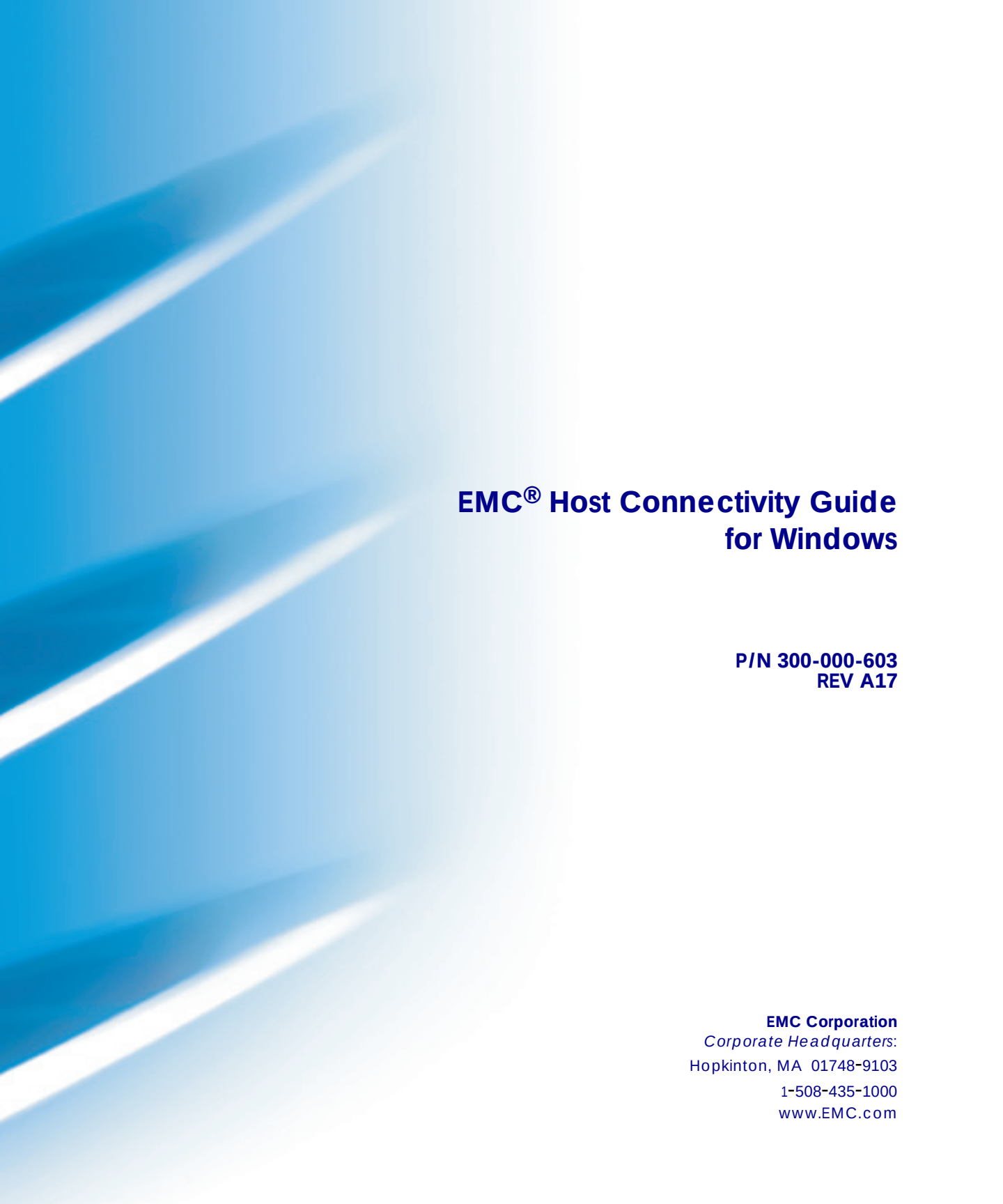




EMC® Host Connectivity Guide for Windows

**P/N 300-000-603
REV A17**

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

Copyright © 2007 EMC Corporation. All rights reserved.

Published May, 2007

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

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

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

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

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

Preface	11
 Chapter 1 General Procedures and Information	
General Windows information	16
Terminology	16
Utilities and functions.....	16
Windows environment.....	17
Hardware connectivity	17
Booting Windows from external storage.....	18
Boot-from-SAN configuration restrictions.....	18
Risks of booting from the storage array	18
How to determine I/O latency and load on the boot	
LUN	19
Boot crashdump save to disk behavior	20
Configuring for Windows 2000/ Windows 2003	21
Configuring the HBA driver	21
Initializing the disks	21
Creating volumes.....	24
System status and error messages.....	26
Adding devices online	27
Editing the Windows I/O timeout value	28
Recovering from drive errors.....	29
Recovery from offline state	29
Rebooting to recover from an offline state.....	31
Recovery from at-risk file system.....	32
Recovery from fail state	33
diskpar and diskpart	34
Microsoft Cluster Server	35
Quorum schemes	35

MSCS in a boot from SAN environment.....	40
Troubleshooting Microsoft cluster issues.....	42
Setting up your networks for the server cluster	42
Configuring TCP/IP settings.....	43
Using a different disk for the quorum resource.....	46
Recovering from a corrupted quorum log or quorum disk	48
Checklist: Installing a physical disk resource	51
Creating a new group in a cluster.....	51
Creating a new resource	53
Best practices for configuring and operating server clusters	54
How to troubleshoot cluster service startup issues.....	59
Cluster disk and drive connection problems	65
Client-to-cluster connectivity problems.....	70
General administrative problems	74
Using Navisphere applications with clusters.....	81
Configuring the Host Agent for clusters.....	81
Configuring the Host Agent	81
Installing the cluster software	82
Adding a generic service resource to the Cluster Group	82

Chapter 2 Fibre Channel Attach Environments

Windows Fibre Channel environment.....	86
Hardware connectivity	86
Boot device support	86
Planning for fabric zoning and connections.....	87
Host configuration with Emulex HBAs	88
Host configuration with QLogic HBAs	89

Chapter 3 iSCSI Attach Environments

Introduction	92
Terminology	92
Software	92
Boot device support	93
Installing the Microsoft iSCSI Initiator	94
Adding or removing components.....	95
Upgrading from version v2.x to v2.03.....	107
Determining the Initiator version	107
Uninstalling the Initiator	108

Using MS iSNS server software with iSCSI configurations	109
Installing iSNS server software	109
Configuring iSNS server	113
Using discovery domains for iSNS	114
iSCSI Boot with the Intel PRO / 1000 family of adapters	119
Preparing your storage array for boot	119
Post installation information	123
Notes on Microsoft iSCSI Initiator	128
Microsoft Cluster Server	128
Dynamic disks	128
Boot	129
NIC teaming	129
Using the Initiator with EMC PowerPath v4.6.x	129
Commonly seen issues	134

Chapter 4 CLARiiON and Symmetrix Information

CLARiiON environment	140
Storage components	140
Required storage system setup	140
Host connectivity	140
LUNZ visibility to host	141
Navisphere Windows host agent and server utility	141
About Navisphere host agent	142
About Navisphere storage system initialization utility	142
Navisphere Management Suite	143
LUN expansion recognition	143
Configuring VERITAS Volume Manager for CLARiiON	145
Symmetrix environment	146
Initial Symmetrix configuration	146
Arbitrated loop addressing	147
Fabric addressing	149
SCSI-3 FCP addressing	150

Chapter 5 PowerPath for Windows

PowerPath and PowerPath iSCSI	154
PowerPath for Windows	155
PowerPath and MSCS	155
Integrating PowerPath into an existing MSCS cluster	155

	EMC PowerPath verification and problem determination.....	158
	Problem determination.....	160
	Making changes to your environment	163
	PowerPath messages.....	164
	PowerPath iSCSI 1.1 for Windows	165
	PowerPath iSCSI 1.1 features.....	165
	PowerPath iSCSI 1.1 restrictions	165
Appendix A	Persistent Binding	
	Understanding persistent binding	170
	Methods of persistent binding.....	173
Appendix B	Solutions Enabler	
	EMC Solutions Enabler	176
Appendix C	General Host Behavior and Limitations	
	Issues	178
	Capabilities and limitations	179
	Operating system / driver capabilities and limitations	179
	How a server responds to a failure in the boot LUN path	183
Appendix D	VERITAS Volume Management Software	
	VERITAS Volume Management software	186
	VERITAS Storage Foundation 5.0.....	186
	VERITAS Storage Foundation 4.3.....	187
	VERITAS Storage Foundation 4.2.....	188
	VERITAS Foundation Suite 4.1	188
	VERITAS Volume Manager 3.1 and VERITAS DMP	188
	VERITAS Volume Manager 3.0	189
	Index.....	191

	Title	Page
1	Disk Management in Computer Management window	22
2	Upgrade to Dynamic Disk	23
3	Healthy volumes and file systems example	25
4	Recovery from offline state	30
5	Offline state requires reboot	31
6	Recovery from at-risk file system	32
7	Recovery from fail state	33
8	MSCS using legacy quorum scheme	36
9	MSCS using MNS quorum scheme	37
10	Software Update Installation Wizard	94
11	iSCSI Initiator Properties window	96
12	iSCSI Initiator Properties, Discovery tab	98
13	Add Target Portal dialog box	99
14	Add iSNS Server dialog box	99
15	iSCSI Initiator Properties, Targets tab	100
16	Log On to Target dialog box	101
17	Advanced Settings window	102
18	iSCSI Initiator Properties, Targets tab	104
19	iSCSI Initiator Properties, Persistent Targets tab	105
20	iSCSI Initiator Properties, Bound Volumes / Devices tab	106
21	Microsoft iSNS Server Installation Wizard	110
22	Installation option dialog	111
23	iSNS DHCP configuration option dialog	111
24	Installation confirmation message	112
25	iSNS General properties	113
26	Target/Initiator Details dialog	114
27	iSNS Server Properties, Discovery Domains tab	115
28	Create Discovery Domain dialog	115
29	iSNS Server Properties, Discovery Domain with members added	116
30	Add registered Initiator or Target to Discovery Domain	117

31	iSNS Server Properties, Discovery Domain Sets tab	118
32	Four paths	124
33	EMC PowerPathAdmin	125
34	iSCSI Initiator Properties dialog box	131
35	Log On to Target dialog box	132
36	Advanced Settings dialog box	133
37	Four paths	134
38	PowerPath Administration icon	158
39	PowerPath Monitor Taskbar icons and status	158
40	One path	159
41	Multiple paths	160
42	Error with an Array port	162
43	Failed HBA path	163
44	Management Snap-In	166
45	Original configuration before the reboot	172
46	Host after the rebooted	172
47	Lun Mapping and Automatic Lun Mapping	180

Tables

	Title	Page
1	FC-AL addressing parameters	149
2	Symmetrix SCSI-3 addressing modes	151
3	Possible failure states	161
4	PowerPath icons	167
5	Server response to failure in the boot LUN path (Single fault)	183

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 Windows 2000 and Windows 2003 host interfaces to EMC storage arrays over Fibre Channel or iSCSI.

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:

- ◆ Symmetrix or CLARiiON system operation
- ◆ Windows 2000 or Windows 2003 operating environment



IMPORTANT

Always consult the *EMC Support Matrix*, available through E-Lab Interoperability Navigator at: <http://elabnavigator.EMC.com> or at <http://www.Powerlink.com>, for the most up-to-date information.

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 font	In running text: • Interface elements (for example, button names, dialog box names) outside of procedures • Items that user selects outside of procedures • Java classes and interface names • Names of resources, attributes, pools, Boolean expressions, buttons, DQL statements, keywords, clauses, environment variables, filenames, functions, menu names, utilities • Pathnames, URLs, filenames, directory names, computer names, links, groups, service keys, file systems, environment variables (for example, command line and text), notifications
bold	• User actions (what the user clicks, presses, or selects) • Interface elements (button names, dialog box names) • Names of keys, commands, programs, scripts, applications, utilities, processes, notifications, system calls, services, applications, and utilities in text
<i>italic</i>	• Book titles • New terms in text • Emphasis in text
Courier	• Prompts • System output • Filenames • Pathnames • URLs • Syntax when shown in command line or other examples
Courier, bold	• User entry • Options in command-line syntax
<i>Courier italic</i>	• Arguments in examples of command-line syntax • Variables in examples of screen or file output • Variables in pathnames

<>	Angle brackets for parameter values (variables) supplied by user.
[]	Square brackets for optional values.
	Vertical bar symbol for alternate selections. The bar means or.
...	Ellipsis for 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

General Procedures and Information

This chapter provides general procedures and information regarding Windows hosts .

◆ General Windows information.....	16
◆ Windows environment.....	17
◆ Booting Windows from external storage	18
◆ Configuring for Windows 2000 / Windows 2003	21
◆ Editing the Windows I/O timeout value	28
◆ Recovering from drive errors	29
◆ diskpar and diskpart	34
◆ Microsoft Cluster Server	35
◆ Troubleshooting Microsoft cluster issues	42

General Windows information

This section provides information that is common to all supported versions of Windows. Please read the entire section before proceeding to the rest of the chapter.

Terminology

You should understand these terms:

- ◆ **Free space** — An unused and unformatted portion of a hard disk that can be partitioned or subpartitioned.
- ◆ **Partition** — A portion of a physical hard disk that functions as though it were a physically separate unit.
- ◆ **Volume** — A partition or collection of partitions that have been formatted for use by a file system. A volume is assigned a drive letter.
- ◆ **Primary partition** — A portion of a physical disk that can be marked for use by an operating system. A physical disk can have up to four primary partitions. A primary partition cannot be subpartitioned.

Utilities and functions

Here are some Windows functions and utilities you can use to define and manage Symmetrix®/CLARiiON® devices. The use of these functions and utilities is optional; they are listed for reference only:

- ◆ **Disk Manager** — Graphical tool for managing disks; for example, partitioning, creating, and deleting volumes.
- ◆ **Registry Editor** — Graphical tool for displaying detailed hardware and software configuration information. Not normally part of the Administrative Tools group, the registry editor, REGEDT.EXE, can be found in the Windows \system32 subdirectory.
- ◆ **Event Viewer** — Graphical tool for viewing system or application errors.

Windows environment

This section lists Fibre Channel support information specific to the Windows environment.

For more information, refer to the appropriate chapter:

- ◆ [Chapter 2, "Fibre Channel Attach Environments"](#)
- ◆ [Chapter 3, "iSCSI Attach Environments"](#)

Hardware connectivity

Refer to the *EMC Support Matrix*, available through E-Lab™ Interoperability Navigator at: <http://elabnavigator.EMC.com> or on <http://www.Powerlink.com>, or contact your EMC representative for the latest information on qualified hosts, host bus adapters, and connectivity equipment.

EMC does not recommend mixing HBAs from different vendors in the same host.

Booting Windows from external storage

Windows hosts have been qualified for booting from EMC array devices interfaced through Fibre Channel as described under "Boot Device Support" in the *EMC Support Matrix*.

Refer to the appropriate Windows HBA Guide, available on <http://Powerlink.EMC.com>, for information on configuring your HBA and installing the Windows operating system to an external storage array.

Boot-from-SAN configuration restrictions

Refer to the *EMC Support Matrix* for any specific boot-from-SAN restrictions. This guide no longer contains restriction information, and the information in the *EMC Support Matrix* supersedes any restriction references found in previous HBA installation guides.

Risks of booting from the storage array

EMC recommends that you do not boot a Windows host from the storage array. However, if it is necessary to use the storage array as a boot disk, EMC recommends shutting down the host server during any maintenance procedures that could make the boot disk unavailable to the host.



CAUTION

Microsoft Windows 2000/Windows 2003 uses virtual memory paging files that reside on the boot disk. If the paging file becomes unavailable to the memory management system when it is needed, the operating system will crash with a blue screen.

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

- ◆ Lost connection to array (pulled or damaged cable connection)
- ◆ Array service/upgrade procedures, such as on-line microcode upgrades and/or configuration changes
- ◆ Array failures, including failed lasers on Fibre Channel ports
- ◆ Array power failure

- ◆ Storage Area Network failures, such as Fibre Channel switches, switch components, or switch power failures
- ◆ Storage Area Network service / upgrade procedures, such as firmware upgrades or hardware replacements

Note: EMC recommends moving the Windows virtual memory paging file to a local disk when booting from the storage array. Consult your Windows manual for instructions on how to move the paging file.

How to determine I/O latency and load on the boot LUN

The current restrictions for boot-from-array configurations listed in the *EMC Support Matrix* represent the maximum configuration that is allowed using typical configurations. There are cases where your applications, host, array, or SAN may already be utilized to a point when these maximum values may not be achieved. Under these conditions, you may wish to reduce the configuration from the maximums listed in the *EMC Support Matrix* for improved performance and functionality.

Here are some general measurements that can be used to determine if your environment may not support the maximum allowed boot-from-array configurations:

- ◆ Using the Windows Performance Monitor, capture and analyze the Physical Disk and Paging File counters for your boot LUN. If response time (sec / operation), or disk queue depth seem to be increasing over time, you should review any additional loading that may be affecting the boot LUN performance (HBA / SAN saturation, failovers, ISL usage, and so forth).
- ◆ Use available Array Performance Management tools to determine that the array configuration, LUN configuration and access is configured optimally for each host.

Possible ways to reduce the load on the boot LUN include:

- ◆ Move application data away from the boot LUN.
- ◆ Reduce the number of LUNs bound to the same physical disks.
- ◆ Select an improved performance RAID type.
- ◆ Contact your EMC support representative for additional information.

Boot crashdump save to disk behavior

If your system is configured to write crashdumps after system failures and the host is configured to boot from the array, you will be able to successfully save the crashdump only on the original available boot device path on which the system started. This is a Windows limitation, and installing PowerPath® will not affect this behavior. At the time a system crash is to be written to disk, Windows has already saved the original boot path, and PowerPath cannot redirect the crashdump file (MEMORY.DMP) to an alternative available device. If you have a configuration for which you want to capture a crashdump, you should ensure that the original primary boot path is available at the time of the crash.

Configuring for Windows 2000/Windows 2003

This section provides information for configuring Windows 2000/Windows 2003 with EMC storage systems.

Configuring the HBA driver

Driver software for Windows hosts is provided with each HBA. Standard Windows driver installation procedures enable use of the HBAs. The driver for the host bus adapter must be installed and configured prior to partitioning the disks. Refer to the appropriate section of [Chapter 2](#) or [Chapter 3](#).

Initializing the disks

To view the drives seen by your host, select **Start, Programs, Administrative Tools, Computer Management** from the Windows desktop. A window with all the physical disks connected to your server (along with its partitions) similar to [Figure 1 on page 22](#) appears.

Select **Disk Management** from the list on the left. In [Figure 1 on page 22](#), the disks have not been formatted or partitioned.

To create volumes and format the disks, you must write signatures on all the disks to make them recognizable by Windows 2000/Windows 2003.

Note: Windows 2000 supports a maximum file system size of 2 TB .

Windows 2003 supports a maximum file system size of 2 TB unless Service Pack 1 or 2 is installed. With SP1 or SP2, the maximum supported physical disk size is 256 TB. Volumes larger than 2 TB must use GPT partitions to support them. Refer to your Windows users guide for information on GPT partitions.)

EMC requires SP1 or SP2 to be on the system.

These limitations should be taken into account when planning your storage environment.

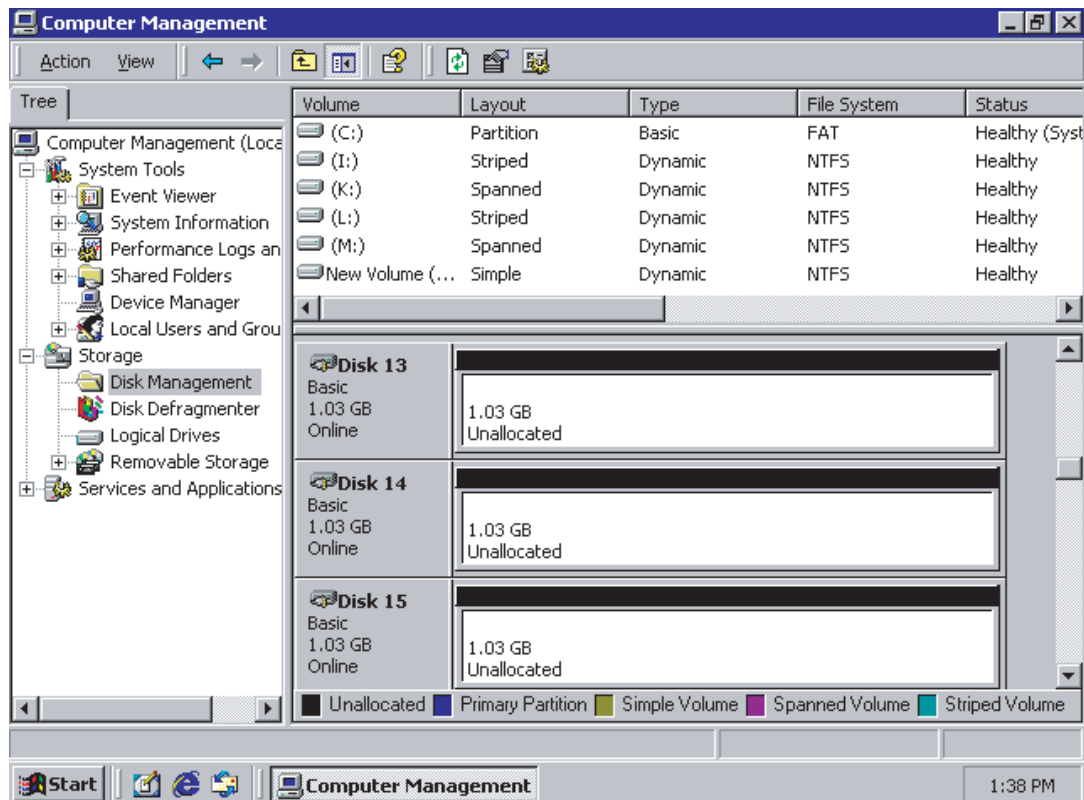


Figure 1 Disk Management in Computer Management window

Disk types

Windows 2000 and Windows 2003 group all disks into two types: *basic* and *dynamic*. This convention is derived from enhancements to the Windows file systems.

Basic disks can be formatted with FAT, FAT32, and NTFS file systems.

Basic disks cannot be used to create striped, spanned, or RAID volumes using Windows Disk Manager. Windows 2000 and Windows 2003 require that basic disks be upgraded to a dynamic status before being used for striped, spanned, or RAID volumes. These volumes are not changed even though the disks that make them up are basic disks.

Refer to [“Upgrading to dynamic disks” on page 23](#).

Upgrading to dynamic disks

In order to use the Windows 2000 or Windows 2003 file systems (spanned, striped, and RAID), you must upgrade your disks to dynamic disks. If you have a basic disk with a partition, you must have at least 2 MB of unallocated free space on the disk to perform the upgrade.



IMPORTANT
Dynamic disks cannot be used for clusters.

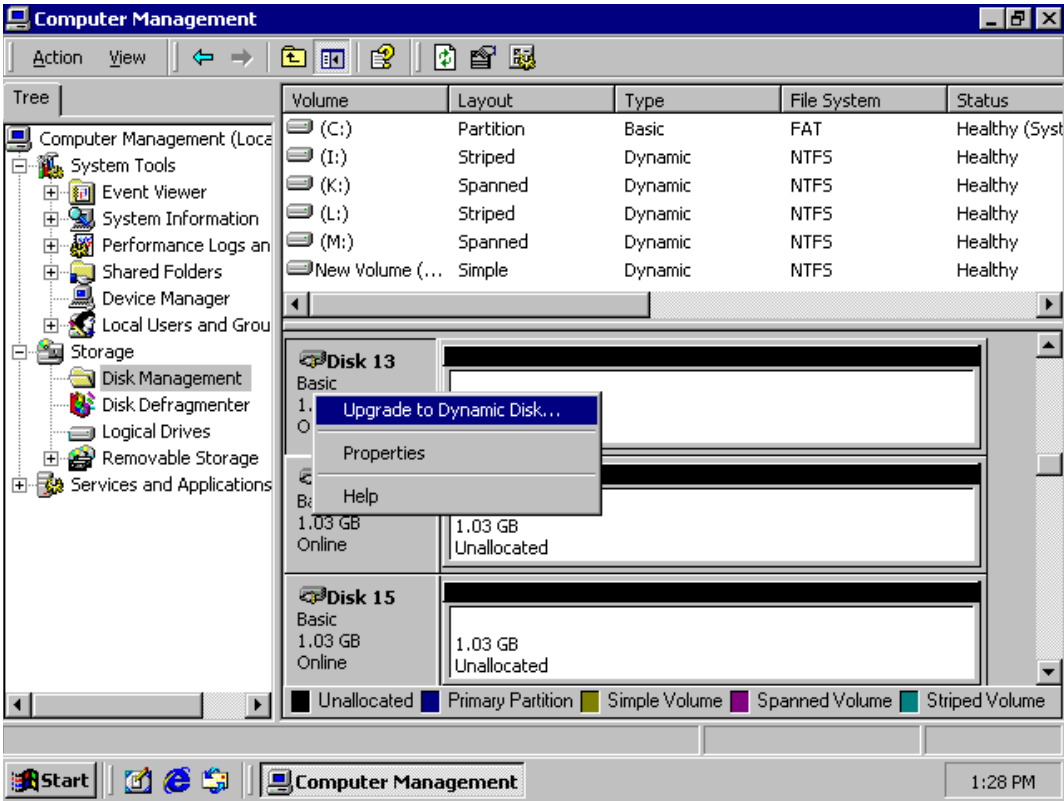


Figure 2 Upgrade to Dynamic Disk

Select the disk to upgrade. In the example in [Figure 2](#), Disk 13 has been selected. Right-click to open the pop-up menu shown in [Figure 2](#). Windows 2000 and Windows 2003 file systems can be used after upgrading.

Creating volumes

Windows 2000 and Windows 2003 allow you to create up to four partitions in the free space of a physical hard disk, create multiple logical drives in the free space, and delete partitions.

You can use `Create Volume Wizard` to create the following:

- ◆ **Spanned disk volume** — Windows 2000 and Windows 2003 create spanned volumes from unpartitioned free space on one or more drives.
- ◆ **Striped volume without parity** — Windows 2000 and Windows 2003 create striped volumes from unpartitioned free space on two or more drives. Unlike spanned volumes, each member partition of the striped volume must be on a different disk. Disk Manager makes partitions on all disks approximately the same size. Up to 32 disks can participate in a striped volume.
- ◆ **Striped volume with parity** — Striped volumes with parity require a minimum of three partitions on different drives.
- ◆ **Host-based mirrored volume** — The procedures for each of the four volumes are explained by the Wizard. For more information about these volumes, click **Help**.

Creating simple, spanned, and stripe volumes

The following steps partition and format one disk at a time:

1. Right-click the Symmetrix/CLARiiON disk you want to partition.
2. Select **Create Volume** from the menu.
3. In the Wizard program, select **Simple Volume**, **Spanned Volume**, or **Striped Volume**.

Note: **Mirrored** and **RAID 5** are not good choices here, because you can mirror disks and create RAID 5 volumes more efficiently on the Symmetrix/CLARiiON system.

4. Enter the size of the partition, or do nothing if you are using the whole disk.
5. Click **OK** to continue the operation.
6. Assign a drive letter and click **OK**.
7. Select the file system type (NTFS or FAT) from the File System menu and type in a label for the partition and click **OK**.

8. Check **Perform a Quick Format** to prevent excessively long waits.
9. Read the summary, and click **OK**.

Disk Manager displays the successful creation of the new drives and volumes like those shown in [Figure 3](#). The color identifies the type of volume.

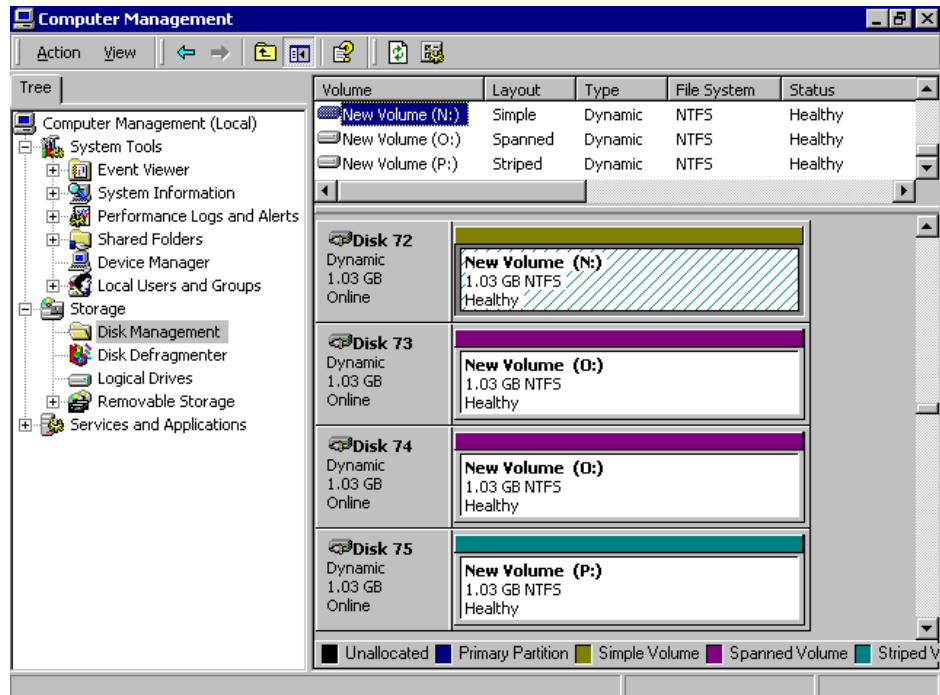


Figure 3 Healthy volumes and file systems example

If file systems are not reported as healthy, refer to [“Symmetrix environment” on page 146](#).

Deleting partitions

Back up the contents of any partition, volume, or logical drive before you delete it:

1. Right-click the partition, volume, or drive to select it.
2. Select **Delete** from the pop-up menu.

Note: Windows warns you that all data will be lost and asks you to confirm.

3. Click **Yes** to delete the selected item. That item becomes free space.

System status and error messages

In a Windows 2000 or Windows 2003 environment you can view, save, or clear system status and error messages. The Windows host logs all errors it detects into a system event log.

To open the event log:

1. Click the **Administrative Tools** icon.
2. Select the **Event Viewer** option.
3. From here:
 - To view an event, double-click the event in the list, or select **View, Details** from the menu bar.
 - a. To save the contents of the system event log:
 - b. Select **Action** on the menu bar
 - c. Click **Save As**, and enter the full path of the location to which the file is to be saved.
 - To clear the contents of the system event log:
 - a. Select **Action** on the menu bar.
 - b. Select **Clear All Events**.
 - c. Click **No** when prompted to save the events.

Note: The **Event Viewer** can also be accessed from the **Computer Management** utility under **System Tools**.

Adding devices online

Whenever devices are added online to the system or the device channel addresses are changed, you must perform the actions described below in order to introduce the new devices to the system.

To add new disk devices while the system remains on line:

1. Add the new disk devices into the storage array configuration using the appropriate storage management software.



IMPORTANT

Always consult the EMC PSE Configuration Group for assistance when working with a system with live data.

2. Run the Disk Management utility and perform a disk rescan, in either of these ways:
 - Right-click the **Disk Management** icon and select **Rescan Disks**.
 - Select **Disk Management**, and **Rescan Disks** from the **Action** menu.

If Disk Manager does not see the new drives, you must reboot the host.

Editing the Windows I/O timeout value

Note: While the EMC-approved drivers for QLogic and Emulex HBAs automatically set the Windows disk timeout value to 60 seconds, some software applications may change it to other values for their own purposes. The follow section describes how to manually set the Windows disk timeout value to 60 seconds.

Some Windows configurations, especially those using the Symmetrix system as a boot device, periodically encounter I/O timeouts, which can cause unusual error messages and performance degradation. To avoid this potential problem, EMC recommends increasing the Windows default I/O timeout value, as follows:

1. Open the Windows registry editor:
 - a. On the Windows taskbar, click **Start**.
 - b. Select **Run**.
 - c. Type **regedt.exe** in the **Open** field, and click **OK**.
2. Follow this path: **HKEY_LOCAL_MACHINE, System, CurrentControlSet, Services, Disk**.
3. Look for **TimeOutValue** in the right pane of the registry editor:
 - If the **TimeOutValue** exists, double-click it, and go to step 4.
 - If the **TimeOutValue** does *not* exist:
 - a. Select **Add Value** from the **Edit** menu.
 - b. In the **Value Name** box, type **TimeOutValue** (exactly as shown).
 - c. For the data type, select **REG_DWORD** from the pull-down menu.
 - d. Click **OK**, and go to step 4.
4. In the **DWORD Editor** window:
 - a. Click **decimal** in the **radix** box.
 - b. Change the value in the **data** box to **60**.
5. Click **OK**.
6. Close the registry editor.
7. Reboot the host.

Recovering from drive errors

This section provides information on recovering from drive errors from Windows 2000 / Windows 2003 hosts. The following sections describe various drive errors and the recovery procedure:

- ◆ [“Recovery from offline state”](#), next
- ◆ [“Rebooting to recover from an offline state”](#) on page 31
- ◆ [“Recovery from at-risk file system”](#) on page 32
- ◆ [“Recovery from fail state”](#) on page 33

Recovery from offline state

When a disk failure occurs from a disconnected cable or a failed storage director, the drives on a Windows 2000 or Windows 2003 system may be put into an offline state. (You can view this by going to the Disk Management utility.) This is a normal behavior that guards against data errors.

To recover from this condition, you must reactivate one offline drive and then perform a disk rescan. (Refer to [Figure 4 on page 30](#).)

This brings up all of the offline drives at once.

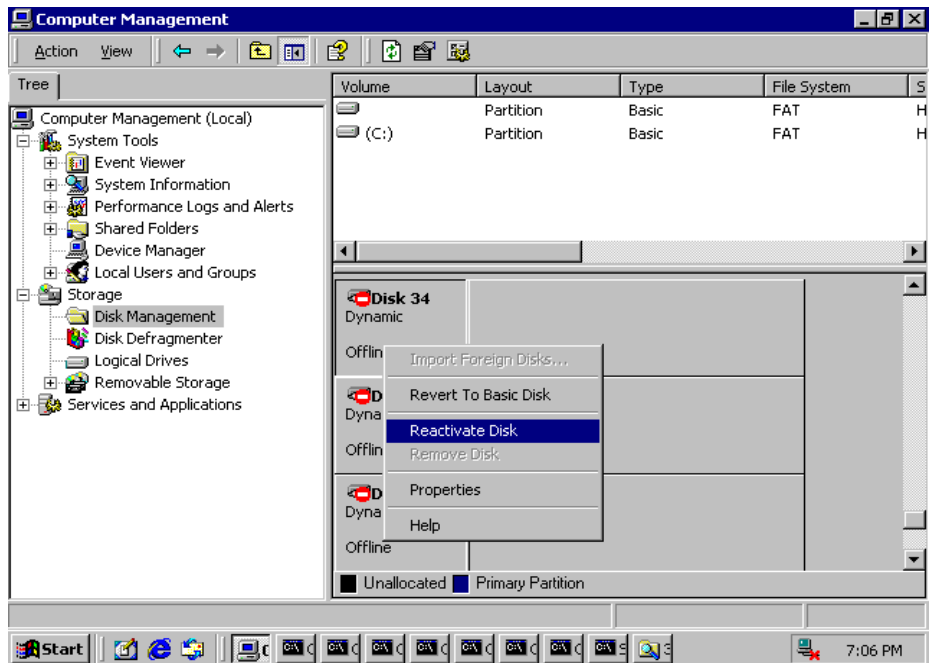


Figure 4 Recovery from offline state

Rebooting to recover from an offline state

If a drive is offline for an extended amount of time, Windows 2000 / 2003 will not let you reactivate the drive. In this case, you must reboot the host to reactivate the drive. (Refer to [Figure 5](#); you cannot select **Reactivate Volume** from the menu shown here, so you must reboot.)

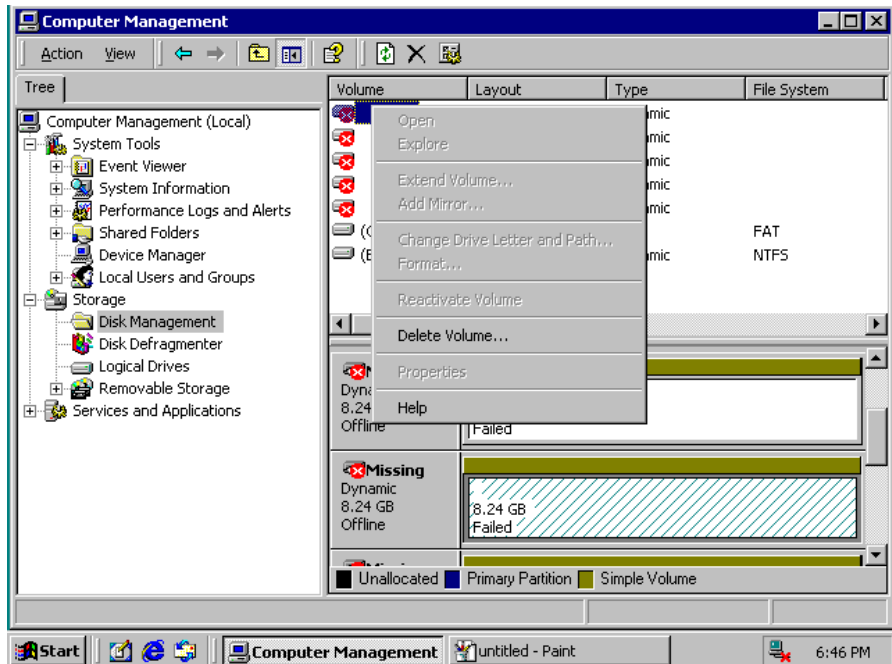


Figure 5 Offline state requires reboot

Recovery from at-risk file system

To recover from an *at-risk file system*, you must reactivate each individual disk. (Refer to [Figure 6](#).)

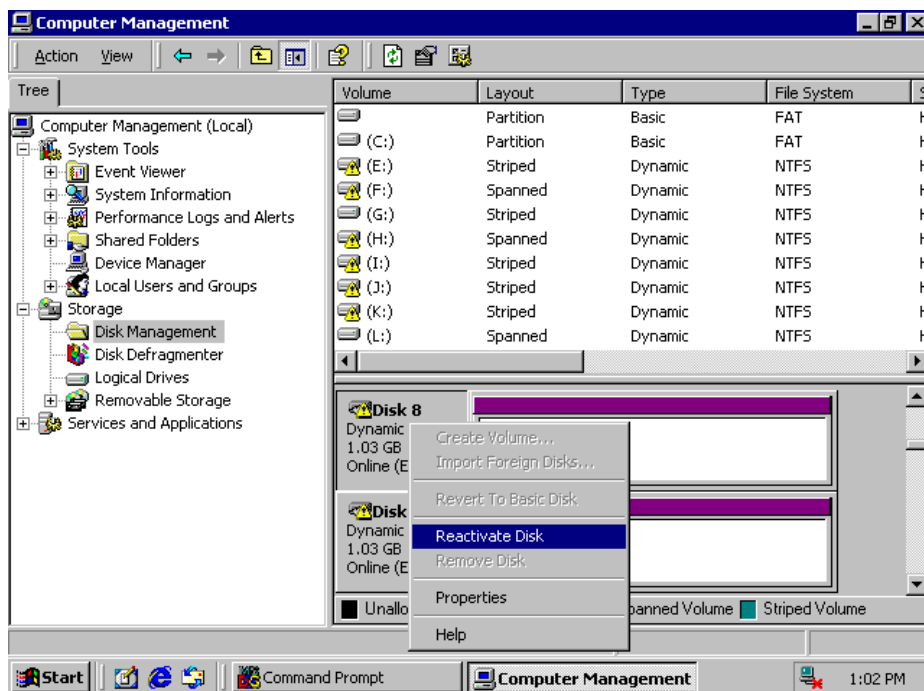


Figure 6 Recovery from at-risk file system

Recovery from fail state

Another case is where the drive is in a *fail* state. To recover, activate the failed volume. (Refer to [Figure 7](#).)

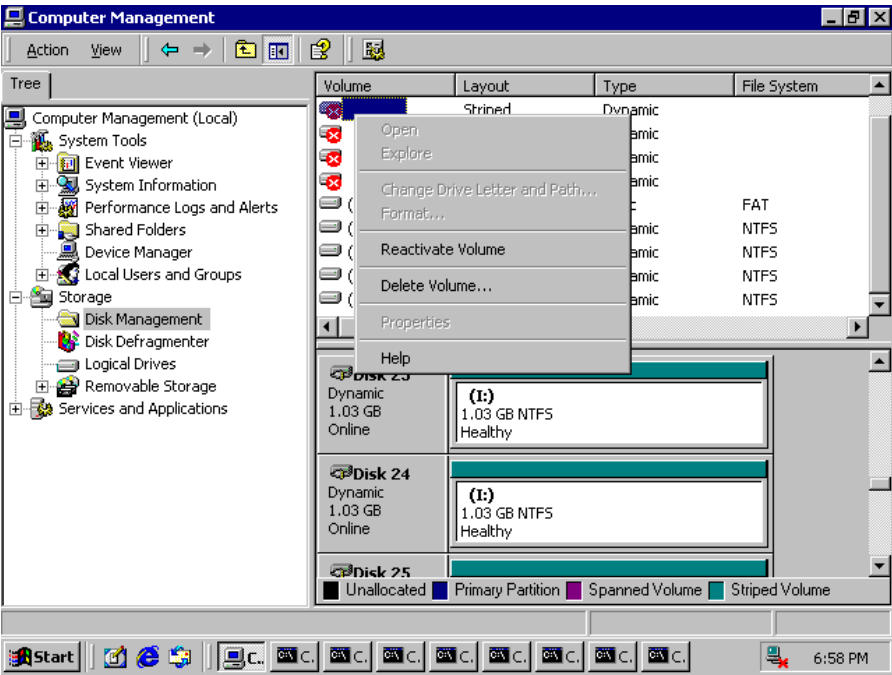


Figure 7 Recovery from fail state

diskpar and diskpart

Current versions of Microsoft Windows Server do not correctly detect disk drive geometry of devices presented from storage area network (SAN) arrays, such as EMC® Symmetrix and CLARiiON. The detection results in a geometry of 63 sectors per track, whereas the correct form of geometry should be 64 sectors per track. Creating partitions and the volumes that they represent without modification will result in suboptimal I/O processes.

In shared storage arrays, inefficient utilization of storage components may not only adversely affect the performance of the host generating the I/O, but also other hosts that share common components.

Optimally configured arrays ensure that all components are used efficiently. Misaligned Windows file systems can result in significant performance degradation. It is not uncommon to see a performance improvement of 20 percent or more (depending on layout and I/O load). As a result of resolving misaligned I/O and its overheads, a more efficient utilization of system resources will occur. This can translate into positive improvements in I/O response for other servers on the shared array.

Two programs are available for aligning partitions. The `diskpar` utility has been around for some time and works on all Windows Systems. The `diskpart` utility has been available since Windows 2000, but could not align partitions until an updated version was included with Windows 2003 SP1. The version of `diskpart` that can align partitions is version 5.2.3790 or higher.

For information on using `diskpar` and `diskpart` for the alignment of disk partitions, refer to the EMC Engineering white paper *Using diskpar and diskpart to Align Partitions on Windows Basic and Dynamic Disks*, available on Powerlink®. This white paper describes how to use `diskpar` and `diskpart` to align partitions on Basic Disks and Dynamic Disks under both Windows 2000 and Windows 2003. The paper provides extensive analysis to demonstrate how `diskpar` works and the limitations of using it.

Microsoft Cluster Server

Microsoft Cluster Service (MSCS) is the native server clustering solution included as part of both Windows 2000, Advanced Server and Datacenter Editions, and Windows Server 2003, Enterprise and Datacenter Editions. A server cluster is used to provide failover support for applications and services.

A server cluster can consist of a varying number of nodes, based on the OS version:

- ◆ Windows 2000 Advanced Server — 2 nodes
- ◆ Windows 2000 Datacenter Server — 4 nodes
- ◆ Windows Server 2003 Enterprise and Datacenter — 8 nodes

Typically, each node is attached to one or more cluster storage devices. Cluster storage devices allow different servers to share the same data. By reading this data, they provide failover for resources.

All EMC arrays provide support for Microsoft Cluster Services over both Fibre Channel and iSCSI attach interfaces.

For a detailed overview and description of MSCS implementations in Windows, refer to

<http://www.microsoft.com/windowsserver2003/techinfo/overview/clustering.mspx>.

Quorum schemes

This section provides information on the differences between shared disk and majority node set and quorum schemes.

Shared disk

Prior to Windows Server 2003, MSCS implemented a single quorum scheme, known as *shared disk* quorum. In this quorum scheme, the quorum device is a configuration database for MSCS, using a quorum *log file* that is located on a LUN residing on a shared storage interconnect, accessible by all member hosts of the cluster (shown in [Figure 8 on page 36](#)).

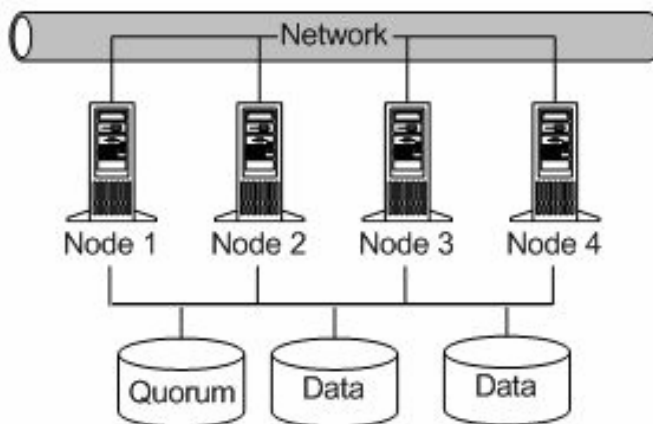


Figure 8 MSCS using legacy quorum scheme

Majority node set (MNS)

Introduced in Windows Server 2003, *majority node set* (MNS) quorum is a new quorum scheme used in a cluster configuration with at least 3 nodes, instead of the classic *shared disk* quorum scheme. The MNS quorum scheme uses quorum data that is locally stored on the system disk of each member host of the cluster, so each member host has a local copy of the quorum data. From a cluster-wide perspective, the individual quorum data stores residing on each member host is transparent; the quorum data appears as a single resource.

In an MNS configuration, a dedicated MNS resource manages the quorum data residing locally on each member host and verifies that each locally-kept copy of the quorum data is consistent across all member nodes (shown in [Figure 9 on page 37](#)).

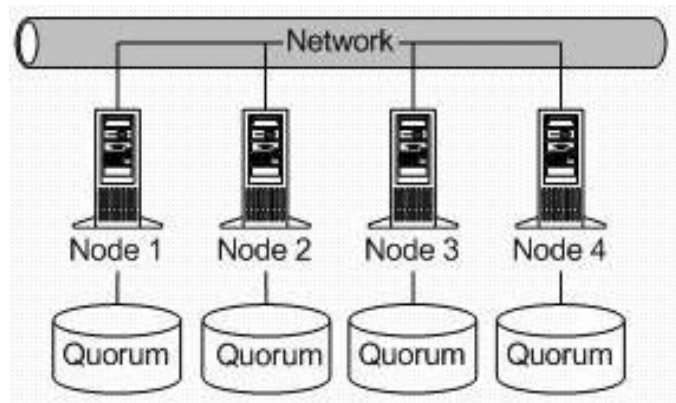


Figure 9 MSCS using MNS quorum scheme

While the disks that make up the MNS can (in theory) be disks on a shared storage fabric, the MNS implementation uses a directory on each node's local system disk to store the quorum data. If the configuration of the cluster changes, that change is reflected across the different disks.

The change is only considered to have been "committed" (that is, made persistent) if that change is made to $(\text{Number of nodes configured in the cluster} / 2) + 1$ nodes of the cluster.

This ensures that a majority (>50%) of the nodes have an up-to-date copy of the quorum data. The cluster service itself will only start up (and bring resources online) if a majority of the nodes configured as part of the cluster are up and running the cluster service. If there are fewer nodes, the cluster is said not to have quorum, and waits (based on the restart properties of the Cluster service) until more nodes attempt to join. Only when a majority (or quorum) of nodes are available, will the cluster service start and bring the resources online. In this way, because the up-to-date configuration is written to a majority of the nodes regardless of node failures, the cluster will always guarantee that it starts up with the latest, most up-to-date configuration.

In the case of a failure or split-brain in an MNS-based cluster, all MSCS instances across all member nodes that do not contain a majority of nodes are terminated. This ensures that if there is an MSCS instance running that contains a majority of the nodes, it can safely start up any resources that are not running on that host safely.

Using shared disk versus MNS quorum schemes

(It can be the only partition in the cluster that is running resources because all other MSCS instances on other hosts are terminated).

Standard shared disk quorum schemes are suitable for the majority of situations that users will require, such as applications that require data to be highly available at a single site (generally utilizing a shared storage array).

MNS-based clusters are ideal for customers with specialized requirements, such as the following:

- ◆ Geographically dispersed cluster environments which do not use third-party replication schemes (such as SRDF/CE for MSCS) to remotely mirror a shared disk quorum and/or data resource. This environment requires a specialized implementation or application specific method to keep data consistent across member hosts.
- ◆ Cluster environments which do not need shared storage, such as applications that do not store persistent data, but need to provide consistent volatile data that is replicated across all member hosts. (Most applications used in a Windows environment do not fall into this category.)

Shared-disk data resources can still be used in a MNS-based cluster; MNS removes the requirement of a shared disk quorum resource only. Shared data disks can still be used in a MNS cluster. For example, clustered applications such as SQL or Exchange can use (and, in fact, require) shared disk storage for data.

Installing and configuring MSCS using shared disk and MNS quorum schemes

The quorum scheme used in an MSCS configuration is chosen during MSCS installation. Installation of a shared disk quorum, and a majority node set quorum is nearly identical, with the exception of modifying the *Quorum resource type* during install

Use the following procedure to configure MSCS on a Windows Server 2003 host:

1. Start **Cluster Administrator**.
2. Under **File**, select **New... Cluster**, and press **Next** in the **New Server Cluster Wizard**.
3. Select the domain name in which the cluster will be created and type a unique name for the cluster.
4. Press **Next** to continue.

5. Enter the name of the system which will be the first node in the cluster. By default, the local server name is entered.
 - If this is correct, press **Next** to continue.
 - If this is *not* correct, type or select the system name to be the first cluster node and press **Next** to continue.

The **New Cluster Server Wizard** begins analyzing the system configuration for cluster feasibility. (This may take several minutes.) If an issue is found, the details of the issue will be detailed on this page. The issue must be corrected before the Wizard allows you to continue. When a **Tasks Completed** message with a green progress bar displays, the analysis has completed. Press **Next** to continue.

6. Enter an IP address that the cluster name will use for network access and press **Next** to continue.
7. Enter the username and password of an authorized account which will be used to manage the cluster. This account must be a member of the domain selected earlier.
8. Press **Next** to continue.
9. The next page reports a **Proposed Cluster Configuration**. Press the **Quorum** button to select the quorum resource or resource type.
 - For a shared-disk quorum scheme, a shared quorum disk is specified (by drive letter / mount point).
 - For a majority node set quorum scheme, select **Majority Node Set** in the drop-down menu. Press **OK** to exit the quorum dialog and press **Next** to continue.

The Wizard begins the cluster configuration. (This may take several minutes.) If an issue is found, the details of the issue will be detailed on this page. The issue must be corrected before the Wizard allows you to continue. When a **Tasks Completed** message with a green progress bar displays, the cluster configuration has completed. Press **Next** to continue.

10. The final screen reports that the **New Server Cluster Wizard** has completed successfully. Press **Finish** to exit the wizard.

The cluster, and its associated resources, can now be managed normally via Cluster Administrator.

MSCS in a boot from SAN environment

This section contains information that should be considered when in a "boot from SAN" environment.

By default, MSCS does not allow configurations where the host is booted from a SAN-attached device, where the boot device (and in most cases, pagefile), and shared cluster devices reside on the same logical bus. For example, installing MSCS on a host where both the SAN-based boot disk and shared disks to be clustered are visible over the same HBA(s), may fail with the following error in Clcfsrv.log:

The cluster cannot manage physical disks that are on the same storage bus as the volume that contains the operating system because other nodes connected to the storage bus cannot distinguish between these volumes and volumes used for data.

This limitation is by design, as cluster-specific disk ownership / arbitration routines may use a variety of LUN, target, or bus resets that can cause delays on any devices attached to the bus. When the boot device is attached to the same bus as the shared cluster disks, such resets could be catastrophic to the OS since they could potentially delay access to the boot device and /or pagefile, resulting in a host crash.

Microsoft provides a registry-based modification for Windows Server 2003 environments which allows the boot device and shared cluster devices to reside on the same logical bus. This modification is provided for those customers who understand the reasons for the original limitation and have mitigated much of the risk of having a boot device and shared cluster disk devices on the same logical bus by implementing STORPort-based HBA miniport drivers.

The STORPort-based miniport drivers allow targeted LUN resets, affecting only the shared cluster devices. This leaves the boot device unaffected, as opposed to bus-wide resets which occur in SCSIPort-based miniport or full-port HBA driver implementations.

Note: STORPort miniport drivers are unavailable in Windows 2000 Server environments.

Refer to the following Microsoft Knowledge Base articles for details on using this registry modification in Windows Server 2003 environments:

<http://support.microsoft.com/kb/886569/en-us>

Note that even in environments using STORPort-based HBA miniport drivers, bus-wide reset may still occur as a "last resort" in rare circumstances, if both the LUN-specific reset, and the subsequent array target reset fails for any reason. Refer to the following Microsoft Knowledge Base article for details regarding the MSCS disk arbitration reset scheme in SCSIPort and STORPort-based miniport environments, respectively:

<http://support.microsoft.com/kb/309186/en-us>

<http://support.microsoft.com/kb/301647/en-us>



IMPORTANT

EMC strongly recommends using the STORPort miniport drivers available for all supported HBAs, in addition to the latest EMC-qualified Microsoft STORPort QFE, for all Windows Server 2003-based installations, including both clustered and non-clustered environments.

Troubleshooting Microsoft cluster issues

This section contains the following information about troubleshooting Microsoft cluster issues:

- ◆ [“Setting up your networks for the server cluster” on page 42](#)
- ◆ [“Configuring TCP / IP settings” on page 43](#)
- ◆ [“Using a different disk for the quorum resource” on page 46](#)
- ◆ [“Recovering from a corrupted quorum log or quorum disk” on page 48](#)
- ◆ [“Checklist: Installing a physical disk resource” on page 51](#)
- ◆ [“Creating a new group in a cluster” on page 51](#)
- ◆ [“Creating a new resource” on page 53](#)
- ◆ [“Best practices for configuring and operating server clusters” on page 54](#)
- ◆ [“How to troubleshoot cluster service startup issues” on page 59](#)
- ◆ [“Cluster disk and drive connection problems” on page 65](#)
- ◆ [“Client-to-cluster connectivity problems” on page 70](#)
- ◆ [“General administrative problems” on page 74](#)
- ◆ [“Using Navisphere applications with clusters” on page 81](#)
- ◆ [“Configuring the Host Agent for clusters” on page 81](#)
- ◆ [“Configuring the Host Agent” on page 81](#)
- ◆ [“Installing the cluster software” on page 82](#)
- ◆ [“Adding a generic service resource to the Cluster Group” on page 82](#)

Setting up your networks for the server cluster

Follow the guidelines in this section to reduce network problems in your server cluster:

- ◆ Use identical network adapters in all cluster nodes, that is, make sure each adapter is the same make, model, and firmware version.

- ◆ Use at least two interconnects. Although a server cluster can function with only one interconnect, at least two interconnects are necessary to eliminate a single point of failure and are required for the verification of original equipment manufacturer (OEM) clusters.
- ◆ Reserve one network exclusively for internal node-to-node communication (the private network). Do not use teaming network adapters on the private networks.
- ◆ Reserve one network exclusively for internal node-to-node communication (the private network). Do not use teaming network adapters on the private networks.
- ◆ Set the order of the network adapter binding as follows:
 1. External public network
 2. Internal private network (Heartbeat)
 3. [Remote Access Connections]

For more information, refer to *Modify the Protocol Bindings Order* in the **TechNet** section of the Microsoft website.

- ◆ Set the speed and duplex mode for multiple speed adapters to the same values and settings. If the adapters are connected to a switch, ensure that the port settings of the switch match those of the adapters.

For more information, refer to *Change Network Adapter Settings* in the **TechNet** section of the Microsoft website.

Configuring TCP/IP settings

To configure TCP/IP settings:

1. Open **Network Connections** (**Start** > **Settings** > **Control Panel** > **Network and Dial-up Connections** from the Windows desktop).
2. Right-click the connection you want to configure, and click **Properties**.
3. Do one of the following:
 - If the connection is a local area connection, on the **General** tab, in **This connection uses the following items**, click **Internet Protocol (TCP/IP)**, and then click **Properties**.

- If the connection is a dial-up or VPN connection, on the **Networking** tab, in **This connection uses the following items**, click **Internet Protocol (TCP/IP)**, and then click **Properties**.
 - If the connection is an incoming connection, on the **Networking** tab, in **Network components**, click **Internet Protocol (TCP/IP)**, and then click **Properties**.
4. Do one of the following:
 - If you want IP settings to be assigned automatically, click **Obtain an IP address automatically**, and then click **OK**.
 - If you want to specify an IP address or a DNS server address:
 - a. Click **Use the following IP address**, and type the IP address in **IP address**.
 - b. Click **Use the following DNS server addresses**, and in **Preferred DNS server** and **Alternate DNS server**, type the addresses of the primary and secondary DNS servers.
 5. To configure DNS, WINS, and IP settings, click **Advanced**.
 6. In a local area connection, selecting the **Obtain an IP address automatically** option enables the **Alternate Configuration** tab. Use this to enter alternate IP settings if your computer is used on more than one network. To configure DNS, WINS, and IP settings, click **User configured** on the **Alternate Configuration** tab.

Notes

Note the following:

- ◆ You should use automated IP settings (DHCP) whenever possible for all connections because they eliminate the need to configure settings such as IP address, DNS server address, and WINS server address.
- ◆ The **Alternate Configuration** settings specify a second set of IP settings that are used when a DHCP server is not available. This is very useful for laptop users who often switch between two different network environments such as DHCP and static IP network environments.
- ◆ Use static IP addresses for each network adapter on each node.

- ◆ For private networks, define the TCP/IP properties for static IP addresses following the guidelines under *Private Network Addressing Options* in the **TechNet** section of the Microsoft website. (That is, specify a class A, B, or C private address.)
- ◆ Do not configure a default gateway or DNS or WINS server on the private network adapters. Also, do not configure private network adapters to use name resolution servers on the public network; otherwise, a name resolution server on the public network might map a name to an IP address on the private network. If a client then received that IP address from the name resolution server, it may fail to reach the address because no route from the client to the private network address exists.
- ◆ Configure WINS and /or DNS servers on the public network adapters. If Network Name resources are used on the public networks, set up the DNS servers to support dynamic updates; otherwise, the Network Name resources may not fail over correctly.

For more information, refer to *Configure TCP/IP settings* in the **TechNet** section of the Microsoft website.

- ◆ Configure a default gateway on the public network adapters. If there are multiple public networks in the cluster, configure a default gateway on only one of these.

For more information, refer to *Configure TCP/IP settings* in the **TechNet** section of the Microsoft website.

- ◆ Clearly identify each network by changing the default name.

For example, you could change the name of the private network connection from the default *Local Area Connection* to *Private Cluster Network*.

- ◆ Change the role of the private network from the default setting of **All communications (mixed network)** to **Internal cluster communications only (private network)**, and verify that each public network is set to **All communications (mixed network)**.

For more information, refer to *Change How the Cluster uses a Network* in the **TechNet** section of the Microsoft website.

- ◆ Place the private network at the top of the Network Priority list for internal node-to-node communication in the cluster.

For more information, refer to *Change Network Priority for Communication between Nodes* in the **TechNet** section of the Microsoft website.

Related information

For more information, also refer to the following in the **TechNet** section of the Microsoft website:

- ◆ *Network Connections Best Practices*
- ◆ *TCP/IP*
- ◆ *New Ways to do Network Connections Tasks*

Using a different disk for the quorum resource

To use a different disk for the quorum resource:

To perform this procedure, you must be a member of the Administrators group on the local computer, or you must have been delegated the appropriate authority. If the computer is joined to a domain, members of the Domain Admins group might be able to perform this procedure.

1. Select the disk you will use for the quorum resource. If the disk has two or more NTFS partitions, ensure that all partitions on the disk are assigned drive letters.
2. Open Cluster Administrator (**Start > Settings > Control Panel > Administrative Tools > Cluster** from the Windows desktop).
3. If one does not already exist, create a Physical Disk or other storage-class resource for the new disk.

For more information, refer to *Checklist: Installing a Physical Disk Resource* in the **TechNet** section of the Microsoft website.

4. In the console tree, click the **Resources** folder.
5. In the details pane, click the resource you will use for the quorum resource.
6. On the **File** menu, click **Take Offline**.
7. In the console tree, click the cluster name.
8. On the **File** menu, click **Properties**.
9. On the **Quorum** tab, click **Quorum resource**, and select a new disk or storage-class resource you want to use as the quorum resource for the cluster.

10. In **Partition**, if the disk has more than one partition, click the partition where you want the cluster specific data kept.
11. In **Root path**, type the path to the folder on the partition.
Example: \MSCS
12. In the console tree, click the **Resources** folder.
13. In the details pane, click the new quorum resource.
14. On the **File** menu, click **Bring Online**.

Cautions Be aware of the following:

- ◆ If Cluster Administrator does not open, it may be because the Cluster service is not started.

To start the Cluster service, refer to *Set the Cluster Service to only Start Manually* in the **TechNet** section of the Microsoft website.

- ◆ If the quorum resource becomes corrupted, you cannot move it until it is repaired.

For more information, refer to *Node-to-Node Connectivity Problems* in the **TechNet** section of the Microsoft website.

- ◆ If the Cluster service is stopped on any other node, or the node is shut down, that node will not be able to form a cluster. Only nodes that are online when this change is made can form the cluster. However, nodes that are offline can still join the cluster and will be able to form a cluster after they have joined the cluster at least once.

Notes Note the following:

- ◆ As a security best practice, consider using **Run as** to perform this procedure. For more information, refer to *Default Local Groups*, *Default Groups*, and *Using Run as* in the **TechNet** section of the Microsoft website.
- ◆ Of the default resource types, only the Physical Disk, Local Quorum, or Majority Node Set resource can be a quorum resource. However, third party vendors can supply other storage class resource types that are quorum-capable.
- ◆ When you are making changes to the quorum, do not add or evict nodes at the same time.

- ◆ If you designate a resource as the quorum resource without first taking that resource offline, restart a node afterwards. This will ensure that normal checkpointing is initiated for the quorum resource. Checkpointing is the saving of an extra copy of cluster configuration data and is normally done at four-hour intervals.

Related information

In addition to the information referenced earlier in this section, refer to the following in the **TechNet** section of the Microsoft website:

- ◆ *Assign, Change, or Remove a Drive Letter*
- ◆ *Quorum Resource*

Recovering from a corrupted quorum log or quorum disk

To recover from a corrupted quorum log or quorum disk:

To perform this procedure, you must be a member of the Administrators group on the local computer, or you must have been delegated the appropriate authority. If the computer is joined to a domain, members of the Domain Admins group might be able to perform this procedure.

1. If the Cluster service is running, open Computer Management (**Start > Settings > Control Panel > Administrative Tools > Computer Management** from the Windows desktop).
2. In the console tree, double-click **Services and Applications**, and then click **Services**.
3. In the details pane, click **Cluster Service**.
4. On the **Action** menu, click **Stop**.
5. Repeat steps 1 through 4 for each remaining node.
6. If you have a backup of the quorum log, restore the log by following the instructions under *Backing Up and Restoring Server Clusters* in the **TechNet** section of the Microsoft website.
7. If you do not have a backup, select any given node. Make sure **Cluster Service** is highlighted in the details pane, and click **Properties** on the **Action** menu.

Under **Service status**, in **Start parameters**, specify **/fixquorum** and click **Start**.

8. Switch from the problematic quorum disk to another quorum resource.

For more information, refer to *Use a Different Disk for the Quorum Resource* in the **TechNet** section of the Microsoft website.

9. In Cluster Administrator, bring the new quorum resource disk online.

For information about how to do this, refer to *Bring a Resource Online* in the **TechNet** section of the Microsoft website.

10. Click **Start > Run**, and type a command with the following syntax:

```
cluster ClusterName res QuorumDiskResourceName /maint:on
```

11. Run `chkdsk`, using the switches `/f` and `/r`, on the quorum resource disk to determine whether the disk is corrupted.

For more information on running `chkdsk`, refer to *chkdsk* in the **TechNet** section of the Microsoft website.

12. Click **Start > Run**, and type a command with the following syntax:

```
cluster ClusterName res QuorumDiskResourceName /maint:off
```

13. The next step depends on whether `chkdsk` detects any corruption on the disk:

- If no corruption is detected, it is likely that the log was corrupted. Proceed to step 15.
- If corruption is detected, check the system log in Event Viewer for possible hardware errors.

Resolve any hardware errors before continuing.

You can use the ClusterRecovery tool, available in the Microsoft Windows Server 2003 Resource Kit, to restore the registry checkpoint files.

14. After `chkdsk` is complete, repeat steps 1 through 4 to stop the Cluster service.
15. Make sure that **Cluster Service** is highlighted in the details pane. Then click **Properties** on the **Action** menu.

Under **Service status**, in **Start parameters**, specify **/resetquorumlog**, and click **Start**. This restores the quorum log from the node's local database.

Important: You must start the Cluster service by clicking **Start** on the service control panel. You cannot click **OK** or **Apply** to commit these changes, as this does not preserve the **/resetquorumlog** parameter.

16. Restart the Cluster service on all other nodes.

Notes Note the following:

- ◆ As a security best practice, consider using **Run as** to perform this procedure. For more information, refer to *Default Local Groups*, *Default Groups*, and *Using Run as* in the **TechNet** section of the Microsoft website.
- ◆ The quorum disk must be formatted with the NTFS file system.
- ◆ If none of the nodes are running, or one node fails while you are changing the quorum resource, only the running nodes are able to form the cluster, and the offline node is only able to join the cluster.

After the offline node has joined the cluster, all nodes are again able to form or join the cluster. This design prevents the offline node from forming the cluster using the old quorum resource.

Related information

In addition to the information referenced earlier in in this section, refer to the following in the **TechNet** section of the Microsoft website:

- ◆ *Start, stop, pause, resume, or restart a service*
- ◆ *Quorum Resource*
- ◆ *Backing Up and Restoring Server Clusters*
- ◆ *Use a Different Disk for the Quorum Resource*
- ◆ *Run a Disk Maintenance Tool, such as Chkdsk on a Physical Disk resource*

Checklist: Installing a physical disk resource

Use the following checklist when installing a physical disk resource.

Step	Reference in TechNet section of Microsoft website
Review the concepts behind cluster resources.	<i>Server Cluster Resources</i>
Review resource groups and resource dependencies.	<i>Server Cluster Groups</i>
Review disk storage concepts.	<i>Disk Management</i>
Plan common resource settings.	<i>Checklist: Creating a New Resource</i>
Make sure the disk has a signature.	<i>Disk Management</i>
Make sure the disk is configured as basic and not dynamic.	<i>Disk Management</i>
Make sure all partitions on the disk are formatted using the NTFS file system. Disk Management	<i>Disk Management</i>
Make sure that all partitions have a drive letter assigned or is set up as a mounted drive. Important: If you use drive letters, assign drive labels that match the drive letters and make sure that you assign the same drive letter to each node in the cluster.	<i>Disk Management</i> <i>Install Local Storage Buses and Devices</i>
Use the New Resource Wizard to create the resource.	<i>Create a New Resource</i>

Creating a new group in a cluster

To create a new group, follow the steps in this section.

To perform this procedure, you must be a member of the Administrators group on the local computer, or you must have been delegated the appropriate authority. If the computer is joined to a domain, members of the Domain Admins group might be able to perform this procedure.

1. Open Cluster Administrator (**Start > Settings > Control Panel > Administrative Tools > Cluster Administrator** from the Windows desktop).
2. On the **File** menu, point to **New**, and click **Group**.
3. In the **New Group Wizard**, in **Name**, type a name for the new group.
4. In **Description**, type any comments you want, and then click **Next**.
5. Under **Available nodes**, click the nodes you want to be the preferred owners for the group, and then click **Add**.

You can also leave **Preferred owners** empty. This means that it does not matter to which node the group fails back after failover takes place. For more information, refer to *Failover and Failback* in the **TechNet** section of the Microsoft website.
6. To change the priority of an owner, select the owner, click **Move Up** or **Move Down**, and then click **Finish**.

Notes Note the following:

- ◆ As a security best practice, consider using **Run as** to perform this procedure. For more information, refer to *Default Local Groups*, *Default Groups*, and *Using Run as* in the **TechNet** section of the Microsoft website.
- ◆ The name you give the group is for administrative purposes only. It is not the same as a network name, which is the name that clients can use to connect to the group. However, make sure this name is different from any other group name in the cluster.
- ◆ Cluster Administrator lists the group description and the group name in the details pane.
- ◆ All resources within the new group fail over together. To have the group fail back to a certain node, specify that node as the preferred owner and enable failback. You can balance groups among all nodes to maximize the performance of the cluster. However, you can choose not to have a preferred owner if the location of the group does not greatly affect performance.

Related information

In addition to the information referenced earlier in this section, refer to the following in the **TechNet** section of the Microsoft website:

- ◆ *Server Cluster Groups*

Creating a new resource

To create a new resource:

To perform this procedure, you must be a member of the Administrators group on the local computer, or you must have been delegated the appropriate authority. If the computer is joined to a domain, members of the Domain Admins group might be able to perform this procedure.

1. Open Cluster Administrator (**Start > Settings > Control Panel > Administrative Tools > Cluster Administrator** from the Windows desktop).
2. In the console tree, double-click the **Groups** folder.
3. In the details pane, click the group to which you want the resource to belong.
4. On the **File** menu, point to **New**, and then click **Resource**.
5. In the **New Resource Wizard**, type the appropriate information in **Name and Description**, click the appropriate information in **Resource type and Group**, and then click **Next**.
6. Add or remove possible owners of the resource, and then click **Next**.
7. To add or remove dependencies:
 - To add dependencies, under **Available resources**, click a resource, and then click **Add**.
 - To remove dependencies, under **Resource dependencies**, click a resource, and then click **Remove**.
8. Repeat step 7 for any other resource dependencies, and then click **Next**.
9. Set resource properties in the *resource***Parameters** dialog box, where *resource* is the name of the resource type.

Different resource types contain different configuration information in their respective dialog boxes.

For more information on setting resource properties, refer to [“Related information” on page 54](#).

Notes Note the following:

- ◆ As a security best practice, consider using **Run as** to perform this procedure. For more information, refer to *Default Local Groups*, *Default Groups*, and *Using Run as* in the **TechNet** section of the Microsoft website.
- ◆ Before adding a resource to your cluster, you must verify that the following are true:
 - The type of resource is either one of the basic types provided with the Windows Server 2003 family or a custom resource type provided by the resource vendor.
 - A group already exists within the cluster to which your resource will belong.
 - All dependent resources have been created.
- ◆ A resource that fails frequently can affect its associated Resource Monitor and other resources interacting with that Resource Manager. For this reason, a separate Resource Monitor is recommended for any resource that has failed repeatedly in the past.

Related information

In addition to the information referenced earlier in this section, refer to the following in the **TechNet** section of the Microsoft website:

- ◆ *Server Cluster Resources*
- ◆ *Resource Types*
- ◆ *Standard Resource Types*
- ◆ *View or Set Resource Properties*
- ◆ *Setting Resource Properties*
- ◆ *Resource Dependencies*
- ◆ *Resource Monitors*

Best practices for configuring and operating server clusters

The guidelines in this section can help you effectively use a server cluster.

Secure your server cluster

To prevent your server cluster from being adversely affected by denial of service attacks, data tampering, and other malicious attacks, it is highly recommended that you plan for and implement the

security measures detailed in *Best practices for securing server clusters* in the **TechNet** section of the Microsoft website.

Verify that your server cluster hardware is listed in the Windows catalog

For Windows Server 2003, Enterprise Edition and Windows Server 2003, Datacenter Edition, Microsoft supports only complete server cluster systems chosen from the Windows Catalog.

To see if your system or hardware components, including your cluster disks, are compatible, refer to *Support Resources* in the **TechNet** section of the Microsoft website.

For a geographically dispersed cluster, both the hardware and software configuration must be certified and listed in the Windows Catalog.

The network interface controllers (NICs) used in certified cluster configurations must be selected from the Windows Catalog.

It is recommended that your cluster configuration consist of identical storage hardware on all cluster nodes to simplify configuration and eliminate potential compatibility problems.

Partition and format disks before adding the first node to your cluster

Partition and format all disks on the cluster storage device before adding the first node to your cluster. You must format the disk that will be the quorum resource. All partitions on the cluster storage device must be formatted with NTFS (they can be either compressed or uncompressed), and all partitions on one disk are managed as one resource and move as a unit between nodes.

Important: Cluster disks on the cluster storage device must be partitioned as master boot record (MBR) and not as GUID partition table (GPT) disks.

Correctly set up your server cluster's networks

Follow these guidelines to reduce network problems in your server cluster:

- ◆ Use identical network adapters in all cluster nodes, that is, make sure each adapter is the same make, model, and firmware version.

- ◆ Use at least two interconnects. Although a server cluster can function with only one interconnect, at least two interconnects are necessary to eliminate a single point of failure and are required for the verification of original equipment manufacturer (OEM) clusters.
- ◆ Reserve one network exclusively for internal node-to-node communication (the private network). Do not use teaming network adapters on the private networks.
- ◆ Set the order of the network adapter binding as follows:
 1. External public network
 2. Internal private network (Heartbeat)
 3. [Remote Access Connections]

For more information, refer to *Modify the Protocol Bindings Order* in the **TechNet** section of the Microsoft website

- ◆ Set the speed and duplex mode for multiple speed adapters to the same values and settings. If the adapters are connected to a switch, ensure that the port settings of the switch match those of the adapters.

For more information, refer to *Change Network Adapter Settings* in the **TechNet** section of the Microsoft website.

- ◆ Use static IP addresses for each network adapter on each node.
- ◆ For private networks, define the TCP/IP properties for static IP addresses following the guidelines under *Private Network Addressing Options* in the **TechNet** section of the Microsoft website. (That is, specify a class A, B, or C private address.)
- ◆ Do not configure a default gateway or DNS or WINS server on the private network adapters. Also, do not configure private network adapters to use name resolution servers on the public network; otherwise, a name resolution server on the public network might map a name to an IP address on the private network. If a client then received that IP address from the name resolution server, it may fail to reach the address because no route from the client to the private network address exists.
- ◆ Configure WINS and /or DNS servers on the public network adapters. If Network Name resources are used on the public networks, set up the DNS servers to support dynamic updates; otherwise, the Network Name resources may not fail over correctly.

For more information, refer to *Configure TCP/IP settings* in the **TechNet** section of the Microsoft website.

- ◆ Configure a default gateway on the public network adapters. If there are multiple public networks in the cluster, configure a default gateway on only one of these.

For more information, refer to *Configure TCP/IP settings* in the **TechNet** section of the Microsoft website.

- ◆ Clearly identify each network by changing the default name.

For example, you could change the name of the private network connection from the default *Local Area Connection* to *Private Cluster Network*.

- ◆ Change the role of the private network from the default setting of **All communications (mixed network)** to **Internal cluster communications only (private network)**, and verify that each public network is set to **All communications (mixed network)**.

For more information, refer to *Change How the Cluster uses a Network* in the **TechNet** section of the Microsoft website.

- ◆ Place the private network at the top of the Network Priority list for internal node-to-node communication in the cluster.

For more information, refer to *Change Network Priority for Communication between Nodes* in the **TechNet** section of the Microsoft website.

Do not install applications into the default Cluster Group

Do not delete or rename the default Cluster Group or remove any resources from that resource group

The default Cluster Group contains the settings for the cluster and some typical resources that provide generic information and failover policies. This group is essential for connectivity to the cluster. It is therefore very important to keep application resources out of the default Cluster Group and so prevent clients from connecting to the Cluster Group's IP address and network name resources. If a resource for an application is added to this group and the resource fails, it may cause the cluster group to fail also, therefore reducing the overall availability of the entire cluster. It is highly recommended that you create separate resource groups for application resources.

For more information, refer to *Planning your Groups and Checklist: Planning and Creating a Server Cluster* in the **TechNet** section of the Microsoft website.

Back up your server cluster

To be able to effectively restore your server cluster in the event of application data or quorum loss, or individual node or complete cluster failure, follow these steps when preparing backups:

1. Perform an Automated System Recovery (ASR) backup on each node in the cluster.
2. Back up the cluster disks from each node.
3. Back up each individual application (for example, Microsoft Exchange Server or Microsoft SQL Server) running on the nodes.

By default, Backup Operators do not have the user rights necessary to create an Automated System Recovery (ASR) backup on a cluster node. However, Backup Operators can perform this procedure if that group is added to the security descriptor for the Cluster service. You can do that using Cluster Administrator or `cluster.exe`. For more information, refer to *Give a User Permissions to Administer a Cluster* and *Cluster* in the **TechNet** section of the Microsoft website.

For more information, refer to *Backing Up and Restoring Server Clusters* in the **TechNet** section of the Microsoft website. For more information on backing up applications in a cluster, see the documentation for that application.

Give the cluster service account full rights to administer computer objects if Kerberos authentication is enabled for virtual servers

If you enable Kerberos authentication for a virtual server's Network Name resource, the Cluster service account does not need full access rights to the computer object associated with that Network Name resource. The Cluster service can use the default access rights given to members of the authenticated users group, but certain operations (for example, renaming the computer object) will be restricted.

It is recommended that you work with your domain administrator to set up appropriate administration rights and permissions for the Cluster service account.

For more information, refer to information about Kerberos authentication under *Virtual Servers* in the **TechNet** section of the Microsoft website.

Do not install scripts used by Generic Script resources on cluster disks

It is recommended that you install script files used by Generic Script resources on local disks, not on cluster disks. Incorrectly written script files can cause the cluster to stop responding. Installing the script files on a local disk makes it easier to recover from this scenario.

For guidelines on writing scripts for the Generic Script resource, refer to the Microsoft Platform Software Development Kit (SDK). For information on troubleshooting Generic Script resource issues, refer to article Q811685, *A Server Cluster with a Generic Script Resource Stops Responding*, in the Microsoft Knowledge Base.

How to troubleshoot cluster service startup issues

This section describes basic troubleshooting steps you can use to diagnose Cluster service startup issues. Although this is not a comprehensive list of all the issues that can cause the Cluster service not to start, it does address about 90 percent of startup issues.

This information was taken from Microsoft Article ID: 266274, revision 3.0.



CAUTION

This section contains information about modifying the registry. Before you modify the registry, back it up, and make sure that you understand how to restore the registry if a problem occurs.

For information about how to back up, restore, and edit the registry, refer to *Description of the Microsoft Windows Registry* in the Microsoft Knowledge Base.

When the Cluster service initially starts, it attempts to join an existing cluster. For this to occur, the Cluster service must be able to contact an existing cluster node. If the join procedure does not succeed, the cluster continues to the form stage; the main requirement of this stage is the ability to mount the quorum device.

Here are the steps in the startup process:

1. Authenticate the Service account.
2. Load the local copy of the cluster database.

3. Use information in the local database to try to contact other nodes to begin the join procedure. If a node is contacted and authentication is successful, the join procedure is successful.
4. If no other node is available, the Cluster service uses the information in the local database to mount the quorum device and updates the local copy of the database by loading the latest checkpoint file and replaying the quorum log.



WARNING

If you use Registry Editor incorrectly, you may cause serious problems that may require you to reinstall your operating system. Microsoft cannot guarantee that you can solve problems that result from using Registry Editor incorrectly. Use Registry Editor at your own risk.

Troubleshooting cluster service startup issues

Follow these steps to troubleshoot issues:

1. Verify that the cluster node that is having problems is able to properly authenticate the Service account. You can determine this by logging on to the computer with the Cluster service account, or by checking the System event log for Cluster service logon problem event messages.

2. Verify that the %SystemRoot%\Cluster folder contains a valid Clusdb file and that the Cluster service attempted to start.

Start the Registry Editor (Regedt32.exe) and verify that the following registry key is valid and loaded:

```
HKEY_LOCAL_MACHINE\Cluster
```

The cluster hive should have a structure that is very similar to Cluster Administrator. Make note of the network and quorum keys. If the database is not valid, you can copy and use the cluster database from a live node. If all nodes do not have a valid cluster database, refer to *How to Use the Cluster TMP file to Replace a Damaged Clusdb File* in the Microsoft Knowledge Base.

3. If the node is not the first node in the cluster, check connectivity to other cluster nodes across all available networks. Use the ping.exe tool to verify TCP/IP connectivity, and use Cluster Administrator to verify that the Cluster service can be contacted.

Use the TCP/IP addresses of the network adapters in the other nodes in the **Connect to** dialog box in Cluster Administrator.

4. If it cannot contact any other node, the service continues with the form phase. It attempts to locate information about the quorum in the local cluster database, and then tries to mount the disk. If the quorum disk cannot be mounted, the service does not start.

If another node has successfully started and has ownership of the quorum, the service does not start. This is usually caused by connectivity or authentication issues. If this is not the case, you can check the status of the quorum device by starting the service with the `-fixquorum` switch, and attempting to bring the quorum disk online, or changing the quorum location for the service. Also check the System event log for disk errors.

If the quorum disk successfully comes online, it is likely that the quorum is corrupted. To correct this issue, refer to the following Microsoft Knowledge Base articles:

- Windows NT 4.0: 172951, *How to Recover from a Corrupted Quorum Log*
 - Windows 2000: 245762, *Recovering from a Lost or Corrupted Quorum Log*
5. Check the attributes of the `Cluster.log` file to make sure that it is not read-only, and make sure that no policy is in effect that prevents modification of the `Cluster.log` file. If either of these conditions exist, the Cluster service cannot start. If these steps do not resolve the problem, you should take additional troubleshooting steps. The cluster log file can be valuable in additional troubleshooting.

By default, cluster logging is enabled on Windows 2000-based computers that are running the Cluster service. To enable cluster logging on Windows NT 4.0-based computers, refer to Microsoft Knowledge Base article 168801, *How to Enable Cluster Logging in Microsoft Cluster Server*.

Recovering from an Event ID 1034 on a server cluster

This information was taken from Microsoft Article ID: 280425, revision 3.3.

Symptoms

A physical disk resource may fail to come online, or the Cluster service may fail to start. The following message is generated in the system event log:

```
Event ID: 1034
Source: ClusDisk
Description: The disk associated with cluster disk
             resource DriveLetter could not be found. The
             expected signature of the disk was DiskSignature.
```

Cause

These issues typically occur if either of the following conditions is true:

- ◆ A disk has become unavailable or inaccessible, and therefore, the Cluster service cannot find it.
- ◆ The signature on the disk has been changed.

The Cluster service recognizes and identifies disks by their disk signatures. Disk signatures are stored on the physical disk in the master boot record (MBR). The MBR is a record that the Cluster service keeps of all the disks that it manages. It uses the MBR to track the disks. During the course of Cluster service operations (start, restart, failover, and so forth), if the Cluster service cannot find a disk that is identified by a particular signature, it will fail to bring the disk online. The cluster component that specifically detects this condition and logs the error is the cluster disk filter driver (*Clusdisk.sys*). The error message provides information on the "missing disk" but does not indicate the reasons that this condition may have occurred.

Resolution

To resolve this problem:

1. Make sure the disk is actually exposed through the shared interconnects and is visible to the operating system. To do this:
 - a. Click **Start**.
 - b. Click **Run**.
 - c. Type **CompMgmt.msc**.
 - d. Click **OK**.
 - e. In **Computer Management** under **System Tools, Device Manager**, look under **Disk Drives**, and you can view all the logical disks that are being presented to the node.

All nodes in a cluster can see the same number of disk drives for disks that are managed by the cluster. For example, if there are 10 disks that are managed by the cluster, all 10 are visible to all nodes in the cluster. If you know the Target ID and LUN of the disk, you can validate them by clicking **Properties** for each disk.

- If the count does not match, the disk is not accessible to that node. Troubleshoot your storage solution to make sure that the disk is accessible and can be mounted by the operating system. When the storage solution is functioning correctly, you can rescan the bus by right-clicking the Device Manager disks.
 - If the count does match, and if the Cluster service is up and running, reduce the complexity, if possible, by moving the all the disk resources (groups that host the resources) to a single node. If the Cluster service has failed, shut down all nodes and restart one node.
2. If the disk signatures have changed, use `dumpcfg.exe` to write the expected signature back to the disk. The signatures of the disks as enumerated by `dumpcfg` should match the list that is derived from the following registry subkey:

`HKLM/System/CurrentControlSet/Services/Clusdisk/Parameters`

Clusdisk uses this information to bind to disks that are managed by the Cluster service.

3. If the signatures in the list do not match the registry subkey list, you must correctly identify the disks that have had their signatures changed and reset them to the expected signatures. To do this:
 - a. Power down all but one node.
 - b. Open Computer Management, double-click **Storage**, and then click **Disk Management**.

In Logical Disk Manager, note the disk number and label that is associated with the failing disk. This information is to the left of the partition information. For example: Disk 0.

Compare the information that is displayed with the message in the *Description* section of the Event ID 1034.

Example: The disk associated with cluster disk resource 'Disk Q:\'. The disk label should not change even if the signature has. The disk label will help you correctly identify the problem disk. Once the disk has been correctly identified its signature can be checked again to validate the mismatch.

- c. If you cannot see the disks in `DiskMgmt.msc`, set the Cluster service and Cluster Disk device to **Manual**, and then restart the node. (All other nodes should remain shut down.)
To do this (noting that step i. may not be necessary):
 - i. Click **Start**, point to **Programs**, point to **Administrative Tools**, and then click **Computer Management**.
 - ii. Click **Device Manager** in the left pane, and then click **Show Hidden Devices** on the **View** menu.
 - iii. In the right pane, view the non-Plug and Play drives section, and then double-click the **Clusdisk** driver.
 - iv. On the **Driver** tab, change the **Startup type** option from **System** to **Disabled**.
 - v. In the left pane, double-click **Services and Applications**, and then click **Services**.
 - vi. In the right pane, double-click the **Cluster** service, and then click **Disabled** in the **Startup type** box.
 - vii. Restart the node, and then repeat step ii if necessary.
- d. Write the signature that the Cluster service expects to the disk:
 - i. Obtain the expected signature from the **Description** section of the Event ID 1034 error message. For example: The expected signature of the disk was 12345678.
 - ii. Copy `DumpCfg.exe` from the Windows 2000 Resource Kit to the local node. At the command prompt, type `dumpcfg.exe`. Under the [DISKS] section, the disk number and signature for all available disks is displayed. Validate the actual disk signature with what the Cluster service
 - iii. Write the expected signature to the disk by using the following command, where 12345678 is the disk signature in hexadecimal, and 0 is the disk number that you replaced (which was obtained from the previous step):

```
dumpcfg.exe -s 12345678 0
```

For more information about using Dumpcfg.exe, type `dumpcfg /?` at the command prompt.

- e. Set the Cluster service back to **Automatic**, and set the Cluster Disk device back to **System** on the node. Start the Cluster Disk device, and then start the Cluster service.
- f. Open Cluster Administrator, and then bring the disk online.
- g. Turn on all other nodes, one at a time, and then test failover.

Cluster disk and drive connection problems

This section describes some problems specific to cluster disk and drive connections.

Problem When the physical disks are not powering up or spinning, Cluster service cannot initialize any quorum resources.

Cause

Cables are not correctly connected, or the physical disks are not configured to spin when they receive power.

Solution

After checking that the cables are correctly connected, check that the physical disks are configured to spin when they receive power.

Problem The Cluster service fails to start and generates an Event ID 1034 in the Event log after you replace a failed hard disk, or change drives for the quorum resource.

Cause

If a hard disk is replaced, or the bus is reenumerated, the Cluster service may not find the expected disk signatures, and consequently may fail to mount the disk.

Solution

Write down the expected signature from the Description section of the Event ID 1034 error message. Then follow these steps:

1. Back up the server cluster.
2. Set the Cluster service to start manually on all nodes, and then turn off all but one node.
3. If necessary, partition the new disk and assign a drive letter.

4. Use the `confdisk.exe` tool (available in the Microsoft Windows Server 2003 Resource Kit) to write that signature to the disk.
5. Start the Cluster service and bring the disk online .
6. If necessary, restore the cluster configuration information.
7. Turn on each node, one at a time.

For information on replacing disks in a server cluster, refer to Microsoft Knowledge Base article Q305793, *How to Replace a Disk with Windows 2000 or Windows Server 2003 family Clusters*.

Problem Drive on the shared storage bus is not recognized.

Cause

Scanning for storage devices is not disabled on each controller on the shared storage bus.

Solution

Verify that scanning for storage devices is disabled on each controller on the shared storage bus. Many times, the second computer you turn on does not recognize the shared storage bus during the BIOS scan if the first computer is running. This situation can manifest itself in a `Device not ready` error being generated by the controller, or in substantial delays during startup. To correct this, disable the option to scan for devices on the shared controller.

This symptom can manifest itself as one of several errors, depending on the attached controller. It is normally accompanied with a one- to two-minute start delay and an error indicating the failure of some device.

Problem Configuration cannot be accessed through Disk Management. Under normal cluster operations, the node that owns a quorum resource locks the drive storing the quorum resource, preventing the other nodes from using the device.

If you find that the cluster node that owns a quorum resource cannot access configuration information through Disk Management, the cause of the problem might be either of the following.

Possible Cause 1

A device does not have physical connectivity and power.

Solution

Reseat controller cards, reseat cables, and make sure the drive spins up when you start.

Possible Cause 2

You attached the cluster storage device to all nodes and started all the nodes before installing the Cluster service on any node.

Solution

After you attach all servers to the cluster drives, you must install the Cluster service on one node before starting all the nodes. Attaching the drive to all the nodes before you have the cluster installed can corrupt the file system on the disk resources on the shared storage bus.

Problem SCSI or Fibre Channel storage devices do not respond.

Possible Cause 1

The SCSI bus is not properly terminated.

Solution

Make sure that the SCSI bus is not terminated early and that the SCSI bus is terminated at both ends.

Possible Cause 2

The SCSI/Fibre Channel cable is longer than the cable specification allows.

Solution

Make sure that the SCSI/Fibre Channel cable is not longer than the cable specification allows.

Possible Cause 3

The SCSI/Fibre Channel cable is damaged.

Solution

Make sure that the SCSI/Fibre Channel cable is not damaged. (For example, check for bent pins and loose connectors on the cable and replace it if necessary.)

Problem Disk groups do not move or stay online pending after move.

Cause

Cables are damaged or not properly installed.

Solution

Check for bent pins on cables and make sure that all cables are firmly anchored to the chassis of the server and drive cabinet.

Problem Disks do not come online or Cluster service does not start when a node is turned off.

Cause

If the quorum log is corrupted, the Cluster service cannot start.

Solution

If you suspect that the quorum resource is corrupted, refer to the information on the problem *Quorum log becomes corrupted* in *Node-to-node connectivity problems* in the **TechNet** section of the Microsoft website.

Problem

Drives do not fail over or come online.

Possible Cause 1

The drive is not on a shared storage bus.

Solution

If drives on the shared storage bus do not fail over or come online, make sure the disk is on a shared storage bus, not on a nonsystem bus.

Possible Cause 2

If you have more than one local storage bus, some drives in **Shared cluster disks** will not be on a shared storage bus.

Solution

If you do not remove these drives from **Shared cluster disks** (in the Cluster Application Wizard), the drives do not fail over, even though you can configure them as resources.

Problem

Mounted drives disappear, do not fail over, or do not come online.

Cause

The clustered mounted drive was not configured correctly.

Solution

Look at the Cluster service errors in the Event Log (**ClusSvc** in the **Source** column). You need to recreate or reconfigure the clustered mounted drive if the description of any Cluster service error is similar to the following:

```
Cluster disk resource disk_resource: Mount point
mount_drive for target volume target_volume is not
acceptable for a clustered disk because reason. This
mount point will not be maintained by the disk
resource.
```

When recreating or reconfiguring the mounted drive(s), follow these guidelines:

- ◆ Make sure that you create unique mounted drives so that they do not conflict with existing local drives on any node in the cluster.
- ◆ Do not create mounted drives between disks on the cluster storage device (cluster disks) and local disks.
- ◆ Do not create a mounted drive from a clustered disk to the cluster disk that contains the quorum resource (the quorum disk). You can, however, create a mounted drive from the quorum disk to a clustered disk.
- ◆ Mounted drives from one cluster disk to another must be in the same cluster resource group, and must be dependent on the root disk.

For more information on viewing the Event Log, refer to *View Event Logs* in the **TechNet** section of the Microsoft website. For more information on creating mounted drives in a server cluster, refer to *Add drives on the shared storage bus* in the **TechNet** section of the Microsoft website.

Problem The cluster quorum disk (containing the quorum resource) becomes disconnected from all nodes in a cluster and you are later unable to add the nodes back to the cluster.

Cause and solution

Refer to identically titled problem in *Node-to-node connectivity problems* in the **TechNet** section of the Microsoft website.

For information about how to obtain product support, refer to *Technical support options* in the **TechNet** section of the Microsoft website.

It is important that you correctly configure the storage topology (for example, SCSI, Fibre Channel, Storage Area Networks) and the storage interconnects (for example, multiple paths) used in your server cluster. Before deploying your server cluster, contact your hardware vendors to ensure that your particular cluster storage configuration is supported at the hardware level.

For descriptions of supported cluster storage topologies, best practices for deploying and managing cluster storage, and a list of cluster storage-related Knowledge Base articles, refer to the cluster storage information in the **TechNet** section of the Microsoft website.

Client-to-cluster connectivity problems

This section describes some problems specific to client-to-cluster connectivity.

Problem Clients cannot connect to virtual servers.

Possible cause 1

The client might not be using the correct network name or IP address to access the cluster.

Solution

Make sure that the client is accessing the cluster using the correct network name or IP address.

Possible cause 2

The client might not have the TCP/IP protocol correctly installed and configured.

Solution

Make sure that the client has the TCP/IP protocol correctly installed and configured. Depending on the application being accessed, the client can address the cluster by specifying either the resource network name or the IP address. In the case of the network name, you can verify proper name resolution by checking the NetBT cache (using the `Nbtstat.exe` utility) to determine whether the name had been previously resolved. Also, confirm proper WINS configuration, both at the client and through the WINS Administrator.

If the client is accessing the resource through a specific IP address, ping the IP address of the cluster resource and cluster nodes from a command prompt.

For more information on `Nbtstat.exe`, refer to *Nbtstat* in the **TechNet** section of the Microsoft website. For instructions on how to troubleshoot connectivity problems, refer to *Troubleshoot client-to-virtual server connectivity* in the **TechNet** section of the Microsoft website.

Possible cause 3

The client may be attempting to connect to an incorrect node after a failover.

Solution

Make sure that the routers are configured correctly. For more information, refer to article Q244331, *MAC Address Changes for Virtual Server During a Failover with Clustering*, in the Microsoft Knowledge Base.

Problem Clients cannot access resources.

Possible cause 1

Both your private network and the primary network are down.

Solution

If your server cluster nodes are multihomed and use a private network to communicate, the primary network can be down and the Cluster service can still function normally. If the private network is lost between nodes, resources that are configured to do so will fail over to the node that has ownership of the quorum resource. Each of the other nodes will stop its Cluster service.

Possible cause 2

A node might have lost connectivity with the network.

Solution

If you suspect a node has lost connectivity with the network:

- ◆ Confirm that the connection configuration settings are correct for each connection. To do this, open Control Panel and double-click **Network Connections**. Right-click a connection and click **Properties**.
- ◆ Check WINS or DNS to make sure that they are properly configured.
- ◆ Confirm that the static IP addresses being used by the Cluster service are still in place and are not being used by other resources on the network.

Problem A node cannot communicate on a network.

Cause

If a node cannot communicate on a network, it might lack a physical connection to the other nodes, or the other nodes might not be connected to one another. If the cabling has failed, the Cluster service might not have received the heartbeat of the node and so failed the resources owned by that node over to other nodes in the cluster. The node can no longer communicate to clients on a network.

Solution

Check the hub and local cabling, both to the network and to the cluster disks.

Problem Clients cannot access a group that has failed over.

Cause

There might not be a physical connection between the cluster nodes and the group that failed over has also failed to come online.

Solution

1. Confirm that you have a physical connection between the cluster nodes:
 - Make sure that the network cabling has not failed or been damaged. If the cabling has failed, the Cluster service might have failed over the resources to another node.
 - Ensure that the storage device cabling has not failed.
 - If the nodes are separated by one or more hubs, check the connectivity through all hubs.
2. Bring the group back online.

Problem Clients cannot attach to a cluster file share resource.

Possible cause 1

WINS or DNS might not be correctly configured.

Solution

Make sure that WINS or DNS are correctly configured. If you are using WINS, run WINS Manager on the WINS server; then, on the **Mappings** menu, click **Show Database**. Make sure that each node and the cluster are registered in the WINS database and that the registrations are active. For more information on WINS and WINS Manager, refer to the *Networking Guide* in the *Microsoft Windows Server 2003 Resource Kit*.

Possible cause 2

Your security policies might not allow the client to access the file share.

Solution

Make sure that your security policies allow the client to access the file share. The client must have the right to log on to the file share, or the Guest account must be enabled for the client to have access.

Problem Clients cannot access a cluster resource.

Possible cause 1

Either the IP Address resource or Network Name resource for the group in which the resource is contained is not online.

Solution

Check the group dependencies; the resource is supposed to be dependent on either the IP Address or the Network Name resource. Ensure that the IP Address and Network Name resources are online in the resource group. From the client computer, try to ping the IP addresses of the virtual server and individual nodes.

Possible cause 2

Either the client or the cluster computer is not configured for either WINS or DNS.

Solution

Make sure that the cluster nodes are configured properly for name resolution using either WINS or DNS, and make sure that clients are configured to use the same form of name resolution.

Possible cause 3

The client is attempting to access the cluster from a different subnet, and DNS is not correctly configured.

Solution

Configure the cluster nodes and client computer to use WINS or DNS. If you use DNS, add a DNS address record for the cluster in the DNS database.

Problem Client can detect all nodes but cannot detect a virtual server.

Possible cause 1

The virtual server might not have its own IP Address and Network Name resources.

Solution

Make sure that the virtual server has its own IP Address and Network Name resources, and that both resources are online. Your servers must have an address and a name on the network for any other server or client to properly recognize that the servers are on the network.

Possible cause 2

One or more nodes are not correctly configured to use WINS or DNS.

Solution

Make sure that all nodes are correctly configured to use WINS or DNS.

Problem

Clients cannot connect to a virtual server and the system event logs on the client computers contain Kerberos authentication errors.

Cause

The cluster might have used a previously created computer object for the virtual server. When the virtual server's Network Name resource is brought online for the first time after Kerberos support has been enabled, the cluster changes the object's password. Active Directory then replicates this change to all domain controllers in the cluster. However, because of replication latencies, clients connecting to this virtual server might receive a Kerberos ticket encrypted with the old password. This can result in Kerberos authentication errors.

Solution

Wait until the new password has been replicated to all domain controllers, or force replication using the `repadmin` and `replmon` command line tools, available in the `\Support\Tools` folder on your operating system installation CD. For information on monitoring and troubleshooting replication issues, refer to *Troubleshooting replication* in the **TechNet** section of the Microsoft website. For information about how to obtain product support, refer to *Technical support options* in the **TechNet** section of the Microsoft website.

General administrative problems

This section describes some general problems not covered in other sections.

Problem

The Cluster service fails and the node cannot detect the network.

Cause

You probably have a configuration problem.

Solution

- ◆ If the node was recently configured, or if you have installed some resource that required you to restart the computer, make sure that the node is still properly configured for the network.

- ◆ Check that the server is properly configured for TCP/IP. Also check that the appropriate services are running. If the node recently failed, there is an instance of failover; but, if the other nodes are misconfigured as well, the failover will be inadequate and client access will fail.

Problem An IP address added to a group in the cluster fails.

Possible cause 1

The IP address is not unique.

Solution

The IP address must be different from every other group IP address and every other IP address on the network.

Possible cause 2

The IP address is not a static IP address.

Solution

The IP addresses must be statically assigned outside of a DHCP scope, or they must be reserved by the network administrator.

Problem An IP address resource is unresponsive when taken offline, for example you are unable to query its properties.

Cause

You may not have waited long enough for the resource to go offline.

Solution

If an IP Address resource is unresponsive when taken offline, make sure that you wait long enough for the resource to go offline. Certain resources take time to go offline. For example, it can take up to three minutes for the IP Address resource to go fully offline.

Problem You receive the error `RPC server is unavailable`.

Cause

The server may not be operational, or the Cluster service and the RPC services may not be running.

Solution

Make sure the server is operational and that both the Cluster service and the RPC services are running. Also, check the name resolution of the cluster; it is possible that you are using the wrong name or that the name is not being properly resolved by WINS or DNS.

Problem Cluster Administrator cannot open a connection to a node.

Cause

The node may not be running.

Solution

If Cluster Administrator cannot open a connection to a node, make sure that the node is running. If it is, confirm that both the Cluster service and the RPC services are running.

Problem An application starts but cannot be closed.

Cause

You may not have taken a resource offline using Cluster Administrator.

Solution

When you bring resources online using Cluster Administrator, you must also take those resources offline using Cluster Administrator; do not attempt to close or exit the application from the application interface.

Problem A resource group has failed over but will not fail back.

Possible cause 1

The hardware and network configurations may not be valid.

Solution

Make sure that the hardware and network configurations are valid. If any interconnect fails, failover can occur because the Cluster service does not detect a heartbeat, or it may not even register that the node was ever online. In this case, the Cluster service fails over the resources to the other nodes in the server cluster, but it cannot fail back because that node is still down.

Possible cause 2

The resource group may not have been configured to fail back immediately, or you are not troubleshooting the problem within the allowable failback hours for the resource.

Solution

Make sure that the resource group is configured to fail back immediately, or that you are troubleshooting the problem within the allowable failback hours for the resource group. A group can be configured to fail back only during specified hours. Often, administrators prevent failback during peak business hours. To check this, use Cluster Administrator to view the resource failback policy.

Possible cause 3

You restarted the node to test the failover policy for the group instead of pressing the reset button.

Solution

Make sure that you press the reset button on the node. The resource group will not failback to the preferred node if you shutdown, then restart the node. For more information on testing failback policies, refer to *Test node failure* in the **TechNet** section of the Microsoft website.

Problem All nodes appear to be functioning correctly, but you cannot access all of the drives from one node.

Possible cause 1

The shared drive may not be functioning.

Solution

Confirm that the shared drive is still functioning. Try to access the drive from another node. If you can do that, check the cable from the device to the node that you cannot perform the access. If the cable is not the problem, restart the computer and then try again to access the device. If you cannot access the drive, check your configuration.

Possible cause 2

The drive has completely failed.

Solution

Determine (from another node) whether the drive is functioning at all. You may have to restart the drive (by restarting the computer) or replace the drive. The hard disk with the resource or a dependency for the resource may have failed. You may have to replace a hard disk. You may also have to reinstall the cluster.

Problem Cluster Administrator update delays.

Cause

If you run Cluster Administrator from a remote computer, it may not display the correct (updated) cluster state when the cluster network name fails over from one node to another node. This can result in Cluster Administrator displaying a node as being online, when it is actually offline.

Solution

To work around this problem, restart Cluster Administrator. You can avoid this problem by connecting to clusters through node names.

However, if the node you are connected to fails, Cluster Administrator stops responding until the RPC connection times out.

Problem Cluster Administrator stops responding when a node fails.

Cause

The Cluster Administrator may be slow in doing dynamic updates.

Solution

If Cluster Administrator stops responding when a node fails, make sure that Cluster Administrator is not just slow in doing dynamic updates. If the Cluster service is running on a remaining node, Cluster Administrator is either not responding or is updating very slowly.

There are two ways to see if the Cluster service is running on a remaining node:

- ◆ Use the TCP/IP Ping utility to ping the cluster name on a remaining node.
- ◆ In Control Panel, double-click **Services**, and check whether the Cluster service is running.

Problem Cannot connect to cluster from recent file list.

Cause

Files listed in the Cluster Administrator recent file list on the **File** menu and in the **Open Connection to Cluster** dialog box) have the cluster name appended to the path. For example, instead of Webclust1, the recent file list may list C:\Windows\Cluster\Webclust1. This problem occurs when Microsoft Visual C++ version 5.0 is installed.

Solution

To work around this problem, manually type the cluster name when you open the connection.

Problem Node performance is sluggish and node fails.

Cause

CPU may be overloaded.

Solution

Check that your node is not processor-bound. That is, that the CPU is not running at 100-percent utilization. If you try to run too many resources for the node capacity, you can overload the CPU.

Also review the size of your paging file. If the paging file is too small, the Cluster service can detect this as a node failure and fail over the groups.

Problem The cluster log contains numerous resource informational messages (for example: Entered LooksAlive, Entered Open, Entered Offline).

Cause

One or more of your Generic Script resources fills the cluster log with multiple copies of Entered LooksAlive, Entered Open, Entered Offline messages.

Solution

When creating a script for a Generic Script resource, do not use the LogInformation method when calling the LooksAlive function.

For more information, refer to the Microsoft Platform Software Development Kit (SDK).

Problem The Cluster service fails to start and returns an error code of ERROR_SHARING_VIOLATION (32) with event ID 1144 (NM_EVENT_REGISTER_NETWORK_FAILED).

Cause

The Internet Assigned Numbers Authority (IANA)-assigned port (3343) used by the cluster network driver (ClusNet) is bound to another process, preventing the Cluster service from starting.

Solution

Use port scanning and process termination utilities to identify and end the process that is bound to port 3343.

To do this:

1. Open a command prompt.
2. Navigate to the %systemroot%\system32 directory.
3. Type **netstat -a -o** and press ENTER.

This displays all listening and connected ports and the process ID of each process bound to that port. Port 3343 appears for each cluster network on the node.

The `-a` option indicates that all connections and listening ports are supposed to be displayed. Server clusters uses UDP so the ports are normally in listening mode rather than in connections.

The `-o` option indicates that the owning process ID is supposed to be displayed.

4. Type `tasklist`.

This displays the IDs for all the processes running on the node, including the process ID that matches the Cluster service (`ClusSvc.exe`).

5. Type `taskkill /pid ID` and press ENTER to terminate the process(es) bound to port 3343 that do not match the ID for the Cluster service.

Problem

You cannot manually restore the cluster database on a local node by copying the `systemroot\cluster\CLUSDB` file from another node.

Cause

If the cluster registry hive is already locked and loaded by the Cluster service, the operating system will prevent you from copying a local CLUSDB file or overwriting an existing CLUSDB file on another node.

Solution

Stop the Cluster service. Then unload the `HKEY_LOCAL_MACHINE\Cluster` hive before restoring the cluster database file.

To do this:

1. Open a command prompt.
2. Type `net stop clussvc` and press ENTER to stop the Cluster service.
3. Use the Registry Editor to unload the hive under `HKEY_LOCAL_MACHINE\Cluster`. For more information, refer to *Unload a hive from the registry* in the **TechNet** section of the Microsoft website.

The operating system will now allow you to copy the CLUSDB file from a node and manually restore it to another node. For information about how to obtain product support, refer to *Technical support options* in the **TechNet** section of the Microsoft website.

Using Navisphere applications with clusters

Navisphere® client applications must use the cluster or network name (found in the default cluster group), rather than the physical server name, to connect to a cluster and manage any FC4500, FC5300, or FC5700 storage systems connected to that cluster. As long as one server in the cluster is running, you can manage any FC4500, FC5300, or FC5700 storage systems. You manage any FC4700 storage systems in the cluster by managing their SP Agents as you normally would were they not in the cluster.

The server that owns the cluster group (containing the cluster name and IP address) receives requests from the Navisphere client application, and therefore must be running the Navisphere Host Agent. The second server in the cluster does not currently own the cluster group so, unless the cluster group fails over to this server, it does not need a running Host Agent. In order to handle the starting and stopping of the proper Host Agent on the servers you must add a Generic Service resource to the cluster group. (Refer to [“Adding a generic service resource to the Cluster Group” on page 82.](#))

Configuring the Host Agent for clusters

You must set the Windows Service Startup for the Host Agent to **Manual** so that the clustering software is able to use the Generic Service resource to start and stop the appropriate Host Agent. (Refer to [“Adding a generic service resource to the Cluster Group” on page 82.](#))

Configuring the Host Agent

To configure the Host Agent:

1. Set the Service Startup for the Host Agent as follows:
 - a. Open the Control Panel dialog box, and double-click the **Services** icon.

- b. Select **Navisphere Agent**, and then click **Startup**.
 - c. Under **Startup Type**, click **Manual**, and then click **OK**.
2. Click **Save**.
3. When the system prompts you to restart the Agent, click **No** if you are going to install the cluster software now; otherwise, click **Yes** to restart the Host Agent.

Installing the cluster software



IMPORTANT

EMC strongly recommends that you stop the Host Agent service before installing the Windows clustering software. Install the clustering software, as described in the clustering software installation instructions.

Adding a generic service resource to the Cluster Group

Before adding the Generic Service resource to the Cluster Group, we recommend that you install the Navisphere Host Agent on both servers in the cluster. Once installed, configure both Agents to work with clusters as described under [“Configuring the Host Agent for clusters” on page 81](#).

You add the Generic Service resource to the cluster group to stop and start the Host Agent on the appropriate server in the cluster. For example, if the cluster group moves from server A to server B, the Generic Service resource for the Host Agent stops the Host Agent on server A, and then starts the Host Agent on server B. The only noticeable interruption may be a time-out in polling while one Host Agent stops and the other one starts.

To add a generic service resource to the cluster group:

1. Open the **Cluster Administrator** dialog box from one of the servers in the cluster.
2. From the **Cluster Administrator** menu bar, follow the path **File, New, Resources**.
3. Enter the name of the resource that controls the Host Agent.

4. Under **Resource Type**, click **Generic Service**; under **Group**, click **Cluster Group**, and click **Next**.

The two servers with the Agent installed should be listed in the **Possible Owners** list. If not, highlight one or both servers, click the right arrow to move them into the **Possible Owners** list, and click **Next**.

The **Resource Dependencies** list should be empty. If not, highlight all the resources in the list, click the left arrow to move them out of the **Resource Dependencies** list, and click **Next**.

5. In **Service Name**, enter **Navisphere_Agent**, and click **Next**.
6. Click **Finish**.

A message appears stating that the service has been successfully created.

7. Return to the main **Cluster Administrator** dialog box, and follow the path **Clustername, Group, Cluster Group**.
8. Right-click the Navisphere Agent resource, and then click **Properties**.
9. Click the **Advanced** tab, select the **Don't restart** check box and click **OK**.
10. Right-click the Navisphere Agent resource and select **Bring Online**.

The Agent service should start on the server that currently owns the cluster group.

Fibre Channel Attach Environments

This chapter provides information specific to Windows hosts over Fibre Channel.

- ◆ Windows Fibre Channel environment 86
- ◆ Planning for fabric zoning and connections 87
- ◆ Host configuration with Emulex HBAs 88
- ◆ Host configuration with QLogic HBAs 89

Windows Fibre Channel environment

This section lists some Fibre Channel support information specific to the Windows environment.

Also refer to [Chapter 1](#) for information that is common to both Fibre Channel and SCSI configurations.

Hardware connectivity

The Fibre Channel implementation supports up to eight Fibre Channel HBAs per host (depending on the number of available PCI slots).

Boot device support

Windows hosts have been qualified for booting from devices interfaced over Fibre Channel as described in the *EMC Support Matrix* and:

- ◆ [“Host configuration with Emulex HBAs” on page 88](#)
- ◆ [“Host configuration with QLogic HBAs” on page 89](#)

Planning for fabric zoning and connections

Note: This section applies only to fabric configurations, as opposed to direct-attach configurations.

Plan for the switch topology, target to hosts mapping, and zoning. Draw the connectivity between the hosts, switch and the Symmetrix system for a verification of the fabric configuration. After the user has decided on the desired connections, the user must configure zoning capability in the switch before proceeding. Refer to the *Connectrix Enterprise Storage Network System Planning Guide* located on <http://Powerlink.EMC.com> for information on zone configuration.

Host configuration with Emulex HBAs

To install one or more EMC-qualified Emulex host bus adapters into a Windows host and configure the host for connection to Symmetrix over Fibre Channel, follow the procedures in *EMC Fibre Channel with Emulex Host Bus Adapters in the Windows Environment*, available in the EMC OEM section of the Emulex website <http://www.emulex.com> or at <http://Powerlink.EMC.com>.

Host configuration with QLogic HBAs

To install one or more EMC-approved QLogic host bus adapters (HBAs) into a Windows host and configure the host for connection to the Symmetrix over a Fibre Channel, follow the procedures in *EMC Fibre Channel and iSCSI with QLogic Host Bus Adapters in the Windows Environment*, available in the EMC OEM section of the QLogic at <http://www.qlogic.com> or at <http://Powerlink.EMC.com>.

This chapter provides information on the Microsoft iSCSI Initiator and the Microsoft Cluster Server.

◆ Introduction	92
◆ Installing the Microsoft iSCSI Initiator	94
◆ Using MS iSNS server software with iSCSI configurations.....	109
◆ iSCSI Boot with the Intel PRO / 1000 family of adapters	119
◆ Notes on Microsoft iSCSI Initiator.....	128

Introduction

The Microsoft iSCSI Software Initiator package adds support to the Windows operating system for using iSCSI targets that support 1.0 of the iSCSI spec. The Microsoft iSCSI Software Initiator package runs with Windows 2000 SP4 or higher, Windows XP SP1 or higher and Windows 2003 Server or higher.

To use the iSCSI software Initiator kernel mode driver, the system must include a qualified Designed for Windows NIC or and iSCSI HBA.

Terminology

You should understand these terms:

- ◆ **Challenge Handshake Access Protocol (CHAP)** — An authentication method that is used during the iSCSI login in both the target discovery and the normal login.
- ◆ **iSCSI Network Portal** — The host NIC IP address that is used for the iSCSI driver to create a session with the storage.

Software

Node-names

The Microsoft iSCSI initiator supports an iSCSI target that is configured with a node-name as the following rules:

- ◆ Node-names are encoded in UTF8 Character set.
- ◆ The length of a node-name should be 223 characters or less.
- ◆ The valid characters are:
- ◆ A name can include any of these characters:
 - **a** through **z** (upper or lower case; uppercase characters are always mapped to lowercase)
 - **0** through **9**
 - **.** (period)
 - **-** (dash)
 - **:** (colon)

Refer to the *Microsoft iSCSI Initiator x.0 User's Guide* for the complete set of rules for setting up the valid initiator and target node-names.

Boot device support

The Microsoft iSCSI initiator does not support booting the iSCSI host from iSCSI storage. Refer to the *EMC Support Matrix* for the latest information about boot device support.

Installing the Microsoft iSCSI Initiator

The Initiator can be found at:

<http://www.microsoft.com/downloads>

At the beginning of the install process, the **Software Update Installation Wizard** screen displays (Figure 10).

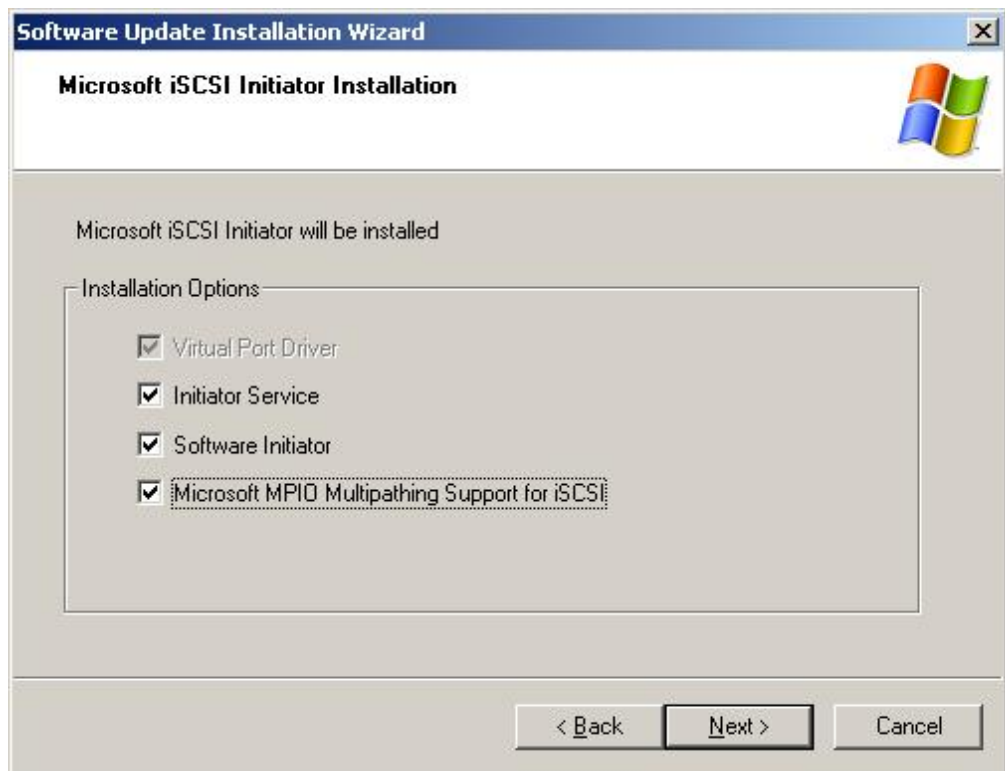


Figure 10 Software Update Installation Wizard

Four installation options are available:

- ◆ Virtual Port Driver (iscsiprt)
This is always checked and cannot be unchecked. All configurations require the port driver and thus it is always installed.
- ◆ Initiator Service (iscsiexe.exe)
This is the kernel mode iSCSI software Initiator driver and is used to connect to iSCSI devices via the Windows TCP/IP stack using NICs. If this option is selected then the Initiator Service option is automatically selected.
- ◆ Software Initiator (msiscsi.exe)
This is the kernel mode iSCSI software Initiator driver and is used to connect to iSCSI devices via the Windows TCP/IP stack using NICs. If this option is selected then the Initiator Service option is also selected automatically.
- ◆ Microsoft MPIO Multipathing Support for iSCSI
This installs the core MS MPIO framework files and the Microsoft iSCSI Device Specific Module (DSM). This will enable the MS iSCSI software Initiator and HBA to perform **session-based multipathing** to a target that supports multiple sessions to a target. If the version of MS MPIO core files is later than the version installed on the computer, the core MS MPIO files will be upgraded to the latest version that is part of the installation package.

Adding or removing components

To add or remove specific MS iSCSI Software Initiator components, run the setup package executable and configure the checkboxes to match the desired installation. The application should autocheck the boxes for components that are already installed. For example, if you want to add the MS MPIO component then you would leave the other checkboxes alone and just check the MS MPIO checkbox.

Note: If the MS MPIO checkbox is *not* checked then the installer will attempt to uninstall the Microsoft iSCSI DSM and the core MS MPIO files. However if there is another DSM installed then the core MS MPIO files will not be uninstalled. The setup application determines if another DSM is installed by checking the MS MPIO supported device list.

The installation will put an icon on your desktop labeled **Microsoft iSCSI Initiator**. To open the iSCSI Initiator Properties window (Figure 11 on page 96), double-click on the icon.

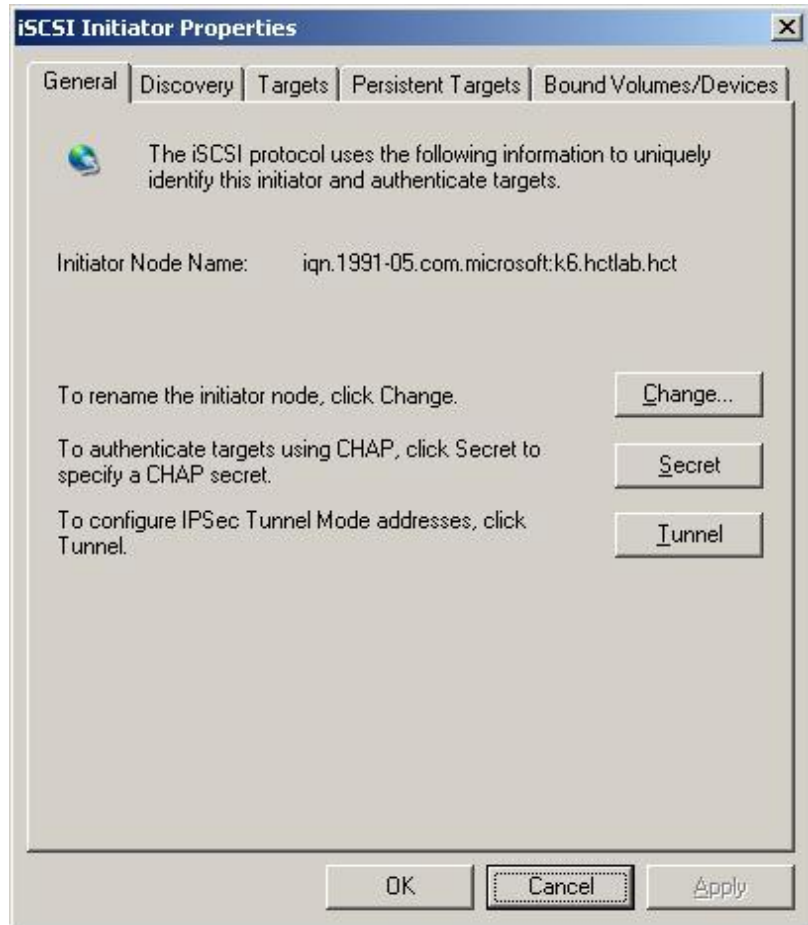


Figure 11 iSCSI Initiator Properties window

The **General** tab (shown in Figure 11) lists the iqn name of your host. This will be the name presented to your storage arrays. You could think of it as the WWN.

Please note the following:

- ◆ There will only be one iqn name for each host even if you have multiple NICs installed.

- ◆ Microsoft requires that the Initiator Service be installed if you are using an iSCSI HBA. Only the service needs to be installed when using an HBA, not the software Initiator.
- ◆ If you are installing the Initiator Service to be used in conjunction with an iSCSI HBA then the IQN name of the HBA is changed to the iqn name shown on the **General** tab.
- ◆ If you have two HBAs installed in your host, BOTH will get the same iqn name.
- ◆ Precautions should be taken when using an iSCSI HBA and the Initiator to boot from array.

For directions on how to boot from array, consult the *EMC Fibre Channel and iSCSI with QLogic Host Bus Adapter Guide*, available at <http://www.qlogic.com> or at <http://Powerlink.EMC.com>.

The **Discovery** tab (Figure 12) shows the two ways to discover targets with the Initiator.

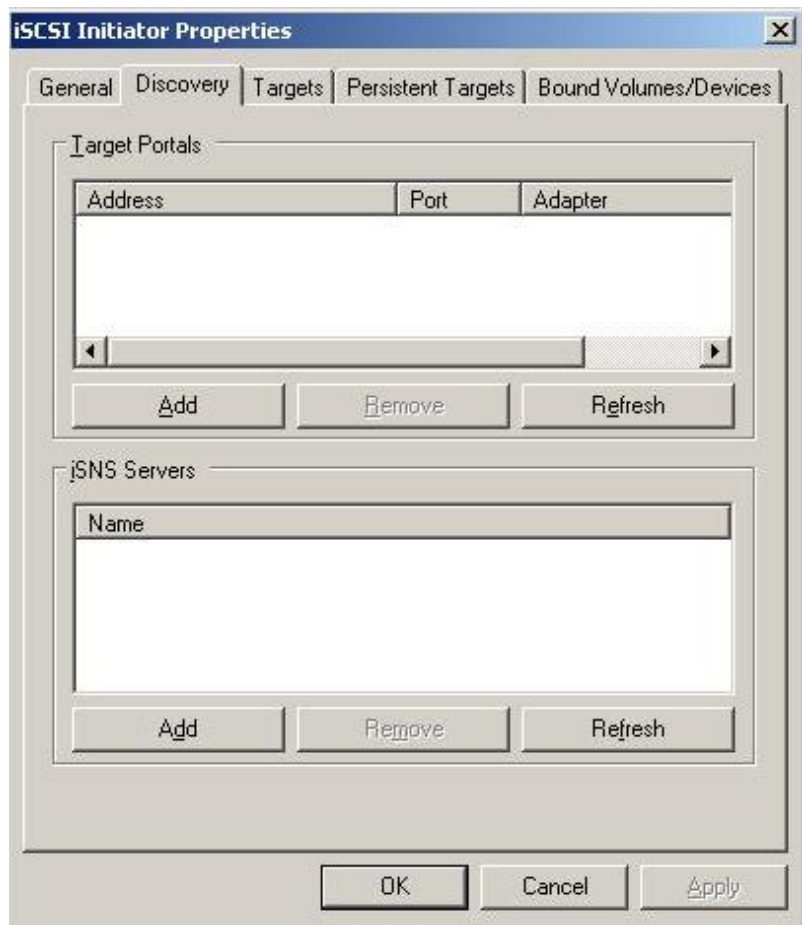


Figure 12 iSCSI Initiator Properties, Discovery tab

The two ways to discover targets with the Initiator are:

- ◆ Explicitly enter the IP address of the target portal.
- ◆ Enter the IP address of the iSNS server.

iSNS is used as a way to register both targets and initiators on a common server. If you enter an IP address for the iSNS server the host systems will register themselves with the iSNS server and will retrieve a list of targets that you can connect to.

More information on iSNS can be found at:

<http://www.microsoft.com/downloads>

Examples

In the following example, we are logging into the target portal at 192.168.150.108 and also registered with the iSNS server at 192.168.8.1.

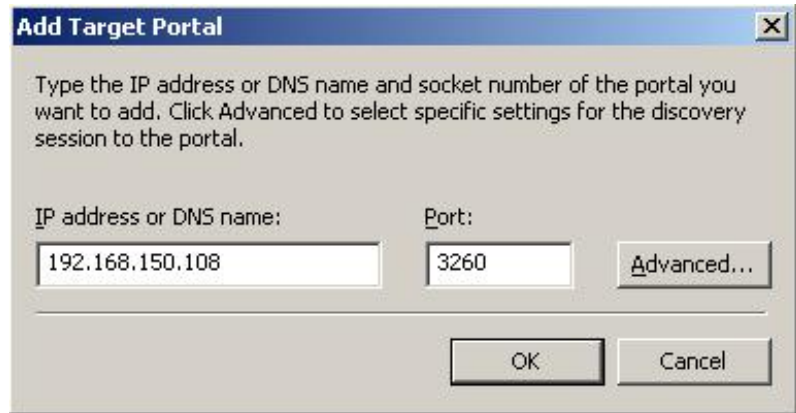


Figure 13 Add Target Portal dialog box

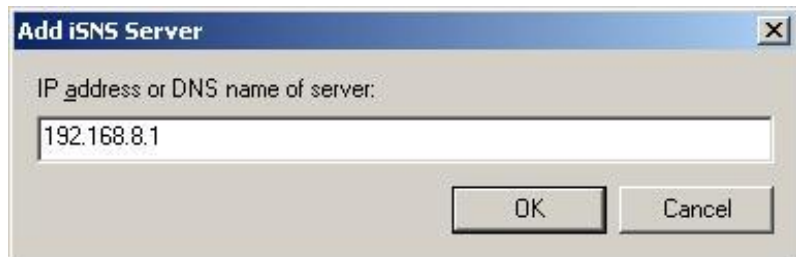


Figure 14 Add iSNS Server dialog box

Through these two discovery mechanisms we get the following information on the **Targets** tab (Figure 15 on page 100).

- ◆ The first entry is a target we discovered by explicitly entering the target portal IP address.
- ◆ All of the other targets listed below are provided by the iSNS server.

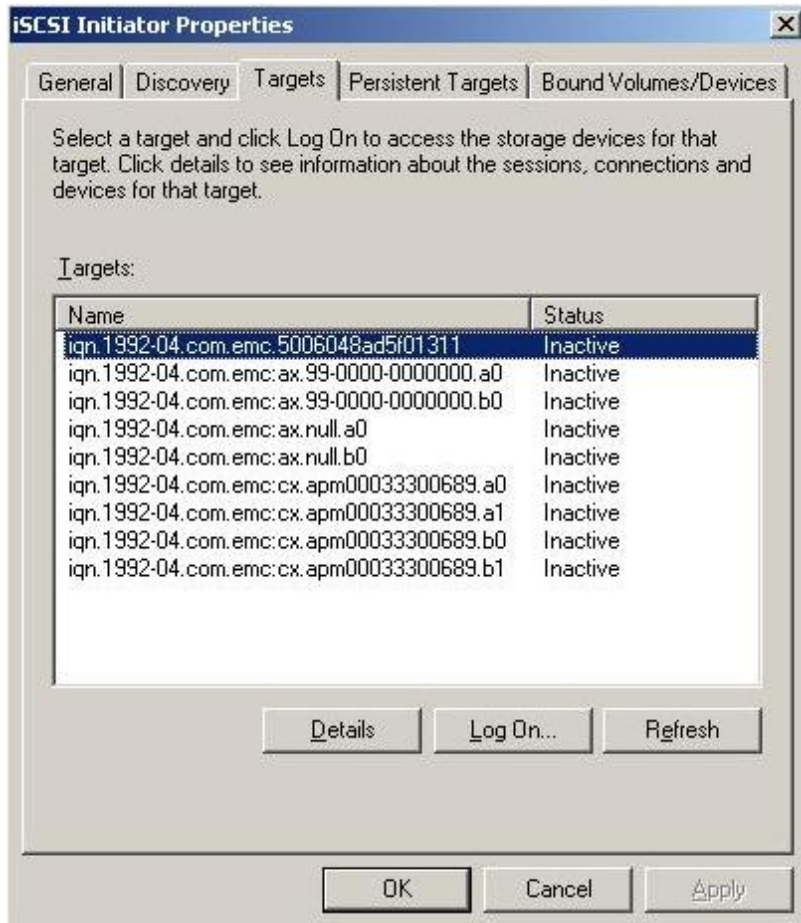


Figure 15 iSCSI Initiator Properties, Targets tab

We will log into the 6th target listed above by highlighting it and clicking **Log On**. The **Log On to Target** dialog box appears ([Figure 16 on page 101](#)).

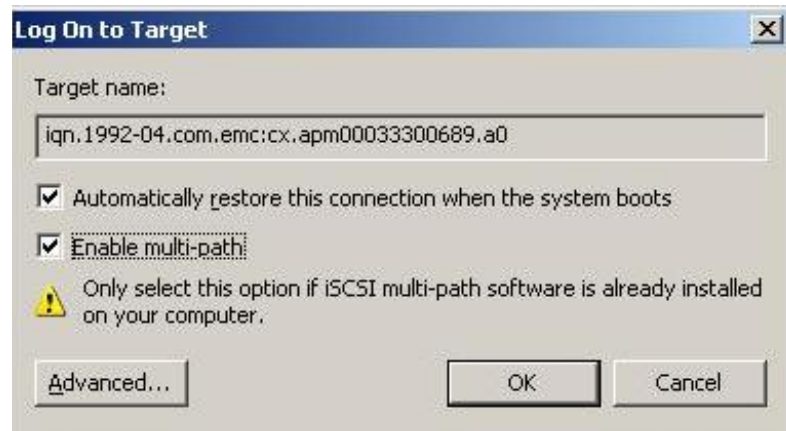


Figure 16 Log On to Target dialog box

Checking **Automatically restore this connection when the system boots** will cause the Initiator to log in to the target after each reboot. If you don't check this box, then the host will not log in to the target and you will lose access to the disk presented from that target.

Checking **Enable multi-path** allows you to create multiple sessions to the target using MPIO. Although you installed the Initiator with Microsoft MPIO Multipathing Support for iSCSI at install time, you still need to check this box if you want to use multipathing with this target. It is recommended to check this option if you are using EMC PowerPath® 4.6.x.

Clicking the **Advanced** button brings up the **Advanced Settings** window (Figure 17).

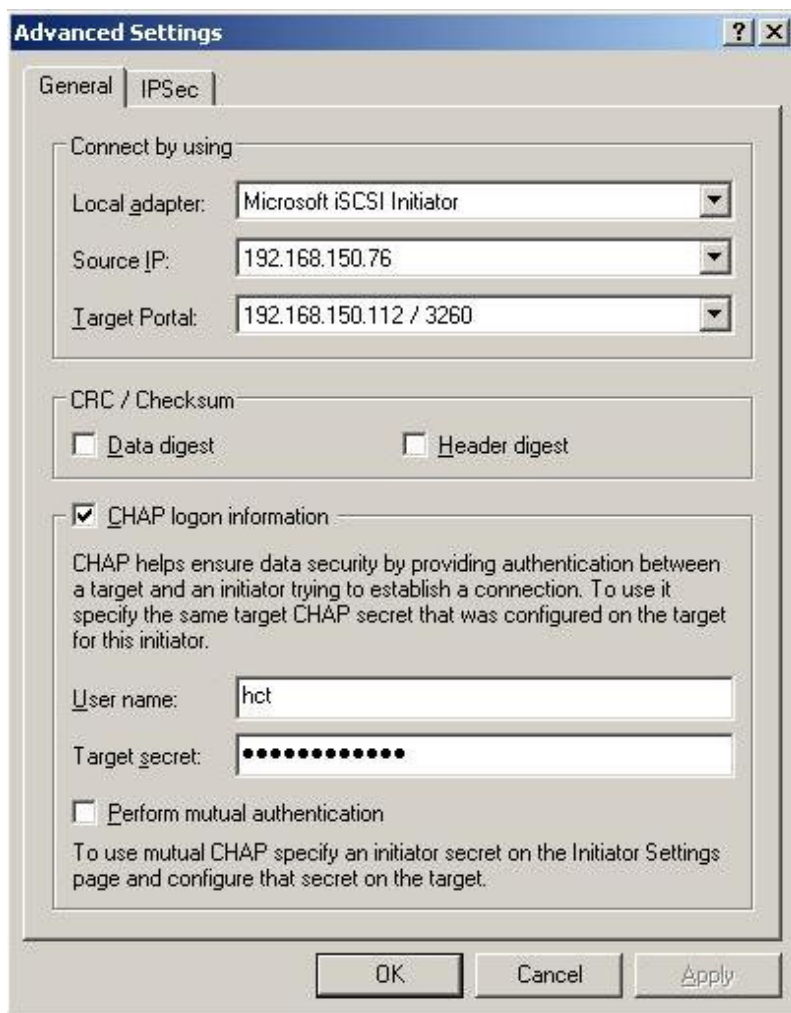


Figure 17 **Advanced Settings window**

In the **Connect by using** section you can select the **Local Adapter**, the **Source IP** and the **Target Portal**. The source IP is the IP address assigned to the NIC installed in the host. If you have multiple NICs, they will be listed here. The Target Portal is the address assigned to the target name you are trying to log in to.

In the **CHAP logon information** section, CHAP logon information is checked. CHAP is a protocol used to authenticate the peer of a connection and is based upon the peer sharing a password or secret. CHAP requires the Initiator to have both a username and secret in order to operate. The CHAP username is typically passed to the target and the target will lookup the secret for that username in its private table. By default the Microsoft iSCSI Initiator service will use the initiator node name as the CHAP username but this can be changed on the array. In this example there is a CHAP username "hct" configured on the array so we connect to the target using HCT and the Target Secret.

The **IPSec** tab is used to configure the parameters for IPSec. This protocol provides authentication and data encryption at the IP packet layer.

Note: At this time IPSec is not supported.

Click **OK** to connect to the target ([Figure 18 on page 104](#)).

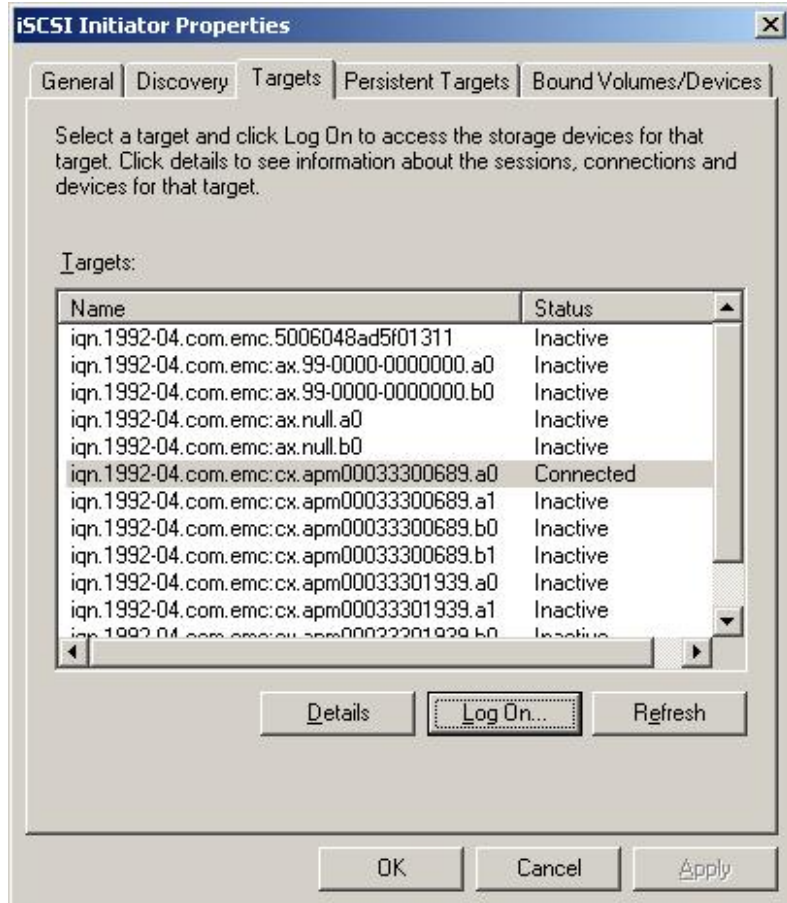


Figure 18 iSCSI Initiator Properties, Targets tab

The **Persistent Targets** tab (Figure 19) will list all the targets that you have chosen to automatically connect to after a reboot.

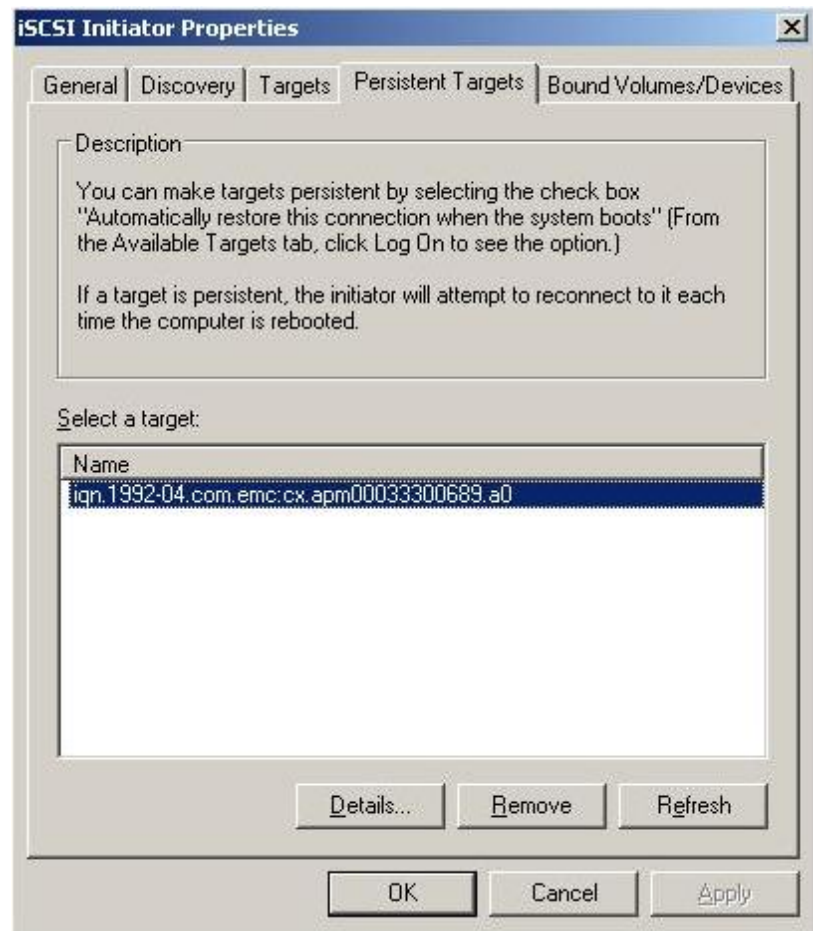


Figure 19 iSCSI Initiator Properties, Persistent Targets tab

The disk startup sequence in Windows, when using the Microsoft iSCSI software Initiator kernel mode driver, is different from the startup sequence when using an iSCSI or other HBA. Disks exposed by the Microsoft iSCSI Initiator kernel mode driver are available for applications and services much later in the boot process and in some cases might not be available until after the service control manager begins to start automatic start services. The Microsoft iSCSI Initiator service includes functionality to synchronize automatic start services and the appearance of iSCSI disk. The iSCSI service can be configured with a list of disk volumes that are required to be present before the start of automatic start services.

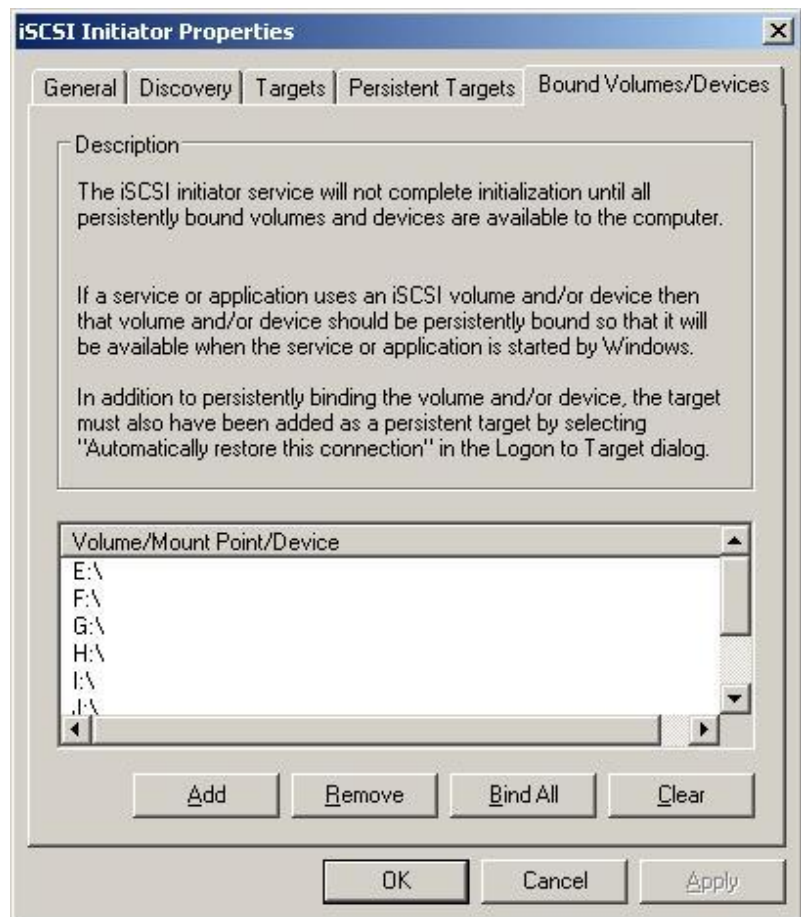


Figure 20 iSCSI Initiator Properties, Bound Volumes/Devices tab

Upgrading from version v2.x to v2.03

Refer to the release notes that come with the install package for 2.03 for the changes made to this version.

If iSCSI 2.0 or higher is installed, the installation of version 2.03 will update just the following files that include fixes:

msiscsi.sys	version 5.2.3790.3099
iscsiwmi.dll	version 5.2.3790.3099
relnotes.txt	version 5.2.3790.3099
mpdev.sys	version 1.17.3790.1759
mpdev.inf	version 1.17.3790.1759
mpspfltr.sy	version 1.17.3790.1759

Please note that the size and date of mpio binaries in the iSCSI Initiator package will not appear to be exactly the same as the size of the files released as 1.17. The build date of the 1.17 binaries maps to the build date of the iSCSI Initiator package. However, the 1.17 binaries included are the same as binaries released as MS MPIO 1.17.

Note: If **Microsoft MPIO Multipathing Support for iSCSI** was selected at install time then a reboot will be required. If you are upgrading the Initiator on a system that is part of a cluster it is recommended that you move all cluster resources off of the node you are upgrading.

Determining the Initiator version

To find the version of the Initiator currently running on your system:

From a command prompt issue the command: **iscsicli versioninfo** and compare the results to the table below:

version 2.0	1653
version 2.02	1941
version 2.03	3099

Uninstalling the Initiator

To completely uninstall the MS iSCSI Software Initiator package, go to the **Add/Remove** program applet in the control panel and click on the **Remove** button for the MS iSCSI Software Initiator package. The uninstall will completely uninstall the iSCSI Initiator package including the kernel mode driver, Initiator service and MS MPIO support, although if there is another DSM then the core MS MPIO files will not be uninstalled. This means that if you have PowerPath 1.1 or 4.6 loaded at the time you uninstall the Initiator, then the MPIO framework will remain on the host.

Using MS iSNS server software with iSCSI configurations

The Microsoft iSNS Server is a Microsoft Windows service that processes iSNS registrations, deregistrations, and queries via TCP/IP from iSNS clients, and also maintains a database of these registrations. The Microsoft iSNS Server package consists of Windows service software, a control-panel applet, a command-line interface tool, and a set of WMI interfaces. Additionally, there are DLLs allowing Microsoft Cluster Server to manage Microsoft iSNS Server as a cluster resource.

When configured properly, the iSNS server allows iSCSI initiators to query for available iSCSI targets that are registered with the iSNS server. The iSNS server also allows administration of iSCSI networks by providing a form of “zoning” in order to allow initiators access only to targets designated by the administrator.

Installing iSNS server software

Microsoft iSNS server is available for download from Microsoft's website in the downloads section. Currently, there is no direct link to the iSNS software.

Prior to running the installation, it is recommended that your iSCSI network interface controller (NIC) be configured to work with your iSCSI network. Symmetrix and CLARiiON iSCSI interfaces must be configured to recognize and register with the iSNS server software. Refer to the *DMX MPCD For iSCSI Version 1.0.0 Technical Notes* (P/N 300-001-170) for information on the configuration of the MPCD for iSCSI into Symmetrix DMX systems. CLARiiON configuration is done by an EMC Customer Engineer (CE) through Navisphere Manager. The CE will configure your CX-Series system settings for each iSCSI port.

After your storage array target ports are configured, install the iSNS server software by starting the installation package downloaded from the Microsoft website.

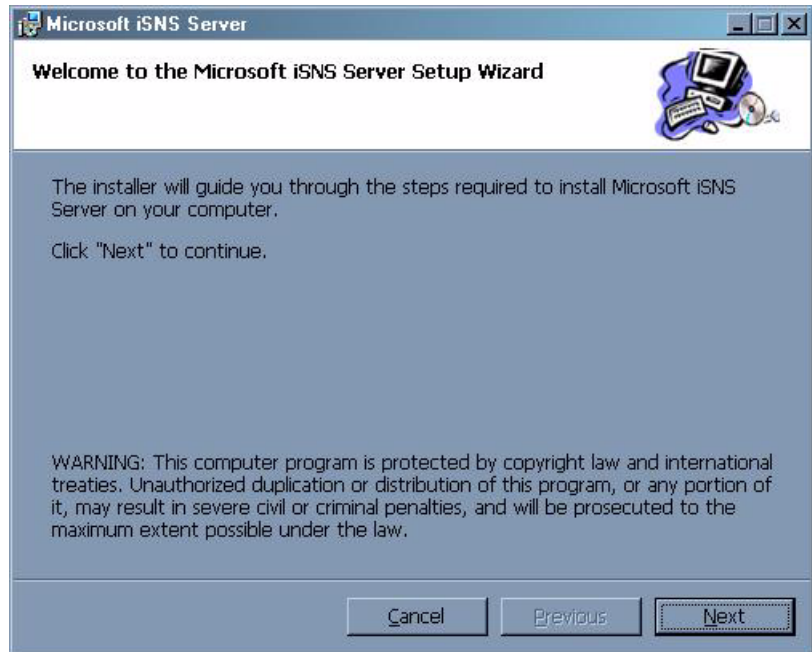


Figure 21 Microsoft iSNS Server Installation Wizard

Follow each step of the iSNS Server software installer ([Figure 21](#)). You are required to accept a user license agreement and to select a location to install the software. Choosing the defaults for each step on the wizard is recommended.

After files are copied, the installation options dialog appears. Here you have the option of installing the iSNS service, uninstalling the iSNS service, or doing nothing. For installation, leave the **Install iSNS Service** button checked, and click **OK** ([Figure 22 on page 111](#)).

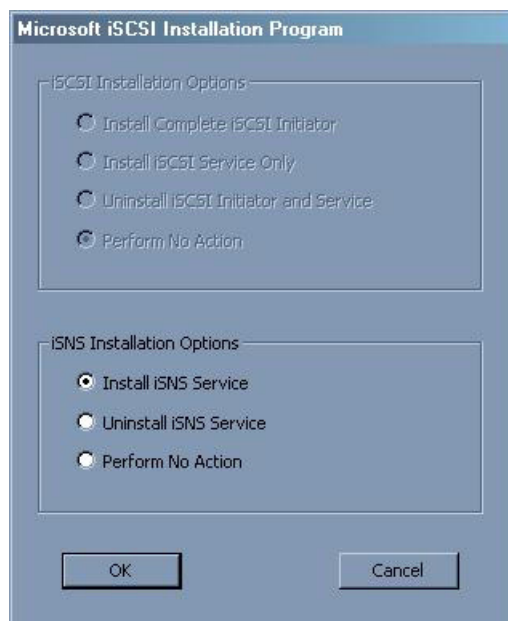


Figure 22 Installation option dialog

Another license agreement dialog appears. You must agree to the license agreement before installation can finish. After the iSNS Service is installed, you are prompted to allow the iSNS DHCP option to be configured (Figure 23). EMC does not recommend using DHCP for iSCSI configurations, but rather using static IP addresses for iSCSI targets and initiators. Select **No** to continue installation.

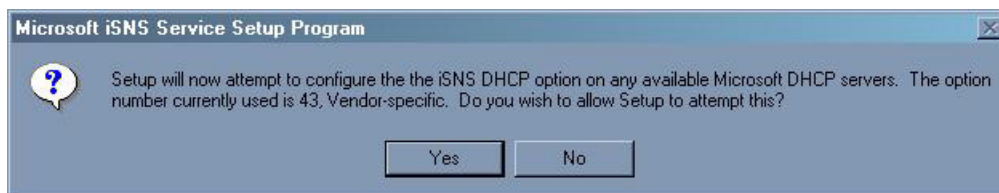


Figure 23 iSNS DHCP configuration option dialog

The installation program displays a message (Figure 24) stating that the iSNS service was installed and started.



Figure 24 **Installation confirmation message**

Click **OK** to see the installation summary screen. Click **Next** to complete installation.

Configuring iSNS server

Using the installed desktop icon, start the iSNS Server GUI interface. When open, the **General** tab (Figure 25) shows any targets and initiators that have registered with the iSNS server.

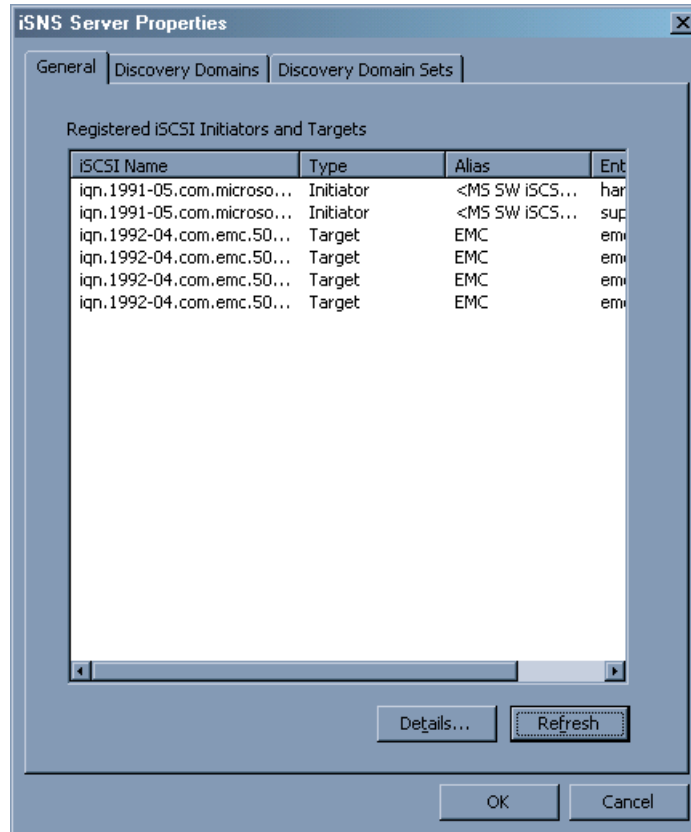
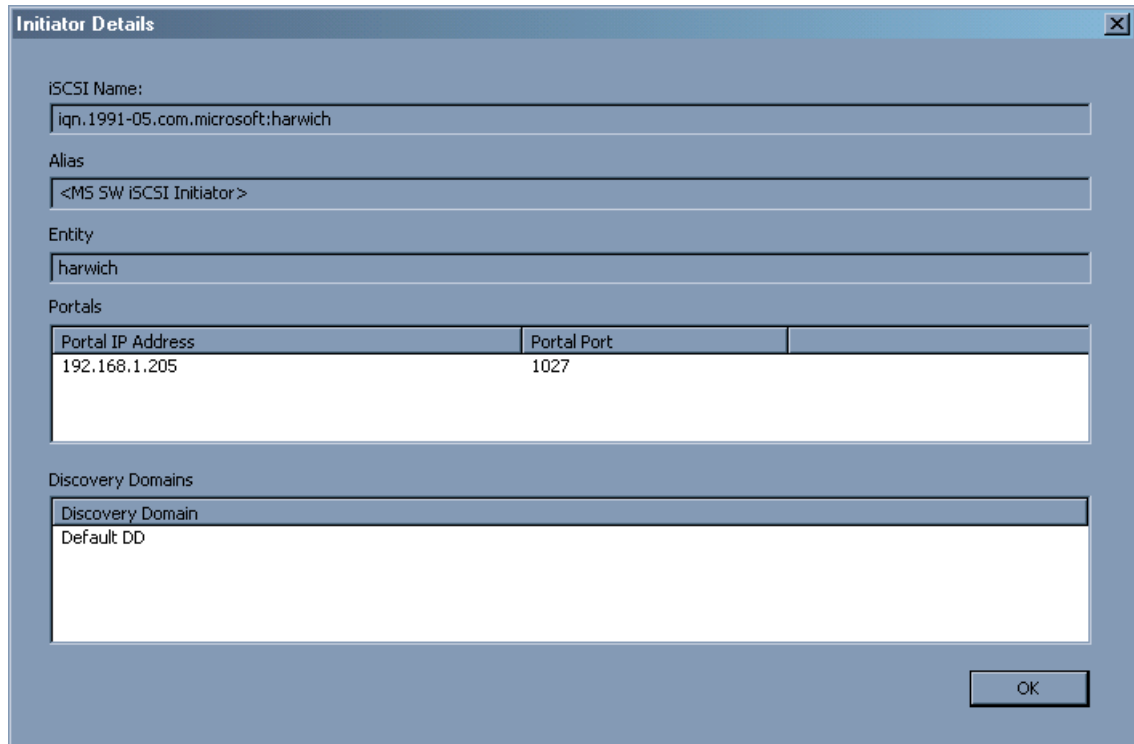


Figure 25 iSNS General properties

By selecting on any individual target or initiator and clicking **Details**, more information about the selected object is displayed (Figure 26 on page 114).



The image shows a Windows-style dialog box titled "Initiator Details". It contains several input fields and a table. The fields are for iSCSI Name, Alias, Entity, and Portals. The Portals section contains a table with two columns: "Portal IP Address" and "Portal Port". The Discovery Domains section contains a list box with "Default DD" selected. An "OK" button is at the bottom right.

Portal IP Address	Portal Port
192.168.1.205	1027

Figure 26 Target/Initiator Details dialog

Using discovery domains for iSNS

The idea of Discovery Domains (DD) in iSNS is to provide a means for partitioning and grouping registered storage nodes for the dual purposes of administrative control and discovery. Microsoft iSNS Server implements the optional Default DD and automatically adds all newly registered storage nodes into the Default DD. However, if a storage node is already a member of a non-Default DD, then it is not automatically placed in the Default DD when it is registered. By automatically adding newly registered storage nodes to the Default DD, the default behavior is that all newly registered storage nodes can discover each other. It is up to an administrator to subsequently partition storage nodes into non-Default Discovery Domains.

1. To create a new DD, click the **Discovery Domains** tab (Figure 27), and then click **Create**.

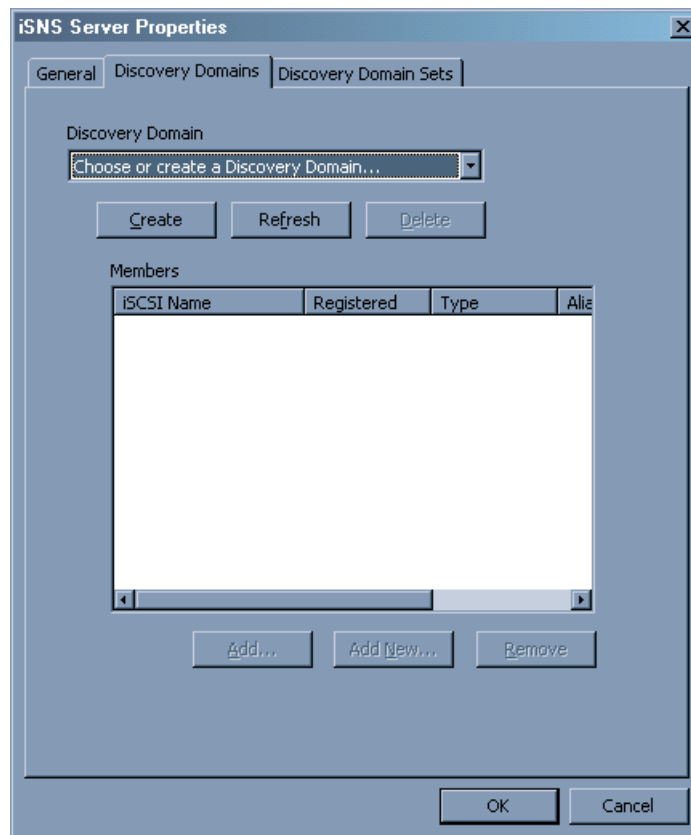


Figure 27 iSNS Server Properties, Discovery Domains tab

2. Type in the name of the DD you want to create (Figure 28), and click **OK**.

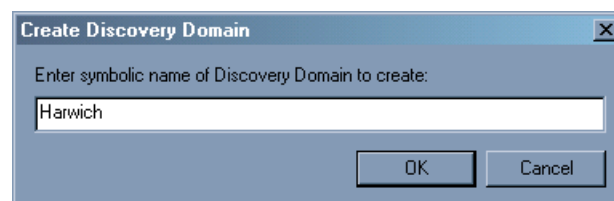


Figure 28 Create Discovery Domain dialog

3. With the DD name selected on the **Discovery Domains** tab (Figure 29), click **Add** to add target and initiator members to the DD (Figure 30 on page 117).

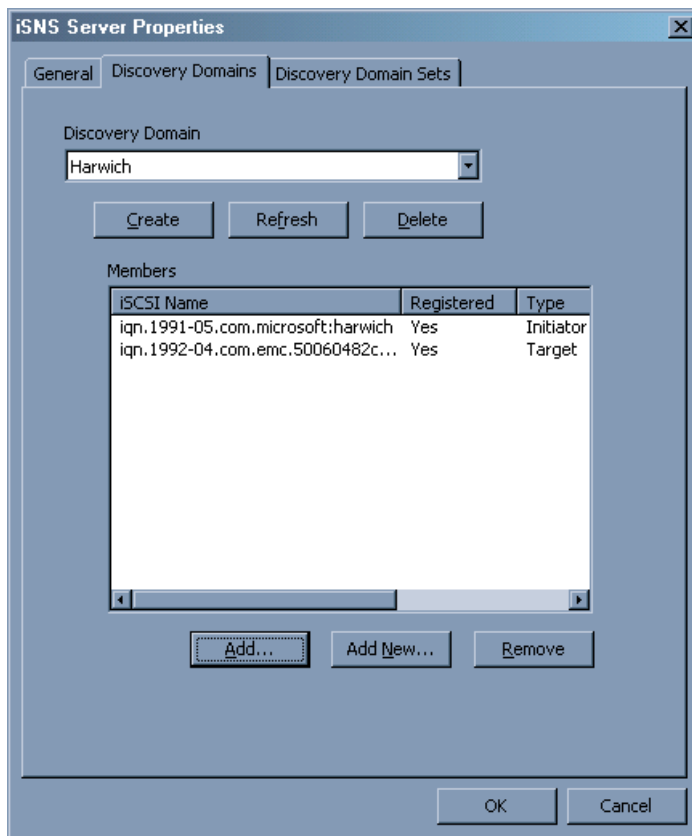


Figure 29 iSNS Server Properties, Discovery Domain with members added

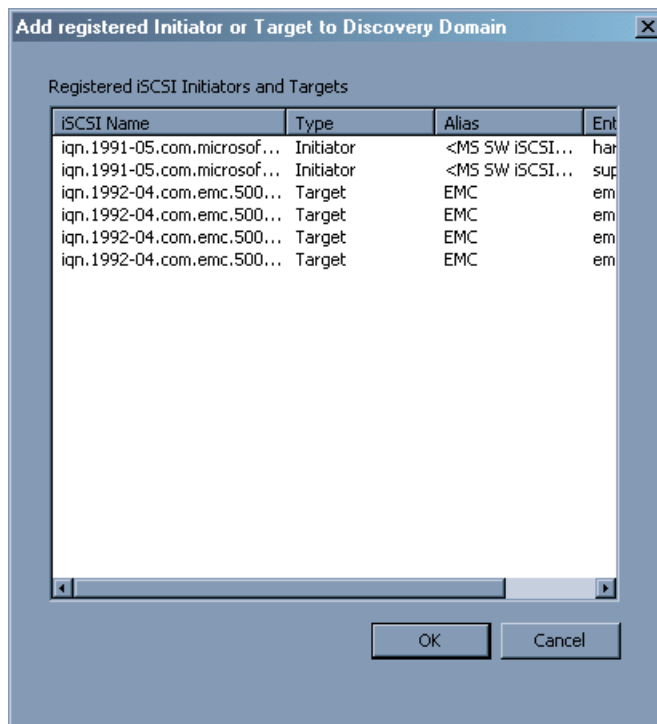


Figure 30 Add registered Initiator or Target to Discovery Domain

4. After your DD is created, you must add it to a Discovery Domain set. Click the **Discovery Domain Sets** tab (Figure 31). Select the Default DDS (Discovery Domain Set) and click **Add** to add your created DD to the DDS.

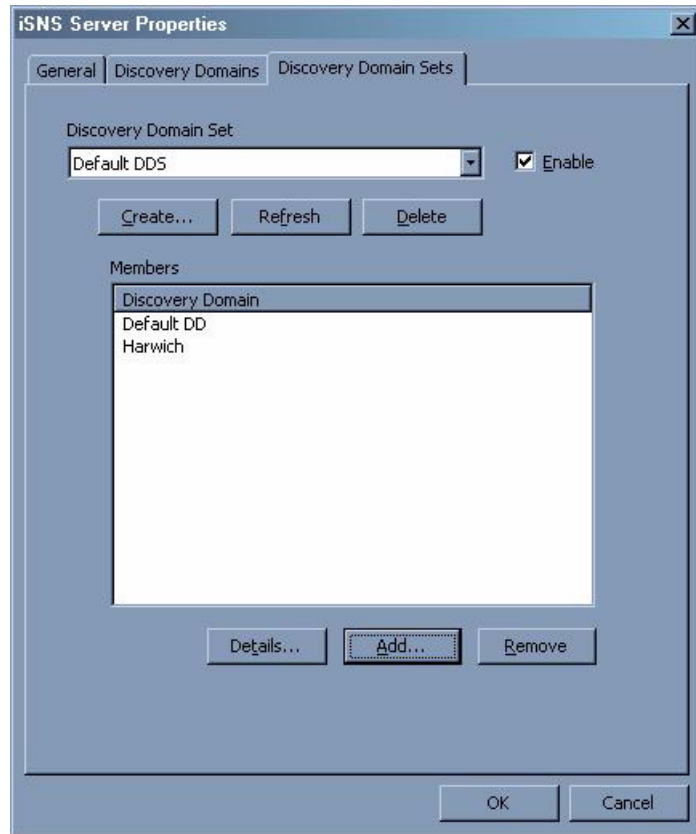


Figure 31 iSNS Server Properties, Discovery Domain Sets tab

Be sure that the desired DDS is enabled by selecting it from the pull down menu and checking the **Enable** box.

These are the basic features that most iSCSI/iSNS users will use in their iSCSI environments. For further features included in the iSNS Server package, refer to the *Microsoft iSNS Server Users Guide* available for download with the iSNS Server software.

iSCSI Boot with the Intel PRO/1000 family of adapters

Intel iSCSI Boot is designed for the Intel PRO/1000 family of PCI-Express Server Adapters. Intel iSCSI Boot provides the capability to boot from a remote iSCSI disk volume located on an iSCSI-based Storage Area Network (SAN).

The basic steps to configuring boot from SAN are:

1. Prepare your storage array for boot from SAN.
2. Install boot-capable hardware in your system.
3. Install the latest Intel iSCSI Boot firmware using the iSCSIUtil DOS utility.
4. Connect the host to a network with the contains the iSCSI target.
5. Configure the iSCSI boot firm ware on the NIC to boot from a pre-configured iSCSI target disk.
6. Configure the host to boot from the iSCSI target.

This section will focus on preparing your array to boot from SAN. The steps listed above are documented on the Intel website and can be downloaded from this link:

http://downloadfinder.intel.com/scripts-df-external/Detail_Desc.aspx?agr=Y&ProductID=2249&DwnldID=12194&strOSs=All&OSFullName=All%20Operating%20Systems&lang=eng

You can also find a list of supported adapters at the above link.

Preparing your storage array for boot

This section explains how to prepare your array in order to successfully present a boot LUN to you host.

The first thing you need to consider is what the host name will be. Using the naming conventions explained in “[Node-names](#)” on [page 92](#), record an appropriate iqname for your host.

In the following example, we will use an iqname of **iqn.1992-05.com.microsoft:intel.hctlab.hct**.

Configuring your CX3 for iSCSI boot

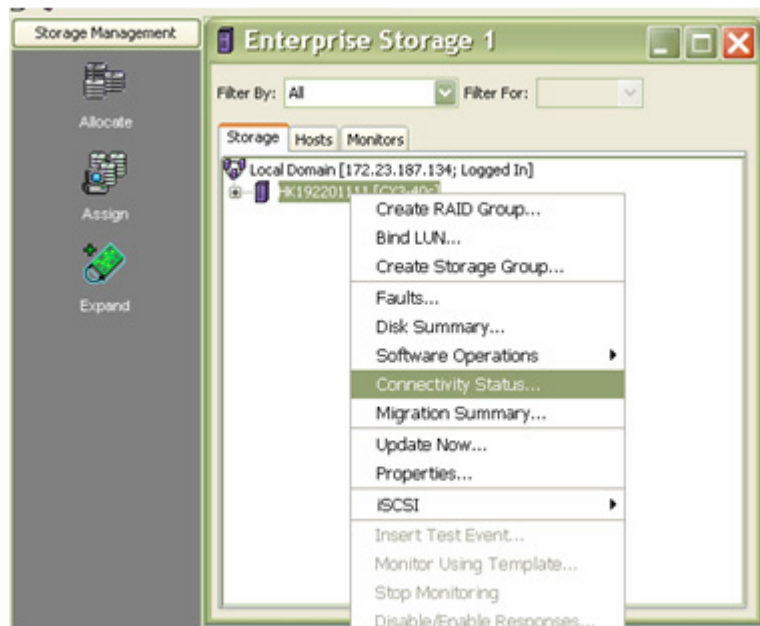
Note: The following example assumes that you are familiar with Navisphere Manager.

In order for boot from SAN to work properly, you first need to present a LUN to the host. To do this on a CX3 using Navisphere Manager:

1. Create new initiator records that identify the host to the array.
2. Create a record for each SP port to you might potentially connect.
3. Create a storage group with the new server and boot LUN. This LUN should be sized properly in order for the OS, and any other applications, to fit properly.

Once this is complete, follow these steps:

1. Right-click on the Storage Array and then select **Connectivity Status**.



2. Click **New** to display the **Create Initiator Record** window.

Create Initiator Record

Initiator Information

Initiator Name: iqn.1991-05.com.microsoft:intel.hctlab.hct

SP - port: B-0 (iSCSI)

Initiator Type: CLARION Open

Unit Serial Number: Array

☒ ArrayCommPath

Failover Mode: 1

HBA Information

Vendor: Model:

Host Information

☒ New Host ☐ Existing Host

Host Name: intel.hctlab.hct

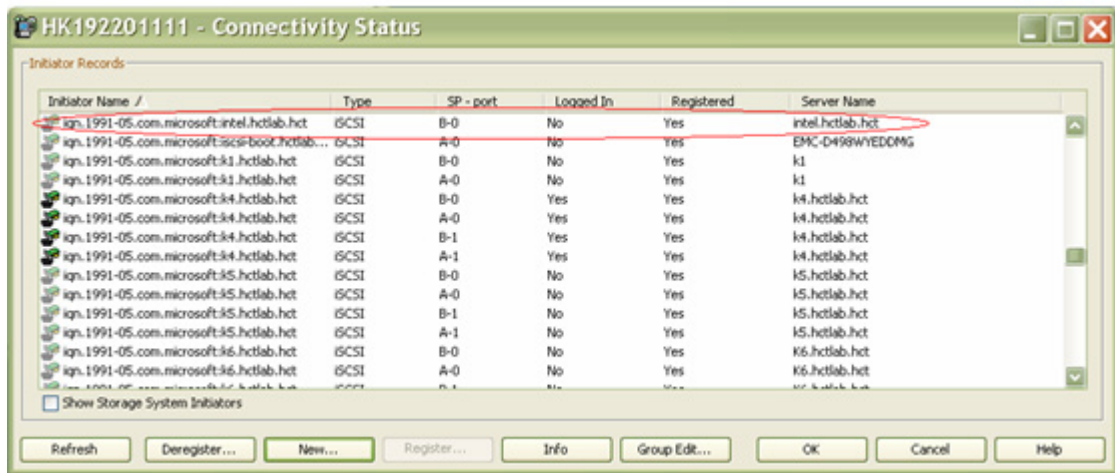
Host: iqn.1991-05.com.microsoft:mongoose-md3

IP Address: 192 . 168 . 1 . 1

OK Cancel Help

3. In the **Create Initiator Record** window:
 - a. Enter the iqn name of the host in the **Initiator Name** field. For our example, use iqn.1992-05.com.microsoft:intel.hctlab.hct.
 - b. Select the SP - port to which the host will connect.
 - c. Enter a host name and ip address in the **Host Information** section.
 - d. Click **OK**.

Once these steps are complete, the host displays in the **Connectivity Status** window.



4. Remember that the above steps create an initiator record that has a path to B0. To create additional initiator records, repeat steps 2 and 3 on page 121.
 - a. Select a different SP - port in the drop-down menu.
 - b. Instead of entering a new host name, select **Existing Host** and choose the host created during the creation of the initial initiator record.

Note: This is critical to provide uninterrupted access if an array side path failure should occur.

5. Once you have created initiator records for each SP port, you will be able to create a storage group as you normally would and assign the newly created host and a boot LUN.

At this point you have configured the array to present a boot LUN to the host at boot up. You can continue with the instructions documented by Intel to install your OS to a local hard drive and then image your host OS to the boot LUN assigned above.

Post installation information

This section contains the following installation:

- ◆ [“Using two Intel NICs in a single host”](#), next
- ◆ [“PowerPath 4.6”](#), on this page
- ◆ [“Known issue” on page 125](#)

Using two Intel NICs in a single host

The above process in [“Configuring your CX3 for iSCSI boot”](#), beginning on [page 119](#), creates connections that can be accessed by the host on all ports available on the CX3 (in our example ports A0, A1, B0, and B1).

By using a dual-port PRO / 1000, or two single-port PRO / 1000s, the Intel BIOS will allow you to set up one port as the primary and another as the secondary. By configuring the primary login to connect to one SP, and the secondary login to connect to the other SP, your host will have access to *both* SPs .

Note: You do not need to configure Microsoft iSCSI Software Initiator for Windows to be able to detect the iSCSI Disk. Microsoft iSCSI Software Initiator automatically retrieves the iSCSI configurations from the PRO / 1000 adapter iSCSI Boot firmware.

PowerPath 4.6

Once you are successfully booting from your array, you can install PowerPath 4.6. With Intel iSCSI Boot from SA, PowerPath 4.6 is limited to providing array-side protection for path failure and SP events.

Before installing PowerPath you will need to make sure your paths between the NICs and the array ports are set up properly. To do this, use a combination of the Intel BIOS logins, documented in the Intel guide in the "Firmware Setup" section, along with additional target logins using the Microsoft Initiator. What you are trying to accomplish is a path setup that looks much like what is discussed in [“Using the Initiator with EMC PowerPath v4.6.x” on page 129](#).

When complete, your paths will look like those shown in [Figure 32 on page 124](#).

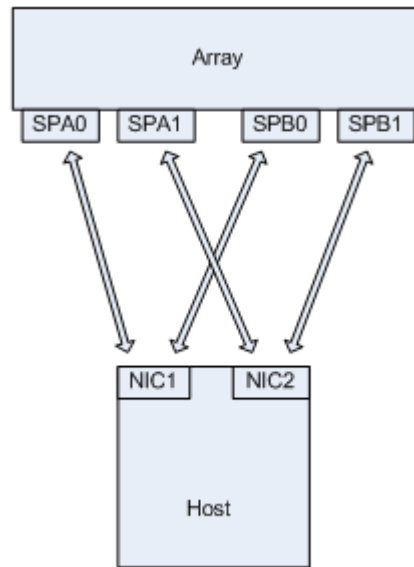


Figure 32 Four paths

Setting this up will be slightly different than what is discussed in [“Using the Initiator with EMC PowerPath v4.6.x” on page 129](#).

Remember that you have already created two paths by configuring the Intel BIOS with primary and secondary logins. So, for example, if you configure the Intel BIOS to connect to A0 and B1, once you boot your host the Microsoft Initiator will show two connected logins to port A0 and B1 on the **Target** tab.

In order to complete the path setup you need to use the process beginning with [Step 1 on page 131](#). When complete, you will have something similar to the following:

- NIC1 → A0
- NIC2 → A1
- NIC1 → B0
- NIC2 → B1

You can now install PowerPath 4.6.x. Once installed, PowerPath administrator will look similar to [Figure 33](#). Failures on the array side (i.e., loss of path between SP port and switch and SP failures) will be managed correctly by PowerPath.

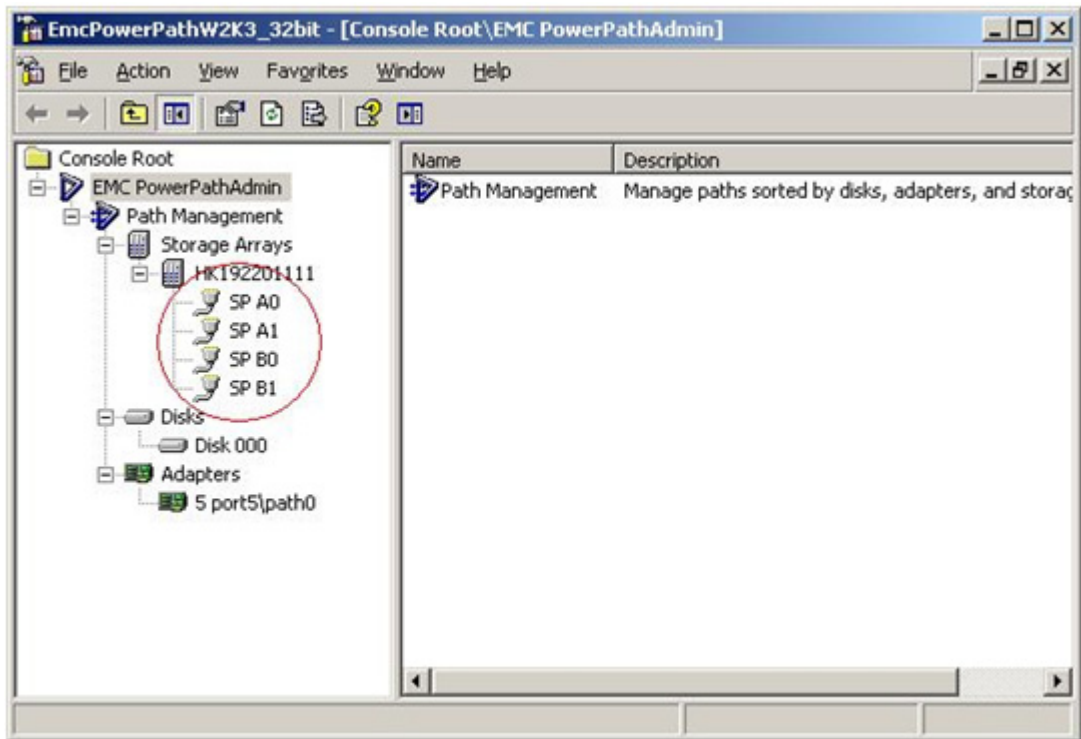


Figure 33 EMC PowerPathAdmin

Known issue Slow performance may occur during network congestion when RFC 1122-delayed acknowledgements extend the error recovery process. In these situations, the default 200 millisecond delay on the acknowledgement can significantly impact read bandwidth. PowerPath, by load-balancing read requests across multiple array ports, increases the likelihood that simultaneous read completions from multiple ports will result in network congestion. This increases the likelihood of experiencing the problem.

As specified in RFC 1122, Microsoft TCP uses delayed acknowledgments to reduce the number of packets that are sent on the media. Instead of sending an acknowledgment for each TCP

segment received, TCP in Windows 2000 and later takes a common approach to implementing delayed acknowledgments. As data is received by TCP on a particular connection, it sends an acknowledgment back only if one of the following conditions is true:

- ◆ No acknowledgment was sent for the previous segment received.
- ◆ A segment is received, but no other segment arrives within 200 milliseconds for that connection.

Typically, an acknowledgment is sent for every other TCP segment that is received on a connection unless the delayed ACK timer (200 milliseconds) expires. You can adjust the delayed ACK timer by editing the registry as outlined in the workaround below.

Workaround

Modify the TCP/IP settings for the network interfaces carrying iSCSI traffic to immediately acknowledge incoming TCP segments. This workaround solves the read performance issue.

The procedure to modify the TCP/IP settings is different for Windows 2000 servers and Windows 2003 servers. Follow the directions appropriate for the version you are running on your servers. The procedures for Windows servers 2003 SP1 or later begins on [page 127](#).

Note: These TCP/IP settings should not be modified for network interfaces not carrying iSCSI traffic as the increased acknowledgement traffic may negatively affect other applications.

Microsoft iSCSI clusters also require array software revision 3.22.xxx.5.511 or 3.24.xxx.5.007 or later.



CAUTION

This workaround contains information about modifying the registry. Before you modify the registry, be sure to back it up and make sure that you understand how to restore the registry if a problem occurs. For information about how to back up, restore, and edit the registry, click the following link to view the following article in the Microsoft Knowledge Base:
<http://support.microsoft.com/kb/256986/>.

The registry change settings in the following steps are recommended for all Microsoft iSCSI configurations.

On a server that runs Windows Server 2003 SP1 or later, follow these steps:

1. Start **Registry Editor** (Regedit.exe).
2. Locate and then click the following registry subkey:
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\Tcpip\Parameters\Interfaces
The interfaces will be listed underneath by automatically generated GUIDs like {064A622F-850B-4C97-96B3-0F0E99162E56}
3. Click each of the interface GUIDs and perform the following steps:
 - a. Check the IPAddress or DhcpIPAddress parameters to determine whether the interface is used for iSCSI traffic. If not, skip to the next interface.
 - b. On the **Edit** menu, point to **New**, and then click **DWORD** value.
 - c. Name the new value **TcpAckFrequency**, and assign it a value of **1**.
4. Exit the **Registry Editor**.
5. Restart Windows for this change to take effect.

Notes on Microsoft iSCSI Initiator

This section contains important information about Microsoft iSCSI Initiator.

Microsoft Cluster Server

Microsoft Cluster Server (MSCS) shared storage (including the quorum disk) can be implemented using iSCSI disk volumes as the shared storage. There is no special iSCSI cluster or application configuration needed to support this scenario. Since the cluster service manages application dependencies, it is not needed to make any cluster managed service (or the cluster service itself) dependent upon the Microsoft iSCSI service.

Microsoft MPIO and the Microsoft iSCSI DSM can be used with MSCS.

Note: Microsoft does not support the use of iSCSI Server clusters with Windows 2000.

File shares/DFS

There is a special requirement if file shares are exposed on iSCSI disk volumes using the Microsoft software iSCSI Initiator. For example, if you have an iSCSI disk volume that is exposed as drive I: and have a file share point I:\Documents then you will need to configure the lanmanager server service to have a dependency on the msiscsi (Microsoft iSCSI Initiator) service on Windows 2000. When using the Microsoft iSCSI Software Initiator 2.0 with Windows 2003, the dependency is not needed.

Dynamic disks

Note: Applies to Windows 2000 and Windows Server 2003.

Configuring volumes on iSCSI disks as Dynamic disk volumes using the Microsoft software iSCSI Initiator is not currently supported. It has been observed that timing issues may prevent dynamic disk volumes on iSCSI disks from being reactivated at system startup.

Boot

Currently, it is not possible to boot a Windows system using an iSCSI disk volume provided by the Microsoft software iSCSI Initiator kernel mode driver. It is possible to boot a Windows system using an iSCSI disk volume provided by an iSCSI HBA. The only currently supported method for booting a Windows system using an iSCSI disk volume is via a supported HBA.

NIC teaming

Microsoft does not support the use of NIC teaming on iSCSI interfaces.

Using the Initiator with EMC PowerPath v4.6.x

This section provides an overview of how to set up the Initiator for use with PowerPath 4.6 and describes how to use the Initiator to create the paths that PowerPath then takes control of. This setup can be compared to creating zones in a Fibre Channel environment.

The Initiator allows you to log in multiple paths to the same target and aggregate the duplicate devices into a single device exposed to Windows. Each path to the target can be established using different NICs, network infrastructure, and target ports. If one path fails then another session can continue processing I/O without interruption to the application. It is PowerPath that aggregates these paths and manages them to provide uninterrupted access to the device.

PowerPath 4.6.x uses the Microsoft MPIO framework that is installed with the Initiator in conjunction with EMC's DSM to provide multipathing functionality.

This section also describes how to use the Initiator to create the paths that PowerPath then takes control of. This setup can be compared to creating zones in a fibre channel environment.

This section is based on the following assumption:

- ◆ The following two network cards are installed in the server:
 - NIC1 (192.168.150.155)
 - NIC2 (192.168.150.156)

- ◆ We will log into the target on four different ports:

- A0 (192.168.150.102)
- A1 (192.168.150.103)
- B0 (192.168.150.104)
- B1 (192.168.150.105).

Logging into a target is described in [“Examples” on page 99](#) of this document.

- ◆ We will log into the target to create four separate paths:

- NIC1 → A0
- NIC2 → A1
- NIC1 → B0
- NIC2 → B1

Follow the next steps to log in to the target.

1. Select the port to log into and click **Log On**.

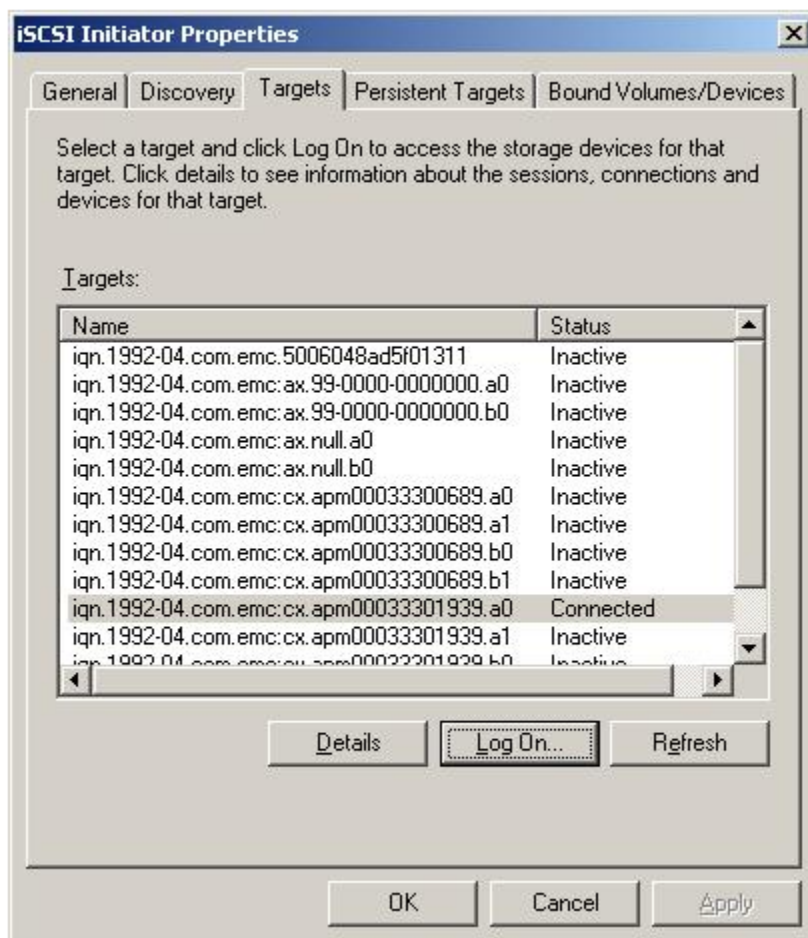


Figure 34 iSCSI Initiator Properties dialog box

The **Log On to Target** dialog box displays.

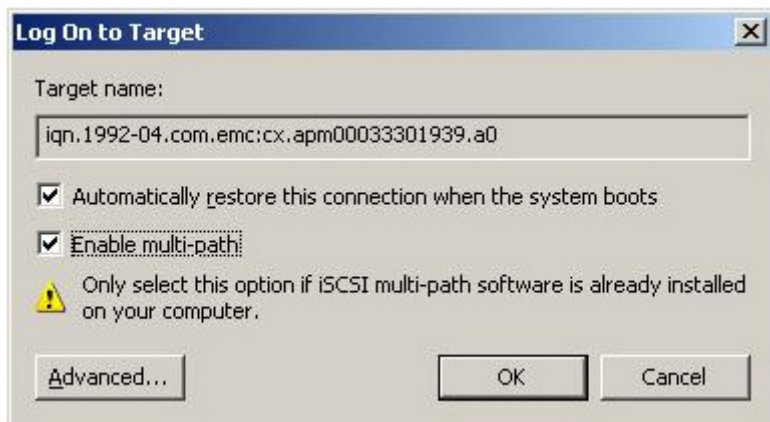


Figure 35 Log On to Target dialog box

2. Click **Advanced**.

The **Advanced Settings** dialog box displays.

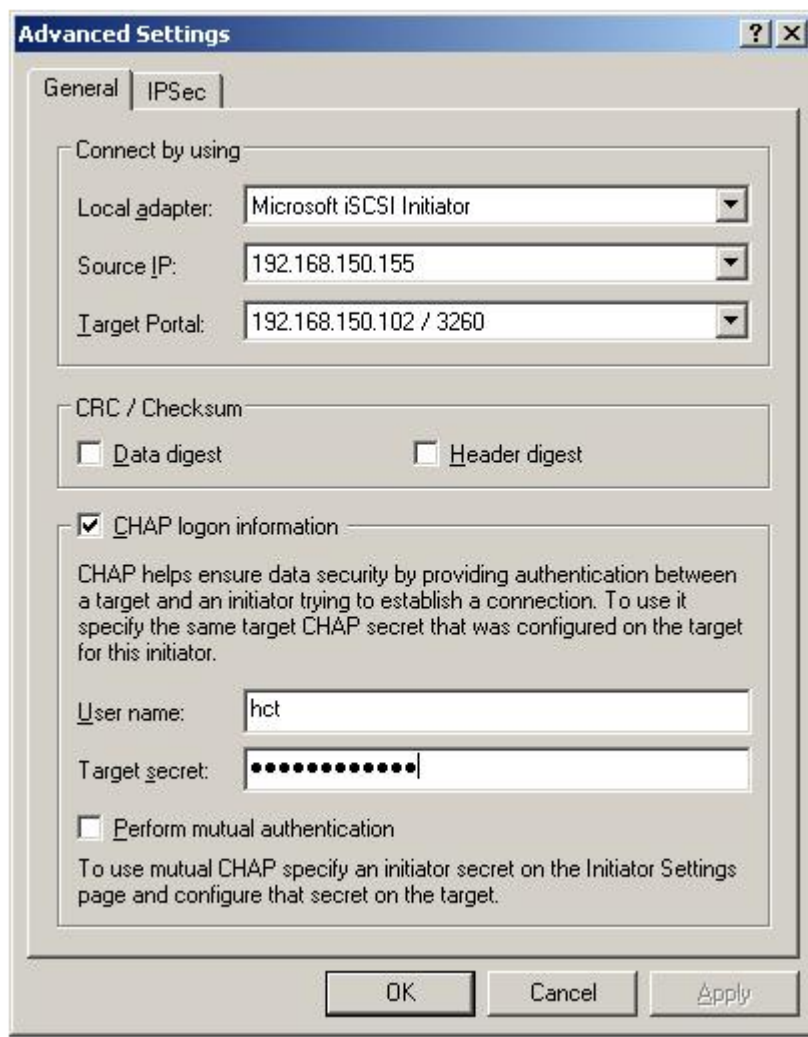


Figure 36 **Advanced Settings dialog box**

3. For the source IP choose the IP address associated with NIC1 and the Target Portal (port) associated with port A0 on the target. Enter CHAP information, if needed, and click **OK**.

To create the other paths to the array repeat the steps above for each path you want to create.

In our example, you would choose a different port on the **Targets** tab (A1, B0....).

Note: You can only log into a port once from a host. Multiple logins are not supported.

On the **Advanced Settings** page choose your source and Target IPs. Once completed, you should have four paths as shown in [Figure 37](#).

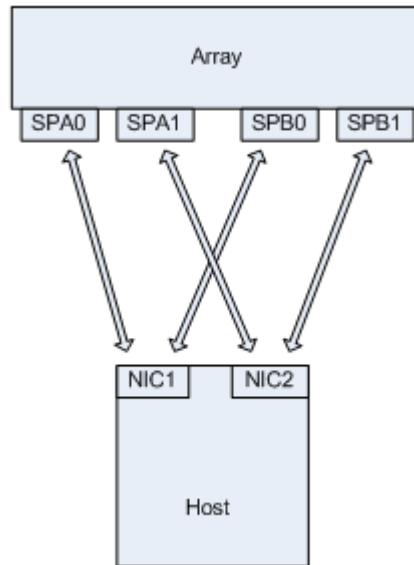


Figure 37 **Four paths**

You can now install PowerPath. PowerPath will aggregate all four paths and present one device to Windows Disk Management. From a host perspective, a failure to a NIC or the path between and NIC and the target will be managed properly by PowerPath and allow access to the device on the target. For example, if NIC1 fails the host would still have an active paths to both SPs since NIC2 has a path to SPA1 and SPB1.

Commonly seen issues

This section lists some of the common issues E-Lab has seen during testing. For a more detailed list, refer to the *Microsoft Initiator 2.0 Users Guide*.

Multipathing

This section lists multipathing errors that may be recorded in the event logs and discusses potential solutions for these error conditions. Even though EMC does not support multiple connections per session (MCS), it is easy to confuse the configuration of MPIO and MCS. The following errors might help point out that you are actually configuring MCS instead of MPIO.

- ◆ Error: "Too many Connections" when you attempt to add a second connection to an existing session.

This issue can occur if the target does not support multiple connections per session (MCS). Consult with the target vendor to see if they plan on adding support for MCS.

- ◆ When you attempt to add a second connection to an existing session, you may notice that the **Add** button within the **Session Connections** window is grayed out.

This issue can occur if you logged onto the target using an iSCSI HBA that doesn't support MCS.

Long boot time

If your system takes a long time to display the login prompt after booting, or it takes a long time to log in after entering your login ID and password, then there may be an issue related to the Microsoft iSCSI Initiator service starting. First see the "Running automatic start services on iSCSI disks" section in the *Microsoft Initiator 2.x Users Guide* for information about persistent volumes and the binding operation. Check the system event log to see if there is an event "Timeout waiting for iSCSI persistently bound volumes...". If this is the case, then one or more of the persistently bound volumes did not reappear after reboot which could be due to network or target error.

Another error you may see on machines that are slow to boot is an event log message "Initiator Service failed to respond in time to a request to encrypt or decrypt data" if you have persistent logins that are configured to use CHAP. Additionally, the persistent login will fail to log in. This is due to a timing issue in the service startup order. To work around this issue, increase the timeout value for the IPsecConfigTimeout value in the registry under:

```
HKLM\SYSTEM\CurrentControlSet\Control\Class\{4D36E97B-E325-11CE-BFC1-08002BE10318}\<Instance Number> \Parameters
```

This has been seen in some cases where clusters are present.

Logging out of target

The MS iSCSI Initiator service will not allow a session to be logged out if there are any handles that are open to the device. In this case, if the session is attempted to be logged out, the error "The session cannot be logged out since a device on that session is currently being used." is reported. This means that there is an application or device which has an open handle to the physical disk on the target. If you look in the system event log you should see an event that has the name of the device with the open handle.

Other event log errors of note

The source for system events logged by Software Initiator will be iScsiPrt. The message in the log would convey the cause of that event.

Some of the common events are listed below. A complete list of events can be found in the *Microsoft Initiator User's Guide 2.x*.

Event ID 1 Initiator failed to connect to the target. Target IP address and TCP Port number are given in dump data.

This event is logged when the Initiator could not make a TCP connection to the given target portal. The dump data in this event will contain the target IP address and TCP port to which Initiator could not make a TCP connection.

Event ID 9 Target did not respond in time for a SCSI request. The CDB is given in the dump data.

This event is logged when the target did not complete a SCSI command within the timeout period specified by SCSI layer. The dump data will contain the SCSI Opcode corresponding to the SCSI command. User can refer to SCSI specification for getting more information about the SCSI command.

Event ID 20 Connection to the target was lost. The Initiator will attempt to retry the connection.

This event is logged when the Initiator loses connection to the target when the connection was in iSCSI Full Feature Phase. This event typically happens when there are network problems, a network cable is removed, a network switch is shutdown, or target resets the connection. In all cases Initiator will attempt to reestablish the TCP connection.

Event ID 34 **A connection to the target was lost, but Initiator successfully reconnected to the target. Dump data contains the target name.**

This event is logged when the Initiator successfully reestablishes a TCP connection to the target.

Troubleshooting

Note the following potential problems and their solutions.

Problem Adding a Symmetrix IP address to the target portals returns `initiator error`.

Solution

- ◆ Verify that the iSCSI login parameter is correct.
- ◆ Verify that the Volume Logix setup is correct.

Problem Adding a Symmetrix IP address to the target portals returns `authentication error`.

Solution

- ◆ Verify that the CHAP feature is enabled, and that the user name and secret are correct.
- ◆ Verify that the Symmetrix system has the correct user name and secret.

Problem Login to the Symmetrix target returns `The target had already been logged via a Symmetrix`.

Solution

- ◆ Press the **Refresh** button to verify that only one iSCSI session is established.
- ◆ Log out of the current iSCSI session to the Symmetrix system, and log in again.

Problem Adding a target portal returns `Connection Failed`.

Solution

- ◆ Make sure the IP address of the Symmetrix system is correct.
- ◆ Make sure the connectivity is correct by using the `ping` utility to the Symmetrix GE port and vice versa.

Problem File shares on iSCSI devices may not be re-created when you restart your computer

Solution

This issue can occur when the iSCSI Initiator service is not initialized when the Server service initializes. The Server service creates file shares. However, because iSCSI disk devices are not available, the Server service cannot create file shares for iSCSI devices until the iSCSI service is initialized. To resolve this issue, follow these steps on the affected server:

1. Make the Server service dependant on the iSCSI Initiator service.
2. Configure the BindPersistentVolumes option for the iSCSI Initiator service.
3. Configure persistent logons to the target. To do this, use one of the following methods.

Method 1

- a. Double-click iSCSI Initiator in Control Panel.
- b. Click the **Available Targets** tab.
- c. Click a target in the **Select** a target list, and then click **Log On**.
- d. Click to select the Automatically restore this connection when the system boots check box.
- e. Click **OK**.

Method 2

- a. Click **Start**, click **Run**, type **cmd**, and then click **OK**.
- b. At the command prompt, type the following command, and then press ENTER:

```
scsicli persistentlogintarget target_iqn T * * * * * * * * * * * * * * 0
```

Note: target_iqn is the iSCSI qualified name (iqn) of the target.

Note: This resolution applies only when you specifically experience this issue with the iSCSI Initiator service. Refer to Microsoft Knowledge Base article 870964 for more information.

CLARiiON and Symmetrix Information

This chapter provides procedures and information specific to CLARiiON and Symmetrix.

- ◆ [CLARiiON environment.....](#) 140
- ◆ [Symmetrix environment](#) 146

CLARiiON environment

This section describes CLARiiON-specific configuration and support information.

Storage components

The basic components of a storage system configuration for CLARiiON arrays are:

- ◆ One or more storage systems.
- ◆ One or more servers connected to the storage system(s), directly or through switches.
- ◆ For Navisphere 6.X, a host that is running an operating system that supports the Navisphere Manager browser-based client and is connected over a LAN to storage-system servers and CX-Series storage systems. For a current list of such operating systems, refer to the *EMC Navisphere Manager 6.X Release Notes* at <http://Powerlink.EMC.com>.

Note: The procedures described in this document assume that all hardware equipment (for example: switches and storage systems) used in the documented configuration is already installed and connected.

Required storage system setup

CLARiiON product documentation and installation procedures for connecting a CLARiiON storage system to a server are available on <http://Powerlink.EMC.com>.

Note: EMC recommends that you download the latest information before you install any server.

Host connectivity

Refer to the *EMC Support Matrix* or contact your EMC representative for the most up-to-date information on qualified switches, hosts, host bus adapters, and software.

The latest EMC-approved HBA drivers and software are available for download at the following websites:

<http://www.emulex.com>

<http://www.qlogic.com>

The EMC HBA installation and configurations guides are available at the EMC-specific download pages of these websites.

LUNZ visibility to host

In the CLARiiON context, LUNZ refers to a fake logical unit zero presented to the host to provide a path for host software to send configuration commands to the array when no physical logical unit zero is available to the host.

When Access Logix™ is used on a CLARiiON array, an agent runs on the host and communicates with the storage system through either LUNZ or a storage device. On a CLARiiON array, the LUNZ device is replaced when a valid LUN is assigned to the HLU LUN0 by the storage group. The agent then communicates through the storage device.

LUNZ has been implemented on CLARiiON arrays to make arrays visible to the host OS and PowerPath when no LUNs are bound on that array. When using a direct-connect configuration, and there is no Navisphere Management station to talk directly to the array over IP, the LUNZ can be used as a pathway for Navisphere CLI to send Bind commands to the array.

Refer to EMC Knowledgebase solution emc65060 for a discussion on this topic.

Navisphere Windows host agent and server utility

The type of storage system your Windows host is attached to will determine the use of the host agent or server utility.

If you have an AX-series storage system running Navisphere Express, you must install the server utility. The host agent is supported only on CX3-Series, CX-Series, and AX-Series storage systems running Navisphere Manager.

About Navisphere host agent

The host agent registers the server's HBA (host bus adapter) with the attached storage system when the host agent service starts. This action sends the initiator records for each HBA to the storage system. Initiator records are used to control access to storage-system data. For legacy storage systems, the host agent will send the initiator records only if Access Logix software is installed. The agent can then automatically retrieve information from the storage system at startup or when requested by Manager or CLI.

The host agent can also:

- ◆ Send drive mapping information to the attached CLARiiON storage systems.
- ◆ Monitor storage system events and notify personnel by email, page, or modem when any designated event occurs.
- ◆ Retrieve LUN WWN (world wide name) and capacity information from Symmetrix storage systems.

The host agent runs on servers attached to CX3-Series and CX-Series. It also runs on servers attached to AX-series storage systems that have been upgraded to Navisphere Manager, (that is, that have the Navisphere Manager enabler installed).

About Navisphere storage system initialization utility

For CX3-series and CX-Series Fibre Channel storage systems, use the utility to discover storage systems, and set network parameters (IP address, subnet mask, and default gateway). In addition, for CX3-series storage systems with iSCSI data ports attached to Windows servers, use the utility to set network parameters for these ports.

Note: For CX-Series storage systems, an authorized service provider must install and run the Initialization Utility.

Detailed information about the Navisphere Host Agent and Navisphere Storage System Initialization Utility can be found in the *EMC CLARiiON Server Support Products for Windows Server Installation Guide* available at <http://Powerlink.EMC.com>. This guide describes how to install and remove the Navisphere Storage System Initialization Utility, Navisphere Host Agent, Navisphere Server

Utility, Navisphere Command Line Interface (CLI), CLARiiON VSS Provider (Windows Server 2003 only), and admhost or admsnap software on a Microsoft server running the Windows Server 2003 or Windows 2000 operating system.

Navisphere Management Suite

The Navisphere Management Suite of integrated software tools allows customers to manage, provision, monitor, and configure CLARiiON systems, bind LUNs, create RAID groups, storage groups, LUN assignment, and control all platform replication applications from an easy-to-use, secure, web-based management console. Navisphere-managed CLARiiON platform applications include Navisphere Analyzer, SnapView™, MirrorView™, and SAN Copy™.

More information about Navisphere is available at <http://Powerlink.EMC.com>.

LUN expansion recognition

LUN expansion is accomplished with Navisphere Manager or CLI using the CLARiiON MetaLUN feature. A MetaLUN is a type of LUN whose capacity is the combined capacities of all the LUNs that comprise it. Currently, MetaLUNs are only supported on CX3-series and CX-Series storage systems. The MetaLUN feature lets you dynamically expand the capacity of a single LUN (base LUN) into a larger unit called a MetaLUN. You do this by adding additional LUNs to the base LUN. You can also add additional LUNs to a MetaLUN to further increase its capacity. During the expansion operation, you can still access the existing data, but you cannot access the additional capacity until the expansion is complete. Like a LUN, a MetaLUN can belong to a Storage Group and can participate in SnapView, MirrorView, and SANCopy sessions.

After expanding one or more LUNs through Navisphere Manager, steps must be taken at the host operating system level and at the volume management level before Windows will recognize the increased LUN capacity. These steps have varying levels of impact on normal user operations.

In Windows, partitions are the building blocks for file system volumes. A LUN is viewed at the operating system level as a physical disk, and can contain partitions based on the classification of disk as assigned by the OS.

In Windows, there are two types of disk groups:

- ◆ Basic
Basic disks have a maximum of four partitions.
- ◆ Dynamic.
Dynamic disks refer to partitions as volumes, and the number of volumes is theoretically unlimited.

These two classifications of disks have different properties with respect to LUN expansion (called Data Volume Extension in Windows documentation).

Refer to the following articles for more information:

- ◆ Microsoft Knowledge Base article 325590, How to use Diskpart.exe to extend a data volume in Windows Server 2003, in Windows XP, and in Windows 2000, available at <http://support.microsoft.com/kb/325590/en-us>.
- ◆ Microsoft Knowledge Base article 175761, Dynamic vs. Basic Storage in Windows 2000 for more information, available at <http://support.microsoft.com/kb/175761>.
- ◆ Microsoft Knowledge Base article 323442, How to use the Disk Management Snap-in to manage Basic and Dynamic Disks in Windows Server 2003, available at <http://support.microsoft.com/kb/323442/>.
- ◆ Microsoft Knowledge Base article 222189, Description of Disk Groups in Windows Disk Management, available at <http://support.microsoft.com/kb/222189>.



CAUTION

Windows 2000 systems must be upgraded to Service Pack 4 before using CLARiiON metaLUN expansion with Microsoft Dynamic Disks. Failure to do so can result in loss of dynamic volumes and/or volume data, and cause system STOP errors. Refer to Microsoft Knowledge Base article 327020 (formerly Q327020), *Error Message Occurs When You Start Disk Management After Extending a Hardware Array*, for more information.

Configuring VERITAS Volume Manager for CLARiiON

When using the VERITAS Volume Manager (VxVM) with Windows servers and CLARiiON, keep in mind the following:

- ◆ The VERITAS Volume Manager replaces the native disk manager that is accessed from the system management console. Refer to VERITAS documentation for installation and operating instructions for your version.
- ◆ If the VERITAS DMP (Dynamic Multi-Pathing) feature is to be used with a CLARiiON array, the minimum supported version of FLARE® that supports DMP must be installed on the array. Refer to the *EMC Support Matrix* for minimum FLARE versions. Also, the Dynamic Multi-Pathing driver update for the EMC CLARiiON CX-Series arrays must be installed on the host server. This driver is available for download from <http://www.semantec.com>.
- ◆ HBA and array settings for DMP are the same as those that are configured for use with EMC PowerPath. Refer to EMC documentation for your HBA on which settings are used for PowerPath configurations.

Refer to the *EMC Support Matrix* for supported VxVM versions, service packs, and hotfixes.

Symmetrix environment

This section describes Symmetrix-specific configuration and support information.

Initial Symmetrix configuration

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

Refer to the *EMC Support Matrix* for required bit settings on Symmetrix channel directors.

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.

Symmetrix SPC-2 director bit 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 which are optimized for a SAN environment (as implemented in Fibre Channel), as opposed to legacy (non SPC-2) functionality, 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, then 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 Management CLI Product Guide*, located on [Powerlink](#), for details regarding setting the SPC-2 bit on a per-initiator basis.

SPC-2 must be enabled for all initiators on a per-host basis, globally; which means 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, from that same host.

SPC-2 conformance is supported for new Symmetrix devices only unless otherwise specified in the *EMC Support Matrix*. Migrations from SPC-2 disabled to SPC-2 enabled configurations, either online or offline, are not yet supported unless otherwise specified in the *EMC Support Matrix* and / or relevant Host Connectivity Guide.

SPC-2 director bit support for Windows

The SPC-2 director bit setting is supported in all EMC-supported Windows OS versions.

The functionality provided by the SPC-2 director bit became a requirement for Windows 2003 Service Pack 1 (or higher) in specific software environments that use Volume Shadow Copy Services. This functionality was not a requirement for Windows 2000 nor the original release of Windows 2003 "RTM" (without an installed Service Pack). Please refer to your software documentation regarding these requirements.

Note: While the SPC-2 functionality may affect only a subset of Windows functionality, all new Enginuity releases have been Microsoft certified for the "Windows Server Catalog" (formerly "Hardware Compatibility List") with the SPC-2 director bit functionality enabled.

Arbitrated loop addressing

The Fibre Channel arbitrated loop (FC-AL) topology defines a method of addressing ports, arbitrating for use of the loop, and establishing a connection between Fibre Channel NL_Ports (level FC-2) on HBAs in the host and Fibre Channel directors (via their adapter cards) in the Symmetrix system. Once loop communications are established between the two NL_Ports, device addressing proceeds in accordance with the SCSI-3 Fibre Channel protocol (SCSI-3 FCP, level FC-4).

The Loop Initialization Process (LIP) assigns a physical address (AL_PA) to each NL_Port in the loop. Ports that have a previously acquired AL_PA are allowed to keep it. If the address is not available,

another address may be assigned, or the port may be set to non-participating mode.

Note: The AL-PA is the low-order 8 bits of the 24-bit address. (The upper 16 bits are used for Fibre Channel fabric addressing only; in FC-AL addresses, these bits are **x'0000'**.)

Symmetrix Fibre Channel director parameter settings, shown in [Table 1 on page 149](#), control how the Symmetrix system responds to the LIP.

After the loop initialization is complete, the Symmetrix port can participate in a logical connection using the hard-assigned or soft-assigned address as its unique AL_PA. If the Symmetrix port is in non-participating mode, it is effectively off line and cannot make a logical connection with any other port.

A host initiating I/O with the Symmetrix system uses the AL_PA to request an open loop between itself and the Symmetrix port. Once the arbitration process has established a logical connection between the Symmetrix system and the host, addressing specific logical devices is done through the SCSI-3 FCP.

FC-AL addressing parameters

Table 1 FC-AL addressing parameters

Parameter	Bit	Description	Default
Disk Array	A	If enabled, the Fibre Channel Director presents the port as a disk array. Refer to “Symmetrix SCSI-3 addressing modes” on page 151 for settings for each addressing mode.	Enabled
Volume Set	V	If enabled <i>and</i> Disk Array is enabled, Volume Set addressing mode is enabled. Refer to “Symmetrix SCSI-3 addressing modes” on page 151 for settings for each addressing mode.	Enabled
Use Hard Addressing	H	If enabled, entering an address (00 through 7D) in the Loop ID field causes the port to attempt to get the AL_PA designated by the Loop ID. If the port does not acquire the AL_PA, the Symmetrix reacts based on the state of the non-participating (NP) bit. If the NP bit is set, the port switches to non-participating mode and is not assigned an address. If non-participating mode is not selected, or if the H-bit was not set, the Symmetrix port accepts the new address that is <i>soft-assigned</i> by the host port.	Enabled
Hard Addressing Non-participating	NP	If enabled <i>and</i> the H bit is set, the director uses <i>only</i> the hard address. If it cannot get this address, it re-initializes and changes its state to non-participating. If the NP bit is not set, the director accepts the soft-assigned address.	Disabled
Loop ID	---	One-byte address (00 through 7D), valid only if the H bit is set.	00
Third-party Logout Across the Port	TP	Allows broadcast of the TPRLO extended link service through all of the FC-AL ports.	Disabled

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 N_Ports can automatically adjust 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 2 on page 151](#), differ in the addressing scheme (target ID, LUN and virtual bus) and number of addressable devices.

Symmetrix SCSI-3 addressing modes

Table 2 Symmetrix SCSI-3 addressing modes

Addressing Mode	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.

PowerPath for Windows

This section provides information regarding PowerPath for Windows.

- ◆ PowerPath and PowerPath iSCSI 154
- ◆ PowerPath for Windows 155
- ◆ EMC PowerPath verification and problem determination 158
- ◆ PowerPath iSCSI 1.1 for Windows 165

PowerPath and PowerPath iSCSI

EMC PowerPath for Windows software is available in two different packages: PowerPath for Windows and PowerPath iSCSI for Windows. It is important to know the differences between the two packages before deploying the software:

PowerPath for Windows: Supports both Fibre Channel and iSCSI environments. PowerPath for Windows is Microsoft digitally certified only for Fibre Channel environments. PowerPath for Windows supports failover path management and load-balancing for up to 32 paths in heterogeneous storage environments. PowerPath for Windows is not currently supported by Microsoft for iSCSI implementations, although it is supported by EMC for EMC iSCSI storage systems.

PowerPath iSCSI for Windows: Supports EMC CLARiiON iSCSI storage systems. PowerPath iSCSI for Windows is Microsoft digitally certified and is built on the Microsoft MPIO framework. PowerPath iSCSI for Windows supports failover path management for up to 8 paths in iSCSI storage environments.

PowerPath for Windows

PowerPath and MSCS

If you are installing PowerPath and MSCS for the first time, install PowerPath first, and then install MSCS. Installing PowerPath first avoids having to disrupt cluster services at a later time.

Integrating PowerPath into an existing MSCS cluster

You can integrate PowerPath into an existing MSCS cluster without shutting down the cluster, if there is close coordination between the nodes and the storage system. Each node in a cluster can own a distinct set of resources. Node A is the primary node for its resources and the failover node for Node B's resources. Conversely, Node B is the primary node for its resources and the failover node for Node A's resources.

If after installing PowerPath on the cluster, you test node failover by disconnecting all cables for a LUN or otherwise disrupting the path between the active host and the array, Windows logs event messages indicating hardware or network failure and possible data loss. If working correctly, the cluster will failover to a node with an active path and you can ignore the messages from the original node as logged in the event log. (You should check the application generating I/O to see if there are any failures. If there are none, everything is working normally.)

Installing PowerPath in a clustered environment requires the following steps:

- ◆ Move all resources to Node A
- ◆ Install PowerPath on Node B
- ◆ Configure additional paths between storage array and Node B
- ◆ Move all resources to Node B
- ◆ Install PowerPath on Node A
- ◆ Configure additional paths between storage array and Node A
- ◆ Return Node A's resources back to Node A

Moving resources to Node A

To move all resources to Node A:

1. Start the MSCS Cluster Administrator utility, select **Start, Programs, Administrative Tools, Cluster Administrator**.
2. In the left pane of the window, select all groups owned by Node B.
3. To move the resources to Node A, select **File, Move Group**. Alternatively, select **Move Group** by right-clicking all group names under **Groups** in the left pane.
4. To pause Node b, click Node B and select **File, Pause Node**. This keeps the node from participating in the cluster during PowerPath installation.

Installing PowerPath onto Node B

To install PowerPath onto Node B:

1. Install PowerPath.
2. Shut down Node B. In a cluster with greater than two nodes, install PowerPath on these other nodes. For example, in a four-node cluster, replace Node B with Nodes B, C, and D in step 4 of the previous section, [“Moving resources to Node A”](#), and also in steps 1 and 2, above.

Configuring additional paths between the storage system and Node B

To configure additional paths:

1. If necessary, reconfigure the storage system so its logical devices appear on multiple ports.
2. If necessary, install additional HBAs on Node b.
3. Connect cables for new paths between Node b and the storage system.
4. Power on Node b.
5. To resume Node b, click **Node b** and select **File, Resume Node**. In a cluster with greater than two nodes, configure additional paths between the storage system and these other nodes. For example, in a four-node cluster, replace Node B with Nodes B, C, and D in steps 2, 3, 4, and 5 above.

Moving resources to Node B

To move all resources to Node B:

1. In the left pane of the Cluster Administrator window, select all groups.

2. To move the resources to Node B, select **File, Move Group**. In a cluster with greater than two nodes, move all resources to any of the remaining nodes. For example, in a four-node cluster, replace Node b with any combination of Nodes B, C, and D to which you want to move resources. For example, you could move resources to Nodes B and C or move them to B, C, and D, or any permutation of Nodes B, C, and D taken alone or together.
3. To pause Node A, click **Node A** and select **File, Pause Node**.

Installing PowerPath onto Node A

To install PowerPath onto Node A:

1. Install PowerPath.
2. Shut down Node A.

Configuring additional paths between the storage system and Node A

To configure additional paths:

1. If necessary, configure the storage system so its logical devices appear on multiple ports.
2. If necessary, install additional HBAs on Node A.
3. Connect cables for new paths between Node A and the storage system.
4. Power on Node A.
5. To resume Node A, click Node A and select **File, Resume Node**.

Returning Node A's resources to Node A

To return Node A's resources:

1. Using the MSCS Cluster Administrator utility, select all groups previously owned by Node A.
2. To move the resources back to Node A, select **File, Move Group**.

EMC PowerPath verification and problem determination

The following section assumes that PowerPath has been installed properly. Please refer to the appropriate *PowerPath Installation and Administration Guide* on Powerlink for instructions on how to install PowerPath. This section will help to verify that PowerPath was installed correctly and help you to recognize some common failure points.

Click the circled icon shown in [Figure 38](#) to access the PowerPath Administration.



Figure 38 PowerPath Administration icon

[Figure 39](#) shows the administration icon and the status it represents.

PowerPath Monitor Taskbar Icons

Taskbar Icon	Status
	Optimal — All configured paths are operating properly.
	Degraded — Some but not all paths have failed.
	All Failed — All configured paths have failed.
	Server Not Responding — Cannot communicate with Monitor service. This occurs if the PowerPath Administrator Suite encountered errors during initialization.

Figure 39 PowerPath Monitor Taskbar icons and status

Figure 40 shows what PowerPath Administrator would look like if installed correctly. Notice that in this case there is one path zoned between the HBA and one port on the storage device.

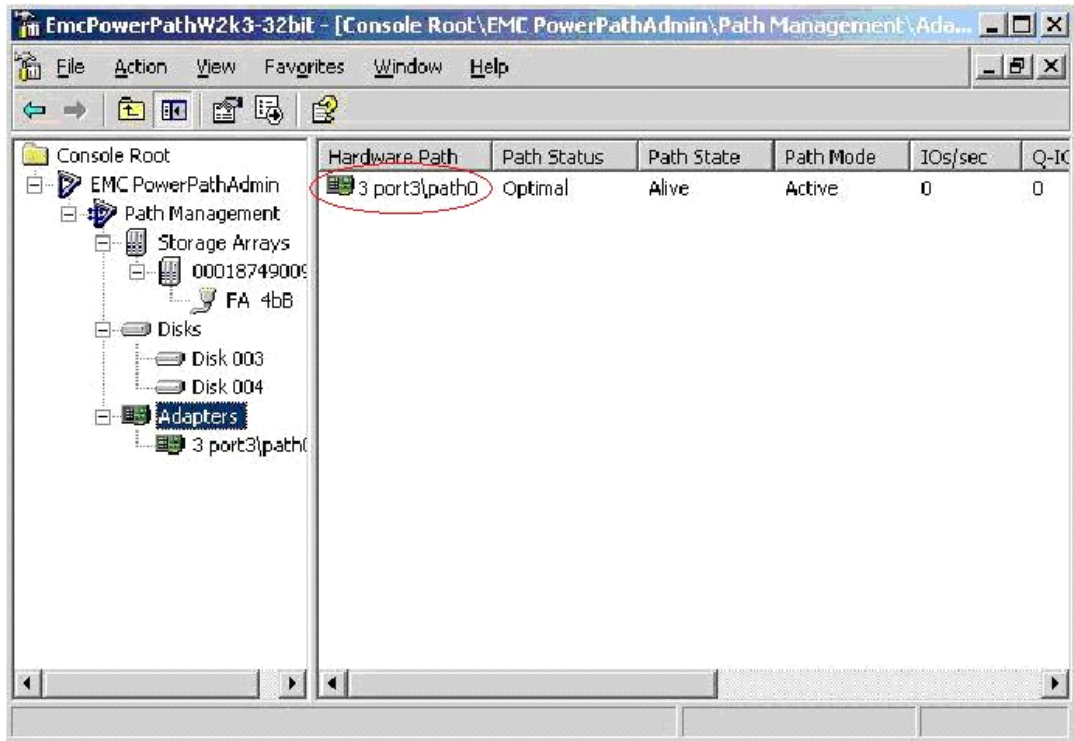


Figure 40 One path

When multiple paths are zoned to your storage device, PowerPath Administrator would look like [Figure 41](#):

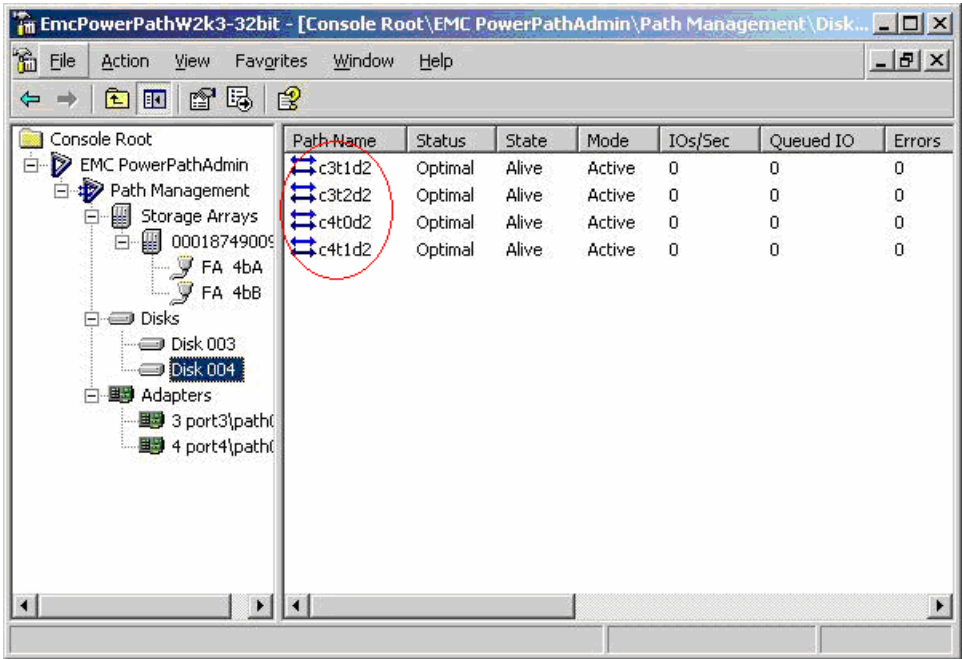


Figure 41 Multiple paths

Problem determination

Determining the cause of loss if connectivity to the storage devices can be simplified by using the PowerPath Administrator. Array ports that are offline, defective HBAs or broken paths show up in the administrator GUI in various ways.

[Table 3 on page 161](#) shows the known possible failure states. Referencing this table can greatly reduce problem determination time.

Table 3 Possible failure states

Object	Degraded Icon	Failed Icon	Unlicensed and Degraded Icon	Unlicensed and Failed Icon
PowerPath device	 One or more (but not all) paths to the disk device have failed.	 All paths to the specific disk device have failed. This disk is not available.	 EMC PowerPath Administrator is unlicensed. One or more (but not all) paths to the disk device have failed.	 EMC PowerPath Administrator is unlicensed. All paths to the specific disk device have failed. This disk is not available.
Adapter	 One or more (but not all) adapters have either failed or degraded. Displays in the Adapters folder in the scope pane. One or more (but not all) paths have failed on the specific adapter. Displays in the individual adapter folders under Adapters in the scope pane or in the result pane when Adapters is selected from the scope pane.	 All paths on this adapter to this disk have failed.	 EMC PowerPath Administrator is unlicensed. One or more (but not all) adapters have either failed or degraded. Displays in the Adapters folder in the scope pane. One or more (but not all) paths have failed on the specific adapter. Displays in the individual adapter folders under Adapters in the scope pane or in the result pane when Adapters is selected from the scope pane.	 EMC PowerPath Administrator is unlicensed. All paths on this adapter to this disk have failed.
Path	Not Applicable.	 One or more of the hardware components that make up the path has failed; therefore, the entire path failed.	Not Applicable.	 EMC PowerPath Administrator is unlicensed and one or more of the hardware components that comprise the path has failed; therefore, the entire path fails.
Root Node	 EMC PowerPath Administrator is identifying a degradation within the managed environment.	 EMC PowerPath Administrator is in a failed state.	 EMC PowerPath Administrator is unlicensed and identifying a degradation within the managed environment.	 EMC PowerPath Administrator is unlicensed and in a failed state.
Storage Array	 One or more (but not all) paths to the storage array have failed or are in a degraded state.	 All paths to the specific array have failed. This array is not available.	 EMC PowerPath Administrator is unlicensed. One or more (but not all) paths to the array have failed or are in a degraded state.	 EMC PowerPath Administrator is unlicensed. All paths to the specific array have failed. This disk is not available.
Storage Array Port	 One or more (but not all) paths to the storage array port have failed or are in a degraded state.	 All paths to the specific array port have failed. This array port is not available.	 EMC PowerPath Administrator is unlicensed. One or more (but not all) paths to the array port have failed or are in a degraded state.	 EMC PowerPath Administrator is unlicensed. All paths to the specific array port have failed. This disk is not available.

Examples of some failures follow:

An error with an Array port, or the path leading to the Array port, is displayed in [Figure 42](#). This is symbolized by the red X through one of the Array ports. Notice that while the Array port is down access to the disk device is still available - degraded access is noted by a red slash.

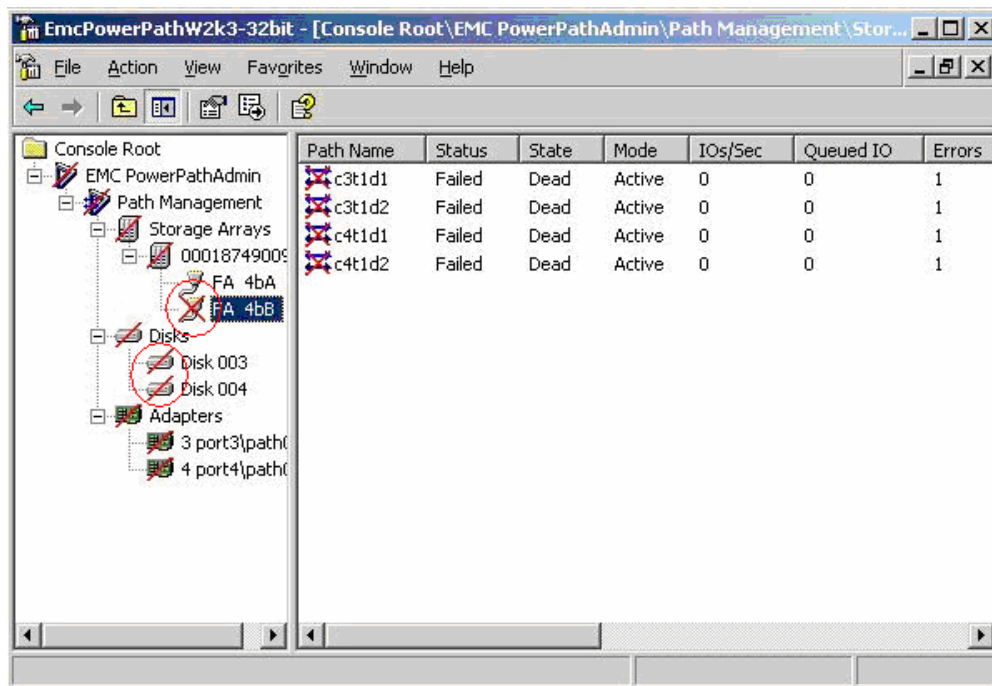


Figure 42 Error with an Array port

Figure 43 shows the result of a problem with one of the HBA or the path leading to the HBA. The failed HBA / path is marked with a red X. Again, notice that access to the disk devices, while degraded, still exists.

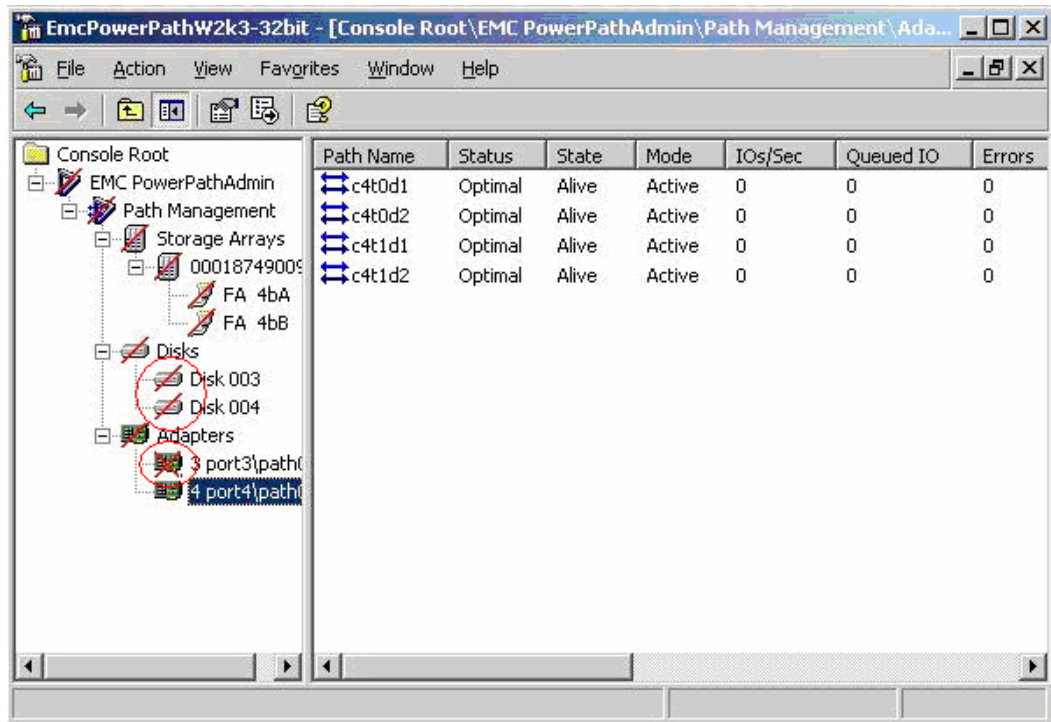


Figure 43 Failed HBA path

Making changes to your environment

You must reconfigure PowerPath after making configuration changes that affect host-to-storage-system connectivity or logical device identification - for example:

- ◆ Fibre Channel switch zone changes
- ◆ Adding or removing Fibre Channel switches
- ◆ Adding or removing HBAs or storage-system ports
- ◆ Adding or removing logical devices

In most cases making changes to your environment will be detected automatically by PowerPath. Depending on the type HBA you are using you may have to scan for new devices in device manager. On some occasions and depending on the operating system version you are running you may need to reboot your system.

PowerPath messages

For a complete list of PowerPath messages and their meanings please refer to the "PowerPath Product Guide - PowerPath Messages" chapter for the version of PowerPath you are running.

PowerPath iSCSI 1.1 for Windows

PowerPath iSCSI 1.1 is EMC's multipathing software designed to work within the Microsoft Multipathing I/O (MPIO) framework. PowerPath 1.1 supports an MC-based snap-in through which you can manage the PowerPath environment.

PowerPath iSCSI 1.1 features

Automatic path failover - PowerPath iSCSI 1.1 supports multiple paths between a CLARiiON LUN and a host. Having multiple paths enables the host to access a LUN even if a specific path is unavailable. PowerPath iSCSI 1.1 automatically detects path failure and redirects I/O from the failed path to an alternate path. This eliminates loss of data and application downtime. Failovers are transparent and non-disruptive to applications.

Autorestore - If LUNs have been relocated, or trespassed, to the passive (unassigned) storage processor because all paths to the active (assigned) storage processor had failed or been taken offline, PowerPath iSCSI 1.1 automatically restores those LUNs to the assigned storage processor when a failed or offline path to that storage processor returns to service.

High-availability cluster support - PowerPath iSCSI 1.1 is particularly beneficial in cluster environments, as it can prevent operational interruptions and costly downtime. PowerPath iSCSI 1.1's path failover capability avoids node failover, maintaining uninterrupted application support on the active node in the event of a path disconnect (as long as another path is available).

PowerPath iSCSI 1.1 restrictions

PowerPath iSCSI 1.1 is supported on the CLARiiON CX and AX storage arrays.

PowerPath iSCSI 1.1 runs with a path restriction of a maximum of 8 to a LUN. If a host has more than 8 paths to the LUN, desired results may not be obtained and is an undesirable and unsupported configuration. Microsoft MPIO Framework version 1.10 is the only supported framework version at this time.

PowerPath iSCSI 1.1 does not support load balancing features.

The management snap-in shown in [Figure 44](#) resides under the Storage component of the Computer Management utility.

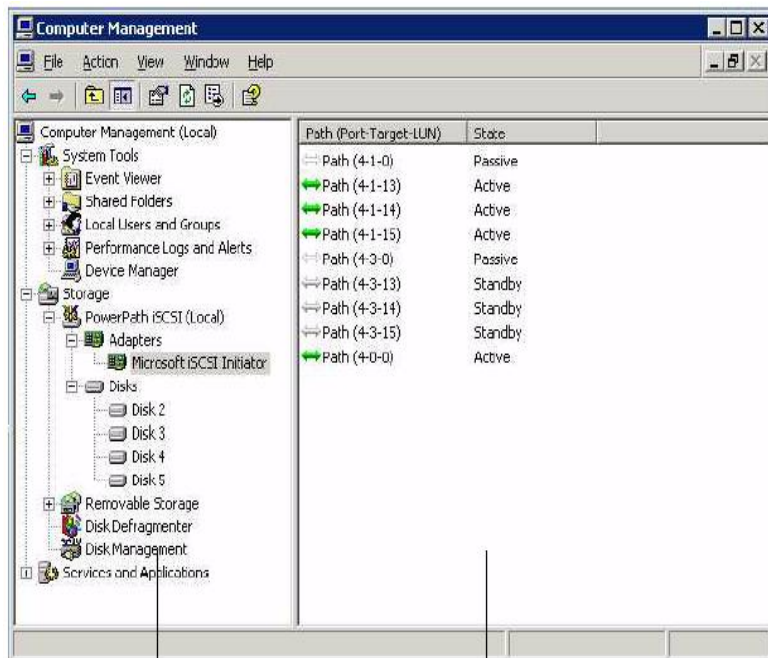


Figure 44 Management Snap-In

Table 4 describes the icons found in PowerPath.

Table 4 PowerPath icons

PowerPath iSCSI 1.1 Icons

Object	Icon
PowerPath iSCSI 1.1 Icon: Managed Paths — Visible within the root level of the MMC snap-in, also referred to as the root node.	
PowerPath iSCSI 1.1 Icon: No Managed Paths — The root node is displayed in light gray to indicate that PowerPath iSCSI 1.1 is not able to locate any MPIO paths to manage.	
Adapter Icon — Identifies an adapter that contains paths that are managed by PowerPath iSCSI 1.1.	
Disk Icon — Identifies a disk managed by PowerPath iSCSI 1.1.	
Active Path Icon — Identifies an active I/O path managed by PowerPath iSCSI 1.1.	
Standby Path Icon — Identifies an I/O-capable path that is inactive.	
Passive Path Icon — Identifies a path connected to the non-active storage processor.	
Unlicensed Path Icon — Identifies a path (other than the single active/passive pair) connected to a host that does not have a valid PowerPath iSCSI 1.1 license.	

Persistent Binding

This appendix provides information on persistent binding.

- ◆ [Understanding persistent binding](#) 170

Understanding persistent binding

Persistent binding is the mechanism to create a continuous logical route from a storage device object in the Windows host to a volume in the EMC storage array across the fabric.

Without a persistent binding mechanism, the host cannot maintain persistent logical routing of the communication from a storage device object across the fabric to an EMC storage array volume. If the physical configuration of the switch is changed (for example, the cable is swapped or the host is rebooted), the logical route becomes inconsistent, causing possible data corruption if the user application is modifying data through inconsistent logical routing of the communication from the driver entry point to a volume in an EMC storage array across the fabric.

The Windows 2000/Windows 2003 operating system (OS) does not provide a satisfactory means to allow persistent binding. Most software applications access storage using file systems managed by the Windows OS. (File systems are represented as <drive letter><colon>, that is, **C:**, **D:**, and so forth) For storage devices containing file systems, Windows writes a disk signature to the disk device. The operating system can then identify, and associate with, a particular drive letter and file system.

Since the signature resides on the disk device, changes can occur on the storage end (a cable swap, for example) that can cause a disk device to be visible to the host server in a new location. However, the OS looks for the disk signature and, providing that nothing on the disk changed, associate the signature with the correct drive letter and file system. This mechanism is strictly an operating system feature and is not influenced by the Fibre Channel device driver.

Some software applications, however, do not use the Windows file systems or drive letters for their storage requirements. Instead they access storage drives directly, using their own built-in “file systems.” Devices accessed in this way are referred to as *raw devices* and are known as *physical drives* in Windows terminology.

The naming convention for physical drives is simple and is always the same for software applications using them. A raw device under Windows is accessed by the name `\\.\PHYSICALDRIVEXXX`, where **XXX** is the drive number.

For example, a system with three hard disks attached using an Emulex Fibre Channel controller assigns the disks the names `\\.\PHYSICALDRIVE0`, `\\.\PHYSICALDRIVE1`, and `\\.\PHYSICALDRIVE2`. The number is assigned during the disk discovery part of the Windows boot process.

During boot-up, the Windows OS loads the driver for the storage HBAs. Once loaded, the OS performs a SCSI **Inquiry** command to obtain information about all the attached storage devices. Each disk drive that it discovers is assigned a number in a *semi-biased* first come, first serve fashion based on HBA. *Semi-biased* means the Windows system always begins with the controller in the lowest-numbered PCI slot where a storage controller resides. Once the driver for the storage controller is loaded, the OS selects the adapter in the lowest-numbered PCI slot to begin the drive discovery process.

It is this naming convention and the process by which drives are discovered that makes persistent binding (by definition) impossible for Windows. Persistent binding requires a continuous logical route from a storage device object in the Windows host to a volume in an EMC storage array across the fabric. As mentioned above, each disk drive is assigned a number in a first-come, first-serve basis. This is where faults can occur.

Example Imagine this scenario: A host system contains controllers in slots 0, 1, and 2. Someone removes a cable from the Emulex controller in host PCI slot 0, then reboots the host.

During reboot, the Windows OS loads the Emulex driver during reboot and begins disk discovery. Under the scenario presented above, there are no devices discovered on controller 0, so the OS moves to the controller in slot 1 and begins naming the disks it finds, starting with `\\.\PHYSICALDRIVE0`. Any software applications accessing `\\.\PHYSICALDRIVE0` before the reboot will be unable to locate their data on the device, because it changed.

[Figure 45 on page 172](#) shows the original configuration before the reboot. HBA0 is in PCI slot 0 of the Windows host. Each HBA has four disk devices connected to it, so Windows has assigned the name `\\.\PHYSICALDRIVE0` to the first disk on HBA0. Each disk after that is assigned a number in sequence as shown in [Figure 45](#).

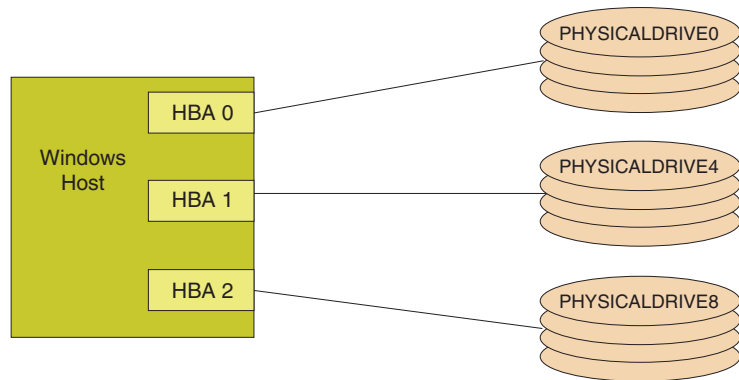


Figure 45 Original configuration before the reboot

Figure 46 shows the same host after the cable attached to HBA0 has been removed and the host rebooted. Since Windows was unable to do a discovery on HBA0, it assigned **\\.\PHYSICALDRIVE0** to the first device it discovered. In this case, that first device is connected to HBA1. Due to the shift, any software application accessing **\\.\PHYSICALDRIVE0** will not find data previously written on the original **\\.\PHYSICALDRIVE0**.

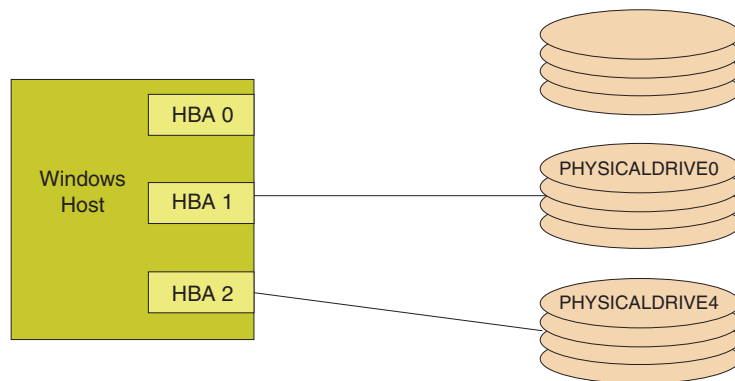


Figure 46 Host after the rebooted

Note: Tape devices are treated the same as disk devices in Windows with respect to persistent binding. Refer to your tape device documentation for more information.

Methods of persistent binding

The Windows device naming convention and disk discovery process does not allow the Windows operating system to establish persistent binding. Therefore, one of these methods must be used:

- ◆ EMC Volume Logix — Provides persistent binding through centralized control by the Symmetrix Fibre Channel fabric ports.
For more information, refer to the *EMC Volume Logix Product Guide*.
- ◆ Switch zone mapping — Provides persistent binding through centralized control by the Fibre Channel switch.
For more information, refer to the *EMC Connectrix Enterprise Storage Network System Planning Guide*.
- ◆ (Emulex HBAs only) Emulex configuration tool mapping — Provides persistent binding of targets through centralized control by the Emulex host adapter. This requires the user to modify the mapping manually.

For more information, refer to *EMC Fibre Channel with Emulex Host Bus Adapters in the Windows Environment*, available on <http://www.emulex.com>:

- a. Click **drivers, software and manuals** at the left side of the screen.
- b. Click **EMC** at the upper center of the next screen.
- c. Click the link to your HBA at the left side of the screen.
- d. Under **EMC Drivers, Software and Manuals**, click the **Installation and Configuration** link under **Drivers for Windows <version>**

This appendix describes the Solutions Enabler.

- ◆ [EMC Solutions Enabler](#) 176

EMC Solutions Enabler

The EMC Solutions Enabler SYMCLI is a specialized library consisting of commands that can be invoked via the command line or within scripts. These commands can be used to:

- ◆ Monitor device configuration and status
- ◆ Perform control operations on devices and data objects within your managed storage complex.

The target storage environments are typically Symmetrix-based, through some features are supported for CLARiiON arrays as well.

For more information, please refer to the *EMC Solutions Enabler Symmetrix Array Management CLI Product Guide*, P / N 300-002-939, available on <http://Powerlink.EMC.com>.

General Host Behavior and Limitations

This appendix describes known issues and limitations.

- ◆ Issues..... 178
- ◆ Capabilities and limitations..... 179
- ◆ How a server responds to a failure in the boot LUN path..... 183

Issues

- ◆ **Basic-to-Dynamic Disk conversion is slower in Windows 2003 with a large number of disks** — The user-interface is slower in Windows 2003 because it converts the disks to dynamic one-by-one, as opposed to the Windows 2000 user interface, which runs the conversions in parallel. The change was required by the introduction of the VDS service in Windows 2003. The behavior is also present when using the `Diskpart` utility. For additional information on this issue, please refer to Microsoft Knowledge Base article 828220.
- ◆ **Windows 2000 Hosts running less than SP4 could be susceptible to data loss during LUN expansion operations** — This issue is corrected by Microsoft in SP4, or by applying hotfix 327020. Refer to EMC Solution IS emc73538 and Microsoft Knowledge Base article 327020 for more information.
- ◆ **To upgrade from Windows 2000 to Windows 2003, you must first uninstall PowerPath** — After upgrading the host, you can reinstall an approved Windows 2003 version. Refer to the EMC PowerPath documentation for further details.
- ◆ **Windows Server 2003 logs Event ID 51 although no data has been lost from the hard disk** — This message may appear during a StorageGroup change, and can be ignored. See Microsoft Knowledge Base article 834910 for details.

Capabilities and limitations

Operating system/driver capabilities and limitations

The following capabilities and limitations should be noted for Windows operating systems.

LUNs Theoretically, Windows supports up to 261,120 total LUNs. This figure is based on the listed support limitations for Windows Server from Microsoft. 8 buses per adapter, 128 targets per bus, and 255 luns per target. However, due to registry hive limitations, a Windows server is most likely to run out of registry space to keep track of these large LUN counts well before reaching this limit.

Microsoft Windows limits the number of LUNs per HBA to 255 (LUNs 00-FE) even though the HBAs are capable of presenting 256 (LUNs 00-FF). It is a function of the operating system that prevents that last LUN from being presented to the user, and not the HBA or driver. This limitation should be taken into consideration when planning your host configuration.

The Emulex SCSIPort driver allows for an increase in the number of LUNs per target that the HBA can report to the OS. Using the ELXCFG configuration tool, it is possible for the HBA to see LUNs with addresses greater than FE. To use these high LUN addresses, set the Max Number of LUNs to your target number (maximum 512), check the Lun Mapping box in the right pane of the configuration tool, and then check the Automatic Lun Mapping box.

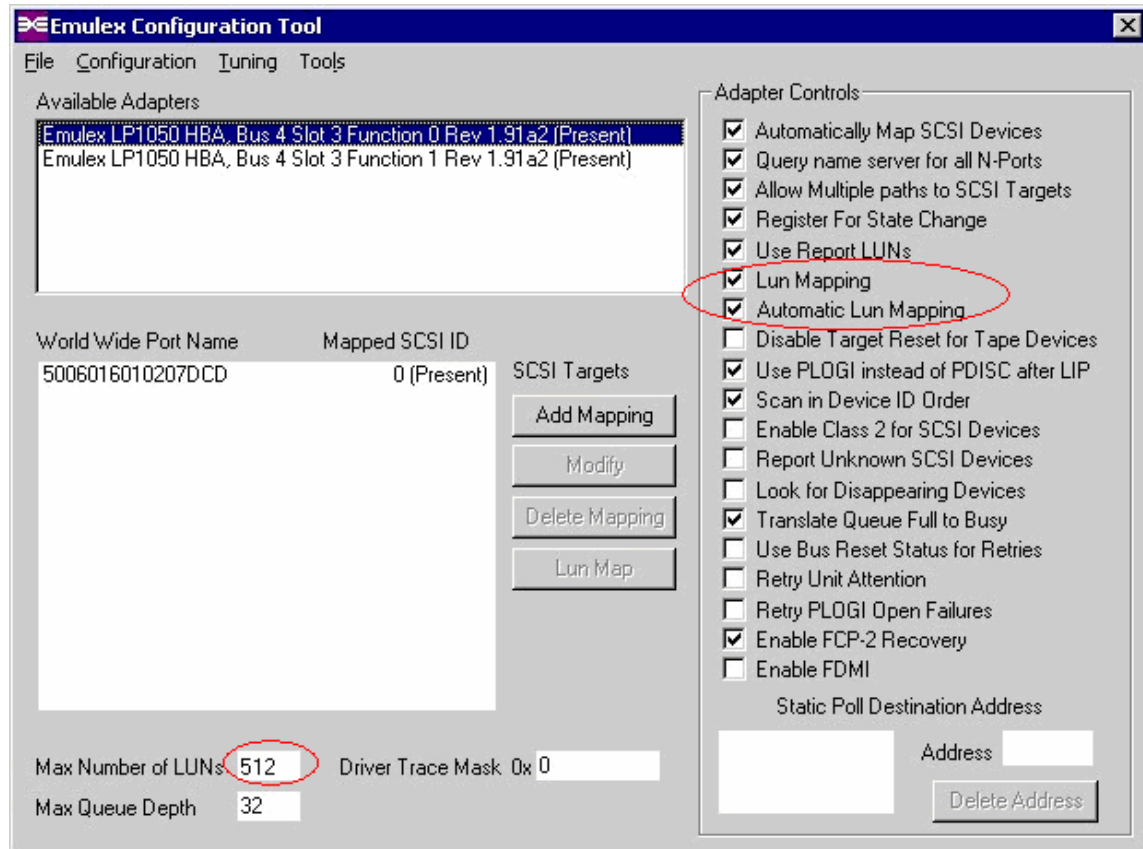


Figure 47 Lun Mapping and Automatic Lun Mapping

Refer to the "Manually Installing the HBA Driver – Advanced Users" section of the *EMC Fibre Channel with Emulex Host Bus Adapters in the Windows Environment* document, available on Powerlink, and your Emulex user guide for more information on using the ELXCFG configuration tool.

The QLogic SCSIPort driver for Windows 2003 is not capable of addressing LUNS above FE as this is a limitation of the operating system. Also, STORPort drivers under Windows 2003 are not capable of addressing LUNs above FE as this is a limitation of the operating system.

Note: This driver is no longer being developed for Windows 2003 configurations.

Note: EMC Storage arrays provide the ability to expand the size of a LUN presented to the host server. Refer to your EMC array's documentation for procedures on expanding LUN sizes. Windows has the ability to recognize the extra space on these expanded LUNs by performing a rescan via the Disk Administrator window.

Volume sizes

Windows 2000 supports a maximum file system size of 2 TB (terabytes.)

Windows 2003 supports a maximum file system size of 2 TB unless Service Pack 1 or 2 is installed. With SP1 or SP2, the maximum supported physical disk size is 256 TB (note that volumes larger than 2 TB must use GPT partitions to support them. Refer to your Windows users guide for information on GPT partitions.)

Note: Windows 2003 for x64 servers does not require SP1 or SP2 to create GPT partitions.

EMC Storage arrays provide the ability to expand the size of a LUN presented to the host server. Refer to your EMC array's documentation for procedures on expanding LUN sizes. Windows has the ability to recognize the extra space on these expanded LUNs by performing a rescan via the Disk Administrator window.

External booting

Bootting externally from an EMC storage array requires an NTFS filesystem if EMC PowerPath is installed on the host. EMC does not support FAT32 filesystems for external boot when using PowerPath.

iSCSI

Microsoft iSCSI Initiator 2.0 known limitations:

- ◆ Dynamic disks on an iSCSI session are not supported.
- ◆ Initiator and target CHAP secrets should each be:
 - 1 through 16 bytes if IPsec is being used.
 - 12 through 16 bytes if IPsec is not being used.
- ◆ The default iSCSI node name is generated from the Windows computer name. If the Windows computer name contains a character that would be invalid in an iSCSI node name(such as the underscore character '_'), the Microsoft iSCSI Initiator service converts the invalid character to a dash (-).

- ◆ If the target returns `CHECK CONDITION` for a SCSI request but does not provide sense data, the initiator completes the request with a status target error. The initiator does not treat the target behavior as a protocol error.
- ◆ The iSCSI control panel applet does not create an icon in the system tray. If the applet is in the background, you can switch to it by using the ALT-TAB key combination or by double-clicking the icon that launches it.

How a server responds to a failure in the boot LUN path

Failure in the path to a SAN-based boot LUN can halt Windows in a fatal error condition. Depending on the failure, Windows may be able to transfer control to another path and continue.

Table 5 shows server reactions to failures in different components.

Table 5 Server response to failure in the boot LUN path (Single fault)

Configuration	Server state	HBA failure	Switch failure	Boot SP/Director Port failure	Boot SP/Director failure	Catastrophic Storage System failure
2 or more HBAs, failover software	Windows running	Multipath	STOP Error ^a	Multipath	Trespass ^b	STOP Error
	Windows booting	Halt	Halt	Halt	Halt	Halt
	Power up	Multipath	No Boot ^a	Multipath	Manual	No Boot
1 HBA, failover software	Windows running	STOP Error	STOP Error	Trespass	Trespass ^b	STOP Error
	Windows booting	Halt	Halt	Halt	Halt	Halt
	Power up	No boot	No boot	Manual	Manual	No boot
1 HBA, no failover software	Windows running	STOP Error	STOP Error	STOP Error	STOP Error	STOP Error
	Windows booting	Halt	Halt	Halt	Halt	Halt
	Power up	No boot	No boot	No boot	No boot	No boot

a. Depending on the fabric configuration, if multiple switches are used, then this behavior would qualify under the Multipath category.

b. CLARiiON only.

Explanations of entries

- ◆ **STOP Error** (fatal blue screen) — Indicates host failure and chance of data corruption.
- ◆ **No boot** — Cannot boot Windows.
- ◆ **Halt** — Windows cannot recover before system has completed startup. (You must reboot and follow the power-up scenario.)
- ◆ **Manual** — Manual intervention is required to continue. (Typically, initiate a LUN trespass using CLI or Manager. With Manager, enable LUN Auto-Assignment in LUN properties.)
- ◆ **Multipath** or **Trespass** — This automatic operation allows no disruption of service. (The delay caused by this operation may affect Windows stability.)

VERITAS Volume Management Software

This appendix contains additional information about VERITAS Volume Management software used with Windows hosts.

- ◆ [VERITAS Volume Management software.....](#) 186

VERITAS Volume Management software

This section contains information on VERITAS Volume Management software for Windows operating systems.

Note: Refer to the latest *EMC Support Matrix* to determine which VERITAS Volume Manager configurations are supported and what service packs may be required.



CAUTION

Configuring large numbers of device paths with VERITAS Volume Manager can cause a Windows system to boot very slowly, and in some cases overrun the NTLDR boot-time registry size and halt. Systems that are configured with more than 512 device paths (total paths to all LUNs) should check with EMC Customer Service before installing VERITAS Volume Manager 3.x.

Note: The C-bit is required on Symmetrix director ports connected to systems running VERITAS DMP. Users of EMC ControlCenter 5.1 and greater should consult their ControlCenter documentation for directions on making this change. Other users must contact their EMC representative to make this change.

Note: For CLARiiON arrays, failover mode one (1) is required for all DMP or MPIO multipathing configurations.

VERITAS Storage Foundation 5.0

Storage Foundation encompasses VERITAS Volume Manager as well as other available volume management software utilities. Refer to the *EMC Support Matrix* for supported features of Foundation Suite.

For version 5.0, Volume Manager and DMP are supported with SCSIPort drivers on Windows 2000 operating systems only. For Windows 2003 VERITAS MPIO multipathing solution is supported with STORPort drivers only and requires an EMC Symmetrix or CLARiiON DSM (device specific module).

At the initial release of version 5.0 in April 2007, DSMs for existing Symmetrix and CLARiiON arrays were included in the native installation of Storage Foundation. Future storage arrays may require an updated DSM for MPIO compatibility. STORPort drivers are also supported in configurations where Powerpath is installed. For Windows 2003 STORPort driver configurations, the Microsoft STORPort hotfix is necessary. Refer to the *EMC Support Matrix* for current STORPort hotfix versions as well as currently supported driver versions.



IMPORTANT

MPIO-based versions of EMC PowerPath cannot be installed at the same time as the VERITAS MPIO multipath solution.

VERITAS MPIO must be disabled when PowerPath is installed.

VERITAS Storage Foundation 4.3

Storage Foundation encompasses VERITAS Volume Manager as well as other available volume management software utilities. Refer to the *EMC Support Matrix* for supported features of Foundation Suite. For version 4.3, Volume Manager and DMP are supported with SCSI Port drivers only. VERITAS MPIO multipathing solution is supported with STORPort drivers only and requires an EMC Symmetrix or CLARiiON DSM (device specific module) to be installed. STORPort drivers are also supported in configurations where Powerpath is installed. For Windows 2003 STORPort driver configurations, the Microsoft STORPort hotfix is necessary. Refer to the *EMC Support Matrix* for current STORPort hotfix versions as well as currently supported driver versions.

Note: VERITAS MPIO configurations do not support load balancing and support a maximum of 16 paths to each device. DMP configurations do not have this 16 path limitation.

VERITAS Storage Foundation 4.2

Storage Foundation encompasses VERITAS Volume Manager as well as other available volume management software utilities. Refer to the *EMC Support Matrix* for supported features of Storage Foundation. For version 4.2, Volume Manager and DMP are supported with SCSIPort drivers only. STORPort drivers are supported in configurations where PowerPath is installed only. For Windows 2003 STORPort driver configurations, the Microsoft STORPort hotfix is necessary. Refer to the *EMC Support Matrix* for current STORPort hotfix versions as well as currently supported driver versions.

VERITAS Foundation Suite 4.1

Foundation suite encompasses VERITAS Volume Manager as well as other available volume management software utilities. Refer to the *EMC Support Matrix* for supported features of Foundation Suite.

For version 4.1, Volume Manager and DMP are supported with SCSIPort drivers only. VERITAS does not support STORPort drivers for Windows 2003 configurations.

VERITAS Volume Manager 3.1 and VERITAS DMP

If using PowerPath with VERITAS Volume Manager 3.1, you also need VERITAS Volume Manager Service Pack 1 or 2.

EMC and VERITAS now provide a Dynamic Multipathing (DMP) Driver Update for VERITAS DMP to interface with CLARiiON CX-Series arrays, providing DMP high-availability capability. Refer to the *EMC Support Matrix* for the minimum supported revisions of VxVM and DMP, as well as the CLARiiON Dynamic Multipathing Driver update.

Note: For more information about DMP and CLARiiON, refer to [page 145](#).

VERITAS Volume Manager 3.0

If using PowerPath with VERITAS Volume Manager 3.0 with Service Pack 1 or 2, you must also make the following registry modifications before PowerPath devices will be available to the VERITAS Enterprise Manager:

Use **regedt32.exe** to set the registry as follows:

```
HKEY_LOCAL_MACHINE\SOFTWARE\VERITAS\VxSvc\CurrentVersion\VolumeManager
    value name = ShowGateKeeperDevices
    data type = REG_DWORD
    value = 0x1
HKEY_LOCAL_MACHINE\SOFTWARE\VERITAS\VxSvc\CurrentVersion\VolumeManager
    value name = ShowEmcHiddenDevices
    data type = REG_DWORD
    value = 0x1
```

After completing these changes, reboot the host system.

A

- addressing, Windows/Symmetrix
 - arbitrated loop 147
 - fabric 149
 - SCSI-3 150

B

- boot support
 - Symmetrix 86

C

- CHAP 92
- CLARiiON
 - support information 140
 - VERITAS DMP 145
 - VERITAS Volume Manager 145
- CLARiiON setup 140
- comments 13
- customer support 13

D

- devices, Symmetrix, adding on line
 - Windows 2000 27
- direct access device 151
- Disk Administrator 16
- disks, Symmetrix
 - initializing 21

E

- error messages, Windows 2000/Windows 2003
 - 26

- errors, drive, recovering from on Windows 2000
 - host 29

F

- failure in boot LUN path, server response to 183
- file shares 137
- functions, Windows 16

H

- help 13

I

- I/O timeout value, adjusting 28
- initializing Symmetrix disks 21
- inq 178
- Inquiry utility 178
- IP address, Symmetrix 137
- iSCSI Network Portal 92
- issues 178

L

- LIP 147
- logical unit addressing 150
- loop initialization process 147

P

- peripheral device addressing 150
- persistent binding 170

S

- SCSI-3 FCP 147
- server response to failure in boot LUN path 183
- service 13
- Solutions Enabler 175
- status messages, Windows 2000/Windows 2003 26
- storage system components
 - CLARiiON 140
- support information
 - CLARiiON 140
 - Fibre Channel 140
- Symmetrix configuration 146
- Symmetrix IP address 137

T

- technical support 13
- terminology, Windows 16
- timeout value, adjusting 28

U

- utilities, Windows 16

V

- VERITAS DMP (Dynamic Multi-Pathing)
 - with CLARiiON 145
- VERITAS Volume Manager (VxVM)
 - with CLARiiON 145
- VERITAS Volume Manager 3.0 189
- volume set addressing 150
- volumes, Symmetrix, creating for Windows 2000 or Windows 2003 24

W

- Windows I/O timeout value, adjusting 28

Z

- zone, planning in Windows/Symmetrix environment 87