Dell™ PowerVault™ NAS Fibre Channel Cluster Systems Installation and Troubleshooting Guide

---

NOTE: A NOTE indicates important information that helps you make better use of your computer.

NOTICE: A NOTICE indicates either potential damage to hardware or loss of data and tells you how to avoid the problem.

CAUTION: A CAUTION indicates a potential for property damage, personal injury, or death.

Initial release: September 2003

Back to Contents Page

Introduction

Dell™ PowerVault™ NAS Fibre Channel Cluster Systems Installation and Troubleshooting Guide

This guide provides information for installing a Dell™ PowerVault™ NAS Fibre Channel cluster solution in a corporate enterprise running the Microsoft® Windows® Storage Server 2003, Enterprise Edition operating system (hereafter known as NAS Manager).

The information in this guide includes:

- Basic Fibre Channel cluster installation procedures, which include:
  - Preparing server and storage systems for clustering
  - Cabling the cluster configuration
  - Configuring the peripherals for the cluster, including the HBA and RAID controllers
- Installation procedures for installing direct-attached and SAN attached cluster configurations in your corporate network
- Cluster upgrading and maintenance procedures
- Information about MSCS, the clustering software built into the Windows Storage Server 2003, Enterprise Edition operating system
- Troubleshooting procedures
- Data sheets for recording critical cluster configuration information

---

Intended Audience

This guide was developed for experienced IT professionals who need to install, cable, and configure a NAS cluster solution in an enterprise environment and for trained service technicians who perform cluster upgrade and maintenance procedures. This guide also addresses readers who are new to clustering technology.

Obtaining More Information

See "Overview" for a general description of NAS clusters and clustering technology.

See "Using MSCS" for an overview of the clustering software built into the Windows Storage Server 2003, Enterprise Edition operating system.

Obtaining Technical Assistance

Dell Enterprise Training and Certification is available; see **www.dell.com/training** for more information. This service may not be offered in all locations.

Overview

This section provides an overview of clustering and the major cluster components used in the Windows Server 2003, Enterprise Edition operating system.

NOTE: In this guide, Microsoft Cluster Service for Windows Storage Server 2003, Enterprise Edition is also referred to as MSCS.

Clustering

Clustering is the process of joining two or more NAS systems together to function as a single system. Using specific hardware and software that is interconnected to work as a single entity, clustering provides an automatic failover solution to hardware or software failures. If one of the clustered systems (also known as cluster nodes) fail for any reason, the user resources running on the failed system are moved (or failed over) to one or more systems in the cluster by the MSCS software—the failover software component in specific versions of the Windows operating system. When the failed system is repaired and brought back online, user resources automatically transfer back to the repaired system or remain on the failover system, depending on how MSCS is configured.

The availability of network services is critical to applications in a client/server environment. Clustering reduces the amount of downtime caused by unexpected failures, providing maximum uptime of mission critical applications—also known as high availability—that surpasses the capabilities of a stand-alone system. Using MSCS, clustering ensures that applications on a failed cluster node continue on the remaining node by migrating and managing the required resource to the remaining node in the cluster. Clusters that reduce the amount of system downtime are known as *high-availability clusters*.

Virtual Servers and Resource Groups

In a standard client/server environment, a user accesses a network resource by connecting to a physical server (such as a NAS system) with a unique IP address and network name. If the server fails for any reason, the user will no longer be able to access the resource. In a cluster environment, a user does not access a physical server, but a virtual server—a network resource managed by MSCS.

Each virtual server, which is transparent to the user, has its own IP address, server name, and disk drive in the shared storage system. MSCS manages the virtual server as a *resource group*, which contains a list of the cluster resources. Virtual servers and resource groups are transparent to the network client.

See "Cluster Resources" in "Using MSCS" for more information on network resources.

Virtual servers are designed to dynamically reconfigure user resources during a connection or hardware failure, providing a higher availability of network resources as compared to a nonclustered NAS system. When MSCS detects a failed cluster node or failed application, MSCS moves the entire virtual server resource group to another cluster node and remaps the virtual server to the new network connection. The network client attached to an application in the virtual server will only experience a momentary delay in accessing their resources while MSCS re-establishes a network connection to the virtual server. This process of moving and restarting a virtual server on a healthy cluster node is called *failover*.

See "Groups" in "Using MSCS" for more information on resource groups.

Failover and Failback

If one of the cluster nodes should fail for any reason, MSCS moves (or *fails over*) the virtual server to another cluster node. After the cluster node is repaired and brought online, MSCS may move (or *fail back*) the virtual server to the original cluster node, if required. This failover capability enables the cluster configuration to keep network resources and application programs running on the network while the failed node is taken offline, repaired, and brought back online. The overall impact of a node failure to network operation is minimal.

See "Failover and Failback" in "Using MSCS" for more information.

EMC® PowerPath®, a software component that is installed on the cluster nodes and stand-alone systems, provides an I/O path failover solution for the I/O traffic between the host system and the storage system. Each cluster node has redundant paths to a storage system in the cluster. If a path between the cluster nodes and the shared storage system fails for any reason, PowerPath reroutes the Fibre Channel I/O traffic to the remaining path(s). After the failed path is repaired, PowerPath can resume network traffic on the original path.

NOTE: The NAS cluster solution requires EMC PowerPath.

See "Storage Management Software" for more information about EMC PowerPath.

Quorum Resource

In every cluster, a single disk resource is designated as the quorum resource. This resource maintains the configuration data necessary for cluster recovery when a cluster node fails. The quorum resource contains the details of all the changes that have been applied to a cluster database.

The quorum resource can be any resource with the following attributes:

- Enables a single node to gain and defend its physical control of the quorum resource.

  For example, SCSI or Fibre Channel disks use reserve and release commands for arbitration.
- Provides physical storage that is accessible by any node in the cluster.
- Uses the NTFS file system.

See "Quorum Resource" in "Using MSCS" and the MSCS online documentation for more information.

NOTE: NAS Fibre Channel clusters do not support the Majority Node Set Quorum resource type.

Shared Storage Systems

In MSCS cluster configurations, Fibre Channel provides the means for the cluster nodes to share access to one or more external storage systems. However, only one of the cluster nodes—one of the servers in the cluster—owns any RAID volume in the external storage system at any point in time. The Cluster Service maintains control over which node has access to each RAID volume in the shared storage system.

See "Fibre Channel Protocol" for more information on Fibre Channel storage systems.

NOTE: The Access Logix™ software is required when the storage system is connected to two or more clusters, two or more nonclustered systems, or a combination of both clustered and nonclustered systems. Access Logix software is not required for one stand-alone system or single cluster configurations.

Each storage system in the cluster is centrally managed by one host system (also called a *management station*) running EMC ControlCenter™ Navisphere Manager™—a storage management application used to configure Dell | EMC storage systems. Navisphere Manager provides centralized storage management and configuration from a single management console. Using a graphical user interface (GUI), you can select a specific view of your storage arrays, as shown in Table 1-1.

Table 1-1. Navisphere Manager Storage Views

View	Description
Storage	Shows a "map" of the hardware configuration to identify hardware faults and shows the logical storage components and their relationships to each other.
Host	Shows the host system's storage group and attached LUNs.
Monitors	Shows all Event Monitor configurations. These configurations include centralized and distributed monitoring configurations.

You can use Navisphere Manager to perform multiple tasks, such as creating RAID groups, binding LUNs, and downloading firmware. Combined with Access Logix, administrators can manage and control data access across multiple heterogeneous hosts in distributed SANs.

Optional software for the shared storage systems include:

- EMC MirrorView™ — Provides synchronous mirroring between two storage systems in a campus environment.
- EMC SnapView™ — Captures point-in-time images of a LUN for backups, decision support, or testing without affecting the contents of the source LUN.
- EMC SAN Copy — Moves data between Dell | EMC storage systems without using host CPU cycles or LAN bandwidth.

See "Installing and Configuring the Shared Storage System" for more information about Navisphere Manager, Access Logix, MirrorView, SnapView, and SAN Copy.

See "Shared Storage Systems" for a description of the shared storage systems for the NAS cluster solution.

NAS Fibre Channel Cluster Solution

The NAS Fibre Channel cluster solution implements two-node clustering technology based on the MSCS software incorporated within the Windows Storage Server 2003, Enterprise Edition operating systems. This cluster solution provides the following benefits in meeting the needs of mission-critical network application programs:

- 2 Gb/s Fibre Channel technology
- High availability of system services and resources to network clients
- Redundant storage for application program data
- Failure recovery for cluster application programs
- Flexible maintenance capabilities, allowing you to repair, maintain, or upgrade a cluster node or cluster storage system without taking the entire cluster offline

Each cluster node is configured with software, storage, and network resources that enable it to monitor and interact with the other nodes to provide mutually redundant operation. If a cluster node fails for any reason, virtual servers and resources groups running client resources are failed over to other healthy cluster nodes. When the failed node is repaired and brought back online, the virtual servers and resource groups are failed back to the repaired node (if desired).

The cluster nodes, therefore, operate as a single resource, rather than a collection of individual systems. Because the cluster nodes interact in this way, they appear as a single system to the network clients.

Storage System

NAS Fibre Channel clusters support the Dell | EMC CX200, CX400, and CX600 systems—modular 2-Gb/s Fibre Channel storage systems that provide scalable storage solutions in a direct-attached or SAN-attached environment.

These storage systems consist of multiple hardware components that work together as a single system. These components include:

- Processor enclosure (DPE or SPE)
- Disk array enclosure (DAE2)
- Standby power supply
- Host bus adapter

Processor Enclosure

The processor enclosure is configured with storage processors that control the RAID arrays in the storage system and provide various storage functionalities, such as snapshots, LUN masking, and remote mirroring.

The CX200 and CX400 are DPEs and are configured with internal hard drives. The CX600 is a storage processor enclosure (SPE) and is not configured with internal hard drives. For primary storage, the SPE must be attached to one or more 2-Gb/s DAEs. The DPE and SPE are controlled by the core software (FLARE), which hosts the Navisphere® management software and manages the workload between the storage processors and the hard drives.

The SPE requires two standby power supplies (SPS) to provide backup power to the SPE storage processors and the first DAE2 or DAE2-Operating System (DAE2-OS) enclosure that is connected to the SPE. If a power failure occurs on the SPE, the backup power allows the SPE to transfer unprocessed write cache to the DAE2 vault disks before it shuts down, thereby preventing data loss. See Table 1-13 in for additional information.

NOTE: The DAE2 and DAE2-OS are 2-Gb/s DAE enclosures. The DAE2-OS enclosure is the first DAE2 enclosure that is connected to the CX600 and has the core software (FLARE) preinstalled on the first five hard drives. These first five hard drives are called the *vault disks*.

The processor enclosure includes the following features:

- Two to four 2-Gb/s front-end ports per storage processor, providing redundant paths to the storage system. If one of the paths fail for any reason, the data is routed through the redundant path until the failed path is repaired.
- EMC ControlCenter Navisphere management software for storage system management, including:
    - SnapView (for point-in-time data copies)
    - MirrorView (for remote synchronous data mirroring)
    - SAN Copy (for moving data between Dell | EMC storage systems)

  NOTE: The CX200 does not support SnapView and Mirrorview.
- Supports nondisruptive upgrades for software and firmware.
- Storage for multiple clustered and nonclustered hosts.
- Supports RAID 0, 1, 1/0, 3, and 5.
- Compatible with the Dell | EMC FC4700 and FC4700-2 storage systems as a MirrorView target or source.

NOTICE: The hardware components in the processor enclosures are not compatible with the components in the Dell | EMC FC4500, FC4700, and FC4700-2 DAEs. Exchanging components between these enclosures could damage the enclosures and void your warranty.

The supported processor enclosures vary in performance and configuration. Table 1-2 provides the features for each processor enclosure.

See Table 1-13 in for more information on the supported processor enclosures.

Table 1-2. Processor Enclosure Features

CX200	CX400	CX600
3-U entry-level modular storage system.	3-U midrange modular storage system.	4-U enterprise modular storage system.
Two Intel® 800-MHz Mobile Pentium® III processors.	Two Intel 800-MHz Pentium III processors.	Two Intel 2-GHz Pentium 4 processors.
1-GB cache.	Up to 2-GB cache.	Up to 8-GB cache.
Supports up to 15 HBAs per SP.	Supports up to 15 HBAs per port	Supports up to 32 HBAs per port.
Up to 200-MB/s data transfer rate.	Up to 680-MB/s data transfer rate.	Up to 1300-MB/s data transfer rate.
Supports up to 15 internal hard drives.	Supports up to 15 internal hard drives.	No internal hard drive support.
Two front-end ports in a hub configuration that connect to the hosts.	Four front-end ports that connect to the hosts.	Eight front-end ports that connect to the hosts.
Two back-end ports that connect to disks through the DAE2 enclosures.	Four back-end ports that connect to disks through the DAE2 enclosures.	Four back-end ports that connect to disks through the DAE2 enclosures.
Supports one DAE2 enclosure for additional data storage.	Supports up to three DAE2 enclosures for additional data storage.	Supports up to 16 DAE2 enclosures for additional storage.
Supports one cluster in a direct-attached cluster configuration.	Supports one cluster in a direct-attached cluster configuration.	Supports up to two clusters in a direct-attached cluster configuration.
Up to 30 drives per array with a maximum storage capacity of 4.4 TB (using fifteen 1-inch 146-GB hard drives for each DAE2 enclosure).	Up to 60 drives per array with a maximum storage capacity of 8.8 TB (using fifteen 1-inch 146-GB hard drives for each DAE2 enclosure).	Up to 240 drives per array with a maximum storage capacity of 35 TB (using fifteen 1-inch 146-GB hard drives for each DAE2 enclosure).
Manages the RAID arrays on the DPE and the DAE2.	Manages the RAID arrays on the DPE and the DAE2.	Manages the RAID arrays on the DAE2.
Core software (FLARE) and vault disks are located on the first five internal hard drives.	Core software (FLARE) and vault disks are located on the first five internal hard drives.	Core software (FLARE) and vault disks are located on the first attached DAE2 (or DAE2-OS).
Supports RAID 0, 1, 1/0, 3, and 5.	Supports RAID 0, 1, 1/0, 3, and 5.	Supports RAID 0, 1, 1/0, 3, and 5.
Supports nondisruptive upgrades for online software and firmware.	Supports nondisruptive upgrades for online software and firmware.	Supports nondisruptive upgrades for online software and firmware.
No support for MirrorView, SnapView, and SAN Copy.	Supports MirrorView, SnapView, and SAN Copy.	Supports MirrorView, SnapView, and SAN Copy.
NOTE: The core software (FLARE) contains the operating system that powers the storage system. The vault disks are the first five hard drives in the DPE (CX200 and CX400) or the first attached DAE2 (CX600) that store unprocessed write cache if the storage system power sources fail.		

Disk Array Enclosure

The NAS cluster solution uses the DAE2—a 2-Gb/s, 3-U DAE that can be configured with up to 15 dual-ported, hot-swappable Fibre Channel hard drives per enclosure. The DAE2 provides a scalable storage solution, allowing you to add

additional hard drives to the enclosure as needed.

The DAE2 is also available as the DAE2-OS—a DAE2 enclosure with the core software (FLARE) preinstalled on the first five hard drives that provides both the primary storage and the operating system for the CX600 storage system.

Additional features include:

- 2-Gb/s Fibre Channel data transfer rate
- Maximum capacity of fifteen 1-inch 146-GB hard drives
- Redundant hot-swappable power supplies, cooling fan assemblies, and link control cards (LCCs)

Table 1-3 lists the number of supported DAE2 enclosures for each storage system.

Table 1-3. Supported DAE2 Enclosures

Storage System	Supported DAE2 Enclosures
CX200	One
CX400	Up to 3
CX600	Up to 16

.

SPS

The SPS is a 1-U power supply that provides backup power to the processor enclosures or the CX600 SPE and its first DAE2 (or DAE2-OS) enclosure. The storage system requires two SPSs, with each SPS connected to a separate power source for redundancy. If both power sources fail, the SPSs enable the SPE to flush the write cache to the vault drives. When the power is restored, the SPE writes the cache content in the vault drives back to the cache. This method ensures that the SPE cached data is not lost during a power failure.

Each SPS is hot-swappable, and two 1000-W DC SPSs are required for the following:

- CX200 — DPE
- CX400 — DPE
- CX600 — SPE and its first DAE2-OS enclosure

The SPS installs in the SPS rack kit below the SPE.

NOTICE: To avoid data loss, do not disconnect any power or SPS cables or turn off the CX200 DPE, CX400 DPE, CX600 SPE, or DAE2 power supplies during normal operation. If you need to shut down the SPE, DPE, or DAE2 enclosure(s), stop all I/O to the storage system, stop the Cluster Service, and then shut down all the host systems attached to the storage systems. After you perform these procedures, turn off the SPS devices by using their power switches.

HBA

The HBA is an I/O adapter that connects the server's PCI bus to the storage components. Two HBAs must be installed in each server in the SAN to manage the I/O data transfer from the server to the storage system. See the *Platform Guide* for a list of supported HBAs.

Fibre Channel Switch

The Fibre Channel switch functions as a director, mapping requests and responses between the interconnected devices and provides a dedicated interconnection between the server and storage system. See the *Platform Guide* for a list of supported Fibre Channel switches.

NOTE: The Dell | EMC DS-24M2 (McData Spherion 4500) Fibre Channel switch supports Flexport configurations, which allows you to vary the number of active switch ports. Flexport configuration is dependent on firmware license key and the appropriate number of SFP modules. When you update the firmware license key, you are provided with the appropriate number of SFP modules and a new key that allows you to access additional ports on the switch.

Hardware and Software Technologies

The NAS cluster solution implements the following hardware and software technologies:

- Clustering technology based on the MSCS software in the Windows Storage Server 2003, Enterprise Edition operating system
- Fibre Channel protocol
- Fibre Channel switch fabric
- Zones
- SAN components
- Storage management software

Clustering Technology

Clustering is the process of connecting multiple servers together to achieve higher availability and performance. MSCS is a software component that is included with the Windows Storage Server 2003, Enterprise Edition operating system that provides failover support for applications and services running on each node.

See "Using MSCS" for more information on the Cluster Service.

NOTE: MSCS and NLB features cannot co-exist on the same cluster node, but can be used together in a multitiered cluster configuration. For more information, see the Dell PowerEdge Clusters website located at **www.dell.com/clusters** or the Microsoft website located at **www.microsoft.com**.

Fibre Channel Protocol

Fibre Channel is a scalable, high-performance data communications technology that allows multiple server systems to share one or more storage systems. Unlike the SCSI technology, which is limited to short-distance connections and direct-attached storage, Fibre Channel provides long-distance connectivity and the high bandwidth needed for transferring data between the cluster nodes and the shared storage devices in a NAS cluster. By employing long-wave fiber optic cable between cascaded switches, systems up to 10 km from the shared storage array can access data as if they are directly attached.

Implementing Fibre Channel technology in the NAS cluster provides you with the following advantages:

- **Flexibility** — Fibre Channel implements both copper and optical cabling, allowing a distance of up to 10 kilometers between switches without signal degradation.
- **Availability** — Fibre Channel components implement redundant connections, providing multiple data paths and greater availability for network clients.
- **Connectivity** — Fibre Channel allows you to connect more devices to each other than SCSI. Because Fibre Channel devices are hot-pluggable, you can add or remove devices from the cluster nodes without bringing down the cluster.

NAS Fibre Channel clusters support 2-Gb Fibre Channel, which provides high-speed data transfer between the cluster nodes and storage systems and supports the smaller SFP connectors for higher port density.

Fibre Channel Switch Fabric

A Fibre Channel switch fabric consists of one or more Fibre Channel switches that provide high-speed connections between servers and storage devices. The switches in a Fibre Channel fabric provide a connection through inbound and outbound points from one device (sender) to another device or switch (receiver) on the network. If the data is sent to another switch, the process repeats itself until a connection is established between the sender and the receiver.

Fibre Channel switches are linked together using ISLs. These ISLs use two Fibre Channel ports to connect the switches together. Fibre Channel fabrics provide you with the ability to set up barriers between different devices and operating environments. These barriers create logical fabric subsets with minimal software and hardware intervention. Similar to VLANs in the client/server network, logical fabric subsets divide a fabric into smaller groups or components, regardless of their proximity to one another. The logical subsets that form these barriers are called *zones*.

Zones

Zones help to segment a SAN into logical fabric subsets by setting up barriers between different operating environments with minimal software and hardware intervention. Similar to subnets in the client/server network, these logical fabric subsets (or zones) divide a fabric into smaller groups or components, regardless of their proximity to one another. By implementing switch zoning together with Access Logix, you can attach multiple clusters or a combination of clusters and stand-alone systems to a SAN.

NOTE: Dell supports only single-initiator zoning for connecting NAS clusters to the CX200, CX400, and CX600 storage systems in a switched environment. A separate zone is created for each HBA, which includes the HBA and one or more SP ports on the storage systems.

SAN

Cluster nodes attach to the storage systems through a cable connection (using a direct-attached configuration) or through a SAN—a high-performance network storage solution. Similar to a LAN, a SAN is used to move data between servers and storage systems, providing a high-speed storage resource for your cluster nodes. Using a specific collection of hardware and software releases set to specific version levels, a SAN bypasses traditional network bottlenecks, providing a high-speed, highly available data consolidation solution for your NAS systems.

See the EMC Support Matrix located in the EMC Technical Library at **www.emc.com** and the *Platform Guide* included with your system documentation located at the Dell Support website at **support.dell.com** for the latest firmware and software requirements.

SAN Components

Table 1-4 lists the key hardware and software components in a SAN.

NOTE: Your SAN may require additional hardware and software components that are not previously listed. See the EMC Support Matrix located in the EMC Technical Library at **www.emc.com** for information about SAN-compliant hardware and software components.

Table 1-4. SAN Components

Component	Description

HBA	Connects the cluster node's PCI bus to the storage system.
Storage system	Provides external storage for the host systems (cluster nodes and stand-alone systems). Supported storage systems include tape libraries, DAEs, and SPEs.
SAN storage management software	Provides centralized control of the SAN for easier management. These software management tools include: • Navisphere Agent • Navisphere Manager • EMC PowerPath • Access Logix (optional) • MirrorView (optional) • SnapView (optional) • Backup software (optional) • SAN Copy (optional) See "Storage Management Software" for a description of these software management tools.
Fibre Channel switch (optional)	Provides a communications link between the servers, storage systems, and tape libraries.
Fabric (optional)	One or more switches, which may be connected together using interswitched links.

Storage Management Software

NAS clusters require additional software components to manage the communications between the cluster nodes and the storage systems in a SAN environment.

Table 1-5 provides a list of Navisphere and Dell OpenManage™ software packages that are used in NAS cluster configurations.

See "Minimum System Requirements" for the storage management software required for your particular cluster configuration.

Table 1-5. Storage Management Software

Software	Function
EMC PowerPath	Detects and re-establishes a failed connection to a processor enclosure caused by a communication failure, such as a failed storage processor, SP port, link control card (LCC), HBA, switch, or fiber optic cable. Without human intervention, PowerPath can automatically reroute the I/O through an alternate path when a failure occurs to provide connectivity for running applications. Additional features include: • Dynamic load balancing • Multiple path detection
Navisphere Manager	Provides centralized storage management and configuration from a single management console. Using a GUI, Navisphere Manager allows you to configure and manage the disks and components in one or more shared storage systems. Additional features include: • Creating, binding, and unbinding LUNs • Changing configuration settings • Monitoring storage systems Navisphere Manager is installed on the Dell | EMC storage systems or a host system that has network access to the storage system.
Navisphere Agent	Provides an interface between the host systems and the Dell | EMC storage system, allowing Navisphere Manager to send and receive information to and from the storage system connected to a host system.
Access Logix (optional)	Allows multiple cluster nodes and servers to share a Fibre Channel shared storage system. Access Logix restricts server access to specific volumes on a shared Dell | EMC storage system and protects your data from unauthorized access.

	Access Logix is required when the storage system is connected to two or more clusters, two or more nonclustered systems, or a combination of both clustered and nonclustered systems.
MirrorView (optional)	Automatically duplicates primary storage system data for a cluster or stand-alone system to a secondary storage system in a campus environment. MirrorView implements synchronous mirroring, which ensures that data is updated on both the primary and secondary storage systems at the same time. You can manage up to 100 mirrors per array or 50 mirrors if the Write Intent Log is initiated. MirrorView is an optional software component that is installed on the storage processors as a nondisruptive upgrade and is transparent to users and software applications. MirrorView can be used in conjunction with SnapView and is managed from within Navisphere Manager.
SnapView (optional)	Captures point-in-time images of a LUN and retains the image independently of subsequent changes to the files. SnapView images can be used to make shared LUNs available to another system for backups, decision support, or testing without affecting the contents of the source LUN. SnapView captures LUN images in one of two formats: • Snapshot — A *point-in-time* image of the source LUN. Snapshots can be accessed by cluster nodes from a different cluster or stand-alone systems as an additional mountable disk volume (or LUN) with read/write access privileges. You can create up to 100 individual snapshots or up to eight snapshots of the same source LUN. You can create up to 100 individual clones or up to eight clones of the same source LUN. • Clone — A *static* image of the original source LUN. Clones can be used to restore corrupted LUNs and migrate LUNs between RAID groups and SPs, allowing you to fine tune your RAID arrays. **NOTE:** Each snapshot or clone must be accessed from a different host. SnapView is an optional software component that is installed on the DPE or SPE storage processors as a nondisruptive upgrade and is transparent to users and software applications. SnapView can be used in conjunction with MirrorView and is managed from within Navisphere Manager.
SAN Copy (optional)	Allows you to move data between Dell | EMC storage systems without using host processor cycles or LAN bandwidth. SAN Copy is an optional software component that is installed on the storage processors as a nondisruptive upgrade and is transparent to users and software applications. SAN Copy can be used in conjunction with SnapView or MirrorView and is managed from within Navisphere Manager.

Optional Cluster Configurations

- Direct-attached cluster
- SAN-attached cluster

The following sections provide detailed information and examples for these options.

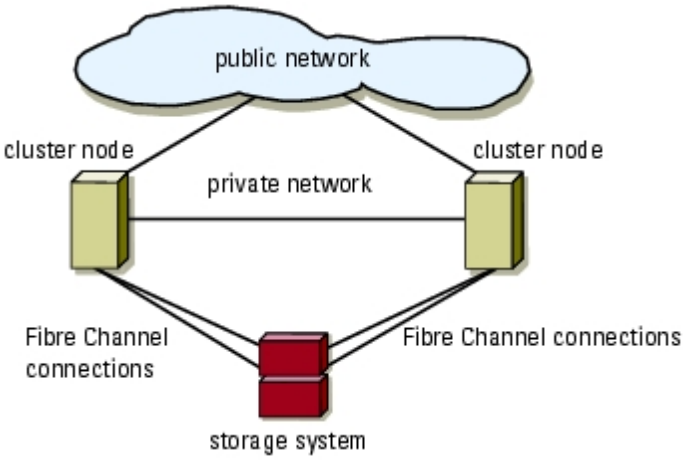
Direct-Attached Cluster

In a direct-attached cluster configuration, both nodes of the cluster are directly attached to a single storage system. In this configuration, the RAID controllers (or storage processors) on the storage systems are connected by cables directly to the Fibre Channel HBAs in the cluster nodes.

NOTE: A direct-attached cluster configuration does not use a SAN.

Figure 1-1 shows a basic direct-attached, single cluster configuration for a NAS Fibre Channel cluster configuration.

Figure 1-1. Direct-Attached, Single Cluster Configuration



Using EMC PowerPath in a Direct-Attach Cluster

EMC PowerPath provides failover capabilities and multiple path detection as well as dynamic load balancing between multiple ports on the same SP. Direct-attached clusters supported by Dell connect to a single port on each SP in the storage system. Because of the single port limitation, PowerPath can only provide failover protection in a direct-attached configuration.

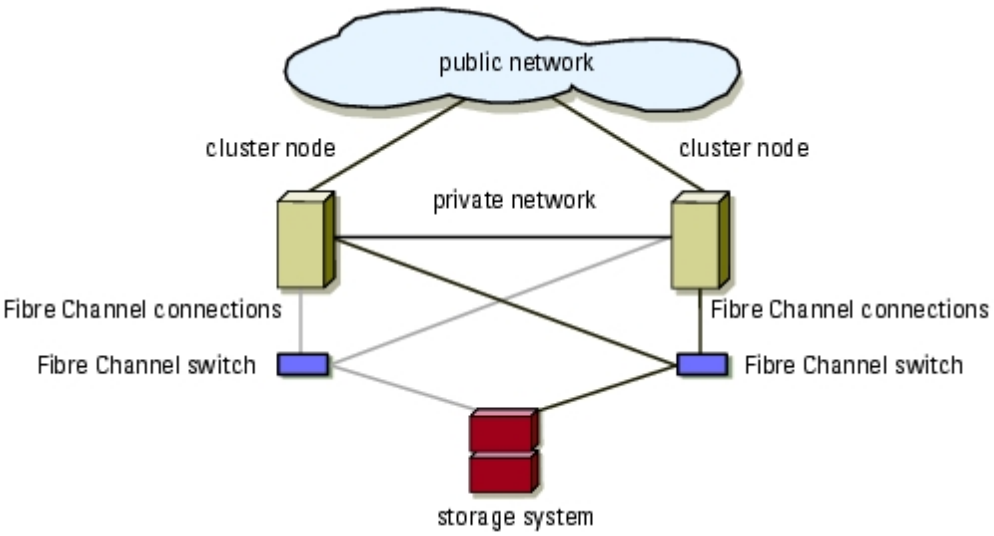
SAN-Attached Cluster

In a SAN-attached cluster configuration, all of the cluster nodes are attached to a single storage system or to multiple storage systems through a SAN using a redundant switch fabric. SAN-attached cluster configurations are superior to direct-attached cluster configurations in configuration flexibility, expandability, and performance.

See "Fibre Channel Switch Fabric" for more information on Fibre Channel switch fabrics.

Figure 1-2 shows a SAN-attached cluster configuration.

Figure 1-2. SAN-Attached Cluster Configuration



Configuring Active and Passive Cluster Nodes

Cluster configurations may include both active and passive cluster nodes. Active nodes are the primary nodes in the cluster. These nodes support the cluster workload by processing application requests and providing client services. Passive nodes are backup nodes that support the active nodes if a hardware or software failure occurs, ensuring that client applications and services are highly available.

NOTE: Passive nodes must be configured with the appropriate processing power and storage capacity to support the resources that are running on the active nodes.

An active/active (active^{x}) configuration is a cluster with virtual servers running separate applications or services on each node. When an application is running on node 1, the remaining cluster node does not have to wait for node 1 to fail. The remaining cluster node can run its own cluster-aware applications (or another instance of the same application) while providing failover capabilities for the resources on node 1. This configuration requires careful planning to ensure that adequate resources are available on each node to handle the increased load if one node fails.

An active/passive (active^{x}/passive^{x}) configuration is a cluster where an *active* cluster node is processing requests for a clustered application while the *passive* cluster node simply waits for the active node to fail. For example, *N* + *I* failover (where *N* = the number of *active* nodes and *I* = the number of *inactive* [or *passive*] nodes) is an active/passive failover policy where a dedicated, passive cluster node provides backup for the active cluster node.

Active/passive configurations are more costly in terms of price and performance because the passive cluster node remains idle all of the time. This configuration is appropriate for business-critical systems because the application can use all of the resources of a standby cluster node if one active cluster node fails.

Table 1-6 provides a description of some active/passive configuration types.

Table 1-6. Active/Passive Configuration Types

Configuration Type	Active Cluster Node(s)	Passive Cluster Node(s)	Description
Active^{1}/Passive^{1}	1	1	The active node(s) processes requests while the passive node waits for the active node to fail.
Active^{2}/Passive^{1}	2	1	
Active^{2}/Passive^{2}	2	2	
Active^{3}/Passive^{1}	3	1	
Active^{3}/Passive^{2}	3	2	
Active^{4}/Passive^{1}	4	1	
Active^{4}/Passive^{2}	4	2	
Active^{5}/Passive^{1}	5	1	
Active^{5}/Passive^{2}	5	2	
Active^{6}/Passive^{1}	6	1	
Active^{6}/Passive^{2}	6	2	
Active^{7}/Passive^{1}	7	1	

Failover and Failback Support

One of the key features of Cluster Service is failover and failback support. Failover is the process of automatically moving resources from a failed cluster node to another cluster node. Failback is the process of moving the resources back to the original cluster node (if required). Failback can be an automatic or manual process, depending on how you configure the

Failover Policies

When implementing a failover policy, configure failback if the cluster node lacks the resources (such as memory or processing power) to support one or more cluster node failures.

Windows Storage Server 2003, Enterprise Edition Cluster Configurations

NAS cluster configurations running Windows Storage Server 2003, Enterprise Edition provides the following failover policies:

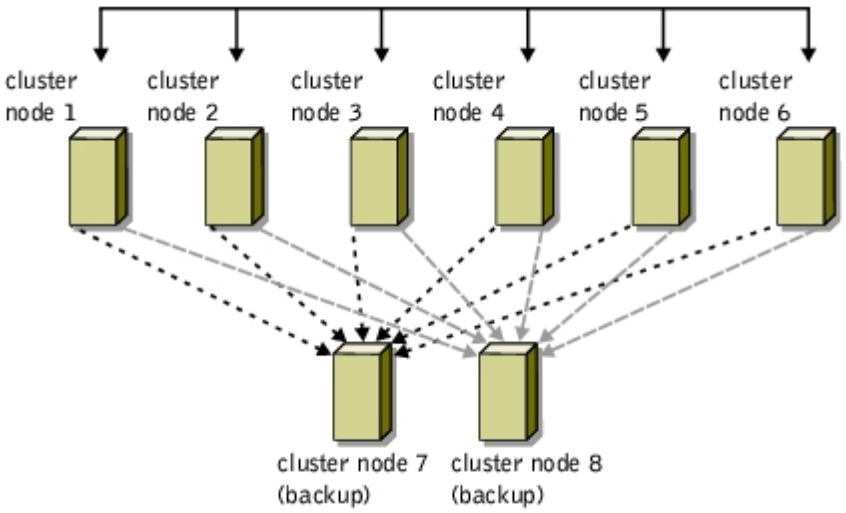
- *N* + *I* failover
- Failover pair
- Multiway failover
- Failover ring

N + I Failover

N + *I* failover (where *N* = the number of *active* nodes and *I* = the number of *inactive* [or *passive*] nodes) is an active/passive failover policy where dedicated, passive cluster node(s) provide backup for the active cluster node(s). This failover solution provides the best solution for critical applications that require dedicated server resources. However, the backup cluster nodes add a higher cost of ownership to the cluster because they remain idle and do not provide the cluster with additional network resources.

Figure 1-3 shows an example of a 6 + 2 (*N* + *I*) failover configuration with six active nodes and two passive nodes. Table 1-7 provides an *N* + *I* failover configuration matrix for Figure 1-3.

Figure 1-3. Example of an N+I Failover Configuration for an Eight-Node Cluster



.

Table 1-7. Example of an N+I Failover Configuration for an Eight-Node Cluster

Cluster Resource Group	Primary Node	AntiAffinityClassNamesValue
A	Node 1	AString
B	Node 2	AString
C	Node 3	AString
D	Node 4	AString
E	Node 5	AString
F	Node 6	AString

Configuring Group Affinity

On *N* + *I* (active/passive) failover clusters running Windows Storage Server 2003, Enterprise Edition, some resource groups may conflict with other groups if they are running on the same node. For example, running more than one Microsoft Exchange virtual server on the same node may generate application conflicts. Using Windows Storage Server 2003, Enterprise Edition, you can assign a public property (or attribute) to a dependency between groups to ensure that they failover to similar or separate nodes. This public property is called *group affinity*.

Group affinity uses the AntiAffinityClassNamesValues public property, which ensures that designated cluster resources are running on *separate nodes*, if possible.

For example, in Table 1-7, the AntiAffinityClassNamesValues string for cluster resource group A and group B are identical (AString), which indicates that these groups are assigned to run on separate cluster nodes, if possible. If node 1 fails, cluster resource group A will failover to the next backup node (node 7). If node 2 then fails, because their AntiAffinityClassNamesValues string value (AString) identifies group A and group B as conflicting groups, group B will skip node 7 and instead fail over to node 8.

To set the public property for the cluster groups shown in Table 1-7:

1. Open a command prompt.
2. Type the following:

```
cluster group "A" /prop AntiAffinityClassNames="AString"
```

3. Repeat step 2 for the remaining cluster groups.

Use the "Cluster Data Sheets" to configure group affinity in your *N* + *I* cluster configuration.

Failover Pair

Failover pair is a failover policy in which each application is allowed to failover between two specific nodes in a multinode cluster. The Possible Owners List in Cluster Administrator is used to determine which nodes will run the failover applications.

This solution is easy to plan and administer, and any applications that do not run well on the same server can easily be moved into separate failover groups. However, because failover pair uses one idle or passive cluster node for each cluster pair, this failover option can be more expensive.

Figure 1-4 shows an example of a failover pair configuration. Table 1-8 provides a failover configuration for the cluster shown in Figure 1-4.

Figure 1-4. Example of a Failover Pair Configuration

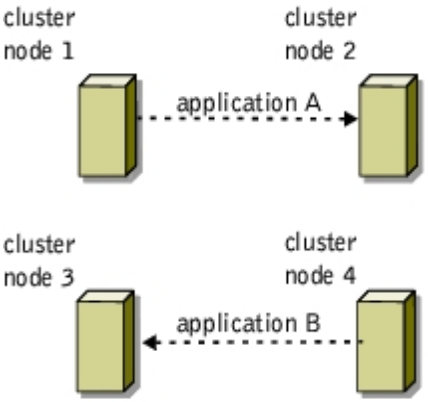


Table 1-8. Example of a Failover Pair Configuration for a Four-Node Cluster

Cluster Resource Group	Possible Owners List
App1	1,2
App2	3,4

Multiway Failover

Multiway failover is an active/active failover policy where running applications from a failed node migrate to multiple nodes in the cluster. This type of failover solution provides automatic failover and load-balancing between cluster nodes. However, you must ensure that the failover cluster nodes have ample resources available to handle the additional workload. Figure 1-5 shows an example of four-node multiway failover configuration.

Table 1-9 provides an example of a four-node multiway failover configuration for the cluster shown in Figure 1-5. For each cluster resource group, the failover order in the Preferred Owners list outlines the order that you want that resource group to failover. In this example, node 1 owns cluster resource groups A, B, and C. If node 1 fails, the cluster resource groups A, B, and C fail over to cluster nodes 2, 4, and 3, respectively. Configure the cluster resource groups similarly on cluster nodes 2, 3, and 4.

When implementing this type of failover policy, configure failback to avoid performance degradation.

Figure 1-5. Example of a Four-Node Multiway Failover Configuration

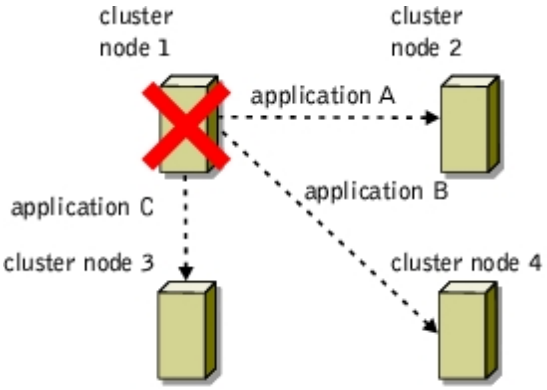


Table 1-9. Example of a Four-Node Multiway Failover Configuration

Cluster Resource Group	Failover Order in the Preferred Owners List
A	Node 1
B	Node 2
C	Node 3
D	Node 4
E	Node 5
F	Node 6

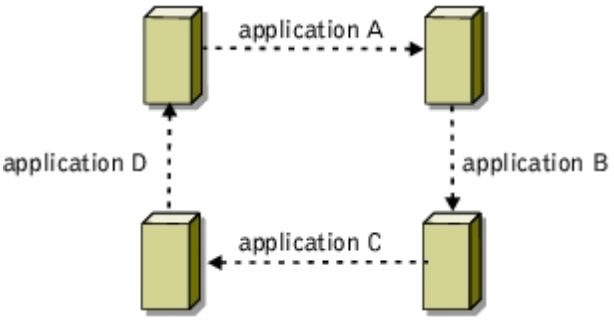
Failover Ring

Failover ring is an active/active failover policy where all running applications migrate from the failed node to the next preassigned cluster node in the Preferred Owners List in Cluster Administrator. If the failing node is the last node in the list, the failed node's resources failover to the first node in the list.

While this type of failover solution provides high resource availability to users, you must ensure that the cluster node next in line for failover has ample resources available to handle the additional workload of the failed node.

Figure 1-6 shows an example of a failover ring configuration.

Figure 1-6. Example of a Four-Node Failover Ring Configuration



Overview

Table 1-10 provides an overview of the failover policies implemented with Windows Storage Server 2003, Enterprise Edition.

Table 1-10. Windows Storage Server 2003, Enterprise Edition Failover Policies

Failover Policy	Description	Advantage	Disadvantage(s)
N + *I*	One or more servers provides backup for multiple servers in the cluster.	Highest resource availability.	• May not handle more than one backup server failure • May not fully utilize all of the servers
Failover pair	Applications can failover between the two nodes in the cluster.	Easy to plan the server capacity of each node.	Half of the server resources are idle.
Multiway	Running applications migrate to multiple nodes in the	Application load balancing.	Must ensure that the failover cluster nodes have ample resources available to handle the additional workload.

	cluster.		
Failover ring	Running applications migrate to the next preassigned cluster node.	Easy to scope server capacity for one server failure.	The cluster node next in line for failover may not have ample resources available to handle the additional workload of the failed node.

Minimum System Requirements

NAS Fibre Channel cluster configurations require the following hardware and software components:

- Cluster nodes
- Cluster storage
- Cluster interconnects (private network)
- Client network connections (public network)
- Operating system and storage management software

Cluster Nodes

Table 1-11 lists the hardware requirements for the cluster nodes.

Table 1-11. Cluster Node Requirements

Component	Minimum Requirement
Cluster nodes	Two supported PowerVault NAS systems running Windows Storage Server 2003, Enterprise Edition.
Processors	At least two processors for each cluster node. **NOTE:** Each cluster node must have identical processors.
RAM	At least 512 MB of RAM installed on each cluster node. **NOTE:** Each cluster node must have identical memory configurations.
HBAs	Two Fibre Channel HBAs for each cluster node. Dell recommends placing your Fibre Channel HBAs on separate PCI buses to improve availability and performance. See the *Platform Guide* for more information about supported NAS systems, specific HBA models supported by the system, and PCI slot configuration guidelines.
NICs	Minimum of two NICs: one NIC for the public network (client LAN connections) and another NIC for the private network (cluster interconnect). You must install identical NICs in each cluster node for the public and private networks.
RAID controller	The operating system volume is located in a RAID 1 configuration on an internal RAID controller on each NAS system.

Cluster Storage

Table 1-12 provides a list of supported storage systems and the configuration requirements for the cluster nodes and stand-alone systems connected to the storage systems. Table 1-13 provides hardware requirements for the processor enclosures (DPE and SPE).

Table 1-12. Cluster Storage Requirements

Hardware Components	Minimum Requirement
Supported storage systems	One Dell | EMC CX200 or CX400 with dual SPSs and at least five internal hard drives on the DPE. OR One Dell | EMC CX600 with dual SPSs and one DAE2 enclosure. At least five hard drives are required on the first DAE2 enclosure that is attached to the CX600 SPE.
Cluster nodes	The cluster nodes must be attached to a single storage system or to multiple storage systems through a SAN using redundant fabric.
Multiple clusters and stand-alone systems	Can share one or more supported storage systems using Access Logix, an optional software component that is available for your storage system. See "Storage Management Software" for information about the Access Logix software.

Table 1-13. Processor Enclosure Requirements

Processor Enclosure	Primary Storage	Secondary Storage	Standby Power Supplies (SPS)
CX200 DPE	Minimum of five internal hard drives	One DAE2 enclosure with a maximum of 15 hard drives	Two per processor enclosure
CX400 DPE	Minimum of five internal hard drives	Up to three DAE2 enclosures with a maximum of 15 hard drives	Two per processor enclosure
CX600 SPE	One DAE2-OS enclosure with a minimum of five hard drives **NOTE:** The CX600 SPE does not support internal hard drives.	Up to 15 DAE2 enclosures with a maximum of 15 hard drives	Two per processor enclosure
NOTE: The DAE2-OS is the first DAE2 enclosure that is connected to the CX600 and has the core software (FLARE) preinstalled on the first five hard drives.			

Client Network Connections (Public Network)

The cluster connections to the public network (for client access of cluster resources) requires one or more identical NICs supported by the system for each cluster node. Configure this network in a mixed mode (**All Communications**) to communicate the cluster heartbeat to the cluster nodes if the private network fails for any reason.

Cluster Interconnects (Private Network)

Table 1-14 provides the minimum requirements for the cluster interconnects (private network).

Table 1-14. Cluster Interconnects (Private Network) Requirements

Hardware Component	Minimum Requirement
NICs	Any NIC supported by the system for each cluster node. The private network NICs must be identical and supported by the system. **NOTE:** Dual-port Fast Ethernet NICs are not recommended for simultaneous cluster connections to the public and private networks. When configured as **All Communications**, the public network can provide redundancy for node-to-node traffic in the case of a failure in the private network segment.

Ethernet switch (optional)	One Ethernet switch for the private network (cluster interconnect).

Ethernet cables	Standard, crossover, or fiber optic Ethernet cables.
	• Standard Ethernet cable (*not included*)
	  ◦ Connects two copper Gigabit Ethernet (1000Base-T) NICs
	  ◦ Connects two Gigabit Ethernet or Fast Ethernet NICs to a switch
	• Crossover Ethernet cable (*not included*) connects two Fast Ethernet NICs
	• Fiber optic Ethernet cable connects Gigabit Ethernet NICs to a switch
	See Table 3-4 for the appropriate connection procedures.

Ethernet switch cabling (optional)	Additional Ethernet cables (not included) may be used to attach to an Ethernet switch for the public network (client connections) and private network (cluster interconnect)

Other Documents You May Need

The *System Information Guide* provides important safety and regulatory information. Warranty information may be included within this document or as a separate document.

- The *Platform Guide* provides information about the platforms that support the NAS cluster configuration.
- The *Rack Installation Guide* included with your rack solution describes how to install your system into a rack.
- The *Setting Up Your System* document provides an overview of initially setting up your system.
- The *Installation and Troubleshooting Guide* describes how to troubleshoot the system and install or replace system components.
- The HBA documentation provides installation instructions for the HBAs.
- Systems management software documentation describes the features, requirements, installation, and basic operation of the software.
- Operating system documentation describes how to install (if necessary), configure, and use the operating system software.
- Documentation for any components you purchased separately provides information to configure and install these options.
- The Dell PowerVault tape library documentation provides information for installing, troubleshooting, and upgrading the tape library.
- The RAID documentation provides information for installing and configuring a RAID controller card.
- The documentation that came with your storage system.
- The EMC PowerPath documentation that came with your HBA kit(s).
- Updates are sometimes included with the system to describe changes to the system, software, and/or documentation.

  NOTE: Always read the updates first because they often supersede information in other documents.

- Release notes or readme files may be included to provide last-minute updates to the system or documentation, or advanced technical reference material intended for experienced users or technicians.

Back to Contents Page

Back to Contents Page

Preparing Your Systems for Clustering

Dell™ PowerVault™ NAS Fibre Channel Cluster Systems Installation and Troubleshooting Guide

---

Before You Begin

Before you configure your storage system hardware and software for a cluster configuration:

1. Ensure that your site can handle the power requirements of the cluster equipment.

   Contact your sales representative for information about your region's power requirements.

   CAUTION: Only trained service technicians are authorized to remove and access any of the components inside the system. See your *System Information Guide* for complete information about safety precautions, working inside the computer, and protecting against electrostatic discharge.

2. Ensure that the following components are installed in your cluster hardware:
   - NICs
   - HBAs
   - SCSI hard drives
   - Fibre Channel hard drives
   - Any additional peripheral components
   - RAID controllers
3. Cable the system hardware for clustering.

   See "Cabling Your Cluster Hardware" for more information.
4. Configure the storage system(s) as described in your Dell | EMC storage system documentation.
5. If you are using hardware-based RAID for the internal SCSI hard drives, configure the hard drives using the controller's BIOS utility.

Installation Overview

This section provides installation overview procedures for configuring your cluster running the Microsoft® Windows® Storage Server 2003 operating system.

1. Ensure that your cluster meets the requirements as described in "Before You Begin."
2. Select a domain model that is appropriate for your corporate network and operating system.

See "Selecting a Domain Model" for more information.

3. Reserve static IP addresses for your cluster resources and components.

   The resources and components include:

   - Public network
   - Private network
   - MSCS software
   - Cluster-aware applications running on the cluster

   You will use these IP addresses when you install or configure MSCS. See "Assigning Static IP Addresses to Your Cluster Resources and Components" for more information.

4. Install or update the HBA drivers.

   The HBA drivers allow your cluster nodes to communicate with the shared storage systems.

   See "Installing the Fibre Channel HBA Drivers" for more information.

   See the *Platform Guide* for the specific HBA driver versions required for your Windows operating system.

5. Install and configure the storage management software for your storage systems.

   These software applications help you manage the communications between the cluster nodes and the storage systems.

   See the documentation included with your Dell | EMC storage system or at the Dell Support website located at **support.dell.com** for more information.

6. Configure the hard drives on the shared storage system(s).

   After you bind the LUNs using EMC® ControlCenter™ Navisphere Manager™ software, you can partition and format the hard drives and assign drive letters.

   See "Configuring the Hard Drives on the Shared Storage System(s)" for more information.

7. Configure the MSCS software.

   The MSCS software is the clustering component of the Windows operating system that provides the failover capabilities for the cluster.

   See "Installing and Configuring MSCS" for more information.

8. Verify cluster functionality. Ensure that:

   - Your cluster components are communicating properly with each other.
   - MSCS is started.

   See "Verifying Cluster Functionality" for more information.

9. Verify cluster resource availability.

   Use Cluster Administrator to check the running state of each resource group.

   See "Verifying Cluster Resource Availability" for more information.

The following sections provide detailed information for each step in the "Installation Overview" that is specific to your Windows operating system.

Selecting a Domain Model

On a cluster running the Windows operating system, the cluster nodes must belong to a common domain or directory model. The following membership configurations are supported:

- All cluster nodes are member servers in a Windows Server 2003 Active Directory domain.
- All cluster nodes are member servers in a Windows Server 2000 Active Directory domain.
- One cluster node is a domain controller in an Active Directory, the remaining cluster nodes are member servers, and there are no other client members.

Configuring the Cluster Nodes as Domain Controllers

If a cluster node is not configured as a domain controller and the node cannot contact a domain controller, the cluster node will not be able to authenticate client requests.

If a cluster node is configured as a domain controller, client access to its cluster resources can continue if the cluster node cannot contact other domain controller(s). However, domain controller functions can cause additional overhead, such as log on, authentication, and replication traffic on the cluster nodes.

Configuring Windows Networking

You must configure the public and private networks in each node before you install the Cluster Service. The following sections introduce you to some principles and procedures necessary to the networking prerequisites.

Assigning Static IP Addresses to Your Cluster Resources and Components

A static IP address is an Internet address that a network administrator assigns exclusively to a system or a resource. The address assignment remains in effect until it is changed by the network administrator.

The IP address assignments for the public LAN segments will depend on the configuration of your environment. If the IP assignments are set up correctly, all of the NIC resources will respond to ping commands and appear online before and after you install MSCS. If the IP assignments are not set up correctly, the cluster nodes may not be able to communicate with the domain. See "Troubleshooting" for more information.

NAS cluster configurations running the Windows operating system require static IP addresses assigned to hardware and software applications in your cluster, as listed in Table 2-1.

Table 2-1. Applications and Hardware Requiring IP Address Assignments

Application / Hardware	Description
Cluster IP address	The cluster IP address is used for cluster management and must correspond to the cluster name. Because each server has at least two NICs, the minimum number of static IP addresses required for a two-node cluster configuration is five (one for each NIC and one for the cluster). Additional static IP addresses are required when MSCS is configured with application programs that require IP addresses, such as file sharing.
Cluster node NICs	The NICs are used to connect to the public and private networks. For cluster operation, two NICs are required: one NIC for the public network (LAN/WAN) and another NIC for the private network (sharing heartbeat information between the cluster nodes). See "Cabling Your Cluster Hardware" for more information about cluster interconnect options. **NOTE:** To ensure cluster operations during a DHCP server failure, Dell recommends using static IP addresses for your cluster.

Configuring IP Addresses for the Private Network

Dell recommends using the static IP address assignments for the NICs used for the private network (cluster interconnect). The IP addresses in Table 2-2 are given as examples only.

Table 2-2. Examples of IP Address Assignments

Usage	Cluster Node 1	Cluster Node 2
Public network static IP address (for client and domain controller communications)	192.168.1.101	192.168.1.102
Public network subnet mask	255.255.255.0	255.255.255.0
Private network static IP address cluster interconnect (for node-to-node communications)	10.0.0.1	10.0.0.2
Private network subnet mask	255.255.255.0	255.255.255.0
Default gateway	192.168.1.1	192.168.1.1
WINS servers	Primary 192.168.1.11 Secondary 192.168.1.12	Primary 192.168.1.11 Secondary 192.168.1.12
DNS servers	Primary 192.168.1.21 Secondary 192.168.1.22	Primary 192.168.1.21 Secondary 192.168.1.22

NOTE: Dell recommends that you do not configure a default gateway, WINS, and DNS on your private network.

If multiple cluster interconnect NICs are connected to a network switch, ensure that all of the private network NICs have a unique address. You can continue the IP address scheme in Table 2-2 with 10.0.0.3, 10.0.0.4, and so on for the private network NICs of the other clusters connected to the same switch.

Additional fault tolerance for the LAN segments can be achieved by using NICs that support adapter teaming or by having multiple LAN segments. To avoid communication problems in the private network, Dell recommends that you *do not* use dual-port NICs for the cluster interconnect.

Creating Separate Subnets for the Public and Private Networks

The public and private network NICs installed in the same cluster node must reside on separate IP subnetworks. Therefore, the private network used to exchange heartbeat information between the cluster nodes must have a separate IP subnet or a different network ID than the public network, which is used for client connections.

Setting the Network Interface Binding Order

1. Click the **Start** button, select **Control Panel**, and then double-click **Network Connections**.
2. From the **Advanced** menu, click **Advanced Settings**.

   The **Advanced Settings** window appears.
3. In the **Connections** box, click the connection you need to modify.
4. Click the up-arrow or down-arrow to adjust the connection priority.
5. Click **OK**.

Using Dual-Port NICs for the Private Network

You can configure your cluster to use the public network as a failover for private network communications. However, dual-port NICs are not supported in the private network.

Verifying Cluster Network Communications

To ensure proper cluster operations, the cluster nodes must be able to communicate with each other through the private network (cluster interconnect). This communication involves the exchange of heartbeat messages, whereby the two cluster nodes inquire about each other's status, or "health," and acknowledge each inquiry.

To verify network communications between the cluster nodes:

1. Open a command prompt on each cluster node.
2. At the prompt, type:

   ```
   ipconfig /all
   ```

3. Press <Enter>.

   All known IP addresses for each local server appear on the screen.

4. Issue the **ping** command from each remote system.

   Ensure that each local server responds to the **ping** command.

Installing the Fibre Channel HBAs

For dual HBA configurations, Dell recommends installing Fibre Channel HBAs on separate PCI buses. Placing the adapters on separate buses improves availability and performance.

See the *Platform Guide* for more information about your system's PCI bus configuration and supported HBAs.

Installing the Fibre Channel HBA Drivers

See the EMC documentation that is included with your HBA kit for more information.

See the Emulex support website located at **www.emulex.com** or the Dell Support website at **support.dell.com** for information about installing and configuring Emulex HBAs and EMC approved drivers.

See the QLogic support website at **www.qlogic.com** for information about installing and configuring QLogic HBAs and EMC approved drivers.

See the *Platform Guide* for information about supported HBA controllers and drivers.

Installing and Configuring the Shared Storage System

See "Storage System" for a list of supported Dell | EMC storage systems.

To install and configure the Dell | EMC storage system in your cluster:

1. Update the core software on your storage system with the EMC Access Logix™ software (optional) and install any additional software options, including EMC SnapView™, EMC MirrorView™, and SAN Copy™. See your Navisphere® documentation for more information.
2. Install the EMC Navisphere Agent and EMC PowerPath® software on each cluster nodes.

See your Navisphere documentation for more information.

3. Update the storage system configuration settings using Navisphere Manager.

   See "Updating the Storage System Configuration Settings" for more information.

The following subsections provide an overview of the storage management software and procedures for connecting the host systems to the storage systems.

Access Logix

Fibre Channel topologies allow multiple clusters and stand-alone systems to share a single storage system. However, if you cannot control access to the shared storage system, you can corrupt your data. To share your Dell | EMC storage system with multiple heterogeneous host systems and restrict access to the shared storage system, you can install and configure the Access Logix software.

Access Logix is an optional software component that restricts LUN access to specific host systems. Using Access Logix software, you can:

- Connect multiple cluster nodes and stand-alone systems to a storage system.
- Create storage groups to simplify LUN management.
- Restrict LUN access to preassigned storage groups for data security.

Access Logix is installed on each storage system in the cluster. Navisphere Manager is installed on one or more storage systems, or one or more hosts attached to the storage system(s).

The storage systems are managed through a *management station*—a local or remote system that communicates with Navisphere Manager and connects to the storage system through an IP address. Using Navisphere Manager, you can secure your storage data by partitioning your storage system arrays into LUNs, assign the LUNs to one or more storage groups, and then restrict access to the LUNs by assigning the storage groups to the appropriate host systems.

Access Logix is required when the storage system is connected to two or more clusters, two or more nonclustered systems, or a combination of both clustered and nonclustered systems. Access Logix is not required for stand-alone systems or single cluster configurations. Access Logix is also required as a prerequisite to using SnapView, MirrorView or SAN Copy. Access Logix is installed by updating the core software (firmware) on your storage system.

Table 2-3 provides a list of cluster and host system configurations and their Access Logix requirement.

Table 2-3. Access Logix Software Requirements

Cluster Configuration	Access Logix Required
Single host or One cluster	No
Two or more clusters or Two or more nonclustered hosts or Any combination of clusters and nonclustered hosts	Yes

Navisphere Agent

Navisphere Agent is installed on the host system and performs the following tasks:

- Registers each host with the storage system
- Communicates configuration information from the host to the storage system

Access Control

Access Control is a feature of Access Logix that connects the host system to the storage system. Enabling **Access Control** prevents all host systems from accessing any data on the storage system until they are given explicit access to a LUN through a storage group. By installing Access Logix on your storage system(s) and enabling **Access Control**, you can prevent the host systems from taking ownership of all LUNs on the storage system and prevent unauthorized access to sensitive information.

Access Control is enabled using Navisphere Manager. After you update the core software on your storage system(s) with Access Logix and connect to the storage system from a management station, **Access Control** appears in the **Storage System Properties** window of Navisphere Manager. After you enable **Access Control** in Navisphere Manager, you are using Access Logix.

See "Storage Management Software" for additional information on Access Logix and Navisphere Manager.

After you enable **Access Control**, the host system can only read and write to specific LUNs on the storage system. This organized group of LUNs and hosts is called a *storage group*.

Storage Groups

Storage groups are a collection of one or more LUNs that are assigned to one or more host systems. Managed by Navisphere Manager, storage groups provide an organized method of assigning multiple LUNs to a host system. After you create LUNs on your storage system, you can assign the LUNs to a storage group in Navisphere Manager and then assign the storage group to a specific host. Because the host can only access its assigned storage group, it cannot access any LUNs assigned to other host systems, thereby protecting your data from unauthorized access.

To create the storage groups on your host systems, you must use Navisphere Manager and enable **Access Control** in the storage system.

NOTE: A host system can access only one storage group per storage system.

Table 2-4 describes the properties in the storage groups.

Table 2-4. Storage Group Properties

Property	Description
Unique ID	A unique identifier that is automatically assigned to the storage group that cannot be changed.
Storage group name	The name of the storage group. The default storage group name is formatted as *Storage Group n*, where *n* equals the existing number of storage groups plus one.
Sharing	Lists whether the storage system is shared between multiple hosts in a cluster or dedicated to one nonclustered system. You can set the sharing state to one of the following: • **Shareable** — If the storage group is for a cluster • **Dedicated** — If the storage group is for a nonclustered server
Connected hosts	Lists the host systems connected to the storage group. Each host entry contains the following fields: • **Name** — Name of the host system • **IP address** — IP address of the host system • **OS** — Operating system that is running on the host system When you connect a host system to a storage group, the host system disconnects the existing storage group through each host system HBA port (or *initiator*) and reconnects to the new storage group. If the host system has dual HBAs and each HBA is connected to one storage system SP, the host system connects to the storage system using redundant paths. **NOTE:** In a clustered environment, all nodes of a cluster must be connected to the same storage group.
LUNs in storage group	Lists the LUNs in the storage group. Each LUN entry contains the following fields: • **Identifier** — LUN icon representing the LUN

	• **Name** — Name of the LUN • **Capacity** — Amount of allocated storage space on the LUN
Used host connection paths	An additional storage group feature that performs the following tasks: • Lists all of the paths from the host server to the storage group • Displays whether the path is enabled or disabled Each path contains the following fields: ◦ **HBA** — Device name of the HBA in the host system ◦ **HBA Port** — Unique ID for the HBA port connected to the storage system ◦ **SP Port** — Unique ID for the storage processor port connected to the HBA port ◦ **SP ID** — ID of the storage processor

Navisphere Manager

Navisphere Manager provides centralized storage management and configuration from a single management console. Using a GUI, Navisphere Manager allows you to configure and manage the disks and components in one or more shared storage systems.

You can access Navisphere Manager through a Web browser. Using Navisphere Manager, you can manage a Dell | EMC storage system locally on the same LAN or through an Internet connection. Navisphere components (Navisphere Manager UI and Storage Management Server) are installed on a Dell | EMC storage system. You can access Navisphere Manager by opening a browser and entering the IP address of the storage system's SP. Navisphere Manager downloads to your system and runs in the web browser.

Optionally, you can run Navisphere Management Server for Windows. This software component installs on a host system connected to a Dell | EMC storage system, allowing you to run Navisphere Storage Management Server on the host system.

Using Navisphere Manager, you can:

- Create storage groups for your host systems
- Create, bind, and unbind LUNs
- Change configuration settings
- Monitor storage systems

See "Storage Management Software" for more information on Navisphere Manager.

EMC PowerPath

PowerPath automatically reroutes Fibre Channel I/O traffic between the host system and a Dell | EMC CX-series storage system to any available path if a primary path fails for any reason. Additionally, PowerPath provides multiple path load balancing, allowing you to balance the I/O traffic across multiple SP ports on a CX400 or CX600.

Updating the Storage System Configuration Settings

Some Dell | EMC storage systems may have shipped with preconfigured settings that may cause the cluster systems to malfunction in some environments. To optimize storage system performance, update the following processor enclosure settings as listed in Table 2-5 using a Web browser and Navisphere Manager.

Table 2-5. Storage System Configuration Settings

Configuration Setting	Description
Cache Page Size	The processor enclosure cache memory allocation size.
High/Low Watermark	The processor enclosure cache level setting—ranging between 0 percent and 100 percent—that determines when to start flushing the cache (high watermark) and when to stop flushing the cache (low watermark).
SPS Test Time	The day and hour (configured in Greenwich Mean Time [GMT]) when the processor enclosure tests the SPS batteries and

	verifies that the SPS monitoring system is functioning correctly.

Update the storage system configuration settings *in order* by performing the steps in the following subsections.

Updating the Cache Page Size and High/Low Watermark Settings

1. Open a Web browser and connect to the storage management server (usually the IP address of the storage processor on a Dell | EMC storage system).
2. Click the **Storage** tab.
3. Right-click the storage device you want to manage and select **Properties**.
4. In the **System Properties** window, click the **Cache** tab.
5. In the **Configuration** box, click the **Cache Page Size** down-arrow and select **8K**.
6. Go to "Updating the High/Low Watermark Settings."

Updating the High/Low Watermark Settings

1. In the **Configuration** box, verify the following:
    a. **Low Watermark** value is set to **60**.
    b. **High Watermark** value is set to **80**.
2. Ensure that the **Enable Watermark Processing** check box is selected.
3. Click **Apply** to apply the new settings.
4. Click **OK**.
5. Go to "Updating the SPS Test Time Setting."

Updating the SPS Test Time Setting

1. In the **Storage** tab, click the plus sign (+) to the left of the storage system.
2. In the **Storage** tab, click the plus sign (+) to the left of **Physical**.
3. In the **Storage** tab, click the plus sign (+) to the left of **Enclosure SPE**.
4. In the Storage tab, click the plus sign (+) to the left of **Standby Power Supplies**.
5. Right-click **Enclosure xPE SPS A** and select **Set Test Time**.

   The **Battery Test Time Dialog** window appears.
6. From the **Test Day** drop-down menu, select the appropriate day to run the SPS battery test. The default setting is **Sunday**.

   NOTE: Because I/O performance may decrease during the SPS test, Dell recommends that you run the SPS battery test during a period of time when your storage system is inactive or when the storage system I/O is minimal.
7. From the **Test Time (GMT)** drop-down menu, select the appropriate time of day to run the SPS battery test.

   The default setting is **01:00 AM**. This value is displayed in GMT.
8. Click **OK**.
9. Close your Web browser.

See the Navisphere Manager online help for more information on updating the SPS test time settings.

Enabling Access Logix and Creating Storage Groups

The following subsection provides the required procedures for creating storage groups and connecting your storage systems to the host systems using the Access Logix software.

Enabling Access Logix and Creating Storage Groups for Systems Using Navisphere 6.*x*

1. Ensure that Navisphere Agent is started on all host systems.
    a. Click the **Start** button and select **Programs**→ **Administrative Tools**, and then select **Services**.
    b. In the **Services** window, verify the following:
        - In the **Name** column, **Navisphere Agent** appears.
        - In the **Status** column, **Navisphere Agent** is set to **Started**.
        - In the **Startup Type** column, **Navisphere Agent** is set to **Automatic**.
2. Open a Web browser.
3. Enter the IP address of the storage management server on your storage system and then press <Enter>.

    NOTE: The storage management server is usually one of the SPs on your storage system.

4. In the **Enterprise Storage** window, click the **Storage** tab.
5. Right-click the icon of your storage system.
6. In the drop-down menu, click **Properties**.

    The **Storage Systems Properties** window appears.
7. Click the **Storage Access** tab.
8. Select the **Access Control Enabled** check box.

    A dialog box appears, prompting you to enable **Access Control**.

NOTICE: Before enabling **Access Control**, ensure that no hosts are attempting to access the storage system. Enabling **Access Control** prevents all hosts from accessing any data until they are given explicit access to a LUN in the appropriate storage group. You must stop all I/O before enabling **Access Control**. Dell recommends shutting down all hosts connected to the storage system during this procedure or data loss may occur. After you enable the **Access Control** software, it cannot be disabled.

9. Click **Yes** to enable **Access Control**.
10. Click **OK**.
11. Right-click the icon of your storage system and select **Create Storage Group**.

    The **Create Storage Group** dialog box appears.
12. In the **Storage Group Name** field, enter a name for the storage group.
13. In the **Sharing State** drop-down menu, select one of the following radio buttons:
    - **Shareable** — If the storage group is for a cluster
    - **Dedicated** — If the storage group is for a nonclustered server
14. Click **Apply**.
15. Add new LUNs to the storage group.

a. Right-click the icon of your storage group and select **Properties**.

b. Click the **LUNs** tab.

c. In the **Available LUNs** window, click an available LUN.

d. Click the right-arrow button to move the selected LUN to the **Selected LUNs** pane.

e. Click **Apply**.

16. Add new hosts to the **Sharable** storage group.

a. In the **Storage Group Properties** dialog box, click the **Hosts** tab.

b. In the **Available Hosts** window pane, click the host system that you want to add to the storage group.

c. Using the right-arrow button, move the selected host to the **Hosts to be Connected** window pane.

d. Repeat step b and step c to add additional hosts.

e. Click **Apply**.

17. Click **OK** to exit the **Storage Group Properties** dialog box.

Configuring the Hard Drives on the Shared Storage System(s)

This section provides information for configuring the hard drives on the shared storage systems. The shared storage system hard drives must be configured before use. The following sections provide information on these configurations.

Configuring and Managing LUNs

Configuring and managing LUNs is accomplished using the Navisphere Manager utility. Before using Navisphere Manager, ensure that the Navisphere Agent service is started on your cluster nodes.

In some cases, the LUNs may have been bound when the system was shipped. It is still important, however, to install the management software and to verify that the desired LUN configuration exists.

You can manage your LUNs remotely using Navisphere Manager. A minimum of one LUN (RAID drive) is required for an active/passive configuration; at least two drives are required for an active/active configuration.

Dell recommends creating at least one LUN or virtual disk for each application. If multiple NTFS partitions are created on a single LUN or virtual disk, these partitions will not be able to fail over individually from node-to-node.

Using the Windows Dynamic Disks and Volumes

The Windows operating system does not support dynamic disks (upgraded disks) or volumes as shared cluster storage. If the shared cluster storage is configured as a dynamic disk, the Cluster Configuration wizard is not able to discover the disks, preventing the cluster and network clients from accessing the disks.

Configuring the RAID Level for the Shared Storage Subsystem

The hard drives in your shared storage subsystem must be configured into LUNs or virtual disks using Navisphere Manager. All LUNs or virtual disks, especially if they are used for the quorum resource, should be bound and incorporate the appropriate RAID level to ensure high availability.

See "Installing the Quorum Resource" for more information on the quorum resource.

NOTE: Dell recommends that you use a RAID level other than RAID 0 (which is commonly called striping). RAID 0 configurations provide very high performance, but do not provide the level of redundancy required for the quorum resource. See the documentation for your storage system for more information about setting up RAID levels for the system.

Naming and Formatting Drives on the Shared Storage System

When the LUNs have completed the binding process, assign drive letters to the LUNs and then format the drives as NTFS drives. Format the drives and assign volume labels from the first cluster node. When completed, the remaining nodes will see the file systems and volume labels.

NOTICE: Accessing the hard drives from multiple cluster nodes may corrupt the file system.

Assigning LUNs to Hosts

If you have **Access Control** enabled in Navisphere Manager, you must create storage groups and assign LUNs to the proper host systems.

Assigning Drive Letters

NOTICE: If the disk letters are manually assigned from the remaining node, the shared disks are simultaneously accessible from both nodes. To ensure file system integrity and prevent possible data loss before you install the MSCS software, prevent any I/O activity to the shared drives by performing this procedure on one node at a time, and ensure that all other nodes are shut down.

The number of drive letters required by individual servers in a cluster may vary. Dell recommends that the shared drives be named in reverse alphabetical order beginning with the letter *z*.

To assign drive letters and format the disks on the shared storage system:

1. With the other node shut down, open Disk Management on node 1.
2. Allow Windows to enter a signature on all new physical or logical drives.

   NOTE: Do not create dynamic disks on your hard drives.
3. Locate the icon for the first unnamed, unformatted drive on the shared storage system.
4. Right-click the icon and select **Create** from the submenu.

   If the unformatted drives are not visible, verify the following:

   - The HBA driver is installed.
   - The storage system is properly cabled to the servers.
   - The LUNs and hosts are assigned through a storage group (if **Access Control** is enabled).
5. In the dialog box, create a partition the size of the entire drive (the default) and then click **OK**.

   NOTE: The MSCS software allows only one node to access a logical drive at a time. If a logical drive is partitioned into multiple disks, only one node is able to access all the partitions for that logical drive. If each node needs to access a separate disk, two or more logical drives must be present in the storage system.
6. Click **Yes** to confirm the partition.
7. With the mouse pointer on the same icon, right-click and select **Change Drive Letter and Path** from the submenu.
8. Assign a drive letter to an NTFS volume.

   To assign a drive letter to an NTFS volume:

   a. Click **Edit** and select the letter you want to assign to the drive (for example, z).

   b. Click **OK**.
9. Click **Yes** to confirm the changes.

10. Right-click the drive icon again and select **Format** from the submenu.
11. Under **Volume Label**, enter a descriptive name for the new volume; for example, `Disk_Z` or `Email_Data`.
12. In the dialog box, change the file system to **NTFS**, select **Quick Format**, and click **Start**.

    NOTE: The NTFS file system is required for shared-disk resources under MSCS.

13. Click **OK** at the warning.
14. Click **OK** to acknowledge that the format is complete.
15. Click **Close** to close the dialog box.
16. Repeat step 3 through step 15 for each remaining drive.
17. Close Disk Management.
18. Shut down node 1.
19. Perform the following steps on the remaining node:
    a. Turn on the node.
    b. Open Disk Management.
    c. Assign the drive letters to the drives.
    d. Reassign the drive letter, if necessary.

       To reassign the drive letter, repeat step 7 through step 9.
    e. Power down the node.

Configuring Hard Drive Letters When Using Multiple Shared Storage Systems

Before installing MSCS, ensure that both nodes have the same view of the shared storage systems. Because each node has access to hard drives that are in a common storage array, each node must have identical drive letters assigned to each hard drive.

See "Assigning Drive Letters" for more information.

NOTE: Drive letters A through D are reserved for the local system.

To ensure that hard drive letter assignments are identical:

1. Ensure that your cables are attached to the shared storage devices in the proper sequence.

   You can view all of the storage devices using Windows Storage Server 2003 Disk Management or Dell OpenManage™ Array Manager.

   Windows 2000 Disk Management displays all of the accessible disks from the first HBA, followed by all of the accessible disks from the second HBA. The disks attached to a port with a lower port number on the Fibre Channel switch are displayed first, followed by those with a higher port number.
2. To maintain proper drive letter assignments, ensure the first HBA detected by each node is connected to the first switch or SP-A and the second detected HBA is connected to the second switch or SP-B.

   See "Cabling Your Power Supplies" in "Cabling Your Cluster Hardware" for the location of SP-A and SP-B on the CX-series storage systems.
3. Go to "Formatting and Assigning Drive Letters and Volume Labels to the Disks."

Formatting and Assigning Drive Letters and Volume Labels to the Disks

1. Shut down all the cluster nodes except node 1.
2. Format the disks, assign the drive letters and volume labels on node 1 by using the Windows Disk Management utility or Array Manager.

   For example, create volumes labeled "Volume Y" for disk Y and "Volume Z" for disk Z.
3. Shut down node 1 and perform the following steps on the remaining node:
   a. Turn on the node.
   b. Open Disk Management.
   c. Assign the drive letters for the drives.
   d. Reassign the drive letter, if necessary.

      To reassign the drive letter:

      - With the mouse pointer on the same icon, right-click and select **Change Drive Letter and Path** from the submenu.
      - Click **Edit**, select the letter you want to assign the drive (for example, "z"), and then click **OK**.
      - Click **Yes** to confirm the changes.

   e. Power down the node.

   If the cables are connected properly, the drive order is the same as on each node, and the drive letter assignments of all the cluster nodes follow the same order as on node 1. The volume labels can also be used to double-check the drive order by ensuring that the disk with volume label "Volume Z" is assigned to drive letter Z and so on for each disk on each node. Assign drive letters on each of the shared disks, even if the disk displays the drive letter correctly.

See your EMC documentation located on the Dell Support website at **support.dell.com** or the EMC support site located at **www.emc.com** for more information on the Navisphere Manager software.

Updating a Dell | EMC Storage System for Clustering

If you are updating an existing Dell | EMC storage system to meet the cluster requirements for the shared storage subsystem, you may need to install additional Fibre Channel disk drives in the shared storage system. The size and number of drives you add depend on the RAID level you want to use and the number of Fibre Channel disk drives currently in your system.

See your storage system's documentation for information on installing Fibre Channel disk drives in your storage system.

You may also need to upgrade the core software version that is running on the storage system. See the *Platform Guide* for specific version requirements.

Installing and Configuring MSCS

MSCS is an integrated service in Windows Storage Server 2003, Enterprise Edition. MSCS performs the basic cluster functionality, which includes membership, communication, and failover management. When MSCS is installed properly, the service starts on each node and responds automatically in the event that one of the nodes fails or goes offline. To provide application failover for the cluster, the MSCS software must be installed on each cluster node. See "Using MSCS" for more information.

The cluster setup files are automatically installed on the system disk. To create a new cluster:

1. Click the **Start** button, select **Programs**→ **Administrative Tools**→ **Cluster Administrator**.
2. From the **File** menu, select **Open Connection**.
3. In the **Action** box of the **Open Connection to Cluster**, select **Create new cluster**.

The **New Server Cluster Wizard** window appears.

4. Click **Next** to continue.
5. Follow the procedures in the wizard, and then click **Finish**.
6. Add the additional node to the cluster:
    a. Turn on the remaining node.
    b. Click the **Start** button, select **Programs→ Administrative Tools,** and then double-click **Cluster Administrator**.
    c. From the **File** menu, select **Open Connection**.
    d. In the **Action** box of the **Open Connection to Cluster**, select **Add nodes to cluster**.
    e. In the **Cluster or server name** box, type the name of the cluster or click **Browse** to select an available cluster from the list, and then click **OK**.

    The **Add Nodes Wizard** window appears.

    If the Add Nodes Wizard *does not* generate a cluster feasibility error, go to step f.

    If the Add Nodes Wizard generates a cluster feasibility error, go to "Adding Cluster Nodes Using the Advanced Configuration Option."
    f. Click **Next** to continue.
    g. Follow the procedures in the wizard and click **Finish**.

Adding Cluster Nodes Using the Advanced Configuration Option

If you are adding additional nodes to the cluster using the Add Nodes wizard and the nodes are not configured with identical internal storage devices, the wizard may generate one or more errors while checking cluster feasibility in the **Analyzing Configuration** menu. If this situation occurs, select **Advanced Configuration Option** in the Add Nodes wizard to add the nodes to the cluster**.**

To add the nodes using the **Advanced Configuration Option**:

1. From the **File** menu in Cluster Administrator, select **Open Connection**.
2. In the **Action** box of the **Open Connection to Cluster**, select **Add nodes to cluster**. and then click **OK**.

    The **Add Nodes Wizard** window appears.
3. Click **Next**.
4. In the **Select Computers** menu, click **Browse**.
5. In the **Enter the object names to select (examples)**, type the names of one to seven systems to add to the cluster, with each system name separated by a semicolon.
6. Click **Check Names**.

    The **Add Nodes Wizard** verifies and underlines each valid system name.
7. Click **OK**.
8. In the **Select Computer**s menu, click **Add**.
9. In the **Advanced Configuration Options** window, click **Advanced (minimum) configuration**, and then click **OK**.
10. In the **Add Nodes** window, click **Next**.
11. In the **Analyzing Configuration** menu, Cluster Administrator analyzes the cluster configuration.

If Cluster Administrator discovers a problem with the cluster configuration, a warning icon appears in the **Checking cluster feasibility** window. Click the plus ("+") sign to review any warnings, if needed.

12. Click **Next** to continue.
13. In the Password field of the **Cluster Service Account** menu, type the password for the account used to run the Cluster Service, and click **Next**.

    The **Proposed Cluster Configuration** menu appears with a summary with the configuration settings for your cluster.
14. Click **Next** to continue.

    The new systems (hosts) are added to the cluster. When completed, **Tasks completed** appears in the **Adding Nodes to the Cluster** menu.

    NOTE: This process may take several minutes to complete.
15. Click **Next** to continue.
16. In the **Completing the Add Nodes Wizard**, click **Finish**.

Verifying Cluster Readiness

To ensure that your server and storage systems are ready for MSCS installation, ensure that these systems are functioning correctly and verify the following:

- All cluster servers are able to log on to the same domain.
- The shared disks are partitioned and formatted, and the same drive letters that reference logical drives on the shared storage system are used on each node.

All IP addresses and network names for each cluster node are communicating with each other and the public network

Installing Applications in the Cluster Group

The Cluster Group contains a network name and IP address resource, which is used to manage the cluster. Because the Cluster Group is dedicated to cluster management and for best cluster performance, Dell recommends that you do not install applications in this group.

Installing the Quorum Resource

When you install a Windows Storage Server 2003 cluster, the installation wizard automatically selects an NTFS disk as the quorum resource for you, which you can modify later. When you complete the procedures in the wizard, you can select another disk for the quorum using **Cluster Administrator**. To prevent quorum resource corruption, Dell and Microsoft recommend that you do not place data on the disk.

Creating a LUN for the Quorum Resource

Dell recommends creating a separate LUN—approximately 1 GB in size—for the quorum resource.

When you create the LUN for the quorum resource:

- Format the LUN with NTFS.
- Use the LUN exclusively for your quorum logs.
- Do not store any application data or user data on the quorum resource.
- To easily identify the quorum resource, Dell recommends that you assign the drive letter "Q" to the quorum resource.

NOTE: The **Majority Node Set Quorum** types for Windows Storage Server 2003 are not supported.

Preventing Quorum Resource Failure

Since the quorum resource plays a crucial role in cluster operation, losing a quorum resource causes the entire cluster to fail. To prevent cluster failure, configure the quorum resource on a RAID volume in the shared storage system.

NOTICE: Dell recommends that you use a RAID level other than RAID 0, which is commonly called striping. RAID 0 configurations provide very high performance, but they do not provide the level of redundancy required for the quorum resource.

Configuring Cluster Networks Running Windows Storage Server 2003

When you install and configure a cluster running Windows Storage Server 2003, the software installation wizard automatically assigns and configures the public and private networks for your cluster. You can rename a network, allow or disallow the cluster to use a particular network, or modify the network role using **Cluster Administrator**. Dell recommends that you configure at least one network for the cluster interconnect (private network) and one network for all communications.

Verifying MSCS Operation

After you install MSCS, verify that the service is operating properly.

If you selected **Cluster Service** when you installed the operating system, see "Obtaining More Information."

If you did not select **Cluster Service** when you installed the operating system:

1. Click the **Start** button and select **Programs**→ **Administrative Tools**, and then select **Services**.
2. In the **Services** window, verify the following:
    - In the **Name** column, **Cluster Service** appears.
    - In the **Status** column, Cluster Service is set to **Started**.
    - In the **Startup Type** column, Cluster Service is set to **Automatic**.

Obtaining More Information

See Microsoft's online help for configuring the Cluster Service.

See "Using MSCS" for more information on the Cluster Service.

Verifying Cluster Functionality

To verify cluster functionality, monitor the cluster network communications to ensure that your cluster components are communicating properly with each other. Also, verify that MSCS is running on the cluster nodes.

Verifying Cluster Resource Availability

In the context of clustering, a resource is a basic unit of failover management. Application programs are made up of resources that are grouped together for recovery purposes. All recovery groups, and therefore the resources that comprise the recovery groups, must be online (or in a ready state) for the cluster to function properly.

To verify that the cluster resources are online:

1. Start **Cluster Administrator** on the monitoring node.
2. Click the **Start** button and select **Programs**→ **Administrative Tools (Common)**→ **Cluster Administrator**.
3. Open a connection to the cluster and observe the running state of each resource group. If a group has failed, one or more of its resources might be offline.

Troubleshooting Failed Resources

Troubleshooting the failed resources is beyond the scope of this document, but examining the properties of each resource and ensuring that the specified parameters are correct are the first two steps in this process. In general, if a resource is offline, it can be brought online by right-clicking the resource and selecting **Bring Online** from the pull-down menu.

See the documentation and online help for Windows Storage Server 2003, Enterprise Edition for information about troubleshooting resource failures.

Using Shadow Copies of Shared Folders

A shadow copy is a point-in-time copy of a shared file or folder. If you change a file on the active file system after making a shadow copy, the shadow copy contains the old version of the file. If an active file gets corrupted or deleted, you can restore the old version by copying the file from the latest shadow copy or restoring a directory or file.

NOTICE: Shadow copies are temporary backups of your data that typically reside on the same volume as your data. If the volume becomes damaged and you lose your data, the shadow copy is also lost. Do not use shadow copies to replace scheduled or regular backups. Table 2-4 provides a summary of shadow copies.

See the *Dell PowerVault 77xN NAS Systems Administrator's Guide* for more information on shadow copies.

You can create shadow copies of shared folders that are located on shared resources, such as a file server. When creating shadow copies of shared folders on a NAS cluster running the Windows Storage Server 2003, Enterprise Edition operating system, note the information listed in Table 2-6.

See the Microsoft Support website at **www.microsoft.com** for more information on shadow copies for shared folders.

Table 2-6. Creating Shadow Copies

Cluster Type/Task	Description	Action
Single quorum device cluster	A cluster with all nodes connected to a storage system with a physical disk resource.	Create and manage shadow copies on the physical disk resource. **NOTE:** The Volume Shadow Copy Service Task resource type can be used to manage shadow copies in a NAS cluster, but requires a dependency on the physical disk resource.
Scheduled tasks that generate volume shadow copies.	Creates a shadow copy of an entire volume.	Run the scheduled task on the same node that owns the volume. **NOTE:** The cluster resource that manages the scheduled task must be able to fail over with the physical disk resource that manages the storage volume.

Shadow Copy Considerations

When using shadow copies, note the following:

- To avoid disabling and re-enabling shadow copies, enable shadow copies after you create your NAS cluster.
- Enable shadow copies in a NAS cluster when user access is minimal—for example, during nonbusiness hours. When you enable shadow copy volumes, the shadow copy volumes and all dependent resources go offline for a brief period of time, which may impact client system access to user resources.

Managing Shadow Copies

You must use the Dell™ PowerVault™ NAS Manager to manage your shadow copies. Using Cluster Administrator or **cluster.exe** to manage shadow copies in a cluster is not supported.

See the *Dell PowerVault 77xN NAS Systems Administrator Guide* for more information on managing shadow copies using NAS Manager.

Enabling Shadow Copies on a Cluster Node

When you enable shadow copies on a cluster node (for example, by using the Configure Shadow Copy user interface through the Computer Management Microsoft Management Console [MMC]), the operating system automatically generates and configures a Volume Shadow Copy Service Task resource and a scheduled task for creating the shadow copy. You are not required to use Cluster Administrator or **cluster.exe** to create the resource. Additionally, the Configure Shadow Copy user interface automatically configures the required resource dependencies.

Table 2-7 provides the default properties of the scheduled task and Volume Shadow Copy Service Task resource.

Table 2-7. Default Properties for the Scheduled Task and Volume Shadow Copy Service Task Resource

Scheduled Task Property	Volume Shadow Copy Service Task Resource (cluster.exe Property)	Default Setting
Name of task	Name of resource (taskname)	ShadowCopyVolume{*VolumeGUID*}
Run	Command to run/Command parameters (ApplicationName/ApplicationParams)	%systemroot%\system32\vssadmin.exe Create Shadow /AutoRetry=5/For=\\[*drive_letter*]\ Volume{VolumeGUID}\
Creator	n/a	Cluster service
Start in	Start in	%systemroot%\system32\
Run as	n/a	Local System
Schedule	Schedule (TriggerArray)	The default settings used by Shadow Copies of Shared Folders

---

Back to Contents Page

Back to Contents Page

Cabling Your Cluster Hardware

Dell™ PowerVault™ NAS Fibre Channel Cluster Systems Installation and Troubleshooting Guide

NAS cluster configurations require cabling for the Fibre Channel storage systems, cluster interconnects, client network connections, and power connections. These systems and components are interconnected to provide four independent functions as listed in Table 3-1, each of which is described in more detail throughout this section.

Table 3-1. Cluster Cabling Components

Components	Description
Shared storage system	Connects the servers' redundant HBAs to the cluster's shared storage system. The server-to-storage connection can be achieved through a direct connection or through a SAN. See "Direct-Attached Cluster" and "SAN-Attached Cluster Configurations" for more information.
Cluster interconnect (private network)	Connects the cluster nodes to each other to exchange cluster information and health status, such as the cluster heartbeat and access to the quorum resource. This connection can be made by using an Ethernet NIC and cabling that is connected to each cluster node. **NOTE:** If more than two nodes are deployed, you will need an Ethernet switch. See the *Platform Guide* for a list of supported NICs for your configuration.
Network connection for public traffic (public network)	Provides a connection between each cluster node and the public network. This connection can be made using an Ethernet NIC and cabling that is connected to the public network. See the *Platform Guide* for a list of supported NICs for your configuration.
Power connection	Provides a connection between the power supplies in your system and the power source. By using power strips or PDUs and separate AC circuits, the cluster can fully utilize the redundant power supplies.

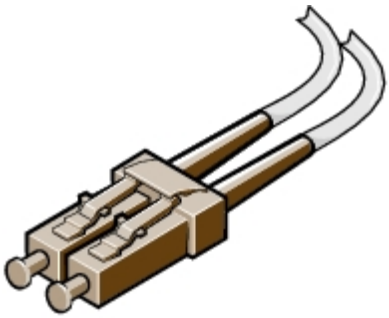
Fibre Channel Cable Connectors

NAS Fibre Channel cluster require two types of fiber optic connectors to connect the various Fibre Channel storage components to the cluster configuration:

- Duplex LC multimode fiber optic connector
- Duplex SC multimode fiber optic connector

The duplex LC multimode fiber optic connector (see Figure 3-1) is used to connect a Fibre Channel switch to a Fibre Channel HBA on a cluster node or to a storage system. This type of connection requires fiber optic cables with duplex LC multimode fibre optic connectors.

Figure 3-1. Duplex LC Multimode Fiber Optic Connector



The duplex SC multimode fiber optic connector (see Figure 3-2) is used to connect a Fibre Channel switch to a Dell™ PowerVault™ 132T or 136T tape library. This type of connection requires fiber optic cables with both duplex LC and SC multimode fiber optic connectors.

Figure 3-2. Duplex SC Multimode Fiber Optic Connector

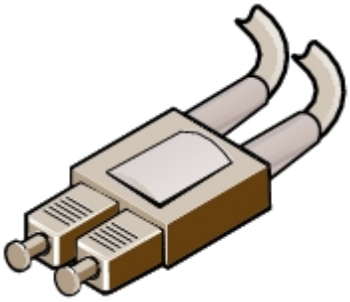


Table 3-2 provides the hardware component applications for the duplex multimode fiber optic connectors.

Table 3-2. Cable Connector Applications

Hardware Component	Duplex Multimode Fiber Optic Connector
PowerVault tape library	SC
Cluster node HBA	LC
CX200 DPE	LC
CX400 DPE	LC
CX600 SPE	LC
Fibre Channel switch	LC

The LC and SC connectors consist of two individual fiber optic connectors. Each connector is indexed and must be inserted and aligned properly in the GBIC or SFP module connector.

When using duplex multimode fiber optic connectors, keep the covers on the connectors until you are ready to insert the connectors into the system. The covers protect the cables and connectors and prevent foreign particles from entering and damaging the connector.

Cabling Your Public and Private Networks

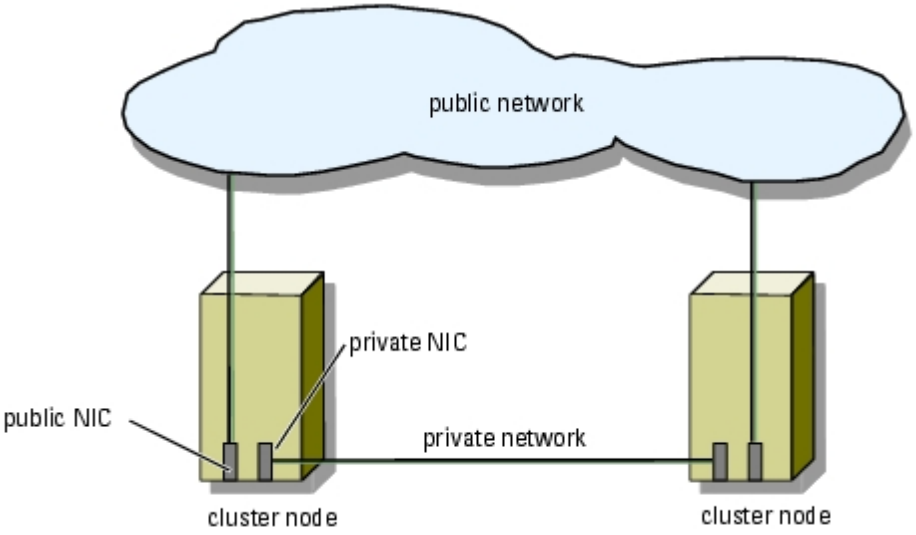
The NICs in the cluster nodes provide at least two network connections for each node. These connections are described in Table 3-3.

Table 3-3. Network Connections

Network Connection	Description
Public network	All connections to the client LAN. At least one public network must be configured for **Mixed mode** for private network failover.
Private network	A dedicated connection for sharing cluster health and status information between the cluster nodes. NICs connected to the LAN can also provide redundancy at the communications level in case the cluster interconnect fails. See your MSCS documentation for more information on private network redundancy.

Figure 3-3 shows an example of NIC cabling in which dedicated NICs in each node are connected to the public network and the remaining NICs are connected to each other (for the private network).

Figure 3-3. Example of Network Cabling Connection



Cabling Your Public Network

The public network connection (client network) to the cluster nodes is provided by a NIC that is installed in each node. Any NIC supported by the system running TCP/IP may be used to connect to the public network segments. Additional NICs may be installed to support additional separate public network segments or to provide redundancy for the public network.

Installing redundant NICs provides your cluster with a failover connection to the public network. If the primary NIC or a switch port fails, your cluster will be able to access the public network through the secondary NIC until the faulty NIC or switch port is repaired.

Using Dual-Port NICs for Your Private Network

You can configure your cluster to use the public network as a failover for private network communications. However, if dual-port NICs are used, they should not be used simultaneously to support both the public and private networks.

Cabling Your Private Network

The private network connection to the cluster nodes is provided by a second or subsequent NIC that is installed in each node. This network is used for intracluster communications.

Table 3-4 lists the required hardware components and connection method for three possible private network configurations.

Table 3-4. Private Network Hardware Components and Connections

Method	Hardware Components	Connection
Network switch	Fast Ethernet or Gigabit Ethernet NICs and switches	Connect *standard* Ethernet cables from the NICs in both cluster nodes to a Fast Ethernet or Gigabit Ethernet switch.
Point-to-Point Fast Ethernet	Fast Ethernet NICs	Connect a *crossover* Ethernet cable between the Fast Ethernet NICs in both cluster nodes.
Point-to-Point Gigabit Ethernet	Copper Gigabit Ethernet NICs	Connect a *standard* Ethernet cable between the Gigabit Ethernet NICs in both cluster nodes.

Cabling the Mouse, Keyboard, and Monitor

If you are installing a NAS cluster configuration in a Dell rack, your cluster will require a switch box to enable the mouse, keyboard, and monitor for your cluster nodes.

See your rack installation documentation included with your rack for instructions on cabling each cluster node's KVM to the mouse/keyboard/monitor switch box in the rack.

Cabling Your Power Supplies

Dell recommends the following guidelines to protect your cluster configuration from power-related failures:

- For cluster nodes with multiple power supply connections, plug each connection into a separate AC circuit.
- Connect UPSs to your server and storage systems (where applicable).
- For some environments, you should consider having backup generators and power from separate electrical stations. See your server and storage system documentation for more information on the power requirements for your cluster's components.

Connecting Standby Power Supplies in the Storage System

NAS Fibre Channel cluster configurations require dual SPSs in the Dell | EMC storage systems. The SPE receives power from redundant SPSs, which are connected to separate power sources. If an SPS power source fails, the redundant SPS and its power source provide backup power to the SPE. If both SPS power sources fail, the SPSs enable the SPE through the SPS sense connectors to flush the write cache to the internal vault drives (CX200 and CX400) or the vault drives in the first attached DAE2 enclosure (CX600). This process ensures that the cached data in the SPE is not lost during a power failure.

See your storage system documentation for additional information about the SPSs.

Figure 3-4, Figure 3-5, and Figure 3-6 show the power cable configuration for the supported storage systems.

NOTE: The CX200 and CX400 enclosure address must be set to "0" for the system to boot and function properly.

Figure 3-4. CX200 Power Cable Configuration

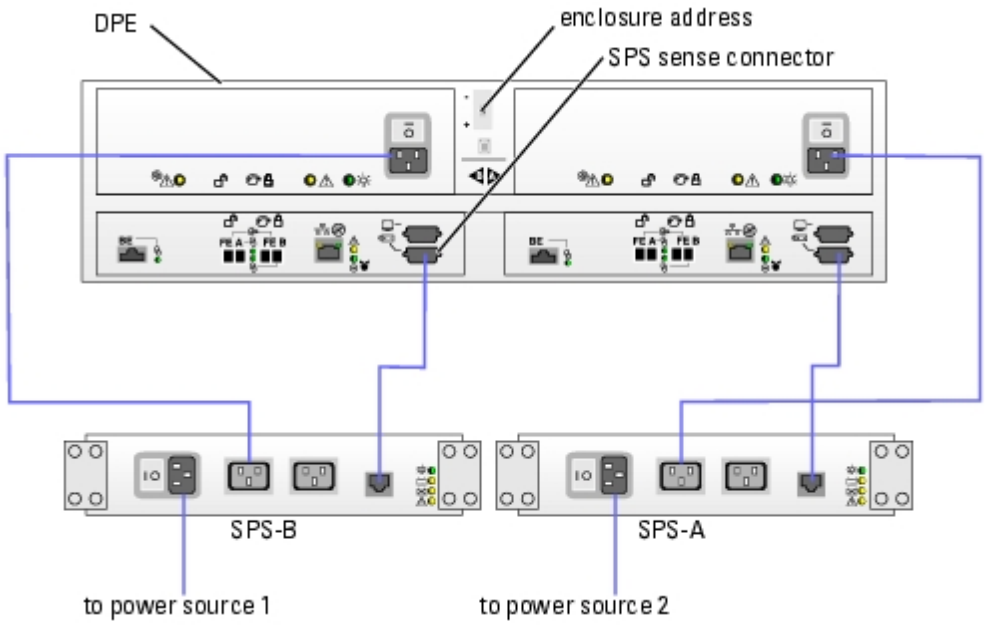


Figure 3-5. CX400 Power Cable Configuration

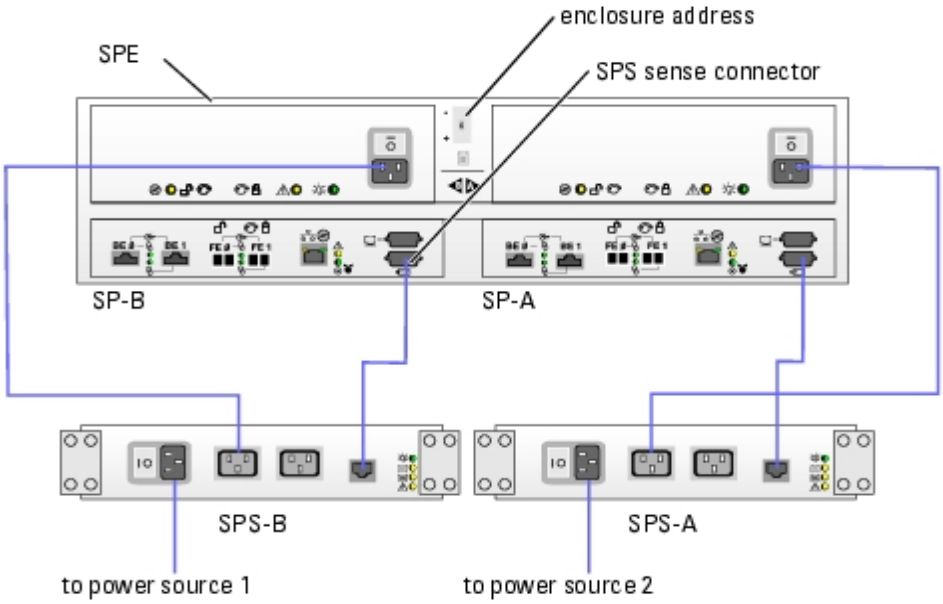
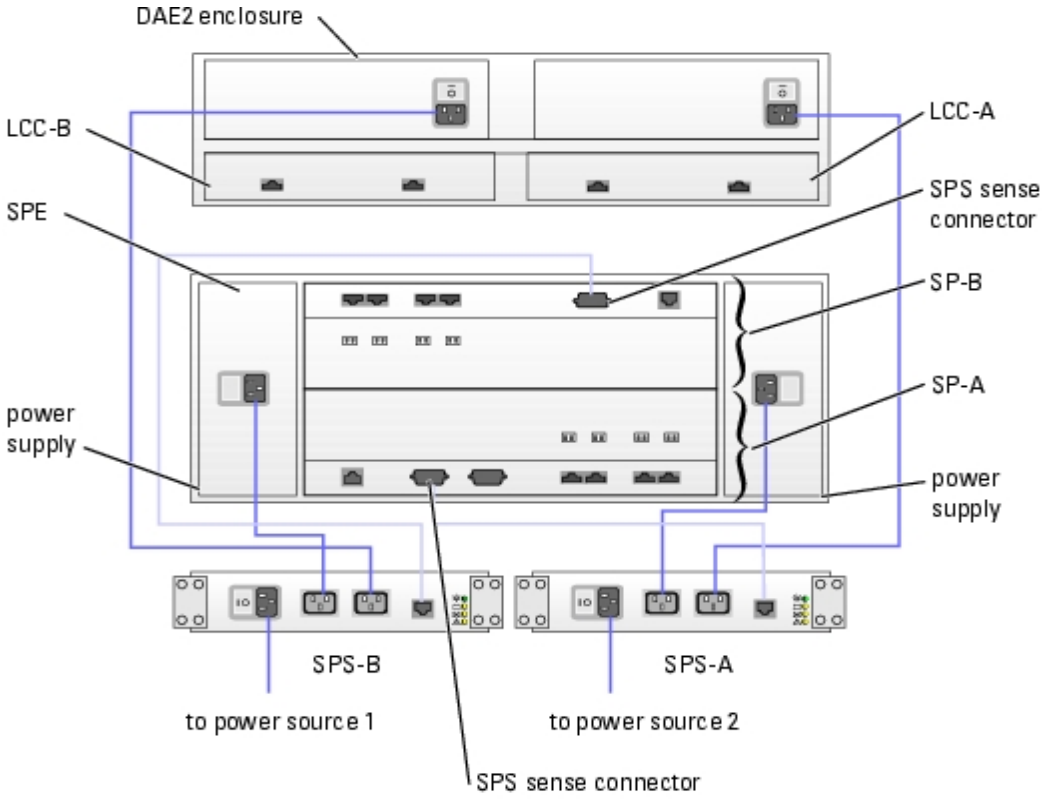


Figure 3-6. CX600 Power Cable Configuration



Back to Contents Page

Back to Contents Page

Installing Your Cluster in a Direct-Attached Environment

Dell™ PowerVault™ NAS Fibre Channel Cluster Systems Installation and Troubleshooting Guide

---

Before You Begin

Verify that your cluster hardware and storage systems installed and configured properly as explained in the following sections:

- "Cabling Your Cluster Hardware"
- "Preparing Your Systems for Clustering"

NOTICE: Microsoft® Windows® standby mode and hibernation mode are not supported in cluster configurations. Do not enable either mode.

Connecting the Storage Systems to Your Cluster

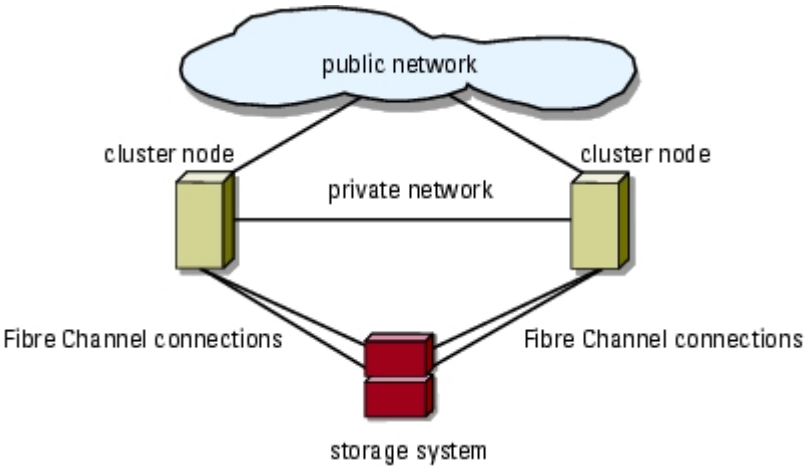
This section provides information for connecting your cluster to one or more storage systems using direct-attached cable connections.

Direct-Attached Cluster

A direct-attached cluster configuration consists of redundant Fibre Channel HBAs cabled directly to a Dell | EMC storage system. Direct-attach configurations are self-contained and do not share any physical resources with other server or storage systems outside of the cluster.

Figure 4-1 shows an example of a direct-attached, single cluster configuration with redundant HBAs installed in each cluster node.

Figure 4-1. Direct-Attached Cluster Configuration



Cabling a Cluster to a Dell | EMC Storage System

Each cluster node attaches to the storage system using two fiber optic cables with duplex LC multimode connectors that attach to the HBAs in the cluster nodes and the SP ports in the Dell | EMC storage system. These connectors consist of two individual fibre optic connectors with indexed tabs that must be aligned properly into the HBAs and SP ports.

NOTICE: Do not remove the connector covers until you are ready to insert the connectors into the system.

The duplex LC multimode connectors attach to the SP ports (SP-A and SP-B) on the back of the storage system.

Cabling a Two-Node Cluster to a CX200 or CX400 Storage System

1. Connect cluster node 1 to the storage system.
    a. Install a cable from cluster node 1 HBA0 to SP-B port FE 0 (or front-end port A [FE A]).
    b. Install a cable from cluster node 1 HBA1 to SP-A port FE 0 (or FE A).
2. Connect cluster node 2 to the storage system.
    a. Install a cable from cluster node 2 HBA0 to SP-B port FE 1 (or FE B).
    b. Install a cable from cluster node 2 HBA1 to SP-A port FE 1 (or FE B).

Figure 4-2 and Figure 4-3 illustrate how to cable a two-node direct-attached cluster to a CX200 and CX400 storage system, respectively.

NOTE: The cables are connected to the storage processor ports in sequential order for illustrative purposes. While the available ports in your storage system may vary, HBA0 and HBA1 must be connected to SP-A and SP-B, respectively.

Figure 4-2. Cabling the Cluster Nodes to a CX200 Storage System

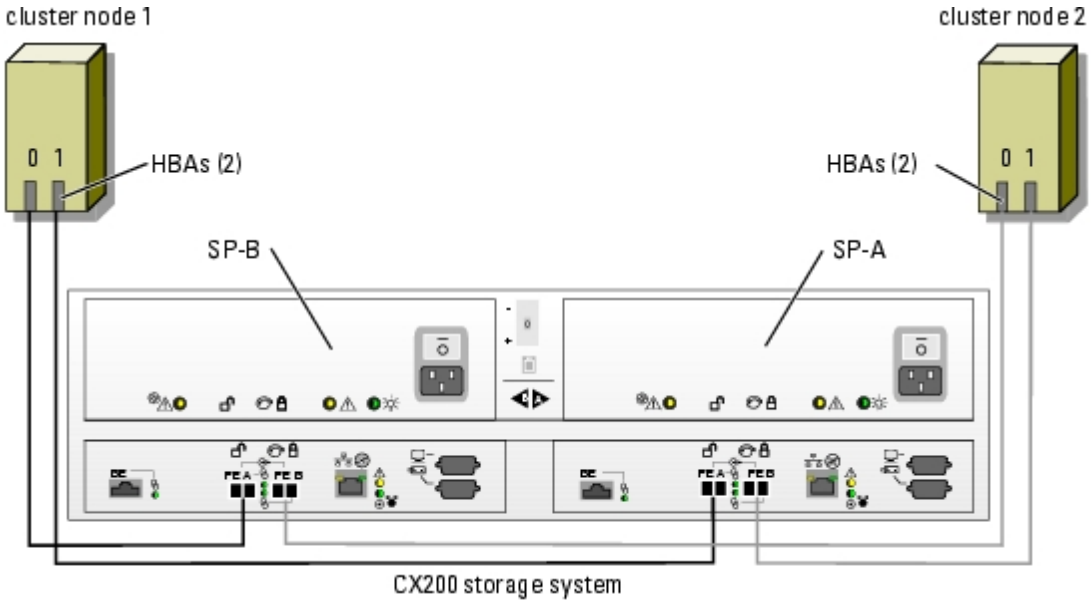
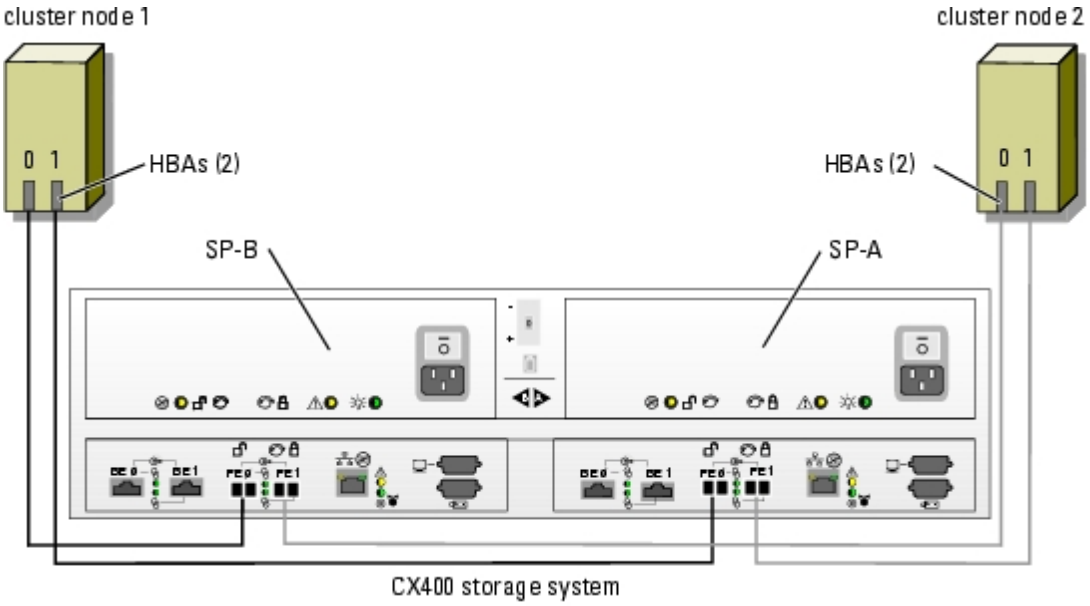


Figure 4-3. Cabling the Cluster Nodes to a CX400 Storage System

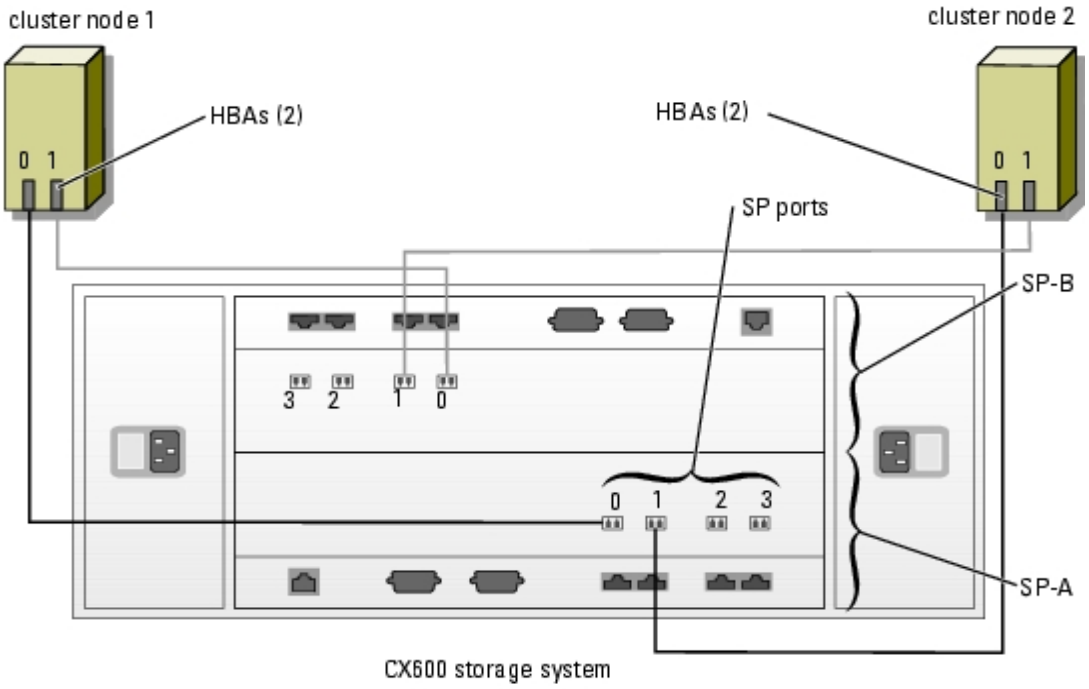


Cabling a Cluster to a CX600 Storage System

1. Connect cluster node 1 to the storage system.
    a. Install a cable from cluster node 1 HBA0 to SP-A port 0.
    b. Install a cable from cluster node 1 HBA1 to SP-B port 0.
2. Connect cluster node 2 to the storage system.
    a. Install a cable from cluster node 2 HBA0 to SP-A port 1.
    b. Install a cable from cluster node 2 HBA1 to SP-B port 1.

Figure 4-4 illustrates how to cable a two-node direct-attached cluster to a CX600 storage system.

Figure 4-4. Cabling the Cluster Nodes to a CX600 Storage System



Cabling a Four-Node Cluster to a CX200 and CX400 Storage System

The CX200 and CX400 storage systems do not support more than two cluster nodes in a direct-attached cluster configuration. Only the CX600 storage system can support a four-node cluster configuration.

Cabling a Four-Node Cluster to a CX600 Storage System

1. Connect cluster node 1 to the storage system.
    a. Install a cable from cluster node 1 HBA0 to SP-A port 0.
    b. Install a cable from cluster node 1 HBA1 to SP-B port 0.
2. Connect cluster node 2 to the storage system.
    a. Install a cable from cluster node 2 HBA0 to SP-A port 1.
    b. Install a cable from cluster node 2 HBA1 to SP-B port 1.
3. Connect cluster node 3 to the storage system.
    a. Install a cable from cluster node 3 HBA0 to SP-A port 2.
    b. Install a cable from cluster node 3 HBA1 to SP-B port 2.
4. Connect cluster node 4 to the storage system.
    a. Install a cable from cluster node 4 HBA0 to SP-A port 3.
    b. Install a cable from cluster node 4 HBA1 to SP-B port 3.

Cabling Two Clusters to a Dell | EMC Storage System

The CX200 and CX400 storage systems do not support more than one direct-attached 2-node cluster.

The CX600 storage system includes four ports on each storage processor, allowing you to connect two 2-node clusters or a single four-node cluster running Windows Storage Server 2003, Enterprise Edition to the storage system in a direct-attached configuration.

NOTE: EMC® Access Logix™ is required if the CX600 storage system is connected to more than one cluster in a direct-attached configuration.

Cabling Two 2-Node Clusters to a Dell | EMC CX600 Storage System

1. In the first cluster, connect cluster node 1 to the storage system.
    a. Install a cable from cluster node 1 HBA0 to SP-A port 0.
    b. Install a cable from cluster node 1 HBA1 to SP-B port 0.
2. In the first cluster, connect cluster node 2 to the storage system.
    a. Install a cable from cluster node 2 HBA0 to SP-A port 1.
    b. Install a cable from cluster node 2 HBA1 to SP-B port 1.
3. In the second cluster, connect cluster node 1 to the storage system.
    a. Install a cable from cluster node 1 HBA0 to SP-A port 2.
    b. Install a cable from cluster node 1 HBA1 to SP-B port 2.
4. In the second cluster, connect cluster node 2 to the storage system.
    a. Install a cable from cluster node 2 HBA0 to SP-A port 3.
    b. Install a cable from cluster node 2 HBA1 to SP-B port 3.

Back to Contents Page

Back to Contents Page

Installing Your Cluster in a SAN Environment

Dell™ PowerVault™ NAS Fibre Channel Cluster Systems Installation and Troubleshooting Guide

---

SAN Overview

A SAN is a high-performance network that is used to move data between server and storage resources, providing scalable storage in a networked environment. Using Fibre Channel technology and fiber optic cabling, a SAN provides the high-speed connection between the cluster nodes and the external storage systems.

Table 5-1 lists the key components in a SAN.

Table 5-1. Key SAN Components

Component	Description
Cluster nodes	Dell™ PowerVault™ systems used as cluster nodes in the cluster. Each system must have dual HBAs installed in the server to manage the I/O data transfer from the server's internal bus to the storage system.
Interconnects	The connections between the cluster nodes and the storage systems. These connections include switches and cabling.
Storage system	Provides external data storage for the host systems (cluster nodes). Supported storage systems include tape libraries, DAEs, DPE, and SPEs.
Fabrics	Private connections of one or more Fibre Channel switches that provide connections between the cluster nodes and storage systems.
SAN management software	Provides centralized control of the SAN for easier management. NAS cluster solutions use EMC® storage management software utilities, which include: • EMC ControlCenter™ Navisphere® Agent • EMC ControlCenter Navisphere Manager™ • EMC PowerPath® • EMC Access Logix™ (optional) • EMC MirrorView™ (optional) • EMC SnapView™ (optional) • EMC SAN Copy™ (optional) See "Storage Management Software" for more information on the storage management software tools.

Connecting the Storage Systems to Your Cluster

This section provides information for connecting your cluster to one or more storage systems using a SAN (Fibre Channel switch fabric).

SAN-Attached Cluster

A SAN-attached cluster is a cluster configuration where all cluster nodes are attached to a single storage system or to multiple storage systems through a SAN using a redundant switch fabric. SAN-attached cluster configurations provide more flexibility, expandability, and performance than direct-attached configurations.

See "Fibre Channel Switch Fabric" for more information on Fibre Channel switch fabrics.

Figure 5-1 shows an example of a SAN-attached cluster. Figure 5-2 shows an example of an eight-node, SAN-attached cluster.

Figure 5-1. SAN-Attached Cluster

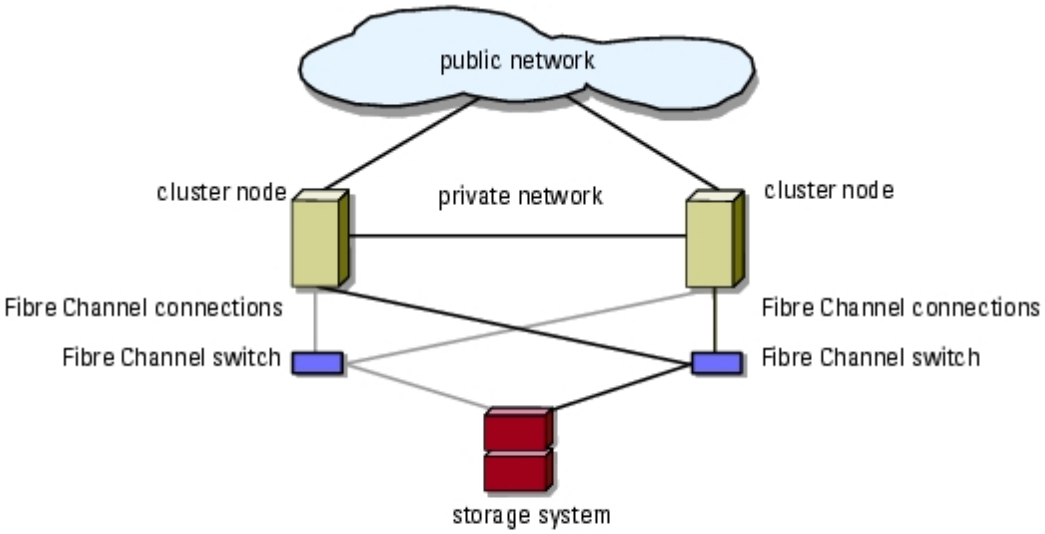
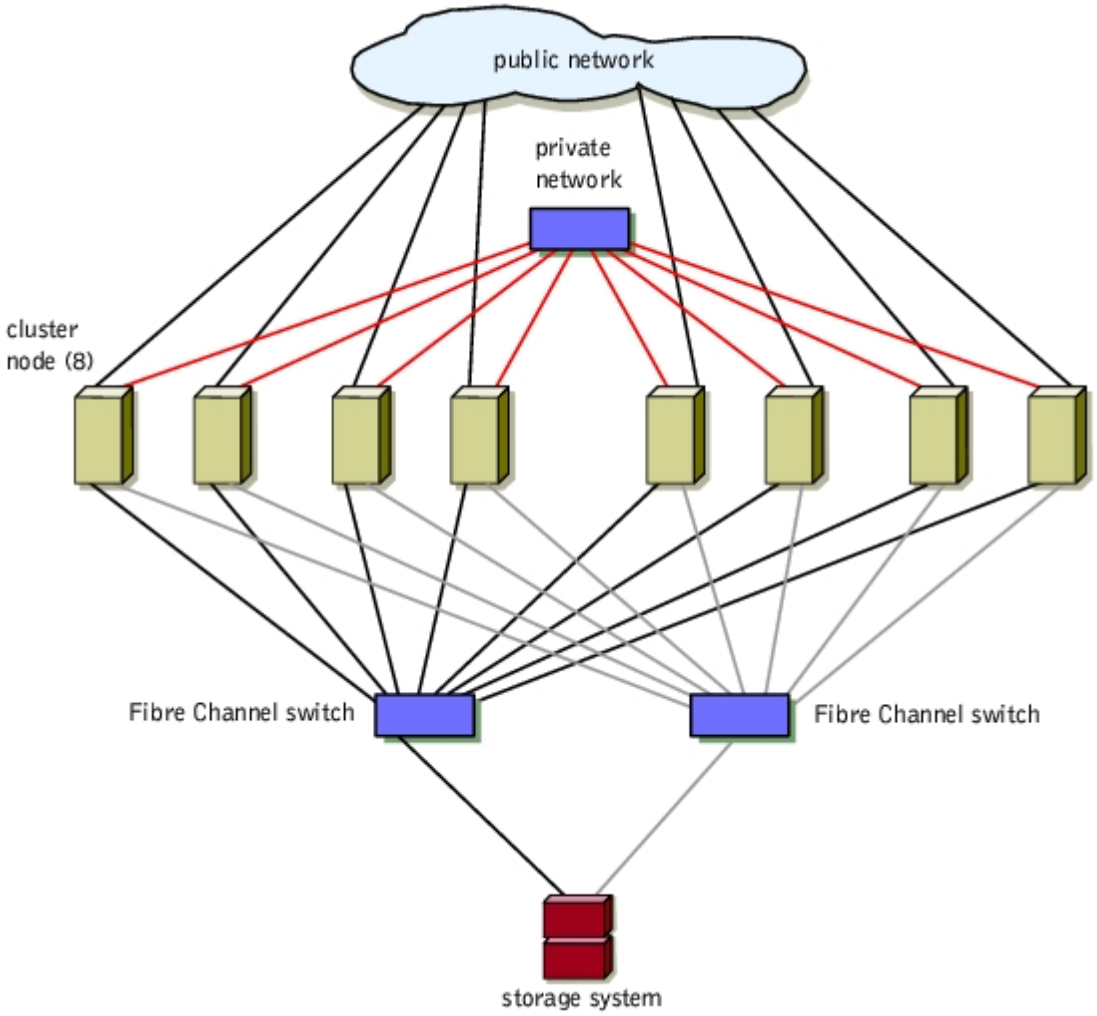


Figure 5-2. SAN-Attached Cluster Running Windows Storage Server 2003, Enterprise Edition



Cabling a SAN-Attached Cluster to a Dell | EMC Storage System

The supported Dell | EMC storage systems are configured as follows:

- CX200 — One DPE, one DAE2 enclosure, and two SPSs
- CX400 — One DPE, one or more DAE2 enclosures, and two SPSs
- CX600 — One SPE, at least one DAE2 (or DAE2-OS) enclosure, and two SPSs

The cluster nodes attach to the storage system using a redundant switch fabric and fiber optic cables with duplex LC multimode connectors.

The switches, the HBAs in the cluster nodes, and the SP ports in the storage system use duplex LC multimode connectors. The connectors consist of two individual fiber optic connectors with indexed tabs that must be inserted and aligned properly in the small form-factor pluggable (SFP) module connectors on the Fibre Channel switches and the connectors on the cluster nodes and storage systems.

See "Fibre Channel Cable Connectors" for more information on the duplex LC multimode fiber optic connector.

Each HBA is cabled to a port on a Fibre Channel switch. One to four cables connect from the outgoing ports on a switch to a storage processor on a Dell | EMC storage system.

Table 5-2 provides information for cabling your storage system to the Fibre Channel switch.

Figure 5-3 and Figure 5-4 illustrate how to cable a SAN-attached cluster to the CX200 and CX400 storage systems, respectively.

Figure 5-5 illustrates how to cable a SAN-attached cluster to a CX600 storage system.

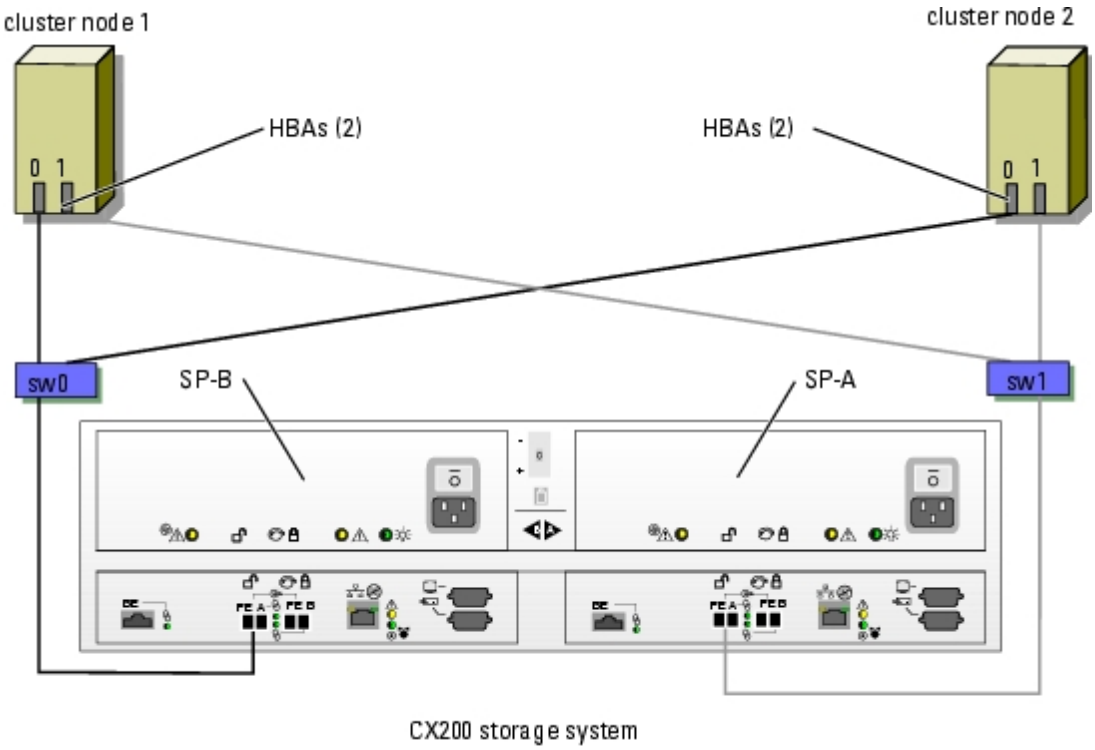
Table 5-2. Storage System Cabling Description

Storage System	SP Ports	Fiber Optic Cables Required	Cabling Description
CX200	One port per storage processor	2	Attach one cable from each storage processor port to the Fibre Channel switch.
CX400	Two ports per storage processor	4	
CX600	Four ports per storage processor	8	

Cabling a SAN-Attached Cluster to a Dell | EMC CX200 Storage System

1. Connect cluster node 1 to the SAN.
    a. Connect a cable from HBA0 to Fibre Channel switch 0 (sw0).
    b. Connect a cable from HBA1 to Fibre Channel switch 1 (sw1).
2. Repeat step 1 for each cluster node.
3. Connect the storage system to the SAN.
    a. Connect a cable from Fibre Channel switch 0 (sw0) to SP-B port FE A.
    b. Connect a cable from Fibre Channel switch 1(sw1) to SP-A port FE A.

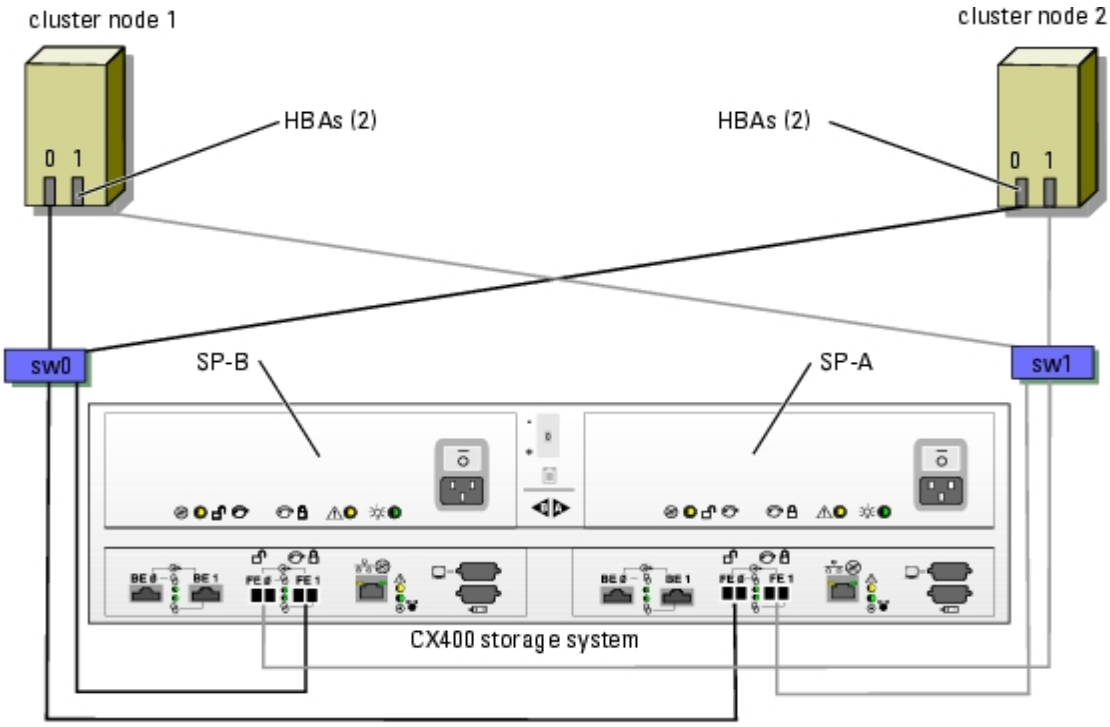
Figure 5-3. Cabling a SAN-Attached Cluster to the Dell | EMC CX200 DPE



Cabling a SAN-Attached Cluster to a Dell | EMC CX400 Storage System

1. Connect cluster node 1 to the SAN.
    a. Connect a cable from HBA0 to Fibre Channel switch 0 (sw0).
    b. Connect a cable from HBA1 to Fibre Channel switch 1 (sw1).
2. Repeat step 1 for each node.
3. Connect the storage system to the SAN.
    a. Connect a cable from Fibre Channel switch 0 (sw0) to SP-A port FE 0.
    b. Connect a cable from Fibre Channel switch 0 (sw0) to SP-B port FE 1.
    c. Connect a cable from Fibre Channel switch 1 (sw1) to SP-A port FE 1.
    d. Connect a cable from Fibre Channel switch 1 (sw1) to SP-B port FE 0.

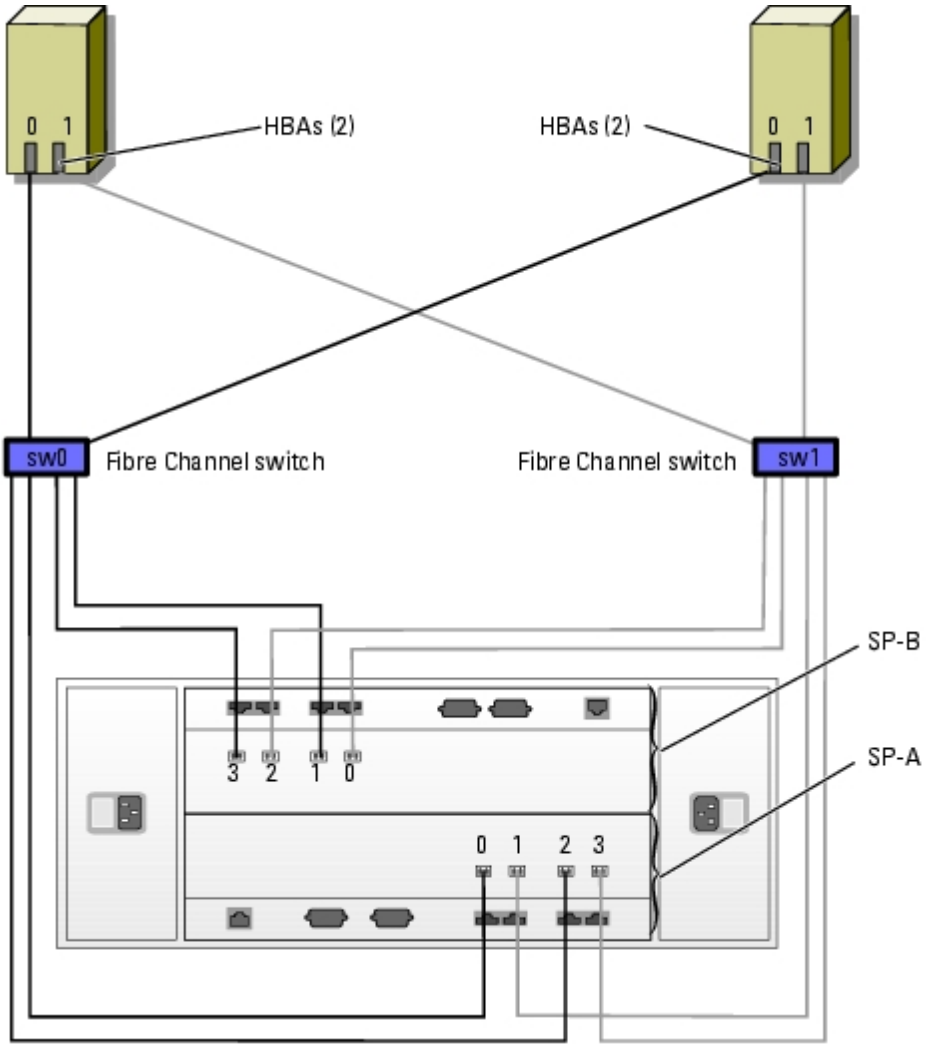
Figure 5-4. Cabling a SAN-Attached Cluster to the CX400 DPE



Cabling a SAN-Attached Cluster to the CX600 Storage System

1. Connect cluster node 1 to the SAN.
    a. Connect a cable from HBA0 to Fibre Channel switch 0 (sw0).
    b. Connect a cable from HBA1 to Fibre Channel switch 1 (sw1).
2. Repeat step 1 for each node.
3. Connect the storage system to the SAN.
    a. Connect a cable from Fibre Channel switch 0 (sw0) to SP-A port 0.
    b. Connect a cable from Fibre Channel switch 0 (sw0) to SP-A port 2.
    c. Connect a cable from Fibre Channel switch 0 (sw0) to SP-B port 1.
    d. Connect a cable from Fibre Channel switch 0 (sw0) to SP-B port 3.
    e. Connect a cable from Fibre Channel switch 1 (sw1) to SP-A port 1.
    f. Connect a cable from Fibre Channel switch 1 (sw1) to SP-A port 3.
    g. Connect a cable from Fibre Channel switch 1 (sw1) to SP-B port 0.
    h. Connect a cable from Fibre Channel switch 1 (sw1) to SP-B port 2.

Figure 5-5. Cabling a SAN-Attached Cluster to the CX600 SPE



Cabling Multiple SAN-Attached Clusters to a Dell | EMC Storage System

To cable two clusters to the storage system, connect the cluster nodes to the appropriate Fibre Channel switches and then connect the Fibre Channel switches to the appropriate storage processors on the processor enclosure.

See the *Platform Guide* for rules and guidelines for SAN-attached clusters.

NOTE: The following procedures use Figure 5-3, Figure 5-4, and Figure 5-5 as examples for cabling additional clusters.

Cabling Multiple SAN-Attached Clusters to the CX200 Storage System

1. In the first cluster, connect cluster node 1 to the SAN.
    a. Connect a cable from HBA0 to Fibre Channel switch 0 (sw0).
    b. Connect a cable from HBA1 to Fibre Channel switch 1 (sw1).
2. In the first cluster, repeat step 1 for each node.
3. For each additional cluster, repeat step 1 and step 2.
4. Connect the storage system to the SAN.
    a. Connect a cable from Fibre Channel switch 0 (sw0) to SP-A port FE A.

b. Connect a cable from Fibre Channel switch 0 (sw0) to SP-B port FE A.

Cabling Multiple SAN-Attached Clusters to the CX400 Storage System

1. In the first cluster, connect cluster node 1 to the SAN.
    a. Connect a cable from HBA0 to Fibre Channel switch 0 (sw0).
    b. Connect a cable from HBA1 to Fibre Channel switch 1 (sw1).
2. In the first cluster, repeat step 1 for each node.
3. For each additional cluster, repeat step 1 and step 2.
4. Connect the storage system to the SAN.
    a. Connect a cable from Fibre Channel switch 0 (sw0) to SP-A port FE 0.
    b. Connect a cable from Fibre Channel switch 0 (sw0) to SP-B port FE 1.
    c. Connect a cable from Fibre Channel switch 1 (sw1) to SP-A port FE 1.
    d. Connect a cable from Fibre Channel switch 0 (sw1) to SP-B port FE 0.

Cabling Multiple SAN-Attached Clusters to the CX600 Storage System

1. In the first cluster, connect cluster node 1 to the SAN.
    a. Connect a cable from HBA0 to Fibre Channel switch 0 (sw0).
    b. Connect a cable from HBA1 to Fibre Channel switch 1 (sw1).
2. In the first cluster, repeat step 1 for each node.
3. For each additional cluster, repeat step 1 and step 2.
4. Connect the storage system to the SAN.
    a. Connect a cable from Fibre Channel switch 0 (sw0) to SP-A port 0.
    b. Connect a cable from Fibre Channel switch 0 (sw0) to SP-A port 2.
    c. Connect a cable from Fibre Channel switch 0 (sw0) to SP-B port 1.
    d. Connect a cable from Fibre Channel switch 0 (sw0) to SP-B port 3.
    e. Connect a cable from Fibre Channel switch 1 (sw1) to SP-A port 1.
    f. Connect a cable from Fibre Channel switch 1 (sw1) to SP-A port 3.
    g. Connect a cable from Fibre Channel switch 1 (sw1) to SP-B port 0.
    h. Connect a cable from Fibre Channel switch 1 (sw1) to SP-B port 2.

Zoning Your Dell | EMC Storage System in a Switched Environment

Dell only supports single-initiator zoning for connecting NAS clusters to a Dell | EMC storage system in a switched environment. When using EMC PowerPath, a separate zone is created from each HBA to the following hardware components:

- CX200 — A single SP port on each SP
- CX400 and CX600 — One or more SP ports

Connecting a Cluster to Multiple Storage Systems

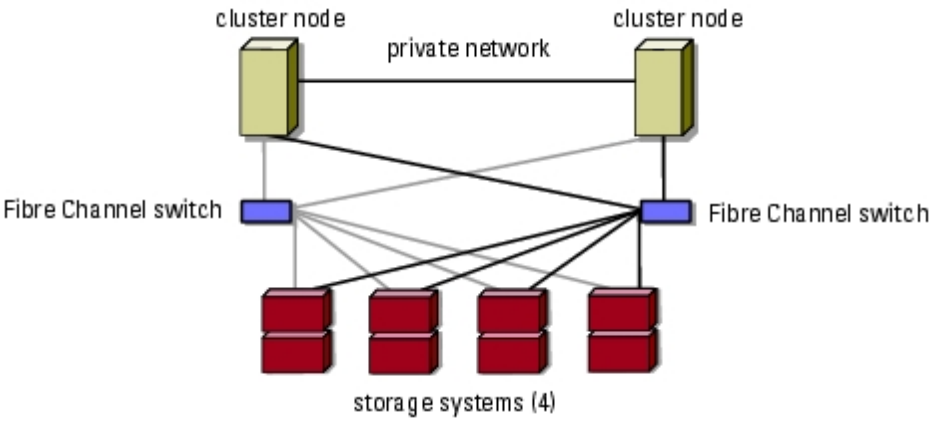
You can increase your cluster storage capacity by attaching multiple storage systems to your cluster using a redundant switch fabric. NAS cluster systems can support configurations with multiple storage units attached to clustered servers. In this scenario, the MSCS software can fail over disk drives in any cluster-attached shared storage array between the cluster nodes.

When attaching multiple storage systems with your cluster, the following rules apply:

- A maximum of four storage systems per cluster.
- The shared storage systems and firmware must be identical. Using dissimilar storage systems and firmware for your shared storage is not supported.
- MSCS is limited to 22 drive letters. Because drive letters A through D are reserved for local disks, a maximum of 22 drive letters (E to Z) can be used for your storage system disks.

Figure 5-6 provides an example of cabling the cluster nodes to four Dell | EMC storage systems. See "Fibre Channel Switch Fabric" for more information.

Figure 5-6. Cluster Nodes Cabled to Four Storage Systems



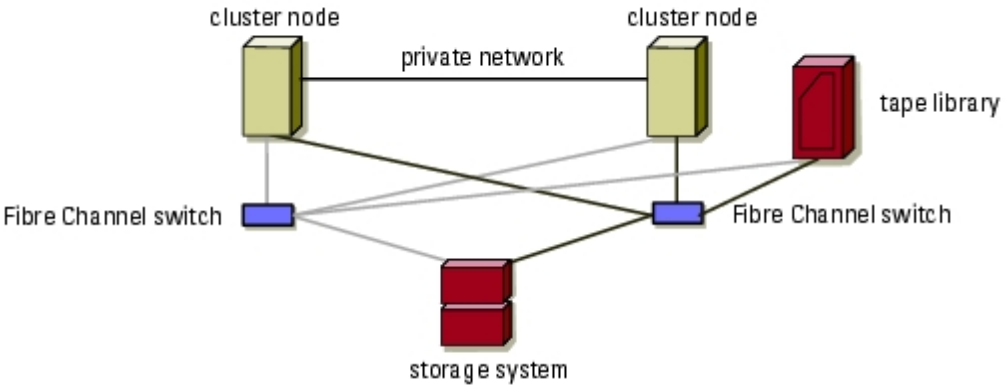
Connecting a Cluster to a Tape Library

To provide additional backup for your cluster, you can add tape backup devices to your cluster configuration. The Dell™ PowerVault™ tape libraries contain an integrated Fibre Channel bridge that connects directly to your Dell | EMC Fibre Channel switch.

Figure 5-7 shows a supported NAS cluster configuration using redundant Fibre Channel switches and a tape library. In this configuration, each of the cluster nodes can access the tape library to provide backup for your local disk resources, as well as your cluster disk resources. Using this configuration allows you to add more servers and storage systems in the future, if needed.

NOTE: While tape libraries can be connected to multiple fabrics, they do not provide path failover.

Figure 5-7. Cabling a Storage System and a Tape Library



Obtaining More Information

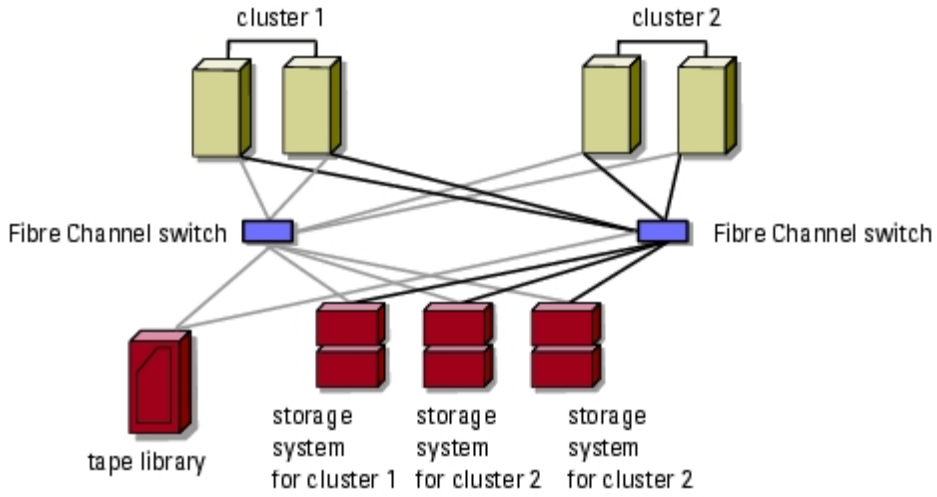
See the storage and tape backup documentation for more information on configuring these components.

Configuring Your Cluster With SAN Backup

You can provide centralized backup for your clusters by sharing your SAN with multiple clusters, storage systems, and a tape library.

Figure 5-8 provides an example of cabling the cluster nodes to your storage systems and SAN backup with a tape library.

Figure 5-8. Cluster Configuration Using SAN-Based Backup



Enabling Access Control and Creating Storage Groups

Fibre Channel allows multiple clusters and stand-alone systems to share a single storage system. To configure cluster and server access to a shared Dell | EMC storage system, you must enable **Access Control** to link the cluster nodes and stand-alone systems to the storage system and then use Navisphere Manager to create storage groups. When you enable **Access Control** and then create storage groups in Navisphere Manager, you are using the Access Logix software.

Access Logix allows multiple cluster nodes and servers to share a Dell | EMC storage system, restricts server access to specific volumes on a shared Dell | EMC storage system, and protects your data from unauthorized access. Access Logix is required when the storage system is connected to two or more clusters, two or more nonclustered systems, or a combination of both clustered and nonclustered systems.

Using Access Logix, you can partition LUNs in a storage system into RAID groups, and then use storage groups to restrict LUN access to the assigned hosts. Access Logix is installed on your Dell | EMC storage system when you update the core software (firmware).

After you update the core software, you can enable Access Control. The default setting is set to **Disabled**. However, after you enable Access Control, it cannot be disabled.

Access Logix is not required on all cluster configurations. Table 5-3 provides a list of cluster configurations and their Access Logix requirement.

Table 5-3. Access Logix Requirement

System Configuration	Access Logix Required
Single host or One cluster	No
Two or more clusters or Two or more nonclustered hosts or Any combination of clusters and nonclustered hosts	Yes

See your Dell | EMC storage system documentation for more information.

To enable Access Control and create storage groups:

1. Open a Web browser window.
2. Enter the IP address of the storage management server on your Dell | EMC storage system.

   NOTE: The storage management server is usually one of the SPs on your storage system, but may also be a host system running the Navisphere Management Server for Windows software.

3. In the **Enterprise Storage** window, click the **Storage** tab.
4. Right-click the icon of your storage system.
5. In the drop-down menu, click **Properties**.
6. In the **Storage Systems Properties** window, click the **Storage Access** tab.

   If the **Storage Access** tab does not appear in the **Storage Systems Properties** window, Access Logix is not installed on your system.

7. In the **Data Access** box, select the **Access Control Enabled** check box.

   A dialog box appears, asking if you want to enable **Access Control**.

NOTICE: Before enabling **Access Control**, ensure that no hosts are attempting to access the storage. Enabling **Access Control** prevents all hosts from accessing any data until they are given explicit access to a LUN in the appropriate storage group. You must stop all I/O before enabling **Access Control**. Dell recommends powering off all hosts connected to the storage system during this procedure or data loss may occur. After you enable **Access Control**, it cannot be disabled.

8. Click **Yes** to enable **Access Control**.
9. Right-click the icon of your storage system and select **Create Storage Group**.

   The **Create Storage Group** dialog box appears.
10. In the **Storage Group Name** field, enter a name for the storage group.
11. In the **Sharing State** drop-down menu, select one of the following:
    - **Shareable** — If the storage group is for a cluster
    - **Dedicated** — If the storage group is for a nonclustered server
12. Click **Apply**.
13. Add new LUNs to the storage group.
    a. Right-click the icon of your storage group and select **Properties**.
    b. Click the **LUNs** tab.
    c. In the **Available LUNs** window, click an available LUN.
    d. Click the right-arrow button to move the selected LUN to the **Selected LUNs** pane.
    e. Click **Apply**.
14. Click the **Hosts** tab.
15. In the **Available Hosts** window, click an available host.
16. Click the right-arrow button to move the selected host to the **Hosts to be Connected** dialog box.
17. Click **OK** to save your configuration and exit.
18. Repeat step 9 through step 17 for each additional cluster or nonclustered hosts that will access your storage system.

Implementing Zoning on a Fibre Channel Switched Fabric

A Fibre Channel switched fabric consists of one or more Fibre Channel switches that provide high-speed connections between servers and storage devices. The switches in a Fibre Channel fabric provide a connection through inbound and outbound points from one device (sender) to another device or switch (receiver) on the network. If the data is sent to another switch, the process repeats itself until a connection is established between the sender and the receiver.

Fibre Channel switches provide you with the ability to set up barriers between different devices and operating environments. These barriers create logical fabric subsets with minimal software and hardware intervention. Similar to subnets in the client/server network, logical fabric subsets divide a fabric into similar groups of components, regardless of their proximity to one another. The logical subsets that form these barriers are called *zones*.

Zoning automatically and transparently enforces access of information to the zone devices. More than one NAS cluster configuration can share a Dell | EMC SAN in a switched fabric using Fibre Channel switch zoning and Access Logix. By using Fibre Channel switches to implement zoning, you can segment the SANs to isolate heterogeneous servers and storage systems from each other.

Using Zoning in SAN Configurations Containing Multiple Hosts

Using the combination of zoning and Access Logix in SAN configurations containing multiple hosts, you can restrict server access to specific volumes on a shared storage system by preventing the hosts from discovering a storage device that belongs to another host. This configuration allows multiple clustered or nonclustered hosts from sharing a storage system.

Using Worldwide Port Name Zoning

NAS cluster configurations support worldwide port name zoning.

A world wide name (WWN) is a unique numeric identifier assigned to Fibre Channel devices, such as HBAs, SPs, and SCSI to Fibre Channel bridges.

A WWN consists of an 8-byte hexadecimal number with each byte separated by a colon. For example, 10:00:00:60:69:00:8a is a valid WWN. Using WWN port name zoning allows you to move cables from port-to-port within the switch fabric without having to update the zones.

Table 5-4 provides a list of WWN identifiers that you can find in the Dell | EMC cluster environment.

Table 5-4. Port Worldwide Names in a SAN Environment

Identifier	Description
*xx:xx:*00:60:69:*xx:xx:xx*	Dell | EMC or Brocade switch
50:06:01:6*x:xx:xx:xx:xx*	Dell | EMC storage processor
*xx:xx:*00:00:C9:*xx:xx:xx*	Emulex LP9002L and LP982 HBAs
*xx:xx:*00:E0:8B:*xx:xx:xx*	QLogic QLA2340 HBA
*xx:xx:xx:*60:45:*xx:xx:xx*	PowerVault 132T and 136T tape libraries
*xx:xx:xx:*E0:02:*xx:xx:xx*	PowerVault 128T tape autoloader
*xx:xx:xx:*00:88:*xx:xx:xx*	McData switch

NOTICE: You must reconfigure the zones if you have replaced Fibre Channel storage components, such as Fibre Channel HBAs or switches. Failure to reconfigure the zones may cause data loss or data corruption.

NOTICE: You must configure your zones before you configure the paths and LUNs. Failure to do so may cause data loss, data corruption, or data unavailability.

Single Initiator Zoning

Each HBA in a SAN must be configured in a separate zone with the storage ports on the switch. This zoning configuration, known as *single initiator zoning*, prevents heterogeneous HBAs from communicating with each other, thereby ensuring that Fibre Channel communications between the HBAs and their target storage systems do not affect one another.

When you zone your HBAs, follow these guidelines:

- Create a zone for each HBA and its target storage devices.
- Each CX200 storage processor can be connected to a maximum of 15 HBAs in a SAN-attached environment
- Each CX400 storage processor port can be connected to a maximum of 15 HBAs in a SAN-attached environment
- Each CX600 storage processor port can be connected to a maximum of 32 HBAs in a SAN-attached environment.

- Each HBA can be connected to a maximum of four storage systems
- The integrated bridge on a tape library can be added to any zone.

NOTE: If you are sharing a storage system with multiple clusters or a combination of clustered and nonclustered systems (hosts), you must install Access Logix and enable Access Control. Otherwise, you can only have one nonclustered system or one NAS cluster attached to the Dell | EMC storage system.

SAN-Attached Cluster Configurations

Certain rules and requirements apply when you configure your SAN-attached cluster. See the *Platform Guide* for information about supported servers, specific HBA models supported by the platform, and PCI slot configuration guidelines.

Table 5-5 provides a list of the documentation you will need to configure a SAN-attached cluster.

Table 5-5. SAN-Attached Cluster Configuration Documentation

Information	Documentation	Location
General rules and guidelines for SAN-attached cluster configurations Rules and requirements for cluster consolidation configurations	This document *Platform Guide* and this document	Included with your NAS cluster or at the Dell Support website at **support.dell.com**
Using Windows Storage Server 2003, Enterprise Edition for NAS cluster products and components	This document	
Installing the Navisphere software	EMC storage system documentation	Included with your storage system or at the EMC Support website at **www.emc.com**
Latest firmware and software revision requirements	EMC Support Matrix	**EMC Technical Library** at **www.emc.com**
	Platform Guide	Included with your NAS cluster, or at the Dell Support website at **support.dell.com**

Back to Contents Page

Back to Contents Page

Maintaining Your Cluster

Dell™ PowerVault™ NAS Fibre Channel Cluster Systems Installation and Troubleshooting Guide

This section provides the following maintenance procedures for systems running the Microsoft® Windows® Storage Server 2003, Enterprise Edition operating system:

- Launching Microsoft Cluster Administrator
- Adding a NIC to a cluster node
- Changing the IP address of a cluster node
- Uninstalling MSCS
- Running **chkdsk /f** on a quorum resource
- Recovering from a corrupt quorum resource
- Replacing a cluster node
- Changing the cluster password
- Reformatting a cluster volume

Launching Microsoft Cluster Administrator

Cluster Administrator is Microsoft's tool for configuring and administering a cluster. The following procedures describe how to run Cluster Administrator locally on a cluster node and how to install the tool on a remote console.

To launch Cluster Administrator, perform the following steps:

1. Click the **Start** button and select **Programs**.
2. Select **Administrative Tools**.
3. Select **Cluster Administrator**.

Adding a NIC to a Cluster Node

This procedure assumes that Windows Storage Server 2003, Enterprise Edition, the current Windows Service Pack, and MSCS are installed on both cluster nodes.

NOTE: The IP addresses used in the following sections are examples only and are not representative of actual addresses to use. The IP addresses are 192.168.1.101 for the NIC in the first node and 192.168.1.102 for the NIC in the second node. The subnet mask for both nodes is 255.255.255.0.

1. Move all cluster resources from the cluster node you are upgrading to another node in the cluster.

   See the Cluster Service documentation for information about moving cluster resources to a specific node.
2. Shut down the cluster node you are upgrading and install the additional NIC in that system.

   See the system *Installation and Troubleshooting Guide* for instructions about installing expansion cards in your system.
3. Boot to the Windows operating system.

Windows Plug and Play detects the new NIC and installs the appropriate drivers.

NOTE: If Plug and Play *does not* detect the new NIC, the NIC is not supported.

a. Update the NIC drivers (if required).

b. After the drivers are installed, click the **Start** button, select **Control Panel**, and then double-click **Network Connections**.

c. In the **Connections** box, locate the new NIC that you installed in the system.

d. Right-click the new NIC, and select **Properties**.

e. Assign a unique static IP address, subnet mask, and gateway.

4. Ensure that the network ID portion of the new NICs IP address is different from the other adapter.

   For example, if the first NIC in the node had an address of 192.168.1.101 with a subnet mask of 255.255.255.0, you might enter the following IP address and subnet mask for the second NIC:

   IP address: `192.168.2.102`

   Subnet mask: `255.255.255.0`

5. Click **OK** and exit NIC properties.
6. On the Windows desktop, click the **Start** button and select **Programs→ Administrative Tools→ Cluster Administrator**.
7. Click the **Network** tab.
8. Verify that a new resource called "New Cluster Network" appears in the window.

   To rename the new resource, right-click the resource and enter a new name.

9. Move all cluster resources to another cluster node.
10. Repeat step 2 through step 9 on each cluster node.

NOTE: Ensure that you assign the new network adapter with the same IP address as the second network adapter on the first node (for example, 192.168.2.101) as you did with the second node.

If the installation and IP address assignments have been performed correctly, all of the new NIC resources appear online and respond successfully to **ping** commands.

Changing the IP Address of a Cluster Node on the Same IP Subnet

NOTE: If you are migrating your cluster nodes to a different subnet, take all cluster resources offline and then migrate all nodes together to the new subnet.

1. Open Cluster Administrator.
2. Stop MSCS on the cluster node.

   The Cluster Administrator utility running on the second cluster node indicates that the first node is down by displaying a red icon in the **Cluster Service** window.

3. Reassign the IP address.
4. If you are running DNS, verify that the DNS entries are correct (if required).
5. Restart MSCS on the cluster node.

   The cluster nodes re-establish their connection and Cluster Administrator changes the node icon back to blue to show that the node is back online.

Removing a Node From a Cluster

1. Take all resource groups offline or move them to another cluster node.
2. Click the **Start** button and select **Programs→ Administrative Tools→ Cluster Administrator**.

3. In **Cluster Administrator**, right-click the icon of the node you want to uninstall and then select **Stop Cluster Service**.
4. In **Cluster Administrator**, right-click the icon of the node you want to uninstall and then select **Evict Node**.

   If you cannot evict the node or the node is the last node in the cluster, perform the following steps:

   NOTICE: To avoid problems, you must follow this procedure if the evicted node is the last node in the cluster

   a. Open a command prompt.

   b. Type the following:

      ```
      cluster node <node_name> /force
      ```

      where <*node_name*> is the cluster node you are evicting from the cluster.
5. Close **Cluster Administrator**.

Running chkdsk /f on a Quorum Resource

You cannot run the **chkdsk** command with the /f (fix) option on a device that has an open file handle active. Because MSCS maintains an open handle on the quorum resource, you cannot run **chkdsk /f** on the hard drive that contains the quorum resource.

To run **chkdsk /f** on a quorum resource's hard drive:

1. Move the quorum resource temporarily to another drive:
   a. Right-click the cluster name and select **Properties**.
   b. Click the **Quorum** tab.
   c. Select another disk as the quorum resource and press <Enter>.
2. Run **chkdsk /f** on the drive that previously stored the quorum resource.
3. Move the quorum resource back to the original drive.

Recovering From a Corrupt Quorum Resource

The quorum resource maintains the configuration data necessary for cluster recovery when a cluster node fails. If the quorum resource is unable to come online, the cluster will not start and all of the shared drives will be unavailable. If this situation occurs, and you need to run **chkdsk** on the quorum resource, you can start the cluster manually from the command line.

To start the cluster manually from a command prompt:

1. Open a command prompt window.
2. Select the cluster folder directory by typing the following:

   ```
   cd\winnt\cluster
   ```
3. Start the cluster in manual mode (on one node only) with no quorum logging by typing the following:

   ```
   Clussvc -debug -noquorumlogging
   ```

   MSCS starts.
4. Run **chkdsk /f** on the disk designated as the quorum resource.

   To run the **chkdsk /f** utility:

   a. Open a second command prompt window.

   b. Type:

```
chkdsk /f q:
```

where *q* is the drive letter assigned to your quorum resource.

5. After the **chkdsk** utility completes, stop MSCS by pressing <Ctrl><c>.
6. Restart the cluster service.

- To restart MSCS from the **Services** console:
  a. Click the **Start** button and select **Programs**→ **Administrative Tools**→ **Services**.
  b. In the **Services** window, right-click **Cluster Service**.
  c. In the drop-down menu, click the **Start** button.
- To restart MSCS from the command prompt:
  a. Open the second command prompt window that you opened in step 4a.
  b. Type the following:

```
Net Start Clussvc
```

The Cluster Service restarts.

See the Microsoft Knowledge Base article 258078 located at the Microsoft Support website at **www.microsoft.com** for more information on recovering from a corrupt quorum resource.

Replacing a Cluster Node

This section assumes that you have a recent tape backup of the cluster node that contains the local registry information.

1. Ensure that the replacement cluster node is physically disconnected from the storage system.
2. Ensure that Windows is installed and configured properly.
3. Install the correct NIC drivers, assign the appropriate IP addresses, and install the driver for the HBAs.
4. If applicable, install the necessary tape device drivers and back up application program.
5. Shut down Windows and turn off the system.
6. Connect the fiber optic cables from each HBA to the appropriate ports on the storage system or Fibre Channel switch.
7. Turn on the replacement cluster node and reassign the LUNs (if applicable).
8. Using the most current tape backup, restore the backup (including the Windows registry files) and restart the system.

   If a backup is not available:
   a. On the remaining cluster node, start **Cluster Administrator**.
   b. Right-click the failed node and select **Evict Node**.
   c. On the replacement node, install and configure MSCS.
   d. Select the option to join an existing cluster.
9. Use Cluster Administrator to verify that the node rejoins the cluster, and check the Windows Event Viewer to ensure errors were not encountered.

NOTICE: If a recent tape backup is not available, all application programs must be reinstalled.

10. Apply any additional service pack.
11. Reinstall any cluster applications (such as Microsoft SQL Server or Exchange Server onto the new node [if required]).

NOTE: You may need to reinstall or configure your cluster applications before moving or testing the failover capabilities of any cluster resources to the new node.

Changing the Cluster Service Account Password

To change the cluster service account password for all nodes in a Windows Storage Server 2003 cluster, open a command prompt and type the following syntax:

```
Cluster /cluster:[cluster_name] /changepass
```

where *cluster_name* is the name of your cluster.

For help with changing the cluster password, type the following:

```
cluster /changepass /help
```

Reformatting a Cluster Disk

NOTE: Ensure that all clients are disconnected from the cluster disk before you perform this procedure.

1. Click the **Start** button and select **Programs**→ **Administrative Tools**→ **Cluster Administrator**.
2. In the **Cluster Administrator** left window pane, expand the **Groups** directory.
3. In the **Groups** directory, right-click a cluster resource group to reformat and select **Take Offline**.
4. In the **Cluster Administrator** right window pane, right-click on the physical disk you are reformatting and select **Bring Online**.
5. In the **Cluster Administer** right window pane, right-click on the physical disk you are reformatting and select **Properties**.

   The **Properties** window appears.
6. Click the **Advanced** tab.
7. In the **Advanced** tab menu in the **"Looks Alive" poll interval** box, select **Specify value**.
8. In the **Specify value** field, type:

   ```
   6000000
   ```

   where 6000000 equals 6000000 milliseconds (or 100 minutes).
9. Click **Apply**.
10. Minimize the **Properties** window.
11. On the Windows desktop, right-click **My Computer** and select **Manage**.

    The **Computer Management** window appears.
12. In the **Computer Management** left window pane, click **Disk Management**.

    The physical disk information appears in the right window pane.
13. Right-click the disk you want to reformat and select **Format**.

    Disk Management reformats the disk.
14. In the **File** menu, select **Exit**.
15. Maximize the **Properties** window.

    The **Properties** window appears.
16. In the **"Looks Alive" poll interval** box, select **Use value from resource type** and click **OK**.
17. In the **Cluster Administrator** left window pane, right-click the cluster group that contains the reformatted disk and select **Bring Online**.
18. In the **File** menu, select **Exit**.

Back to Contents Page

Back to Contents Page

Using MSCS

Dell™ PowerVault™ NAS Fibre Channel Cluster Systems Installation and Troubleshooting Guide

This section provides information about Microsoft® Cluster Service (MSCS). This section is intended to be an overview of MSCS and provides information about the following:

- Cluster objects
- Cluster networks
- Network interfaces
- Cluster nodes
- Groups
- Cluster resources
- Failover and failback

For information about specific MSCS procedures, see the MSCS online help.

Cluster Objects

Cluster objects are the physical and logical units managed by MSCS. Each object is associated with the following:

- One or more properties, or attributes, that define the object and its behavior within the cluster.
- A set of cluster control codes used to manipulate the object's properties.
- A set of object management functions used to manage the object through MSCS.

Cluster Networks

A network performs one of the following roles in a cluster:

- A network that carries internal cluster communication
- A public network that provides client systems with access to cluster application services
- A public-and-private network that carries both internal cluster communication and connects client systems to cluster application services
- Neither a public nor private network that carries traffic unrelated to cluster operation

Preventing Network Failure

The Cluster Service uses all available private and public-and-private networks for internal communication. Configure multiple networks as private or public-and-private to protect the cluster from a single network failure. If there is only one such network available and it fails, the cluster nodes stop communicating with each other. When two nodes are unable to communicate, they are partitioned and the Cluster Service automatically shuts down on one node. While this shutdown guarantees the consistency of application data and the cluster configuration, it can make cluster resources unavailable.

For example, if each node has only one network adapter, and the network cable on one of the nodes fails, each node, (because it is unable to communicate with the other), attempts to take control of the quorum resource. There is no guarantee that the node with a functioning network connection will gain control of the quorum resource. If the node with the failed network cable gains control, the entire cluster is unavailable to network clients. To avoid this problem, ensure that all nodes have at least two networks and are configured to use both networks for the private network (internal communications).

Node-to-Node Communication

The Cluster Service does not use public only networks for internal communication. For example, a cluster has Network A configured as private and Network B configured as public. If Network A fails, the Cluster Service does not use Network B because it is public; the nodes stop communicating and one node terminates its Cluster Service.

Network Interfaces

The Microsoft Windows® operating system keeps track of all NICs in a server cluster. This tracking system allows you to view the state of all cluster network interfaces from a cluster management application, such as Cluster Administrator.

Cluster Nodes

A cluster node is a system in a server cluster that has a working installation of the Windows operating system and the Cluster Service.

Cluster nodes have the following characteristics:

- Every node is attached to one or more cluster storage devices. Each cluster storage device attaches to one or more disks. The disks store all of the cluster's configuration and resource data. Each disk can be owned by only one node at any point in time, but ownership can be transferred between nodes. The result is that each node has access to all cluster configuration data.
- Every node communicates with the other nodes in the cluster through one or more NICs that attach nodes to networks.
- Every node in the cluster is aware of another system joining or leaving the cluster.
- Every node in the cluster is aware of the resources that are running on all nodes in the cluster.
- All nodes in the cluster are grouped under a common cluster name, which is used when accessing and managing the cluster.

Table 7-1 defines various states of a node that can occur in cluster operation.

Table 7-1. Node States and Definitions

State	Definition
Down	The node is not actively participating in cluster operations.
Joining	The node is in the process of becoming an active participant in the cluster operations.
Paused	The node is actively participating in cluster operations but cannot take ownership of resource groups and cannot bring resources online.
Up	The node is actively participating in all cluster operations, including hosting cluster groups.
Unknown	The state cannot be determined.

When the Cluster Service is configured for the first time on a node, the administrator must choose whether that node forms its own cluster or joins an existing cluster. When the Cluster Service is started on a node, that node searches for other active nodes on networks enabled for internal cluster communications.

Forming a New Cluster

If a cluster cannot be joined, the node attempts to form the cluster by gaining control of the quorum resource. If the node gains control of the quorum resource, the node forms the cluster and uses the recovery logs in the quorum resource to update its cluster database. The Cluster Service maintains a consistent, updated copy of the cluster database on all active nodes.

Joining an Existing Cluster

A node can join an existing cluster if it can communicate with another cluster node. If a cluster exists and the joining node finds an active node, it attempts to join that node's cluster. If it succeeds, the Cluster Service then validates the node's name and verifies version compatibility. If the validation process succeeds, the node joins the cluster. The node is updated with the latest copy of the cluster database.

Groups

A group is a collection of cluster resources with the following characteristics:

- All of the resources in the group are moved to the alternate node when one resource in a group fails and it is necessary to move the resource to an alternate node.
- A group is always owned by one node at any point in time, and a resource is always a member of a single group. Therefore, all of a group's resources reside on the same node.

Groups enable resources to be combined into larger logical units. Typically a group is made up of related or dependent resources, such as applications and their associated peripherals and data. However, groups can also be established with resources that are unrelated and nondependent to balance the load or for administrative convenience.

Every group maintains a prioritized list of the nodes that can and should act as its host. The preferred nodes list is generated by the Cluster Service. Cluster Service produces a list of preferred nodes for a group from the list of possible owners that is maintained by the group's resources and can be modified by an Administrator.

To maximize the processing power of a cluster, establish at least as many groups as there are nodes in the cluster.

Cluster Resources

A cluster resource is any physical or logical component that has the following characteristics:

- Can be brought online and taken offline
- Can be managed in a server cluster
- Can be hosted (owned) by only one node at a time

To manage resources, the Cluster Service communicates to a resource DLL through a Resource Monitor. When the Cluster Service makes a request of a resource, the Resource Monitor calls the appropriate entry-point function in the resource DLL to check and control the resource's state.

Dependent Resources

A dependent resource requires another resource to operate. For example, a network name must be associated with an IP address. Because of this requirement, a network name resource is dependent on an IP address resource. A resource can specify one or more resources on which it is dependent. A resource can also specify a list of nodes on which it is able to run. Preferred nodes and dependencies are important considerations when administrators organize resources into groups.

Dependent resources are taken offline before the resources upon which they depend are taken offline, likewise, they are brought online after the resources on which they depend are brought online.

Setting Resource Properties

Using the resource Properties dialog box, you can perform the following tasks:

- View or change the resource name
- View or change the resource description and possible owners
- Assign a separate memory space for the resource
- View the resource type, group ownership, and resource state
- View which node currently owns the resource
- View pre-existing dependencies and modify resource dependencies
- Specify whether to restart a resource and the settings used to restart the resource (if required)
- Check the online state of the resource by configuring the **Looks Alive** and **Is Alive** polling intervals in the Cluster Service
- Specify the time requirement for resolving a resource in a pending state (**Online Pending** or **Offline Pending**) before the Cluster Service places the resource in **Offline** or **Failed** status
- Set specific resource parameters

The General, Dependencies, and Advanced tabs are the same for every resource. Some resource types support additional tabs.

Properties of a cluster object should not be updated on multiple nodes simultaneously. See the MSCS online documentation for more information.

Resource Dependencies

Groups function properly only if resource dependencies are configured correctly. The Cluster Service uses the dependencies list when bringing resources online and offline. For example, if a group in which a physical disk and a file share are located is brought online, the physical disk containing the file share must be brought online before the file share.

Table 7-2 shows resources and their dependencies. The resources in the right column must be configured before you create the resource.

Table 7-2. Cluster Resources and Required Dependencies

Resource	Required Dependencies
File share	Network name (only if configured as a distributed file system [DFS] root)
IP address	None
Network name	IP address that corresponds to the network name
Physical disk	None

Setting Advanced Resource Properties

You can configure the advanced resource properties using the **Advanced** tab in the resource **Properties** dialog box. Use the **Advanced** tab to have the Cluster Service perform the following tasks:

- Restart a resource or allow the resource to fail.
    - To restart the resource, select **Affect the group** (if applicable).
    - To failover the resource group to another cluster node when the resource fails, select **Affect the group** and then enter the appropriate values in **Threshold** and **Period**. If you do not select **Affect the group**, the resource group will not failover to the healthy cluster node.

  The **Threshold** value determines the number of attempts by the Cluster Service to restart the resource before the resource fails over to a healthy cluster node.

  The **Period** value assigns a time requirement for the **Threshold** value to restart the resource.
- Adjust the time parameters for **Looks Alive** (general check of the resource) or **Is Alive** (**detailed check of the resource**) to determine if the resource is in the online state.
- Select the default number for the resource type.

  To apply default number, select **Use resource type value**.
- Specify the time parameter for a resource in a pending state (**Online Pending** or **Offline Pending**) to resolve its status before moving the resource to **Offline** or **Failed** status.

Resource Parameters

The **Parameters** tab in the **Properties** dialog box is available for most resources. Table 7-3 lists each resource and its configurable parameters.

Table 7-3. Resources and Configurable Parameters

Resource	Configurable Parameters
File share	Share permissions and number of simultaneous users Share name (clients will detect the name in their browse or explore lists) Share comment Shared file path
IP address	IP address Subnet mask Network parameters for the IP address resource (specify the correct cluster network)
Network name	System name
Physical disk	Drive for the physical disk resource (the drive cannot be changed after the resource is created)

Quorum Resource

The quorum resource is a common resource in the cluster that is accessible by all of the cluster nodes. Normally a physical disk on the shared storage, the quorum resource maintains data integrity, cluster unity, and cluster operations—such as forming or joining a cluster—by performing the following tasks:

- **Enables a single node to gain and defend its physical control of the quorum resource** — When the cluster is formed or when the cluster nodes fail to communicate, the quorum resource guarantees that only one set of active, communicating nodes is allowed to form a cluster.
- **Maintains cluster unity** — The quorum resource allows cluster nodes that can communicate with the node containing the quorum resource to remain in the cluster. If a cluster node fails for any reason and the cluster node containing the quorum resource is unable to communicate with the remaining nodes in the cluster, MSCS automatically shuts down the node that does not control the quorum resource.
- **Stores the most current version of the cluster configuration database and state data** — If a cluster node fails, the configuration database helps the cluster recover a failed resource or recreate the cluster in its current configuration.

The physical disk is the only type of resource supported by the NAS cluster solution that can act as a quorum resource.

NOTE: The Majority Node Set Quorum resource type is not supported.

Using the Quorum Resource for Cluster Integrity

The quorum resource is also used to ensure cluster integrity by performing the following functions:

- Maintaining the cluster node database
- Ensuring cluster unity

When a node joins or forms a cluster, the Cluster Service must update the node's private copy of the cluster database. When a node joins an existing cluster, the Cluster Service can retrieve the data from the other active nodes. However, when a node forms a cluster, no other node is available. The Cluster Service uses the quorum resource's recovery logs to update the node's cluster database, thereby maintaining the correct version of the cluster database and ensuring that the cluster is intact.

For example, if node 1 fails, node 2 continues to operate, writing changes to the cluster database. Before you can restart node 1, node 2 fails. When node 1 becomes active, it updates its private copy of the cluster database with the changes made by node 2 using the quorum resource's recovery logs to perform the update.

To ensure cluster unity, the operating system uses the quorum resource to ensure that only one set of active, communicating nodes is allowed to operate as a cluster. A node can form a cluster only if it can gain control of the quorum resource. A node can join a cluster or remain in an existing cluster only if it can communicate with the node that controls the quorum resource.

For example, if the private network (cluster interconnect) between cluster nodes 1 and 2 fails, each node assumes that the other node has failed, causing both nodes to continue operating as the cluster. If both nodes were allowed to operate as the cluster, the result would be two separate clusters using the same cluster name and competing for the same resources. To solve this problem, MSCS uses the node that owns the quorum resource to maintain cluster unity and solve this problem. In this scenario, the node that gains control of the quorum resource is allowed to form a cluster, and the other fails over its resources and becomes inactive.

Resource Failure

A failed resource is not operational on the current host node. At periodic intervals, the Cluster Service checks to see if the resource appears operational by periodically invoking the Resource Monitor. The Resource Monitor uses the resource DLL for each resource to detect if the resource is functioning properly. The resource DLL communicates the results back through the Resource Monitor to the Cluster Service.

Adjusting the Poll Intervals

You can specify how frequently the Cluster Service checks for failed resources by setting the **Looks Alive (general resource check)** and **Is Alive (detailed resource check)** poll intervals. The Cluster Service requests a more thorough check of the resource's state at each **Is Alive** interval than it does at each **Looks Alive** interval; therefore, the **Is Alive** poll interval is typically longer than the **Looks Alive** poll interval.

NOTE: Do not adjust the **Looks Alive** and **Is Alive** settings unless instructed by technical support.

Adjusting the Threshold and Period Values

If the resource DLL reports that the resource is not operational, the Cluster Service attempts to restart the resource. You can specify the number of times the Cluster Service can attempt to restart a resource in a given time interval. If the Cluster Service exceeds the maximum number of restart attempts (**Threshold** value) within the specified time period (**Period** value), and the resource is still not operational, the Cluster Service considers the resource to be failed.

NOTE: See "Setting Advanced Resource Properties" to configure the **Looks alive**, **Is alive, Threshold,** and **Period** values for a particular resource.

NOTE: Do not adjust the **Threshold** and **Period** values settings unless instructed by technical support.

Configuring Failover

You can configure a resource to failover an entire group to another node when a resource in that group fails for any reason. If the failed resource is configured to cause the group that contains the resource to failover to another node, Cluster Service will attempt a failover. If the number of failover attempts exceeds the group's threshold and the resource is still in a failed state, the Cluster Service will attempt to restart the resource. The restart attempt will be made after a period of time specified by the resource's **Retry Period On Failure** property, a property common to all resources.

When you configure the **Retry Period On Failure** properly, consider the following guidelines:

- Select a unit value of minutes, rather than milliseconds (the default value is milliseconds).
- Select a value that is greater or equal to the value of the resource's restart period property. This rule is enforced by the Cluster Service.

NOTE: Do not adjust the **Retry Period On Failure** settings unless instructed by technical support.

Resource Dependencies

A dependent resource requires—or depends on—another resource to operate. For example, if a Generic Application resource requires access to clustered physical storage, it would depend on a physical disk resource.

The following terms describe resources in a dependency relationship:

- **Dependent resource** — A resource that depends on other resources (the dependencies).
- **Dependency** — A resource on which another resource depends.
- **Dependency tree** — A series of dependency relationships such that resource A depends on resource B, resource B depends on resource C, and so on.

  Resources in a dependency tree obey the following rules:

  - A dependent resource and all of its dependencies must be in the same group.
  - The Cluster Service takes a dependent resource offline before any of its dependencies are taken offline, and brings a dependent resource online after all its dependencies are online, as determined by the dependency hierarchy.

Creating a New Resource

Before you add a resource to your NAS cluster, you must verify that the following elements exist in your cluster:

- The type of resource is either one of the basic types provided with MSCS or a custom resource type provided by the application vendor, Microsoft, or a third party vendor.
- A group that contains the resource already exists within your cluster.
- All dependent resources have been created.
- A separate Resource Monitor—recommended for any resource that has caused problems in the past.

To create a new resource:

1. Click the **Start** button and select **Programs**→ **Administrative Tools**→ **Cluster Administrator**.

   The **Cluster Administrator** window appears.
2. In the console tree (usually the left pane), double-click the **Groups** folder.
3. In the details pane (usually the right 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**, and click the appropriate information in **Resource type** and **Group**.
6. Click **Next**.
7. Add or remove possible owners of the resource, and then click **Next**.

8. The **New Resource** window appears with **Available resources** and **Resource dependencies** selections.

   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**.

9. Repeat step 7 for any other resource dependencies, and then click **Finish**.
10. Set the resource properties.

    For more information on setting resource properties, see the MSCS online help.

Deleting a Resource

1. Click the **Start** button and select **Programs**→ **Administrative Tools**→ **Cluster Administrator**.

   The **Cluster Administrator** window appears.

2. In the console tree (usually the left pane), click the **Resources** folder.
3. In the details pane (usually the right pane), click the resource you want to remove.
4. In the **File** menu, click **Delete**.

When you delete a resource, Cluster Administrator also deletes all the resources that have a dependency on the deleted resource.

File Share Resources

Creating a Cluster-Managed File Share

1. Launch **Windows Explorer**.
2. On a shared volume, create a new folder for the file share.

   NOTE: Do not create a share for this folder.

3. Right-click the folder and select **Properties**.
4. In the **Properties** window, click the **Security** tab.
5. In the **Group or users names** box, verify that the **Cluster Service** account has **Full Control** rights to this folder for the NTFS file system.
6. Close **Windows Explorer**.
7. Click the **Start** button and select **Programs**→ **Administrative**→ **Tools**→ **Cluster Administrator**.
8. In the Cluster Administrator left window pane, ensure that a physical disk resource exists in the cluster.
9. In the Cluster Administrator left or right window pane, right-click and select **New**→ **Resource**.

10. In the **New Resource** window, perform the following steps:
    a. In the **Name** field, type a name for the new share.
    b. In the **Description** field, type a description of the new share (if required).
    c. In the **Resource type** drop-down menu, select **File Share**.
    d. In the **Group** drop-down menu, select the appropriate virtual server for your file share.
11. Click **Next**.

    The **Possible Owners** window appears.
12. Select the appropriate cluster node(s) in the **Available nodes** box on which this resource can be brought online.
13. Click the **Add** button to move the cluster node(s) to the **Possible owners** menu.
14. Click **Next**.

    The **Dependencies** window appears.
15. In the **Available resources** menu, select the appropriate resource dependencies which must be brought online first by the Cluster Service.
16. Click the **Add** button to move the resources to the **Resource dependencies** menu.
17. Click **Next**.

    The **File Share Parameters** window appears.
18. Perform the following steps:
    a. In the **Share name** field, type the name of the file share.
    b. In the **Path** field, type the path to the file share.
    c. In the **Comment** field, enter any additional information about the file share (if required).
    d. Click **Permissions** and apply the appropriate group or user names and permissions for the file share (if required), and then click **OK**.
    e. Click **Advanced** and select the appropriate file share properties (if required), and then click **OK**.

    See "File Share Resource Types" for more information.
19. Click **Finish**.

    The Cluster Administrator window appears.
20. In the right window pane, right-click the share and select **Bring Online**.

Deleting a File Share

1. Click the **Start** button and select **Programs**→ **Administrative**→ **Tools**→ **Cluster Administrator**.
2. In the **Cluster Administrator** window console tree, click the **Resources** folder.

3. In the right window pane, right-click the file share you want to remove and select **Delete**.

NOTE: When you delete a resource, Cluster Administrator automatically deletes all the resources that have a a dependency on the deleted resource.

DFS File Shares

You can use the **File Share** resource type selection in Cluster Administrator to create a resource that manages a stand-alone DFS root; however, fault-tolerant DFS roots cannot be managed by this resource. The DFS root **File Share** resource has required dependencies on a network name and an IP address. The network name can be either the cluster name or any other network name for a virtual server.

A cluster-managed DFS root is different from an Active Directory (or domain-based) DFS root. If the data set does not change very often, using and replicating a domain-based DFS root can be a better selection than a cluster-managed DFS root for providing high availability. If the data set changes frequently, replication is not recommended, and a cluster-managed DFS root is the better solution.

Table 7-4 provides a summary for choosing the appropriate DFS root management scheme.

See the *Dell PowerVault 77xN NAS Systems Administrator's Guide* for more information.

Table 7-4. Selecting the Appropriate DFS Root Management Scheme

Data Set Activity	DFS Root Management
Data changes often	Domain-based
Data does not change very often	Cluster-managed

NOTE: Microsoft Windows Storage Server 2003, Enterprise Edition supports multiple stand-alone DFS roots. The DFS roots can exist in multiple resource groups and each group can be hosted on a different node in the cluster.

File Share Resource Types

If you want to use a Dell™ PowerVault™ NAS cluster as a high-availability file server, you will need to select the type of file share for your resource. Three ways to use this resource type are available:

- **Basic file share** — Publishes a single file folder to the network under a single name.
- **Share subdirectories** — Publishes several network names—one for each file folder and all of its immediate subfolders. This method is an efficient way to create large numbers of related file shares on a single file server.

  For example, you can create a file share for each user with files on the cluster node.
- **DFS root** — Creates a resource that manages a stand-alone DFS root. Fault tolerant DFS roots cannot be managed by this resource. A DFS root file share resource has required dependencies on a network name and an IP address. The network name can be either the cluster name or any other network name for a virtual server.

Enabling Cluster NFS File Share Capabilities

After you add a node to the cluster, enable the NFS file sharing capabilities by performing the following steps.

NOTE: Perform this procedure on one cluster node after you configure the cluster.

1. Open a command prompt.
2. At the prompt, type:

   ```
   c:\dell\util\cluster
   ```

3. In the **cluster** directory, run the **NFSShareEnable.bat** file.

Failover and Failback

This section provides information about the failover and failback capabilities of the Cluster Service.

Failover

When an individual application or user resource (also known as a cluster resource) fails on a cluster node, the Cluster Service will detect the application failure and try to restart the application on the cluster node. If the restart attempt reaches a preset threshold, the Cluster Service brings the running application offline, moves the application and its resources to another cluster node, and restarts the application on the other cluster node. This process of automatically moving resources from a failed cluster node to another healthy cluster node is called *failover*.

In order to failover and failback running applications, cluster resources are placed together in a group so the Cluster Service can move the cluster resources as a combined unit. This process ensures that the failover and/or failback procedures transfers all of the user resources as transparently as possible.

After failover, the Cluster Administrator can reset the following recovery policies:

- Application dependencies
- Application restart on the same cluster node
- Workload rebalancing (or failback) when a failed cluster node is repaired and brought back online

Failover Process

The Cluster Service attempts to fail over a group when any of the following conditions occur:

- The node currently hosting the group becomes inactive for any reason.
- One of the resources within the group fails, and it is configured to affect the group.
- Failover is forced by the System Administrator.

When a failover occurs, the Cluster Service attempts to perform the following procedures:

- The group's resources are taken offline.

  The resources in the group are taken offline by the Cluster Service in the order determined by the group's dependency hierarchy: dependent resources first, followed by the resources on which they depend.

For example, if an application depends on a Physical Disk resource, the Cluster Service takes the application offline first, allowing the application to write changes to the disk before the disk is taken offline.

- The resource is taken offline.

  Cluster Service takes a resource offline by invoking, through the Resource Monitor, the resource DLL that manages the resource. If the resource does not shut down within a specified time limit, the Cluster Service forces the resource to shut down.

- The group is transferred to the next preferred host node.

  When all of the resources are offline, the Cluster Service attempts to transfer the group to the node that is listed next on the group's list of preferred host nodes.

  For example, if cluster node 1 fails, the Cluster Service moves the resources to the next cluster node number, which is cluster node 2.

- The group's resources are brought back online.

  If the Cluster Service successfully moves the group to another node, it tries to bring all of the group's resources online. Failover is complete when all of the group's resources are online on the new node.

The Cluster Service continues to try and fail over a group until it succeeds or until the number of attempts occurs within a predetermined time span. A group's failover policy specifies the maximum number of failover attempts that can occur in an interval of time. The Cluster Service will discontinue the failover process when it exceeds the number of attempts in the group's failover policy.

Modifying Your Failover Policy

Because a group's failover policy provides a framework for the failover process, make sure that your failover policy is appropriate for your particular needs. When you modify your failover policy, consider the following guidelines:

- Define the method in which the Cluster Service detects and responds to individual resource failures in a group.
- Establish dependency relationships between the cluster resources to control the order in which the Cluster Service takes resources offline.
- Specify **Time-out**, failover **Threshold**, and failover **Period** for your cluster resources
  - **Time-out** controls how long the Cluster Service waits for the resource to shut down.
  - **Threshold** and **Period** control how many times the Cluster Service attempts to fail over a resource in a particular period of time.
- Specify a **Possible owner list** for your cluster resources. The **Possible owner list** for a resource controls which cluster nodes are allowed to host the resource.

Failback

When the System Administrator repairs and restarts the failed cluster node, the opposite process may occur. After the original cluster node has been restarted and rejoins the cluster, the Cluster Service will bring the running application and its resources offline, move them from the failover cluster node to the original cluster node, and then restart the application. This process of returning the resources back to their original cluster node is called failback.

You can configure failback to occur immediately, at any given time, or not at all. However, be sure to configure the failback time during your offpeak hours to minimize the effect on user, as they may see a delay in service until the resources come back online.

Back to Contents Page

Back to Contents Page

Troubleshooting

Dell™ PowerVault™ NAS Fibre Channel Cluster Systems Installation and Troubleshooting Guide

This appendix provides troubleshooting information for NAS cluster configurations.

Table A-1 describes general cluster problems you may encounter and the probable causes and solutions for each problem.

Table A-1. General Cluster Troubleshooting

Problem	Probable Cause	Corrective Action
The nodes cannot access the storage system, or the cluster software is not functioning with the storage system.	The storage system is not cabled properly to the nodes or the cabling between the storage components is incorrect. The length of the interface cables exceeds the maximum allowable length. One of the cables is faulty. **Access Control** is not enabled correctly.	Ensure that the cables are connected properly from the node to the storage system. See "Cabling Your Public and Private Networks" for more information. Ensure that the fiber optic cables do not exceed 300 m (multimode) or 10 km (single mode switch-to-switch connections only). Replace the faulty cable. Verify the following: • All switched zones are configured correctly. • The EMC® Access Logix™ software is installed on the storage system. • All LUNs and hosts are assigned to the proper storage groups.
One of the nodes takes a long time to join the cluster.	The node-to-node network has failed due to a cabling or hardware failure. Long delays in node-to-node communications may be normal.	Check the network cabling. Ensure that the node-to-node interconnection and the public network are connected to the correct NICs. Verify that the nodes can communicate with each other by running the ping command from each node to the other node. Try both the host name and IP address when using the ping command.
Attempts to connect to a cluster using Cluster Administrator fail.	The Cluster Service has not been started. A cluster has not been formed on the system. The system has just been booted and services are still starting.	Verify that the Cluster Service is running and that a cluster has been formed. Use the Event Viewer and look for the following events logged by the Cluster Service: Microsoft Cluster Service successfully formed a cluster on this node. or Microsoft Cluster Service successfully joined the cluster. If these events do not appear in Event Viewer, see the *Microsoft Cluster Service Administrator's Guide* for instructions on setting up the cluster on your system and starting the Cluster Service.
Using Microsoft® Windows NT® 4.0 to remotely administer a cluster generates error messages.	Normal. Some resources in Windows® Storage Server 2003, Enterprise Edition are not supported in Windows NT 4.0.	Dell strongly recommends that you use Windows XP Professional or Windows Server 2003 for remote administration of a NAS cluster.
Unable to add a node to the	The new node cannot access the	Ensure that the new cluster node can enumerate the

cluster.	shared disks. The shared disks are enumerated by the operating system differently on the cluster nodes.	cluster disks using Windows Disk Administration. If the disks do not appear in Disk Administration, check the following: • Check all cable connections. • Check all zone configurations. • Check the Access Control settings on the attached storage systems. Use the "Advanced" with "Minimum" option.
The disks on the shared cluster storage appear unreadable or uninitialized in Windows Disk Administration	This situation is normal if you stopped the Cluster Service or if the cluster node does not own the cluster disk.	No action required.
The **Create NFS Share** option does not exist.	The Enable NFS Share utility is not installed on one of the cluster nodes.	Run the Enable NFS File Share utility. See "Enabling Cluster NFS File Share Capabilities" for more information.

Back to Contents Page

Back to Contents Page

Cluster Data Sheets

Dell™ PowerVault™ NAS Fibre Channel Cluster Systems Installation and Troubleshooting Guide

The configuration matrix and data sheets on the following pages are provided for the system installer to record pertinent information about Dell™ PowerVault™ NAS cluster configurations. Make a copy of the appropriate data sheet to use for the installation or upgrade, complete the requested information on the sheet, and have the completed sheet available if you need to call Dell for technical assistance. If you have more than one cluster, complete a copy of the form for each cluster.

NAS Cluster Configuration Matrix Form

You can attach the following form to the back of each cluster node or rack. The system installer may want to use the form to record important information about the hardware on each cluster component. Make additional copies as needed for each cluster node and have copies of each form available any time you call Dell for technical support.

Cluster Type	NAS Cluster Solution
Cluster name	
Installer	
Date installed	
Applications	
Location	
Notes	

Node (Server Name)	Server Type	Cluster Name	Service Tag Number

Storage Array	Description of Installed Items (Drive letters, RAID types, applications/data)	Service Tag Number
Storage		
Storage		
Storage		
Storage		

System	Storage Port Connection on Server
Node 1, HBA 1	
Node 1, HBA 2	
Node 2, HBA 1	
Node 2, HBA 2	

Zoning Configuration Matrix Form

Switch	Storage Port	Storage 1	Storage 2	Storage 3	Storage 4

Back to Contents Page

Back to Contents Page

Abbreviations and Acronyms

Dell™ PowerVault™ NAS Fibre Channel Cluster Systems Installation and Troubleshooting Guide

AC

alternating current

BIOS

basic input/output system

CD

compact disc

DAE

disk array enclosure

DAE2-OS

disk array enclosure 2 (Gb)—operating system

DDNS

dynamic domain naming system

DFS

distributed file system

DHCP

dynamic host configuration protocol

DLL

dynamic link library

DLT

digital linear tape

DNS

domain naming system

DPE

disk processor enclosure

ESD

electrostatic discharge

GB

gigabyte

Gb

gigabit

GBIC

gigabit interface converter

Gb/s

gigabits per second

GMT

Greenwich Mean Time

GUI

graphical user interface

HBA

host bus adapter

Hz

hertz

ID

identification

IIS

Internet Information Server

I/O

input/output

IP

Internet Protocol

ISL

interswitched link

km

kilometers

LAN

local area network

LCC

link control card

LUN

logical unit number

m

meter

MB

megabyte(s)

MB/sec

megabyte(s) per second

MHz

megahertz

MSCS

Microsoft® Cluster Service

MS-DOS

Microsoft Disk Operating System

NetBIOS

network basic input/output system

NLB

network load balancing

NIC

network interface controller

NTFS

NT File System

OS

operating system

PDU

power distribution unit

PERC

PowerEdge™ Expandable RAID Controller

PCI

Peripheral Component Interconnect

PCI-X

Peripheral Component Interconnect eXtended

POST

power-on self-test

PSM

persistent storage manager

RAID

redundant array of independent disks

RAM

random access memory

SAN

storage area network

SCSI

small computer serial interface

SES

SCSI enclosure services

SFP

small form-factor pluggable

SMP

symmetric multiprocessing

SP

storage processor

SPE

storage processor enclosure

SPS

standby power supply

SQL

Structured Query Language

SNMP

Simple Network Management Protocol

TB

terabyte

TCP/IP

Transmission Control Protocol/Internet Protocol

UPS

uninterruptible power supply

VLAN

virtual local area network

Back to Contents Page