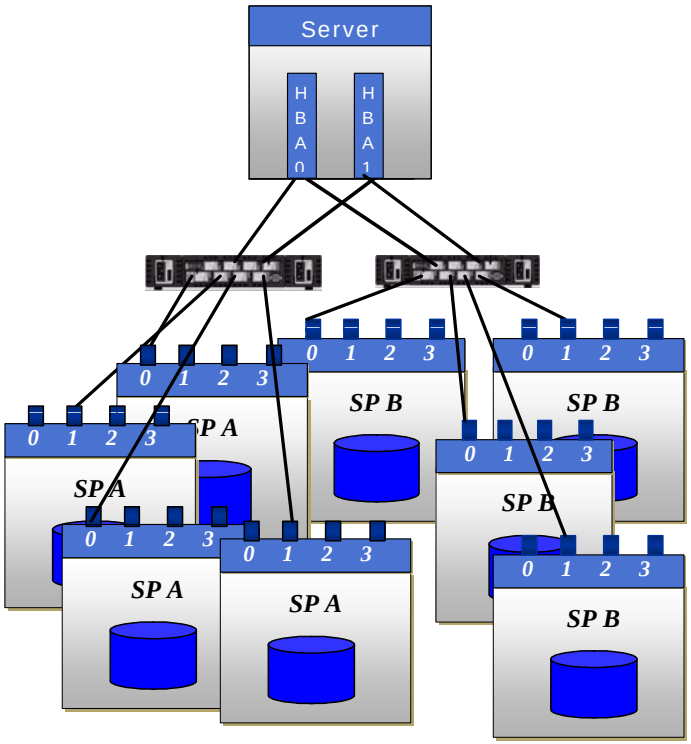

Figure 19 **One host, two switches, and four CLARiiON arrays**

Note: All qualified HBAs are listed in the *EMC Support Matrix*.

When assigning CLARiiON LUNs to a VMware ESX Server host, the LUNs may be assigned to only one SP or their assignments may be split between the two SPs. Either configuration is valid.

Booting the VMware ESX Server v2.5.x from CLARiiON storage arrays

VMware ESX Server v2.5.x hosts have been qualified for booting from CLARiiON devices interfaced through Fibre Channel as specified in the *EMC Support Matrix*.

EMC does not recommend booting VMware ESX Server from the CLARiiON storage array unless the host is using VMware's native failover functionality.

Note: If VMware's native failover functionality is not used and it is necessary to use a CLARiiON device as a boot disk, you should shut down the Virtual Machines and the ESX Server during any maintenance procedures that might cause the boot disk to become unavailable to the host.

The CLARiiON device that is to contain the Master Boot Record (MBR) for the host must have a lower logical unit number (LUN) than any other device visible to the host. This device must be mapped as /dev/sda by the VMware ESX Server operating system for the boot to succeed from the device.

Note that this includes the case of the CLARiiON ghost LUN. If the LUN 0 presented to a SCSI path is not owned by the SP attached to that path, a disconnected LUN 0 (ghost LUN) will be presented. The ghost LUN is not write-enabled so that the MBR cannot be written to it. Always ensure that the boot LUN is owned by the correct SP and does not trespass at any time; trespass of the boot LUN would result in a system crash.

When attaching the VMware host to the CLARiiON storage array, use the adapter in the lowest-numbered PCI slot in the server. To keep the configuration and installation simple, it is recommended that only that HBA be cabled to the array. Ensure that the boot BIOS or firmware has been applied to the HBA and rhat the boot BIOS has been enabled on the HBA to be used for SAN boot.

Prior to the installation, the VMware ESX Server must have been registered on the array and assigned to a Storage Group. If using version v2.1.x and prior, the host's HBAs will need to be manually registered on the CLARiiON Storage Array. The VMware ESX Server should be attached to the lowest numbered SP with visible LUNS.

At least one LUN must be bound to the host's Storage Group and owned by the SP connected to the HBA being used for the fabric boot. The lowest-numbered path to the boot LUN must be the active path.

Since the lowest numbered path to the boot LUN must be active, attaching to the CLARiiON via direct connect or FC-AL is not supported. Only FC-SW environments are supported when booting from the SAN.

It is required that the boot LUN be assigned Host LUN ID 0.

During the installation procedure, it is recommended, but not required, that only one LUN be assigned to the Storage Group for ease of use. After the installation has completed, additional LUNs can be added to the Storage Group.

When performing the installation, you will need to manually type **`bootfromsan`** or **`bootfromsan-text`** at the installation prompt.

By selecting this option, the HBAs will automatically be set to be shared between the Service Console and the Virtual Machines. Remember that this parameter should not be changed when creating your VMware ESX Server configuration.

Cautions

CAUTION

If VMware loses all paths to the array for a long enough period, the disks disappear from the system. A hard reboot is required to bring the system back to a usable state.

Restrictions

- Raw disk mapping is not supported when booting from the fabric.
- Only FC-SW environments are supported when booting from the SAN.

Fabric zoning

When using VMware ESX Server hosts in a fabric environment, the zoning must be set up as single initiator and single target zoning. A single initiator / single target zone is composed of one HBA and one CLARiiON storage array port.

Note: Multi-initiator zones are not supported in a VMware ESX Server fabric environment.

Required storage system configuration

CLARiiON configuration is performed by an EMC Customer Engineer (CE) through Navisphere Manager. The CE will configure the CLARiiON Storage Array for each Fibre Channel port.

The procedures in this document assume that any switches and storage systems to be used in this configuration have been installed, and that the Symmetrix Fibre Channel Adapter ports have been connected to the switch ports.

Note: EMC highly recommends using Access Logix to mask LUNs.

To verify that the VMware ESX Server host can see all of the CLARiiON target devices, configure the host as described in the remainder of this chapter.

Useful VMware ESX Server utilities and functions

Table 4 describes useful utilities and functions.

Table 4 Useful VMware ESX Server utilities and functions

Utility/Function	Description
fdisk	Command used to create and manipulate partition tables.
grep	Command used to search through a file or files to find specific text.
vmkpcidivy	Command used to create a boot configuration for the VMkernel.
vmkfstools	Command used to create and manipulate files on LUNs owned by the VMware ESX Server host.
vmkload_mod	Command used to view, load, remove driver modules in the VMkernel.
vmkmultipath	Command used to display and set the configuration on a particular device and the paths attached to that device.
vm-support	Command used to gather information about the VMware ESX Server itself and virtual machines to assist in debugging issues or to obtain performance information for the virtual machines.
vmkdump	Command used to manage the VMkernel's dump partition.

Failover functionality in VMware ESX Server v2.x for attach to CLARiiON storage arrays

Native failover functionality has been incorporated into the VMkernel beginning with VMware ESX Server v2.1.0. The VMkernel itself will see the multiple paths configured to the storage for the host. However, the Virtual Machines will be presented only a single SCSI path to the devices that are allocated or assigned to each Virtual Machine.

VMware ESX Server v2.x does support multiple paths to a Symmetrix / CLARiiON array, but load balancing is not currently supported. However, there are two alternate path polices that may be implemented. The path policies are MRU (Most Recently Used) and Fixed.

Note: PowerPath is not available for VMware ESX Server.

The default policy is Fixed for active / active disk arrays and MRU for active / passive disk arrays.

The MRU policy is strongly recommended for CLARiiON arrays and any other active / passive arrays. The reason for this recommendation is the higher likelihood of "path thrashing" occurring on active / passive types of arrays when using the Fixed failover policy. Path thrashing occurs when two hosts are accessing the same LUN via different storage processors on a disk array. In the case of the CLARiiON, this may cause a loop where a LUN is trespassed from one storage processor to the other and back again. Such thrashing behavior will cause poor performance and may cause errors.

With the MRU policy, VMware ESX Server hosts will use one active path to a storage processor. A trespass will be forced to the other storage processor only when the initial path has failed and there is no access to devices via that path. Therefore, VMware ESX Server hosts using the MRU policy will quickly settle on accessing a LUN through the storage processor that is accessible to all hosts.

The MRU policy will not automatically restore the I / O back to the initial path. This limitation means that there is not an easy way to restore the paths initially used by the VMware ESX Server hosts to access the LUNs. The MRU policy will choose the first available path that avoids performing a trespass.

The trespass back to the original path will need to be performed manually using the Navisphere Manager.

The Fixed path policy allows you to manually configure the path to be used. If the preferred path fails, the Fixed path policy will transfer I/O to the live path until the time when the preferred path is restored. When the preferred path is restored, all I/O will be transferred back to the preferred path.

The paths currently being used can be viewed and changed using the command `vmkmultipath` or using the MUI. When using the MUI, select the **Options** tab and then select **Storage Management**, then **Failover Paths**.

Figure 20 is an example of the information reported with the **Failover Paths** dialog box:

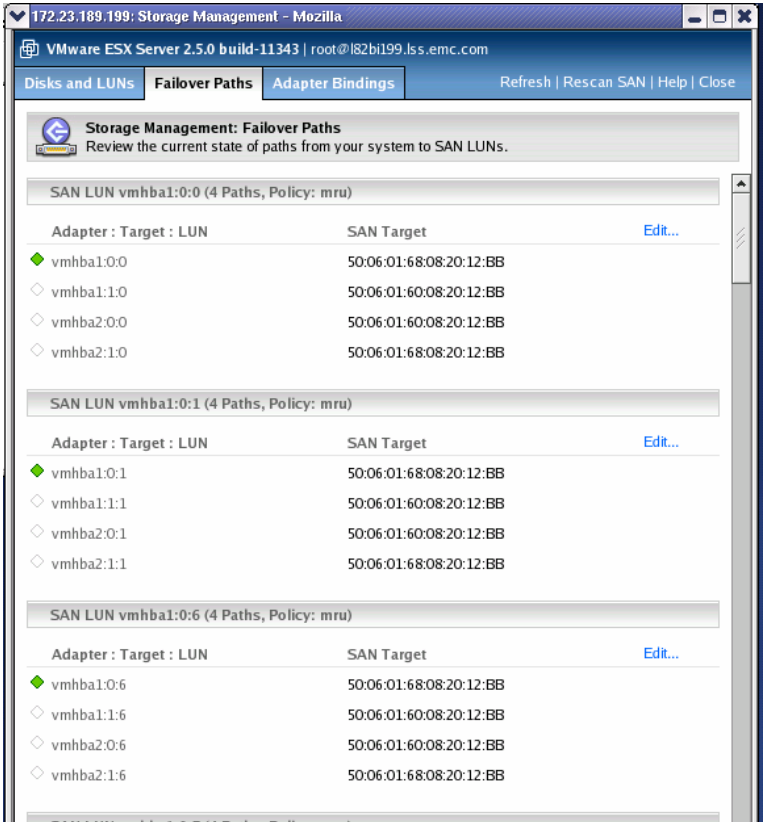

Figure 20 Information reported with Failover Paths dialog box

The **Failover Paths** illustrates the SAN LUN information, the paths being used, and the failover policy. As can be seen in the figure, the LUNs attached to the VMware ESX Server host are individually listed with the pertinent HBA, target, and LUN information. The current HBA being used and the failover policy being used may be edited here on a per LUN basis.

When a cable is pulled, I/O from the VMware ESX Server and the Virtual Machines essentially will be frozen for approximately a minute. The Fibre Channel HBA driver within VMware ESX Server must identify that the link is down before a failover will occur. During the time that the link is down, the Virtual Machines will not be responsive, but they should resume to their normal behavior once the failover has occurred. In the event that all paths to the storage have failed, then I/O errors will be reported by the Virtual Machines and access to the `/vmfs` directory will also fail.

Manual registration of VMware ESX Server v2.x for attach to CLARiiON storage arrays

Currently, the VMware ESX Server host must be manually registered on the CLARiiON array as the currently released version of the Naviagent does not include support for VMware ESX.

In order to manually register the host, you must first identify the WWNs of the HBAs. The WWNs are used when creating zones on the switches and in registering the VMware ESX Server host onto the CLARiiON arrays. Zoning by WWN is not required as zoning by either WWN or by switch port is supported.

- The recommended method to discover WWNs is to run the `wwpn.pl` command for either QLogic or Emulex HBAs. For each vmhba instance, the `wwpn.pl` will provide the corresponding QLogic or Emulex WWPNs.

  For example:

  ```
   [root@l82bi199 /]# /usr/sbin/wwpn.pl
  vmhba2: 210000e08b0910a7 (Qlogic) 6:1:0
  vmhba3: 210100e08b2910a7 (Qlogic) 6:1:1
  ```

- An alternate method to obtain Emulex HBAs' initiator and target information is to refer to `/proc/scsi/lpfcdd/`*`N`* (where *`N`* indicates the file for each adapter in the system) when the driver is loaded. By `grep`'ing the file(s), the necessary information to register to host will be reported.

  grep the file to obtain the initiator and target information.

  For example,

  ```
  grep DID /proc/scsi/lpfcdd/1
  ```

  produces output similar to the following for the first Emulex HBA:

```
lpfc0t00 DID 060300 WWPN 50:06:01:61:10:60:12:5c WWNN 50:06:01:60:90:60:12:5c
lpfc0t01 DID 060400 WWPN 50:06:01:69:10:60:12:5c WWNN 50:06:01:60:90:60:12:5c
```

- Also, when using QLogic HBAs, the same information is logged in `/proc/scsi/qla2x00/`*`N`* (where *`N`* indicates the file for each adapter in the system) when the driver is loaded.

  grep the file to obtain the initiator and target information.

For example, for a host with QLA23xx HBAs:

```
grep scsi-qla /proc/scsi/qla2300/0
```

produces output similar to the following for the first QLogic HBA:

```
scsi-qla0-adapter-node=200000e08b0910a7;
scsi-qla0-adapter-port=210000e08b0910a7;
scsi-qla0-target-0=5006016810601270;
scsi-qla0-target-1=5006016010601270;
scsi-qla0-target-2=50060160082012bb;
scsi-qla0-target-3=50060169082012bb;
```

Now that the WWNs have been identified, the VMware ESX Server host can now be registered to the CLARiiON.

The following section will describe the manual registration process.

In order to manually register the host on the CLARiiON array, perform the following steps:

1. Start the Navisphere Manager in a web browser on a host to be used for management purposes.
2. Select the **Storage** tab so that the arrays being managed by the Navisphere Manager are displayed.
3. Right click on the appropriate array, and select the **Connectivity Status** option.
4. The **Connectivity Status** dialog for that array will show the Initiator WWNs for each host logged into the array.

An example of the **Connectivity Status** dialog can be seen Figure 21 on page 84.

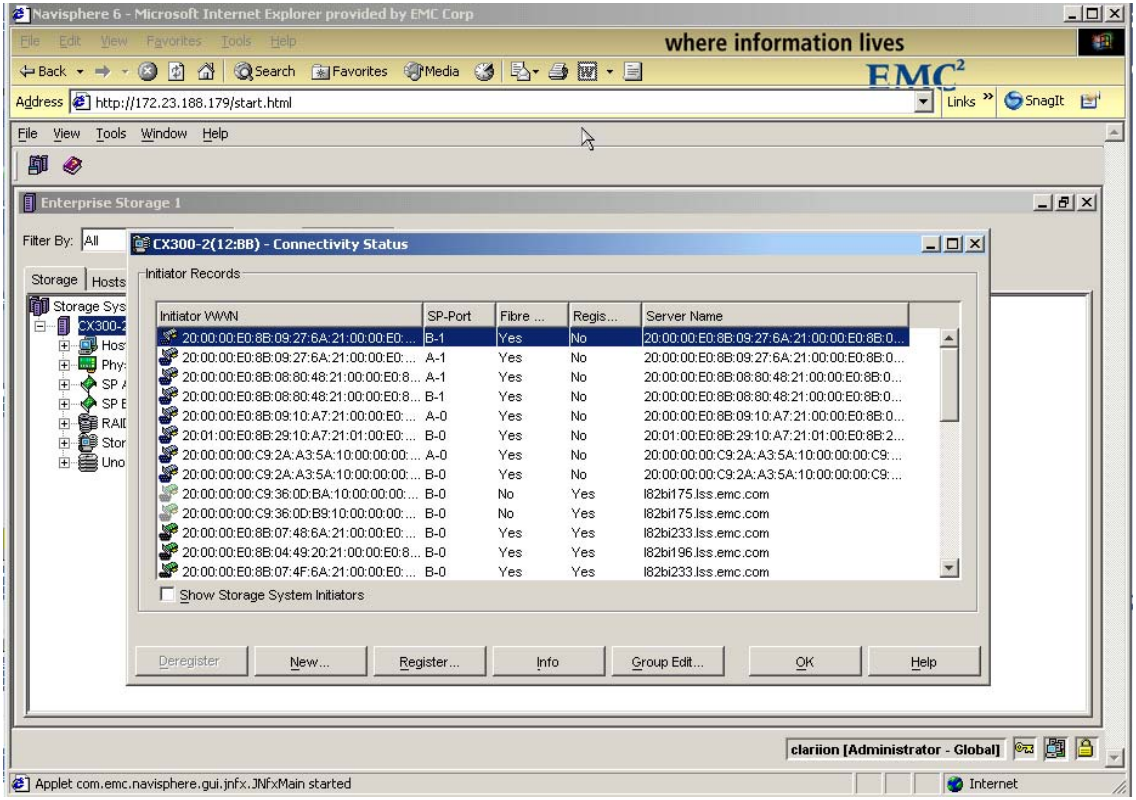

Figure 21 Connectivity Status dialog

5. In order to manually register your host's HBAs, select the WWN of the HBA and click on **Register**. You will be prompted to add the server information such as the server name and IP address. Repeat for each HBA instance for this VMware ESX Server host. This will register this HBA or HBAs for your VMware host.

6. Another dialog will appear called **Register Initiator Record**. When attaching a VMware ESX Server to a CLARiiON, the standard default Linux director bit settings should be used.

   An example of a non-failover environment is a host with one HBA zoned only to one SP port, such as SPA2.

For a non-failover environment, the required settings are as follows:

- Initiator Type: CLARiiON Open
- ArrayCommPath: Disabled
- FailOverMode: 0
- Unit Serial Number: Array

For a failover-enabled environment, the required settings are as follows:

- Initiator Type: CLARiiON Open
- ArrayCommPath: Enabled
- FailOverMode: 1
- Unit Serial Number: Array

Please note that the failover functionality referred to here is the native failover functionality incorporated into the VMkernel, not PowerPath. PowerPath is not available for the VMkernel.

Figure 22 on page 86 shows an example of registering for a failover-enabled environment.

Note: The box for the ArrayCommPath parameter is checked and the Failover Mode is set to 1.

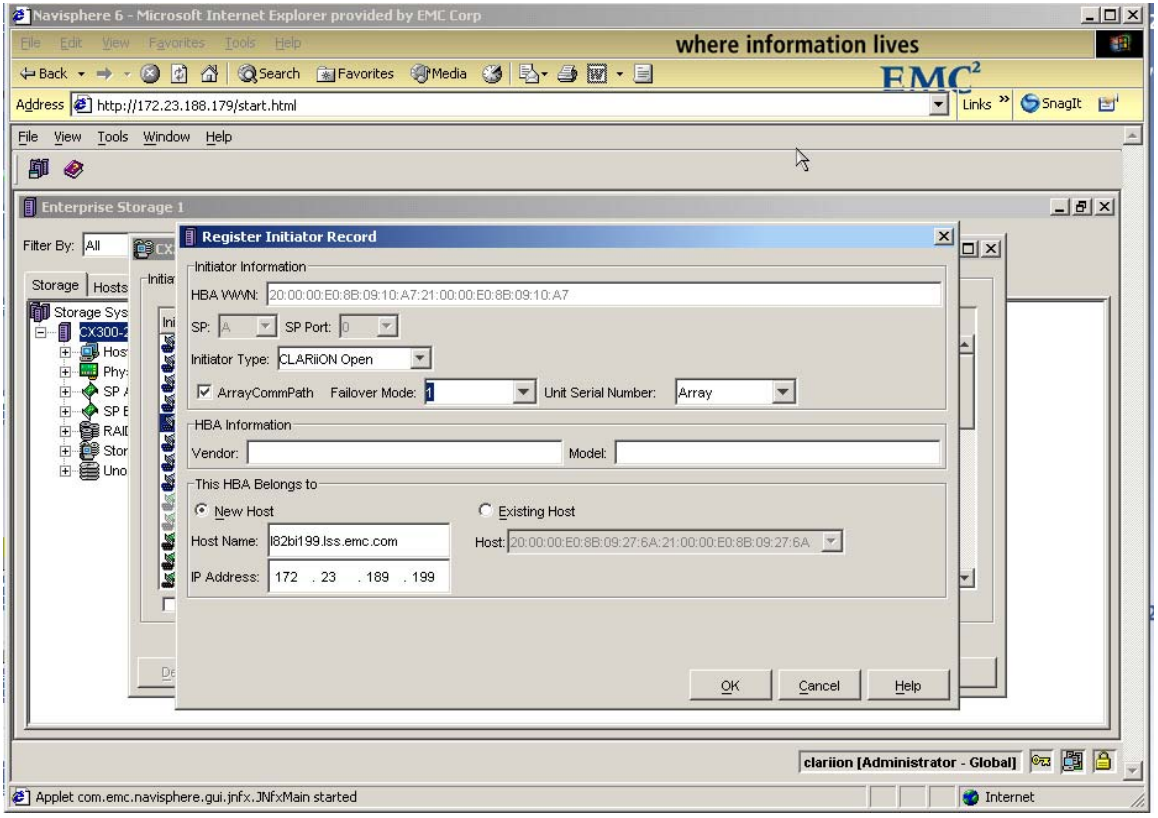

Figure 22 Registering for a failover-enabled environment example

Because no Naviagent is used on the VMware ESX Server, you will receive a warning message when registering the host.

Figure 23 on page 87 shows an example of the warning message.

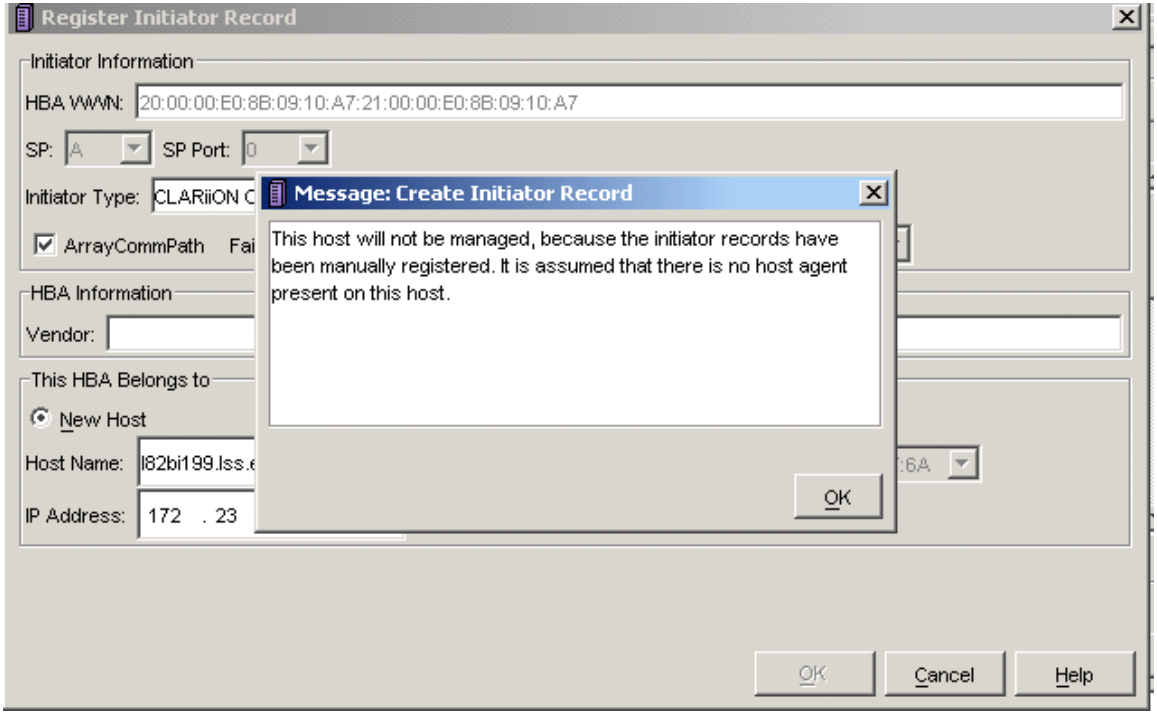

Figure 23 Warning message example

Because the Naviagent is not being used, this warning is to be expected and is acceptable.

7. Repeat steps 1 through 4 for each HBA in the VMware ESX Server system.
8. To verify that your host has been properly registered, right click on the Array and go to the **Hosts** tab.

   The host will be reported as attached, but manually registered as in the example of the system named `l82bi199.lss.emc.com` as show in Figure 24 on page 88.

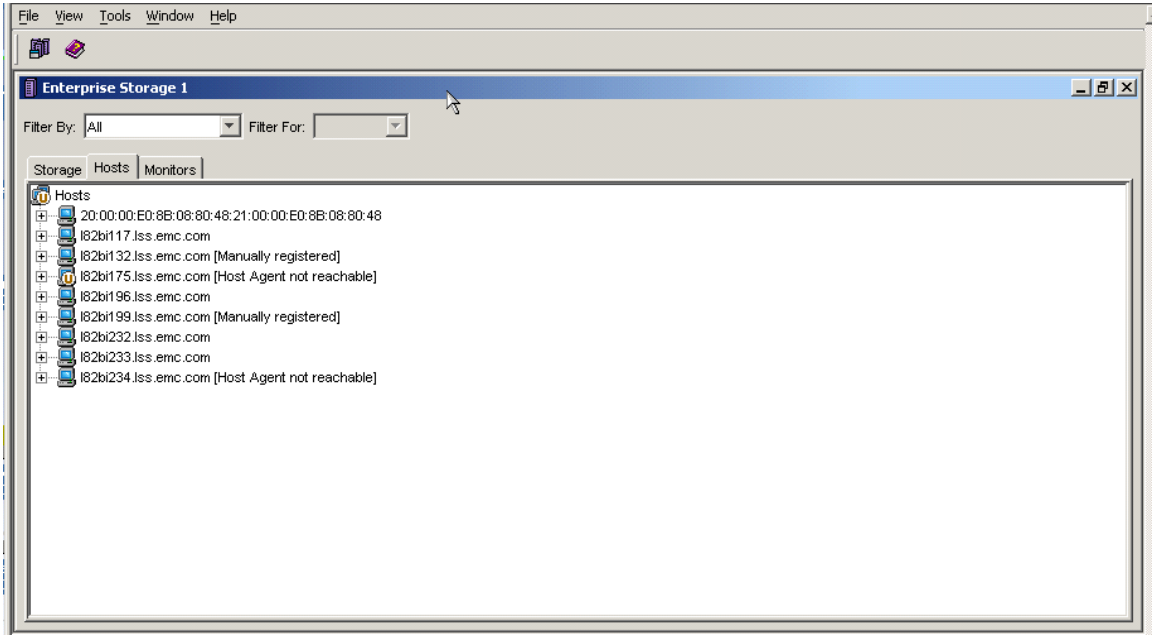

Figure 24 Host reported as attached, but manually registered example

Adding the VMware ESX Server host to a storage group

After the host is registered manually, the host can be added to a Storage Group so that it can be allocated devices and can access those devices.

1. Right click on the array, and select the option **Create Storage Group**. Provide it with a name such as the name or IP address of the host or a nickname.
2. After the Storage Group is created, then LUNs and the host may be added to the Storage Group. Select the newly created Storage Group from the list, and right click on it.
3. Select the **Properties** option and select LUNs to add to the group and select the host to own the group.

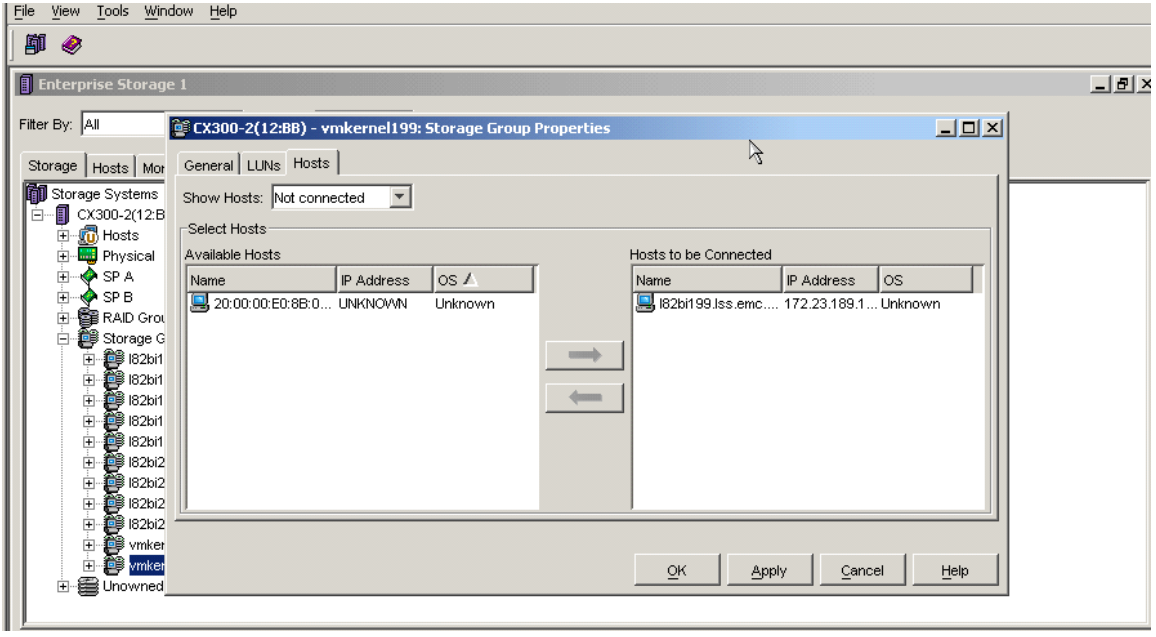

Figure 25 Storage Group Properties window

4. The VMware host will be listed under the **Hosts To be Connected** and the OS will report itself as 'Unknown'. This is to be expected since the host was manually registered on the array and is acceptable behavior.

Performing a CLARiiON non-disruptive upgrade (NDU) with VMware ESX Server hosts

The CLARiiON FLARE, or storage-system software, is installed via a non-disruptive upgrade or NDU. This process allows users to upgrade the FLARE without having to take their hosts offline.

Note: Before performing a non-disruptive upgrade (NDU) on a CLARiiON array attached to an ESX Server, keep track of what LUNs are owned by what Storage Processor. During an NDU, since the MRU failover policy is used with CLARiiON systems, all LUNs may end up on a single SP after the upgrade.

Once the upgrade is complete, manually trespass each LUN using Navisphere Manager or CLI to the Storage Processor that owned the LUN before the NDU process. As a best practice, place LUNs on their default Storage Processor. This ensures a uniform balance between the two Storage Processors.

To perform an NDU.

Before starting the NDU procedure, ensure that there are no faults reported and no failed SPs on the array to be upgraded.

1. Verify the version of FLARE currently running on the CLARiiON array.
2. Select the desired software bundle to which the CLARiiON array will be upgraded.
3. Download the software bundle and extract on the package on the desktop. A packages directory and an ".lst" file will be extracted.
4. Log into the CLARiiON array via a web browser such as Microsoft Internet Explorer.
5. Right click on the array and select the Software Operations and Software Installation Wizard to begin the NDU process.

   For an example, please see Figure 26 on page 91.

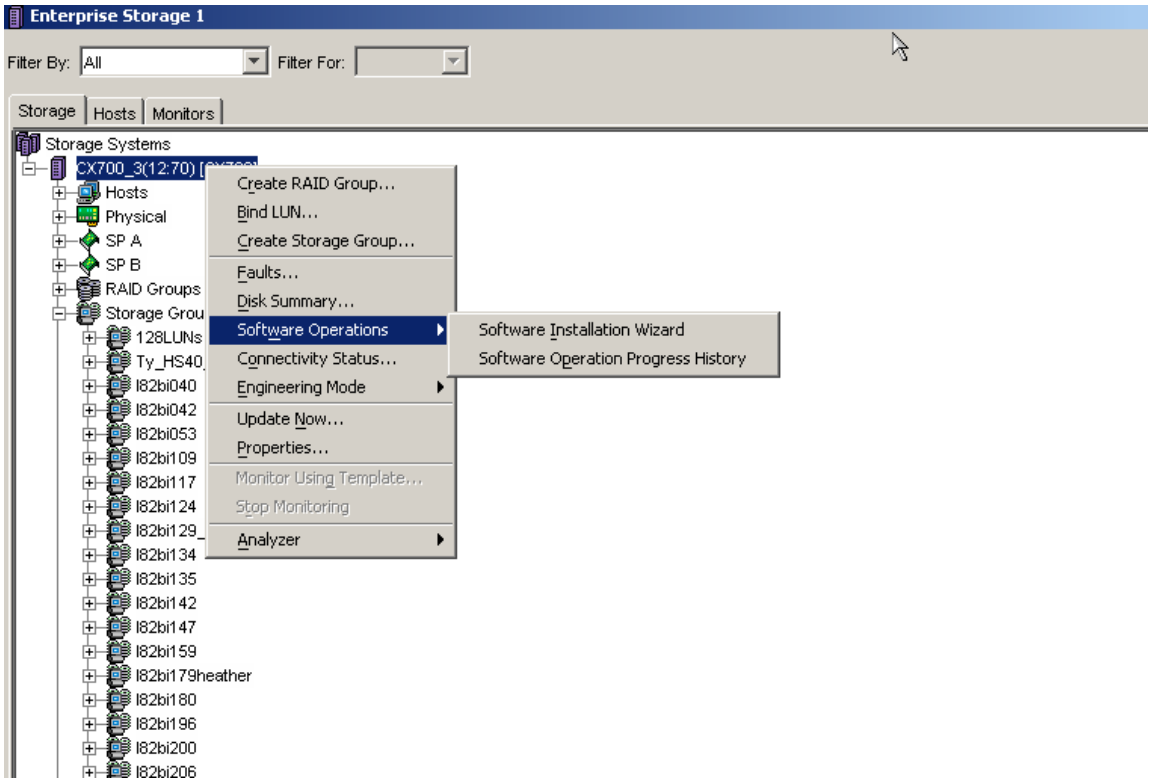

Figure 26 Software Installation Wizard

The Software Installation Wizard will guide you through the upgrade.

6. After selecting the Software Installation Wizard, the window as shown in Figure 27 on page 92 will display to report the steps that are to be taken by the wizard during the upgrade of the FLARE.

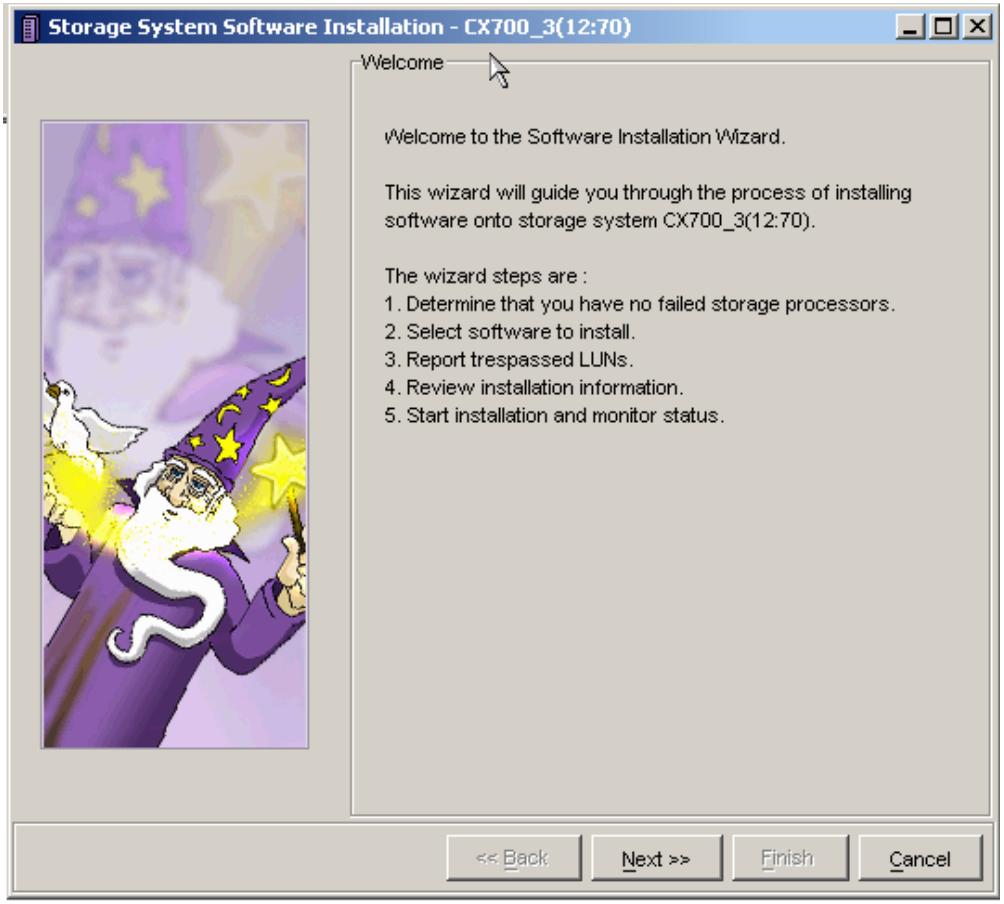

Figure 27 Software Installation Wizard window

Click **Next** to continue.

7. The next window shown by the Wizard (Figure 28 on page 93) allows a change to the default 360 second delay. This value may be changed as indicated in the window.

Note: Please note that reducing the value to one that is too low may incur data loss.

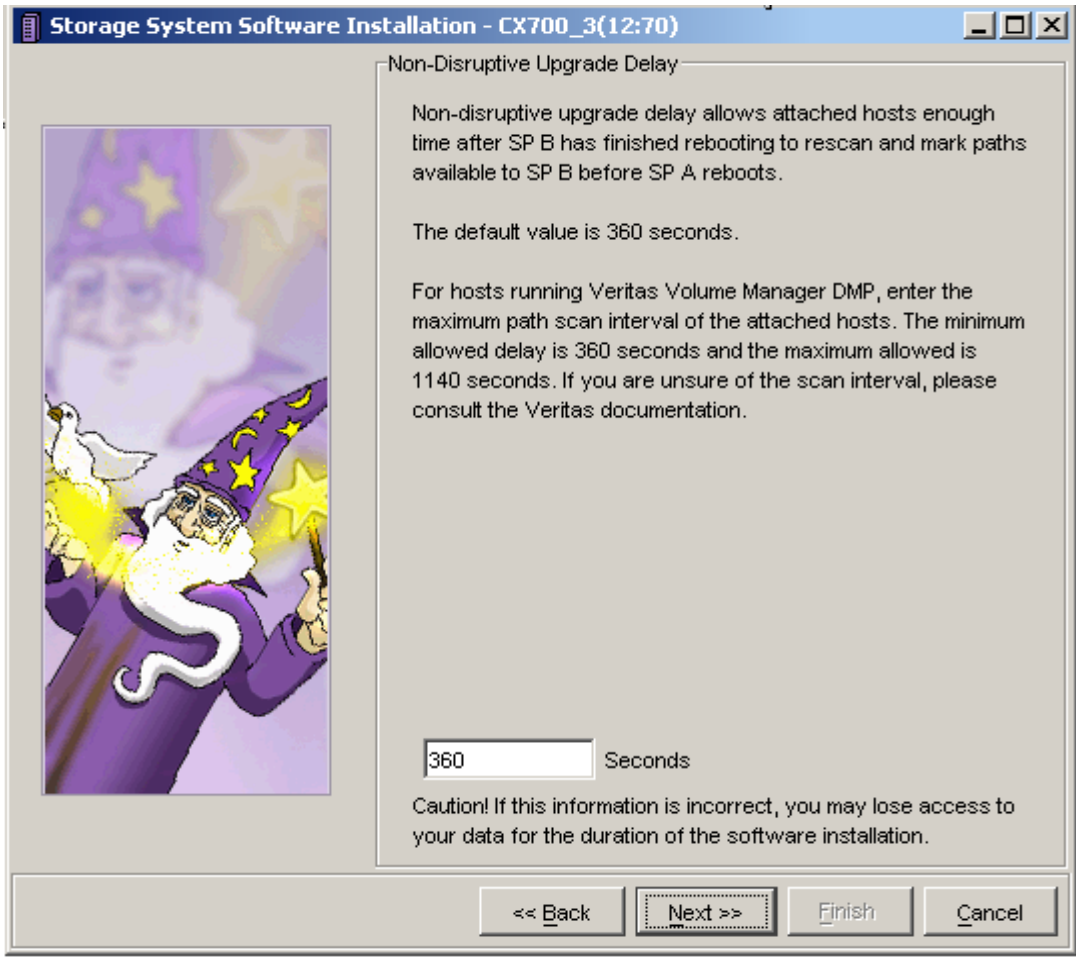

Figure 28 Non-Disruptive Upgrade Delay window

During an NDU, SP B is upgraded first and rebooted. After SP B has finished rebooting, the Wizard provides a built-in delay of 360 seconds. This delay should provide enough time for the hosts attached to the array to rescan the SCSI bus and mark the paths to SP B available before rebooting SPA.

When SP B is rebooted, all of the LUNs owned by SP B will be trespassed to SPA. The VMware ESX Server native failover functionality will handle this trespass so that I/O may continue on the same LUNs now available via SPA. When SPA is rebooted, the reverse will occur.

The **Failover Paths** window on the VMware ESX Server may be used to monitor graphically the failovers from one path to another.

8. The installation wizard shows the current version of FLARE as well as the software enablers currently installed on the array. The software enablers should be upgraded when the FLARE is upgraded so ensure that the downloaded package includes the software enablers corresponding to those currently installed.

   Use the Browse option to search for the downloaded software bundle to be used for the upgrade. Select the index file (*.1st) and click **Next** to begin the NDU. The .lst file will automatically go to the packages directory and will pull the necessary files from that directory.

   Under the package name, select FLARE operating environment (Figure 29 on page 95).

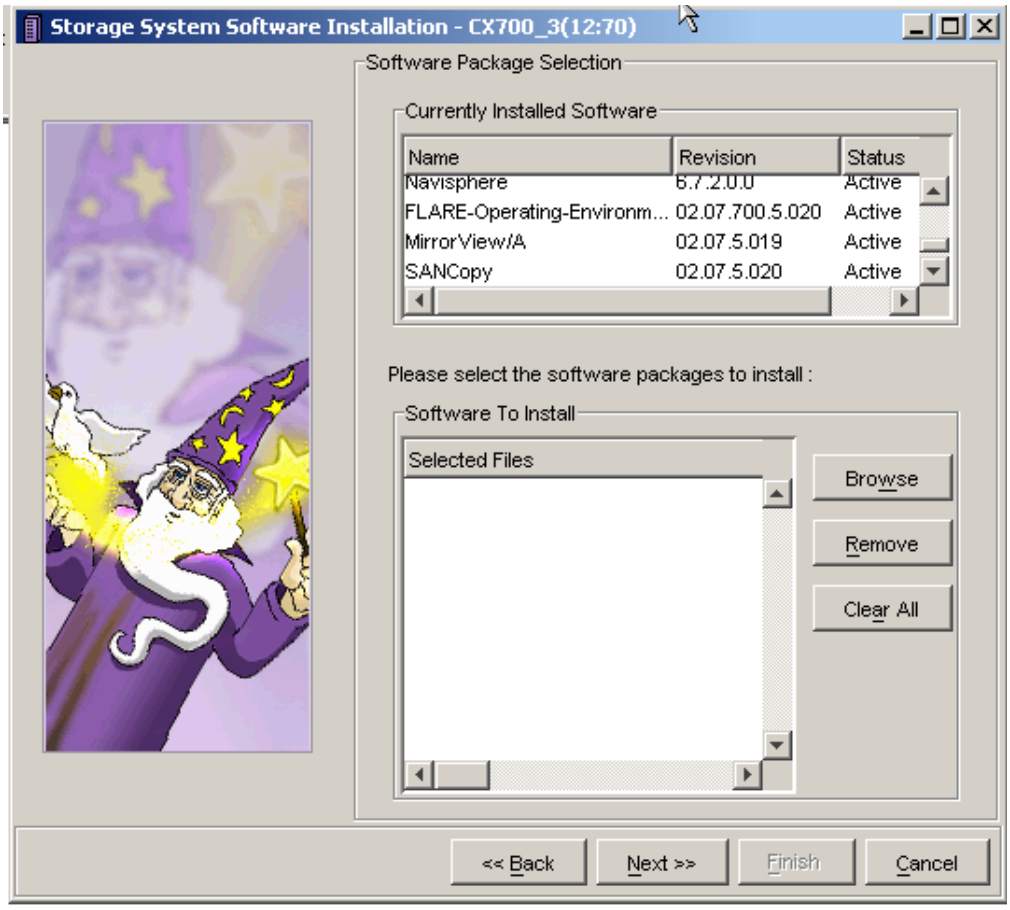

Figure 29 Software Package Selection window

The progress of the NDU may be monitored via the Software Operation Progress History window. In the window, each step completed will be marked with a check mark as shown in Figure 30 on page 96.

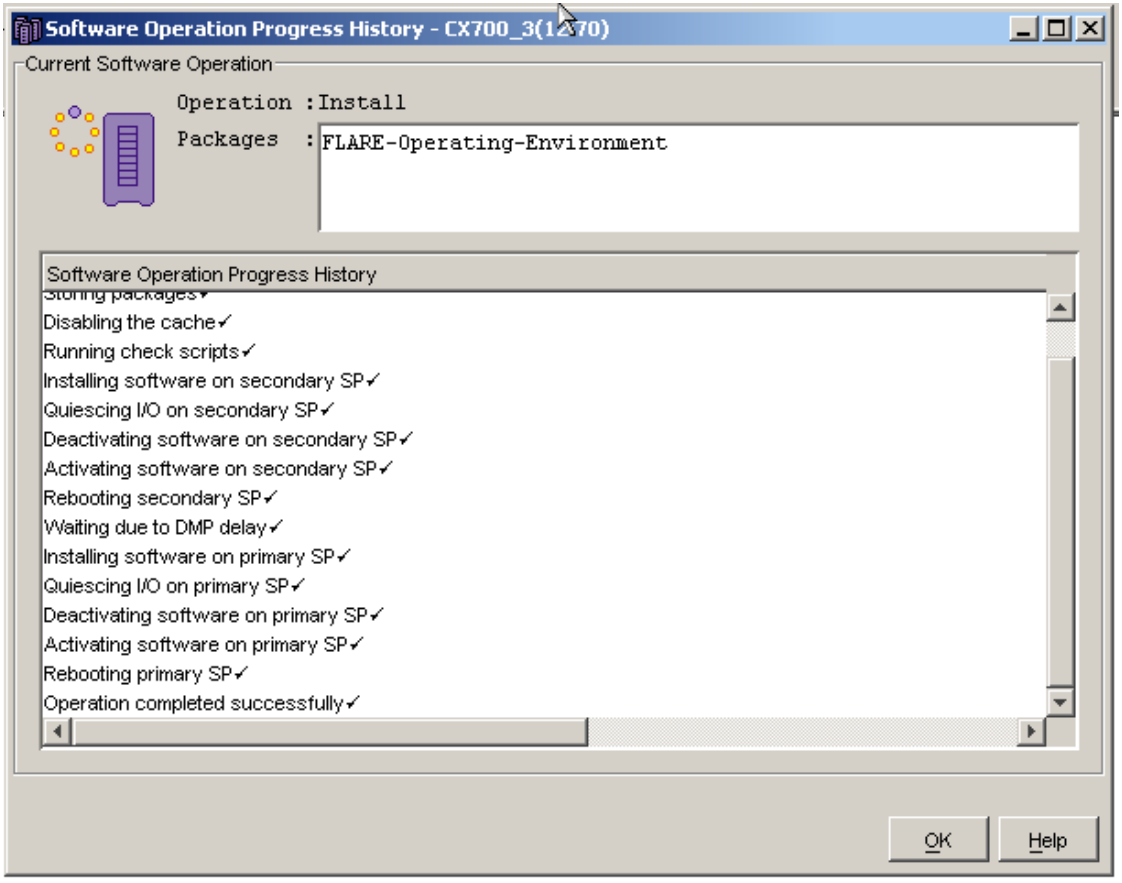

Figure 30 Software Operation Progress History window

Once the NDU has completed, return to the main screen and right click on the array. Select the Storage System Properties option and then select Software. Select the FLARE Operating Environment so it is highlighted and select the **Commit** button. The **Commit** button will activate the new FLARE on the array.

Manual trespass on CLARiiON arrays to recover the original path

In some cases when using the MRU policy, it may be desirable to trespass the LUNs back to their original path. This may be performed manually using the Navisphere Manager. The instructions for a manual LUN trespass are below.

Note: When ESX Server is booting from SAN, on several occasions the CLARiiON LUN that the ESX Server in installed on trespasses. For example, if your ESX Server is booting from the CLARiiON array that an NDU is being performed on, the NDU may cause the ESX Server boot LUN to trespass. In this case, if you happen to reboot the ESX Server after the NDU, it cannot find the disk to boot from. This is expected behavior because of the way the boot LUN setup is done in the QLogic HBA BIOS. The fix is to use the Navisphere GUI to trespass the boot LUN back to the default owner.

1. Open a web browser and type in the IP address of the CLARiiON array your host is attached to.
2. Once the GUI is updated with the log-in window, type in the user name and password for the array.
3. Click on the array to expand the listing of properties.
4. Click on the SP that holds the active path for the LUNs. The LUNs allocated to that VMware ESX Server host should be listed and should be associated with the host's name and the vmhba device numbers.

For example, refer to the LUNs 192, 193, 196, and 197 owned by host l82bi199 in Figure 31.

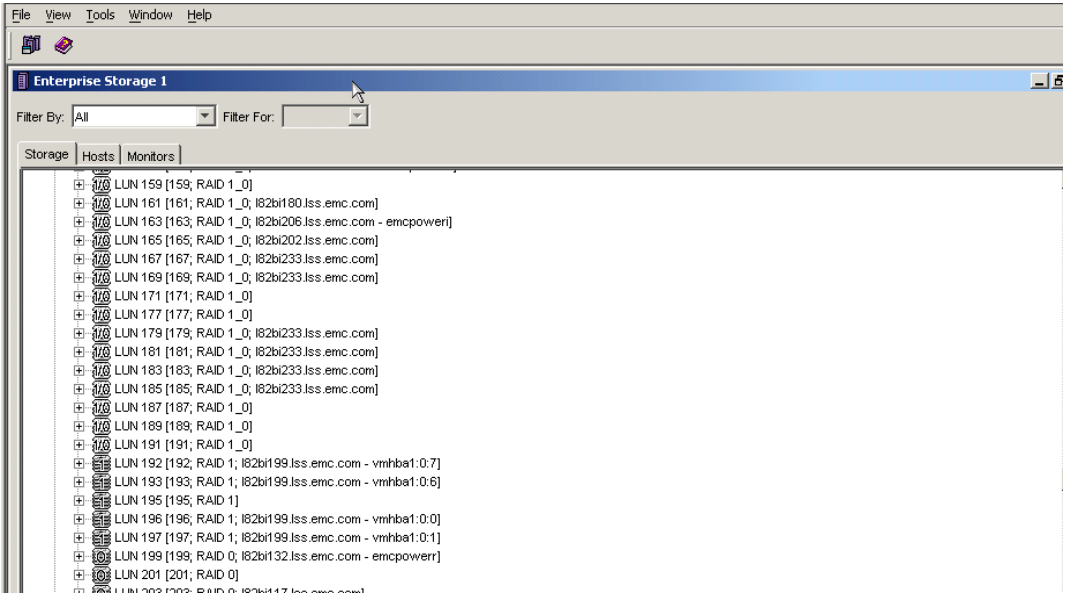

Figure 31 LUNs example

5. To view the Properties of the LUNs, right click on the LUN. A window will provide the specific properties of the LUN.

For example, if you were to select LUN 196 that is allocated to the host, then you would see information as shown in Figure 32:

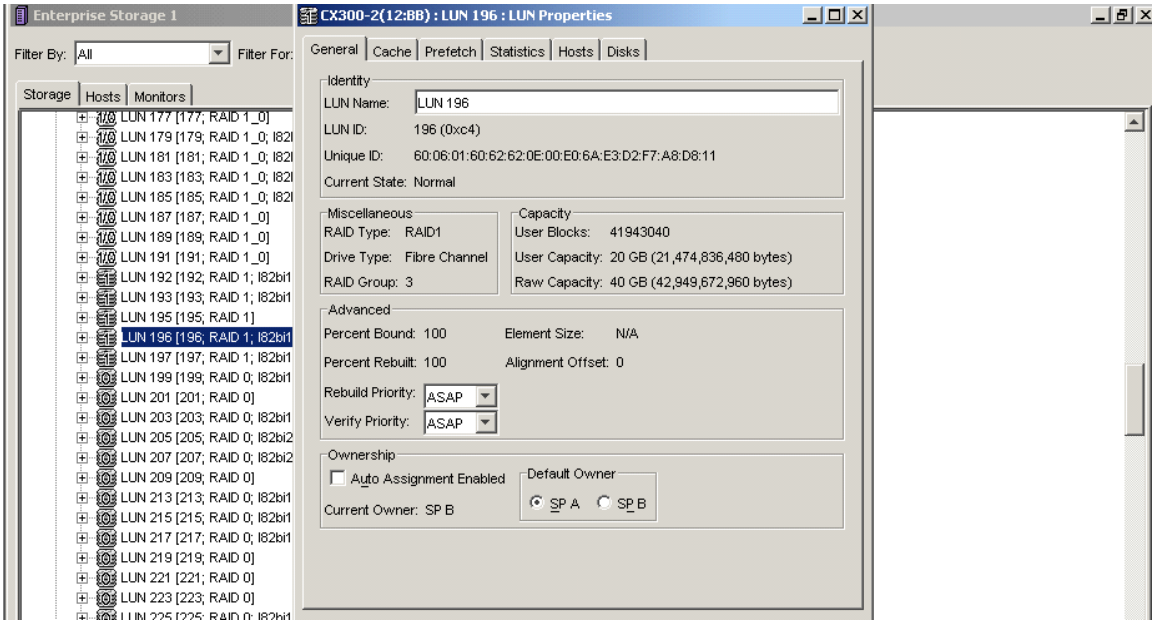

Figure 32 View LUN properties

As can be seen in the **General** tab above, LUN 196 is originally owned by SP A. In this instance, LUN 196 has been trespassed to SP B.

6. For the host in this example, both LUN 192 and 196 are owned by SP A while LUNs 193 and 197 are owned by SP B. A trespass has occurred and currently, all four LUNs are located on SP B.

In order to trespass LUNs 192 and 196 back to SP A, right click on those LUNs. In the menu that is presented (see Figure 33 for an example), select the option to **Trespass**.

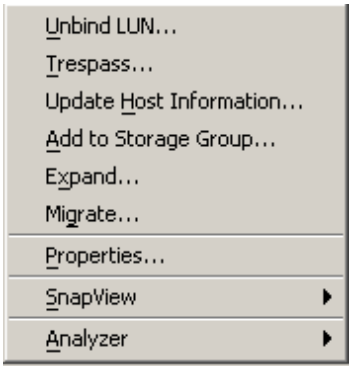

Figure 33 Menu

7. With the appropriate LUNs are highlighted, select the option to Trespass. You will be prompted to confirm that the LUNs should be trespassed as can be seen in Figure 34 on page 101.

   Select **Yes** to continue and to trespass the LUNs.

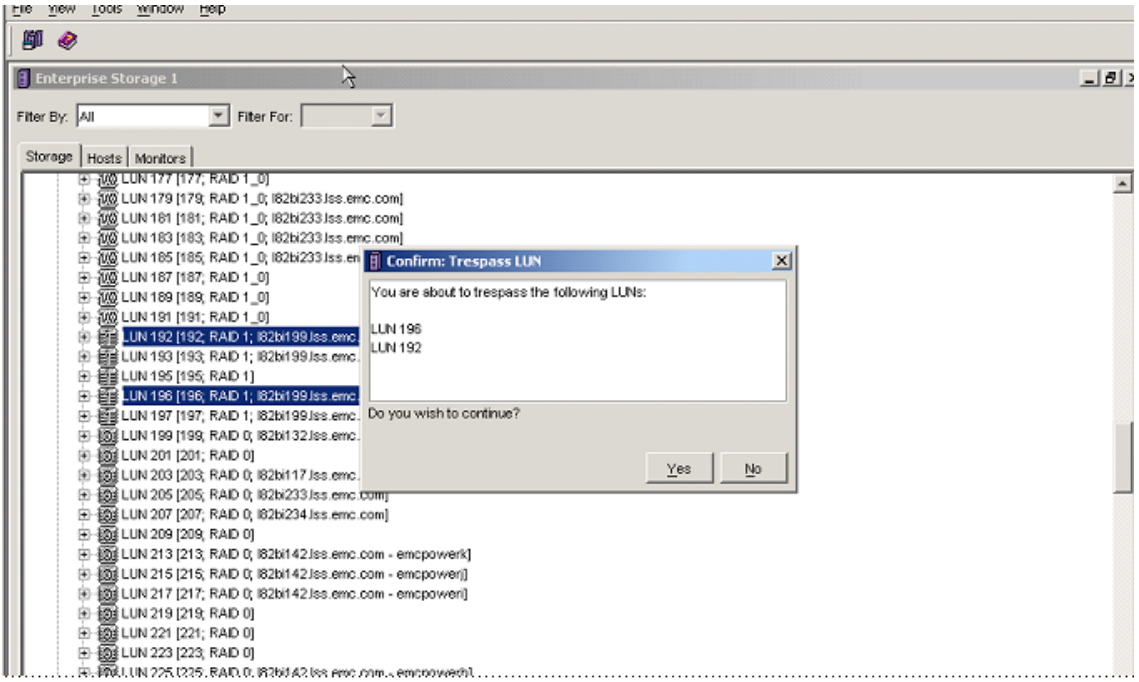

Figure 34 Confirm trespass LUNs

8. After the trespass has been completed, only LUNs will be reported on SP B as is demonstrated in Figure 35.

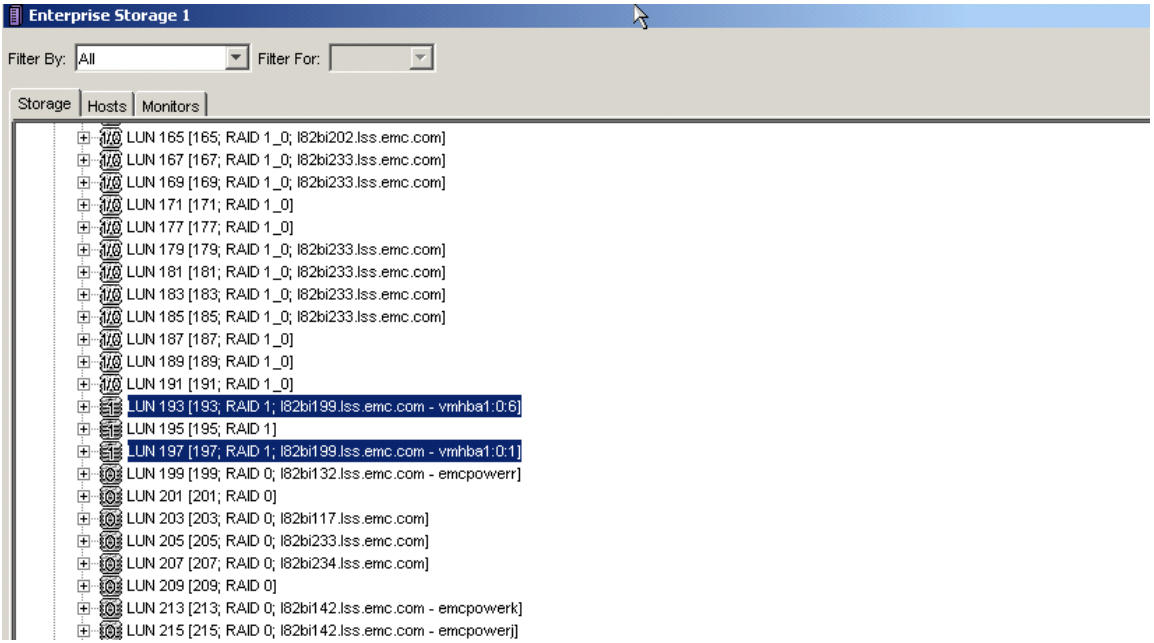

Figure 35 Report on LUNs

5

VMware ESX Server v3.x with EMC Storage

This chapter contains information on VMware ESX Server 3.x.

Overview

VMware introduced the new VMware Infrastructure 3 in 2006. VMware Infrastructure 3 is a feature-rich suite which includes VMware ESX Server (versions 3.0 and 3.0.1) and VMware VirtualCenter (versions 2.0 and 2.0.1) plus new distributed services including VMware HA, VMware DRS, and VMware Consolidated Backup.

VMware ESX Server 3.0 supports both Fibre Channel storage attach and iSCSI storage attach via the in-kernel software iSCSI initiator. VMware ESX Server 3.0.1 supports Fibre Channel storage attach and iSCSI storage attach via the in-kernel software iSCSI initiator as well as the QLogic hardware iSCSI initiator.

Keep the following in mind when setting up ESX Server 3.0 or 3.0.1 in a customer environment:

- Please refer to the latest *EMC Support Matrix* and make sure that the server, HBAs, and switches you are planning to use are supported.

  Note: Please note that a number of servers are supported with ESX Server 3.0.1 but *not* with ESX Server 3.0.

- The following are the Symmetrix FA port settings that are required for ESX Server 3.0 and 3.0.1:

  - C, VCM, SC3, SPC-2, UWN: Enabled

  If you are upgrading from ESX Server 2.x to ESX Server 3.x, you should keep in mind that ESX Server 2.x did not require the SPC-2 to be set. In this case, the bit can be enabled as part of the upgrade from ESX Server 2.x or when new devices are presented. However, enabling the SPC-2 bit in an existing VMware infrastructure can cause disruption. For instructions, please refer to the white paper, *Enabling SPC-2 Compliancy on EMC Symmetrix DMX Devices Connected to VMware VI3 Environments*, available at http://powerlink.emc.com/km/live1/en_US/Offering_Technical/White_Paper/H4116-enabling-spc2-compl-emc-symmetrix-dmx-vmware-envnmt-wp.pdf.

- The Disk.MaxLUN parameter in ESX Server 3 has been set to 256 by default (in ESX Server 2.x it was 8), so you do *not* have to change this parameter for your host to access more than 8 LUNs.

Maximum storage specifications supported with ESX Server 3.x can be found in the VMware *Storage SAN Compatibility Guide for ESX Server 3.x* at http://www.vmware.com/pdf/vi3_san_guide.pdf.

Configuring iSCSI initiator ports

This section describes the steps required to configure iSCSI initiator ports.

Step 1: Use the VMkernel to create a vSwitch and assign an IP address to the NICs or iSCSI HBAs

To assign an IP address to a NIC in an ESX Server host:

While configuring a software iSCSI initiator using NICs, VMware recommends NIC *teaming* to provide failover between multiple NICs, which means the NICs are on the *same* vSwitch using the *same* IP address.

1. Log in to the VMware VI client as administrator.
2. From the inventory panel, select the server with the iSCSI initiator to which you want to assign an IP address.
3. Click the **Configuration** tab and click **Networking**.
4. Click **Add Networking**.
5. In **Connection Type**, select **VMkernel** and click **Next**.

   The **Network Access** page appears.
6. Either select a vSwitch or click **Create a virtual switch** to create a new vSwitch.
7. Check the box next to each adapter that you want to connect to the vSwitch. The adapters that you select appear in the **Preview** pane. Click **Next**.
8. Under **Port Group Properties**, select or enter a network label.
9. Under **IP Settings**, enter the adapter IP address and subnet mask for the VMkernel.
10. If necessary, set the **VMkernel Default Gateway.**
11. Click **Next**. Review the **Summary** and if all of the settings are correct, click **Finish**.

To assign an IP address to an iSCSI HBA in an ESX Server host

1. Log in to the VMware VI client as administrator.
2. From the inventory panel, select the server with the iSCSI HBA to which you want to assign an IP address.

3. Click the **Configuration** tab and then click **Storage Adapters**.
4. Select the iSCSI HBA initiator whose IP address you want to assign and click **Properties**.
5. Click **Configure**.
6. If you desire, in **iSCSI Alias** enter the user-friendly name that you use to identify the iSCSI hardware initiator.
7. Under **Hardware Initiator Properties** choose one of the IP settings options.
8. Click **OK**. and then click **Next**. Review the summary, and if all of the settings are correct, click **Finish**.

Step 2: Enable services and agents in the initiator firewall

Configure the service console firewall to accept services and installed management agents, enabling the services and agents to access the ESX Server, as follows:

1. Log in to the VMware VI client as administrator.
2. In the VI Client, under **Hosts & Clusters,** click the server.
3. Click the **Configuration** tab and then click **Security Profile**.
4. Click **Properties** to open the **Firewall Properties** dialog box. This dialog box lists services and management agents.
5. If not already checked, enable **Software iSCSI Client** by checking the box.
6. Click **OK**.

Step 3: Configure ESX Server 3.X for iSCSI software initiators or hardware initiators

Configure ESX Server 3.X for a software initiator (NIC) or hardware initiator (iSCSI HBA) as follows:

1. Log into VMware VI Client as administrator and go to the **iSCSI Initiator Properties** page for the initiator that you want to configure:
   a. From the inventory panel, select the server with the initiator that you want to configure.
   b. Click the **Configuration** tab, and click **Storage Adapters**.
   c. Select the iSCSI initiator that you want to configure, and click **Properties**.

2. In the **iSCSI Initiator Properties** page, click the **General** tab and then click **Configure**.
3. For an iSCSI *software* initiator, check the **Enabled** box to enable the initiator; *or*

   For an iSCSI *hardware* initiator, deselect **Enabled** box to disable software initiators.
4. Under **iSCSI Properties**, you can enter an **iSCSI name** and **iSCSI Alias** for the software initiator and then click **OK**.
5. Add target addresses for the software or hardware initiator:
   a. Click the **Dynamic Discovery** tab and then click **Add**.
   b. Enter the send targets server information and click **OK** to add target information from a selected storage system.
6. Click **Close** to close the **iSCSI Initiator Properties** page.

Once the iSCSI initiator ports on the ESX Server are configured, iSCSI storage must be presented to the ESX Server. Please refer to the latest *EMC Support Matrix* for EMC arrays that are currently supported via iSCSI attach to VMware ESX Server 3.x.

Fibre Channel over Ethernet Converged Network Adapters

EMC is now supporting Fibre Channel over Ethernet (FCoE) Converged Network Adapter (CNA) offerings with VMware ESX Servers. FCoE adapters represent a method to converge both Fibre Channel and Ethernet traffic over a single physical link to a switch infrastructure that manages both storage (SAN) and network (IP) connectivity within a single unit.

The benefits of FCoE technology become apparent in large data centers:

- Where dense, rack-mounted and blade server chassis exist.
- Where physical cable topology simplification is a priority.
- In virtualization environments, where several physical storage and network links are commonly required.

The installation of a FCoE adapter provides the host with an Intel-based 10 gigabit ethernet interface (using the existing "in-box" drivers), and a Fibre Channel HBA interface, which will require the installation of the supported driver revision.

Refer to the *EMC Support Matrix* for the most up-to-date list of supported FCoE Converged Network Adapter (CNA) offerings.

This section describes the steps required to install and configure FCoE CNAs in a VMware ESX Server for Emulex and QLogic:

- "Configuring Emulex FCoE CNAs with VMware ESX Server" on page 110
- "Configuring QLogic FCoE CNAs with VMware ESX server" on page 113

Important points to consider

- Currently, the only VMware ESX Server version that supports Fibre Channel over Ethernet is ESX Server 3.5 Update 2. Always refer to the *EMC Support Matrix* to verify which servers are supported in FCoE configurations with ESX Server.
- At this time, ESX Server boot from SAN via FCoE CNAs is not supported. However, virtual machines booting from the SAN are supported.
- iSCSI-attach to a SAN via FCoE CNAs is not supported.

Configuring Emulex FCoE CNAs with VMware ESX Server

This section provides information on configuring Emulex FCoE CNAs with the VMware ESX Server, including:

Installing the CNA

To install the CNA:

1. Install VMware ESX Server 3.5 update 2 on your server.

   For instructions on installing the operating system, refer to the *ESX Server 3 and VirtualCenter Installation Guide* at http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_installation_guide.pdf.

2. Power off the host system.

3. With the host system powered off, install the CNA and make the necessary cable connections as instructed in the accompanying CNA documentation (or on the Emulex website at www.emulex.com).

   The card installs into a single PCI bus slot.

4. Reapply power and allow the system to boot normally.

Installing the Emulex CNA driver

Using the emulex adapter with the VMware ESX Server requires CNA driver software. Refer to the latest *EMC Support Matrix* for supported driver versions.

To install the Emulex CNA driver on the VMware ESX Server:

Download the appropriate CNA driver ISO image from the "VMware ESX 3.5 U2 Driver CD for driver lpfc_740 version elx_7.4.0.31" section on http://www.vmware.com/download/vi/drivers_tools.html.

Install the driver by following instructions under the "Driver CD as an Update CD Only" section on the driver download page.

Following the installation of the proper driver for the FCoE adapter, the Fibre Channel interface will function identically to that of a standard Emulex Fibre Channel HBA, as the LP21000/LP21002 simply encapsulates Fibre Channel traffic within ethernet frames.

Updating the Emulex CNA firmware and boot BIOS

To update the Emulex CNA firmware and boot BIOS:

1. If you do not already have the Emulex HBAnyware Applications kit installed on the ESX Server, download the latest available Emulex HBAnyware Applications kit that is compatible with the ESX Server version that you are running from www.emulex.com.

   Note: If the ESX Server has a /usr/sbin/hbanyware directory, the Emulex HBAnyware Applications kit is already installed.

   To install the HBAnyware CLI on the ESX Server:

   a. Log in as 'root' at the ESX Server console or in a secure shell session to the ESX Server.

   b. Copy the elxvmwarecorekit <*kit version*>.rpm file to a directory on the install machine.

   c. Using the **cd** command, change to the directory where you copied the rpmfile.

   d. Install the rpm. Type: **rpm -i elxvmwarecorekit <kit version>.rpm**

      – The rpmcontents are installed in /usr/sbin/hbanyware.
      – The hbacmd utility is also located in this directory.

2. Copy the appropriate firmware and BIOS files from the CD-ROM or the Emulex website to the appropriate directory, /usr/sbin/hbanyware.

   - The firmware file for Emulex LP21000 CNA is **ad100a5.all**
   - The firmware file for Emulex LP21002 CNA is **af100a5.all**
   - The BIOS file for Emulex LP2100x CNA is **AB202A2.PRG**

3. Execute the command **/usr/sbin/hbanyware/hbacmd ListHBAs** and record the Port WWN information for all adapters within the host.

4. To update the Emulex CNA firmware:

   a. Execute the command **/usr/sbin/hbanyware/hbacmd Download <*WWPN*> <*firmwarefile*>**, where WWPN is the first Port WWN recorded from Step 3, and firmwarefile is the firmware file name listed in Step 2.

      The utility will report "Download Complete" when the firmware download has completed successfully.

   b. If the installed adapter is a dual-ported model, then the firmware does not need to be downloaded to each adapter port; downloading it to either port results in both ports being updated with the latest firmware. Otherwise, for multiple single-port adapter models, repeat Step 4a for each adapter port WWN reported in Step 3 to ensure that all adapter instances within the host have been updated.

5. To update the Emulex CNA Boot BIOS:

   a. Execute the command **/usr/sbin/hbanyware/hbacmd Download <*WWPN*> <*BIOSfile*>**, where WWPN is the first Port WWN recorded from Step 3, and BIOSfile is the BIOS file name listed in Step 2.

      The utility will report "Download Complete" when the BIOS download has completed successfully.

   b. If the installed adapter is a dual-ported model, then the boot BIOS does not need to be downloaded to each adapter port; downloading it to either port results in both ports being updated with the latest BIOS. Otherwise, for multiple single-port adapter models, repeat Step 5a for each adapter port WWN reported in Step 3, to ensure that all adapter instances within the host have been updated.

Updating the Emulex CEE/Menlo firmware

FCoE adapters include an additional chip component that requires the latest supported firmware. This chip is commonly referred to as a *CEE* (converged enhanced ethernet), or *Menlo* chip, the purpose of which is to handle the convergence of storage (FC) and network (IP) traffic over a single ethernet interface.

To update the CEE/Menlo firmware on the CNAs:

1. Copy the appropriate CEE/Menlo firmware file from the CD-ROM or the EMC-approved section of the Emulex website to the appropriate directory, /usr/sbin/hbanyware.

The CEE firmware file for Emulex LP2100x CNA is **/usr/sbin/hbanyware**.

2. Execute the command **/usr/sbin/hbanyware/hbacmd ListHBA** and record the Port WWN information for all adapters within the host.

3. Execute the command **/usr/sbin/hbanyware/hbacmd CEEdownload <*WWPN*>**, where WWPN is the first Port WWN recorded from Step 2, and BIOSfile is the BIOS file name listed in Step 1.

   The utility will report "Download Complete" when the BIOS download has completed successfully.

4. If the installed adapter is a dual-ported model, then the CEE/Menlo firmware does not need to be downloaded to each adapter port; downloading it to either port results in both ports being updated with the latest BIOS. Otherwise, for multiple single-port adapter models, repeat Step 3 for each adapter port WWN reported in Step 2, to ensure that all adapter instances within the host have been updated.

Configuring QLogic FCoE CNAs with VMware ESX server

This section provides information on configuring QLogic FCoE CNAs with the VMware ESX server, including:

Installing the CNA

To install the CNA:

1. Install VMware ESX Server 3.5 update 2 on your server.

   For instructions on installing the operating system, refer to the ESX Server 3 and VirtualCenter Installation Guide at http://www.vmware.com/pdf/vi3_35/esx_3/r35u2/vi3_35_25_u2_installation_guide.pdf.

2. Power off the host system.

3. With the host system powered off, install the CNA and make the necessary cable connections as instructed in the accompanying CNA documentation (or on the QLogic website at www.QLogic.com).

   The card installs into a single PCI bus slot.

4. Reapply power and allow the system to boot normally.

Installing the QLogic CNA driver

Using the QLogic adapter with the VMware ESX Server requires CNA driver software. Refer to the latest *EMC Support Matrix* for supported driver versions.

To install the QLogic CNA driver on VMware ESX Server:

Download the appropriate CNA driver ISO image from the "VMware ESX 3.5 U2 Driver CD for driver qla2300_707_vmw version 7.08-vm60" section on http://www.vmware.com/download/vi/drivers_tools.html.

Install the driver by following instructions under the "Driver CD as an Update CD Only" section on the driver download page.

Following the installation of the proper driver for the FCoE adapter, the Fibre Channel interface will function identically to that of a standard QLogic Fibre Channel HBA, as the QLE8042 simply encapsulates Fibre Channel traffic within ethernet frames.

Updating the QLogic CNA boot BIOS

To update the QLogic CNA boot BIOS:

1. If you do not already have the QLogic SANsurfer FC HBA CLI package installed on the ESX Server, download the latest available version compatible with the ESX Server version that you are running from www.QLogic.com.

   Note: If the ESX Server has a /opt/QLogic_Corporation/SANsurferCLI/ directory, the QLogic SANsurfer FC HBA CLI package is already installed.

   To install the HBAnyware CLI on the ESX Server:

   a. Log in as 'root' at the ESX Server console or in a secure shell session to the ESX Server.

   b. Copy the scli-<*kit version*>.rpm.gz file to a directory on the install machine.

c. Using the **cd** command, change to the directory where you copied the rpm.gz file.

d. Unzip the file. Type: **gunzip scli-<kit version>.rpm.gz**

e. Install the rpm. Type: **rpm -i scli-<kit version>.rpm**

- The rpm contents are installed in /opt/QLogic_Corporation/SANsurferCLI/.
- The scli utility is also located in this directory.

2. From the EMC-approved section of the QLogic Download site, download the boot code zip file for your CNA.

   Unzip the boot code zip file. The file with the .dat contains the EMC-specific parameters. The file with the .bin extension is the BIOS file.

3. To update the BIOS, execute the command

   /opt/QLogic_Corporation/SANsurferCLI/scli.

4. Select Options **7** (Utilities), then **1** (Flash Update), then **1** (FCoE Engine), then **1** (Update Entire Image).

   The utility prompts you to enter a file name.

5. Enter the name of the BIOS file.

   The utility should report the update has completed successfully.

6. Press **Enter** to continue, then Select Options **0** (Return to Main Menu), then **12** to exit the SANsurferCLI utility.

7. To install EMC-specific parameters on your CNA, execute the command

   /opt/QLogic_Corporation/SANsurferCLI/scli -r all <*File Name*>

   Where *File Name* is the .dat file mentioned in Step 2.

Updating the QLogic CEE/Menlo firmware

FCoE adapters include an additional chip component that requires the latest supported firmware. This chip is commonly referred to as a (converged enhanced ethernet), or *Menlo* chip, the purpose of which is to handle the convergence of storage (FC) and network (IP) traffic over a single ethernet interface.

To update the CEE/Menlo firmware on the CNAs:

1. Copy the appropriate CEE/Menlo firmware file from the CD-ROM or the EMC-approved section of the QLogic website to the appropriate directory, /opt/QLogic_Corporation/SANsurferCLI.
2. Execute the command **/opt/QLogic_Corporation/SANsurferCLI/scli**.

   Select Options **7** (Utilities), then **6** (FCoE Utility), then **1** (FCoE Engine), then **1** (Update Firmware).
3. The utility prompts to enter a file name. Enter the name of the CEE/Menlo firmware file listed in Step 1. The utility should report the update has completed successfully.

   Press **Enter** to continue, then Select Options **0** (Return to Main Menu), then **12** to exit the SANsurferCLI utility.
4. Reboot the host for the new CEE/Menlo firmware update to take effect.

6

Symmetrix SPC-2 Director Bit Considerations

This chapter contains information to consider for the Symmetrix SPC-2 director.

Considerations

Enginuity™ code versions 5671.58.64 (and later) for DMX and DMX-2, and 5771.87.95 (and later) for DMX-3, provide support for compliance with newer SCSI protocol specifications; specifically, SCSI Primary Commands - 2 (SPC-2) as defined in the SCSI document at http://www.t10.org/ftp/t10/drafts/spc2/spc2r20.pdf.

The SPC-2 implementation in Enginuity includes functionality which, based on OS and application support, may enhance disk-attach behavior to use newer SCSI commands optimized for a SAN environment (as implemented in Fibre Channel), as opposed to legacy (non SPC-2) functionality, which was targeted for older SCSI implementations utilizing physical SCSI bus-based connectivity (which cannot leverage the enhanced functionality of newer SCSI specifications).

In environments sharing director ports between hosts with multiple vendor operating systems, ensure that all hosts' operating systems are capable of supporting the SPC-2 functionality before enabling it on the port. If any OS sharing the affected director port does not support SPC-2 functionality, the SPC-2 bit cannot be set on a per-port basis and must be set on a per-initiator basis using Solutions Enabler 6.4 CLI. Refer to the *EMC Solutions Enabler Symmetrix Array Controls CLI v6.4 CLI Product Guide*, P/N 300-002-940, available on Powerlink, for details regarding how to set the SPC-2 bit on a per-initiator basis.

SPC-2 must be enabled for all initiators on a per-host basis, globally, so if SPC-2 conformance is enabled for a specific Symmetrix device visible to a specific host, SPC-2 conformance must be enabled for all paths to that same device and from that same host.

Offline and online migrations from SPC-2 disabled to SPC-2 enabled configurations are discussed in the white paper, *Enabling SPC-2 Compliancy on EMC Symmetrix DMX Devices Connected to VMware VI3 Environments*, available at http://powerlink.emc.com/km/live1/en_US/Offering_Technical/White_Paper/H4116-enabling-spc2-compl-emc-symmetrix-dmx-vmware-envnmt-wp.pdf.

A

Migration Considerations

Data migrations from one array to another can be done using SAN-based functionality (using Open Replicator, SAN Copy, MirrorView, SRDF, etc.) or using host-based functionality. This appendix provides pointers for a seamless VMware ESX Server host-based array data migration.

ESX 2.5

For host-based migration for the ESX 2.5, use the **cp** command.

To perform the migration, use the following steps:

1. Present the source and target disk(s) to the ESX Server.
2. Create VMFS on the target disk(s) and assign an appropriate label.
3. Power off the virtual machines.
4. Use the **cp** command to copy virtual disks from the source VMFS to the target VMFS.
5. Remove access to the source disk(s).
6. Rescan the SAN fabric.
7. Re-label the VMFS on the target disk(s) to the original label.
8. Power on the virtual machines.

ESX 3.x

For host-based migration for the ESX 3.x, use the **vmkfstools** command.

To perform the migration, use the following steps:

1. Present the source and target disk(s) to the ESX Server.
2. Create VMFS on the target disk(s) and assign an appropriate label.
3. Use VC to create VMFS since it automatically aligns VMFS volumes.
4. Create directories on the target VMFS to match the source VMFS.
5. Copy the configuration files from the source VMFS to the target VMFS.
6. Power off the virtual machines.
7. Copy the virtual disks using the **vmkfstools** command.
8. Remove access to the source disk(s).
9. Rescan the SAN fabric.
10. Un register virtual machines from the VC.
11. Delete the source VMFS information from the VC database.
12. Re-label the target VMFS to the original source VMFS label name.
13. Re-register and power on the virtual machines.

ESX 3.5

For host-based migration for the ESX 3.5, use storage Vmotion.

Using Storage VMotion, you can migrate a virtual machine and its disk files from one datastore to another while the virtual machine is running. You can choose to place the virtual machine and all its disks in a single location, or select separate locations for the virtual machine configuration file and each virtual disk. The virtual machine does not change execution host during a migration with Storage VMotion.

For details, see the ESX 3.5 *Basic System Administration* manual, available at http://www.vmware.com.

Index