COMPAQ

White Paper

March 2001
14AK-0301A-WWEN
Prepared by:
Industry Standard Server Division

Compaq Computer Corporation

Contents

TaskSmart N-Series Cluster Administration

Abstract: The material contained in this abbreviated administration guide is intended as a supplement to the complete *Compaq TaskSmart N2400 Administration Guide*. After an introduction on clustering, sections on basic and advanced administration tasks describe differences in the operation of a Compaq TaskSmart N-Series Cluster as compared to a single TaskSmart N2400 appliance. These sections also provide an overview of the hardware and software components that are unique to the TaskSmart N-Series Cluster. Sections on hardware failure and troubleshooting detail the processes and expected behavior for various types of failures, and provide basic troubleshooting assistance.

Notice

Compaq shall not be liable for technical or editorial errors or omissions contained herein. The information in this document is provided "as is" without warranty of any kind and is subject to change without notice. The warranties for Compaq products are set forth in the express limited warranty statements accompanying such products. Nothing herein should be construed as constituting an additional warranty.

Introduction

General Overview

This document provides a starting point for administration of the Compaq *TaskSmart*™ N-Series Cluster. It covers cluster-specific administrative procedures that have been added to the Compaq TaskSmart N2400 appliance. This document contains information on additions such as SecurePath and *StorageWorks*™ Command Console, as well as information on how clustering affects previous TaskSmart N-Series software such as Compaq *SANworks*™ Virtual Replicator. For more information regarding standard administrative tasks for the TaskSmart N-Series Cluster, see the *Compaq TaskSmart N2400 Administration Guide*.

Key Benefits

The TaskSmart N-Series Cluster has the following benefits over general-purpose servers and nonclustered file servers:

- Reduces planned and unplanned fileserver downtime
- Achieves higher aggregate performance with active/active clustering
- Achieves much greater disk capacity over a single TaskSmart N2400 appliance
- Works seamlessly with industry-standard backup, virus protection, and management applications
- Uses standard disks and enclosures that ship with the TaskSmart N2400 appliance

Key Features

The TaskSmart N-Series Cluster has the following standard features:

- Dual active cluster nodes serving Common Internet File System (CIFS) protocol (Windows) clients
- Familiar Microsoft Cluster Service interface
- Optimized industry-standard operating system (OS): Microsoft Windows Powered OS
- Storage pooling and virtualization that promote simplified storage management
- Snapshots that provide temporary, point-in-time copies of data
- Enterprise-class storage subsystem and industry-standard clustering technology united in a fileserver appliance with the ability to endure failures of disks, storage enclosures, SCSI cables, RAID controllers, SAN switches, fiber connects, Fibre Channel controllers, and complete system OS or hardware failure on a single cluster node

References

This document references the following documents and websites:

- *Compaq TaskSmart N-Series Cluster Installation Guide*
- *Compaq TaskSmart N-Series Cluster Planning Guide*

- *Compaq TaskSmart N2400 Administration Guide*
- *Compaq SANworks Virtual Replicator Version 2.0 System Administrator's Guide*
- *SecurePath Version 3.1 for Microsoft Windows Installation and Reference Guide*

  This guide is available on the following website:

  www.compaq.com/products/storageworks/techdoc/storagemgtsoftware/AA-RL4SC-TE.html
- *HSG80 Array Controller ACS Version 8.5 CLI Reference Guide*

  This guide is available on the following website:

  www5.compaq.com/products/storageworks/techdoc/raidstorage/EK-HSG85-RG-A01.html
- *Compaq StorageWorks Command Console V2.3 User Guide*

  This guide is available on the following website:

  storage.inet.cpqcorp.net/Document_Storage/TechDocs/AA-RFA2G-TE.pdf
- *Windows Clustering Technologies: Cluster Service Architecture*

  This guide is available on the following Microsoft website:

  www.microsoft.com/WINDOWS2000/library/howitworks/cluster/clusterarch.asp

Note: Some information in these documents is of a general nature, and does not pertain to the TaskSmart N-Series Cluster.

Cluster Administration

Cluster Overview

The TaskSmart N-Series Cluster is based on Microsoft Cluster Service (MSCS). This section provides a brief overview of MSCS, and includes definitions of terms used in the administration of the cluster. This information is provided to facilitate the administration of the TaskSmart N-Series Cluster. For more in-depth information, refer to the Compaq and Microsoft white papers on clustering.

The most basic part of a cluster is the server heads. A server head must be a TaskSmart N-Series Cluster appliance. In MSCS, the server heads are referred to as nodes.

Hardware and software components that are managed by the cluster service are called cluster resources. Resources have three defining characteristics:

- They can be brought online and taken offline.
- They can be managed in a server cluster.
- They can be owned by only one node at a time.

Examples of cluster resources are file shares, IP addresses, and network names.

A cluster group is defined as a collection of resources managed by the cluster service as a single, logical unit. A cluster group defines units of failover and dependencies between resources. Groups are always owned by only one node at any point in time. They enable resources to be combined into logical units.

Failover of cluster groups and resources happens when one of the following three events occurs:

- When a node hosting the group becomes inactive for any reason. For example, a node could become inactive due to a shutdown of cluster service or loss of power. The group is then failed over to the other node in the cluster.
- If all of the resources within the group are dependent on one resource and that resource fails, the entire group is failed over to the other node.
- An administrator can manually initiate a failover.

When a resource is failed over, the cluster service goes through certain steps. First, all of the resources are taken offline in the order defined by the resource dependencies. Secondly, the cluster service attempts to transfer the group to the next node on the preferred owners list. If the transfer is successful, the resources are brought online in accordance with the dependency structure of the resource.

When using MSCS, a failover policy is created that defines how the server cluster detects and responds to the failure of individual resources in the group. After a failover occurs and the cluster is brought back to its original state, failback can occur automatically based on the policy. After a previously failed node comes back online, the cluster service can fail back the groups to the original host. The failback policy must be set before the failover occurs for failback to work as intended.

Each cluster must have a shared disk called the Quorum Disk. This physical disk in the common cluster disk array plays a critical role in cluster operations. It offers a means of persistent storage. The disk must provide physical storage that can be accessed by any node in the cluster. The Quorum Disk maintains data integrity by the following methods:

- The Quorum Disk stores the most current version of the cluster database.
- The Quorum Disk guarantees that only one set of active communicating nodes is allowed to operate as a cluster.
- A node can establish the cluster if that node gains control of the Quorum Disk upon startup.
- A node can join or remain in the cluster if that node can communicate with the node that owns the Quorum Disk.

Each resource in the cluster is dependent on another resource to function, and may have resources dependent on it. A resource and all of its dependencies must be located in the same group so that if a resource fails over, all of its dependent resources fail over.

The nodes are connected to each other by a crossover cable, a network switch, or a network hub. This connection allows communication between the nodes to track the state of each cluster node. Each node sends out periodic messages to the other node. These messages are called heartbeats. If a node stops sending messages, the cluster service will fail over any resources that the node owns. For example, if the node that owns the Quorum Disk is shut down for any reason, its heartbeat will stop. The other node detects the lack of the heartbeat and takes over ownership of the Quorum Disk and the cluster. In case the cluster interconnect fails, Compaq recommends setting up the public network as a secondary choice for the heartbeat. However, this setup leads to increased traffic on the public network.

Figure 1 shows how the TaskSmart N-Series Cluster should be set up.

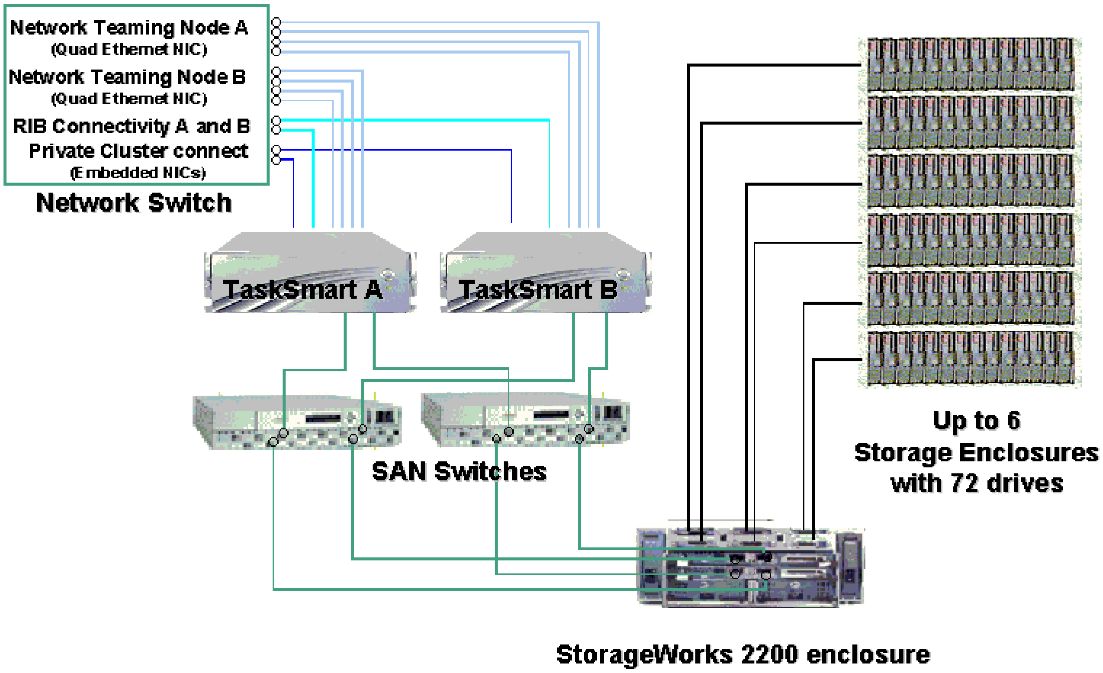


Figure 1: TaskSmart N-Series Cluster diagram

Cluster Administrator

The Microsoft Cluster Administrator is an application for configuring, controlling, and monitoring the TaskSmart N-Series Cluster. Cluster Administrator provides information about cluster groups and resources. Administrators can use Cluster Administrator to create cluster groups, create virtual servers, create file share resources, manage cluster objects (including performing a manual failover), and monitor cluster activity.

Creating Cluster Groups

A default cluster group is automatically created when the cluster is first created. This default cluster group contains an Internet Protocol (IP) Address resource, a Network Name resource, and the Quorum Disk resource. When the new cluster is created, the (IP) address and the cluster name that were specified during setup are set up as the IP address and network name of this default cluster group.

IMPORTANT: Do not delete or rename the Cluster Group or IP Address, because doing so will result in losing the cluster and will require reinstallation of the cluster.

When initially creating groups, the first priority of the administrator is to gain an understanding of how to manage resource groups and their resources. Administrators may choose to create a resource group and a virtual server (IP Address resource and Network Name resource) for each node that will contain all resources owned by that node, OR the administrator may choose to create a resource group and virtual server for each pool created in SANworks Virtual Replicator (SWVR). For detailed information on SWVR resources, see the “SANworks Virtual Replicator” section later in this document.

Creating only one resource group and one virtual server for each node facilitates group and resource administration. This setup allows administrators to include all file share resources under one group. Clients access all of the resources owned by one node through a virtual server name.

Alternatively, creating one resource group and one virtual server for each pool (SCE Pool resource) that is created in SANworks Virtual Replicator allows administrators to dedicate groups to specific departments. For example, an administrator can create the following resource groups:

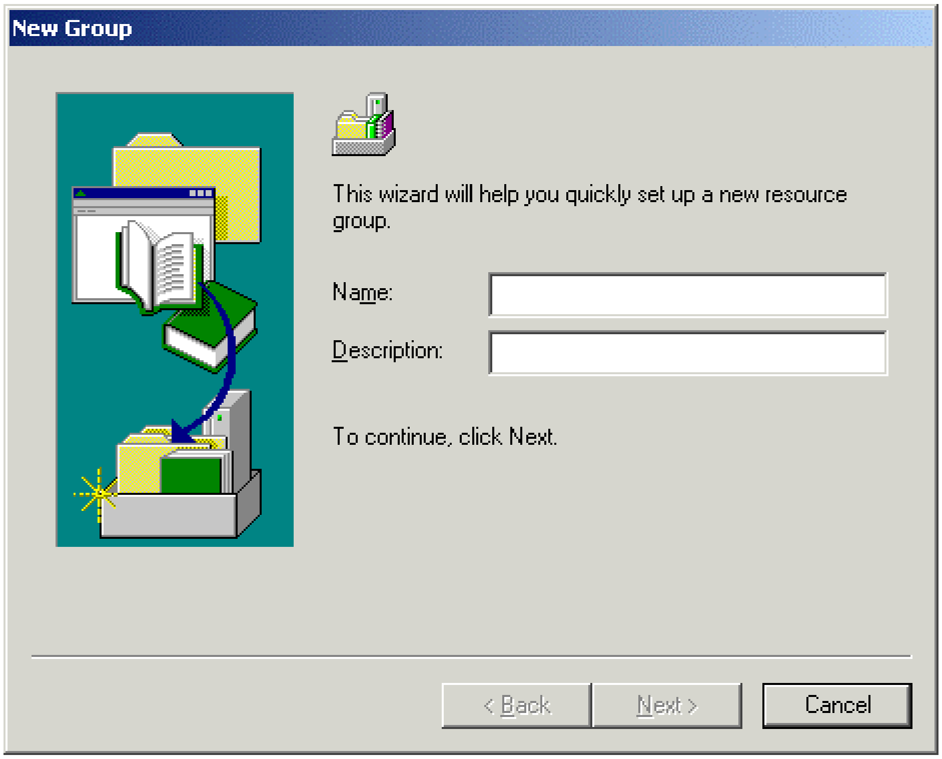
- Finance
- Marketing

The administrator can then provide each department with a unique IP address and network name. This way if one of the resource groups becomes unavailable, the other one is still available. This configuration provides administrators with more granular control of the resource groups. However, in this type of configuration, every time a new SWVR pool is created, a new virtual server must be created. This leads to performance degradation due to an increase in resource overhead.

The following section demonstrates how to create a resource group and a virtual server. These steps must to be repeated any time a group or resource is created.

Procedure for Creating a Resource Group

1. Start Cluster Administrator.
2. Select the cluster, right click the cluster, and then select **New Group**.


New Group

This wizard will help you quickly set up a new resource group.

Name:

Description:

To continue, click Next.

Figure 2: Creating a new resource group

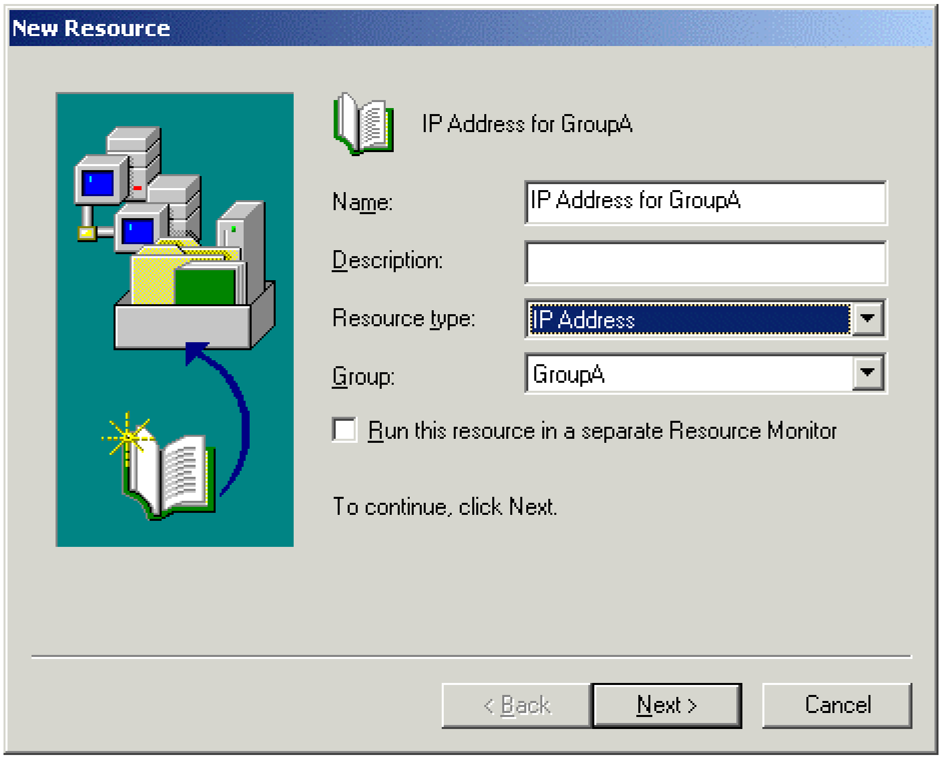
3. Enter the group name and its description.
4. Select **Preferred Owners**.

Note: The first node listed in preferred owners should be the node on which the group is being created. The preferred owners list can be reordered later.

Procedure for Creating an IP Resource

Note: Both nodes must be up and functioning normally before resources can be created.

1. From Cluster Administrator, select **New Resource.**
2. Enter a name for the resource.
3. Select **IP Address** from the **Resource type** drop-down menu.
4. Select the group you just created from the **Group** drop-down menu.



New Resource

IP Address for GroupA

Name: IP Address for GroupA

Description:

Resource type: IP Address

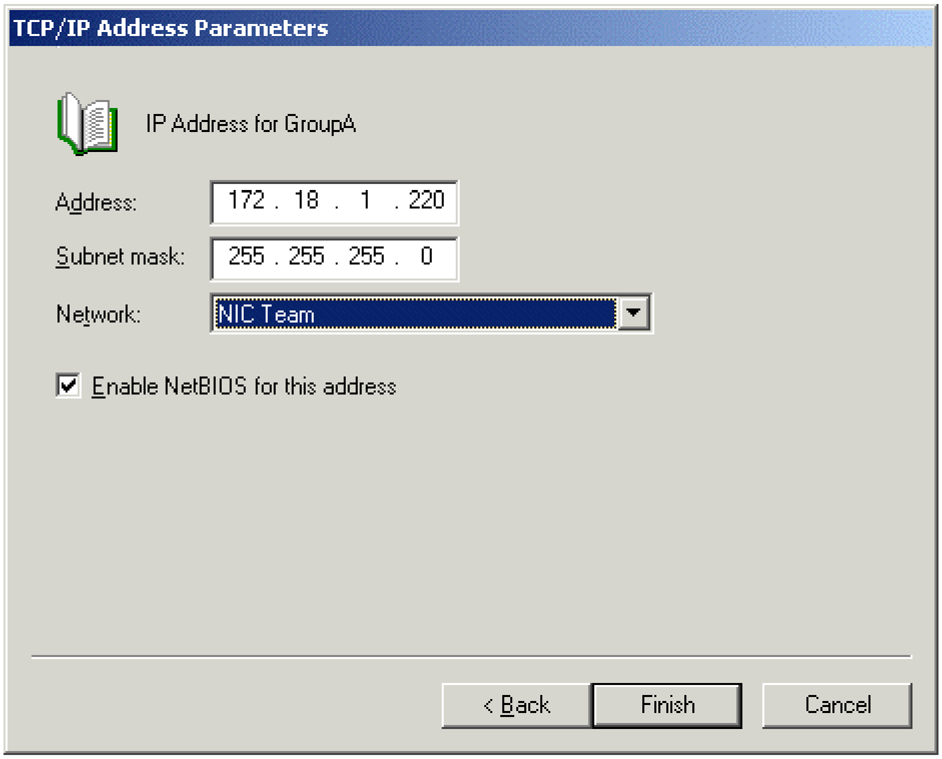
Group: GroupA

☐ Run this resource in a separate Resource Monitor

To continue, click Next.

Figure 3: Creating an IP Address resource

5. Click **Next** to display the **Possible Owners** screen. Both nodes must be listed as possible owners.
6. Click **Next** to display the **Dependencies** screen.
7. Click **Next** to display the **TCP/IP Address Parameters** screen.
8. Select the address and the subnet mask.
9. Select the public network from the **Network** drop-down menu.
10. Click **Finish**.



TCP/IP Address Parameters

IP Address for GroupA

Address: 172 . 18 . 1 . 220

Subnet mask: 255 . 255 . 255 . 0

Network: NIC Team

☑ Enable NetBIOS for this address

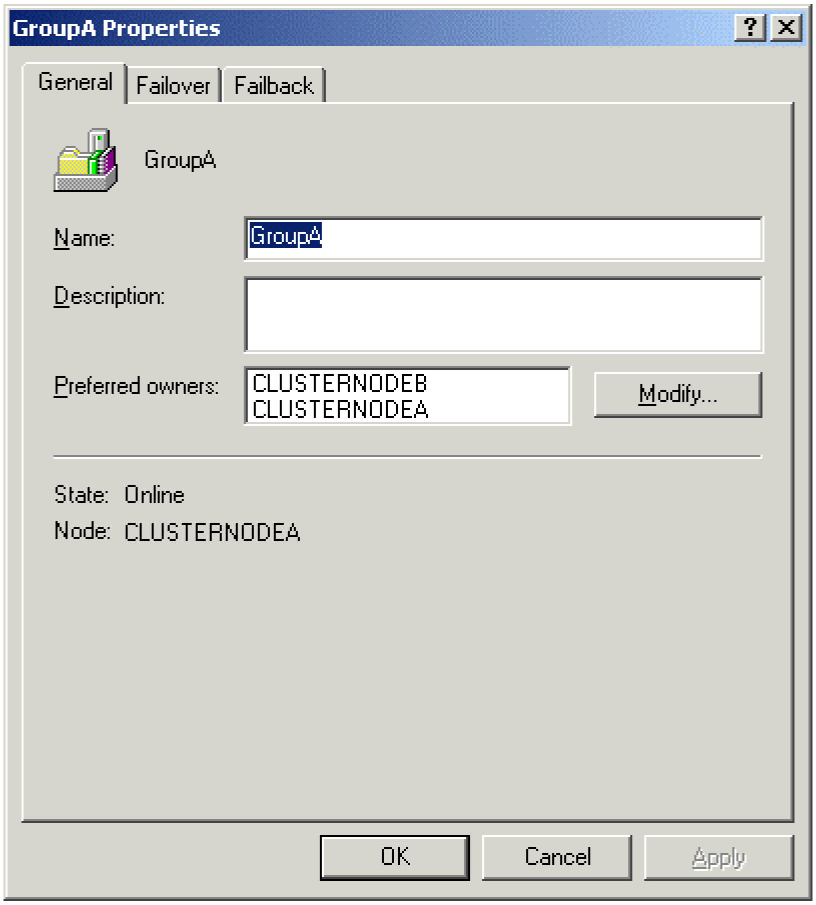
Figure 4: Configuring an IP address resource

Procedure for Creating a Network Name Resource

1. From Cluster Administrator, select **New Resource**.
2. Enter a name for the resource.
3. Select **Network Name** from the **Resource type** drop-down menu.
4. Select the group you just created from the **Group** drop-down menu.
5. Click **Next** in the **Possible Owners** screen.
6. Add the IP Address resource created previously to the **Resource Dependencies** drop-down menu in the **Dependencies** screen.
7. Enter a network name in the **Network Name** parameters screen.

Group Properties

Now that the group resource has been created, the administrator can configure the group properties. The Group Resource Properties screen provides users with three tabs. The tabs are **General**, **Failover**, and **Failback**. To display the group resource properties, right click the group in the **Name** field, and then select **Properties**.



GroupA

Name: GroupA

Description:

Preferred owners: CLUSTERNODEB
CLUSTERNODEA

State: Online

Node: CLUSTERNODEA

Figure 5: Group properties

General Information

Administrators may find it necessary to change group information such as the group name, description, and preferred owners. The **General** tab provides users with the ability to change this information.

IMPORTANT: Do not delete or rename the Cluster Group or IP Address, because doing so will result in losing the cluster and will require reinstallation of the cluster.

Failover

The **Failover** tab consists of two settings. (**Threshold** and **Period**) The **Threshold** setting is the maximum amount of times the group is allowed to fail over in the amount of hours set in the **Period** setting. If a group fails over the maximum amount of times in the hours set in the period setting, the group will be taken completely offline. The default setting for **Threshold** is ten times, and the default setting for **Period** is six hours. Users can adjust this setting to meet their organizational requirements.

Failback

The **Failback** tab has two settings that can be set by the user (**Prevent Failback** and **Allow Failback**). In the case of a node failure, the group will fail over to the available node. The **Prevent Failback** setting prevents the group resource from being moved back to its preferred node when the node becomes available. In this case, users will have to manually move the group resource back. The **Allow Failback** setting allows users to choose immediate failback or to set a time interval for failback.

The first setting in the **Failback** screen is **Immediately,** which allows Cluster Administrator to fail back the group resource as soon as the failed node has become available again. The **Failback Between** setting is used to set the time interval for failback. The numbers that can be set are between 0 and 23 for the beginning and end of the interval. The numbers correspond to the local time of the cluster group. The numbers are read on a 24-hour clock.

Note: Users can set the first number in the range to be higher than the second number if they want the interval to occur on the following day.

SANworks Virtual Replicator

Compaq SANworks Virtual Replicator (SWVR) is an application that provides advanced, centralized storage management capabilities for the TaskSmart N-Series Cluster solution. SWVR runs on each node in the cluster. SWVR is used to create pools, virtual disks, and snapshots.

Pool

Creating a pool from the Snapshot Manager is exactly the same as creating a pool in the TaskSmart N2400 appliance server. The only difference is that SWVR automatically creates a cluster group and an SCE (Storage and Clusters Extension) Pool resource within Cluster Administrator.

SCE Pool

Snapshot Manager creates an SCE Pool resource in Cluster Administrator whenever a new pool is created. This resource is required for SWVR to successfully work in a cluster environment. Compaq requires that administrators use Snapshot Manager to create pools. Do not use Cluster Administrator to create SCE Pool resources.

IMPORTANT: Do not rename any SCE Pool resource. This causes loss or corruption of data.

Virtual Disks

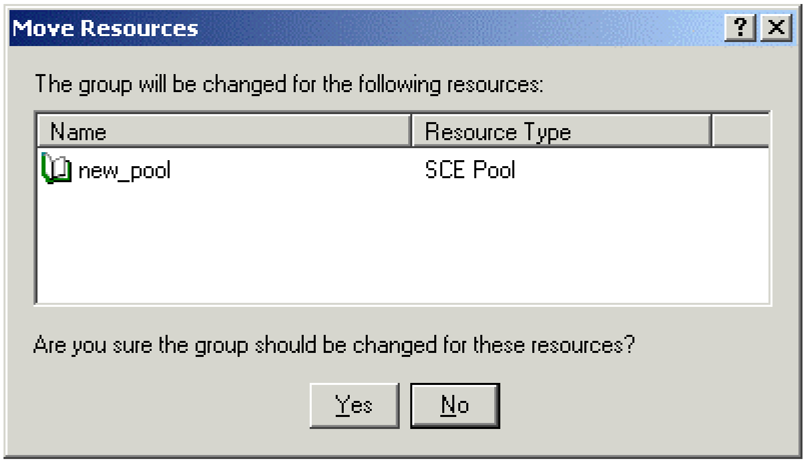
Creating a Virtual disk from the Snapshot Manager is exactly the same as creating a Virtual disk in the TaskSmart N2400 appliance server. Virtual disks do not create any cluster resources.

Snapshots

Creating a snapshot from the Snapshot Manager is exactly the same as creating a snapshot in the TaskSmart N2400 appliance server. Snapshots do not create any cluster resources.

Moving Pools to a Group

In this section, it is assumed that administrators have decided to create one resource group and one virtual server for each node. In this scenario, administrators will have to move their SCE Pool resources to the resource group. To do this, administrators will have to move the SCE Pool resource from the default cluster group to the resource group. When the pool resource has been moved to the desired group, the empty default cluster group of the pool must be deleted.


Move Resources

The group will be changed for the following resources:

Name	Resource Type
new_pool	SCE Pool

Are you sure the group should be changed for these resources?

Figure 6: Moving SCE Pool resource

Creating a Virtual Disk File Share Resource

When the virtual disk is created and mapped to a drive letter, file share resources are created through the Cluster Administrator interface. The first thing administrators must do is determine how they are going to share a virtual disk. Administrators have three options: basic file shares, multiple file shares, or share subdirectories.

Basic File Shares

Administrators can choose to create a file share resource type to provide a single folder that is published to the network under a single name. Keep in mind that the file share resource must be created as a resource under the resource group, where the virtual server resources and SCE Pool resources are located. The root of the virtual disk may be shared out, but it is recommended that a subdirectory of the root of the virtual disk be shared out instead.

Multiple File Shares

Administrators may choose to create multiple folders on a virtual disk and share each folder independently, granting more granular control over how storage is shared and accessed.

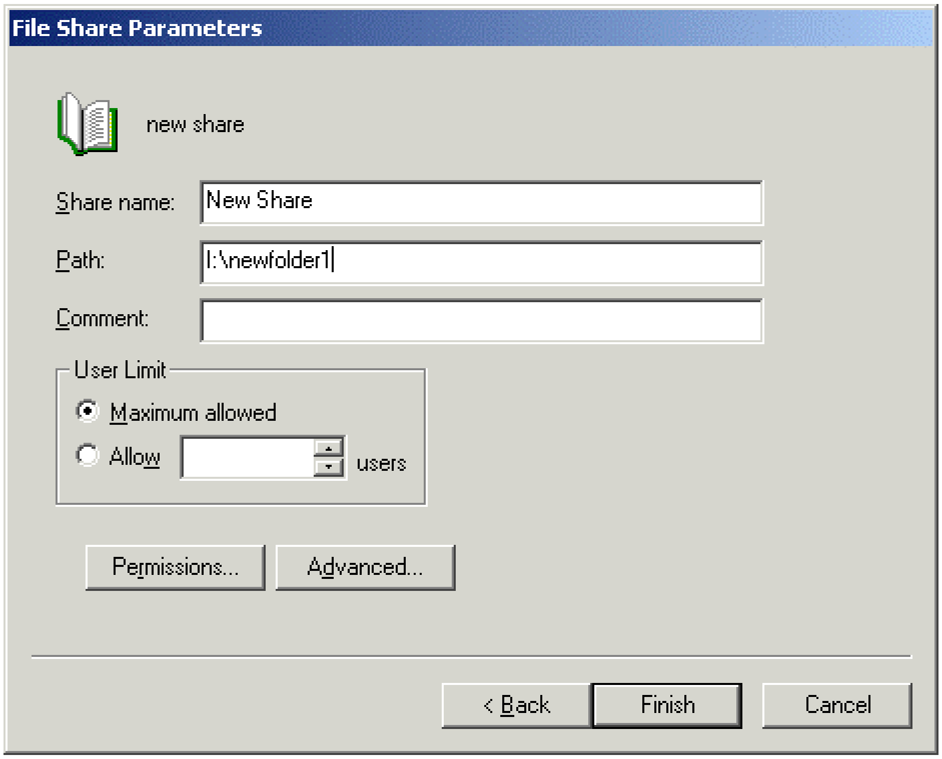
Share Subdirectories

Administrators can also choose to create a single file-share resource under which all subdirectories are shared. This is done by establishing a network file-share name for a file folder and all of its immediate subfolders. This procedure allows administrators to create large numbers of file shares without incurring additional clustering performance penalties.

After an administrator determines how he is going to create the file share, the administrator must create the necessary folder or folders prior to creating the file-share resource.

Procedure for Creating a File-Share Resource

1. From Cluster Administrator, select **New Resource**.
2. Enter a name for the resource.
3. Select **File Share** from the **Resource type** drop-down menu.
4. Select the resource group from the **Group** drop-down menu.
5. Click **Next** in the **Possible Owners** window.
6. Add the appropriate SCE Pool to the **Resource Dependencies** field in the **Dependencies** window.
7. In the **File Share Parameters** window, define the share name, path, and user limit.



Share name: New Share

Path: I:\newfolder1

Comment:

User Limit

● Maximum allowed

○ Allow users

Figure 7: Creating a file share resource

Bringing a Resource Group Online

When all file share resources have been created, administrators can bring the resource group online. To do this, simply right click the resource group, and then select **Bring Online**. This procedure will bring the resource group and its resources online.

Manipulating Nodes, Groups, and Resources

Some of the Cluster Administrator management options include:

- Pausing and resuming a node
- Evicting a node
- Stopping and starting a node

- Moving groups
- Changing the online status, including manually initiating a failover
- Renaming a resource or group
- Deleting a resource or group

These tasks are discussed in the following sections.

Pausing and Resuming a Node

In certain situations administrators may need to perform maintenance or testing on one of the nodes in the cluster. When this is necessary, administrators must pause one of the nodes. Pausing a node maintains the current resources online, but any additional groups or resources cannot be brought online until the node is resumed.

CAUTION: While a node is paused, the cluster is operating in a non-fault-tolerant state. A paused node cannot take ownership of the cluster group.

To pause a node, perform the following steps:

1. Open Cluster Administrator.
2. On the console tree, right click the node, and then select **Pause Node**.

When the maintenance or testing is completed, the node can be resumed.

To resume a node, perform the following steps:

1. Open Cluster Administrator.
2. On the console tree, right click the node, and then select **Resume Node**.

Evicting a Node

In certain situations, administrators may have to evict a node to perform maintenance or to use the QuickRestore program.

Note: If a cluster node has failed, it is not necessary to stop the cluster service to evict the node.

To evict a node, perform the following steps:

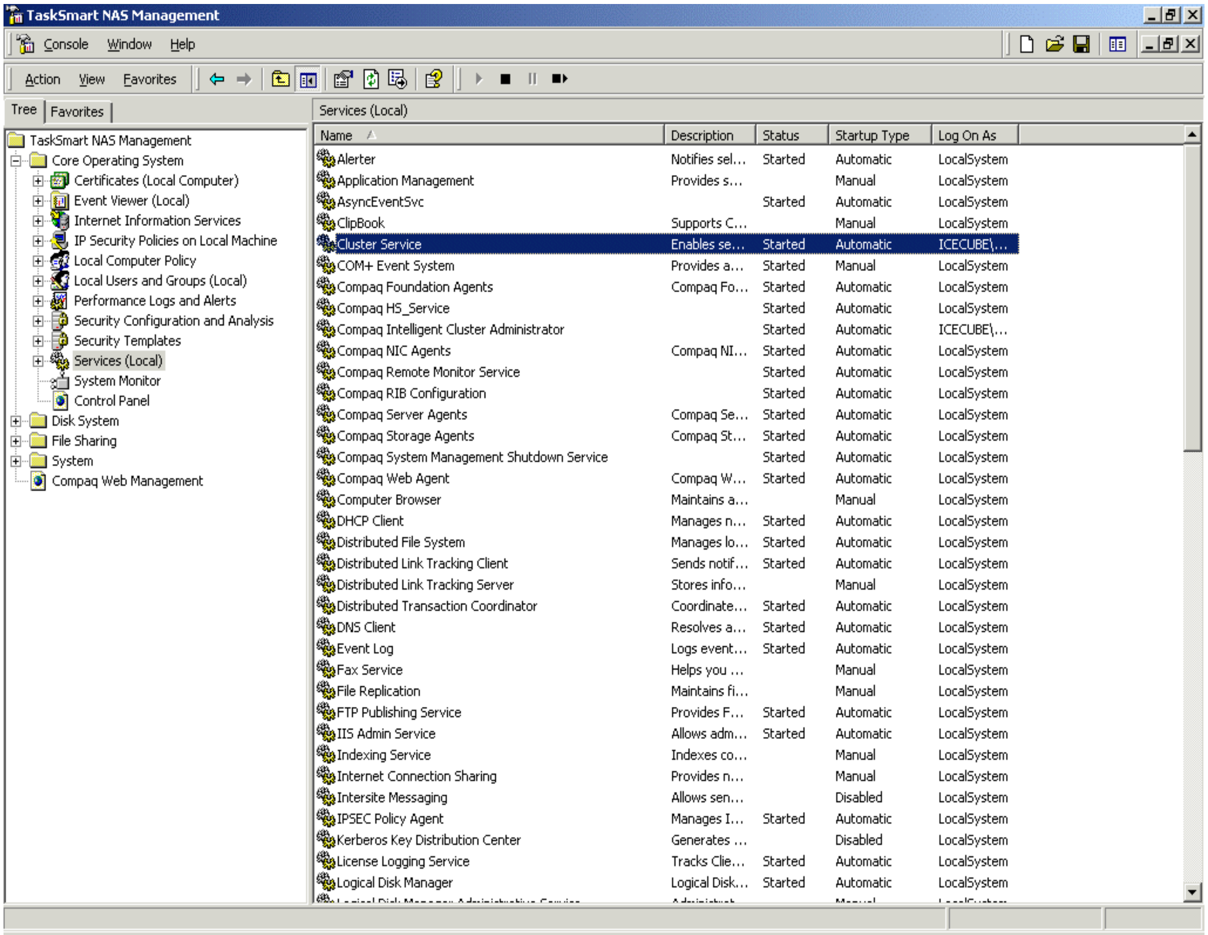
1. Open Cluster Administrator.
2. Stop the cluster service.
3. On the console tree, right click the node, and then select **Evict Node**.

Stopping and Starting the Cluster Service

To stop and start Cluster Service from the TaskSmart NAS Management Console, perform the following steps:

1. Open the TaskSmart N-Series Console.
2. Click **Microsoft Management Console**.
3. Expand the Core Operating System tree.
4. Click **Services**.
5. Right click **Cluster Service** in the right panel.

6. Then, select either **Stop**, **Restart**, or **Start**, accordingly.



Name	Description	Status	Startup Type	Log On As
Alerter	Notifies sel…	Started	Automatic	LocalSystem
Application Management	Provides s…		Manual	LocalSystem
AsyncEventSvc		Started	Automatic	LocalSystem
ClipBook	Supports C…		Manual	LocalSystem
Cluster Service	Enables se…	Started	Automatic	ICECUBE\…
COM+ Event System	Provides a…	Started	Manual	LocalSystem
Compaq Foundation Agents	Compaq Fo…	Started	Automatic	LocalSystem
Compaq HS_Service		Started	Automatic	LocalSystem
Compaq Intelligent Cluster Administrator		Started	Automatic	ICECUBE\…
Compaq NIC Agents	Compaq NI…	Started	Automatic	LocalSystem
Compaq Remote Monitor Service		Started	Automatic	LocalSystem
Compaq RIB Configuration		Started	Automatic	LocalSystem
Compaq Server Agents	Compaq Se…	Started	Automatic	LocalSystem
Compaq Storage Agents	Compaq St…	Started	Automatic	LocalSystem
Compaq System Management Shutdown Service		Started	Automatic	LocalSystem
Compaq Web Agent	Compaq W…	Started	Automatic	LocalSystem
Computer Browser	Maintains a…		Manual	LocalSystem
DHCP Client	Manages n…	Started	Automatic	LocalSystem
Distributed File System	Manages lo…	Started	Automatic	LocalSystem
Distributed Link Tracking Client	Sends notif…	Started	Automatic	LocalSystem
Distributed Link Tracking Server	Stores info…		Manual	LocalSystem
Distributed Transaction Coordinator	Coordinate…	Started	Automatic	LocalSystem
DNS Client	Resolves a…	Started	Automatic	LocalSystem
Event Log	Logs event…	Started	Automatic	LocalSystem
Fax Service	Helps you …		Manual	LocalSystem
File Replication	Maintains fi…		Manual	LocalSystem
FTP Publishing Service	Provides F…	Started	Automatic	LocalSystem
IIS Admin Service	Allows adm…	Started	Automatic	LocalSystem
Indexing Service	Indexes co…		Manual	LocalSystem
Internet Connection Sharing	Provides n…		Manual	LocalSystem
Intersite Messaging	Allows sen…		Disabled	LocalSystem
IPSEC Policy Agent	Manages I…	Started	Automatic	LocalSystem
Kerberos Key Distribution Center	Generates …		Disabled	LocalSystem
License Logging Service	Tracks Clie…	Started	Automatic	LocalSystem
Logical Disk Manager	Logical Disk…	Started	Automatic	LocalSystem

Figure 8: Cluster service

Manually Moving Groups

When a failure occurs, Cluster Administrator automatically moves groups to the other node. In certain situations, groups may have to be moved manually. Move the group manually by using the following steps:

1. Open Cluster Administrator.
2. On the console tree, right click the group to move, and then select **Move Group.**

Changing Online Status, Including Manually Initiating a Failover for Groups and Resources

Cluster Administrator provides administrators with the ability to change group and resource status. This procedure is normally done for testing and maintenance purposes. Administrators can take a group or resource offline. They can also initiate a failure on a resource. Initiating a failure on a resource is helpful for testing the restart and failover policies of resources and groups.

To manually take a group offline, do the following:

1. Open Cluster Administrator.
2. On the console tree, right click the group to take offline, and then select **Take Offline**.

To manually take a resource offline, do the following:

1. Open Cluster Administrator.
2. On the console tree, double click the resource group.
3. The resources for the resource group are displayed on the right panel.
4. Right click the resource to take offline, and then select **Take Offline**.

To manually initiate a failover for a resource, do the following:

1. Open Cluster Administrator.
2. On the console tree, double click the resource group.
3. The resources for the resource group are displayed on the right panel.
4. Right click the appropriate resource on which to initiate a failure, and then select **Initiate Failure**.

Renaming a Resource or Group

To rename a resource, do the following:

1. Open Cluster Administrator.
2. On the console tree, double click the resource group.
3. The resources for the resource group are displayed on the right panel.
4. Right click the resource to rename, and then select **Rename**.

Note: Changing the name of a resource does not affect the failover policy for that resource or the group to which it belongs.

Note: The resource name must be unique.

To rename a group, do the following:

1. Open Cluster Administrator.
2. On the console tree, right click the group to rename, and then select **Rename**.

Note: The group name must be unique.

Deleting a Resource or Group

Before deleting a resource or group, the group or resource must be taken offline.

IMPORTANT: Do not delete or rename the cluster group, because doing so will result in losing the cluster and will require reinstallation of the cluster.

To delete resources, do the following:

1. Open Cluster Administrator.
2. On the console tree, double click the resource group.
3. The resources for the resource group are displayed on the right panel.

4. Right click the resource to delete, and then select **Delete**.

Note: When a resource is deleted, Cluster Administrator also deletes all the resources that have a dependency on the deleted resource.

To delete a group, do the following:

1. Open Cluster Administrator.
2. On the console tree, right click the group to delete, and then select **Delete**.

IMPORTANT: Do not delete SCE Pool resources using Cluster Administrator. SWVR Snapshot Manager must be used to delete pools, after deleting any underlying snapshots and virtual disks. Deleting the pool through Snapshot Manager automatically deletes the cluster SCE Pool resource in Cluster Administrator.

Monitoring Cluster Activity

Event Log

MSCS outputs cluster errors to the Event Log. Any errors that are placed in the event log on one node by MSCS are replicated to the other node.

Quorum Log

On the Quorum Disk, MSCS writes a recovery log so that both nodes in the cluster can access the cluster system state information. When the node that owns the Quorum Disk fails or is brought offline, MSCS gets the current state of the cluster from the failed node and places this information in the recovery log. The remaining node can now access the Quorum Disk to update its cluster database and gain control of the cluster. This information in the quorum log can provide information used to diagnose a failure that occurred.

Basic Administration

Basic administration of the TaskSmart N-Series Cluster will be familiar to administrators of the TaskSmart N2400 appliance, but there are important differences. The *Compaq TaskSmart N2400 Administration Guide* provides important information regarding storage management, file permissions, CIFS file sharing, user and group management, and remote administration techniques. These functions have not changed. Administrators should continue to rely on the *Compaq TaskSmart N2400 Administration Guide* for these topics. This abbreviated guide addresses the features that have either changed or are unique to the TaskSmart N-Series Cluster.

Changes to basic administrative tasks include the following:

- The **TaskSmart N-Series Console**, **Microsoft Management Console (MMC)** interface, and the **Start** menu have been changed to include new items for clustering.
- Power up and power down procedures are more complex.
- The tool used to create physical disk arrays is the StorageWorks Command Console (SWCC) rather than the Compaq Array Configuration Utility (ACU) that was used with the TaskSmart N2400 appliance.

- The TaskSmart N-Series Cluster can now be managed and monitored as a cluster rather than as standalone servers through *Compaq Insight Manage*r^{TM}.
- SecurePath software is loaded to ensure a reliable data path, even if a component fails.

This section details these differences, identifies their improvements, and gives an overview of the essentials for administering the cluster as it differs from the basic TaskSmart N2400 appliance.

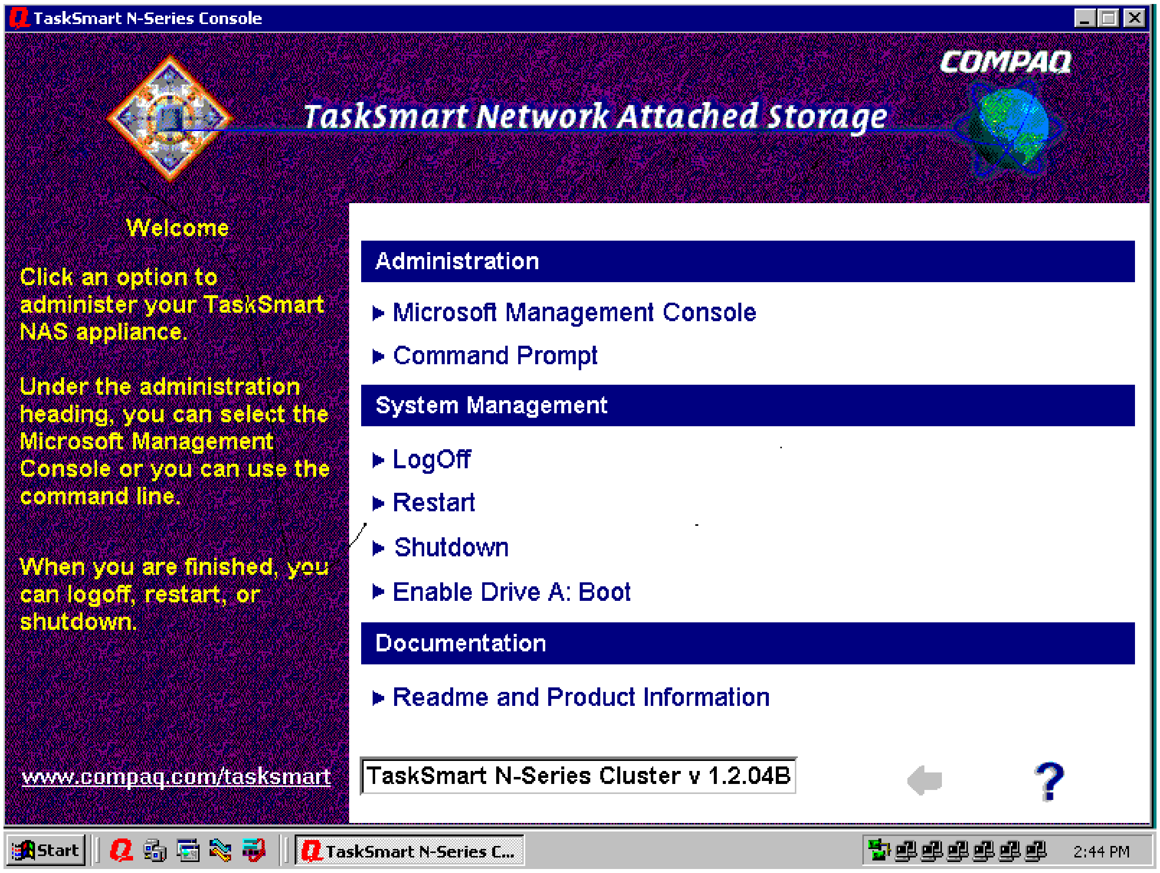
TaskSmart N-Series Console

A special interface called the TaskSmart N-Series Console provides links to many major administrative interfaces. This interface makes administration easier by showing only the items required for fileserver administration, bypassing the normal Windows desktop. The desktop interface is still available for those who are more comfortable with the traditional Microsoft administrative interface.

After logging on, the **TaskSmart N-Series Console** dialog box is displayed, containing options for the following tasks:

- Microsoft Management Console
- Command Prompt
- LogOff
- Restart
- Shutdown
- Enable Drive A: Boot
- ReadMe and Product Information

The TaskSmart N-Series console is shown in Figure 9.



Welcome

Click an option to administer your TaskSmart NAS appliance.

Under the administration heading, you can select the Microsoft Management Console or you can use the command line.

When you are finished, you can logoff, restart, or shutdown.

Figure 9: The TaskSmart N-Series Console

The TaskSmart N-Series Console functions are detailed below:

Microsoft Management Console

The **Microsoft Management Console (MMC)** option launches an MMC control tailored specifically for the TaskSmart N-Series Cluster.

Command Prompt

The **Command Prompt** option launches a command interface for running command-line utilities and programs such as net view, net use, ipconfig, ping, and cd.

Logoff

The **Logoff** option operates as on the TaskSmart N2400 appliance, requiring confirmation before logging off of this node of the system.

Restart

The **Restart** option launches a dialog box that restarts this node of the system. The option is disabled if the current user is not authorized to restart the system. A dialog box is displayed that requires positive confirmation before restarting the system.

IMPORTANT: Restarting a cluster node should be done only after confirming that the other node in the cluster is functioning normally. Adequate warning should be given to users connected to resources of the node being restarted.

Restarting a cluster node causes all file shares served by that node to fail over to the other node in the cluster. Until the failover process completes, any currently executing read and write operations will fail. The other node will be placed under a heavier load by the extra work until the restarted node restarts.

Shutdown

The **Shutdown** option launches a dialog to confirm the need to shut down the cluster node. If the user is not authorized to shut down the system, the **Shutdown** option will be disabled.

IMPORTANT: Shutting down a cluster node must be done only after confirming that the other node in the cluster is functioning normally. Adequate warning should be given to users connected to resources of the node being restarted.

Shutting down a cluster node will cause file shares served by that node to fail over to the other node. This will cause any currently executing client read and write operations to fail until the cluster failover process completes. The other node will be placed under a heavier load by the extra work until the second node is powered up and rejoins the cluster.

If it is your intent to shut down both nodes for maintenance, be aware that a complete shutdown process involves not only shutting down the servers themselves, but also the storage subsystem. Use the procedures outlined in the “Powering Down” section of this document. Failure to properly shut down the storage subsystem could lead to loss of data.

Enable Drive A: Boot

This option sets a configuration parameter so that the server will boot from a diskette on the next boot cycle only. Normally, the system will not boot from a diskette.

ReadMe and Product Information

The **ReadMe and Product Information** option launches a page with links to help files for various system components. Cluster-specific items have been added to the help document list, including help on HSG80 Array Controllers and SecurePath software. The link to the online *Compaq TaskSmart N2400 Administration Guide* has been removed, because this document replaces that guide.

Powering Up the TaskSmart N-Series Cluster

The power up process for the TaskSmart N-Series Cluster is more complex than it was for the TaskSmart N2400 appliance, because extra care must be taken with the storage subsystem.

The power up process can be divided into three segments: supplying power, verifying the storage subsystem, and powering up the cluster nodes. These procedures are detailed as follows:

Supplying Power

1. Plug in and power up the Ethernet switch, if it has not already been done.

2. Plug in and power up the SAN switch. Note that it may take up to 10 minutes before the switch is ready. See the switch documentation for details. See the following note.
3. Plug in the storage enclosures containing the disks. See the following note.
4. Plug in the StorageWorks enclosure model 2200. See the following note.
5. Plug in, but **do not turn on**, the TaskSmart N2400 server heads.
6. Plug in the power supply for the Compaq Remote Insight Lights-Out Edition board located in slot 1 of each cluster node.

Note: All major components of the TaskSmart N-Series Cluster have redundant power supplies. To achieve maximum protection and fault tolerance, the two power supplies for each component (TaskSmart N2400 server heads, SAN switches, StorageWorks enclosure model 2200, and storage enclosures) should be plugged into separate power circuits, preferably separate uninterruptible power supply (UPS) sources on separate power circuits. Should one UPS or power circuit fail, each component will continue to function by relying on its redundant power supply connected to the second UPS and circuit.

Verifying the Storage Subsystem

1. Connect the RJ-11 connector on the serial cable to the port on the top HSG80 Array Controller in the Model 2200 enclosure. Connect the other end of the serial cable to the nine-pin serial connector on a Microsoft Windows-based computer running Microsoft Windows NT or Windows 2000.
2. Log on to the Windows-based computer that is connected to the serial cable from the HSG80 Array Controller.
3. Verify that power is supplied to the Model 2200 enclosure before attempting to connect to the HSG80 Array Controller through the serial connection.

Run **HyperTerminal** and create a new connection with the following input for each option:

Baud rate	9600
Data bits	8
Parity	none
Stop bits	1
Flow control	hardware

4. Connect using the new connection.
5. After pressing the **Enter** key several times, the controller prompt is displayed, indicating that the controller is online.
6. Review the configurations for dual-redundant configuration. Enter the following in HyperTerminal to show the status of the bottom controller:

```
prompt > show other_controller
```

7. Scan the output for the phrases **Configured for multi-bus failover** and **In dual-redundant configuration**.

8. After you see these phrases, enter the following in HyperTerminal, to display the status of the controller to which the serial cable is connected:

   ```
   prompt > show this_controller
   ```

9. Scan the output for the phrases **Configured for multi-bus failover** and **In dual-redundant configuration**.

   If the expected output is not displayed, there is either a problem with the power or with the controller configuration. Refer to the *Compaq TaskSmart N-Series Cluster Installation Guide* for details on proper controller configuration.

Powering Up the Cluster Nodes

IMPORTANT: Do not power up the TaskSmart N-Series Cluster nodes without first powering up the storage subsystem, and verifying it according to the previous procedure.

1. The TaskSmart N-Series Cluster nodes may be powered up after power has been supplied, and the storage subsystem is confirmed to be operating normally.
2. Power up a single node by pressing the power button on the front of the server. If both nodes are powered up at the same time, the first node that completes the sequence will be the cluster quorum owner–the node that controls the cluster database. Some customers may wish to designate a particular server node as the usual cluster quorum owner. If so, care should be taken to power up that node first, and let it boot completely before powering up the second cluster node.
3. Power up the second cluster node by pressing the power button on the front of the server.
4. Both cluster nodes will boot up and display the logon dialog. Background processes will start the cluster service and form the cluster.

Powering Down the TaskSmart N-Series Cluster

The power down process for the TaskSmart N-Series Cluster is similar to the process for the TaskSmart N2400 appliance, but with the cluster, extra care must be taken with the storage subsystem.

The power down process is divided into three main steps: powering down the cluster nodes, shutting down the storage subsystem, and removing power. These steps are detailed as follows.

Powering Down the Cluster Nodes

IMPORTANT: Before powering down the cluster nodes, follow the proper shutdown procedure by using the **Shutdown** button on the **TaskSmart N-Series Console** or from the **Start** menu. Note the cautions described in the previous subsection of the “Basic Administration” section titled “Shutdown.”

A TaskSmart N-Series Cluster node automatically powers down when the shutdown procedure has been performed as described in the following section, “Shutting Down the Storage Subsystem.” The shutdown procedure automatically turns off power to the server, putting it in standby mode, denoted by a yellow LED. After both nodes have reached this point, it is safe to continue the process of powering down the cluster.

Shutting Down the Storage Subsystem

1. Connect the RJ-11 connector on the serial cable to the port on the top HSG80 Array Controller in the StorageWorks enclosure model 2200. Connect the other end of the serial cable to the nine-pin serial connector on a Microsoft Windows-based computer running Microsoft Windows NT or Windows 2000.
2. Log on to the Windows-based computer that is connected to the serial cable from the HSG80 Array Controller.
3. Run HyperTerminal and if necessary, create a new connection with the following input for each option:

Baud rate	9600
Data bits	8
Parity	None
Stop bits	1
Flow control	Hardware

4. Connect using the new or previously defined connection.
5. After pressing the **Enter** key several times, the controller prompt is displayed, indicating that the controller is online.
6. Shut down the bottom controller first by entering the following in HyperTerminal:

```
prompt > shutdown other_controller immediate
```

7. Now shut down the top controller by entering the following in HyperTerminal:

```
prompt > shutdown this_controller immediate
```

8. Observe the back of the Model 2200 enclosure. Locate the reset button to the left of the fiber ports on each of the two HSG80 Array Controllers. Wait until the reset button and the leftmost three of the six small LEDs next to the reset button are all green on both HSG80 Array Controllers.

Removing Power

IMPORTANT: Do not power down the storage subsystem components (Model 2200 enclosure, SAN switches, or storage enclosures) of the TaskSmart N-Series Cluster without following the previous procedures. Improperly powering down can cause corruption and loss of data, and could require reinstallation of the cluster.

1. Unplug the TaskSmart N-Series cluster nodes and the Remote Insight Lights-Out Edition power supplies.
2. Power down and unplug the Ethernet switch, unless other computer systems are connected to it.
3. Power down and unplug the StorageWorks SAN switch.
4. Unplug the storage enclosures containing the disks.
5. Unplug the Model 2200 enclosure.

IMPORTANT: The StorageWorks enclosure model 2200 contains cache batteries to ensure that contents are preserved during power failures. If the system will be powered down for more than a few minutes, use the procedures outlined in the Model 2200 enclosure reference guide to ensure that the cache batteries do not fully discharge.

StorageWorks Command Console

StorageWorks Command Console (SWCC) replaces the Compaq Array Configuration Utility (ACU) on the cluster as the tool for creating external disk arrays and logical disks. Even though these activities may be done through a command-line interface via a serial cable to the storage subsystem, SWCC is a more intuitive, graphical interface for managing the storage, receiving event notifications, and monitoring the storage subsystem of the TaskSmart N-Series Cluster.

The information provided here is not intended to replace the *Compaq StorageWorks Command Console User's Guide*, but to introduce SWCC features and give an overview of how this software is used on the TaskSmart N-Series Cluster. The "Advanced Administration" section of this guide provides information on how to use SWCC for creating storage arrays and disk units.

Starting SWCC

To start using StorageWorks Command Console:

1. Click **Start** and select **Programs**.
2. Select **Compaq Cluster Tools** from the **Programs** menu.
3. Select **StorageWorks Command Console** from the next menu.
4. SWCC will start up, and display the main window, as shown in Figure 10.

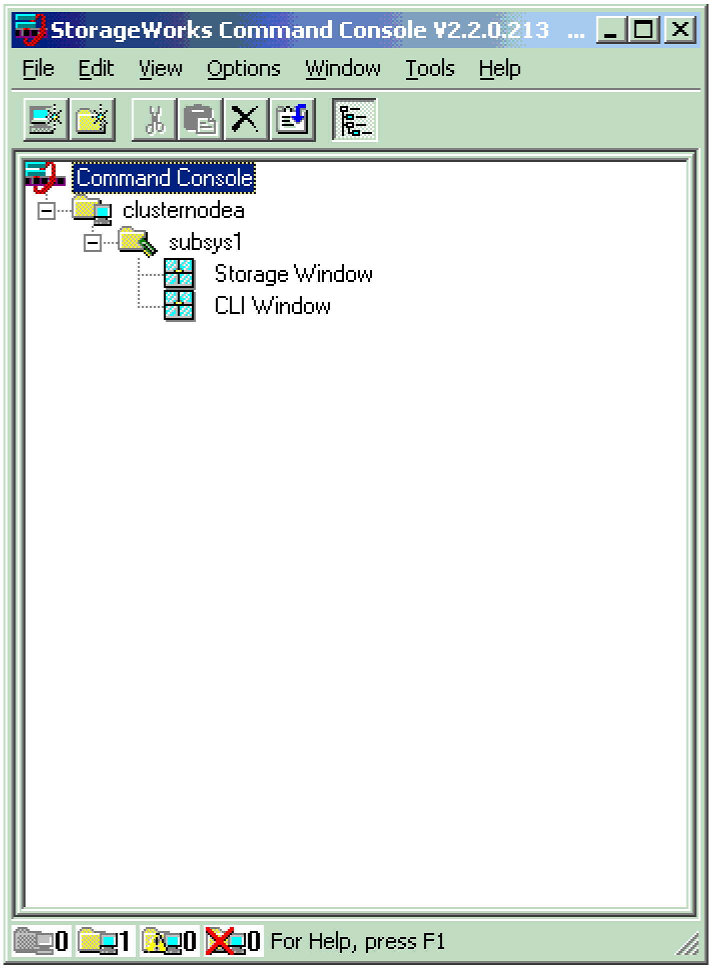

StorageWorks Command Console V2.2.0.213 ...
File Edit View Options Window Tools Help
Command Console
clusternodea
subsys1
Storage Window
CLI Window
0 1 0 0 For Help, press F1

Figure 10: StorageWorks Command Console window

SWCC Components

The StorageWorks Command Console has several main parts, the most important of which are:

- **Navigation window**: This is the main window for SWCC. A tree-structured interface expands to show the cluster, the subsystem, and the **Storage** and **CLI** windows.
- **Storage window**: This is the interface that allows the administrator to monitor and control the storage subsystem graphically, through the connection to the subsystem.
- **CLI window**: This window allows the administrator to issue monitoring and configuration commands in a command-line interface to the storage subsystem.

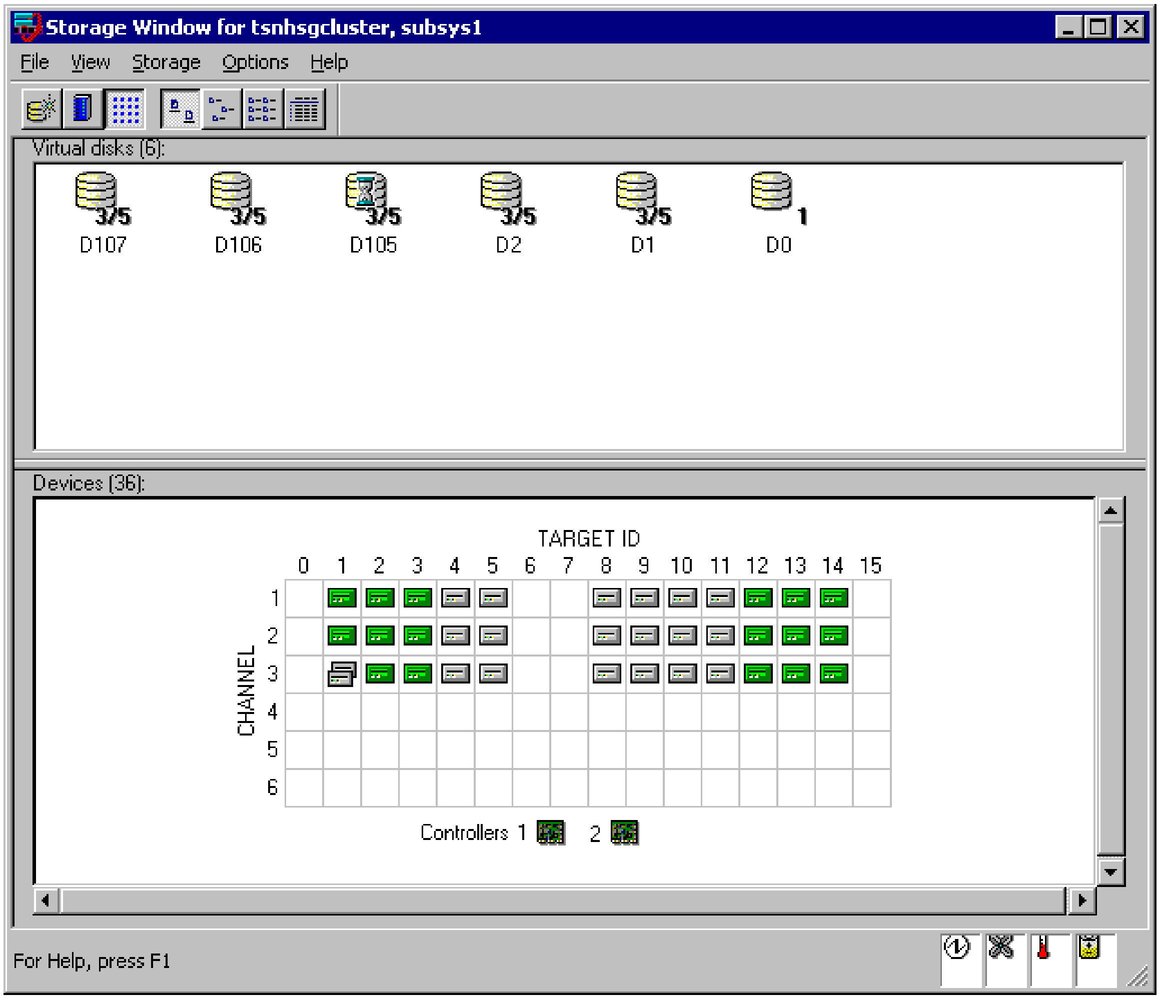


Figure 11: StorageWorks Command Console Storage window

SWCC Storage Window

The **Storage** window is split into two main panes. The top pane displays the set of disk units and their current state. The bottom pane is a representation of the physical storage subsystem in either a simple grid view or an actual physical view of the storage enclosures and the Model 2200 enclosure.

In Figure 11, note that the top pane depicts several important facts about the disks on the storage subsystem.

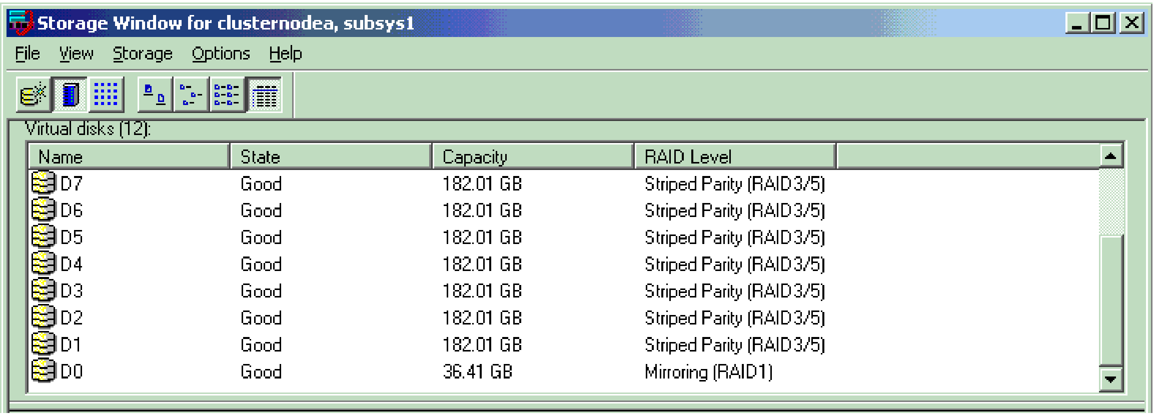
- There are five RAIDsets and one mirrorset.
- RAIDsets are depicted by a disk icon with a “3/5” label, meaning RAID level 3 or 5.
- Mirrorsets are depicted by a disk icon with a “1” label, meaning RAID 1.
- Disk D105 is depicted with an hourglass symbol, meaning that it is in a reconstructing state. Other informative symbols are also used, and are fully defined in the SWCC help menu.

Figure 11 also shows a grid view of the storage subsystem in the bottom pane. From this view, the following facts are presented about the storage subsystem:

- There are three connected storage enclosures.
- There are a total of 36 disk devices.
- There are 12 disks in each enclosure, with the first and last drive bay of each enclosure remaining empty.

Changing the View from the Storage Window

Alternative views are available for both the top and bottom panes. Using the buttons at the top of the **Storage** window, it is possible to change to a more detailed presentation of the virtual disks, as shown in Figure 12.



Virtual disks (12):

Name	State	Capacity	RAID Level
D7	Good	182.01 GB	Striped Parity (RAID3/5)
D6	Good	182.01 GB	Striped Parity (RAID3/5)
D5	Good	182.01 GB	Striped Parity (RAID3/5)
D4	Good	182.01 GB	Striped Parity (RAID3/5)
D3	Good	182.01 GB	Striped Parity (RAID3/5)
D2	Good	182.01 GB	Striped Parity (RAID3/5)
D1	Good	182.01 GB	Striped Parity (RAID3/5)
D0	Good	36.41 GB	Mirroring (RAID1)

Figure 12: View of the Storage window top pane showing a detailed view of the devices

Note: The term "virtual disk" as used by StorageWorks Command Console GUI is not the same as the virtual disk created by SANworks Virtual Replicator (SWVR).

Similarly, the bottom pane of the **Storage** window can be changed using the **View** menu to present a more realistic view of the storage subsystem and disk enclosures. See Figure 13.

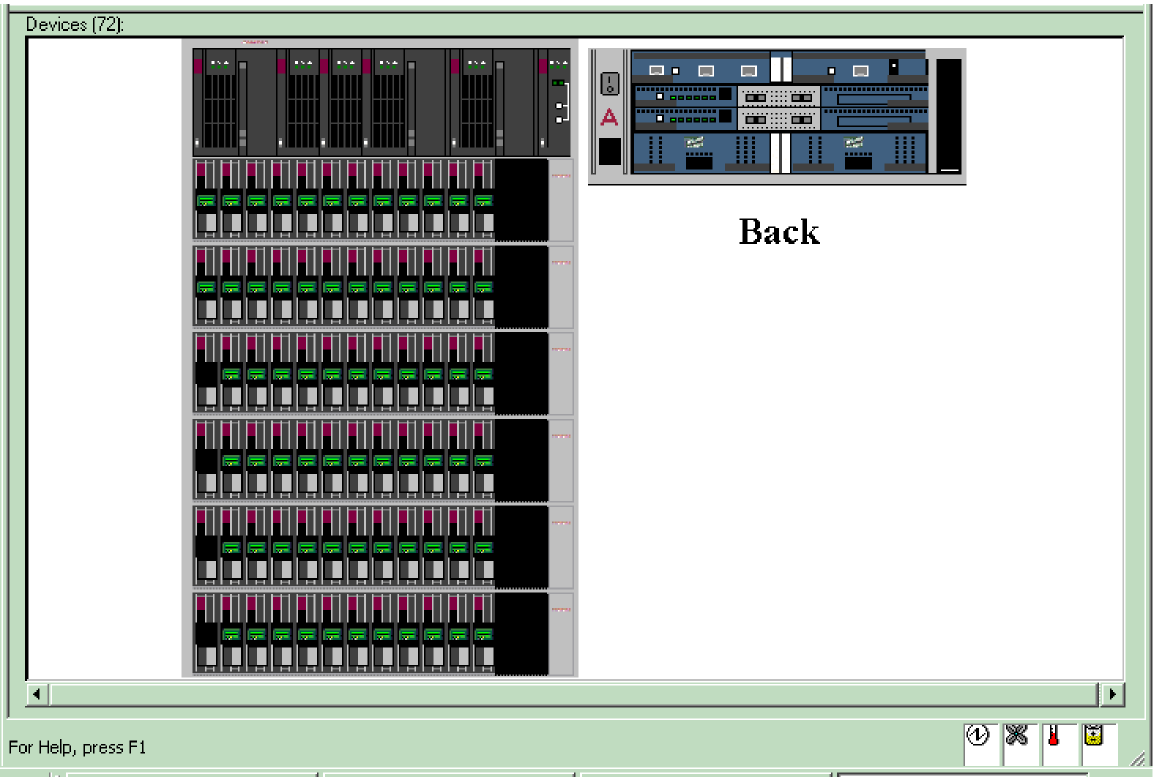


Figure 13: View of the Storage window bottom pane with a physical view of the subsystem

Monitoring Status with the SWCC Storage Window

As described previously, the administrator can monitor the status of SWCC virtual disks, controllers, and individual physical disks through the **Storage** window. Additionally, the bottom of the storage window has indicators for depicting the status of power, fans, temperature, and cache battery. The legend in Figure 14 shows all the possible conditions of each component:

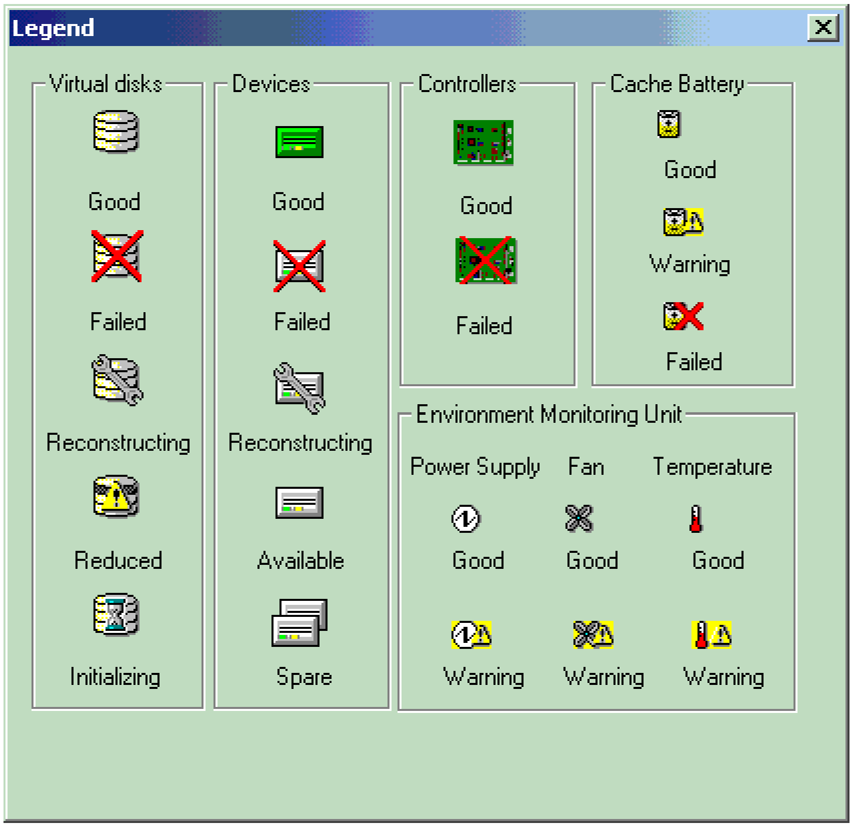


Figure 14: StorageWorks Command Console Legend

CAUTION: Failed components should be replaced as soon as possible using the procedures detailed in the "Hardware Failure" section of this guide and in the *Compaq StorageWorks 2200 Enclosure* reference guide. Failure to properly replace a failed component could lead to cluster failure and data loss or corruption.

Verifying Controller Settings with SWCC

The administrator can check and change the properties of the HSG80 Array Controllers by accessing their property sheets in SWCC. To access a property sheet, click the controller in the interface, and then right click and select **Properties** from the menu. The property sheet contains information such as:

- World Wide identifier
- Controller type and firmware version
- Operational state
- Connections
- Battery status
- Port topology

Adding a New Disk with SWCC

There is more to adding a disk device to the storage subsystem than just sliding a supported disk into an available disk bay. The administrator must also make the system aware of the addition. This must take place before attempting to use the disks to create an array, or RAIDset, as SWCC refers to them. The steps for adding a device are:

1. From the **Navigation** window, open the **Storage** window for the storage subsystem.
2. In the graphical view in the bottom pane, select a slot for the new disk by clicking the slot.
3. From the **Storage** menu, select **Device**, and then select **Add**.
4. Enter the SWCC password in the dialog box that is displayed.
5. Physically insert the device into the chosen slot.
6. A dialog box is displayed, asking you to confirm disk insertion. Click **OK** when this step is completed.
7. The subsystem will now scan for the new disk, after which it can be used in a new RAIDset, mirrorset, or spareset.

Identifying Disks and RAIDsets

It is sometimes necessary to visually identify a particular disk or set of disks in the storage subsystem. The following process facilitates that task:

1. From the **Storage** window, select a virtual disk in the top pane, or a particular physical disk in the bottom pane.
2. If you selected a virtual disk in the top pane, the member disks of the underlying RAIDset or mirrorset are highlighted in the bottom pane of the **Storage** window.
3. Select the **Storage** menu at the top of the **Storage** window, and then select **Device**.
4. From the **Device** menu, select **Locate**.
5. Observe the storage subsystem. Disk members of the object selected in step 1 can be identified by the blinking drive failure indicator lights on each drive.
6. Click **OK** on the dialog box to discontinue the locate task.

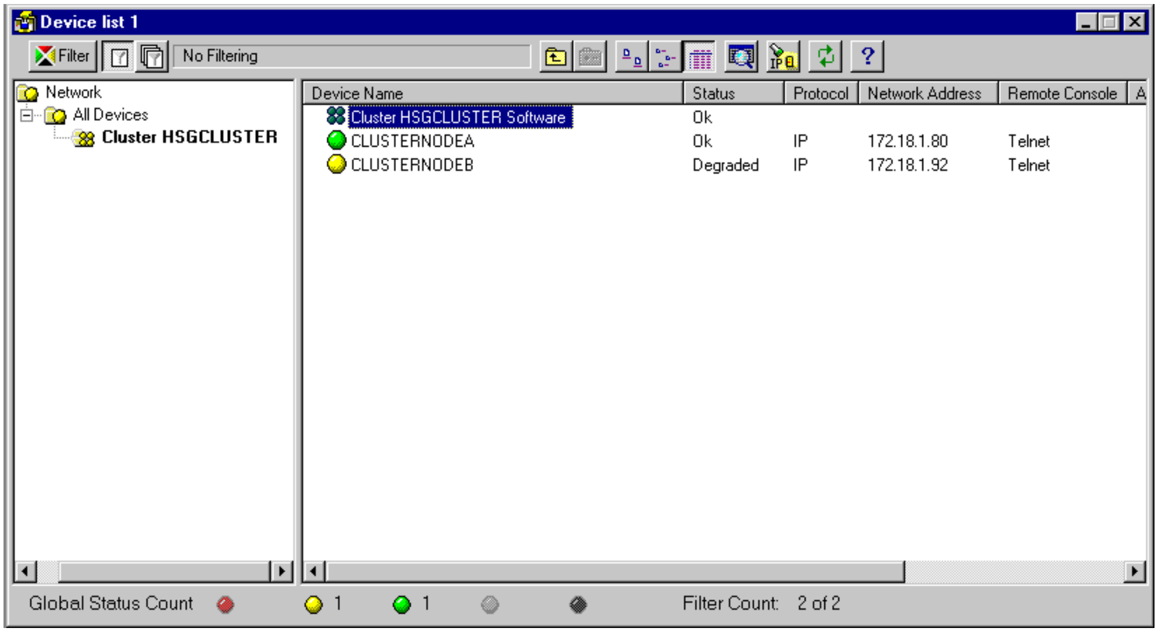
Compaq Insight Manager

Just as in the TaskSmart N2400 appliance, the Compaq TaskSmart N-Series Cluster is equipped with the latest Compaq Insight Management Agents for servers. Compaq Insight Manager allows administrators to manage the TaskSmart N-Series Cluster from its Windows-based console application and from a remote location through the Insight Manager Agent Web Interface.

Compaq Insight Manager Console

Managing the Compaq TaskSmart N-Series Cluster from the Compaq Insight Manager console is the same as in the TaskSmart N2400 appliance. To manage the cluster, identify one of the nodes in the cluster and add it to the list of all devices. When one of the nodes is discovered by Compaq Insight Manager, Compaq Insight Manager will automatically show a cluster icon in the list of devices. To manage the other node in the cluster, that node must also be added to the list of all devices.

IMPORTANT: The administrator must identify the IP address of the node and not the IP address of the cluster or virtual server.



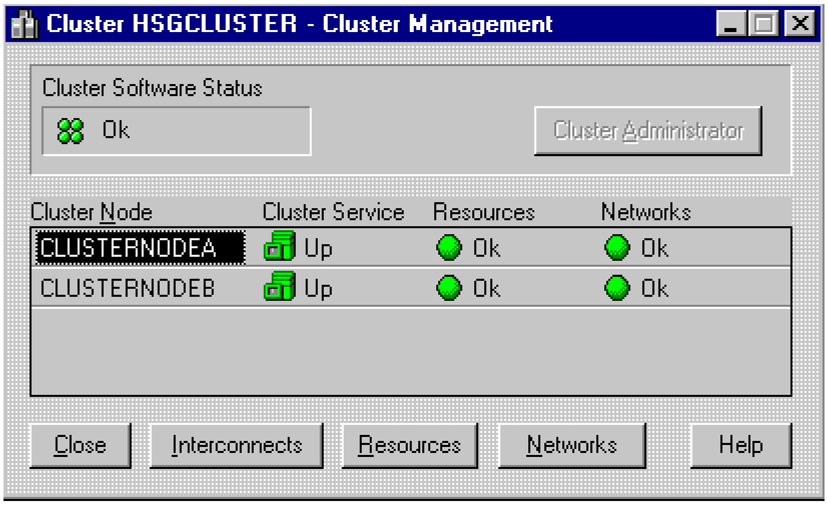
Device Name	Status	Protocol	Network Address	Remote Console	A
Cluster HSGCLUSTER Software	Ok				
CLUSTERNODEA	Ok	IP	172.18.1.80	Telnet	
CLUSTERNODEB	Degraded	IP	172.18.1.92	Telnet	

Global Status Count ● ● 1 ● 1 ● ● Filter Count: 2 of 2

Figure 15: Compaq Insight Manager Console

Cluster Management Window

Compaq Insight Manager provides a **Cluster Management** window that allows users to view the state of the cluster service running on each node



Cluster Software Status: Ok

Cluster Node	Cluster Service	Resources	Networks
CLUSTERNODEA	Up	Ok	Ok
CLUSTERNODEB	Up	Ok	Ok

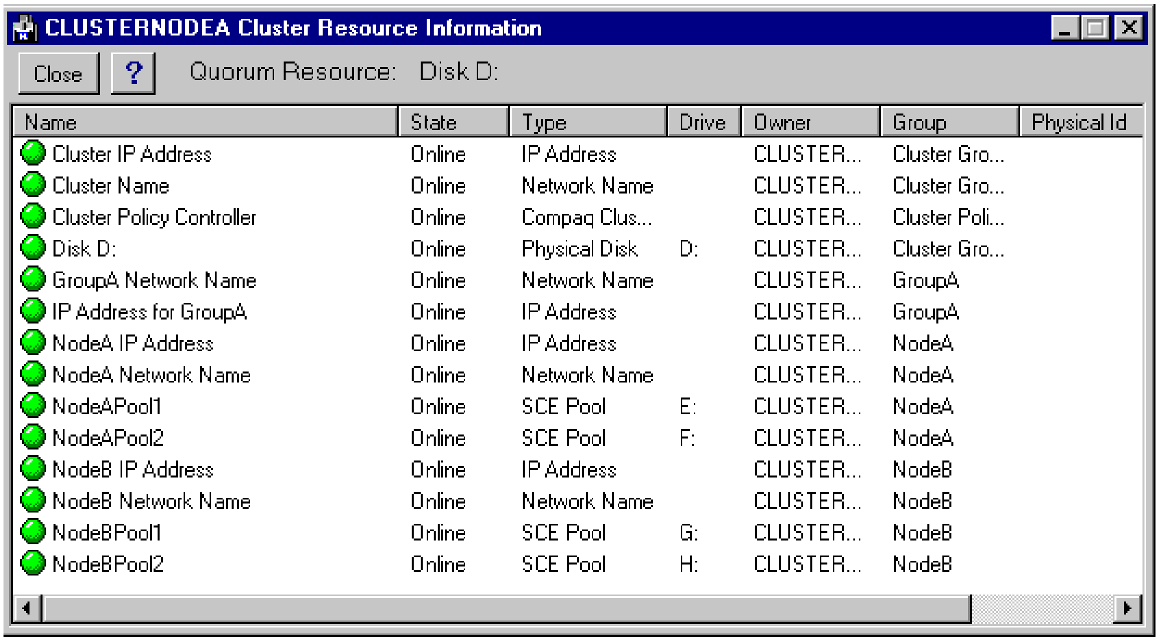
Figure 16: Cluster Management window

The following items are included in the **Cluster Management** window:

Cluster Administrator button-Starts the Microsoft Cluster Administrator application if it is installed. If the application is not installed, the button will be disabled.

Interconnects button-Launches a **Cluster Interconnect Information** window. The button is disabled if no cluster nodes are selected.

Resources button-Displays the status of the shared resources for each node in the cluster. See Figure 17 for an example of the **Cluster Resource Information** window.



CLUSTERNODEA Cluster Resource Information

Quorum Resource: Disk D:

Name	State	Type	Drive	Owner	Group	Physical Id
Cluster IP Address	Online	IP Address		CLUSTER...	Cluster Gro...	
Cluster Name	Online	Network Name		CLUSTER...	Cluster Gro...	
Cluster Policy Controller	Online	Compaq Clus...		CLUSTER...	Cluster Poli...	
Disk D:	Online	Physical Disk	D:	CLUSTER...	Cluster Gro...	
GroupA Network Name	Online	Network Name		CLUSTER...	GroupA	
IP Address for GroupA	Online	IP Address		CLUSTER...	GroupA	
NodeA IP Address	Online	IP Address		CLUSTER...	NodeA	
NodeA Network Name	Online	Network Name		CLUSTER...	NodeA	
NodeAPool1	Online	SCE Pool	E:	CLUSTER...	NodeA	
NodeAPool2	Online	SCE Pool	F:	CLUSTER...	NodeA	
NodeB IP Address	Online	IP Address		CLUSTER...	NodeB	
NodeB Network Name	Online	Network Name		CLUSTER...	NodeB	
NodeBPool1	Online	SCE Pool	G:	CLUSTER...	NodeB	
NodeBPool2	Online	SCE Pool	H:	CLUSTER...	NodeB	

Figure 17: Cluster Resource Information window

Networks button - Launches a **Cluster Network Information** window. The button is disabled if no cluster nodes are selected.

For more information about using Compaq Insight Manager, visit the Compaq System Management website:

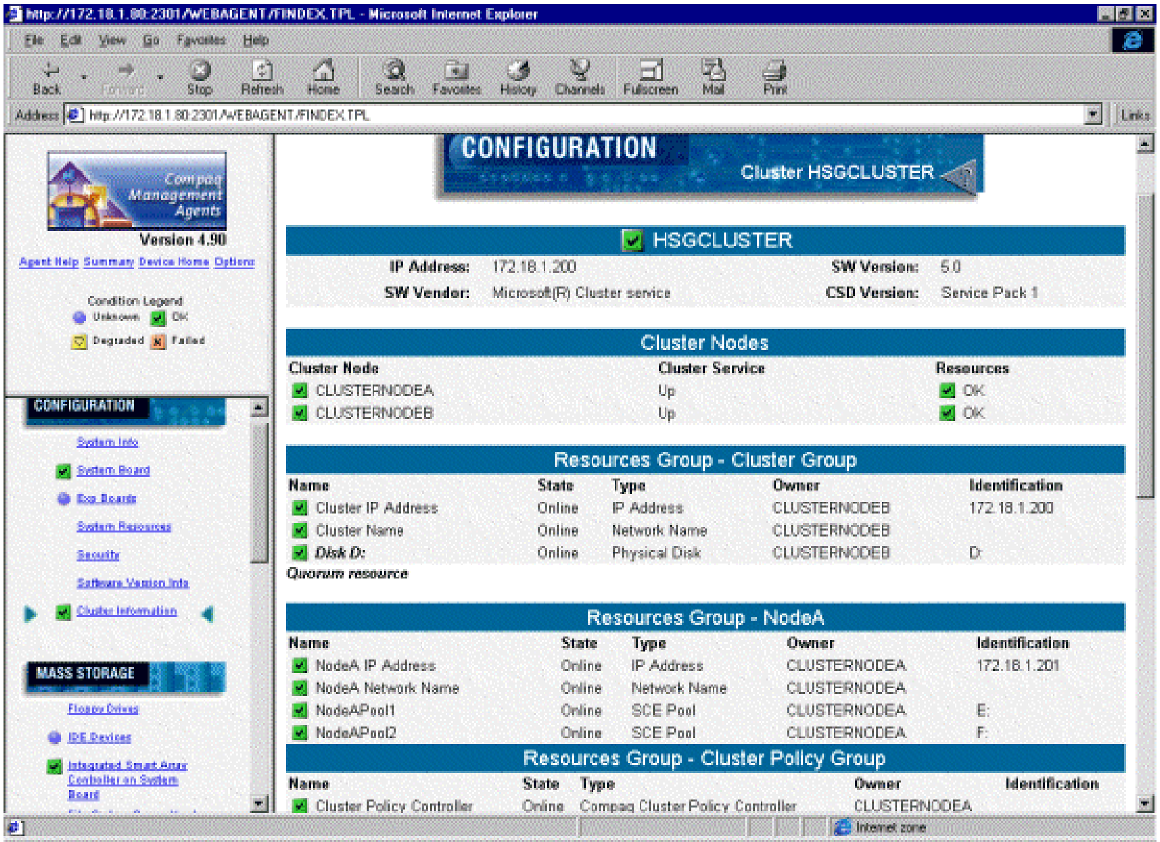
www.compaq.com/products/servers/management/

Web Agent Interface

Accessing the Compaq Insight Manager Agent Web Interface is the same as accessing it from the TaskSmart N2400 appliance.

The two options for accessing the Compaq Insight Manager Agent Web Interface are:

1. From the Compaq Insight Manager console, right click the device name and select **View Web Data**. The Agent Web Interface of the server launches in a new window within Compaq Insight Manager.
2. Open a Web browser and enter the server IP address using port 2301. An example IP address is http://122.18.1.14:2301. The default logon account is “Anonymous.” Click the account name to log on as Administrator. The default username and password are both “administrator,” in lower case type. After the user is logged on as an Administrator, the user can change the password.



CONFIGURATION

Cluster HSGCLUSTER

✓ HSGCLUSTER

IP Address:	172.18.1.200	**SW Version:**	5.0
SW Vendor:	Microsoft(R) Cluster service	**CSD Version:**	Service Pack 1

Cluster Nodes

Cluster Node	Cluster Service	Resources
✓ CLUSTERNODEA	Up	✓ OK
✓ CLUSTERNODEB	Up	✓ OK

Resources Group - Cluster Group

Name	State	Type	Owner	Identification
✓ Cluster IP Address	Online	IP Address	CLUSTERNODEB	172.18.1.200
✓ Cluster Name	Online	Network Name	CLUSTERNODEB	
✓ *Disk D:*	Online	Physical Disk	CLUSTERNODEB	D:

Quorum resource

Resources Group - NodeA

Name	State	Type	Owner	Identification
✓ NodeA IP Address	Online	IP Address	CLUSTERNODEA	172.18.1.201
✓ NodeA Network Name	Online	Network Name	CLUSTERNODEA	
✓ NodeAPool1	Online	SCE Pool	CLUSTERNODEA	E:
✓ NodeAPool2	Online	SCE Pool	CLUSTERNODEA	F:

Resources Group - Cluster Policy Group

Name	State	Type	Owner	Identification
✓ Cluster Policy Controller	Online	Compaq Cluster Policy Controller	CLUSTERNODEA	

Figure 18: Compaq Insight Manager Web Agent Interface

SecurePath and SecurePath Manager

This section provides a basic overview of SecurePath and SecurePath Manager. It discusses the following topics:

- Basic Overview of SecurePath
- Basic Overview of SecurePath Manager
- Functionality within SecurePath Manager

For configuration of the SecurePath agent and configuring a SecurePath Manager profile, refer to the *Compaq TaskSmart N-Series Cluster Installation Guide*.

Basic Overview of SecurePath

SecurePath is a high-availability software product that provides continuous data access to the TaskSmart N-Series Cluster. Redundant hardware, advanced RAID technology, and automated failover capability are also used in the TaskSmart N-Series Cluster configuration to further enhance fault tolerance and availability. SecurePath eliminates the host bus adapter and interconnect hardware as single points of failure in the storage subsystem. SecurePath allows the StorageWorks dual-controller RAID subsystem to be cabled to two or more Fibre Channel Switch paths using two separate host bus adapters in each cluster node. SecurePath monitors each path and automatically reroutes I/O to the functioning alternate path(s) should an adapter, cable, switch, or controller failure occur.

Failure detection is designed to prevent false or unnecessary failovers. The SecurePath utility provides continuous monitoring capability and identifies failed paths and storage units that have been failed over from the other controller. A storage unit refers to a logical container in which RAID arrays reside.

SecurePath consistently monitors the “health” of available storage units and physical paths through its path verification process. A redundant physical connection defines a physical “path” in SecurePath Manager. Each path originates at a unique host bus adapter (HBA) port on a cluster node, and ends at a unique port on the Model 2200 enclosure storage system. SecurePath can also perform static load balancing by distributing devices (LUNs) between the two controllers.

Basic Overview of SecurePath Manager

SecurePath Manager (SPM) is used to monitor and manage the SecurePath environment. SPM displays specific information about the state of RAID storage systems and I/O paths, both of which are configured through the HSG80 Array Controllers. Use SPM to set various properties and modes associated with a managed storage profile, and to set failback policy. SPM automatically detects and indicates path failures, and provides the capability to distribute and move RAIDsets across controller pairs for static load balancing.

Functionality Within SecurePath Manager

Within SecurePath Manager, physical storage objects are displayed in the frame on the left side, and the paths to those components are displayed in a frame on the right side (See Figure 19). The administrator selects the method SPM uses to identify storagesets with the **View** pull-down menu. These options include:

- **Disk LUN UID** – A unique 128-bit value assigned by SecurePath
- **Disk Number** – The logical disk number assigned by the Windows Disk Administrator
- **Drive Letter** – The logical drive letter assigned by the Windows Disk Administrator
- **Bus/Target/LUN** – The physical address representing the connection to the host server
- **Volume Label -** The volume label assigned by the user with Windows Explorer or Disk Administrator.

Note: If the drive letter or disk number options are selected, then the presentation of information is not accurate with regards to SANworks Virtual Replicator. Because advanced disk virtualization is being utilized, the pool and virtual disk information is not represented within SecurePath Manager. If the volume label option is selected, only the labels of disks created within logical volume manager are displayed. In this case, that would be only the Quorum Disk.

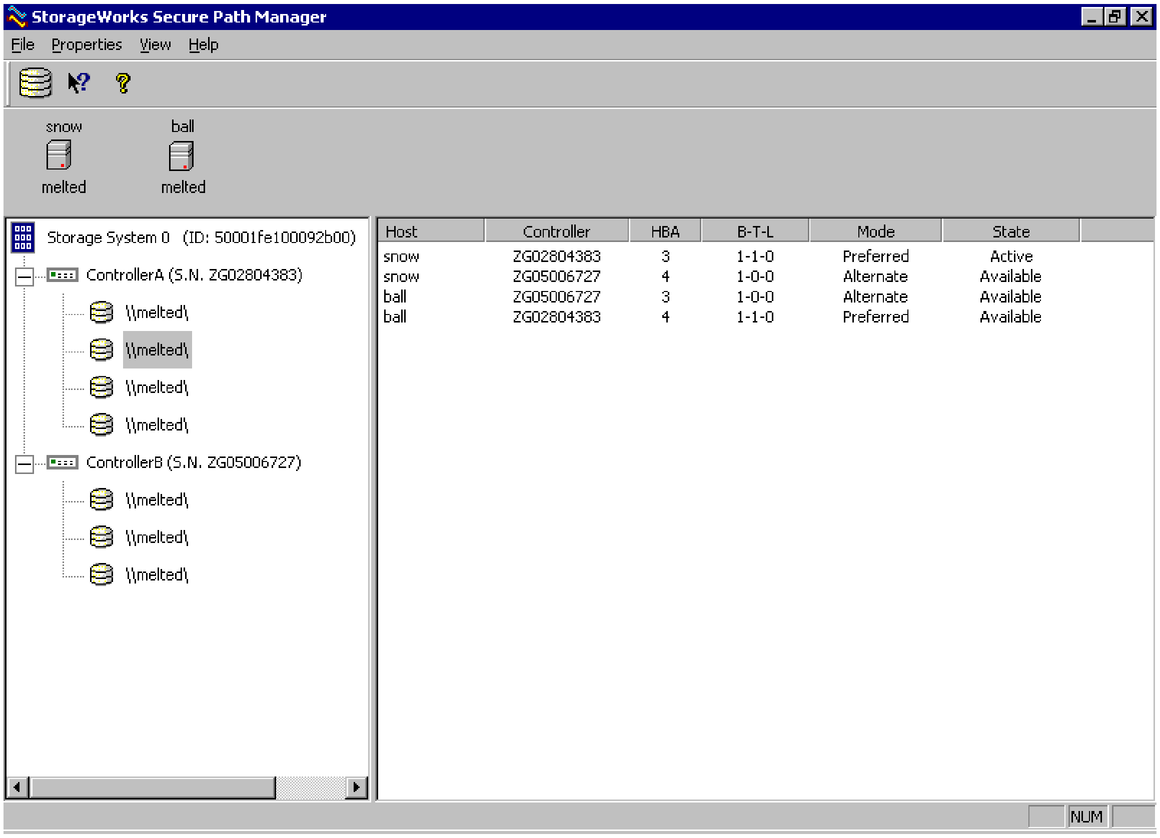
SPM always displays the owning host name, or cluster name (for clustered hosts), along with the chosen storageset identifier. When a storageset from the Storage System view is highlighted, SPM displays information about the physical paths that have been configured for access to that storageset in the frame on the right. The Physical Path view includes the following information for each path:

- **Host** – The SecurePath host system, with an established access path to the storageset
- **Controller** – The RAID storage system controller servicing the path
- **HBA** – The physical port number of the host bus adapter (HBA) servicing the path.

  The HBA is a relative number determined by Windows "order of discovery" for adapters on that host.
- **B-T-L** – The physical bus, target, and logical unit number (LUN) describing the path address for the storageset.
- **Mode** – A user-selectable parameter that specifies path behavior during normal and failure conditions

  Path mode can be set to Preferred, Alternate, Pre-Offline (Preferred and Offline), or Alt-Offline (Alternate and Offline).
- **State** – The current status of the path

Note: The Quorum Disk always displays only one active path, while other storagesets display two active paths. This is directly related to the fact that the Quorum Disk is owned by only one node at a time, therefore limiting the paths to one active, one preferred, and two alternates.



Storage System 0 (ID: 50001fe100092b00)

- ControllerA (S.N. ZG02804383)
  - \\melted\
  - \\melted\
  - \\melted\
  - \\melted\
- ControllerB (S.N. ZG05006727)
  - \\melted\
  - \\melted\
  - \\melted\

Host	Controller	HBA	B-T-L	Mode	State
snow	ZG02804383	3	1-1-0	Preferred	Active
snow	ZG05006727	4	1-0-0	Alternate	Available
ball	ZG05006727	3	1-0-0	Alternate	Available
ball	ZG02804383	4	1-1-0	Preferred	Available

Figure 19: SecurePath Manager

Advanced Administration

Storage Subsystem

The storage subsystem for the TaskSmart N-Series Cluster is managed by dual HSG80 Array Controllers, which are installed in a StorageWorks enclosure model 2200. This duplication provides fault tolerance. The disk storage enclosures are then attached to the Model 2200 enclosure via a SCSI cable. The Model 2200 enclosure supports six storage cabinets and a total of 72 disk devices. This section covers advanced administration procedures for the configuration and maintenance of the storage subsystem.

Units and Disksets

A unit is a container in which a diskset physically resides. The controller maps requests from a cluster node to the logical unit number. The unit number is the designation by which the controller keeps track of the unit. Units can be created from all diskset types. When configuring the disk devices to be presented to the cluster nodes, they are first placed into RAIDsets, mirrorsets, or stripesets. The diskset is then assigned to a unit. Units are presented as physical disk devices in SANworks Virtual Replicator (SWVR). When a new pool is created in SWVR, a list of disk devices composed of those units is displayed that can be used to create the new pool. Those disk devices are the units that were created on the HSG80 Array Controller. For example, to create a RAIDset named r1 with 6 disks, the following command is used:

```
prompt > add raidset r1 disk10000 disk10100 disk10200 disk10300 disk10400
disk10500
```

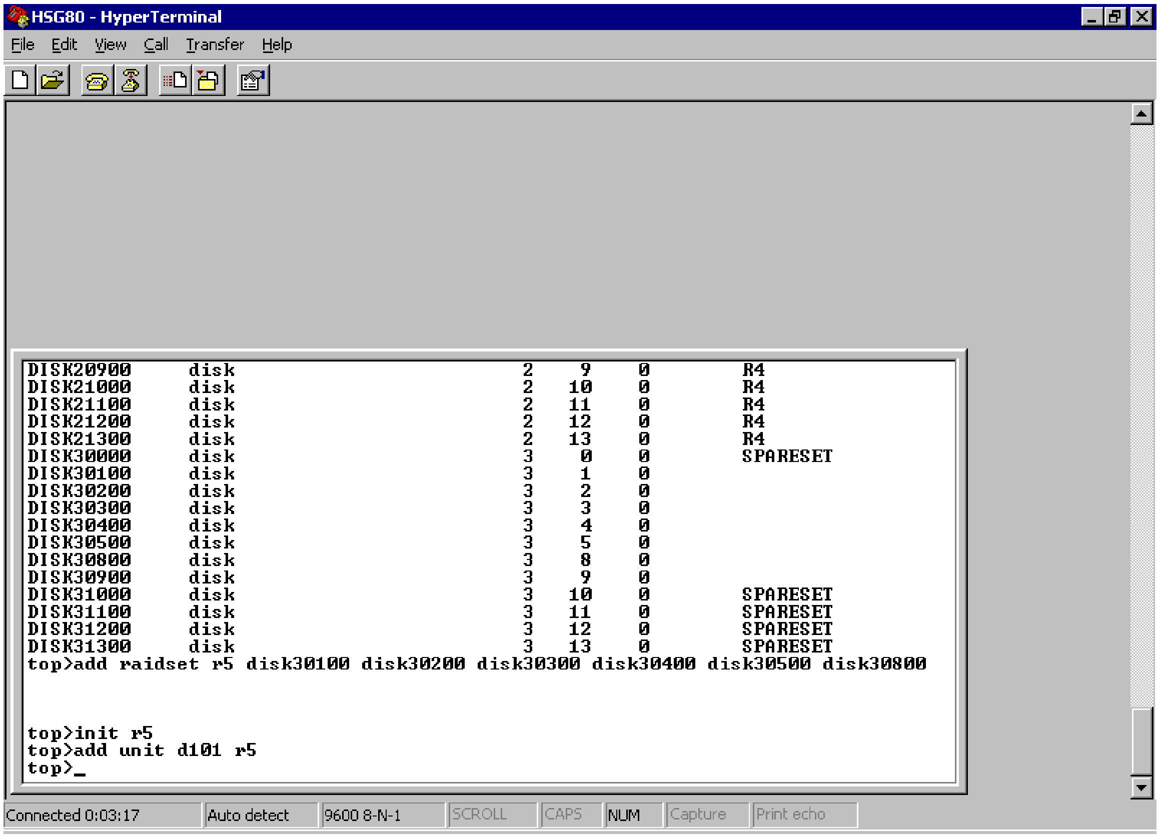
After the RAIDset has been created, it is necessary to initialize the RAIDset. The initialize command destroys any existing metadata or data on the specified disks in the RAIDset, and reserves a small amount of disk space for new metadata. The following command is used to initialize a RAIDset named r1:

```
prompt > initialize r1
```

After the RAIDset has been initialized, it is necessary to create a unit and assign the RAIDset to that unit. The following command is used to create a unit called d1 and assign the RAIDset r1 to it:

```
prompt > add unit d1 r1
```

After the unit has been created and a diskset has been assigned to it, the disk device is listed in SWVR and a pool can be created from that disk device using the New Pool wizard.



```
DISK20900     disk                          2    9    0        R4
DISK21000     disk                          2   10    0        R4
DISK21100     disk                          2   11    0        R4
DISK21200     disk                          2   12    0        R4
DISK21300     disk                          2   13    0        R4
DISK30000     disk                          3    0    0        SPARESET
DISK30100     disk                          3    1    0
DISK30200     disk                          3    2    0
DISK30300     disk                          3    3    0
DISK30400     disk                          3    4    0
DISK30500     disk                          3    5    0
DISK30800     disk                          3    8    0
DISK30900     disk                          3    9    0
DISK31000     disk                          3   10    0        SPARESET
DISK31100     disk                          3   11    0        SPARESET
DISK31200     disk                          3   12    0        SPARESET
DISK31300     disk                          3   13    0        SPARESET
top>add raidset r5 disk30100 disk30200 disk30300 disk30400 disk30500 disk30800



top>init r5
top>add unit d101 r5
top>_
```

Figure 20: Creation of a RAIDset and a unit

After a unit is created, there are several switches that can be specified. This section provides a brief description of the switches available with the creation of a unit. For further information on these commands, refer to the documentation provided with the HSG80 Array Controllers. Some of the switch commands are:

- enable_access_path (default) and disable_access_path

  These commands determine which host connections can access the unit.

  - The default condition is that access paths to all host connections are enabled.
  - To restrict host access to a set of host connections, specify **disable_access_path=all** when the unit is added, and then use the **set unit** command to specify the set of host connections that are to have access to the unit.

- maximum_cached_transfer

  This switch affects both read and write-back cache.

  - Sets the largest number of write blocks to be cached by the controller.
  - The controller will not cache any transfers over the set size.
  - Acceptable write block sizes are 1 through 1024.

- preferred_path=other_controller and preferred_path=this_controller

  - The preferred controller is the controller through which the unit is initially on line.
  - The default setting is that there is no preferred path.
  - Specify a preferred path when load distribution is desired.

- read_cache (default) and noread_cache

  If the read cache switch is selected, when the controller receives a read request from the host, the controller reads the data from the disk drives, delivers it to the host, and stores the data in its cache module.

  - Subsequent reads for the same data take the data from cache rather than accessing the data from the disks.
  - Read caching improves performance in almost all situations.
  - Under certain conditions, such as when performing a backup, read caching may not be necessary because only a small amount of data is cached.

- readahead_cache (default) and noreadahead_cache

  The read ahead cache switch enables the controller to keep track of read I/Os.

  - If the controller detects sequential read I/Os from the host, the controller will then try to keep ahead of the host by reading the next sequential blocks of data (those the host has not yet requiesced), and put the data in cache.
  - The controller can detect multiple sequential I/O requests across multiple units.
  - Read-ahead caching improves host application performance because the data is read from the controller cache rather than from the disk array.
  - Read-ahead caching is the default switch setting for units. If you are adding a unit that is not expected to receive sequential I/O requests, select noreadahead_cache for the unit.

- write_protect (default) and nowrite_protect

  The write protect switch specifies whether data contained on the selected unit can be overwritten.

  - Specify write_protect to prevent host write operations to the unit. Some exceptions are that the controller may still write to a write-protected RAIDset to satisfy a reconstruct pass or to reconstruct a newly replaced member. Additionally, metadata, reconstruct, and copy writes are still allowed to RAIDsets and mirrorsets.

  - Specify `nowrite_protect` to allow the host to write data to the unit. This command allows the controller to overwrite existing data. `nowrite_protect` is the default for transportable disks.

- `writebac_cache` and `nowriteback_cache`
  - The writeback cache switch allows the controller to declare the write operation "complete" as soon as the data reaches its cache memory. The controller performs the slower operation of writing the data to the disk drives at a later time. The no writeback cache switch enables only write-through chaching.
  - In write-through caching, when the controller receives a write request from the host, the controller places the data in its cache module, writes the data to the disk drives, and then notifies the host when the write operation is complete
  - This process is called write-through caching because the data passes through and is stored in the cache memory on its way to the disk drives. Write-through caching is enabled only when write-back caching is disabled.

CAUTION: Though there is built-in redundancy to protect data contained in cache, allowing data to be written to write-back cache may result in the loss of data.

Unit Offsets and Connections

Each path between a host computer Fibre Channel HBA and an active host port on a controller is a connection. In the TaskSmart N-Series Cluster, each cluster node has two host bus adapters (HBA). This creates a total of eight connections: two for each host bus adapter and four for each cluster node. Data is accessed through these connections between the cluster nodes and the controllers. Each controller has two ports: port 1 and port 2. Unit offsets are used to establish the beginning of the range of units that a host connection can access. They define and restrict access to a contiguous group of unit numbers. Host connections on port 1 are set to an offset of 0; port 1 connections can see units 0 through 99. Host connections on port 2 are set to an offset of 100; port 2 connections can see units 100 through 199. For example:

- Conn1, conn2, conn3, and conn4 are all connected to port 1 on each of the two controllers. They each have a unit offset of 0. All units in the d0 through d7 range use port 1 and the connections associated with port 1 (conn1, conn2, conn3, and conn4).
- Conn5, conn6, conn7, and conn8 are all connected to port 2 on each of the two controllers. They each have a unit offset of 100. All units in the d100 through d107 range use port 2 and the connections associated with port 2 (conn5, conn6, conn7, and conn8).

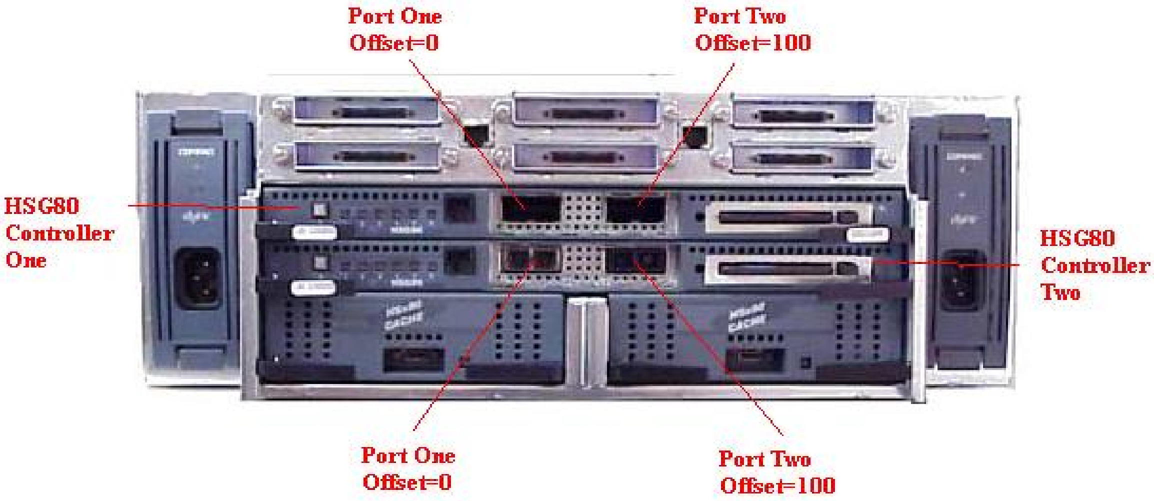


Figure 21: View of the back of a StorageWorks enclosure model 2200

Moving Disk Drives

Disk drives should not be moved or rearranged in the storage enclosures. It is not possible to migrate data by moving a storage enclosure from a TaskSmart N2400 appliance to the TaskSmart N-Series Cluster. Data integrity will not be maintained. The reason for this limitation is that when the controller configures drives in a RAIDset or mirrorset, the controller reserves a small section for metadata. This section of metadata is the method by which the controller identifies the RAIDsets or mirrorsets. If a new storage enclosure is moved into the TaskSmart N-Series Cluster, then that metadata section is missing, and the controller must initialize any new RAIDsets or mirrorsets. The initialization of RAIDsets and mirrorsets is data destructive and all data will be lost. It is possible to move a storage enclosure into the TaskSmart N-Series Cluster if data integrity is not a requirement. To accomplish this, connect the new storage enclosure to the Model 2200 enclosure and use the command `run config` to scan for the new disk devices. The controller will recognize the new disk devices which can be distributed into storagesets. The procedure for moving a single disk out of a RAIDset or mirrorset, is described in the following section entitled “Replacing Disk Drives.”

Replacing Disk Drives

When a disk drive fails or must be replaced, there is a specific procedure for disk removal and replacement. When a disk fails, the controller automatically removes the disk from the RAIDset and places it into a failedset. If the diskset has a disk replacement policy, and a spareset is present, then the controller automatically inserts the spareset into the RAIDset. There are three options for disk replacement policies:

- `policy=best_fit` (default) chooses a replacement disk drive from the spareset that equals or exceeds the base member size (the smallest disk drive at the time the mirrorset was initialized). If there is more than one disk drive in the spareset that meets the criteria, the controller selects a disk drive with the best performance.

- policy=best_performance allows the software to choose a replacement disk drive from the spareset with the best performance. The controller attempts to select a disk on a different port than existing mirrorset members. If there is more than one disk drive in the spareset matching the best performance criteria, the controller selects a disk drive that equals or exceeds the base member size.
- nopolicy prevents the controller from automatically replacing a failed disk device. The mirrorset operates in a reduced state until the administrator selects either policy=best_fit, or policy=best_performance, or a member is manually placed in the mirrorset.

  An example of policy command is:

  ```
  add raidset r8 disk10100 disk20100 disk30100 disk40100 policy=best_fit
  ```

  To change the policy of a RAIDset after it has been created, the following command can be used:

  ```
  set r8 policy=best_fit
  ```

The following scenarios outline the methods that should be used to replace a failed disk on storage enclosure that is managed by an HSG80 Array Controller.

Scenario 1: Mirrorset or RAIDset with a Failed Disk

When a disk fails, the HSG80 Array Controller marks that disk as failed and places it into a failedset. To confirm this, enter **Show Devices** in the HyperTerminal console. If disk 10100 is failed, it will be in a failedset.

Note: For this example: r1 is a RAIDset and disk10100 is a failed disk from r1.

1. Physically remove the failed disk.
2. At the HyperTerminal console, enter the following commands:
   - show r1 (This command shows the mirrorset policy settings. Record this information.)
   - delete failedset disk10100 (This command deletes the failed disk from the failedset.)
3. Physically insert the new disk.
4. At the HyperTerminal console, enter the following commands:
   - set r1 nopolicy (This command disables any controller replacement mechanisms.)
   - run config (This command causes the controller to scan for disk devices.)
   - set r1 replace=disk10100 (This command adds diskxxxxx to the mirrorset.)
   - set r1 policy=*yyyyy* (This command sets the replacement policy of the RAIDset. *yyyyy* represents the previously recorded policy settings.)

Scenario 2: Mirrorset or RAIDset with a Failed Disk and a Spareset

Note: For this example, r1 is a RAIDset, disk10100 is a failed disk from r1, and disk11300 is a spareset.

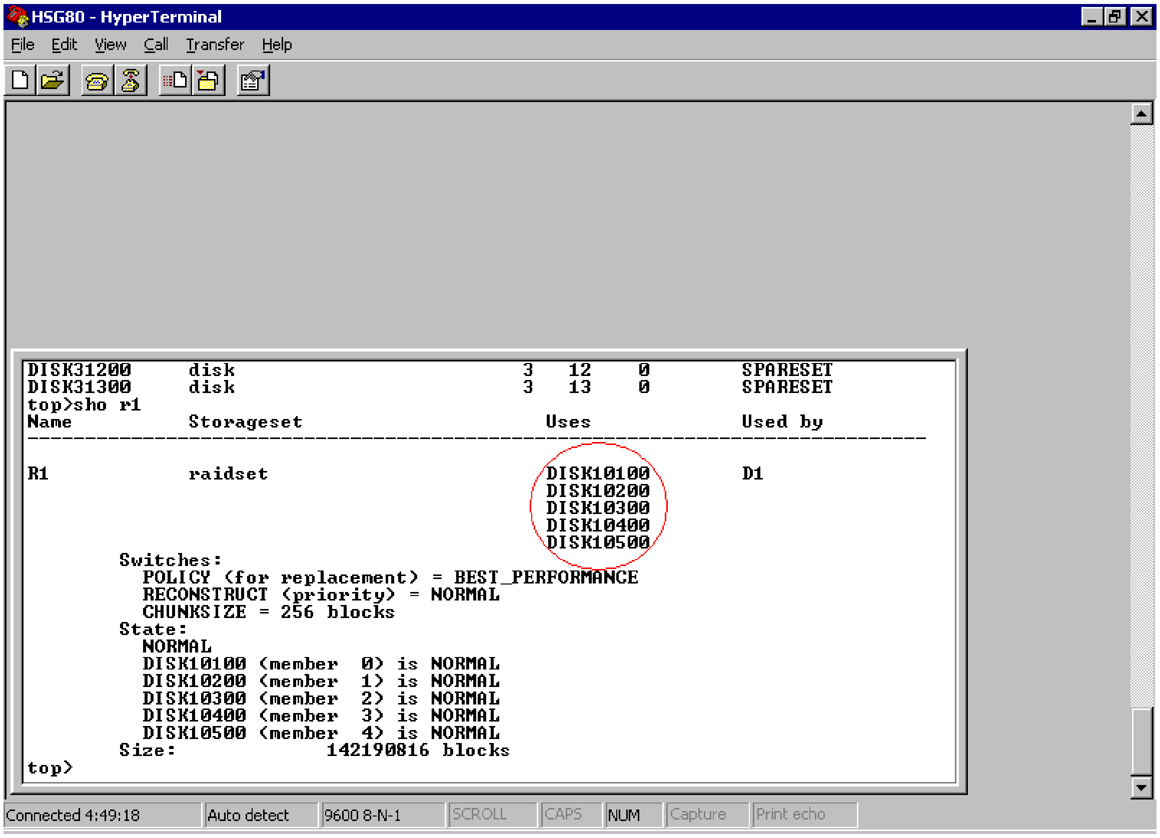
1. Physically remove the failed disk.
2. At the HyperTerminal console, enter the following command:
   - show r1 (This command shows the mirrorset policy settings. Record this information.)

- delete failedset disk10100 (This command deletes the failed disk from the failedset.)

3. Physically insert the new disk.
4. At the HyperTerminal console, enter the following commands:
   - run config (This command causes the controller to scan for any new disk devices.)
   - set r1 nopolicy(This command disables any controller replacement mechanisms.)
   - set r1 remove=disk11300 (This command removes the spare disk that was automatically added into the set.)
   - set r1 replace=disk10100 (This command adds the original disk back into the set.)
   - delete failedset disk11300 (This command deletes the failed disk from the failedset.)
   - add spareset disk11300 (This command adds the original spareset disk back into a spareset.)
   - set r1 policy=*yyyyy* (This command sets the replacement policy of the RAIDset. *yyyyy* represents the preferred settings as recorded in step 1.)

Moving Arrays

Arrays or RAIDsets should not be moved within the storage subsystem. When an array has been created, the array is dependent on the disk locations that were specified at the creation of the array. The exception to this rule is when a drive in the array fails and a replacement policy is in effect. In this case, a spare is automatically inserted into the array. It is also not possible to add another disk into an array for capacity expansion. Figure 22 shows the configuration of a RAIDset called r1. The show r1 command was used to display the information about r1.



```
DISK31200     disk                          3   12     0        SPARESET
DISK31300     disk                          3   13     0        SPARESET
top>sho r1
Name          Storageset                     Uses               Used by
------------------------------------------------------------------------------

R1            raidset                        DISK10100          D1
                                             DISK10200
                                             DISK10300
                                             DISK10400
                                             DISK10500
        Switches:
          POLICY (for replacement) = BEST_PERFORMANCE
          RECONSTRUCT (priority) = NORMAL
          CHUNKSIZE = 256 blocks
        State:
          NORMAL
          DISK10100 (member  0) is NORMAL
          DISK10200 (member  1) is NORMAL
          DISK10300 (member  2) is NORMAL
          DISK10400 (member  3) is NORMAL
          DISK10500 (member  4) is NORMAL
        Size:            142190816 blocks
top>
```

Figure 22: RAIDsets are dependent on particular disk locations

The controllers write data to the specific port-target-LUN that is specified in the name of the disk device. For a disk named 10100, the first number designates port 1, the second two numbers designate target 01, and the last two numbers designate LUN 00. It is not possible to physically move an entire RAIDset or mirrorset to another location in a storage enclosure. If a RAIDset were to be moved, the controller would not be able to read the metadata on the disks and the RAIDset would be rendered useless. The RAIDset would have to be recreated and data would be lost. If it is necessary to rearrange disks in an enclosure, a full backup of all data is required.

Capacity Expansion

Expanding the storage capacity of the TaskSmart N-Series Cluster is something that must be planned in advance. Unplanned additions cannot be expected to work efficiently or effectively. Adding additional StorageWorks 4314 storage enclosures and drives is simple, but effectively employing them requires more forethought and configuration effort.

The simplest example is when the horizontal array configuration is being used. In this case, it is a simple matter to add the additional storage enclosures and drives, and then configure the new RAIDsets and units.

The more complicated example is when the vertical array configuration is being used. Vertical array configurations provide the highest fault tolerance. See the *Compaq TaskSmart N-Series Cluster Planning Guide* for details. When using vertical arrays, the administrator is left with only three options for capacity expansion:

- The first option is if the storage was planned and configured properly for future expansion. Consult the *Compaq TaskSmart N-Series Cluster Planning Guide* for details of this method. Also refer to the following section, “Expanding RAIDsets,” on how to perform the expansion.
- The second option is to use horizontal array configurations for the additional StorageWorks 4314 storage enclosures and their drives. This option entails giving up some of the high-availability and performance capabilities of vertical array configurations, but allows the administrator to nondestructively add additional storage.
- The third and most drastic option is to back up all of the data in the TaskSmart N-Series Cluster, verify that the backup completed successfully and the data can be restored from it, and then destroy and recreate the array configuration. This method is preferred from a storage standpoint because it allows the administrator to maintain a vertical array configuration. However, it requires significant down time to reconstruct, and may not be feasible.

For details on the proper planning of storage strategies and configuration for future growth, consult the *Compaq TaskSmart N-Series Cluster Planning Guide*. For the mechanical details of how to expand this configuration, see the “Expanding RAIDsets” subsection later in this guide.

Expanding RAIDsets

This section details the process of expanding RAIDsets in a vertical array configuration. The administrator must have configured the RAIDsets in the manner documented in the *Compaq TaskSmart N-Series Cluster Planning Guide*. See this guide for details of the planning and setup.

Earlier in this guide, the “Replacing Disk Drives” section detailed how to replace failed drives. The method for expanding the RAIDsets vertically is very similar. In essence, the administrator must manually fail one of the drives in a storage enclosure that contains more than one RAIDset member disk. The administrator then replaces the failed drive with a drive from the new storage enclosure. The administrator waits for the RAIDsets to rebuild, and then fails the next drive. Figure 23 is included in the *Compaq TaskSmart N-Series Cluster Planning Guide*, but is also included here to help illustrate this process.

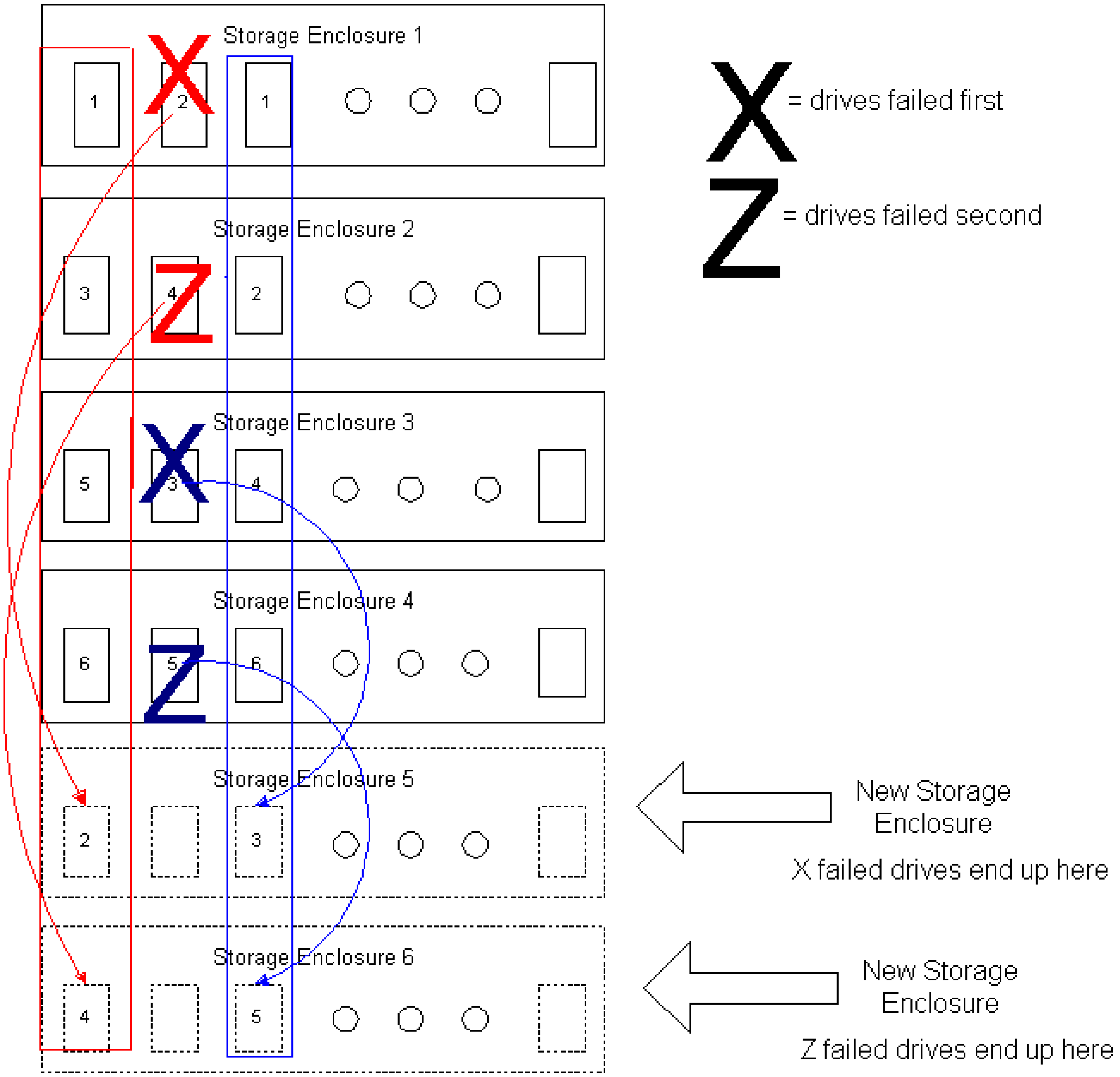


Figure 23: Expanding vertical RAIDsets

In the example in Figure 23, the process would be:

1. Using HyperTerminal, connect to the TOP controller.
2. At the TOP> prompt, set the replacement policy for RAIDsets r1 and r2 to nopolicy using the following command:

   prompt> set r2 nopolicy (Change **r2** to proper RAIDset name.)

Note: Record the original policy before setting it to **nopolicy**, so that the administrator can restore the original setting when this process is complete.

3. Issue a remove command on disk 2 in r1 and disk 3 in r2. This places the removed disk into a failed set. Use the following command:

   prompt> set *rx* remove = disk*yyyyy* (Change *rx* and *yyyyy* to proper names.)

4. Physically remove the failed disk.

5. Delete the failed disk by entering the following command:

   `prompt> delete failedset disk`*yyyyy* (Change *yyyyy* to the proper name.)

6. Issue a `replace` command on a disk in storage enclosure 5 that is in line with each array. Figure 23 lists the drives as their original drive numbers, in this case 2 and 3. Enter the following command:

   `prompt> set` *rx* `replace = disk`*yyyyy* (Change *rx* and *yyyyy* to proper names.)

7. Physically reinsert the disk that was removed in step 4.

8. Wait for the RAIDsets to rebuild. Issue the `show rx` command, to display the status. (**x** is the RAIDset that is being rebuilt.)

9. Repeat steps 1 through 8 for disk 4 in r1 and disk 5 in r2.

10. Set the replacement policy for units r1 and r2 to back to their original setting:

    `prompt> set` *Rx* `policy=`*xxxxx* (Change *rx* to the proper RAIDset name, and change *xxxxx* to the RAIDset original policy as recorded previously.)

Administrators cannot adjust these general steps. However, the RAIDset configuration will vary depending on the environment. For example, the current setup could already have five StorageWorks 4314 storage enclosures, or four StorageWorks 4314 storage enclosures with the addition of a single new StorageWorks 4314 storage enclosure.

See the *Compaq TaskSmart N-Series Cluster Planning Guide* for information on how to initially configure the RAIDsets, and why an administrator would want to use this configuration.

Port Security

Because the storage subsystem has more than one host connection, each access path must be specified carefully to avoid giving undesirable host connections access to units. The default condition is that access paths to all host connections are enabled. To restrict host access to a set of host connections, specify `disable_access_path=all` when the unit is added, and then use the **set** unit command to specify the set of host connections that are to have access to the unit. This command grants cluster nodes sole access to the storage enclosures. Keep in mind that a connection is established from the HBA of a cluster node to the HSG80 Array Controller. This means that someone can not plug a fiber connection into the SAN switch and still have access to the storage enclosures.

For example, to create unit d5 from mirrorset m1, with four host connections (conn1, conn2, conn3, conn4) having access to this unit, enter:

```
add unit d5 m1 disable_access_path=all

set unit d5 enable_access_path=conn1, conn2, conn3, conn4
```

Connections are granted by the port and the unit offset for that port. All port 1 connections have a unit offset of 0 and all port 2 connections have a unit offset of 100. All units from D0 through D7 use the unit offset of 0 and all units D100 through D107 use the unit offset of 100. This means that all units D0 through D7 grant access to all port 1 connections, and all units D100 through D107 grant access to all port 2 connections.

SAN Switch Zoning

Compaq StorageWorks SAN switches provide a fabric management feature called SAN switch zoning, which provides the ability to split the switch into zones. Each zone is essentially a virtual fabric. When used in addition to the port security feature of the HSG80 Array Controllers, switch zoning adds an additional level of information management and data security. A node can access only other nodes that are in the same zone. Any node outside of the zone is unaware of the existence of the nodes inside the zone. All devices connected to a fabric may be configured into one or more zones.

Every zone has a name that begins with a letter, which can be followed by letters, digits and the underscore character "_." Names are case sensitive. For example, "Zone_1" and "zone_1" are different zones. Spaces are not allowed. Every zone has a member list, consisting of one or more members. Empty zones are not allowed. The maximum number of zones and the maximum number of members in a zone are constrained by memory usage. Because these limits are far larger than the number of devices connected to a fabric, they are effectively unlimited.

Zone definitions are persistent. This means that the definition remains in effect across restarts and power cycles until it is deleted or changed. A device can be a member of multiple zones. Zoning management is performed using Telnet via either out-of-band or in-band communication by logging onto a switch. Any switch in the Fabric may be used. A change made to the Zoning information on one switch is replicated through all fabric switches. For more information on SAN switch zoning refer to the *Compaq StorageWorks Fibre Channel Switch Management Guide* available at the following link:

www.compaq.com/products/storageworks/techdoc/hubs-and-bridges/AA-RMMJA-TE.html

StorageWorks Command Console

In this section, the advanced uses of StorageWorks Command Console (SWCC) are discussed. The StorageWorks Command Console offers administrators two separate applications. When an administrator starts SWCC, the following window is displayed.

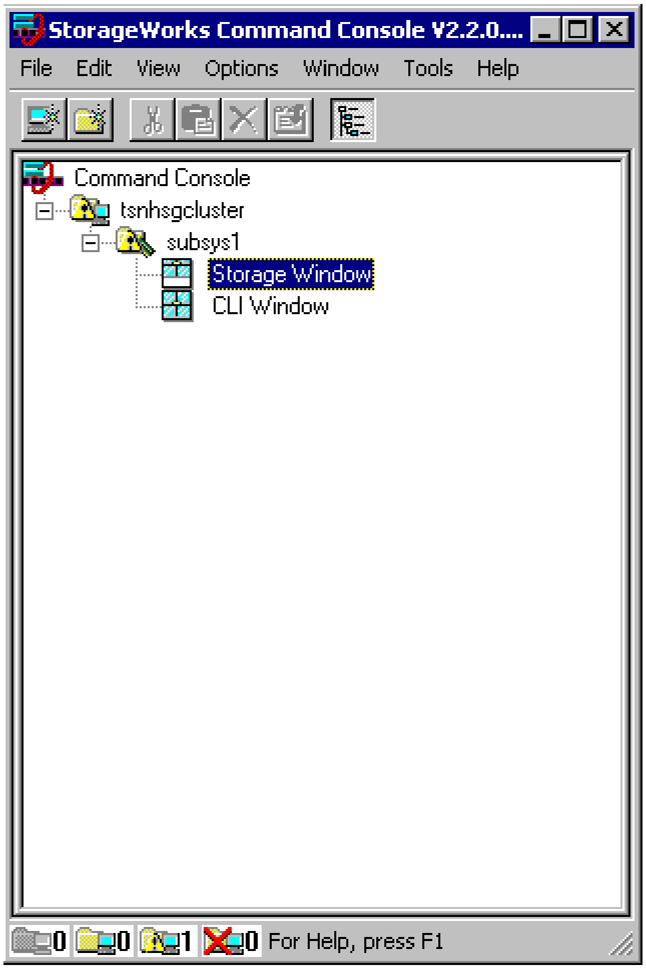


Figure 24: Navigation window

The **Storage** window and the **CLI** window are launched from the **Navigation** window.

Storage Window

The **Storage** window is a graphical interface that depicts all of the disks that are known by the system. From this window, the administrator creates RAIDsets, updates controller software, alters the controller configuration, and modifies devices.

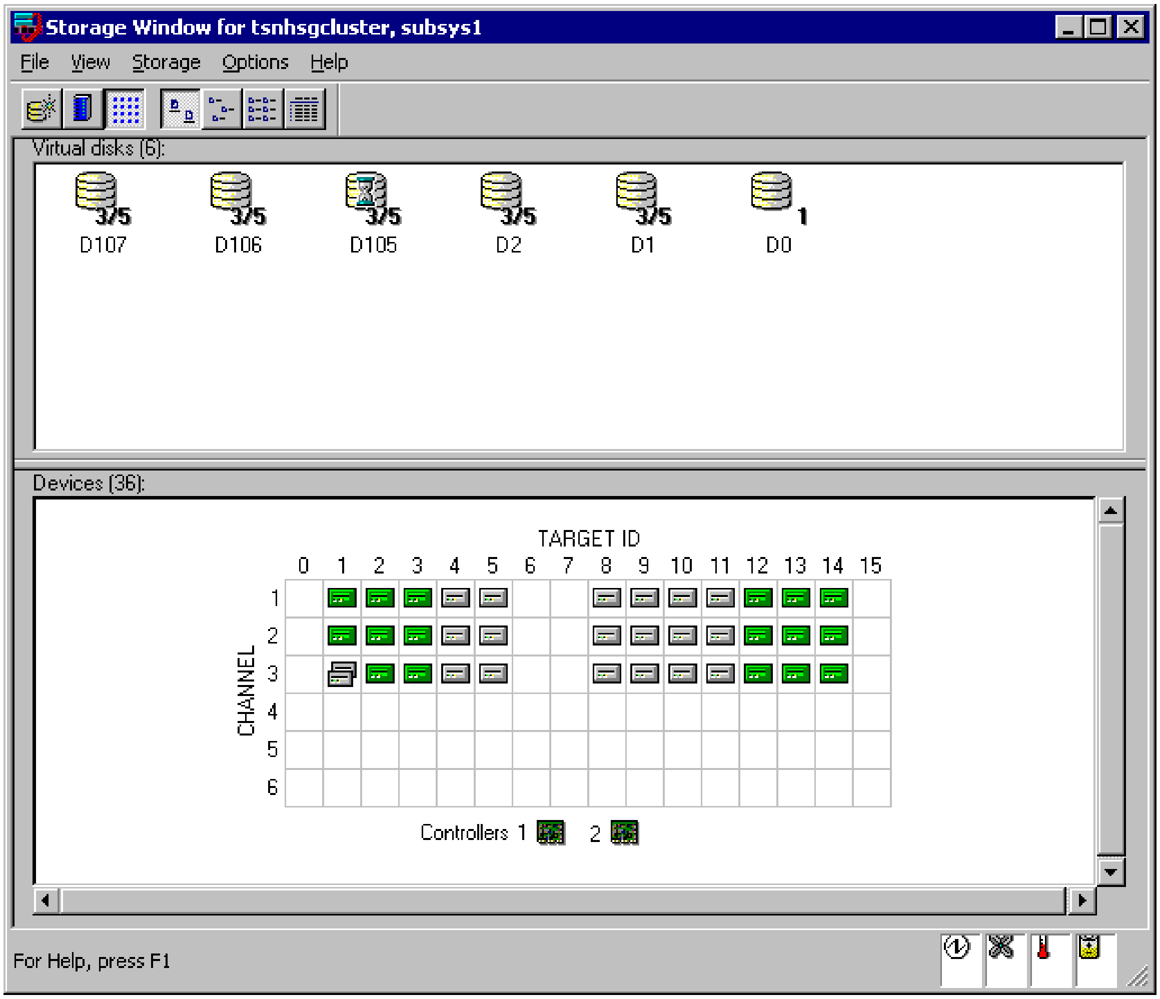
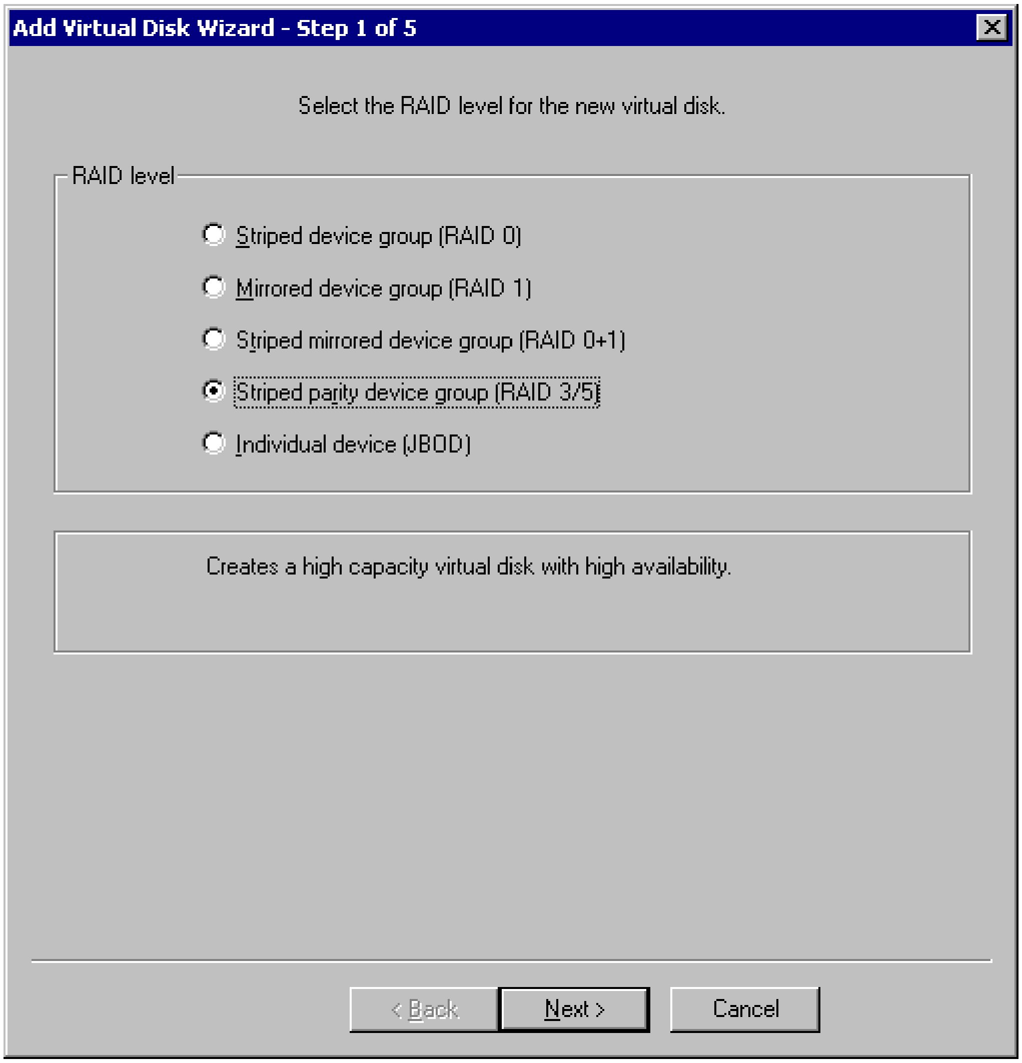


Figure 25: Storage window

In SWCC, RAIDsets are called virtual disks, which are not the same as SWVR virtual disks. To create RAIDsets, click the **Storage** menu and choose **add virtual disks**. If this is the first use of this instance of the **Storage** window, a dialog box is displayed asking for the password that was set up for the StorageWorks Agent. At the first screen, choose the RAID level that needs to be created.

The **Devices** menu of the **Storage** menu facilitates the adding and deleting of disk drives, making and removing spares, and locating the disks. A physical disk or disk unit must be selected before you attempt to use these commands. Another way to add or remove spares is to right click the disk, and then select **Add Spare** or **Remove Spare**. The administrator can also update the controller software through the **Storage** menu. The controller configuration can be restarted, loaded, and saved through the **Storage** menu.


Add Virtual Disk Wizard - Step 1 of 5

Select the RAID level for the new virtual disk.

RAID level

○ Striped device group (RAID 0)

○ Mirrored device group (RAID 1)

○ Striped mirrored device group (RAID 0+1)

● Striped parity device group (RAID 3/5)

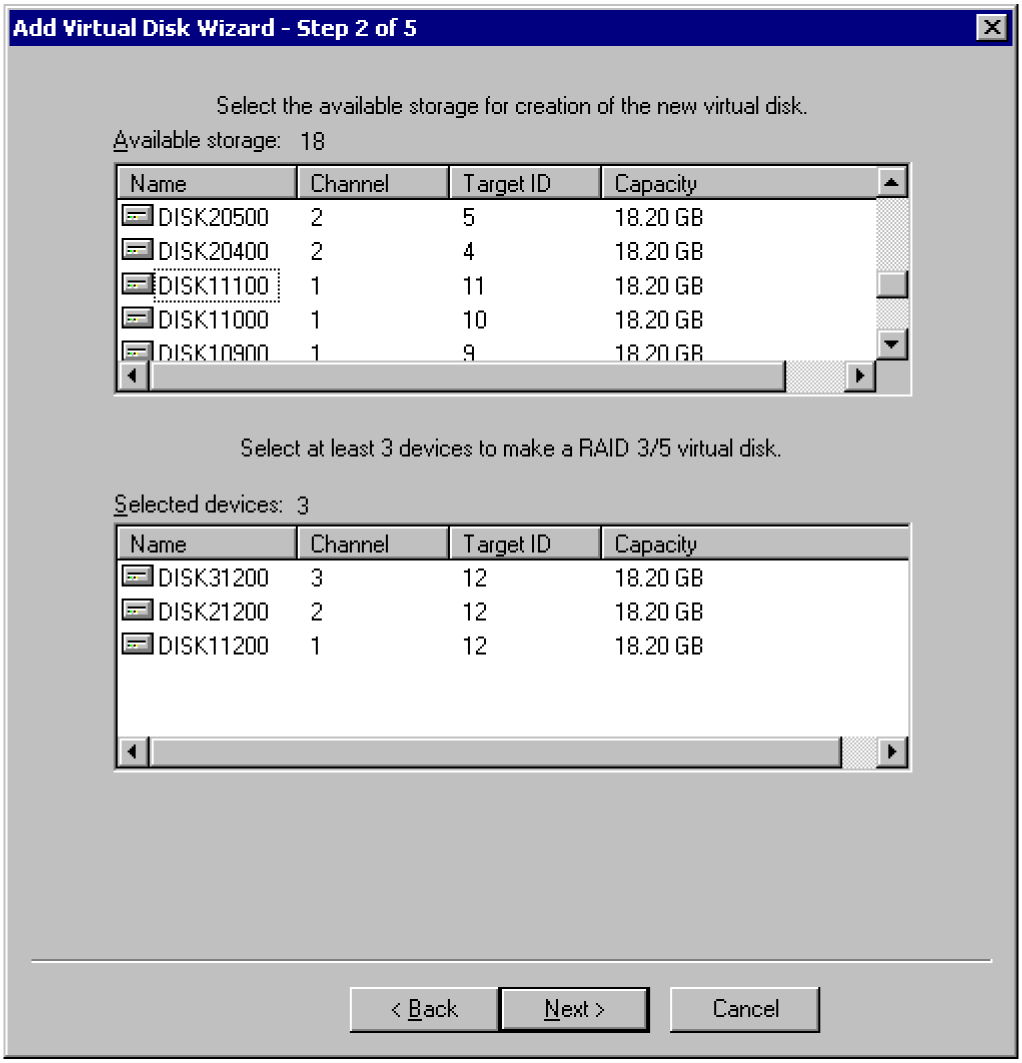
○ Individual device (JBOD)

Creates a high capacity virtual disk with high availability.

Figure 26: SWCC Virtual Disk wizard - Step 1 of 5

SWCC allows RAID 0, RAID 1, RAID 0+1, RAID 3/5, and JBOD to be created as shown in Figure 26. At the next screen, select the disks to be included in the RAIDset.

If the proper number of disks is not selected, SWCC displays an error requiring the administrator to select more disks. On the next screen, the minimum and maximum capacities of the virtual disk are shown.



Select the available storage for creation of the new virtual disk.

Available storage: 18

Name	Channel	Target ID	Capacity
DISK20500	2	5	18.20 GB
DISK20400	2	4	18.20 GB
DISK11100	1	11	18.20 GB
DISK11000	1	10	18.20 GB
DISK10900	1	9	18.20 GB

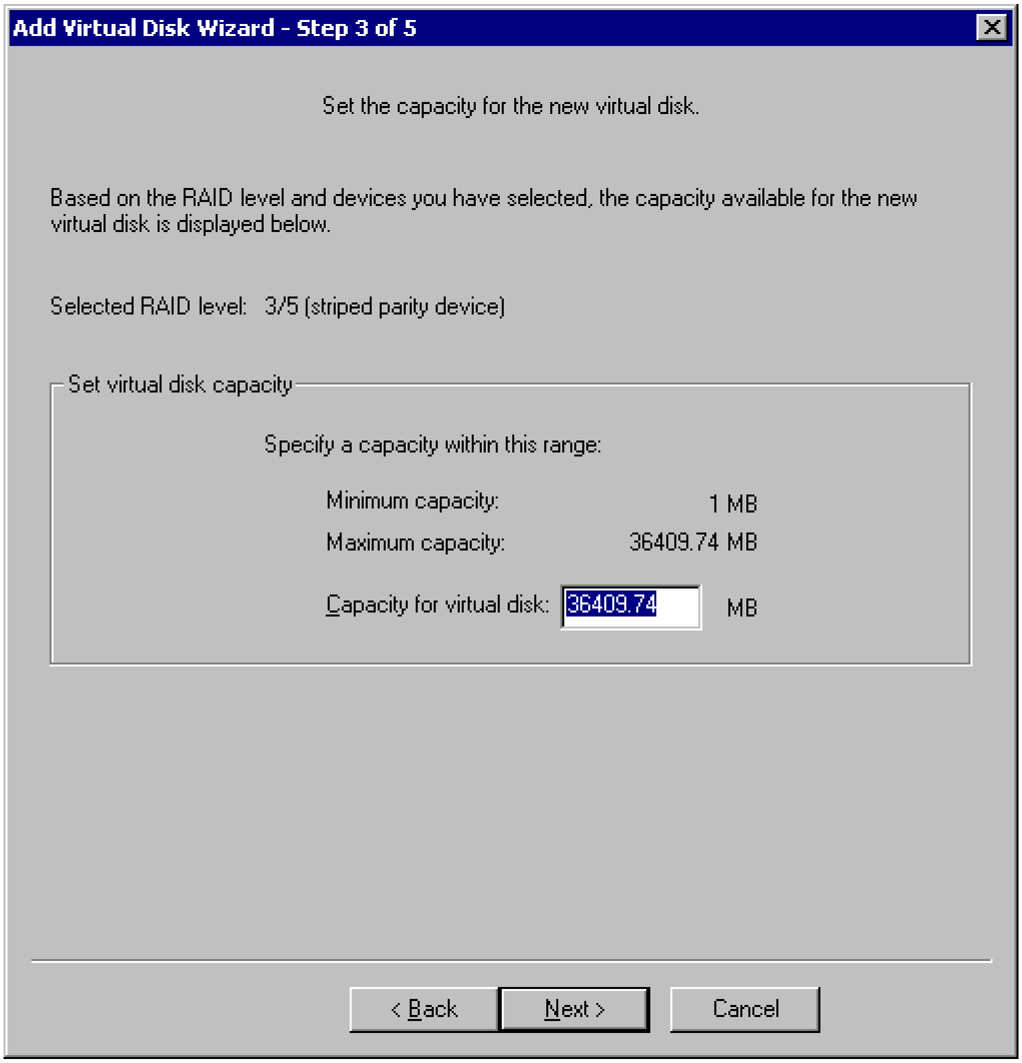
Select at least 3 devices to make a RAID 3/5 virtual disk.

Selected devices: 3

Name	Channel	Target ID	Capacity
DISK31200	3	12	18.20 GB
DISK21200	2	12	18.20 GB
DISK11200	1	12	18.20 GB

Figure 27: SWCC Virtual Disk wizard - Setup 2 of 5

Note: JBODs are not supported for use on the Compaq TaskSmart N-Series Cluster.



Set the capacity for the new virtual disk.

Based on the RAID level and devices you have selected, the capacity available for the new virtual disk is displayed below.

Selected RAID level: 3/5 (striped parity device)

Set virtual disk capacity

Specify a capacity within this range:

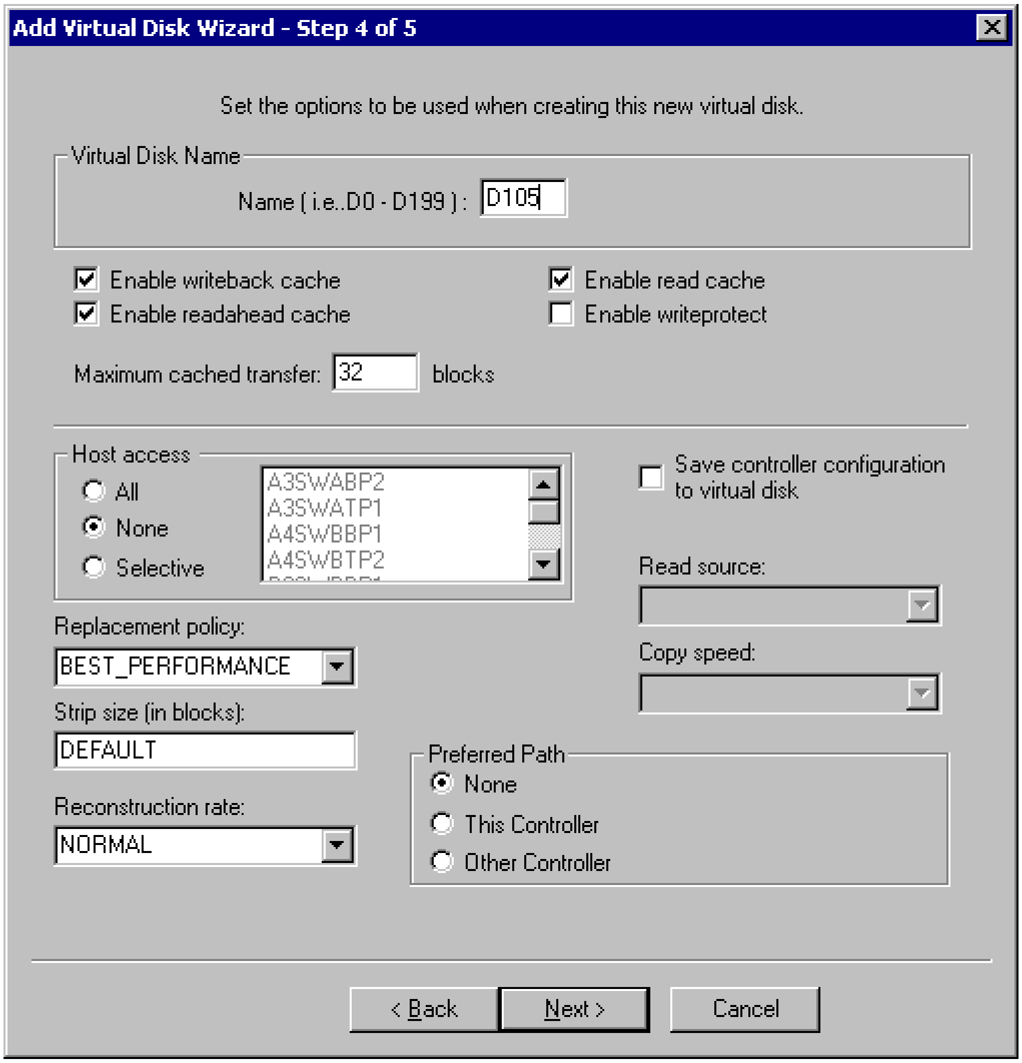
Minimum capacity: 1 MB

Maximum capacity: 36409.74 MB

Capacity for virtual disk: 36409.74 MB

Figure 28: SWCC Virtual Disk wizard - Step 3 of 5 (selecting the disk capacity)

Use the entire capacity of the virtual disk, because there is no advantage to dividing the space on a RAIDset. At the next screen, the names of the disk and host access are configured.



Add Virtual Disk Wizard - Step 4 of 5

Set the options to be used when creating this new virtual disk.

Virtual Disk Name

Name (i.e..D0 - D199) : D105

☑ Enable writeback cache
☑ Enable readahead cache
☑ Enable read cache
☐ Enable writeprotect

Maximum cached transfer: 32 blocks

Host access
○ All
● None
○ Selective

A3SWABP2
A3SWATP1
A4SWBBP1
A4SWBTP2

☐ Save controller configuration to virtual disk

Read source:

Copy speed:

Replacement policy: BEST_PERFORMANCE

Strip size (in blocks): DEFAULT

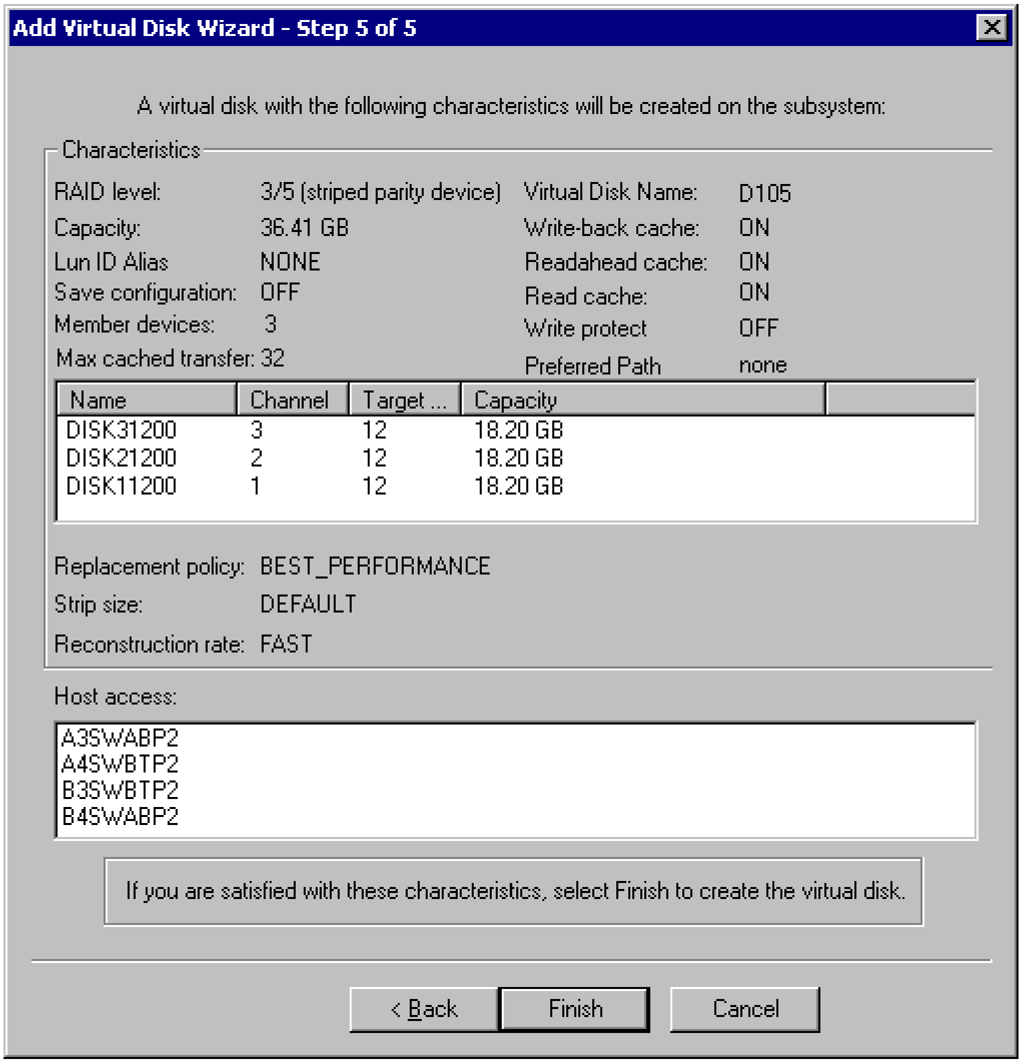
Reconstruction rate: NORMAL

Preferred Path
● None
○ This Controller
○ Other Controller

Figure 29: SWCC Virtual Disk wizard - Step 4 of 5 (setting virtual disk options)

In this screen, the administrator sets all of the policies for the disk. The final screen of the wizard shows all of the characteristics of the virtual disk that was created.

Note: A disk unit should not be used until it is finished initializing. This initialization may take up to several hours, depending on how many physical disks are members of the underlying RAIDset.



A virtual disk with the following characteristics will be created on the subsystem:

Characteristics

RAID level:	3/5 (striped parity device)	Virtual Disk Name:	D105
Capacity:	36.41 GB	Write-back cache:	ON
Lun ID Alias	NONE	Readahead cache:	ON
Save configuration:	OFF	Read cache:	ON
Member devices:	3	Write protect	OFF
Max cached transfer: 32		Preferred Path	none

Name	Channel	Target ...	Capacity
DISK31200	3	12	18.20 GB
DISK21200	2	12	18.20 GB
DISK11200	1	12	18.20 GB

Replacement policy: BEST_PERFORMANCE

Strip size: DEFAULT

Reconstruction rate: FAST

Host access:

A3SWABP2
A4SWBTP2
B3SWBTP2
B4SWABP2

If you are satisfied with these characteristics, select Finish to create the virtual disk.

Figure 30: SWCC Virtual disk wizard - step 5 of 5

Deleting System-Created 8-MB Disk Partitions

When configuring drives with the HGS80 Array Controller, the system may create a small 8-MB partition on each drive. While this partition may be viewed in the Logical Volume Manager (LVM), it cannot be deleted through the LVM. Compaq has provided a tool to delete these partitions. This tool is called **diskpart**.

CAUTION: If the **dilx** program has previously been executed on the drives, do not run the **dilx** program again. Executing the **dilx** program again will cause data destruction.

To run the **diskpart** program, do the following:

1. Go to the **Start** menu. Select **Programs**, **Tools**, and then click **diskpart**.
2. To display all of the volumes on the system, enter the following command:

   ```
   List Volume
   ```

   Then, press the **Enter** key. All volumes on the system are displayed.

3. Select the 8-MB partition to delete by entering the following command:

   select volume *x* (*x* represents the volume number of the partition to be deleted.)

   Then, press the **Enter** key.

CAUTION: The **diskpart** program is data destructive. If the wrong volume or partition is accidentally selected, the data residing on that volume or partition will be destroyed.

4. Delete the selected partition by entering the following command:

   delete volume

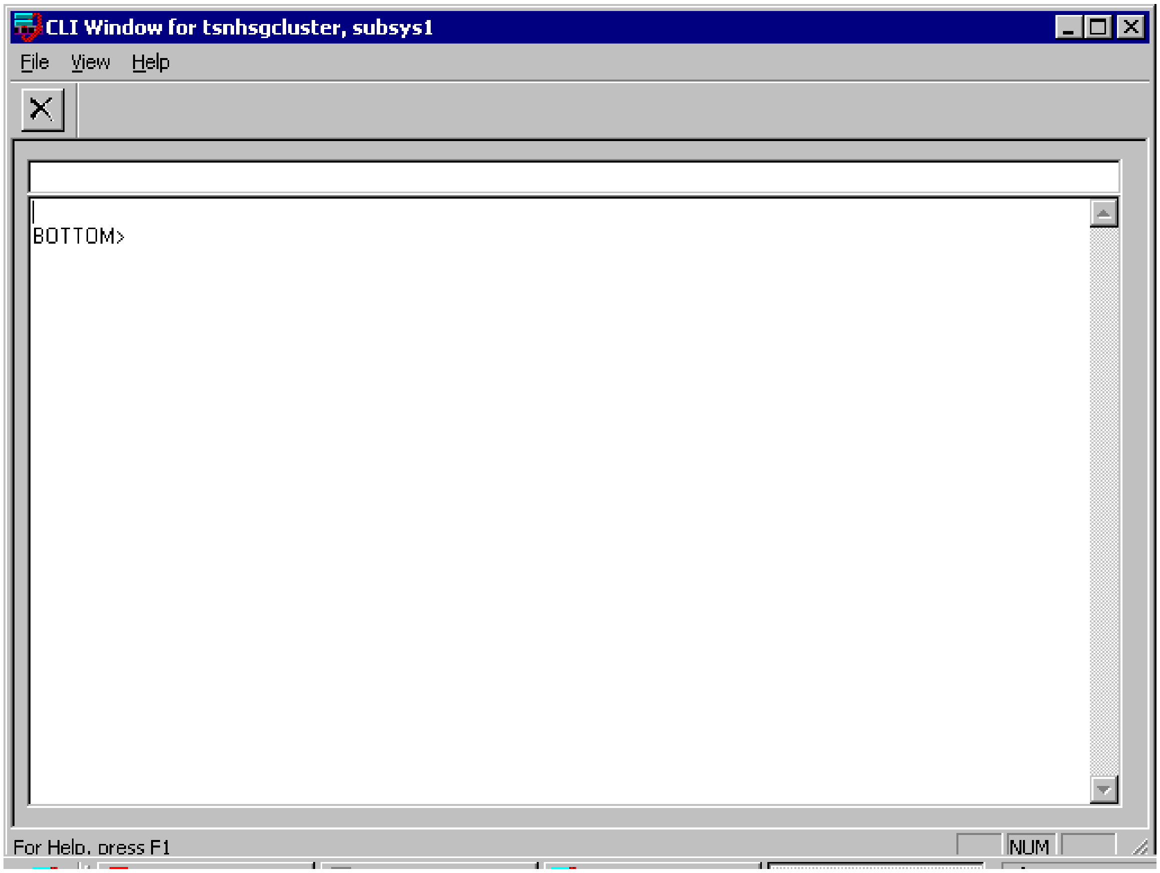
   Then, press the **Enter** key. A message is displayed indicating the name of the volume that has been deleted.

5. Repeat steps 3 through 4 for each 8-MB partition listed.

After diskpart has completed, all existing volumes can be seen by SWVR.

CLI Window

The **CLI** window offers a command line to configure the HSG80 Array Controller. This window can replace the use of HyperTerminal. The **CLI** window provides more commands and applications for configuration than the SWCC. Figure 31 depicts the **CLI** window.



```
BOTTOM>
```

Figure 31: CLI window

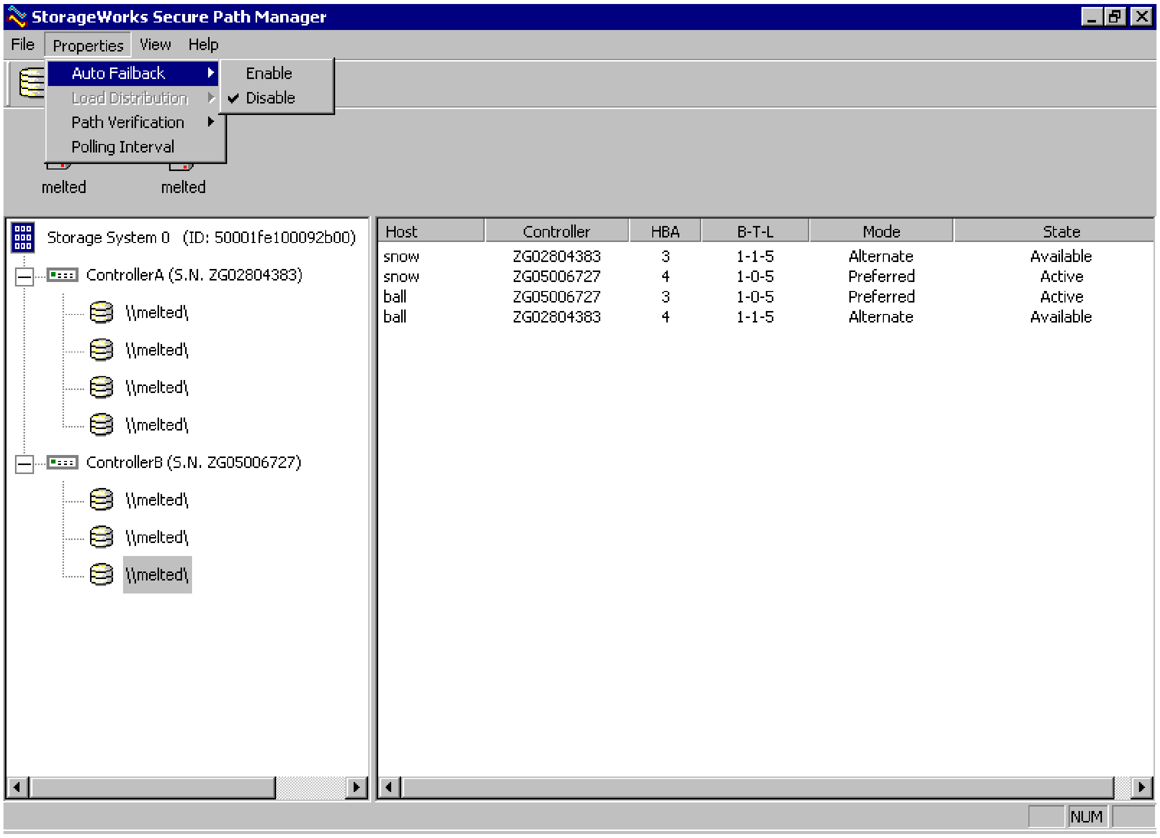
SecurePath and SecurePath Manager

This section discusses some advanced levels of administration using SecurePath Manager (SPM). The following are the **Properties** settings that can be altered for a storage profile. It is important to note that these properties have a global effect on all resources managed by an SPM storage profile. (See Figure 32 for details.) Use the **Properties** pull-down menu to:

- Enable or Disable the **Auto Failback** policy (default = disabled*)*. When Auto Failback is enabled, all storagesets that have failed over to an alternate path will automatically fail back to their preferred path when access to that path is restored. Storagesets fail back automatically only if I/O operations to those storagesets are in progress. Auto failback enabled in conjunction with Path Verification permits failback to occur for inactive storagesets.
- Enable or Disable the **Load Distribution** policy (default = disabled). Load Distribution allows multiple paths between a host and a specific storageset to be used in parallel for I/O, maximizing performance potential.

Note: Load Distribution is disabled in Microsoft Cluster Server (MSCS).

- Enable or Disable the **Path Verification** policy (default = enabled). With Path Verification enabled, SecurePath periodically runs diagnostics on all preferred and alternate paths to determine their current state. If a path is diagnosed with a problem that would prevent reliable I/O operations to complete, it is marked as FAILED and no further I/O operations are permitted on that path.
- Set the **Polling Interval** policy (default = 90 seconds). This setting determines the rate at which SPM will request configuration change information from the SecurePath Agent(s) in the storage profile. The Polling Interval setting affects only the rate at which displayed information is updated and has no effect on the current configuration. The polling interval is user selectable from a minimum of 5 seconds to a maximum of 30 minutes.



Storage System 0 (ID: 50001fe100092b00)

- ControllerA (S.N. ZG02804383)
  - \\melted\
  - \\melted\
  - \\melted\
  - \\melted\
- ControllerB (S.N. ZG05006727)
  - \\melted\
  - \\melted\
  - \\melted\

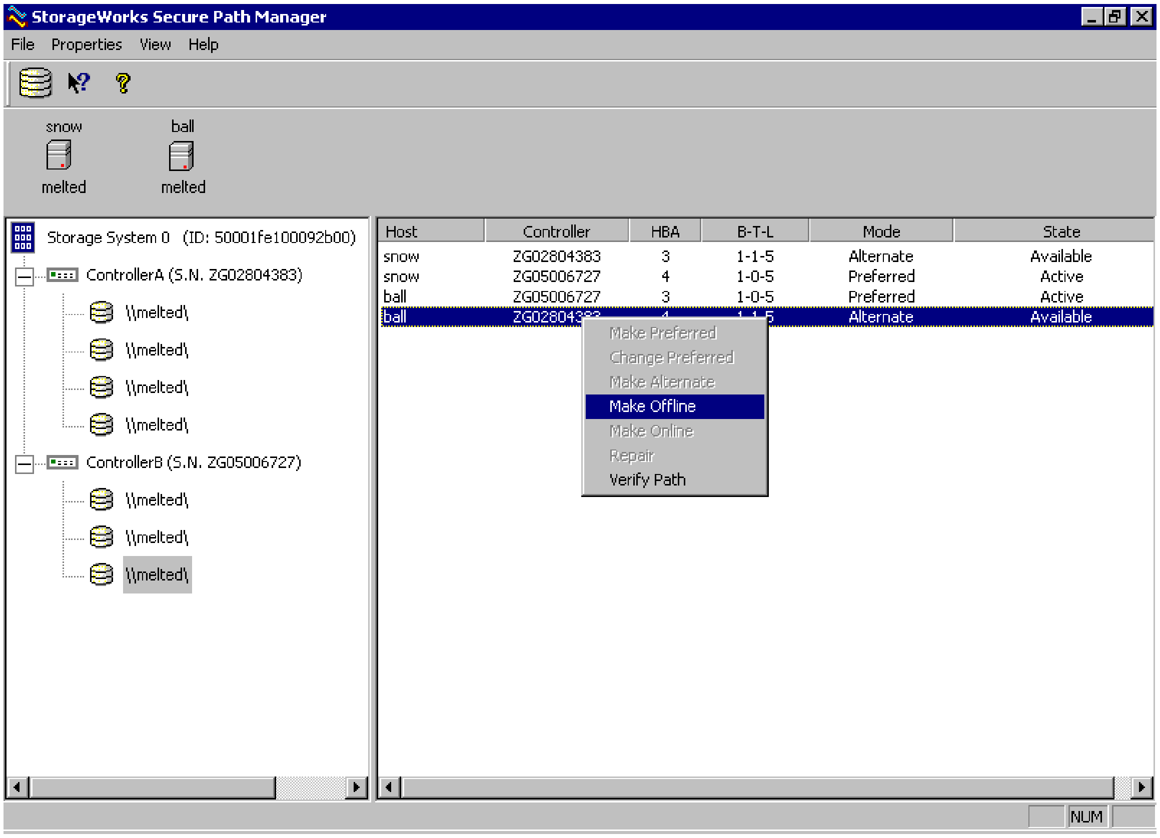
Host	Controller	HBA	B-T-L	Mode	State
snow	ZG02804383	3	1-1-5	Alternate	Available
snow	ZG05006727	4	1-0-5	Preferred	Active
ball	ZG05006727	3	1-0-5	Preferred	Active
ball	ZG02804383	4	1-1-5	Alternate	Available

Figure 32: SecurePath Manager properties

The following actions can be performed on the storagesets and paths managed by SPM:

- Moving A Storageset
- Making A Path Alternate
- Making A Preferred Path
- Changing A Preferred Path
- Making A Path Offline
- Making A Path Online
- Verifying A Path
- Repairing A Path

Figure 33 displays all of the options. (Only the **Make Offline** and **Verify Path** options are highlighted in this image.)



Host	Controller	HBA	B-T-L	Mode	State
snow	ZG02804383	3	1-1-5	Alternate	Available
snow	ZG05006727	4	1-0-5	Preferred	Active
ball	ZG05006727	3	1-0-5	Preferred	Active
ball	ZG02804383	4	1-1-5	Alternate	Available

Figure 33: Action options for a path

Moving a Storageset

Select **Move a Storageset** to change the ownership from the current RAID Array controller to the other. This action is useful to load balance I/O across controllers or to manually return a failed over storageset to its preferred path when Auto-Failback has been disabled. There are two methods available to move a storageset.

- Click the drive to highlight it in the storage system view.
- Drag the drive to the other controller, or right click to select **Move To Other Controller**.

Making a Path Alternate

To disable I/O operations to one or more paths, select **Make a Path Alternate** when Load Distribution is enabled. To make a path alternate:

1. Click the preferred path to change.
2. Select **Make Alternate**.

Making a Path Preferred

To re-enable I/O operations to a path that has previously been disabled using **Make Alternate**, select **Make a Path Preferred** when Load Distribution is enabled. To make a path preferred:

1. Click the alternate path to change.
2. Select **Make Preferred**.

Changing a Preferred Path

Select **Change a Preferred Path** when Load Distribution is disabled. When there are multiple paths available to a storageset on the same controller, and a new preferred path is needed for normal I/O operations, use the following steps:

1. Click the alternate path to change to Preferred.
2. Right click **Change Preferred**.

Making a Path Offline

Select **Make a Path Offline** to prevent that path from being used for any I/O operations under any circumstances. For instance, use the offline mode to replace or repair a storage interconnect component. To make a path offline:

1. Click the path to take offline.
2. Right click to select **Make Offline**. If the path was an alternate, its mode will change to Alt-Offline. If the path was preferred, its mode will change to Pre-Offline.

Making a Path Online

Select **Make a Path Online** to return a path that is currently in the alt-offline or pre-offline mode to its original mode. To make a path online:

1. Click a path in the Alt-Offline or Alt-Online mode.
2. Right click to select **Make Online**. If the path was Alt-Online, its mode will change to Alternate. If the path was Pre-Offline, its path will change to Preferred.

Verifying a Path

Select **Verify a Path** to have SPM determine the current state of a path. To verify a path:

1. Click the path to view.
2. Right click **Verify Path**. SPM will generate a pop-up message when the verification completes to indicate the result of the operation. No state change occurs as a result of this operation.

Repairing a Path

Select **Repair a Path** to have SPM restore access to a failed path after the problem has been corrected. To repair a path:

1. Click a path that is in a failed state.
2. Right click **Repair Path**. If the repair action completes successfully, the path state changes to Available if its mode is Alternate. The path state changes to Active if its mode is Preferred.

Fault Tolerance

The fault tolerance of the TaskSmart N2400 appliance has been greatly improved by the TaskSmart N-Series Cluster. Each server head contains two fiber HBAs for dual redundancy. Each HBA connects to one of the two SAN switches via a fiber cable. In turn, the SAN switches connect to a port on each of the HSG80 Array Controllers. The TaskSmart N-Series Cluster provides multiple paths to the data through these dual redundant fiber connections.

To maintain the best possible fault tolerance, Compaq recommends the following practices. To carve the disks in the best fault-tolerant configuration, consult the *Compaq TaskSmart N-Series Planning Guide*. When using NIC teaming, Compaq recommends teaming either four ports or two ports on the four-port NIC. Each port of a NIC team must connect to the same network segment. When creating pools, it is best to not put all of the disk space into a single pool.

When creating the cluster groups, two levels of granularity can be achieved. Each level has its advantages and disadvantages in administrating the cluster.

- One way to group resources is to create two groups, one for each node in the cluster. Each group has its own network name and IP address. The administrator decides on which node to place each pool resource. This configuration provides a very coarse level of granularity. Resources have to remain on the same node as the rest of the group. Only two IP addresses and network names are required. This configuration creates less overhead for resource and network administration. A possible disadvantage of this approach is that the resource groups can potentially grow large when many file shares are created.
- A second way to group resources is to create a group for each pool that is created. For each group, a network name and IP address must be created. Each group contains a single pool, network name, IP address, and any file share resources. The administrator chooses which node owns each pool. However, the more groups that are created, the more network names and IP addresses the administrator has to manage. Remembering where a certain resource is located requires more administrative overhead and forces clients to map more drive letters.

Hardware Failure

This section discusses cluster behavior during various potential failure scenarios. For information on client behaviors during those failures, refer to the “Expected Client Behaviors During Failovers” section later in this document

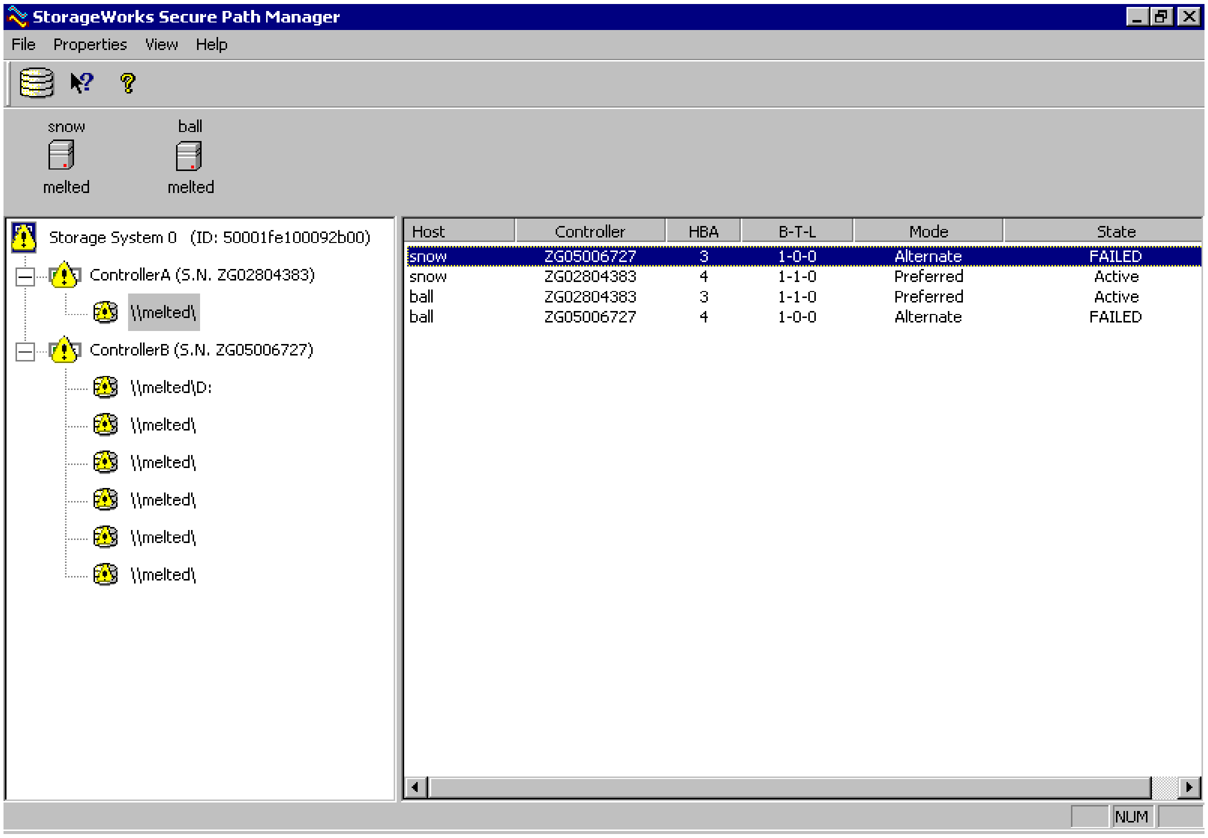
SAN Switch Failure

When a SAN switch fails, there is minimal impact on the existing cluster environment. This section covers the expected cluster behavior in the event of a SAN switch failure and how to properly replace the switch while the cluster is still up and running.

Expected Behaviors During a SAN Switch Failure

When a SAN switch fails, two areas within the cluster detect the failure and respond as follows:

- SecurePath detects the failed SAN switch as a failure on two of the four paths to each unit (LUN). This process takes about 10-15 minutes depending on the number of units the cluster controls. Figure 34 shows SecurePath Manager after it has detected the failed paths.

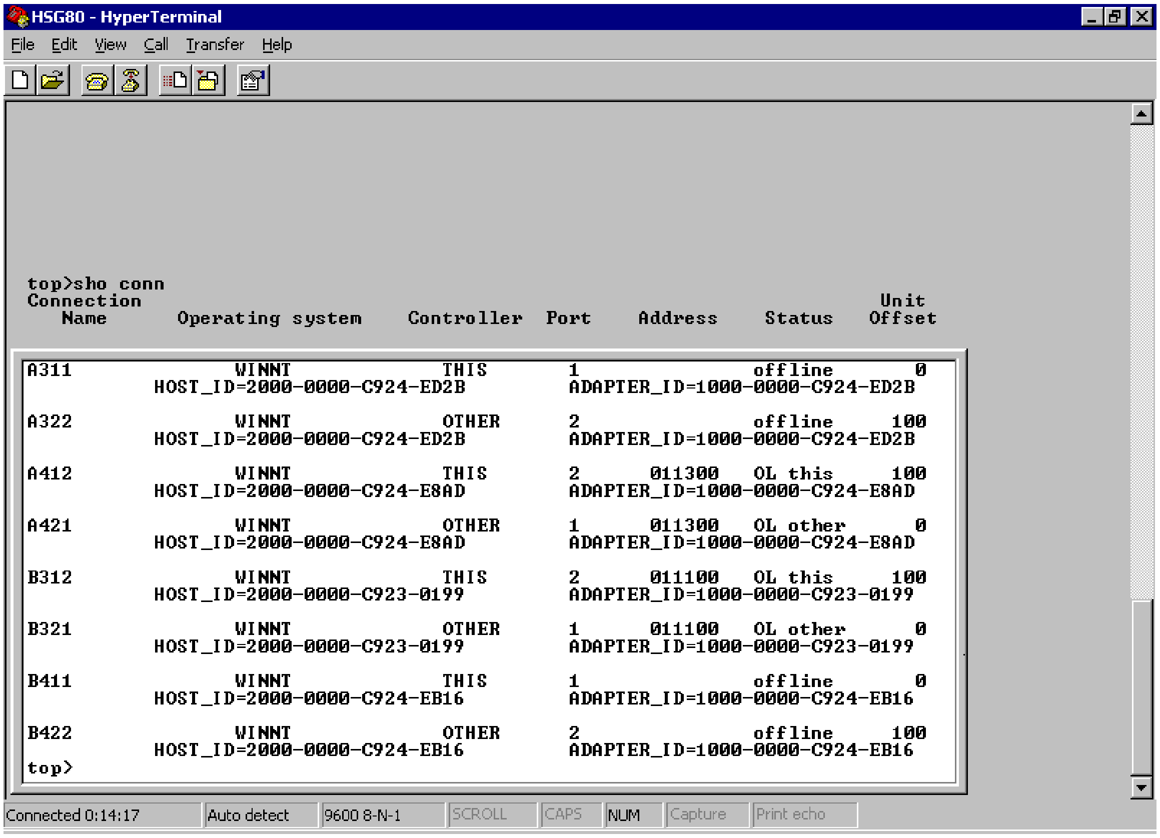


Storage System 0 (ID: 50001fe100092b00)
- ControllerA (S.N. ZG02804383)
  - \\melted\
- ControllerB (S.N. ZG05006727)
  - \\melted\D:
  - \\melted\
  - \\melted\
  - \\melted\
  - \\melted\
  - \\melted\

Host	Controller	HBA	B-T-L	Mode	State
snow	ZG05006727	3	1-0-0	Alternate	FAILED
snow	ZG02804383	4	1-1-0	Preferred	Active
ball	ZG02804383	3	1-1-0	Preferred	Active
ball	ZG05006727	4	1-0-0	Alternate	FAILED

Figure 34: Display of a failed switch in SecurePath Manager

- The HSG80 Array Controller will show four of the eight connections as being offline. In a HyperTerminal connection, enter the command `show connections` to view their status.



```
top>sho conn
Connection                                                               Unit
   Name       Operating system    Controller  Port    Address    Status   Offset

A311              WINNT               THIS        1                 offline       0
           HOST_ID=2000-0000-C924-ED2B            ADAPTER_ID=1000-0000-C924-ED2B

A322              WINNT               OTHER       2                 offline     100
           HOST_ID=2000-0000-C924-ED2B            ADAPTER_ID=1000-0000-C924-ED2B

A412              WINNT               THIS        2       011300    OL this     100
           HOST_ID=2000-0000-C924-E8AD            ADAPTER_ID=1000-0000-C924-E8AD

A421              WINNT               OTHER       1       011300    OL other      0
           HOST_ID=2000-0000-C924-E8AD            ADAPTER_ID=1000-0000-C924-E8AD

B312              WINNT               THIS        2       011100    OL this     100
           HOST_ID=2000-0000-C923-0199            ADAPTER_ID=1000-0000-C923-0199

B321              WINNT               OTHER       1       011100    OL other      0
           HOST_ID=2000-0000-C923-0199            ADAPTER_ID=1000-0000-C923-0199

B411              WINNT               THIS        1                 offline       0
           HOST_ID=2000-0000-C924-EB16            ADAPTER_ID=1000-0000-C924-EB16

B422              WINNT               OTHER       2                 offline     100
           HOST_ID=2000-0000-C924-EB16            ADAPTER_ID=1000-0000-C924-EB16
top>
```

Figure 35: Display of a failed switch on the HSG80 Array Controller

Replacement Procedures for a Failed SAN Switch

To replace a failed SAN switch, use the following steps:

1. Shut down the failed switch and remove the fiber and GBIC connections.
2. Install the new SAN switch, and then insert the GBIC and fiber connections back into the switch. Do not turn the switch on.
3. Shut down Node B, or the node that does *not* own the Quorum Disks.
4. Turn on the new SAN switch. This process takes approximately five minutes.
5. When the SAN switch is back up, restart the node that was powered down.
6. Confirm that the connections are back online on the HSG80 Array Controller. In a HyperTerminal connection, enter the command `show connections`. All connections should report either `OL this` or `OL other`.

HBA Failure

The process for replacing a failed HBA is more involved than for a failed switch. When an HBA fails and is replaced, a new connection will be generated on the HSG80 Array Controller. Connections are established through the HBA on a controller, and when an HBA goes down, the connections associated with it are taken offline. When a new HBA is installed in place of the failed HBA, two new connections are established. This section covers the procedures for replacing a failed HBA.

1. Shut down the node with the HBA failure.
2. Remove the failed HBA and replace it.
3. Restart the node.
4. Open a connection to the HSG80 Array Controller through a HyperTerminal connection.
5. Enter the following command:

   `show connections`

   Two connections should be displayed as being offline and two new connections will be present. If only one new connection is present, a controller restart is necessary. Use the `Restart This` command to restart the controller.
6. Delete the old connections that are now offline. Enter the following command:

   `delete` *`xxxx`* (*`xxxx`* represents the connection name.)
7. Rename the new connections to match the naming convention that was used during the initial installation. Further information on naming conventions for connections can be found in the *Compaq TaskSmart N-Series Cluster Installation Guide*. Enter the following command:

   `rename` *`xxxx`* (*`xxxx`* represents the new connection name.)
8. A connection with an access path explicitly enabled on a unit cannot be deleted.

   The access path is enabled explicitly through the `enable_access_path` qualifier of the `add` unit or `set` unit commands. If Access Path is generically enabled for all connections through the `enable_access_path=all` command, then any connection can be deleted. A connection with explicit access path must have the access path disabled before the connection can be deleted. Use the `disable_access_path` qualifier of the `set` unit command to disable the access path.
9. Enter the following command to display the path access settings:

   `show d`*`x`* (*`x`* represents the unit number.)

   If the screen displays `access=all`, then path access has been generically enabled. However, if the screen displays `access=xxxxx,xxxxx,xxxxx`. (`xxxxx` represents the connection names.) , then it will be necessary to delete the old connection and add a new connection to this unit. To confirm which connection will need to be added to the unit, refer to the naming convention for each connection. All port 2 connections are granted access to units in the 100 unit offset range, or all units named d100 through d107 are granted access from all port 2 connections.
10. Open SecurePath Manager and confirm that there are four paths for each unit.

Node Failure

One of the main benefits of running a clustered server is the ability to lose one node and still be able to service requests for resources normally served from that node. This is accomplished through failover. When the node fails, all of its resources fail over to the other node in the cluster. This section covers the recovery procedures for single-node and dual-node failures.

Recovery from Single Node Failure

If a single node fails within your cluster, several tools and procedures can be used for a quick and clean recovery. The following procedures describe the necessary steps to perform a complete recovery and rejoin the cluster.

From the node that is still active:

1. Start Cluster Administrator, and then right click the node that has failed. Then, select **Evict Node**.
2. Perform a QuickRestore on the failed node by selecting **Full**.
3. Configure the network settings.
4. Install cluster services by doing the following procedures:
   - Select **Start**, **Settings**, and then **Control Panel.**
   - Select **Add/Remove Programs**, and then select **Add/Remove Windows Components**.
   - Click **Configure** for Cluster Service.
   - Select **Join As Second Node In A Cluster** to install cluster services.
5. Install SANworks Virtual Replicator by doing the following:
   - Insert the SANworks Virtual Replicator CD into the CD ROM drive.
   - Choose **Complete Setup**.
   - Upon completion, restart the node.
6. When the node comes online, open Cluster Administrator and confirm that the node successfully joined the cluster. Move a group or a cluster group over to the other node to confirm functionality.
7. Select **Start**, **Programs**, and then **Compaq SANworks Virtual Replicator** (SWVR) to launch Snapshot Manager.

   The SWVR resources are displayed. If they are not:

   - From the **Action** menu, select **Connect to Another Computer**.
   - Specify the name of the cluster. This process refreshes all of the SANworks Virtual Replicator information in Snapshot Manager.

Recovery From Dual-Node Failure

If both nodes of a cluster fail, then it is critical that a current backup of both the system state and the Quorum Disk is available for the restore process. This section covers the recovery procedures using NTBackup and Legato NetWorker 6.0. If a backup software other than NTBackup or Legato NetWorker is used, consult the documentation for the procedures for backing up and restoring a Quorum Disk in a Microsoft Cluster Services environment.

Recovering from Dual-Node Failure Using NTBackup

If both nodes of the cluster fail, the following procedures can recover the nodes along with the Quorum Disk. These procedures restore the cluster to its working state before the disaster. It is assumed that a current backup of the System State and the Quorum Disk is available to perform this recovery, and that no configuration changes have occurred since the backup. It is critical that the data backup, and the System State, and the Quorum Disk backup have the same time stamp.

CAUTION: If changes have been made to SWVR resources that are not in the current System State and Quorum Disk backups, then data loss can occur. Ensure that a backup of the System State and the Quorum disks occur frequently.

Note: It is imperative that a current backup of the System State and the Quorum Disk is kept in an offsite location to be used in the event of a disaster.

1. Use the TaskSmart N-Series Cluster QuickRestore CD to perform a QuickRestore on both nodes. Select **Full** from the menu.
2. When the QuickRestore process is complete, reconfigure the cluster interconnect private network settings and the public network settings on Node A. It is not necessary for the cluster to rejoin the domain on Node A. It is reconfigured by the System State restore. If NIC teaming is installed and configured, then a restart will be necessary.
3. Install the Clusrest tool. (This tool is required to restore the Quorum Disk.)

   This tool is available on the following website:

   www.microsoft.com/WINDOWS2000/library/resources/reskit/tools/existing/clusrest-o.asp

   To download and install the Clusrest tool, perform the following steps:

   – Click the **Download This Tool** link in the previous step.

   – In the **File Download** dialog box, select **Save This Program To Disk**.

   – Select a location on your computer to save the file, and then click **Save**.

   – In Windows Explorer, go to the location where the downloaded file is saved, double click the file to start the installation process, and then follow the onscreen instructions.

   – The downloaded file is a self-extracting executable (.exe) file. Running the file installs the tool and documentation on the system.
4. Open NTBackup. The following prompt is displayed:

   – **Import Media Present**

   – Select **Allocate All Compatible Import Media To Backup**.
5. On Node A, restore both the System State and the Quorum Disk. A restart is required upon completion.
6. When the restart is complete, log onto the domain, and then run the Clusrest tool:

   – Open a command prompt.

   – Change to the directory where the Clusrest tool is installed. The default directory is c:\program files\resource kit.

   – Run the Clusrest tool by entering `clusrest`.

- Upon completion, close the command prompt, and then restart Node A.

7. Reinstall Compaq SANworks Virtual Replicator:
   - Insert the SANworks Virtual Replicator CD into the CD ROM drive.
   - Select **Repair** from the **SWVR Installation** menu.
8. Open **Snapshot Manager** of SANworks Virtual Replicator to ensure that all pools, virtual disks, and snapshots are present.
9. Open Cluster Administrator.
10. Evict Node B by right clicking Node B and selecting **Evict**.
11. Power up Node B. Install and configure all network configurations, including domain configuration.
12. Install cluster services:
    - Select **Start, Settings,** and then **Control Panel**.
    - Select **Add/Remove Programs**, and then select **Add/Remove Windows Components.**
    - Click **Configure** for Cluster Service.
    - Select **Join As Second Node In A Cluster** to install cluster services.
13. Install SANworks Virtual Replicator:
    - Insert the SANworks Virtual Replicator CD into the CD ROM drive.
    - Select **Complete Setup.**
14. Upon completion, restart the node. After restarting, open SANworks Virtual Replicator.

    If the resources are not displayed, complete the following steps:
    - From the **Action** menu, select **Connect to Another Computer**.
    - Specify the name of the cluster. This refreshes all of the SANworks Virtual Replicator information in Snapshot Manager.
15. Configure any appropriate changes in Cluster Administrator. Because Node B was evicted and then rejoined the cluster, any configurations specific to Node B must be re-established.

    Configurations that may need to be altered include:
    - Preferred nodes for group resources

Note: Ensure that the pool resource within each group has both nodes as a possible owner.

- Failover or failback settings for each resource

Recovering From Dual Node Failure Using Legato NetWorker 6.0

This section provides general guidelines for performing a cluster recovery in the case of a failure of both cluster nodes. In this scenario, the operating system is unusable on each node. Therefore, this recovery procedure includes a QuickRestore of each node, as well as NetWorker recovery of the cluster database from a previous backup.

IMPORTANT: Because cluster configurations may vary, it is not possible to provide cluster disaster recovery procedures for every situation. Depending on the particular cluster configuration and the nature of the failure, it might be necessary to vary some of the procedures described in this section.

To perform a complete cluster recovery in a situation in which both cluster nodes, Node A and Node B, have failed, do the following procedures:

1. Perform a QuickRestore on each node by selecting **Full**. During the QuickRestore setup, verify that the domain controller is available and that each potential node is able to join.
2. Delete the MSCS folder on the Quorum Disk.
3. On the Quorum Disk, run `chkdsk`.
4. Shut down Node B, and start the MSCS installation on Node A. From the **Control Panel,** select **Add/Remove Programs,** and then **Add/Remove Windows Components**. Select **Configure** for Cluster Services. The MSCS Cluster wizard is displayed and provides guidance through the setup process.
5. During the setup process using Cluster wizard, enter the same configuration information that was used prior to the failure of the cluster nodes (including user account, IP addresses, and cluster name.)
6. Restart Node A.
7. Install MSCS on Node B, joining A.
8. Restart Node B.
9. On Node B, run the `net stop clussvc` command from a command prompt.
10. Install the NetWorker client software on Node A.
11. From NetWorker User on Node A, select the **System Files** and **System State** save sets to recover to the desired point in time prior to the cluster failure. Click **Start** to begin the recovery process.
12. After the recovery on Node A is complete, restart Node A.
13. Reinstall Compaq SANworks Virtual Replicator:
    - Insert the SANworks Virtual Replicator CD into the CD ROM drive.
    - Select **Repair** from the **SWVR Installation** menu.
14. On Node A, run Cluster Administrator to confirm that the states of the cluster resources were restored to the desired point in time.
15. Start the cluster service on Node B. To do so, at Node B, run the `net start clussvc` command from a command prompt. (This command can also be executed from the **Computer Management, Services** menu.)
16. From Cluster Administrator on Node B, verify that the cluster group can be moved between the nodes by right clicking the group, and then selecting **Move Group**.
17. Install SANworks Virtual Replicator:
    - Insert the SANworks Virtual Replicator CD into the CD ROM drive.
    - Select **Complete Setup**.

18. Reinstall the NetWorker client software on Node B.

Controller Failure

The TaskSmart N-Series Cluster is configured with dual controllers for the StorageWorks enclosure model 2200. If one of the two controllers were to fail, normal cluster operations would continue. When a controller is taken offline or fails in a Model 2200 enclosure, the resources that were stored on that controller move over to the other controller. During this time, one controller must manage all resources, and the resources are not balanced over two controllers. It is necessary to replace the failed controller as soon as possible. This section covers the procedures for replacing a controller and its respective cache module in the event of a failure.

Making a File Copy of System Settings

Documentation is a crucial part of any recovery solution. A copy of all system settings should be stored in a location different from the servers it represents. In the event that both controllers fail in a disaster, thorough documentation is the key to quick recovery. This section demonstrates how to extract the configuration of the HSG80 Array Controller into a text document and what information should be included.

Use the following steps to create this file:

In HyperTerminal:

1. Select **Transfer**, and then select **Capture Text**.
2. Select the location where the file is to be saved. It is recommended that this file is not saved on any of the disks managed by the controllers. Save the file or its printout in a remote location.
3. The following are the recommended commands to enter to extract the output to the specified file. When these commands are executed, the output will be saved in the specified text file.

Show Connections	Displays the connections to the controllers and their unit offset
Show Devices	Displays all disks attached to the HGS80 Array Controllers
Show Mirrorsets	Displays all mirrorsets configured on the controllers along with the associated disks
Show Other	Displays the current configuration for the secondary controller
Show RAIDsets	Displays all RAIDsets configured on the controllers along with the associated disks.
Show Sparesets	Displays all sparesets configured on the controllers along with the associated disks.
Show This	Displays the current configuration of the controller that has an established connection
Show Units	Displays the LUN on the controllers and what disksets are associated with them

When both controllers have been replaced, it is necessary to complete the configuration process that is explained in the installation guide. It is critical that up to date and complete documentation be available to provide quick and efficient recovery.

Single Controller Failure

In the event that a single controller fails, the recovery process is simple. Because the Model 2200 enclosure is configured with redundant controllers, the failure of one controller will not result in data loss. During a controller failure, all units are moved over to the active controller, and access to the cluster is not altered. Keep in mind that units are not balanced over the two controllers anymore and there may be a small decrease in performance until the failed controller is replaced. This replacement process includes two steps: removing the old controller, and installing the new controller.

IMPORTANT: New controller hardware must be compatible with the remaining controller hardware before beginning replacement procedures. See the product-specific release notes that accompanied the software release for information regarding hardware compatibility. The software versions and patch levels must be the same on both controllers. The new cache module must contain the same memory configuration as the module being replaced.

Replacing a Controller and Cache Module in a Dual-Redundant Controller Configuration

Use the following steps in “Removing a Controller and Cache Module in a Dual-Redundant Controller Configuration” and “Installing a Controller and its Cache Module in a Dual-Redundant Controller Configuration” to replace a controller and its cache module. Both cache modules must contain the same cache memory configuration.

Removing a Controller and Cache Module in a Dual-Redundant Controller Configuration

Use the following steps to remove a controller and its cache module:

1. Connect a computer or terminal to the maintenance port of the operational controller. The controller connected to the computer or terminal becomes “this controller;” the controller being removed becomes the “other controller.”
2. Disable failover and take the controllers out of the dual-redundant configuration with the following command:

   ```
   set nofailover
   ```

3. Remove the program card electrostatic discharge (ESD) cover and the program card from the “other controller.” Save them in a static-free place to use with the replacement controller.
4. Start the field replacement utility (frutil) with the following command:

   ```
   run frutil
   ```

   Follow the onscreen instructions to remove the elements.

CAUTION: The device ports must be idle before the controller is removed. This is indicated by an “All device ports quiesced” message. Failure to allow the ports to become inactive might result in data loss. This process might take several minutes.

Note: A countdown timer allows a total of four minutes to remove both the controller and cache module. After four minutes, “this controller” exits the frutil program and resumes operations. If this time limit is exceeded, return to step 4 and proceed.

5. When instructed to remove the controller and cache module, use the following steps:
   - Disconnect all fiber cables from the “other controller.” For cables without extender clips, use thin needle-nose pliers to disconnect each cable.
   - Disengage both retaining levers on the controller, and remove the “other controller.” Then, place the controller in an antistatic bag or on a grounded antistatic mat.
   - Disengage both retaining levers on the controller, and remove the “other controller” cache module, and then place the cache module on a grounded antistatic mat or an antistatic bag.
6. If the replacement controller and cache module are available, remove the DIMMs from the “other controller” cache module for installation in the replacement cache module. Do this using the following steps:
   - Press the DIMM retaining clips down at both ends of the DIMM being removed. To make pressing down on the DIMM retaining clips easier, consider using the eraser end of a pencil or a small screwdriver.
   - Gently remove the DIMM from the DIMM slot, and then place it in an antistatic bag or on a grounded antistatic mat.
   - Repeat these procedures in step 6 for each DIMM.
7. If the replacement cache module is available, insert each DIMM into the new module. Do this by using the following steps:
   - Insert each DIMM straight into the appropriate slot of the replacement cache module, ensuring that the notches in the DIMM align with the tabs in the slot.
   - Press the DIMM gently into the slot until seated at both ends.
   - Engage the two retaining clips for the DIMM.
   - Make sure that both ends of the DIMM are firmly seated in the slot and both retaining levers engage the DIMM.
   - Repeat these procedures in step 7 for each DIMM.
8. The frutil utility program will prompt: “Is a replacement controller and cache module available now?”

   If the replacement is not available at this time, enter **n** to indicate no, and exit out of frutil. Disconnect the computer or terminal from the controller maintenance port.

   If the replacement controller and cache module is available, enter **y** to indicate yes. Follow the onscreen directions. When finished, exit out of frutil. The other controller will start. Disconnect the computer or terminal from the controller maintenance port.

Note: When fully seated, the replacement controller restarts automatically. The reset LED turns on.

Installing a Controller and its Cache Module in a Dual-Redundant Controller Configuration

Use the following steps to install a controller and its cache module.

1. Install the DIMMs that were removed from the old cache module onto the replacement cache module. To do this, do the following:
   - Insert each DIMM straight into the appropriate slot of the replacement cache module, ensuring that the notches in the DIMM align with the tabs in the slot.
   - Press the DIMM gently into the slot until seated at both ends.
   - Engage the two retaining clips for the DIMM.
   - Make sure that both ends of the DIMM are firmly seated in the slot and both retaining levers engage the DIMM.
   - Repeat these procedures of step 1 for each DIMM.
2. Connect a computer or terminal to the maintenance port of the operational controller. The controller connected to the computer or terminal becomes “this controller”; the controller being installed becomes the “other controller.”
3. Start the field replacement utility (frutil) with the following command:

   ```
   run frutil
   ```

   Follow the onscreen instructions to install the elements.

Note: A countdown timer allows a total of four minutes to install both the cache module and controller. After four minutes, “this controller” will exit frutil and resume operation. If this time limit is exceeded, return to step 3 and proceed.

CAUTION: Carefully align the cache module and controller in the appropriate guide rails. Misalignment might damage the backplane.

Note: When the replacement cache module and controller are fully seated, the replacement controller restarts automatically. The reset LED turns on.

4. Press the **Enter** key to continue.

   The “other controller” restarts and frutil exits.

Note: When a controller restarts, a visual indication is the temporary cycling of the port LEDs and a flashing reset button.

5. If the “other controller” did not restart, follow these steps:
   - Press and hold the “other controller” reset button.
   - Reseat the “other controller” program card.
   - Release the reset button.

Note: In mirrored mode, frutil initializes the mirrored portion of the new cache module, checks for old data on the cache module, and then restarts all device ports. After the device ports restart, frutil tests the cache module and the ECB. After the test completes, the device ports are idled and a mirror copy of the cache module data is created on the newly installed cache module.

6. Replace the program card ESD cover.

IMPORTANT: If the controller being installed was previously used in another subsystem, purging the controller of the old configuration is required. (See the `config reset` command in the controller CLI reference guide).

7. Enable failover and re-establish the dual-redundant controller configuration with the following command:

   ```
   set failover copy=this_controller
   ```

   This command copies the subsystem configuration from “this controller” to the new controller.
8. Verify the failover configuration with the following command:

   ```
   show this_controller full
   ```
9. Configure the controller using the *Compaq TaskSmart N-Series Cluster Installation Guide* as a reference.
10. Connect all host bus cables to the new controller.
11. Disconnect the computer or terminal from the controller maintenance port.

Replacing a Controller in a Dual-Redundant Controller Configuration

Use the following steps in “Removing a Controller in a Dual-Redundant Controller Configuration” and “Installing a Controller in a Dual-Redundant Controller Configuration” to replace a controller.

Removing a Controller in a Dual-Redundant Controller Configuration

Use the following steps to remove a controller:

1. Connect a computer or terminal to the maintenance port of the operational controller. The controller connected to the computer or terminal becomes “this controller”; the controller being removed becomes the “other controller.”
2. Disable failover and take the controllers out of the dual-redundant configuration with the following command:

   ```
   set nofailover
   ```
3. Remove the program card ESD cover and the program card from the “other controller.” Save them in a static-free place to use with the replacement controller.
4. Start the field replacement utility (frutil) with the following command:

   ```
   run frutil
   ```

   Follow the onscreen instructions to remove the controller.

CAUTION: The device ports must be idle before the controller is removed. This is indicated by an "All device ports quiesced" message. Failure to allow the ports to go inactive might result in data loss. This process might take several minutes.

Note: A countdown timer allows a total of four minutes to remove the controller. After four minutes, "this controller" will exit frutil and resume operation. If the time limit is exceeded, return to step 4 and proceed.

5. When instructed to remove the "other controller," use the following steps:
   - Disconnect all fiber cables from the "other controller." For cables without extender clips, use thin needle-nose pliers to disconnect each cable.
   - Disengage both retaining levers on the controller and remove the "other controller." Then, place the controller in an antistatic bag or on a grounded antistatic mat.
6. The frutil utility program will prompt: "Is a replacement controller available now?"

   If the replacement is not available at this time, enter **n** to indicate no, and exit out of frutil. Disconnect the computer or terminal from the controller maintenance port.

   If the replacement controller is available, enter **y** to indicate yes. Follow the onscreen directions. When finished, exit out of frutil. The other controller will start. Disconnect the computer or terminal from the controller maintenance port.

Note: When fully seated, the replacement controller restarts automatically. The reset LED turns on.

Installing a Controller in a Dual-Redundant Controller Configuration

Use the following steps to install a controller:

1. Connect a computer or terminal to the maintenance port of the operational controller. The controller connected to the computer or terminal becomes "this controller;" the controller being installed becomes the "other controller."
2. Start the field replacement utility (frutil) with the following command:

   ```
   run frutil
   ```

   Follow the onscreen instructions to install the controller.

Note: A countdown timer allows a total of four minutes to install the controller. After four minutes, "this controller" will exit frutil and resume operation. If this time limit is exceeded, return to step 2 and proceed.

CAUTION: Carefully align the controller in the appropriate guide rails. Misalignment might damage the backplane.

Note: When fully seated, the replacement controller restarts automatically. The reset LED turns on.

3. Press the **Enter** key to continue. The "other controller" restarts and frutil exits.

Note: When a controller restarts, a visual indication is the temporary cycling of the port LEDs and a flashing reset button.

4. If the "other controller" did not restart, follow these steps:
   – Press and hold the "other controller" reset button.
   – Reseat the "other controller" program card.
   – Release the reset button.
5. Replace the program card ESD cover.

IMPORTANT: If the controller being installed was previously used in another subsystem, purging the controller of the old configuration is required. (See the `config reset` command in the controller CLI reference guide.)

6. Enable failover and re-establish the dual-redundant controller configuration with the following command:

   ```
   set failover copy=this_controller
   ```

   This command copies the subsystem configuration from "this controller" to the new controller.
7. Verify the failover configuration with the following command:

   ```
   show this_controller full
   ```

8. Configure the controller using the *Compaq TaskSmart N-Series Cluster Installation Guide* as a reference.
9. Connect all host bus cables to the new controller.
10. Disconnect the computer or terminal from the controller maintenance port.

Dual Controller Failure

If both nodes are unaffected by the dual controller failure, shut them down and use a HyperTerminal connection from another computer. It is best if the nodes are not up when configuration of the controller takes place.

It is extremely important that the controllers are configured exactly the way they were prior to disaster. The previous section titled "Making a File Copy Of System Settings" details the process for documenting the controller settings.

It is also necessary to replace all damaged components. When both controllers have been replaced, it is necessary to complete the configuration process that is explained in the *Compaq TaskSmart N-Series Cluster Installation Guide*. It is critical that up to date and complete information is present to provide quick and efficient recovery time.

This section assumes the following conditions:

- Both nodes are in working condition and the cluster state is unharmed.
- The storage cabinets and the data residing on them are still intact.
- Current documentation of the controller configuration is available.
- A current backup of all data is available.

Use the following steps for recovery from dual controller failure:

1. Shut down both nodes.
2. Replace both controllers and their cache modules. (This procedure is documented in the previous section titled “Single Controller Failure.”)

CAUTION: Do not initialize the disk sets as they are created. This will result in total data loss.

3. Reconfigure the controllers to their original configuration according to the documentation that was extracted from the controllers before failure. It is very important that all units, disksets, and any specific configurations pertaining to them are restored exactly to their previous state.
4. Disable access to all units, with the exception of the Quorum Disk unit. Use the following command:

   ```
   set d1 disable_access=all
   ```

   This command must be entered for all units, with the exception of the Quorum Disk unit. For more information on enabling and disabling access to units, see the *Compaq TaskSmart N-Series Installation Guide*.
5. Restart Node A. Do not restart Node B at this time.

   Confirm that Cluster Services has started. Do this by opening Cluster Administrator and confirm that the cluster resources are online. The pool and share resources will be offline at this point. This is expected behavior.
6. Enable access to each unit, one at a time. After access has been enabled to the unit, rescan for new devices in device manager. Follow this procedure until access has been enabled for all units. To enable access to a unit, use the following command:

   ```
   set d1 enable=!newcon00,!newcon03,!newcon04,!newcon05
   ```

   For more information on enabling and disabling access to units, see the *Compaq TaskSmart N-Series Cluster Installation Guide*.
7. Restart Node A.
8. Confirm that all SANworks Virtual Replicator resources are online. Confirm this in both Cluster Administrator and Snapshot Manager.
9. Restart Node B.
10. Confirm that Node B successfully joined the cluster by doing the following in Cluster Administrator:
    - Move a cluster group to Node B.
    - Move a pool group to Node B.

Ethernet Switch Failure

There are three uses of an Ethernet Switch. Failure may occur in any of the following scenarios:

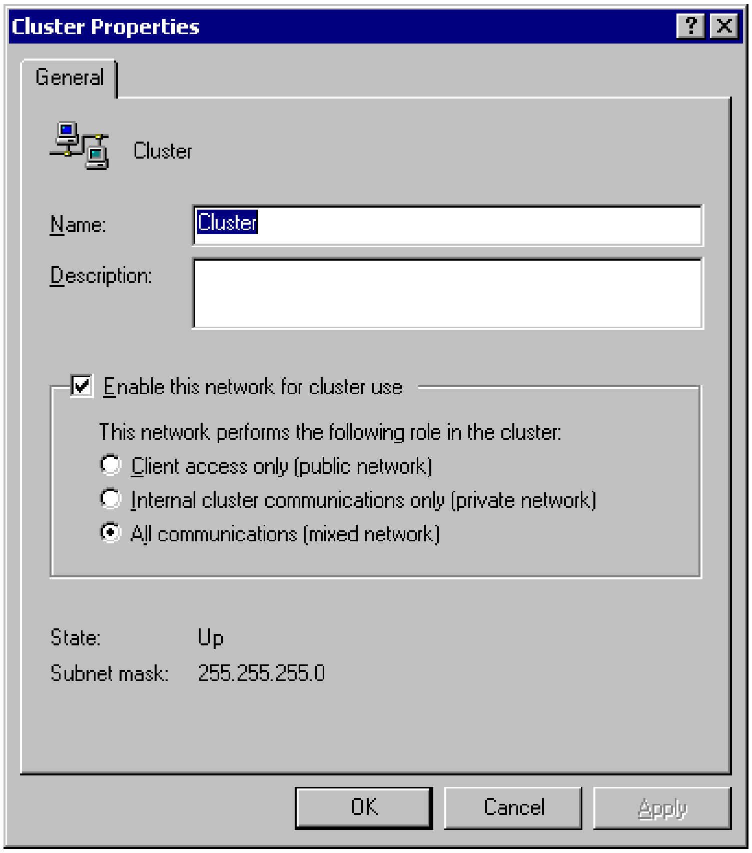
- The switch was being used to connect the public infrastructure to the cluster.
- The switch was being used to interconnect the cluster nodes and the public infrastructure.
- The switch was being used to interconnect the cluster nodes.

This section discusses the methods of recovery from failure for all three scenarios.

If the failed switch was being used to connect the public infrastructure, then all communications to the cluster will be lost. Likewise, if IP address resources have been configured for the virtual servers on that public infrastructure, then those resources will fail. The best repair procedure for this scenario is to replace the failed switch as soon as possible. After the failed switch has been replaced, it is necessary to open Cluster Administrator and confirm that the IP address resources have come back online. If the IP address resources have not come back online, right click the resources, and select **Bring Online**. If the first two procedures do not work, you must restart each node in the cluster.

If the failed switch was being used to interconnect the cluster nodes and a public infrastructure, access to the cluster will be lost and the second node in the cluster will fail. The IP address resources for the virtual servers will also fail because the network they are associated with is no longer functioning. In this scenario, it is imperative that the failed switch be replaced immediately. When the failed switch has been replaced, open Cluster Administrator and confirm that the failed IP address resources have come back online. The node that failed should come back online. If not, either the cluster service or the node must be restarted.

If the failed switch was being used for the cluster interconnect, normal operations may continue. This is the case if the public infrastructure connection was configured for mixed communications during the initial setup of Cluster Administrator. To confirm this, open Cluster Administrator, select the cluster name, and then select **File**, **Properties**, **Network Priority**. Highlight the connection name and select **Properties**. Figure 36 displays the **Cluster Properties** screen for the connection.



Cluster

Name: Cluster

Description:

☑ Enable this network for cluster use

This network performs the following role in the cluster:

○ Client access only (public network)

○ Internal cluster communications only (private network)

● All communications (mixed network)

State: Up

Subnet mask: 255.255.255.0

Figure 36: Public connection cluster properties

The public infrastructure connection must be set to the **All communications (mixed network)** setting. This setting provides fault tolerance for the interconnect. If this setting is the original configuration, then simply replace the failed switch and resume normal operations. If the public network connection was not originally configured this way, then the interconnect has no fault-tolerant connection and will fail. In this event, one node of the cluster will fail. It will be necessary to replace the failed switch, and then restart the failed node to restore the cluster.

NIC Failure

A failure of the network interface card impacts the ability of the cluster to balance its resources across both nodes. When a cluster group is configured, it is necessary to create a virtual server. Within that virtual server, the IP address resource is tied to the public infrastructure that is associated with the NIC card. When the NIC card goes offline, the IP address resource fails over to the other node that still has public network communications. When the IP address resource fails over, all members of the group associated with it also fail over.

To recover from a NIC failure, ensure that the cluster group is not owned by the node with the failed NIC card. If the cluster group is owned by the node with the failed NIC card, move it over to the other cluster node. Shut down the node with the failed NIC card, and then replace the card. Bring the node back up, and ensure that the node has rejoined the cluster.

StorageWorks Enclosure Model 2200 Failure

When the Model 2200 enclosure fails, the symptoms will be the same as losing both controllers simultaneously. When the proper recovery procedures are followed, access to all resources will be lost, but the data is still maintained. Shut down both nodes of the cluster while the repair process takes place. The following steps describe the recovery procedures for when a Model 2200 enclosure fails.

1. Shut down both nodes of the cluster.
2. Replace all damaged components of the Model 2200 enclosure. When removing and replacing the HSG80 Array Controllers, keep the corresponding cache modules with their respective controllers.

IMPORTANT: HSG80 Array Controllers and their corresponding cache modules form a pair and must be kept together.

CAUTION: Do not initialize the disksets as they are created. This will result in total data loss.

3. Reconfigure the controllers to their original configuration. It is critical that all units, disk sets, and any specific configurations pertaining to them are restored exactly to their previous state.
4. Disable access to all units, with the exception of the Quorum Disk unit. Use the following command:

   ```
   set d1 disable_access=all
   ```

   This command must be entered for all units, with the exception of the Quorum Disk unit. For more information on enabling and disabling access to units, see the *Compaq TaskSmart N-Series Cluster Installation Guide.*

5. Restart Node A. Do not restart Node B at this time.

   Confirm that Cluster Service has started. Open Cluster Administrator and confirm that the cluster resources are online. The pool and share resources will be offline at this point. This is expected behavior.

6. Enable access to each unit, one at a time. After access to the unit has been enabled, rescan for new devices in device manager. Follow this procedure until access has been enabled for all units. To enable access to a unit, use the following command:

   ```
   set d1 enable=!newcon00,!newcon03,!newcon04,!newcon05
   ```

   More information on enabling and disabling access to units is found in the “Port Security” section of this guide.

7. Restart Node A.

8. Confirm that all SANworks Virtual Replicator resources are online. Confirm this in both Cluster Administrator and Snapshot Manager.

9. Restart Node B. Confirm that Node B successfully joined the cluster by moving the cluster group to Node B. Use Cluster Administrator to make these moves.

Expected Client Behavior During Failovers

This section describes client behavior during a failover in the cluster. A failure is anything that causes a group and resource failover from one node to the other. Several scenarios are used to depict client behavior when connected to a virtual server file share during a failure in the cluster. These scenarios include simple connection, copying files to and from the virtual server, reading files, writing files, deleting files, and audio and video streaming.

Simple Connection

In a simple connection, a client is connected to the virtual server by mapping a drive letter to a file share. In this case, the client detects a failure in the session and continuously attempts to re-establish a connection to the file share. After a short delay, the IP address of the failed drive is made available on the other node and the connection is re-established.

Copying Files to and from the Virtual Server

Clients who copy files to or from the file share will experience an error when the connection is lost to the virtual servers. When a cluster failure occurs and one of the nodes goes offline, clients who are copying to or from the share encounter an error indicating that the share is no longer available. After a short delay, the IP address of the virtual server will fail over to the other node, and the service is resumed. Any file copied during the loss of connection could be in an inconsistent state, and must be copied again. Users who experience a cluster failure during file copy operations will have to restart the process.

Reading Files

Clients who are connected to a file share of a virtual server and are reading applications may encounter a short delay during the failover process. After the resources fail over, normal application use resumes.

Writing Files

Clients who are connected to a file share of a virtual server and are writing to an application will encounter an error indicating that the share is no longer available and that it cannot write to the application. After a short delay, the resources will fail over to the other node, and the share will be available again. Users can resume writing to the application.

Deleting Files

Clients who are connected to file share of a virtual server and are deleting files will experience an error when the connection is lost to the virtual server. When one of the cluster nodes goes down, clients who are deleting files encounter an error indicating that the share is no longer available. After a short delay, the resources will fail over to the other node and the service will be renewed. Any file that was deleted during the loss of connection may not have been completely deleted, and must be deleted again to complete the request. Users who experience a cluster failure while deleting files must restart the delete process.

Audio and Video Streaming

Clients who are connected to a file share of a virtual server and are streaming audio or video will experience an error when the connection is lost to the virtual servers. When one of the nodes fails, clients who are streaming audio or video will lose the connection and cause an error on the application running the audio or video. After a short delay, the resources fail over to the other node, and the connection to the share will be available again. Users must restart their audio or video streaming application.

Troubleshooting

This troubleshooting section describes problems unique to the TaskSmart N-Series Cluster. Additional troubleshooting information is available in the *Compaq TaskSmart N2400 Administration Guide.*

FAQ

Q: Can I integrate the TaskSmart N-Series Cluster into my existing SAN?

A: In this release, the TaskSmart N-series Cluster is deployed as a SAN island. This means that the storage of the cluster is isolated from other servers, and is available only to the two TaskSmart N-Series Cluster nodes.

Q: Can I use DRM, EVM, and other SAN features with the TaskSmart N-Series Cluster?

A: Any SAN functionality beyond basic disk-block serving is currently not supported. The version of the HSG firmware that ships with the TaskSmart N-Series Cluster is version 8.5F, which does not include any of the advanced SAN functionality, such as DRM and EVM.

Q: What file-serving protocols are supported on the TaskSmart N-Series Cluster?

A: Currently, only the CIFS file-serving protocol is supported. The NFS package, SFU, did not support active/active clustering at the time of the release of the TaskSmart N-Series Cluster.

Q: Because the HSG80 Array Controller supports storage enclosures not made by Compaq, does the TaskSmart N-Series Cluster support storage enclosures not made by Compaq?

A: No, the TaskSmart N-Series Cluster supports only the storage subsystem defined in the *Compaq TaskSmart N-Series Installation Guide.*

Storage Subsystem Troubleshooting

SCSI Device problems

Troubleshooting the SCSI portion of the TaskSmart N-Series Cluster storage subsystem is similar to troubleshooting any other SCSI-based storage. Possible points of failure include the Model 2200 enclosure, the SCSI ports on the back of the Model 2200 enclosure, the SCSI cable connections from the Model 2200 enclosure to the StorageWorks 4314 external storage enclosures, the StorageWorks 4314 external storage enclosures themselves, and of course, the drives inside the StorageWorks 4314 external storage enclosures.

The procedure is very simple technically, but can be time consuming. The procedure is to replace individual components with other components that are known to be in good working order to determine if the original component is faulty. This process is repeated along the component chain until the faulty component is identified. For example, start by replacing the SCSI cable with another SCSI cable that is known to work. One way to accomplish this is to use a SCSI cable that is connected to one of the other StorageWorks 4314 external storage enclosures which is working, provided of course that this will not take down a production environment.

For more information about troubleshooting SCSI device problems, see the *Compaq TaskSmart N2400 Administration Guide.*

Fiber Device problems

Troubleshooting the fiber portion of the TaskSmart N-Series Cluster storage subsystem is very similar to troubleshooting the SCSI portion. Some possible fiber hardware failures include the HBA card, the fiber cable, the GBIC in the SAN switch, the SAN switch itself, and the fiber port on the HSG80 Array Controller in the Model 2200 enclosure.

The troubleshooter must work through the chain, from the HBA back to the fibre switch, and if necessary, from there to the HSG80 Array Controller in the Model 2200 enclosure. Each component must be replaced with another one that is known to be in good working order, until the source of the problem is located.

If the problem is traced back to the HBA from the SAN switch, the troubleshooter should confirm that the HBA firmware is up to date. This requires enabling diskette boot, restarting the cluster node, booting with the firmware update diskette, and verifying in the update program X86DNLD.EXE that the firmware version displayed is up to date. The HBA firmware version for the current shipping version of the TaskSmart N-Series Cluster is DS3.81A1. Also, verify that the KGPSA driver version is 5.4.2a.7.

TaskSmart N-Series Cluster Troubleshooting

General Cluster Troubleshooting

This section details some of the more common TaskSmart N-Series Cluster problems, as they relate to Cluster Administrator and the cluster itself. This list of problems is not a complete list of all possible problems. Rather, it covers the most common problems an administrator may encounter.

Second Node Cannot Join or Rejoin Cluster Upon Setup

Verify in Cluster Administrator that there is no second node already defined, either as a server or as a possible owner for the resources. If there is already a second node, another node will not be able to join the cluster. Evict the existing second node, and then try to rejoin the cluster. For more details on how to evict a node, see the earlier subsection titled "Evicting a Node."

Another item to check is the cluster service account. The cluster service uses this account to log on to the domain. Verify that the cluster service account is the same on both nodes.

Second Node Cannot Rejoin Cluster after Failure or Restart

Check the cluster service to see if it is running. If it has stopped, attempt to restart the service. If it stops again, view the event log for error messages. Also, check the cluster service account, and verify that it is the same for both nodes.

Cluster Resources Will Not Fail Over to the Other Node

Check to make sure that the other node is up and running. Make sure that the other node is shown as online in Cluster Administrator. Check the cluster resource you are attempting to fail over, and make sure that the other node is listed as a possible owner. For more details about possible owners, see the previous section titled "Group Properties."

"This Node Does Not Own the Group" Error

A node must be the owner of a group to be able to manage resources, move resources between groups, and move the group to the other node. For example, an SCE Pool resource cannot be moved to a group that is not owned by the current node. Use the node that is the owner of the group to make configuration changes.

One way to simplify cluster configuration during initial setup is to move all of the groups to a single node. Then, all of the configuration steps can be done on this one node. When the configuration is complete, the groups can be moved to their designated nodes.

SWVR Troubleshooting

This section details some of the more common TaskSmart N-Series Cluster problems, as they relate to SANworks Virtual Replicator. This list of problems is not a complete list of all possible problems. Rather, it covers the most common problems an administrator may encounter.

Accidental Deletion of Pool from Cluster Administrator rather than SWVR Snapshot Manager

All snapshots, virtual disks, and pools should be deleted using the SWVR Snapshot Manager, which automatically deletes the corresponding SCE Pool resource within Cluster Administrator. If Cluster Administrator is accidentally used, there is currently no way to reverse the deletion. The component resources are still available, and can be used in another pool, but the cluster cannot use that pool name again. The administrator will have to create a new pool with a different name. The same units can be used.

Snapshot Manager Does Not Display SWVR Resources

First, verify that the cluster service has started. Make sure that it is running and that there are no cluster errors in the event log. Next, verify that the unavailable resources are not owned by another node. Finally, from the **Action** menu, select **Connect To Another Computer**, and specify the name of the cluster.

Unable to Move Pool Resources to a Newly Restored Node

Verify that SANworks Virtual Replicator has been installed on the newly restored node. Next, open Cluster Administrator and select the **Properties** tab on the pool resource. Verify that both nodes are listed as possible owners for that pool resource.

SecurePath Troubleshooting

This section details some of the more common TaskSmart N-Series Cluster problems, as they relate to SecurePath. This list of problems is not a complete list of all possible problems. Rather, it covers the most common problems an administrator will encounter.

SecurePath GUI Does Not Display a Failure, Even Though SecurePath Detected the Failure

SecurePath detected the failure and reacted appropriately. The problem is a GUI refresh problem. The administrator must stop and then restart the SecurePath service. When the service is restarted, SecurePath will verify all the data paths again. The amount of time this takes depends on the polling interval setting in SecurePath. As SecurePath verifies each path, the failures show up in the GUI.

A Path Failure Does Not Show Up in SecurePath Manager

If you did not enable **Path Verification**, SecurePath will detect failures only for paths with active I/O. This means that it is possible that one or more paths may be failed to other storagesets owned by the same controller, but not yet detected by SecurePath. If you have Path Verification enabled, SecurePath will automatically detect the failure of paths to all of the affected storagesets on the controller, and immediately perform whatever path or controller failover activity is necessary to maintain availability. To identify the source of path failover activity, first check the Storage System view for path failed icons. Then, examine the Physical Path view of the affected storageset. Check for paths that indicate failed status. SecurePath displays one or more paths to a particular storageset in the failed state depending upon the following conditions:

- Was I/O active on the affected storageset?

  SecurePath determines path failures by detecting the failure of I/O operations. If I/O was not active on a broken preferred path, the fault is not detected. The state will not be marked as failed until I/O operations occur and cannot be properly executed.

- Is Path Verification enabled?

  Path Verification periodically tests the viability of all paths, and automatically detects faults on all preferred and alternate paths. A controller failover on installations with multiple paths to a storageset results in failed states for both the preferred and the alternate paths to the failed controller.

- Is Load Distribution enabled with more than one preferred path?

  When the load distribution property is enabled, SecurePath makes each available path to a storageset through the owning controller a preferred path. When load distribution is enabled and a single path failure occurs, SecurePath changes only the failed preferred path to the failed state. When a controller failover occurs, SecurePath changes each of the preferred paths to failed state.

Ethernet Connectivity Troubleshooting

Ethernet connectivity problems can manifest themselves in a variety of ways. The most common symptoms are the inability of the TaskSmart N-Series Cluster server to access resources on the network, the inability of network resources to access the server, and the inability of the server to connect to any resource other than the server itself. One example of a network connectivity problem is the inability of a client to access a virtual server that seems to be properly configured and working.

As with most of the hardware troubleshooting detailed in this section, the troubleshooting process involves working down the chain of possible failures, replacing each component with another one that is known to be in working condition, until the faulty component is located. In the case of Ethernet connectivity, possible points of failure include the Ethernet switches, the network cables, the NIC ports, and the NIC teaming configuration. Each of the following sections addresses these points of failure.

Ethernet Switches

Ethernet switches can experience several different types of failures, from loss of power or configuration information, and hardware failures such as port failures.

First, verify that the switch is powered up. See if any link lights are active on the switch. These lights will show that the switch is detecting and providing network connections.

Next, verify that the switch port to which the server is connected is showing an active link. If not, unplug a cable that is showing an active connection, and plug the cable of the TaskSmart N-Series Cluster server into that port. Then, plug the newly unplugged cable into the original switch port of the server.

If the new switch port to which the TaskSmart N-Series Cluster server is plugged lights up as active, and if the original port stays off, there is a problem with that port on the switch. Move the other connection to another port that works, and then verify that the TaskSmart N-Series Cluster server has network connectivity.

If the link light does not activate for the server, but does activate for the other connection, then there is a problem with the TaskSmart N-Series Cluster network cable or configuration. Proceed to the next sections.

Network Cables

Network cables are sensitive to temperature and physical hardships such as being stretched, bent, or crimped. Damage from such treatment may not manifest itself immediately. Nor does it always result in complete loss of network connectivity. It may show up much later as degraded network performance.

To eliminate the cable as the problem, replace it with another cable. Rerun the cable from the network switch to the NIC port of the TaskSmart N-Series Cluster server. View the link light on the back of the NIC card and on the Ethernet switch port. If the link light is active, test the network connectivity again.

If the new cable does not fix the problem, proceed to the next section, entitled “NIC Ports.”

NIC Ports

Although it is rare, it is possible that the NIC port or the NIC controller itself has failed. To test this possibility, swap the cables between the suspect NIC port and a working NIC port. Check to see if the link light is active on the suspect NIC port and on the working NIC port.

If the link light activates on the suspect NIC port, and if network connectivity is re-established, then there is a problem with either the network cable or the switch. (See the previous sections.)

If the link light does not activate on the suspect NIC port, but the working NIC port still works and has network connectivity, then the NIC port itself may either be configured improperly or may be faulty.

Finally, replace the NIC controller with one that is known to work. If this NIC controller also does not work, then the problem is most likely with the configuration of the NIC. Verify the networking settings on the NIC controller and on the network adapters and protocols. If this NIC controller does work, then the problem is that the original NIC controller is faulty and must be replaced.

NIC Teaming Configuration

One of the most common configuration problems is with the setup of NIC teaming on the TaskSmart N-Series Cluster. A few simple rules should be followed to ensure that NIC teaming will operate properly.

First, when NIC teaming is initially installed, the TaskSmart N-Series Cluster server should be restarted prior to configuration of the virtual network adapter or NIC teaming. This first restart completes the installation of the drivers into the network stack. Attempting to configure the network adapter before the installation is complete can result in error messages and unpredictable behavior. When the TaskSmart N-Series Cluster server restarts, network configurations can be applied to the virtual adapter and network teams can be created. When this setup is complete, the TaskSmart N-Series Cluster server must be restarted one more time for these final settings to take effect.

Second, IP addresses and other network configurations must not be applied to the component NIC ports of the NIC team. For example, do not configure the four NIC ports with IP addresses on the same subset, or assign IP addresses or any other network settings, and then attempt to combine the NIC ports into a NIC team. Doing so can result in errors and unpredictable behavior. The proper way to configure the NIC team is to combine all four NIC ports into a single NIC team, and then assign an IP address to the network adapters of the NIC team.

Third, NIC teaming should not be used to combine NIC ports from separate network boards. For example, NIC teaming should never be configured to combine the embedded NIC with one or more ports on the four-port network controller in slot 3. Network teams should be used only on the four-port network controller.

Following these rules will prevent NIC teaming configuration problems. If the network team is not functioning properly, check the configuration of the NIC teams and their network settings. Also, verify the physical connections, as detailed in the earlier sections.

QuickRestore

The most common cause of failure for the QuickRestore CD is damage to the media itself. This includes scratches on the CD surface, fingerprints or dust on the CD surface, or a defect introduced during the CD replication process.

This failure manifests itself during the QuickRestore process as a CD that will not boot, error messages that are displayed during the boot or Quick Restore process, or the Quick Restore process halting or locking-up at any point during the restore.

To determine whether the problem is a media defect, try using a different QuickRestore CD. If the new CD works, then the original media was the problem. If the problem persists with the new CD, then additional troubleshooting and investigation is required. In this case, the problem will most likely be related to the hardware or hardware settings.

Another less common problem is that of the QuickRestore process returning an error message saying that it is being used on the wrong version of the NAS product. In this case, the problem is most likely with the system ROM. The QuickRestore CD looks for the correct ROM type and revision number to ensure that it is not being used on the wrong product, the wrong product version, or on a non-TaskSmart N-Series server. The QuickRestore software has either detected the wrong ROM type, or the wrong ROM revision, on the unit. Verify during the boot process that the ROM type and date are correct for the QuickRestore version being used.

Glossary

Allocation Unit

The file system has a minimum disk block that it can read from or write to. A physical disk is divided into small blocks of 512 bytes each, called “sectors.” For convenience of the operating system (OS) and speed of access, these sectors are grouped together in blocks called file allocation units. These units are marked by the OS to indicate whether a unit is free for data to be written to, which units are in use, and which units are grouped together to make up a file.

In general, the larger the allocation unit, the faster the I/O. By using a larger allocation unit, the OS does not have to track the space and usage to the smallest detail. However, more space may be wasted. If 16 K allocation units are being used and the file to be written is 24 K, two 16 K allocation units are allocated to it, meaning 32 K of space is being used to store a 24 K file.

Array

See “RAID Array.”

Disk Spindle

See “Spindle.”

Disk Storage

Refers to a total amount of storage; for example, “make sure that you have enough disk storage to accommodate the data migration.” Disk storage amounts often refer to raw storage space, not usable storage space.

Drive

This can be either a physical drive or a disk that is being referenced through a drive letter. Two examples are: “insert the drive into the StorageWorks 4314 storage enclosure,” and “copy the file to the C: drive,” respectively.

Drive Array

See “Array” or “RAID Array.”

Drive Letter

A name or label that the Windows OS assigns to a physical device, which allows the OS and the users to interact with the storage space. For example, the boot drive on the server is labeled drive “C:,” the diskette drive is labeled drive “A:,” and the CDROM drive is labeled “B:.”

There are a finite number of drive letters. Choices are alpha letters ranging from A: to Z:, with some already in use. A physical disk does not have to have an assigned drive letter, and in the future may have one assigned to it or removed from it. However, a physical disk which does not have a drive letter assigned to it cannot be used by the OS, the users, or anyone else.

Drive Slots

Also called drive bays, these are the openings in a storage enclosure that allow an installer to insert a physical disk. For example, the TaskSmart N-Series appliance server has four drive slots. Two drive slots are used by the OS drives, and two drive slots are filled by blanks. The StorageWorks 4314 storage enclosure has 14 drive slots, therefore supporting 14 physical drives.

File Share

A resource shared out on the network for file storage. For example, by sharing out the directory “data” in CIFS, the administrator has created a file share called “data.”

File Share Resource

There are two parts to understanding this term:

A *resource* is an object that a cluster can manage and monitor for failure.

A *file share resource* is a resource in a cluster that pertains specifically to a file share. File shares are resources, just like physical disks and IP addresses. The cluster must be able to monitor a file share to manage it and fail it over.

General Use Share

A file share that is being used by everyone for any user-desired purpose. This share is not restricted to a specific user, department, application, or purpose.

Horizontal Array

A RAID array that resides entirely within a single StorageWorks 4314 storage enclosure. The 4314 Enclosure contains all the of the drives in the array. Refer to *the TaskSmart N-Series Cluster Planning Guide* for additional details.

Horizontal Method

See "Horizontal Array."

Logical Drive

The meaning of this term depends on the context in which it is used:

Compaq Array Configuration Utility (ACU)

In the ACU, RAID arrays are carved into separate logical drive partitions (logical drives). Each of these partitions is presented to the OS as a physical drive. Even though they are all included in the same RAID array, they might not be on the same physical disks.

SWVR

"Logical Drive" can refer to a logical unit number (LUN) that is being incorporated into a pool, or could refer to pool space that is being presented to the OS as a physical drive.

Windows

"Logical Drive" can refer to a device that Windows recognizes as a physical disk, a RAID array, or a partition that is being mounted on a drive letter through Logical Volume Manager (LVM).

LUN

Logical unit number (LUN). The actual meaning of this acronym depends on the context in which it is being used. The following list consists of several contexts, and the meaning of LUN in each one.

HSG80 Array Controller

Each HSG80 Array Controller presents a unit as a single physical device to the next higher layer. Each unit has its own unique identifier, and it is presented as a single physical device to the next higher level (the OS). A unit can contain one or more RAIDsets. (See RAIDsets.). The clustered NAS does not support the units of more than one RAIDset.

SCSI (physical device level)

Each physical drive must have a unique number (LUN). The controller uses these LUNs to distinguish each individual drive for reading and writing data.

SWVR

Any device that is presented as a physical disk drive to the OS is considered a LUN. It does not matter if it is an individual physical disk, a RAID array, or a logical drive. If it is presented to the OS as a physical disk that can be used for storage, SWVR calls it a LUN.

Windows

A LUN is any device that is presented as a physical disk to the OS.

Network Share

See "File Share."

Physical Disk

An individual hard drive. These are the individual drives that are pulled out of, or pushed into, a StorageWorks 4314 storage enclosure. The 4314 Enclosure can hold up to 14 physical disks.

Physical Drive

See "Physical Disk."

Pool

SANworks Virtual Replicator (SWVR) combines physical disks or RAID arrays into a single, large collection of space, which is called a "pool." From this large grouping of storage space, individual disks called "virtual disks" are carved out. A virtual disk is presented to the OS exactly like a physical disk.

For example, an SWVR pool might consist of two RAID 5 arrays, totaling ~ 360 GB (180 GB per RAID 5 array). If two 100 GB virtual disks are created from the pool, it is presented to the operating system as if two physical 100 GB disk drives were installed in the server.

RAID Array

A collection of physical disk drives that are combined together by the ACU to be presented to the hardware layer as, and be accessible as, one much larger physical disk. There are different RAID levels, which have different performance and fault tolerance characteristics. (See "Disk Spindle.")

RAIDset

This is the HSG80 Array Controller name for a RAID array.

Segment

The meaning of this term depends on the context in which it is used:

Physical disks

Refers to the smallest data block to which data may be written. For example, "the smallest segment" means the smallest disk block unit. (See "Allocation Unit.")

Network

An individual, unique subnet. For example, using a class C address (255.255.255.0 netmask), 172.1.1.x would be one network segment, 172.1.2.x another, and 173.1.1.x yet another.

SWVR

SWVR has a setting similar to the allocation unit. (See "Allocation Unit.") SWVR adds another layer of mapping to abstract out the physical storage from the virtual storage (the pool). The selected segment size determines how much space can be addressed by the map, and thus determines the maximum size to which a pool may grow. A 32 K segment size means only 250 GB can be in a pool, 64 K means 500 GB (1/2 TB), and 128 K means a full 1 TB.

Share

See "File Share."

Slots

See "Drive Slots."

Snapshot

A point in time copy of a virtual disk. This is a feature of the SWVR software package. As changes are made to the data on the original virtual disk, the Snapshot Manager makes copies of the original data. The snapshot seems to the users to be an exact copy of the data at the exact point in time that the snapshot was created. Changes to the data after the snapshot is created are not reflected in the snapshot's copy of the data.

Spindle

A single drive is called a "spindle," due to the rotating center post inside the drive. (See "Spindle Count" and "Disk Spindle.")

Spindle Count

The total number of drives, usually in reference to a RAID array. The maximum spindle count supported in a RAID 5 array by the StorageWorks 4314 storage enclosure is 14.

Storage Array

See "RAID Array."

Storage Pool

See "Pool."

Storage Unit

This is an SWVR term for the drives or RAID arrays that are used to make up an SWVR pool.

Unit

A unit is the container in which a RAIDset or mirrorset resides. A unit is a single item that is presented to the next higher level (the OS) as a physical drive. For example, unit D1 appears to the OS as if it were a single physical drive, without regard to the fact that it may consist of one or more RAIDsets. Multiple RAIDsets per unit is not a supported configuration.

Vertical Array

A RAID array that is configured to be contained within multiple StorageWorks 4314 storage enclosures, with the implication that the drives are evenly divided up among the 4314 Enclosures. For example, if there are three 4314 Enclosures, and a six-drive array, there are two drives in each of the 4314 Enclosures. NOTE: This is only an example, not a supported configuration. The only supported configuration uses vertical arrays configured with no more than one member disk of an array in each 4314 Enclosure. Refer to the *Compaq TaskSmart N-Series Cluster Planning Guide* for additional information.

Vertical Configuration

See “Vertical Array.”

Virtual Disk

The meaning of this term depends on the context:

SANworks Virtual Replicator (SWVR)

A virtual disk is a portion of space that is carved out of the pool, and is presented to the OS as if it were a single physical drive of the total size. It does not matter if the virtual disk spans across RAID arrays, or if it consists of more than a single RAID array. See “Pools” for an example.

StorageWorks Command Console(SWCC)

In SWCC, a virtual disk is a logical partition of space in a RAID array. This is very similar to the logical disk function in the Compaq Array Configuration Utility (ACU). A single RAID array can have its storage space divided up into one or more separate partitions. NOTE: Compaq does not support using SWCC virtual disks that are smaller than the full size of the RAIDset.

Virtual Server

This is a clustering term. A virtual server allows the administrator to present a cluster group and all of its resources as if it were a physical server on the network. Therefore, a single physical server can be presented as multiple servers on the network. During a cluster failover, the virtual server names and IP addresses of the failed node transfer over to the remaining cluster node, so that clients may still access its resources.