Sun Cluster 3.0 U1 System Administration Guide

Adobe PostScript

Contents

Preface

The *Sun Cluster 3.0 U1 System Administration Guide* provides procedures for administering a Sun™ Cluster 3.0 configuration.

This document is intended for experienced system administrators with extensive knowledge of Sun software and hardware. This document is not to be used as a planning or presales guide; you should have already determined your system requirements and purchased the appropriate equipment and software before reading this document.

The instructions in this book assume knowledge of the Solaris™ operating environment and expertise with the volume manager software used with Sun Cluster.

Using UNIX Commands

This document contains information on commands specific to administering a Sun Cluster configuration. This document may not contain complete information on basic UNIX® commands and procedures.

See one or more of the following for this information:

- AnswerBook2™ online documentation for the Solaris software environment
- Other software documentation that you received with your system
- Solaris operating environment man pages

Typographic Conventions

Typeface or Symbol	Meaning	Examples
`AaBbCc123`	The names of commands, files, and directories; on-screen computer output	Edit your `.login` file. Use `ls -a` to list all files. `% You have mail.`
`AaBbCc123`	What you type, when contrasted with on-screen computer output	`%` **`su`** `Password:`
AaBbCc123	Book titles, new words or terms, words to be emphasized	Read Chapter 6 in the *User's Guide*. These are called *class* options. You *must* be superuser to do this.
	Command-line variable; replace with a real name or value	To delete a file, type `rm` *filename*.

Shell Prompts

Shell	Prompt
C shell	*machine_name*%
C shell superuser	*machine_name*#
Bourne shell and Korn shell	$
Bourne shell and Korn shell superuser	#

Related Documentation

Application	Title
Hardware	*Sun Cluster 3.0 U1 Hardware Guide*
Software Installation	*Sun Cluster 3.0 U1 Installation Guide*
Data Services	*Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide*
API Development	*Sun Cluster 3.0 Data Services Developers' Guide*
Concepts	*Sun Cluster 3.0 U1 Concepts*
Error Messages	*Sun Cluster 3.0 U1 Error Messages Manual*
Release Notes	*Sun Cluster 3.0 U1 Release Notes*

Ordering Sun Documentation

Fatbrain.com, an Internet professional bookstore, stocks select product
documentation from Sun Microsystems, Inc.

For a list of documents and how to order them, visit the Sun Documentation Center
on Fatbrain.com at:

`http://www1.fatbrain.com/documentation/sun`

Accessing Sun Documentation Online

The `docs.sun.com`[SM] web site enables you to access Sun technical documentation
on the Web. You can browse the `docs.sun.com` archive or search for a specific book
title or subject at:

`http://docs.sun.com`

Getting Help

If you have problems installing or using Sun Cluster, contact your service provider and provide the following information:

- Your name and email address (if available)
- Your company name, address, and phone number
- The model and serial numbers of your systems
- The release number of the operating environment (for example, Solaris 8)
- The release number of Sun Cluster (for example, Sun Cluster 3.0)

Use the following commands to gather information on your system for your service provider:

Command	Function
prtconf -v	Displays the size of the system memory and reports information about peripheral devices
psrinfo -v	Displays information about processors
showrev -p	Reports which patches are installed
prtdiag -v	Displays system diagnostic information
/usr/cluster/bin/scinstall -pv	Displays Sun Cluster release and package version information

Also have available the contents of the /var/adm/messages file.

Introduction to Administering Sun Cluster

This chapter provides information on preparing to administer the cluster and the procedures for using Sun Cluster administration tools.

This is a list of the procedures in this chapter.

Administering Sun Cluster Overview

Sun Cluster's highly-available environment ensures that critical applications are available to end users. The system administrator's job is to make sure that Sun Cluster is stable and operational.

Before undertaking an administrative task, you should have familiarized yourself with the planning information in the *Sun Cluster 3.0 U1 Installation Guide* and the glossary in the *Sun Cluster 3.0 U1 Concepts* document. Sun Cluster administration is organized into tasks among the following manuals.

- Standard tasks, used to administer and maintain the cluster on a regular, perhaps daily basis. These tasks are described in this guide.

- Data service tasks, such as installation, configuration, and changing properties. These tasks are described in the *Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide*.
- Service tasks, such as adding or repairing storage or network hardware. These tasks are described in the *Sun Cluster 3.0 U1 Hardware Guide*.

For the most part, you can perform Sun Cluster administration tasks while the cluster is operational, with the impact on cluster operation limited to a single node. For those procedures that require that the entire cluster be shut down, schedule downtime for off hours, when there is minimal impact on the system. If you plan to take down the cluster or a cluster node, notify users ahead of time.

Administration Tools

You can perform administrative tasks on Sun Cluster using a Graphical User Interface (GUI) or using the command line. This section provides an overview of these tools.

Graphical User Interface

Sun Cluster supports two Graphical User Interface (GUI) tools that you can use to perform various administrative tasks on your cluster. These GUI tools are SunPlex Manager and Sun Management Center. See Chapter 9 for more information and for procedures about configuring SunPlex Manager and Sun Management Center. For specific information about how to use these tools, see the online help for each GUI.

Command Line Interface

You can perform most Sun Cluster administration tasks interactively through the `scsetup(1M)` utility. Whenever possible, administration procedures in this guide are described using `scsetup`.

You can administer the following items through the `scsetup` utility.

- Quorum
- Resource groups
- Cluster interconnect
- Device groups and volumes
- Private hostnames
- New nodes

■ Other cluster properties

Listed here are the other commands you use to administer Sun Cluster. See the man pages for more detailed information.

TABLE 1-1   Sun Cluster Command-Line Interface Commands

Command	Description
ccp(1M)	Starts remote console access to the cluster.
pmfadm(1M)	Provides administrative access to the process monitor facility.
pnmset(1M)	Configures Public Network Management (PNM).
pnmstat(1M)	Reports the status of Network Adapter Failover (NAFO) groups monitored by PNM.
sccheck(1M)	Checks and validates the global mount entries in the /etc/vfstab file.
scconf(1M)	Updates a Sun Cluster configuration. The -p option lists cluster configuration information.
scdidadm(1M)	Provides administrative access to the device ID configuration.
scgdevs(1M)	Runs the global device namespace administration script.
scinstall(1M)	Installs and configures Sun Cluster software; can be run interactively or non-interactively. The -p option displays release and package version information for the Sun Cluster software.
scrgadm(1M)	Manages the registration of resource types, the creation of resource groups, and the activation of resources within a resource group. The -p option displays information on installed resources, resource groups, and resource types.
scsetup(1M)	Runs the interactive cluster configuration utility, which generates the scconf command and its various options.
scshutdown(1M)	Shuts down the entire cluster.
scstat(1M)	Provides a snapshot of the cluster status.
scswitch(1M)	Performs changes affecting node mastery and states for resource groups and disk device groups.

In addition, you use commands to administer the volume manager portion of Sun Cluster. These commands depend on the specific volume manager used in your cluster, either Solstice DiskSuite™ or VERITAS Volume Manager.

Preparing to Administer the Cluster

This section describes what to do to prepare for administering your cluster.

Documenting a Sun Cluster Hardware Configuration

As your Sun Cluster configuration grows and changes, documenting the hardware aspects that are unique to your site saves administration time when it becomes necessary to change or upgrade the cluster. Labeling cables and connections between the various cluster components can also make administration easier.

Keeping records of your original cluster configuration, and subsequent changes, can also help to reduce the time required by a third-party service provider when servicing your cluster.

Using an Administrative Console

You can use a dedicated SPARC workstation, known as the *administrative console*, to administer the active cluster. Typically, you install and run the Cluster Control Panel (CCP) and graphical user interface (GUI) tools on the administrative console. For more information on the CCP, see "How to Remotely Log In to Sun Cluster" on page 6. For instructions on installing the Sun Management Center and SunPlex Manager GUI tools, see the *Sun Cluster 3.0 U1 Installation Guide*.

The administrative console is not a cluster node. The administrative console is used for remote access to the cluster nodes, either over the public network or through a network-based terminal concentrator.

If your cluster consists of a Sun Enterprise™ 10000 server, you must have the ability to log in from the administrative console to the System Service Processor (SSP) and connect using the netcon(1M) command. The default method for netcon to connect with a Sun Enterprise 10000 domain is through the network interface. If the network is inaccessible, the cluster console (cconsole) access through the network connection will hang. To prevent this, you can use netcon in "exclusive" mode by setting the -f option or by sending ~* during a normal netcon session. This gives you the option of toggling to the serial interface if the network becomes unreachable. Refer to netcon(1M) for more information.

Sun Cluster does not require a dedicated administrative console, but using one provides these benefits:

- Enables centralized cluster management by grouping console and management tools on the same machine
- Provides potentially quicker problem resolution by Enterprise Services or your service provider

Backing Up the Cluster

It is important to back up your cluster on a regular basis. Even though Sun Cluster provides an HA environment, with mirrored copies of data on the storage devices, do not consider this to be a replacement for regular backups. Sun Cluster can survive multiple failures, but it does not protect against user or program error, or catastrophic failure. Therefore, you must have a backup procedure in place to protect against data loss.

The following information should be included as part of your backup.

- All file system partitions
- All database data if you are running DBMS data services
- Disk partition information for all cluster disks
- The `md.tab` file if you are using Solstice DiskSuite as your volume manager

Beginning to Administer the Cluster

TABLE 1-2 provides a starting point for administering your cluster.

TABLE 1-2   Sun Cluster 3.0 Administration Tools

If You Want To...	Then...	For More Information Go To...
Remotely Log in to the Cluster	Use the `ccp` command to launch the Cluster Control Panel (CCP). Then select one of the following icons: `cconsole`, `crlogin`, or `ctelnet`.	"How to Remotely Log In to Sun Cluster" on page 6
Interactively Configure the Cluster	Launch the `scsetup` utility.	"How to Access the scsetup Utility" on page 8
Display Sun Cluster Release Number and Version Information	Use the `scinstall` command with either the `-p` or `-pv` options.	"How to Display Sun Cluster Release and Version Information" on page 8

TABLE 1-2   Sun Cluster 3.0 Administration Tools

If You Want To...	Then...	For More Information Go To...
Display Installed Resources, Resource Groups, and Resource Types	Use the `scrgadm -p` command.	"How to Display Configured Resource Types, Resource Groups, and Resources" on page 9
Graphically Monitor Cluster Components	Use SunPlex Manager or the Sun Cluster module for Sun Management Center.	SunPlex Manager or Sun Cluster module for Sun Management Center online help
Graphically Administer Some Cluster Components	Use SunPlex Manager or the Sun Cluster module for Sun Management Center.	SunPlex Manager or Sun Cluster module for Sun Management Center online help
Check the Status of Cluster Components	Use the `scstat` command.	"How to Check the Status of Cluster Components" on page 10
View the Cluster Configuration	Use the `scconf -p` command.	"How to View the Cluster Configuration" on page 13
Check Global Mount Points	Use the `sccheck` command.	"How to Check the Global Mount Points" on page 15
Look at Sun Cluster System Messages	Examine the `/var/adm/messages` file.	Solaris system administration documentation
Monitor the Status of Solstice DiskSuite	Use the `metastat` commands.	Solstice DiskSuite documentation
Monitor the Status of VERITAS Volume Manager	Use the `vxstat` or `vxva` commands.	VERITAS Volume Manager documentation

▼ How to Remotely Log In to Sun Cluster

The Cluster Control Panel (CCP) provides a launch pad for `cconsole`, `crlogin`, and `ctelnet` tools. All three tools start a multiple window connection to a set of specified nodes. The multiple-window connection consists of a host window for each of the specified nodes and a common window. Input directed into the common window is sent to each of these host windows, allowing you to run commands simultaneously on all nodes of the cluster. See the `ccp(1M)` and `cconsole(1M)` man pages for more information.

1. Verify that the following prerequisites are met before starting the CCP.

   ■ Install the appropriate Sun Cluster software (`SUNWccon` package) on the administrative console.

- Make sure the PATH variable on the administrative console includes the Sun Cluster tools directory, /opt/SUNWcluster/bin, and /usr/cluster/bin. You can specify an alternate location for the tools directory by setting the $CLUSTER_HOME environment variable.

- Configure the clusters file, the serialports file, and the nsswitch.conf file if using a terminal concentrator. These can be either /etc files or NIS/NIS+ databases. See clusters(4) and serialports(4) for more information.

2. **Determine if you have a Sun Enterprise E10000 server platform.**

   - If no, proceed to .

   - If yes, log into the System Service Processor (SSP) and connect by using the netcon command. Once connected, type Shift~@ to unlock the console and gain write access.

3. **Start the CCP launch pad.**

   From the administrative console, type the following command.

   ```
   # ccp clustername
   ```

   The CCP launch pad is displayed.

4. **To start a remote session with the cluster, click the appropriate icon (cconsole, crlogin, or ctelnet) in the CCP launch pad.**

Example

The following example shows the Cluster Control Panel.
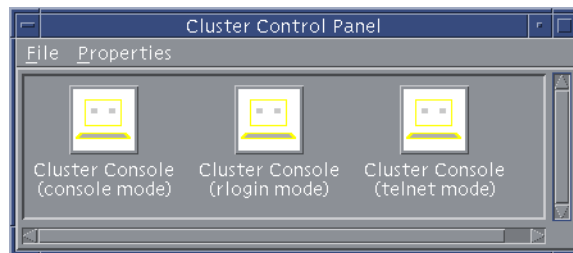
FIGURE 1-1   Cluster Control Panel

Where to Go From Here

You can also start `cconsole`, `crlogin`, or `ctelnet` sessions from the command line. See `cconsole(1M)` for more information.

▼ How to Access the `scsetup` Utility

The `scsetup(1M)` utility enables you to interactively configure quorum, resource group, cluster transport, private hostname, device group, and new node options for the cluster.

1. **Become superuser on any node in the cluster.**

2. **Enter the** `scsetup` **utility.**

   ```
   # scsetup
   ```

   The Main Menu is displayed.

3. **Make your selection from the menu and follow the onscreen instructions.**

   See the `scsetup` online help for more information.

▼ How to Display Sun Cluster Release and Version Information

You do not need to be logged in as superuser to perform these procedures.

● **Display the Sun Cluster patch numbers.**

   Sun Cluster update releases are identified by the main product patch number plus the update version, which is 110648-05 for Sun Cluster 3.0 U1.

   ```
   % showrev -p
   ```

● **Display the Sun Cluster release number and version strings for all Sun Cluster packages.**

   ```
   % scinstall -pv
   ```

Example—Displaying the Sun Cluster Release Number

The following example displays the cluster's release number.

```
% showrev -p | grep 110648
Patch: 110648-05 Obsoletes:  Requires:  Incompatibles:  Packages:
```

Example—Displaying Sun Cluster Release and Version Information

The following example displays the cluster's release information and version information for all packages.

```
% scinstall -pv
SunCluster 3.0
SUNWscr:      3.0.0,REV=2000.10.01.01.00
SUNWscdev:    3.0.0,REV=2000.10.01.01.00
SUNWscu:      3.0.0,REV=2000.10.01.01.00
SUNWscman:    3.0.0,REV=2000.10.01.01.00
SUNWscsal:    3.0.0,REV=2000.10.01.01.00
SUNWscsam:    3.0.0,REV=2000.10.01.01.00
SUNWscvm:     3.0.0,REV=2000.10.01.01.00
SUNWmdm:      4.2.1,REV=2000.08.08.10.01
```

▼ How to Display Configured Resource Types, Resource Groups, and Resources

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

You do not need to be logged in as superuser to perform this procedure.

● **Display the cluster's configured resource types, resource groups, and resources.**

```
% scrgadm -p
```

Example—Displaying Configured Resource Types, Resource Groups, and Resources

The following example shows the resource types (RT Name), resource groups (RG Name), and resources (RS Name) configured for the cluster schost.

```
% scrgadm -p
RT Name: SUNW.SharedAddress
  RT Description: HA Shared Address Resource Type
RT Name: SUNW.LogicalHostname
  RT Description: Logical Hostname Resource Type
RG Name: schost-sa-1
  RG Description:
    RS Name: schost-1
      RS Description:
      RS Type: SUNW.SharedAddress
      RS Resource Group: schost-sa-1
RG Name: schost-lh-1
  RG Description:
    RS Name: schost-3
      RS Description:
      RS Type: SUNW.LogicalHostname
      RS Resource Group: schost-lh-1
```

▼ How to Check the Status of Cluster Components

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

You do not need to be logged in as superuser to perform this procedure.

● **Check the status of cluster components.**

```
% scstat -p
```

Example—Checking the Status of Cluster Components

The following example provides a sample of status information for cluster components returned by scstat(1M).

```
% scstat -p
-- Cluster Nodes --

                    Node name          Status
                    ---------          ------
  Cluster node:     phys-schost-1      Online
  Cluster node:     phys-schost-2      Online
  Cluster node:     phys-schost-3      Online
  Cluster node:     phys-schost-4      Online

------------------------------------------------------------------


-- Cluster Transport Paths --

                    Endpoint           Endpoint          Status
                    --------           --------          ------
  Transport path:   phys-schost-1:qfe1 phys-schost-4:qfe1 Path online
  Transport path:   phys-schost-1:hme1 phys-schost-4:hme1 Path online
...

------------------------------------------------------------------


-- Quorum Summary --

  Quorum votes possible:      6
  Quorum votes needed:        4
  Quorum votes present:       6


-- Quorum Votes by Node --

                    Node Name          Present Possible Status
                    ---------          ------- -------- ------
  Node votes:       phys-schost-1      1       1        Online
  Node votes:       phys-schost-2      1       1        Online
...
```

```
-- Quorum Votes by Device --

                      Device Name         Present Possible Status
                      -----------         ------- -------- ------
  Device votes:       /dev/did/rdsk/d2s2  1          1        Online
  Device votes:       /dev/did/rdsk/d8s2  1          1        Online
...


-- Device Group Servers --

                         Device Group        Primary            Secondary
                         ------------         -------            ---------
  Device group servers:  rmt/1               -                  -
  Device group servers:  rmt/2               -                  -
  Device group servers:  schost-1            phys-schost-2      phys-schost-1
  Device group servers:  schost-3            -                  -


-- Device Group Status --

                            Device Group        Status
                            ------------         ------
  Device group status:      rmt/1               Offline
  Device group status:      rmt/2               Offline
  Device group status:      schost-1            Online
  Device group status:      schost-3            Offline

------------------------------------------------------------------


-- Resource Groups and Resources --

           Group Name           Resources
           ----------           ---------
 Resources: test-rg             test_1
 Resources: real-property-rg    -
 Resources: failover-rg         -
 Resources: descript-rg-1       -
...
```

```
-- Resource Groups --

          Group Name          Node Name          State
          ----------          ---------          -----
    Group: test-rg            phys-schost-1      Offline
    Group: test-rg            phys-schost-2      Offline
...


-- Resources --

          Resource Name       Node Name          State      Status Message
          -------------       ---------          -----      --------------
  Resource: test_1            phys-schost-1      Offline    Offline
  Resource: test_1            phys-schost-2      Offline    Offline
```

▼ How to View the Cluster Configuration

You can also accomplish this procedure by using the SunPlex Manager GUI. See the
SunPlex Manager online help for more information.

You do not need to be logged in as superuser to perform this procedure.

● **View the cluster configuration.**

```
% scconf -p
```

Example—Viewing the Cluster Configuration

The following example lists the cluster configuration.

```
% scconf -p
Cluster name:                      cluster-1
Cluster ID:                        0x3908EE1C
Cluster install mode:              disabled
Cluster private net:               172.16.0.0
Cluster private netmask:           255.255.0.0
Cluster new node authentication:   unix
Cluster new node list:             <NULL - Allow any node>
Cluster nodes:                     phys-schost-1 phys-schost-2 phys-schost-3
phys-schost-4

Cluster node name:                            phys-schost-1
  Node ID:                                    1
  Node enabled:                               yes
  Node private hostname:                      clusternode1-priv
  Node quorum vote count:                     1
  Node reservation key:                       0x3908EE1C00000001
  Node transport adapters:                    hme1 qfe1 qfe2


Node transport adapter:                       hme1
    Adapter enabled:                          yes
    Adapter transport type:                   dlpi
    Adapter property:                         device_name=hme
    Adapter property:                         device_instance=1
    Adapter property:                         dlpi_heartbeat_timeout=10000
...
```

```
Cluster transport junctions:                              hub0 hub1 hub2

Cluster transport junction:                               hub0
  Junction enabled:                                       yes
  Junction type:                                          switch
  Junction port names:                                    1 2 3 4
...

Junction port:                                    1
    Port enabled:                                         yes

Junction port:                                    2
    Port enabled:                                         yes
...
Cluster transport cables
                   Endpoint          Endpoint          State
                   --------          --------          -----
  Transport cable:  phys-schost-1:hme1@0 hub0@1         Enabled
  Transport cable:  phys-schost-1:qfe1@0 hub1@1         Enabled
  Transport cable:  phys-schost-1:qfe2@0 hub2@1         Enabled
  Transport cable:  phys-schost-2:hme1@0 hub0@2         Enabled
...
Quorum devices:                                           d2 d8

Quorum device name:                                       d2
  Quorum device votes:                                    1
  Quorum device enabled:                                  yes
  Quorum device name:                                     /dev/did/rdsk/d2s2
  Quorum device hosts (enabled):                          phys-schost-1
 phys-schost-2
  Quorum device hosts (disabled):
...

Device group name:                                        schost-3
  Device group type:                                      SDS
  Device group failback enabled:                          no
  Device group node list:                                 phys-schost-3, phys-schost-4
  Diskset name:                                           schost-3
```

▼ How to Check the Global Mount Points

The sccheck(1M) command checks the /etc/vfstab file for configuration errors
with the cluster file system and its global mount points. The sccheck command
only returns errors. If no errors are found, sccheck merely returns to the shell
prompt.

> **Note –** Run sccheck after making cluster configuration changes that have affected
> devices or volume management components.

1. **Become superuser on any node in the cluster.**

2. **Verify the cluster configuration.**

   ```
   # sccheck
   ```

Example—Verifying the Cluster Configuration

The following example shows that the node phys-schost-3 is missing the mount
point /global/schost-1.

```
# sccheck
vfstab-check: WARNING - phys-schost-3 - Missing mount point /global/schost-1
```

Shutting Down and Booting a Cluster

This chapter provides the procedures for shutting down and booting a cluster and individual cluster nodes.

This is a list of the procedures in this chapter.

For a high-level description of the related procedures in this chapter, see TABLE 2-1 and TABLE 2-2.

Shutting Down and Booting a Cluster Overview

The Sun Cluster scshutdown(1M) command stops cluster services in an orderly fashion and cleanly shuts down the entire cluster. You might do this when moving a cluster from one location to another or if there was data corruption caused by an application error.

Note – Use scshutdown instead of the shutdown or halt commands to ensure proper shutdown of the entire cluster. The Solaris shutdown command is used with the scswitch command to shut down individual nodes. See "How to Shut Down a Cluster" on page 19 or "Shutting Down and Booting a Single Cluster Node" on page 24 for more information.

The scshutdown command stops all nodes in a cluster by:

1. Taking all running resource groups offline.

2. Unmounting all cluster file systems.

3. Shutting down active device services.

4. Running init 0 and bringing all nodes to the OBP ok prompt.

Note – If necessary, you can boot a node in non-cluster mode so that it does not participate in cluster membership. This is useful when installing cluster software or for performing certain administrative procedures. See "How to Boot a Cluster Node in Non-Cluster Mode" on page 30 for more information.

TABLE 2-1   Task List: Shutting Down and Booting a Cluster

Task	For Instructions, Go To...
Stop the cluster - Use scshutdown	"How to Shut Down a Cluster" on page 19
Start the cluster by booting all nodes. The nodes must have a working connection to the cluster interconnect to attain cluster membership.	"How to Boot a Cluster" on page 20
Shut down and reboot the cluster - Use scshutdown At the ok prompt, boot each node individually with the boot command. The nodes must have a working connection to the cluster interconnect to attain cluster membership.	"How to Reboot a Cluster" on page 21

▼ How to Shut Down a Cluster

> ⚠ **Caution –** Do not use `send brk` on a cluster console to shut down a cluster node. This functionality is not supported within a cluster. If you use `send brk` and then enter `go` at the `ok` prompt to reboot, the node will panic.

1. **If your cluster is running Oracle Parallel Server (OPS), shut down all OPS database instances.**

   Refer to the Oracle Parallel Server product documentation for shutdown procedures.

2. **Become superuser on any node in the cluster.**

3. **Shut down the cluster immediately to OBP.**

   From a single node in the cluster, type the following command.

   ```
   # scshutdown -g0 -y
   ```

4. **Verify that all nodes have reached the** `ok` **prompt.**

   Do not power off any nodes until all cluster nodes are at the `ok` prompt.

5. **If necessary, power off the nodes.**

Example—Shutting Down a Cluster

The following example shows the console output when stopping normal cluster operation and bringing down all nodes to the `ok` prompt. The `-g 0` option sets the shutdown grace period to zero, `-y` provides an automatic `yes` response to the confirmation question. Shutdown messages also appear on the consoles of the other nodes in the cluster.

```
# scshutdown -g0 -y
May 2 10:08:46 phys-schost-1 cl_runtime: WARNING: CMM monitoring disabled.
phys-schost-1#
INIT: New run level: 0
The system is coming down.  Please wait.
System services are now being stopped.
/etc/rc0.d/K05initrgm: Calling scswitch -S (evacuate)
The system is down.
syncing file systems... done
Program terminated
ok
```

Where to Go From Here

See to restart a cluster that has been shut down.

▼ How to Boot a Cluster

1. **To start a cluster whose nodes have been shut down and are at the** ok **prompt, boot each node.**

   If you make configuration changes between shutdowns, start the node with the most current configuration first. Except in this situation, the boot order of the nodes does not matter.

   ```
   ok boot
   ```

   Messages are displayed on the booted nodes' consoles as cluster components are activated.

   Note – Cluster nodes must have a working connection to the cluster interconnect to attain cluster membership.

   ---

2. **Verify that the nodes booted without error and are online.**

   The scstat(1M) command reports the nodes' status.

   ```
   # scstat -n
   ```

   Note – If a cluster node's /var file system fills up, Sun Cluster might not be able to restart on that node. If this problem arises, see "How to Repair a Full /var File System" on page 32.

Example—Booting a Cluster

The following example shows the console output when booting node `phys-schost-1` into the cluster. Similar messages appear on the consoles of the other nodes in the cluster.

```
ok boot
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
NOTICE: Node 1 with votecount = 1 added.
NOTICE: Node 2 with votecount = 1 added.
NOTICE: Node 3 with votecount = 1 added.
...
NOTICE: Node 1: attempting to join cluster
...
NOTICE: Node 2 (incarnation # 937690106) has become reachable.
NOTICE: Node 3 (incarnation # 937690290) has become reachable.
NOTICE: cluster has reached quorum.
NOTICE: node 1 is up; new incarnation number = 937846227.
NOTICE: node 2 is up; new incarnation number = 937690106.
NOTICE: node 3 is up; new incarnation number = 937690290.
NOTICE: Cluster members:    1   2   3
...
```

▼ How to Reboot a Cluster

Run the `scshutdown`(1M) command to shut down the cluster, then boot the cluster with the `boot` command on each node.

1. **(Optional). For a cluster running Oracle Parallel Server (OPS), shut down all OPS database instances.**

   Refer to the Oracle Parallel Server product documentation for shutdown procedures.

2. **Become superuser on any node in the cluster.**

3. **Shut down the cluster to OBP.**

   From a single node in the cluster, type the following command.

   ```
   # scshutdown -g0 -y
   ```

   This shuts down each node to the `ok` prompt.

> **Note –** Cluster nodes must have a working connection to the cluster interconnect to attain cluster membership.

4. Boot each node.

The order in which the nodes are booted does not matter unless you make configuration changes between shutdowns. In this case, you should start the nodes such that the one with the most current configuration boots first.

```
ok boot
```

Messages appear on the booted nodes' consoles as cluster components are activated.

5. Verify that the nodes booted without error and are online.

The scstat command reports the nodes' status.

```
# scstat -n
```

> **Note –** If a cluster node's /var file system fills up, Sun Cluster might not be able to restart on that node. If this problem arises, see "How to Repair a Full /var File System" on page 32.

Example—Rebooting a Cluster

The following example shows the console output when stopping normal cluster operation, bringing down all nodes to the ok prompt, then restarting the cluster. The -g 0 option sets the grace period to zero, -y provides an automatic yes response to the confirmation question. Shutdown messages also appear on the consoles of other nodes in the cluster.

```
# scshutdown -g0 -y
May 2 10:08:46 phys-schost-1 cl_runtime: WARNING: CMM monitoring disabled.
phys-schost-1#
INIT: New run level: 0
The system is coming down.  Please wait.
...
The system is down.
syncing file systems... done
Program terminated

ok boot
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
...
NOTICE: Node 1: attempting to join cluster
...
NOTICE: Node 2 (incarnation # 937690106) has become reachable.
NOTICE: Node 3 (incarnation # 937690290) has become reachable.
NOTICE: cluster has reached quorum.
...

NOTICE: Cluster members:   1  2  3
...
NOTICE: Node 1: joined cluster
...
The system is coming up.  Please wait.
checking ufs filesystems
...
```

```
reservation program successfully exiting
Print services started.
volume management starting.
The system is ready.
phys-schost-1 console login:

NOTICE: Node 1: joined cluster
...
The system is coming up.   Please wait.
checking ufs filesystems
...

reservation program successfully exiting
Print services started.
volume management starting.
The system is ready.
phys-schost-1 console login:
```

Shutting Down and Booting a Single Cluster Node

Note – Use the scswitch command in conjunction with the Solaris shutdown command to shut down an individual node. Use the scshutdown command only when shutting down an entire cluster.

TABLE 2-2 Task Map: Shutting Down and Booting a Cluster Node

Task	For Instructions, Go To...
Stop a cluster node - Use scswitch(1M) and shutdown(1M)	"How to Shut Down a Cluster Node" on page 25

TABLE 2-2   Task Map: Shutting Down and Booting a Cluster Node

Task	For Instructions, Go To...
Start a node by booting it The node must have a working connection to the cluster interconnect to attain cluster membership.	"How to Boot a Cluster Node" on page 27
Stop and restart (reboot) a cluster node - Use scswitch and shutdown The node must have a working connection to the cluster interconnect to attain cluster membership.	"How to Reboot a Cluster Node" on page 28
Boot a node so that it does not participate in cluster membership - Use scswitch and shutdown, then boot -x	"How to Boot a Cluster Node in Non-Cluster Mode" on page 30

▼ How to Shut Down a Cluster Node

Caution – Do not use send brk on a cluster console to shut down a cluster node. If you use send brk and then enter go at the ok prompt to reboot, the node will panic. This functionality is not supported within a cluster.

1. **If you are running Oracle Parallel Server (OPS), shut down all OPS database instances.**

   Refer to the Oracle Parallel Server product documentation for shutdown procedures.

2. **Become superuser on the cluster node to be shut down.**

3. **Switch all resource groups, resources, and device groups from the node being shut down to other cluster members.**

   On the node to be shut down, type the following command.

   ```
   # scswitch -S -h nodelist
   ```

   -S   Evacuates all device services and resource groups from the specified node.

   -h *nodelist*   Specifies the node from which you are switching resource groups and device groups.

4. **Shut down the cluster node to OBP.**

   On the node to be shut down, type the following command.

   ```
   # shutdown -g0 -y -i0
   ```

5. **Verify that the cluster node has reached the** ok **prompt.**

6. **If necessary, power off the node.**

Example—Shutting Down a Cluster Node

The following example shows the console output when shutting down node phys-schost-1. The -g0 option sets the grace period to zero, -y provides an automatic yes response to the confirmation question, and -i0 invokes run level 0 (zero). Shutdown messages for this node appear on the consoles of other nodes in the cluster.

```
# scswitch -S -h phys-schost-1
# shutdown -g0 -y -i0
May 2 10:08:46 phys-schost-1 cl_runtime: WARNING: CMM monitoring disabled.
phys-schost-1#
INIT: New run level: 0
The system is coming down.  Please wait.
Notice: rgmd is being stopped.
Notice: rpc.pmfd is being stopped.
Notice: rpc.fed is being stopped.
umount: /global/.devices/node@1 busy
umount: /global/phys-schost-1 busy
The system is down.
syncing file systems... done
Program terminated
ok
```

Where to Go From Here

See to restart a cluster node that has been shut down.

▼ How to Boot a Cluster Node

> **Note –** Starting a cluster node can be affected by the quorum configuration. In a two-node cluster, you must have a quorum device configured so that the total quorum count for the cluster is three (one for each node and one for the quorum device). In this situation, if the first node is shut down, the second node continues to have quorum and runs as the sole cluster member. For the first node to come back in the cluster as a cluster node, the second node must be up and running and the required cluster quorum count (two) must be present.

1. **To start a cluster node that has been shut down, boot the node.**

```
ok boot
```

   Messages are displayed on all node consoles as cluster components are activated.

   > **Note –** A cluster node must have a working connection to the cluster interconnect to attain cluster membership.

2. **Verify that the node has booted without error, and is online.**

   The scstat(1M) command reports the status of a node.

```
# scstat -n
```

   > **Note –** If a cluster node's /var file system fills up, Sun Cluster might not be able to restart on that node. If this problem arises, see "How to Repair a Full /var File System" on page 32.

Example—Booting a Cluster Node

The following example shows the console output when booting node
`phys-schost-1` into the cluster.

```
ok boot
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
...
NOTICE: Node 1: attempting to join cluster
...
NOTICE: Node 1: joined cluster
...
The system is coming up.  Please wait.
checking ufs filesystems
...
reservation program successfully exiting
Print services started.
volume management starting.
The system is ready.
phys-schost-1 console login:
```

▼ How to Reboot a Cluster Node

1. **If the cluster node is running Oracle Parallel Server (OPS), shut down all OPS database instances.**

   Refer to the Oracle Parallel Server product documentation for shutdown procedures.

2. **Become superuser on the cluster node to be shut down.**

3. **Shut down the cluster node by using the** `scswitch` **and** `shutdown` **commands.**

   Enter these commands on the node to be shut down. The `-i  6` option with the
   `shutdown` command causes the node to reboot after it shuts down to the `ok` prompt.

   ```
   # scswitch -S -h nodelist
   # shutdown -g0 -y -i6
   ```

Note – Cluster nodes must have a working connection to the cluster interconnect to
attain cluster membership.

---

4. **Verify that the node has booted without error, and is online.**

```
# scstat -n
```

Example—Rebooting a Cluster Node

The following example shows the console output when rebooting node
phys-schost-1. Shutdown and startup messages for this node appear on the
consoles of other nodes in the cluster.

```
# scswitch -S -h phys-schost-1
# shutdown -g0 -y -i6
May 2 10:08:46 phys-schost-1 cl_runtime: WARNING: CMM monitoring disabled.
phys-schost-1#
INIT: New run level: 6
The system is coming down.  Please wait.
System services are now being stopped.
Notice: rgmd is being stopped.
Notice: rpc.pmfd is being stopped.
Notice: rpc.fed is being stopped.
umount: /global/.devices/node@1 busy
umount: /global/phys-schost-1 busy
The system is down.
syncing file systems... done
rebooting...
Resetting ...
,,,
Sun Ultra 1 SBus (UltraSPARC 143MHz), No Keyboard
OpenBoot 3.11, 128 MB memory installed, Serial #5932401.
Ethernet address 8:8:20:99:ab:77, Host ID: 8899ab77.
...
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
...
NOTICE: Node 1: attempting to join cluster
...
NOTICE: Node 1: joined cluster
...
The system is coming up.  Please wait.
The system is ready.
phys-schost-1 console login:
```

▼ How to Boot a Cluster Node in Non-Cluster Mode

You can boot a node so that it does not participate in the cluster membership, that is, in non-cluster mode. This is useful when installing the cluster software or for performing certain administrative procedures, such as patching a node.

1. **Become superuser on the cluster node to be started in non-cluster mode.**

2. **Shut down the node by using the** scswitch **and** shutdown **commands.**

   ```
   # scswitch -S -h nodelist
   # shutdown -g0 -y -i0
   ```

3. **Verify that the node is at the** ok **prompt.**

4. **Boot the node in non-cluster mode by using the** boot(1M) **command with the** -x **option.**

   ```
   ok boot -x
   ```

   Messages appear on the node's console stating that the node is not part of the cluster.

Example—Booting a Cluster Node in Non-Cluster Mode

The following example shows the console output when shutting down node `phys-schost-1` then restarting it in non-cluster mode. The `-g0` option sets the grace period to zero, `-y` provides an automatic `yes` response to the confirmation question, and `-i0` invokes run level 0 (zero). Shutdown messages for this node appear on the consoles of other nodes in the cluster.

```
# scswitch -S -h phys-schost-1
# shutdown -g0 -y -i0
May 2 10:08:46 phys-schost-1 cl_runtime: WARNING: CMM monitoring disabled.
phys-schost-1#
...
rg_name = schost-sa-1 ...
offline node = phys-schost-2 ...
num of  node = 0 ...
phys-schost-1#
INIT: New run level: 0
The system is coming down.  Please wait.
System services are now being stopped.
Print services stopped.
syslogd: going down on signal 15
...
The system is down.
syncing file systems... done
WARNING: node 1 is being shut down.
Program terminated

ok boot -x
...
Not booting as part of cluster
...
The system is ready.
phys-schost-1 console login:
```

Repairing a Full /var File System

Both Solaris and Sun Cluster software write error messages to the /var/adm/messages file, which over time can fill the /var file system. If a cluster node's /var file system fills up, Sun Cluster might not be able to restart on that node. Additionally, you might not be able to log in to the node.

▼ How to Repair a Full /var File System

If a node reports a full /var file system and continues to run Sun Cluster services, use this procedure to clear the full file system.

1. **Become superuser on the cluster node with the full /var file system.**

2. **Clear the full file system.**

   For example, delete nonessential files contained in the file system.

Administering Global Devices and Cluster File Systems

This chapter provides the procedures for administering global devices and cluster file systems.

This is a list of the procedures in this chapter.

For a high-level description of the related procedures in this chapter, see TABLE 3-1.

See the *Sun Cluster 3.0 U1 Concepts* document for conceptual information related to global devices, the global namespace, disk device groups, and the cluster file system.

Administering Global Devices and the Global Namespace Overview

Administration of Sun Cluster disk device groups depends on the volume manager installed on the cluster. Solstice DiskSuite is "cluster-aware," so you add, register, and remove disk device groups by using the Solstice DiskSuite `metaset(1M)` command. With VERITAS Volume Manager (VxVM), you create disk groups by using VxVM commands. Then you register the disk groups as Sun Cluster disk device groups through the `scsetup(1M)` utility. When removing VxVM disk device groups, you use both the `scsetup` utility and VxVM commands.

When administering disk device groups, or volume manager disk groups, you need to be on the cluster node that is the primary node for the group.

Normally, you do not need to administer the global device namespace because the global namespace is automatically set up during installation and automatically updated during Solaris operating environment reconfiguration reboots. However, if the global namespace needs to be regenerated or updated, you can run the `scgdevs(1M)` command from any cluster node. This causes the global namespace to be updated on all other cluster node members, as well as on nodes that might join the cluster in the future.

Global Device Permissions for Solstice DiskSuite

Changes made to global device permissions are not automatically propagated to all the nodes in the cluster for Solstice DiskSuite and disk devices. If you want to change permissions on global devices, you must manually change the permissions on all the nodes in the cluster. For example, if you want to change permissions on global device `/dev/global/dsk/d3s0` to 644, you must execute

```
# chmod 644 /dev/global/dsk/d3s0
```

on all nodes in the cluster.

VxVM does not support the `chmod` command. To change global device permissions in VxVM, consult the VxVM administrator's guide.

VERITAS Volume Manager Administration Considerations

For Sun Cluster to maintain the VxVM namespace, you must register any VxVM disk group or volume changes as Sun Cluster disk device group configuration changes. Registering these changes ensures that the namespace on all cluster nodes is updated. Examples of configuration changes that impact the namespace include adding, removing, or renaming a volume; and changing the volume permissions, owner, or group ID.

Note – Never import or deport VxVM disk groups using VxVM commands once the disk group has been registered with the cluster as a Sun Cluster disk device group. The Sun Cluster software will handle all cases where disk groups need to be imported or deported.

Each VxVM disk group must have a cluster-wide unique minor number. By default, when a disk group is created, VxVM chooses a random number that is a multiple of 1000 as that disk group's base minor number. For most configurations with only a small number of disk groups, this is sufficient to guarantee uniqueness. However, it is possible that the minor number for a newly-created disk group will conflict with the minor number of a pre-existing disk group imported on a different cluster node. In this case, attempting to register the Sun Cluster disk device group will fail. To fix this problem, the new disk group should be given a new minor number that is a unique value and then registered as a Sun Cluster disk device group.

If you are setting up a mirrored volume, Dirty Region Logging (DRL) can be used to decrease volume recovery time in the event of a system crash. Use of DRL is strongly recommended, although it could decrease I/O throughput.

Administering Cluster File Systems Overview

Use standard Solaris file system commands, such as `mount`, `newfs`, and so on, to administer a cluster file system. Mount cluster file systems by specifying the `-g` option to the `mount` command. Cluster file systems can also be automatically mounted at boot.

Note – No special Sun Cluster commands are necessary for cluster file system administration. Administer a cluster file system as you would any other Solaris file system.

Administering Disk Device Groups

The `scsetup(1M)` utility is an interactive interface to the `scconf(1M)` command. When `scsetup` runs, it generates `scconf` commands. These generated commands are shown in the examples at the end of some procedures.

TABLE 3-1   Task List: Administering Disk Device Groups

Task	For Instructions, Go To...
Update the global device namespace (without a reconfiguration reboot) - Use `scgdevs`	"How to Update the Global Device Namespace" on page 38
Add Solstice DiskSuite disksets and register them as disk device groups - Use `metaset`	"How to Add and Register a Disk Device Group (Solstice DiskSuite)" on page 38
Remove Solstice DiskSuite disk device groups from the configuration - Use `metaset` and `metaclear`	"How to Remove and Unregister a Disk Device Group (Solstice DiskSuite)" on page 39
Remove a node from a Solstice DiskSuite disk device group - Use `metaset`	"How to Remove a Node From a Disk Device Group (Solstice DiskSuite)" on page 40

TABLE 3-1 Task List: Administering Disk Device Groups *(Continued)*

Task	For Instructions, Go To…
Add VERITAS Volume Manager disk groups as disk device groups - Use VxVM commands and `scsetup`	"How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)" on page 41 "How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)" on page 42 "How to Add a New Volume to an Existing Disk Device Group (VERITAS Volume Manager)" on page 43 "How to Make an Existing Disk Group Into a Disk Device Group (VERITAS Volume Manager)" on page 44 "How to Assign a New Minor Number to a Disk Device Group (VERITAS Volume Manager)" on page 45 "How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)" on page 46 "How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 48
Remove VERITAS Volume Manager disk device groups from the configuration - Use `scsetup` (to generate scconf)	"How to Remove a Volume From a Disk Device Group (VERITAS Volume Manager)" on page 49 "How to Remove and Unregister a Disk Device Group (VERITAS Volume Manager)" on page 50
Add a node to a VERITAS Volume Manager disk device group - Use `scsetup` (to generate scconf)	"How to Add a Node to a Disk Device Group (VERITAS Volume Manager)" on page 51
Remove a node from a VERITAS Volume Manager disk device group - Use scsetup (to generate scconf)	"How to Remove a Node From a Disk Device Group (VERITAS Volume Manager)" on page 53
Change disk device group properties - Use scsetup (to generate scconf)	"How to Change Disk Device Properties" on page 55
Display disk device groups and properties - Use scconf	"How to List a Disk Device Group Configuration" on page 56
Switch the primary for a disk device group - Use `scswitch`	"How to Switch the Primary for a Device Group" on page 57
Put a disk device group in maintenance state - Use `metaset` or `vxdg`	"How to Put a Disk Device Group in Maintenance State" on page 59

▼ How to Update the Global Device Namespace

When adding a new global device, manually update the global device namespace by running `scgdevs(1M)`.

Note – The `scgdevs` command does not have any effect if the node running the command is not currently a cluster member or if the `/global/.devices/node@`*nodeID* file system is not mounted.

---

1. **Become superuser on any node of the cluster.**

2. **Use the** `scgdevs` **command to reconfigure the namespace.**

```
# scgdevs
```

Example—Updating the Global Device Namespace

The following example shows output generated by a successful run of `scgdevs`.

```
# scgdevs
Configuring the /dev/global directory (global devices)...
obtaining access to all attached disks
reservation program successfully exiting
```

▼ How to Add and Register a Disk Device Group (Solstice DiskSuite)

Use the `metaset(1M)` command to create a Solstice DiskSuite diskset and register it as a Sun Cluster disk device group. When you register the diskset, the name you assigned to the diskset will automatically be assigned to the disk device group.

1. **Become superuser on the node connected to the disks where you want to create the diskset.**

2. **Use the** `metaset` **command to add the Solstice DiskSuite diskset and register it as a disk device group with Sun Cluster.**

```
# metaset -s diskset -a -h nodelist
```

| -s *diskset* | Specifies the diskset to be created. |
| -a -h *nodelist* | Adds the list of nodes that can master the diskset. |

3. **Verify that the disk device group has been added.**

The disk device group name will match the diskset name specified with metaset.

```
# scconf -p | grep disk-device-group
```

Example—Adding a Solstice DiskSuite Disk Device Group

The following example shows the creation of the diskset and disk device group and verifies that the disk device group has been created.

```
# metaset -s dg-schost-1
# scconf -p | grep dg-schost-1
Device group name: dg-schost-1
```

▼ How to Remove and Unregister a Disk Device Group (Solstice DiskSuite)

Disk device groups are Solstice DiskSuite disksets that have been registered with Sun Cluster. To remove a disk device group, use the metaclear(1M) and metaset(1M) commands. These commands unregister and remove the disk device group from the cluster configuration, and return the disk device group to a non-cluster diskset.

Refer to the Solstice DiskSuite documentation for the steps to remove a diskset.

▼ How to Remove a Node From a Disk Device Group (Solstice DiskSuite)

Use this procedure to remove a cluster node from the list of potential primaries of a disk device group, on a cluster running Solstice DiskSuite. A node can belong to more than one disk device group at a time, so repeat the `metaset` command for each disk device group from which you want to remove the node.

1. Determine the disk device group(s) of which the node to be removed is a member.

```
# scstat -D
```

2. Become superuser on the node that currently owns the disk device group you want to modify.

3. Delete the node's hostname from the disk device group.

```
# metaset -s setname -d -f -h nodelist
```

-s *setname*	Specifies the disk device group name
-d	Deletes from the disk device group the nodes identified with -h
-f	Force
-h *nodelist*	Removes the node from the list of nodes that can master the disk device group

---

Note – The update can take several minutes to complete.

---

4. Repeat Step 3 for each disk device group from which the node is being removed as a potential primary.

5. Verify that the node has been removed from the disk device group.

The disk device group name will match the diskset name specified with `metaset`.

```
# scstat -D
```

Example—Removing a Node From a Disk Device Group (SDS)

The following example shows the removal of the host name `phys-schost-2` from a disk device group configuration. This eliminates `phys-schost-2` as a potential primary for the designated disk device group. Verify removal of the node by running the `scstat -D` command and checking that the removed node is no longer displayed in the screen text.

```
[Determine the disk device group(s) for the node:]
# scstat -D
  -- Device Group Servers --
                       Device Group   Primary        Secondary
                       ------------   -------        ---------
  Device group servers: dg-schost-1  phys-schost-1  phys-schost-2

[Become superuser.]

[Remove the hostname from all disk device groups:]
# metaset -s dg-schost-1 -d -f -h phys-schost-2

[Verify removal of the node:]
# scstat -D
  -- Device Group Servers --
                       Device Group   Primary        Secondary
                       ------------   -------        ---------
  Device group servers: dg-schost-1  phys-schost-1  -
```

▼ How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)

> **Note –** This procedure is only for initializing disks. If you are encapsulating disks, use the procedure "How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)" on page 42.

After adding the VxVM disk group, you need to register the disk device group.

1. **Become superuser on any node of the cluster that is *physically connected* to the disks that make up the disk group being added.**

2. **Create the VxVM disk group and volume.**

   Use your preferred method to create the disk group and volume.

> **Note –** If you are setting up a mirrored volume, use Dirty Region Logging (DRL) to decrease volume recovery time in the event of a node failure. However, DRL might decrease I/O throughput.

See the VERITAS Volume Manager documentation for the procedures to complete this step.

If you are installing Oracle Parallel Server, create shared VxVM disk groups using the cluster functionality of VxVM as described in the VERITAS Volume Manager Administrator's Reference Guide. Otherwise, create VxVM disk groups using the standard procedures documented in the VERITAS documentation.

3. **Register the VxVM disk group as a Sun Cluster disk device group.**

   See "How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)" on page 46.

▼ How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)

> **Note –** This procedure is only for encapsulating disks. If you are initializing disks, use the procedure "How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)" on page 41.

You can make non-root disks into Sun Cluster disk device groups by first encapsulating them as VxVM disk groups, then registering them as Sun Cluster disk device groups.

Disk encapsulation is only supported during initial creation of a VxVM disk group. Once a VxVM disk group is created and registered as a Sun Cluster disk device group, only disks which can be initialized should be added to the disk group.

1. **Become superuser on any node of the cluster.**

2. **If the disk being encapsulated has file system entries in the** `/etc/vfstab` **file, make sure that the** `mount at boot` **option is set to** `no`**.**

   This can be set back to `yes` once the disk has been encapsulated and registered as a Sun Cluster disk device group.

3. **Encapsulate the disks.**

   Use vxdiskadm menus or the graphical user interface to encapsulate the disks. VxVM requires two free partitions as well as unassigned cylinders at the beginning or the end of the disk. Slice 2 must also be set to the entire disk. See the vxdiskadm(1M) man page for more information.

4. **Shut down and restart the node.**

   The scswitch(1M) command switches all resource groups and device groups from the primary node to the next preferred node. Then shutdown(1M) is used to shut down and restart the node.

   ```
   # scswitch -S -h nodelist
   # shutdown -g0 -y -i6
   ```

5. **If necessary, switch all resource groups and device groups back to the original node.**

   If the resource groups and device groups were initially configured to fail back to the primary node, this step is not necessary.

   ```
   # scswitch -z -h nodelist -D disk-device-group
   # scswitch -z -h nodelist -g resource-group
   ```

6. **Register the VxVM disk group as a Sun Cluster disk device group.**

   See "How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)" on page 46.

▼ How to Add a New Volume to an Existing Disk Device Group (VERITAS Volume Manager)

> **Note –** After adding the volume, you need to register the configuration change by using the procedure "How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 48.

When you add a new volume to an existing VxVM disk device group, you need to perform the procedure from the primary node for the disk device group, and the disk device group must be online.

1. **Become superuser on any node of the cluster.**

2. **Determine the primary node and status for the disk device group to which you are adding the new volume.**

```
# scstat -D
```

3. **If the disk device group is offline, bring it online.**

```
# scswitch -z -D disk-device-group -h nodelist
```

-z -D *disk-device-group*	Switches the specified device group.
-h *nodelist*	Specifies the name of the node to switch the disk device group to. This node becomes the new primary.

4. **From the primary node (the node currently mastering the disk device group), create the VxVM volume in the disk group.**

   Refer to your VERITAS Volume Manager documentation for the procedure used to create the VxVM volume.

5. **Register the VxVM disk group changes so the global namespace gets updated.**

   See "How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 48.

▼ How to Make an Existing Disk Group Into a Disk Device Group (VERITAS Volume Manager)

You can make an existing VxVM disk group into a Sun Cluster disk device group by first importing the disk group onto the current node, then registering the disk group as a Sun Cluster disk device group.

1. **Become superuser on any node of the cluster.**

2. **Import the VxVM disk group onto the current node.**

```
# vxdg import diskgroup
```

3. **Register the VxVM disk group as a Sun Cluster disk device group.**

   See "How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)" on page 46.

▼ How to Assign a New Minor Number to a Disk Device Group (VERITAS Volume Manager)

If disk device group registration fails because of a minor number conflict with another disk group, you must assign the new disk group a new, unused minor number. After assigning the new minor number, rerun the procedure to register the disk group as a Sun Cluster disk device group.

1. Become superuser on any node of the cluster.

2. Determine the minor numbers in use.

```
# ls -l /global/.devices/node@nodeid/dev/vx/dsk/*
```

3. Choose any other multiple of 1000 that is not in use as the base minor number for the new disk group.

4. Assign the new minor number to the disk group.

```
# vxdg reminor diskgroup base-minor-number
```

5. Register the VxVM disk group as a Sun Cluster disk device group.

See "How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)" on page 46.

Example—How to Assign a New Minor Number to a Disk Device Group

This example uses the minor numbers 16000-16002 and 4000-4001. The vxdg reminor command is used to assign the base minor number 5000 to the new disk device group.

```
# ls -l /global/.devices/node@nodeid/dev/vx/dsk/*
/global/.devices/node@nodeid/dev/vx/dsk/dg1
brw-------   1 root     root      56,16000 Oct   7 11:32 dg1v1
brw-------   1 root     root      56,16001 Oct   7 11:32 dg1v2
brw-------   1 root     root      56,16002 Oct   7 11:32 dg1v3

/global/.devices/node@nodeid/dev/vx/dsk/dg2
brw-------   1 root     root      56,4000 Oct   7 11:32 dg2v1
brw-------   1 root     root      56,4001 Oct   7 11:32 dg2v2
# vxdg reminor dg3 5000
```

▼ How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)

This procedure uses the scsetup(1M) utility to register the associated VxVM disk group as a Sun Cluster disk device group.

Note – Once a disk device group has been registered with the cluster, never import or deport a VxVM disk group using VxVM commands. If you make a change to the VxVM disk group or volume, use the procedure "How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 48 to register the disk device group configuration changes. This will ensure that the global namespace is in the correct state.

The prerequisites to register a VxVM disk device group are:

- Superuser privilege on a node in the cluster.
- The name of the VxVM disk group to be registered as a disk device group.
- A preferred order of nodes to master the disk device group.

When you define the preference order, you also specify whether you want the disk device group to be switched back to the most preferred node in the event that the most preferred node goes down and later returns to the cluster.

See scconf(1M) for more information on node preference and failback options.

1. **Become superuser on any node of the cluster.**

2. **Enter the** scsetup **utility.**

   ```
   # scsetup
   ```

   The Main Menu is displayed.

3. **To work with VxVM disk device groups, type** 4 **(Device groups and volumes).**
   The Device Groups Menu is displayed.

4. **To register a VxVM disk device group, type** 1 **(Register a VxVM disk group as a device group).**
   Follow the instructions and enter the name of the VxVM disk group to be registered as a Sun Cluster disk device group.

5. **If you encounter the following error while attempting to register the disk device group, reminor the disk device group.**

```
scconf: Failed to add device group - in use
```

To reminor the disk device group, use the procedure "How to Assign a New Minor Number to a Disk Device Group (VERITAS Volume Manager)" on page 45. This procedure enables you to assign a new minor number that does not conflict with a minor number used by an existing disk device group.

6. **Verify that the disk device group is registered and online.**

If the disk device group is properly registered, information for the new disk device group displays when using the following command.

```
# scstat -D
```

Note – If you change any configuration information for a VxVM disk group or volume that is registered with the cluster, you must reregister the disk device group by using scsetup. Such configuration changes include adding or removing volumes, as well as changing the group, owner, or permissions of existing volumes. Reregistration after configuration changes ensures that the global namespace is in the correct state. See "How to Update the Global Device Namespace" on page 38.

Example—Registering a VERITAS Volume Manager Disk Device Group

The following example shows the `scconf` command generated by `scsetup` when it registers a VxVM disk device group (`dg1`), and the verification step. This example assumes that the VxVM disk group and volume were created previously.

```
# scsetup

scconf -a -D type=vxvm,name=dg1,nodelist=phys-schost-1:phys-schost-2

# scstat -D
-- Device Group Servers --
                        Device Group      Primary         Secondary
                        ------------      -------         ---------
Device group servers:   dg1               phys-schost-1   phys-schost-2

-- Device Group Status --
                          Device Group      Status
                          ------------      ------
  Device group status:    dg1               Online
```

Where to Go From Here

To create a cluster file system on the VxVM disk device group, see "How to Add a Cluster File System" on page 61.

If there are problems with the minor number, see "How to Assign a New Minor Number to a Disk Device Group (VERITAS Volume Manager)" on page 45.

▼ How to Register Disk Group Configuration Changes (VERITAS Volume Manager)

When you change any configuration information for a VxVM disk group or volume, you need to register the configuration changes for the Sun Cluster disk device group. This ensures that the global namespace is in the correct state.

1. **Become superuser on any node in the cluster.**

2. **Enter the** `scsetup`**(1M) utility.**

```
# scsetup
```

The Main Menu is displayed.

3. **To work with VxVM disk device groups, type** 4 **(Device groups and volumes).**

The Device Groups Menu is displayed.

4. **To register configuration changes, type** 2 **(Synchronize volume information for a VxVM device group).**

Follow the instructions and enter the VxVM disk group that has changed configuration.

Example—Registering VERITAS Volume Manager Disk Group Configuration Changes

The following example shows the `scconf` command generated by `scsetup` when it registers a changed VxVM disk device group (`dg1`). This example assumes that the VxVM disk group and volume were created previously.

```
# scsetup

scconf -c -D name=dg1,sync
```

▼ How to Remove a Volume From a Disk Device Group (VERITAS Volume Manager)

> **Note –** After removing the volume from the disk device group, you must register the configuration changes to the disk device group using the procedure "How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 48.

1. **Become superuser on any node of the cluster.**

2. **Determine the primary node and status for the disk device group.**

```
# scstat -D
```

3. **If the disk device group is offline, bring it online.**

```
# scswitch -z -D disk-device-group -h nodelist
```

-z	Performs the switch.
-D *disk-device-group*	Specifies the device group to switch.
-h *nodelist*	Specifies the name of the node to switch to. This node becomes the new primary.

4. **From the primary node (the node currently mastering the disk device group), remove the VxVM volume in the disk group.**

```
# vxedit -g diskgroup -rf rm volume
```

-g *diskgroup*	Specifies the VxVM disk group containing the volume.
-rf rm *volume*	Removes the specified volume.

5. **Register the disk device group configuration changes to update the global namespace, using** scsetup.

See "How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 48.

▼ How to Remove and Unregister a Disk Device Group (VERITAS Volume Manager)

Removing a Sun Cluster disk device group will cause the corresponding VxVM disk group to be deported, not destroyed. However, even though the VxVM disk group still exists, it cannot be used in the cluster unless re-registered.

This procedure uses the scsetup(1M) utility to remove a VxVM disk group and unregister it as a Sun Cluster disk device group.

1. **Become superuser on any node of the cluster.**

2. **Take the disk device group offline.**

```
# scswitch -F -D disk-device-group
```

−F	Places the disk device group offline.
−D *disk-device-group*	Specifies the device group to take offline.

3. **Enter the** scsetup **utility.**

The Main Menu is displayed.

```
# scsetup
```

4. **To work with VxVM device groups, type** 4 **(Device groups and volumes).**

The Device Groups Menu is displayed.

5. **To unregister a VxVM disk group, type** 3 **(Unregister a VxVM device group).**

Follow the instructions and enter the VxVM disk group to be unregistered.

Example—Removing and Unregistering a VERITAS Volume Manager Disk Device Group

The following example shows the VxVM disk device group dg1 taken offline, and the scconf(1M) command generated by scsetup when it removes and unregisters the disk device group.

```
# scswitch -F -D dg1
# scsetup

  scconf -r -D name=dg1
```

▼ How to Add a Node to a Disk Device Group (VERITAS Volume Manager)

This procedure adds a node to a disk device group using the scsetup(1M) utility.

The prerequisites to add a node to a VxVM disk device group are:

- Superuser privilege on a node in the cluster
- The name of the VxVM device group to which the node will be added
- The name or node ID of the nodes to add

1. **Become superuser on any node of the cluster.**

2. **Enter the** `scsetup(1M)` **utility**

   The Main Menu is displayed.

   ```
   # scsetup
   ```

3. **To work with VxVM disk device groups, type** 4 **(Device groups and volumes).**

   The Device Groups Menu is displayed.

4. **To add a node to a VxVM disk device group, type** 4 **(Add a node to a VxVM device group).**

   Follow the instructions and enter the device group and node names.

5. **Verify that the node has been added.**

   Look for the device group information for the new disk displayed by the following command.

   ```
   # scconf -p
   ```

Example—Adding a Node to a VERITAS Volume Manager Disk Device Group

The following example shows the `scconf` command generated by `scsetup` when it adds a node (`phys-schost-3`) to a VxVM disk device group (`dg1`), and the verification step.

```
# scsetup

scconf a D type=vxvm,name=dg1,nodelist=phys-schost-3

# scconf -p
Device group name:                           dg1
   Device group type:                        VXVM
   Device group failback enabled:            yes
   Device group node list:                   phys-schost-1, phys-schost-3
```

▼ How to Remove a Node From a Disk Device Group (VERITAS Volume Manager)

Use this procedure to remove a cluster node from an existing cluster disk device group (disk group) running VERITAS Volume Manager (VxVM).

1. **Determine the disk device group of which the node to be removed is a member.**

```
# scstat -D
```

2. **Become superuser on a current cluster member node.**

3. **Execute the** scsetup **utility.**

```
# scsetup
```

The Main Menu is displayed.

4. **To reconfigure a disk device group, type** 4 **(Device groups and volumes).**

5. **To remove the node from the VxVM disk device group, type** 5 **(Remove a node from a VxVM device group).**

Follow the prompts to remove the cluster node from the disk device group. You will be asked for information about the following:

VxVM device group

Node name

6. **Verify that the node has been removed from the VxVM disk device group:**

```
# scconf -p | grep Device
```

Example—Removing a Node From a Disk Device Group (VxVM)

This example shows removal of the node named `phys-schost-4` from the `dg1` VxVM disk device group.

```
[Determine the disk device group for the node:]
# scstat -D
  -- Device Group Servers --
                      Device Group  Primary         Secondary
                      ------------  -------         ---------
  Device group servers: dg-schost-1  phys-schost-1  phys-schost-2

[Become superuser and execute the scsetup utility:]
# scsetup

 Select Device groups and volumes>Remove a node from a VxVM device group.

Answer the questions when prompted.
You will need the following information.
  You Will Need:            Example:
  VxVM device group name    dg1
  node names                phys-schost-1

[Verify that the scconf command executed properly:]

scconf -r -D name=dg1,nodelist=phys-schost-4

    Command completed successfully.

Quit the scsetup Device Groups Menu and Main Menu.

[Verify that the node was removed:]
# scconf -p | grep Device
  Device group name:              dg1
    Device group type:            VxVM
    Device group failback enabled: no
    Device group node list:       phys-schost-3
    Device group diskset name:    dg1
```

▼ How to Change Disk Device Properties

The method for establishing the primary ownership of a disk device group is based on the setting of an ownership preference attribute called `preferenced`. If the attribute is not set, the primary owner of an otherwise unowned disk device group is the first node that attempts to access a disk in that group. However, if this attribute is set, you must specify the preferred order in which nodes attempt to establish ownership.

If you disable the `preferenced` attribute, then the `failback` attribute is also automatically disabled. However, if you attempt to enable or re-enable the `preferenced` attribute, you have the choice of enabling or disabling the `failback` attribute.

If the `preferenced` attribute is either enabled or re-enabled, you are required to re-establish the order of nodes in the primary ownership preference list.

This procedure uses `scsetup(1M)` to set or unset the `preferenced` attribute and the `failback` attribute for Solstice DiskSuite or VxVM disk device groups.

To run this procedure, you need the name of the disk device group for which you are changing attribute values.

1. **Become superuser on any node of the cluster.**

2. **Enter the** `scsetup(1M)` **utility**

   The Main Menu is displayed.

   ```
   # scsetup
   ```

3. **To work with disk device groups, type** 4 **(Device groups and volumes).**

   The Device Groups Menu is displayed.

4. **To change a device group property, type** 6 **(Change key properties of a VxVM or Solstice DiskSuite device group).**

   Follow the instructions to set the `preferenced` and `failback` options for a device group.

5. **Verify that the disk device group attributes have been changed.**

   Look for the device group information displayed by the following command.

   ```
   # scconf -p
   ```

Example—Changing Disk Device Group Properties

The following example shows the `scconf` command generated by `scsetup` when it sets the attribute values for a disk device group (`dg-schost-1`).

```
# scconf -c -D name=dg-schost-1,nodelist=phys-schost-1:phys-schost-2,\
preferenced=true,failback=enabled

# scconf -p | grep Device
Device group name:                          dg-schost-1
   Device group type:                       SDS
   Device group failback enabled:           yes
   Device group node list:                  phys-schost-1, phys-schost-2
   Device group ordered node list:          yes
   Device group diskset name:               dg-schost-1
```

▼ How to List a Disk Device Group Configuration

You do not need to be superuser to list the configuration.

There are three ways you can list disk device group configuration information.

● **Use the SunPlex Manager GUI.**

See the SunPlex Manager online help for more information.

● **Use** `scstat`**(1M) to list the disk device group configuration.**

```
% scstat -D
```

● **Use** `scconf`**(1M) to list the disk device group configuration.**

```
% scconf -p
```

Example—Listing the Disk Device Group Configuration By Using `scstat`

Using the `scstat -D` command displays the following information.

```
-- Device Group Servers --
                     Device Group       Primary              Secondary
                     ------------       -------              ---------
  Device group servers:  schost-2       -                    -
  Device group servers:  schost-1       phys-schost-2        phys-schost-3
  Device group servers:  schost-3       -                    -
-- Device Group Status --
                         Device Group       Status
                         ------------       ------
  Device group status:      schost-2        Offline
  Device group status:      schost-1        Online
  Device group status:      schost-3        Offline
```

Example—Listing the Disk Device Group Configuration By Using `scconf`

When using the `scconf` command, look for the information listed under device groups.

```
# scconf -p
...
Device group name: dg-schost-1
    Device group type:            SDS
    Device group failback enabled:  yes
    Device group node list:         phys-schost-2, phys-schost-3
    Device group diskset name:      dg-schost-1
```

▼ How to Switch the Primary for a Device Group

This procedure can also be used to start (bring online) an inactive device group.

You can also bring an inactive device group online, or switch the primary for a device group, by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

1. **Become superuser on any node of the cluster.**

2. **Use** scswitch(1M) **to switch the disk device group primary.**

```
# scswitch -z -D disk-device-group -h nodelist
```

-z	Performs the switch.
-D *disk-device-group*	Specifies the device group to switch.
-h *nodelist*	Specifies the name of the node to switch to. This node become the new primary.

3. **Verify that the disk device group has been switched to the new primary.**

If the disk device group is properly registered, information for the new disk device group displays when using the following command.

```
# scstat -D
```

Example—Switching the Primary for a Disk Device Group

The following example shows how to switch the primary for a disk device group and verify the change.

```
# scswitch -z -D dg-schost-1 -h phys-schost-1
# scstat -D

-- Device Group Servers --

                        Device Group        Primary              Secondary
                        ------------        -------              ---------
Device group servers:   dg1                 phys-schost-1        phys-schost-2

-- Device Group Status --

                        Device Group        Status
                        ------------        ------
  Device group status:  dg1                 Online
```

▼ How to Put a Disk Device Group in Maintenance State

Putting a device group in maintenance state prevents that device group from automatically being brought online whenever one of its devices is accessed. You should put a device group in maintenance state when completing repair procedures that require that all I/O activity be quiesced until completion of the repair. Putting a device group in maintenance state also helps prevent data lost by ensuring that a disk device group is not brought online on one node while the diskset or disk group is being repaired on another node.

1. **Place the device group in maintenance state.**

   ```
   # scswitch -m -D disk-device-group
   ```

2. **If the repair procedure being performed requires ownership of a diskset or disk group, manually import that diskset or disk group.**

   ■ For Solstice DiskSuite:

   ```
   # metaset -C take -f -s diskset
   ```

> ⚠ **Caution –** If you are taking ownership of an SDS diskset, the `metaset -C take` command *must* be used when the device group is in maintenance state. Using `metaset -t` will bring the device group online as part of taking ownership. If you are importing a VxVM disk group, the `-t` flag must be used when importing the disk group. This prevents the disk group from automatically being imported if this node is rebooted.

   ■ For VERITAS Volume Manager:

   ```
   # vxdg -t import disk-group-name
   ```

3. **Complete whatever repair procedure you need to perform.**

4. **Release ownership of the diskset or disk group.**

> ⚠ **Caution –** Before taking the disk device group out of maintenance state, you must release ownership of the diskset or disk group. Failure to do so may result in data loss.

- For Solstice DiskSuite:

```
# metaset -C release -s diskset
```

- For VERITAS Volume Manager:

```
# vxdg deport disk-group-name
```

5. **Bring the disk device group online.**

```
# scswitch -z -D disk-device-group -h nodelist
```

Example—Putting a Disk Device Group in Maintenance State

This example shows how to put disk device group dg-schost-1 into maintenance state, and remove the disk device group from maintenance state.

```
[Place the disk device group in maintenance state.]
# scswitch -m -D dg-schost-1


[If needed, manually import the diskset or disk group.]
For Solstice DiskSuite:
  # metaset -C take -f -s dg-schost-1
For VERITAS Volume Manager:
  # vxdg -t import dg1


[Complete all necessary repair procedures.]


[Release ownership.]
For Solstice DiskSuite:
  # metaset -C release -s dg-schost-1
For VERITAS Volume Manager:
  # vxdg deport dg1


[Bring the disk device group online.]
# scswitch -z -D dg-schost-1 -h phys-schost-1
```

Administering Cluster File Systems

TABLE 3-2   Task Map: Administering Cluster File Systems

Task	For Instructions, Go To...
Add cluster file systems after the initial Sun Cluster installation - Use `newfs` and `makedir`	"How to Add a Cluster File System" on page 61
Remove a cluster file system - Use `fuser` and `umount`	"How to Remove a Cluster File System" on page 65
Check global mount points in a cluster for consistency across nodes - Use `sccheck`	"How to Check Global Mounts in a Cluster" on page 68

▼ How to Add a Cluster File System

Perform this task for each cluster file system you create after your initial Sun Cluster installation.

Caution – Be sure you specify the correct disk device name. Creating a cluster file system destroys any data on the disks. If you specify the wrong device name, you will erase data that you may not intend to delete.

The prerequisites to add an additional cluster file system are:

- Superuser privilege on a node in the cluster.
- A device group (Solstice DiskSuite device group or a VxVM device group) or block disk slice upon which to create the cluster file system.

1. Become superuser on any node in the cluster.

Tip – For faster file system creation, become superuser on the current primary of the global device for which you are creating a file system.

2. **Create a file system using the** `newfs`**(1M) command.**

```
# newfs raw-disk-device
```

The following table shows examples of names for the *raw-disk-device* argument. Note that naming conventions differ for each volume manager.

TABLE 3-3   Sample Raw Disk Device Names

If Your Volume Manager Is …	A Disk Device Name Might Be …	Description
Solstice DiskSuite	/dev/md/oracle/rdsk/d1	Raw disk device d1 within the oracle diskset.
VERITAS Volume Manager	/dev/vx/rdsk/oradg/vol01	Raw disk device vol01 within the oradg disk group.
None	/dev/global/rdsk/d1s3	Raw disk device for block slice d1s3.

3. **On each node in the cluster, create a mount point directory for the cluster file system.**

A mount point is required *on each node*, even if the cluster file system will not be accessed on that node.

```
# mkdir -p /global/device-group/mount-point
```

device-group   Name of the directory that corresponds to the name of the device group that contains the device.

mount-point   Name of the directory on which to mount the cluster file system.

Tip – For ease of administration, create the mount point in the /global/*device-group* directory. Using this location enables you to easily distinguish cluster file systems, which are globally available, from local file systems.

---

4. **On each node in the cluster, add an entry to the** /etc/vfstab **file for the mount point.**

   a. **Use the following required mount options.**

> **Note –** Logging is required for all cluster file systems.

- **Solaris UFS logging –** Use the `global,logging` mount options. See the `mount_ufs`(1M) man page for more information about UFS mount options.

> **Note –** The `syncdir` mount option is not required for UFS cluster file systems. If you specify `syncdir`, you are guaranteed POSIX-compliant file system behavior. If you do not, you will have the same behavior that is seen with UFS file systems. When you do not specify `syncdir`, performance of writes that allocate disk blocks, such as when appending data to a file, can significantly improve. However, in some cases, without `syncdir` you would not discover an out-of-space condition until you close a file. The cases in which you could have problems if you do not specify `syncdir` are rare. With `syncdir` (and POSIX behavior), the out-of-space condition would be discovered before the close.

- **Solstice DiskSuite trans metadevice –** Use the `global` mount option (do not use the `logging` mount option). See your Solstice DiskSuite documentation for information about setting up trans metadevices.

b. **To automatically mount the cluster file system, set the** `mount at boot` **field to** `yes`**.**

c. **Ensure that, for each cluster file system, the information in its** `/etc/vfstab` **entry is identical on each node.**

d. **Ensure that the entries in each node's** `/etc/vfstab` **file list devices in the same order.**

e. **Check the boot order dependencies of the file systems.**

For example, consider the scenario where `phys-schost-1` mounts disk device `d0` on `/global/oracle`, and `phys-schost-2` mounts disk device `d1` on `/global/oracle/logs`. With this configuration, `phys-schost-2` can boot and mount `/global/oracle/logs` only after `phys-schost-1` boots and mounts `/global/oracle`.

See the `vfstab`(4) man page for details.

5. **On any node in the cluster, verify that mount points exist and** `/etc/vfstab` **file entries are correct on all nodes of the cluster.**

```
# sccheck
```

If there are no errors, nothing is returned.

6. **From any node in the cluster, mount the cluster file system.**

```
# mount /global/device-group/mount-point
```

7. **On each node of the cluster, verify that the cluster file system is mounted.**

You can use either the df(1M) or mount(1M) command to list mounted file systems.

Example—Adding a Cluster File System

The following example creates a UFS cluster file system on the Solstice DiskSuite metadevice /dev/md/oracle/rdsk/d1.

```
# newfs /dev/md/oracle/rdsk/d1
...

[on each node:]
# mkdir -p /global/oracle/d1

# vi /etc/vfstab
#device                device                      mount           FS   fsck  moun
t         mount
#to mount             to fsck                      point           type pass  at
boot      options
#
/dev/md/oracle/dsk/d1 /dev/md/oracle/rdsk/d1 /global/oracle/d1 ufs  2     yes
        global,logging
[save and exit]

[on one node:]
# sccheck
# mount /global/oracle/d1
# mount
...
/global/oracle/d1 on /dev/md/oracle/dsk/d1 read/write/setuid/global/logging/
largefiles on Sun May 3 08:56:16 2001
```

▼ How to Remove a Cluster File System

You 'remove' a cluster file system by merely unmounting it. If you want to also remove or delete the data, remove the underlying disk device (or metadevice or volume) from the system.

Note – Cluster file systems are automatically unmounted as part of the system shutdown that occurs when you run scshutdown(1M) to stop the entire cluster. A cluster file system is not unmounted when you run shutdown to stop a single node. However, if the node being shut down is the only node with a connection to the disk, any attempt to access the cluster file system on that disk results in an error.

The prerequisites to unmount cluster file systems are:

- Superuser privilege on a node in the cluster.
- The file system cannot be busy. A file system is considered busy if a user is in a directory in the file system, or if a program has a file open in that file system. The user or program could be running on any node in the cluster.

1. **Become superuser on any node in the cluster.**

2. **Determine which cluster file systems are mounted.**

```
# mount -v
```

3. **On each node, list all processes that are using the cluster file system, so you know which processes you are going to stop.**

```
# fuser -c [ -u ] mount-point
```

-c	Reports on files that are mount points for file systems and any files within those mounted file systems.
-u	(Optional) Displays the user login name for each process ID.
mount-point	Specifies the name of the cluster file system for which you want to stop processes.

4. **On each node, stop all processes for the cluster file system.**

   Use your preferred method for stopping processes. If necessary, use the following command to force termination of processes associated with the cluster file system.

   ```
   # fuser -c -k mount-point
   ```

   A SIGKILL is sent to each process using the cluster file system.

5. **On each node, verify that no processes are using the file system.**

   ```
   # fuser -c mount-point
   ```

6. **From just one node, umount the file system.**

   ```
   # umount mount-point
   ```

   mount-point   Specifies the name of the cluster file system you want to unmount. This can be either the directory name where the cluster file system is mounted, or the device name path of the file system.

7. **(Optional) Edit the** /etc/vfstab **file to delete the entry for the cluster file system being removed.**

   Perform this step on each cluster node that has an entry for this cluster file system in its /etc/vfstab file.

8. **(Optional) Remove the disk device group/metadevice/plex.**

   See your volume manager documentation for more information.

Example—Removing a Cluster File System

The following example removes a UFS cluster file system mounted on the Solstice DiskSuite metadevice `/dev/md/oracle/rdsk/d1`.

```
# mount -v
...
/global/oracle/d1 on /dev/md/oracle/dsk/d1
read/write/setuid/global/logging/largefiles on Sun Oct  3 08:56:16 1999
# fuser -c /global/oracle/d1
/global/oracle/d1: 4006c
# fuser -c -k /global/oracle/d1
/global/oracle/d1: 4006c
# fuser -c /global/oracle/d1
/global/oracle/d1:
# umount /global/oracle/d1

(on each node, remove the highlighted entry:)
# vi /etc/vfstab
#device          device          mount   FS      fsck    mount   mount
#to mount        to fsck         point   type    pass    at boot options
#
/dev/md/oracle/dsk/d1 /dev/md/oracle/rdsk/d1 /global/oracle/d1 ufs 2 yes
global,logging
[Save and exit.]
```

Note – To remove the data on the cluster file system, remove the underlying device. See your volume manager documentation for more information.

▼ How to Check Global Mounts in a Cluster

The sccheck(1M) utility verifies the syntax of the entries for cluster file systems in the /etc/vfstab file. If there are no errors, nothing is returned.

Note – Run sccheck after making cluster configuration changes, such as removing a cluster file system, that have affected devices or volume management components.

---

1. **Become superuser on any node in the cluster.**

2. **Check the cluster global mounts.**

   ```
   # sccheck
   ```

Administering Quorum

This chapter provides the procedures for administering quorum within Sun Cluster.

This is a list of the procedures in this chapter.

- "How to Add a Quorum Device" on page 71
- "How to Remove a Quorum Device" on page 73
- "How to Remove the Last Quorum Device From a Cluster" on page 75
- "How to Replace a Quorum Device" on page 77
- "How to Put a Quorum Device Into Maintenance State" on page 77
- "How to Bring a Quorum Device Out of Maintenance State" on page 79
- "How to List the Quorum Configuration" on page 80

Most examples shown in this chapter are from a three-node cluster.

See the *Sun Cluster 3.0 U1 Concepts* document for a conceptual overview of quorum and quorum devices.

Administering Quorum Overview

You can use the `scconf(1M)` command to perform all quorum administrative procedures. In addition, you can accomplish some procedures by using the `scsetup(1M)` interactive utility or the SunPlex Manager GUI. Whenever possible, quorum procedures are described in this chapter using `scsetup`. The SunPlex Manager online help describes how to perform quorum procedures using the GUI.

If any quorum-related `scconf` command is interrupted or fails, the quorum configuration information can become inconsistent in the cluster configuration database. If this occurs, either rerun the command or run `scconf` with the `reset` option to reset the quorum configuration.

Note – The scsetup(1M) utility is an interactive interface to the scconf(1M) command. When scsetup runs, it generates scconf commands. These generated commands are shown in the examples at the end of the procedures.

There are two commands that can be used to view the quorum configuration: scstat -q option and scconf -p. Most of the verification steps in this chapter use scconf, but you can substitute scstat -q if you find its output more useful.

TABLE 4-1 Task List: Administering Quorum

Task	For Instructions, Go To...
Add a quorum device to a cluster - Use scsetup	"How to Add a Quorum Device" on page 71
Remove a quorum device from a cluster - Use scsetup (to generate scconf)	"How to Remove a Quorum Device" on page 73
Remove the last quorum device from a cluster - Use scsetup (to generate scconf)	"How to Remove the Last Quorum Device From a Cluster" on page 75
Replace a quorum device in a cluster - Use the add and remove procedures	"How to Replace a Quorum Device" on page 77
Put a quorum device into maintenance state (While in maintenance state, the quorum device does not participate in voting to establish the quorum.) - Use scsetup (to generate scconf)	"How to Put a Quorum Device Into Maintenance State" on page 77
Reset the quorum configuration to its default state - Use scsetup (to generate scconf)	"How to Bring a Quorum Device Out of Maintenance State" on page 79
List the quorum devices and vote counts - Use scconf	"How to List the Quorum Configuration" on page 80

▼ How to Add a Quorum Device

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

To complete this procedure, identify a disk drive by its device ID (DID), which is shared by the nodes. Use the scdidadm -L command to see the list of DID names. Refer to the scdidadm(1M) man page for additional information.

See the scsetup(1M) and scconf(1M) man pages for information on the commands used in this procedure.

1. **Become superuser on any node of the cluster.**

2. **Enter the scsetup utility.**

   ```
   # scsetup
   ```

   The Main Menu is displayed.

3. **To work with quorum devices, type 1 (Quorum).**

   The Quorum Menu is displayed.

4. **To add a quorum device, type 1 (Add a quorum disk).**

   Follow the instructions and type the name of the device to be used as the quorum device.

5. **Verify that the quorum device has been added.**

   ```
   # scstat -q
   ```

6. **Repeat Step 3 through Step 5 for each group of nodes that shares a storage enclosure.**

Example—Adding a Quorum Device

The following example shows the scconf command generated by scsetup when it adds a quorum device and a verification step.

```
Become superuser on any cluster node.

[Execute the scsetup utility:]
# scsetup

Select Quorum>Add a quorum disk.

Answer the questions when prompted.
[Verify that the scconf command completed successfully:]

scconf -a -q globaldev=d20

    Command completed successfully.

Quit the scsetup Quorum Menu and Main Menu.

[Verify the quorum device is removed:]
# scstat -q

-- Quorum Summary --

  Quorum votes possible: 4
  Quorum votes needed:   3
  Quorum votes present:  4

-- Quorum Votes by Node --

                  Node Name          Present Possible Status
                  ---------          ------- -------- ------
  Node votes:     phys-schost-1        1        1      Online
  Node votes:     phys-schost-2        1        1      Online

-- Quorum Votes by Device --

                 Device Name         Present Possible Status
                 -----------         ------- -------- ------
  Device votes:  /dev/did/rdsk/d3s2 1        1        Online
  Device votes:  /dev/did/rdsk/d4s2 1        1        Online
```

▼ How to Remove a Quorum Device

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

When a quorum disk is removed, it no longer participates in the voting to establish quorum. Note that all two-node clusters require that at least one quorum device be configured. If this is the last quorum device on a cluster, scconf(1M) will fail to remove the device from the configuration.

Note – If the device you intend to remove is the last quorum device in the cluster, see the procedure "How to Remove the Last Quorum Device From a Cluster" on page 75.

---

1. **Become superuser on any node in the cluster.**

2. **Determine the quorum device to be removed.**

   ```
   # scconf -pv | grep Quorum
   ```

3. **Execute the scsetup(1M) utility.**

   ```
   # scsetup
   ```

   The Main Menu is displayed.

4. **To work with the quorum device, type 1 (Quorum).**

5. **To remove the quorum device, type 2 (Remove a quorum disk).**
   Answer the questions displayed during the removal process.

6. **Quit scsetup.**

7. **Verify that the quorum device is removed.**

   ```
   # scstat -q
   ```

Example—Removing a Quorum Device

This example shows how to remove a quorum device from a cluster with two or more quorum devices configured.

```
Become superuser on any node and place the node to be removed in maintenance
state.

[Determine the quorum device to be removed:]
# scconf -pv | grep Quorum

[Execute the scsetup utility:]
# scsetup

Select Quorum>Remove a quorum disk.

Answer the questions when prompted.
[Verify that the scconf command completed successfully:]

scconf -r -q globaldev=d4

    Command completed successfully.

Quit the scsetup Quorum Menu and Main Menu.

[Verify the quorum device is removed:]
# scstat -q

-- Quorum Summary --

  Quorum votes possible:      3
  Quorum votes needed:        2
  Quorum votes present:       3

-- Quorum Votes by Node --

                   Node Name          Present Possible Status
                   ---------          ------- -------- ------
  Node votes:      phys-schost-1      1         1        Online
  Node votes:      phys-schost-2      1         1        Online

-- Quorum Votes by Device --

               Device Name          Present Possible Status
               -----------          ------- -------- ------
  Device votes: /dev/did/rdsk/d3s2  1        1        Online
```

▼ How to Remove the Last Quorum Device From a Cluster

If the device you intend to remove is not the last quorum device in the cluster, use the previous procedure, "How to Remove a Quorum Device" on page 73.

Note – All two-node clusters require at least one configured quorum device. If this is the last quorum device on a two-node cluster, the cluster must be placed into install mode before scconf(1M) allows you to remove the device from the configuration. This should only be done if a node is being removed from the cluster.

---

1. **Become superuser on any node in the cluster and place the node to be removed in maintenance state.**

   See "How to Put a Node Into Maintenance State" on page 114.

2. **Place the cluster in install mode.**

   ```
   # scconf -c -q installmode
   ```

3. **Remove the quorum device using the** scconf **command.**

   ```
   # scconf -r -q globaldev=device
   ```

4. **Verify that the quorum device has been removed.**

   ```
   # scstat -q
   ```

Example—Removing the Last Quorum Device

This example shows how to remove the last remaining quorum device in a cluster configuration.

```
[Become superuser on any node.]

[Place the cluster in install mode:]
# scconf -c -q installmode

[Remove the quorum device:]
# scconf -r -q globaldev=d3

[Verify the quorum device has been removed:]
# scstat -q

-- Quorum Summary --

  Quorum votes possible:      2
  Quorum votes needed:        2
  Quorum votes present:       2

-- Quorum Votes by Node --

                   Node Name          Present Possible Status
                   ---------          ------- -------- ------
  Node votes:      phys-schost-1      1       1        Online
  Node votes:      phys-schost-2      1       1        Online

-- Quorum Votes by Device --

                   Device Name        Present Possible Status
                   -----------        ------- -------- ------
```

▼ How to Replace a Quorum Device

1. **Configure a new quorum device on the storage enclosure that currently includes the disk to be replaced.**

   You need to first add a new quorum device to the configuration to take the place of the old device. See "How to Add a Quorum Device" on page 71 to add a new quorum device to the cluster.

2. **Remove the failed disk as a quorum device.**

   See "How to Remove a Quorum Device" on page 73 to remove the old quorum device from the configuration.

3. **Replace the failed disk.**

   Refer to the hardware procedures for your disk enclosure in the *Sun Cluster 3.0 U1 Hardware Guide*.

▼ How to Put a Quorum Device Into Maintenance State

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

Put a quorum device into maintenance state when taking the quorum device out of service for an extended period of time. This way, the disk device's quorum vote count is set to zero and does not contribute to the quorum count while the device is being serviced.

Use the scconf(1M) command to put a quorum device into maintenance state. The scsetup utility does not currently have this capability.

Note – All two-node clusters require at least one configured quorum device. If this is the last quorum device on a two-node cluster, scconf will fail to put the device into maintenance state.

To put a cluster node into maintenance state, see "How to Put a Node Into Maintenance State" on page 114.

1. **Become superuser on any node of the cluster.**

2. **Put the quorum device into the maintenance state.**

   ```
   # scconf -c -q globaldev=device,maintstate
   ```

`-c`	Specifies the change form of the `scconf` command.
`-q`	Manages the quorum options.
`globaldev=`*device*	Specifies the DID name of the disk device to change, for example, `d4`.
`maintstate`	Puts the shared quorum device into maintenance state.

3. **Verify that the quorum device is now in maintenance state.**

   The output for the device you placed in maintenance state should read zero for the Quorum Device Votes.

   ```
   # scconf -p | grep -i quorum
   ```

Example—Putting a Quorum Device Into Maintenance State

The following example shows how to put a quorum device into maintenance state and how to verify the results.

```
# scconf -c -q globaldev=d20,maintstate
# scconf -p | grep -i quorum
  Node quorum vote count:                        1
  Node quorum vote count:                        1
Quorum devices:                                  d20
Quorum device name:                              d20
  Quorum device votes:                           0
  Quorum device enabled:                         no
  Quorum device name:                            /dev/did/rdsk/d20s2
  Quorum device hosts (enabled):                 phys-schost-2 phys-schost-3
  Quorum device hosts (disabled):
```

Where to Go From Here

To re-enable the quorum device, see "How to Bring a Quorum Device Out of Maintenance State" on page 79.

To put a node into maintenance state, see "How to Put a Node Into Maintenance State" on page 114.

▼ How to Bring a Quorum Device Out of Maintenance State

After a quorum device has been in maintenance state, and is being brought back online, use the following procedure to reset the quorum vote count to the default. For cluster nodes, the default quorum count is one. For quorum devices, the default quorum count is *N*-1, where *N* is the number of nodes with non-zero vote counts that have ports to the quorum device.

You should run this procedure any time a quorum device has been put into maintenance state and you are bringing it out of maintenance state.

> **Caution –** If you do not specify either the globaldev or node options, the quorum count is reset for the entire cluster.

To bring a cluster node as well as its associated quorum devices out of maintenance state, see "How to Bring a Node Out of Maintenance State" on page 116.

1. **Become superuser on any node of the cluster.**

2. **Reset the quorum count.**

   ```
   # scconf -c -q globaldev=device,reset
   ```

   -c   Specifies the change form of the scconf command.

   -q   Manages the quorum options.

   globaldev=*device*   Specifies the DID name of the quorum device to reset, for example, d4.

   reset   The change flag that resets quorum.

3. **If you are resetting the quorum count because a node was in maintenance state, reboot the node.**

4. **Verify the quorum vote count.**

   ```
   # scconf -p | grep -i quorum
   ```

Example—Resetting the Quorum Vote Count (Quorum Device)

The following example resets the quorum count for a quorum device back to the default and verifies the result.

```
# scconf -c -q globaldev=d20,reset
# scconf -p | grep -i quorum
  Node quorum vote count:                               1
  Node quorum vote count:                               1
Quorum devices:                                         d20
Quorum device name:                                     d20
  Quorum device votes:                                  1
  Quorum device enabled:                                yes
  Quorum device name:                                   /dev/did/rdsk/d20s2
  Quorum device hosts (enabled):                        phys-schost-2 phys-schost-3
  Quorum device hosts (disabled):
```

▼ How to List the Quorum Configuration

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

You do not need to be superuser to list the quorum configuration.

● **Use** scconf(1M) **to list the quorum configuration.**

```
# scconf -p | grep -i quorum
```

Example—Listing the Quorum Configuration

```
# scconf -p | grep "Quorum | vote"
 Node quorum vote count:                               1
 Node quorum vote count:                               1
Quorum devices:                                        d20
Quorum device name:                                    d20
  Quorum device votes:                                 1
  Quorum device enabled:                               yes
  Quorum device name:                                  /dev/did/rdsk/d20s2
  Quorum device hosts (enabled):                       phys-schost-2 phys-schost-3
  Quorum device hosts (disabled):
```

Administering Cluster Interconnects and Public Networks

This chapter provides the software procedures for administering the Sun Cluster interconnects and public networks.

Administering the cluster interconnects and public networks consists of both hardware and software procedures. Typically, you configure the cluster interconnects and public networks, including NAFO groups, when you initially install and configure the cluster. If you later need to alter a cluster interconnect or public network configuration, you can use the software procedures in this chapter.

This is a list of the procedures in this chapter.

For a high-level description of the related procedures in this chapter, see TABLE 5-1 and TABLE 5-2.

Refer to the *Sun Cluster 3.0 U1 Concepts* document for background and overview information on the cluster interconnects and public networks.

Administering the Cluster Interconnects

This section provides the procedures for reconfiguring cluster interconnects, such as cluster transport adapters and cluster transport cables. These procedures require that you install Sun Cluster software.

Most of the time, you can use the `scsetup` utility to administer the cluster transport for the cluster interconnects. See the `scsetup`(1M) man page for more information.

For cluster software installation procedures, see the *Sun Cluster 3.0 U1 Installation Guide*. For procedures about servicing cluster hardware components, see the *Sun Cluster 3.0 U1 Hardware Guide*.

TABLE 5-1 Task List: Administering the Cluster Interconnect

Task	For Instructions, Go To...
Administer the cluster transport - Use `scsetup`	"How to Access the scsetup Utility" on page 8
Check the status of the cluster interconnect - Use `scstat`	"How to Check the Status of the Cluster Interconnect" on page 83
Add a cluster transport cable, transport adapter, or transport junction - Use `scsetup`	"How to Add Cluster Transport Cables, Transport Adapters, or Transport Junctions" on page 84
Remove a cluster transport cable, transport adapter, or transport junction - Use `scsetup`	"How to Remove Cluster Transport Cables, Transport Adapters, and Transport Junctions" on page 86
Enable a cluster transport cable - Use `scsetup`	"How to Enable a Cluster Transport Cable" on page 89
Disable a cluster transport cable - Use `scsetup`	"How to Disable a Cluster Transport Cable" on page 90

▼ How to Check the Status of the Cluster Interconnect

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

You do not need to be logged in as superuser to perform this procedure.

1. **Check the status of the cluster interconnect.**

```
# scstat -W
```

2. **Refer to the following for common status messages.**

Status Message	Description and Possible Action
Path online	The path is currently functioning correctly. No action is necessary.
Path waiting	The path is currently being initialized. No action is necessary.
Path faulted	The path is not functioning. This can be a transient state when paths are going between the waiting and online state. If the message persists when scstat -W is rerun, take corrective action.

Example—Checking the Status of the Cluster Interconnect

The following example shows the status of a functioning cluster interconnect.

```
# scstat -W
-- Cluster Transport Paths --
                     Endpoint            Endpoint            Status
                     --------            --------            ------
  Transport path:    phys-schost-1:qfe1  phys-schost-2:qfe1  Path online
  Transport path:    phys-schost-1:qfe0  phys-schost-2:qfe0  Path online
  Transport path:    phys-schost-1:qfe1  phys-schost-3:qfe1  Path online
  Transport path:    phys-schost-1:qfe0  phys-schost-3:qfe0  Path online
  Transport path:    phys-schost-2:qfe1  phys-schost-3:qfe1  Path online
  Transport path:    phys-schost-2:qfe0  phys-schost-3:qfe0  Path online
```

▼ How to Add Cluster Transport Cables, Transport Adapters, or Transport Junctions

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

1. Ensure that the physical cluster transport cables are installed.

For the procedure on installing a cluster transport cable, see the *Sun Cluster 3.0 U1 Hardware Guide*.

2. Become superuser on any node in the cluster.

3. Enter the scsetup **utility.**

```
# scsetup
```

The Main Menu is displayed.

4. Access the Cluster Interconnect Menu by typing 3 **(Cluster interconnect).**

5. Add the transport cable by typing 1 **(Add a transport cable).**

Follow the instructions and enter the requested information.

6. Add the transport adapter by typing 2 **(Add a transport adapter to a node).**

Follow the instructions and enter the requested information.

7. Add the transport junction by typing 3 **(Add a transport junction).**

Follow the instructions and enter the requested information.

8. Verify that the cluster transport cable, transport adapter, or transport junction is added.

```
# scconf -p | grep cable
# scconf -p | grep adapter
# scconf -p | grep junction
```

Example—Adding a Cluster Transport Cable, Transport Adapter, or Transport Junction

The following example shows how to add a transport cable, transport adapter, or transport junction to a node using the `scsetup` command.

```
[Ensure the physical cable is installed.]


Become superuser on any node and place the node to be removed in maintenance
state.


# scsetup


Select Cluster interconnect.
Select either Add a transport cable, Add a transport adapter to a node, or Add
a transport junction.


Answer the questions when prompted.
    You Will Need:           Example:
    node names               phys-schost-1
    adapter names            qfe2
    junction names           hub2
    transport type           dlpi


[Verify that the scconf command completed successfully:]
Command completed successfully.


Quit the scsetup Cluster Interconnect Menu and Main Menu.


[Verify that the cable, adapter, and junction are added:]
# scconf -p | grep cable
  Transport cable:    phys-schost-2:qfe0@1 ethernet-1@2    Enabled
  Transport cable:    phys-schost-3:qfe0@1 ethernet-1@3    Enabled
  Transport cable:    phys-schost-1:qfe0@0 ethernet-1@1    Enabled


# scconf -p | grep adapter
  Node transport adapters:                       qfe2 hme1 qfe0
  Node transport adapter:                        qfe0
  Node transport adapters:                       qfe0 qfe2 hme1
  Node transport adapter:                        qfe0
  Node transport adapters:                       qfe0 qfe2 hme1
  Node transport adapter:                        qfe0
```

```
# scconf -p | grep junction
  Cluster transport junctions:                      hub0 hub1 hub2
  Cluster transport junction:                       hub0
  Cluster transport junction:                       hub1
  Cluster transport junction:                       hub2
```

▼ How to Remove Cluster Transport Cables, Transport Adapters, and Transport Junctions

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

Use the following procedure to remove cluster transport cables, transport adapters, and transport junctions from a node configuration. When a cable is disabled, the two endpoints of the cable remain configured. An adapter cannot be removed if it is still in use as an endpoint on a transport cable.

> ⚠ **Caution –** A cluster node needs at least one functioning cable attached to the cluster interconnect to be a member of the cluster. Always verify the status of a node's cluster interconnect before disabling a cable. Only disable a cable connection after you have verified that it is redundant; that is, that another connection is available. Disabling a node's last remaining working cable takes the node out of cluster membership.

1. **Become superuser on any node in the cluster.**

2. **Check the status of the remaining cluster transport path.**

```
# scstat -W
```

> ⚠ **Caution –** If you receive an error such as "path faulted" while attempting to remove one node of a two-node cluster, investigate the problem before continuing with this procedure. Such a problem could indicate that a node path is unavailable. Removing the remaining good path takes the node out of cluster membership and could result in a cluster reconfiguration.

3. **Enter the** `scsetup` **utility.**

```
# scsetup
```

The Main Menu is displayed.

4. **Access the Cluster Interconnect Menu by typing** 3 **(Cluster interconnect).**

5. **Remove the cable by typing** 4 **(Remove a transport cable).**

Follow the instructions and enter the requested information.

Note – If you are removing a physical cable, disconnect the cable between the port and the destination device.

---

6. **Remove the adapter by typing** 5 **(Remove a transport adapter from a node).**

Follow the instructions and enter the requested information.

Note – If you are removing a physical adapter from a node, see the *Sun Cluster 3.0 U1 Hardware Guide* for hardware service procedures.

---

7. **Remove the junction by typing** 6 **(Remove a transport junction).**

Follow the instructions and enter the requested information.

Note – A junction cannot be removed if any of the ports are still in use as endpoints on any transport cables.

---

8. **Verify that the cable or the adapter has been removed.**

```
# scconf -p | grep cable
# scconf -p | grep adapter
# scconf -p | grep junction
```

The transport cable or adapter removed from the given node should not appear in the output from this command.

Example—Removing a Transport Cable, Transport Adapter, or Transport Junction

The following example shows how to remove a transport cable, transport adapter, or transport junction using the scsetup command.

```
[Become superuser on any node in the cluster.]

[Enter the scsetup utility:]
# scsetup

Select Cluster interconnect.
Select either Add a transport cable, Add a transport adapter to a node, or Add
a transport junction.

Answer the questions when prompted.
    You Will Need:         Example:
    node names             phys-schost-1
    adapter names          qfe1
    junction names         hub1

[Verify that the scconf command completed successfully:]
"Command completed successfully."

Quit the scsetup Cluster Interconnect Menu and Main Menu.

[Verify that the cable, adapter, or junction is removed:]
# scconf -p | grep cable
  Transport cable:    phys-schost-2:qfe0@1 ethernet-1@2    Enabled
  Transport cable:    phys-schost-3:qfe0@1 ethernet-1@3    Enabled
  Transport cable:    phys-schost-1:qfe0@0 ethernet-1@1    Enabled
# scconf -p | grep adapter
  Node transport adapters:   qfe2 hme1 qfe0
  Node transport adapter:    qfe0
  Node transport adapters:   qfe0 qfe2 hme1
  Node transport adapter:    qfe0
  Node transport adapters:   qfe0 qfe2 hme1
  Node transport adapter:    qfe0
# scconf -p | grep junction
  Cluster transport junctions:    hub0 hub2
  Cluster transport junction:     hub0
  Cluster transport junction:     hub2
```

▼ How to Enable a Cluster Transport Cable

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

This option is used to enable an already existing cluster transport cable.

1. **Become superuser on any node in the cluster.**

2. **Enter the** scsetup **utility.**

```
# scsetup
```

The Main Menu is displayed.

3. **Access the Cluster Interconnect Menu by typing** 2 **(Cluster interconnect).**

4. **Enable the transport cable by typing** 7 **(Enable a transport cable).**

   Follow the instructions when prompted. You need to enter both the node and the adapter names of one of the endpoints of the cable you are trying to identify.

5. **Verify that the cable is enabled.**

```
# scconf -p | grep cable
```

Example—Enabling a Cluster Transport Cable

This example shows how to enable a cluster transport cable on adapter qfe-1 located on the node phys-schost-2.

```
[Become superuser on any node.]


[Enter the scsetup utility:]
# scsetup


Select Cluster interconnect>Enable a transport cable.


Answer the questions when prompted.
You will need the following information.
    You Will Need:          Example:
   node names              phys-schost-2
   adapter names           qfe1
   junction names          hub1
```

```
[Verify that the scconf command completed successfully:]

scconf -c -m endpoint=phys-schost-2:qfe1,state=enabled

Command completed successfully.

Quit the scsetup Cluster Interconnect Menu and Main Menu.

[Verify that the cable is enabled:]
# scconf -p | grep cable
  Transport cable:    phys-schost-2:qfe1@0 ethernet-1@2    Enabled
  Transport cable:    phys-schost-3:qfe0@1 ethernet-1@3    Enabled
  Transport cable:    phys-schost-1:qfe0@0 ethernet-1@1    Enabled
```

▼ How to Disable a Cluster Transport Cable

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

You might need to disable a cluster transport cable to temporarily shut down a cluster interconnect path. This is useful when troubleshooting a cluster interconnect problem or when replacing cluster interconnect hardware.

When a cable is disabled, the two endpoints of the cable remain configured. An adapter cannot be removed if it is still in use as an endpoint in a transport cable.

<div style="border: 1px solid; padding: 0.5em;">

⚠ **Caution –** A cluster node needs at least one functioning cable attached to the cluster interconnect to be a member of the cluster. Always verify the status of a node's cluster interconnect before disabling a cable. Only disable a cable connection after you have verified that it is redundant; that is, that another connection is available. Disabling a node's last remaining working cable takes the node out of cluster membership.

</div>

1. **Become superuser on any node in the cluster.**

2. **Check the status of the cluster interconnect before disabling a cable.**

   ```
   # scstat -W
   ```

> **⚠ Caution –** If you receive an error such as "path faulted" while attempting to remove one node of a two-node cluster, investigate the problem before continuing with this procedure. Such a problem could indicate that a node path is unavailable. Removing the remaining good path takes the node out of cluster membership and could result in a cluster reconfiguration.

3. **Enter the** scsetup **utility.**

```
# scsetup
```

The Main Menu is displayed.

4. **Access the Cluster Interconnect Menu by typing** 3 **(Cluster interconnect).**

5. **Disable the cable by typing** 8 **(Disable a transport cable).**

   Follow the instructions and enter the requested information. All of the components on this cluster interconnect will be disabled. You need to enter both the node and the adapter names of one of the endpoints of the cable you are trying to identify.

6. **Verify that the cable is disabled.**

```
# scconf -p | grep cable
```

Example—Disabling a Cluster Transport Cable

This example shows how to disable a cluster transport cable on adapter qfe-1 located on the node phys-schost-2.

```
[Become superuser on any node.]


[Enter the scsetup utility:]
# scsetup


Select Cluster interconnect>Disable a transport cable.


Answer the questions when prompted.
You will need the following information.
   You Will Need:         Example:
   node names             phys-schost-2
   adapter names          qfe1
   junction names         hub1
```

```
[Verify that the scconf command completed successfully:]

scconf -c -m endpoint=phys-schost-2:qfe1,state=disabled

Command completed successfully.

Quit the scsetup Cluster Interconnect Menu and Main Menu.

[Verify that the cable is disabled:]
# scconf -p | grep cable
  Transport cable:    phys-schost-2:qfe1@0 ethernet-1@2    Disabled
  Transport cable:    phys-schost-3:qfe0@1 ethernet-1@3    Enabled
  Transport cable:    phys-schost-1:qfe0@0 ethernet-1@1    Enabled
```

Administering the Public Network

If you need to alter a public network configuration, you can use the software procedures in this section.

When administering public network adapters, pay attention to the following points.

- Avoid unconfiguring (unplumbing) or bringing down the active adapter of a NAFO group without first switching over the active adapter to a backup adapter in the group. See "How to Switch a NAFO Group's Active Adapter" on page 100.
- Avoid rewiring backup adapters to different subnets without first removing them from their respective NAFO groups.
- Logical adapter operations can be done on the active adapter even if monitoring is on for the group.
- You must maintain at least one public network connection for each node in the cluster. The cluster is inaccessible without a public network connection.

For cluster software installation procedures, see the *Sun Cluster 3.0 U1 Installation Guide*. For procedures about servicing public networking hardware components, see the *Sun Cluster 3.0 U1 Hardware Guide*.

TABLE 5-2 Task Map: Administering the Public Network

Task	For Instructions, Go To...
Create a NAFO group on a node	"How to Create a NAFO Group" on page 94
Add more public network adapters to a node	"How to Add an Adapter to a NAFO Group" on page 96
Delete a NAFO group	"How to Delete a NAFO Group" on page 97
Remove backup adapters from an existing NAFO group	"How to Remove an Adapter From a NAFO Group" on page 99
Switch the active adapter to a backup adapter	"How to Switch a NAFO Group's Active Adapter" on page 100
Check the status of NAFO groups	"How to Check the Status of NAFO Groups" on page 102
Change parameters to tune the PNM fault detection and failover process	"How to Change Public Network Management Tunable Parameters" on page 103

▼ How to Create a NAFO Group

You can also accomplish this procedure by using the SunPlex Manager GUI. See the
SunPlex Manager online help for more information.

1. **Note the following requirements for creating a NAFO group:**
   - All public network adapters must be configured to belong to a NAFO group.
   - For any given node, there can be at most one NAFO group on a given subnet.
   - All adapters in a given NAFO group must be connected to the same subnet.
   - Only one adapter in a given NAFO group can have a hostname association, that
     is, an `/etc/hostname.`*adapter* file.
   - A public network adapter can belong to only one NAFO group.

2. **Become superuser on the node being configured for a NAFO group.**

3. **For this node, find out the public network adapters that are physically connected
   to the same subnet.**

   These adapters form the backup adapters for the NAFO group.

4. **Create the** `/etc/hostname.`*adapter* **file for one of the public network adapters, if
   the file does not already exist.**

   The adapter specified in this file will be the default active adapter for the NAFO
   group.

   ```
   # vi /etc/hostname.<adapter>
   phys-schost-1
   ```

5. **Edit the** `/etc/inet/hosts` **file to add the IP address and corresponding
   hostname assigned to the public network adapter.**

   For example, the following shows the IP address `192.29.75.101` and hostname
   `phys-schost-1` added to the `/etc/inet/hosts` file.

   ```
   # vi /etc/inet/hosts
   192.29.75.101   phys-schost-1
   ```

   Note – If a naming service is used, this information should also exist in the naming
   service database.

   ---

6. **Create the NAFO group.**

   ```
   # pnmset -c nafo-group -o create adapter [adapter ...]
   ```

–c *nafo-group*	Performs a configuration subcommand for the specified NAFO group. NAFO groups must be named nafo*N*, where *N* is a nonnegative integer identifier for the group. Group names are local to each node. Thus, the same NAFO group name can be used on multiple nodes.
-o create	Creates the new NAFO group.
adapter [*adapter* ...]	Specifies the public network adapter(s) that serves as the backup adapter. See Step 3.

Note – If an adapter is already configured, it will be chosen as the active adapter and the pnmset command does not alter its state. Otherwise, one of the backup adapters will be configured and assigned the IP address found in the /etc/hostname.*adapter* file for the NAFO group.

---

7. **Verify the status of the NAFO group.**

```
# pnmstat -l
```

Example—Creating a NAFO Group

The following example shows the creation of a NAFO group (nafo0) configured with two network adapters (qfe0 and qfe1).

```
# pnmstat -l
# vi /etc/hostname.qfe0
phys-schost-1
# vi /etc/inet/hosts
192.29.75.101   phys-schost-1
# pnmset -c nafo0 -o create qfe0 qfe1
# pnmstat -l
group   adapters        status   fo_time   act_adp
nafo0   qfe0:qfe1        OK       NEVER     qfe0
```

▼ How to Add an Adapter to a NAFO Group

You can add adapters to an existing NAFO group to provide additional backup adapters for the NAFO group and thereby increase the availability of public network connectivity for the cluster node.

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

1. **Do you need to install the new public network adapter card(s) in the node(s)?**
   - If yes, see the *Sun Cluster 3.0 U1 Hardware Guide* for instructions.
   - If no, proceed to Step 2.

2. **Make sure the adapter to be added to the NAFO group is connected to the same subnet as the active adapter for the NAFO group.**

3. **Make sure the adapter is not plumbed, and that it does not have an associated** /etc/hostname.*adapter* **file.**

4. **Become superuser on the node that contains the NAFO group to which the new adapter is being added.**

5. **Add the adapter to the NAFO group.**

   ```
   # pnmset -c nafo-group -o add adapter
   ```

   -c *nafo-group*   Specifies the NAFO group to which the new adapter is being added.

   -o add *adapter* Specifies the public network adapter being added to the named NAFO group.

6. **Verify the status of the NAFO group.**

   ```
   # pnmstat -l
   ```

Example—Adding an Additional Public Network Adapter to a NAFO Group

The following example shows the addition of adapter `qfe2` to NAFO group `nafo0` which already contained two adapters (`qfe0`, `qfe1`).

```
# pnmstat -l
group   adapters        status   fo_time   act_adp
nafo0   qfe0:qfe1        OK       NEVER     qfe0
# pnmset -c nafo0 -o add qfe2
# pnmstat -l
group   adapters        status   fo_time   act_adp
nafo0   qfe0:qfe1:qfe2   OK       NEVER     qfe0
```

▼ How to Delete a NAFO Group

Delete a NAFO group when you do not want monitoring and failover for any adapter in the group. To be deleted, a NAFO group cannot be in use by logical host resource groups or shared address resource groups.

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

1. **Become superuser on the node that contains the NAFO group that is being deleted.**

2. **Identify whether the NAFO group is being used by any logical host or shared address resources.**

   ```
   # scrgadm -pv
   ```

   You can also use `scrgadm -pvv` (with two v flags) to locate the resources that are using the NAFO group you are going to delete.

3. **Switch the logical host resource groups and shared address resource groups that use this NAFO group.**

   ```
   # scswitch -z -g resource-group -h nodelist
   ```

`-z -g` *resource-group*		Switches the specified resource group.
`-h` *nodelist*		Specifies the name of the node to switch the resource group to.

4. Delete the NAFO group.

```
# pnmset -c nafo-group -o delete
```

`-c` *nafo-group*	Specifies the NAFO group to be deleted.
`-o delete`	Deletes the NAFO group.

5. Verify the status of the NAFO group.

The deleted NAFO group should not appear in the listing.

```
# pnmstat -l
```

Example—Deleting a NAFO Group

The following example shows the NAFO group named `nafo1` deleted from the system. Logical host resource group `lh-rg-1`, which uses this NAFO group, is first switched to a different node.

```
# scswitch -z -g lh-rg-1 -h phys-schost-2
# pnmstat -l
group   adapters        status   fo_time   act_adp
nafo0   qfe0:qfe1        OK       NEVER     qfe0
nafo1   qfe2            OK       NEVER     qfe2
# pnmset -c nafo1 -o delete
# pnmstat -l
group   adapters        status   fo_time   act_adp
nafo0   qfe0:qfe1        OK       NEVER     qfe0
```

▼ How to Remove an Adapter From a NAFO Group

Remove backup adapters from an existing NAFO group to enable the adapter to be removed from the system, to be replaced, or to be reconnected to a different subnet and used as backup for another NAFO group.

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

> **Caution –** Removing the last backup adapter from a NAFO group results in no protection against faults detected on the active adapter, reducing public network availability for the cluster node.

1. **If you want to remove the active adapter, first switch to another adapter in the group.**

   See "How to Switch a NAFO Group's Active Adapter" on page 100.

2. **As superuser, remove the adapter from the NAFO group.**

   ```
   # pnmset -c nafo-group -o remove adapter
   ```

   -c *nafo-group*   Specifies the NAFO group from which to remove the adapter.

   -o remove *adapter*   Removes the adapter from the NAFO group.

3. **Verify the status of the NAFO group.**

   The deleted adapter should not appear in the listing for the NAFO group.

   ```
   # pnmstat -l
   ```

Example—Removing an Adapter From a NAFO Group

The following example removes adapter `qfe2` from NAFO group `nafo0`.

```
# pnmstat -l
group   adapters       status   fo_time  act_adp
nafo0   qfe0:qfe1:qfe2  OK       NEVER    qfe0
# pnmset -c nafo0 -o remove qfe2
# pnmstat -l
group   adapters       status   fo_time  act_adp
nafo0   qfe0:qfe1       OK       NEVER    qfe0
```

▼ How to Switch a NAFO Group's Active Adapter

Switch the active adapter to a backup adapter so that the current active adapter can be removed from the NAFO group. The `pnmd(1M)` daemon moves all IP addresses hosted by the current active adapter to the new active adapter in a similar fashion as a fault-triggered adapter failover.

You can also accomplish this procedure by using the SunPlex Manager GUI. See the SunPlex Manager online help for more information.

Note – Connections can experience a delay while the switchover is taking place. This delay can last for several minutes. Otherwise, the operation is transparent to higher-level applications.

---

1. **Ensure the physical connectivity of the new active adapter is identical to that of the current active adapter.**

   If the new active adapter fails to host some of the IP addresses as the current active adapter, network and data services that depend on those IP addresses are interrupted until the physical connectivity is fixed or a subsequent successful failover occurs.

2. **Become superuser on the node that contains the NAFO group whose active adapter you want to switch.**

3. **Switch the active adapter.**

   ```
   # pnmset -c nafo-group -o switch adapter
   ```

-c *nafo-group*	Specifies the NAFO group containing the adapter to switch.
-o switch *adapter*	Makes the specified adapter the active adapter in the NAFO group.

- If this change is temporary and you do not want the change to survive a reboot of the node, go to Step 5.
- If this change is permanent and you do want the change to survive a reboot of the node, go to Step 4.

4. **Rename the** /etc/hostname.*adapter* **file for the old active adapter to reflect the new active adapter.**

```
# mv /etc/hostname.<old_adapter> /etc/hostname.<new_adapter>
```

5. **Verify the status of the NAFO group.**

The "switched-to" adapter should now appear as the active adapter.

```
# pnmstat -l
```

Example—Switching a NAFO Group's Active Adapter

The following example switches the active adapter to qfe1 from qfe0.

```
# pnmstat -l
group    adapters        status   fo_time  act_adp
nafo0    qfe0:qfe1         OK       NEVER    qfe0
# pnmset -c nafo0 -o switch qfe1
# mv /etc/hostname.qfe0 /etc/hostname.qfe1
# pnmstat -l
group    adapters        status   fo_time  act_adp
nafo0    qfe0:qfe1         OK       11       qfe1
```

▼ How to Check the Status of NAFO Groups

You can also accomplish this procedure by using the SunPlex Manager GUI. See the
SunPlex Manager online help for more information.

● **Run the** `pnmstat(1M)` **command to list information about the current setup and
status of all NAFO groups on a node.**

```
# pnmstat -l
```

You can also use the `pnmptor(1M)` and `pnmrtop(1M)` commands to get
information on adapters.

Examples—Checking the Status of NAFO Groups

The following example shows the status of a node's three NAFO groups.

```
# pnmstat -l
Group    adapters         status  fo_time act_adp
nafo0    qfe5             OK      NEVER   qfe5
nafo1    qfe6             OK      NEVER   qfe6
nafo2    qfe7             OK      NEVER   qfe7
```

Examples—Finding the Active Adapter For a NAFO Group

The following example shows that the active adapter in NAFO group `nafo0` is
adapter `qfe5`.

```
# pnmptor nafo0
qfe5
```

Examples—Finding an Adapter's NAFO Group

The following example shows that adapter `qfe5` belongs to NAFO group `nafo0`.

```
# pnmrtop qfe5
nafo0
```

▼ How to Change Public Network Management Tunable Parameters

There are four tunable parameters in this algorithm.

- `inactive_time`
- `ping_timeout`
- `repeat_test`
- `slow_network`

These parameters provide an adjustable trade-off between speed and correctness of fault detection. See TABLE 5-3 for more information.

Use this procedure to change the default Public Network Management (PNM) values for the pnmd(1M) daemon.

1. **Become superuser on any node in the cluster.**

2. **If it doesn't already exist, create the** pnmparams **file.**

```
# vi /etc/cluster/pnmparams
```

3. **Use the following table to set PNM parameters.**

Note – Settings in the `/etc/cluster/pnmparams` file apply to all NAFO groups on the node. Lines started with a pound sign (#) are ignored. Other lines in the file must be of the format: *variable=value*.

TABLE 5-3   Public Network Management Tunable Parameters

Parameter	Description
`inactive_time`	Number of seconds between successive probes of the packet counters of the current active adapter. Default is 5.
`ping_timeout`	Time-out value in seconds for the `ALL_HOST_MULTICAST` and subnet broadcast pings. Default is 4.
`repeat_test`	Number of times to do the ping sequence before declaring that the active adapter is faulty and failover is triggered. Default is 3.
`slow_network`	Number of seconds waited after each ping sequence before checking packet counters for any change. Default is 2.
`warmup_time`	Number of seconds waited after failover to a backup adapter before resuming fault monitoring. This allows extra time for any slow driver or port initialization. Default is 0.

Note – The changes don't take affect until the next time the pnmd daemon starts up.

Example—Changing Public Network Management Tunable Parameters

The following shows a sample /etc/cluster/pnmparams file, with two parameters changed from their default values.

```
inactive_time=3
repeat_test=5
```

Administering the Cluster

This chapter provides the procedures for administering items that affect the entire cluster.

This is a list of the procedures in this chapter.

Administering the Cluster Overview

TABLE 6-1   Task List: Administering the Cluster

Task	For Instructions, Go To
Change the name of the cluster	"How to Change the Cluster Name" on page 106
List node IDs and their corresponding node names	"How to Map Node ID to Node Name" on page 107
Permit or deny new nodes to add themselves to the cluster	"How to Work With New Cluster Node Authentication" on page 108
Change the time for a cluster using the Network Time Protocol (NTP)	"How to Reset the Time of Day in a Cluster" on page 109
Bring down a node and enter the OpenBoot™ PROM	"How to Enter the OpenBoot PROM (OBP) on a Node" on page 110
Change the private hostname	"How to Change the Private Hostname" on page 111
Put a cluster node in maintenance state	"How to Put a Node Into Maintenance State" on page 114
Bring a cluster node out of maintenance state	"How to Bring a Node Out of Maintenance State" on page 116
Add a node to a cluster	"How to Add a Cluster Node to the Authorized Node List" on page 119
Remove a node from a cluster	"How to Add a Cluster Node to the Authorized Node List" on page 119

▼ How to Change the Cluster Name

If necessary, you can change the cluster name after initial installation.

1. **Become superuser on any node in the cluster.**

2. **Enter the** scsetup**(1M) utility.**

```
# scsetup
```

The Main Menu is displayed.

3. **To change the cluster name, type** 6 **(Other cluster properties).**
   The Other Cluster Properties menu is displayed.

4. **Make your selection from the menu and follow the onscreen instructions.**

Example—Changing the Cluster Name

The following example shows the scconf(1M) command generated from the scsetup utility to change to the new cluster name, dromedary.

```
# scconf -c -C cluster=dromedary
```

▼ How to Map Node ID to Node Name

During Sun Cluster installation, each node is automatically assigned a unique node ID number. The node ID number is assigned to a node in the order in which it joins the cluster for the first time; once assigned, the number cannot be changed. The node ID number is often used in error messages to identify which cluster node the message concerns. Use this procedure to determine the mapping between node IDs and node names.

You do not need to be superuser to list configuration information.

1. **Use the** scconf**(1M) command to list the cluster configuration information.**

```
% scconf -pv | grep "Node ID"
```

Example—Mapping the Node ID to the Node Name

The following example shows the node ID assignments

```
% scconf -pv | grep "Node ID"
    (phys-schost-1) Node ID:                      1
    (phys-schost-2) Node ID:                      2
    (phys-schost-3) Node ID:                      3
```

▼ How to Work With New Cluster Node Authentication

Sun Cluster enables you to determine if new nodes can add themselves to the cluster and with what type of authentication. You can permit any new node to join the cluster over the public network, deny new nodes from joining the cluster, or indicate a specific node that can join the cluster. New nodes can be authenticated by using either standard UNIX or Diffie-Hellman (DES) authentication. If you select DES authentication, you must also configure all necessary encryption keys before a node can join. See the keyserv(1M) and publickey(4) man pages for more information.

1. Become superuser on any node in the cluster.

2. Enter the scsetup(1M) utility.

```
# scsetup
```

The Main Menu is displayed.

3. To work with cluster authentication, type 6 (New nodes).

The New Nodes menu is displayed.

4. Make your selection from the menu and follow the onscreen instructions.

Examples—Preventing New Machines From Being Added to the Cluster

The following example shows the scconf(1M) command generated from the scsetup utility that would prevent new machines from being added to the cluster.

```
# scconf -a -T node=.
```

Examples—Permitting All New Machines to Be Added to the Cluster

The following example shows the scconf command generated from the scsetup utility that would enable all new machines to be added to the cluster.

```
# scconf -r -T all
```

Examples—Specifying a New Machine to Be Added to the Cluster

The following example shows the scconf command generated from the scsetup utility to enable a single new machine to be added to the cluster.

```
# scconf -a -T node=phys-schost-4
```

Examples—Setting the Authentication to Standard UNIX

The following example shows the scconf command generated from the scsetup utility to reset to standard UNIX authentication for new nodes joining the cluster.

```
# scconf -c -T authtype=unix
```

Examples—Setting the Authentication to DES

The following example shows the scconf command generated from the scsetup utility to use DES authentication for new nodes joining the cluster.

```
# scconf -c -T authtype=des
```

Note – When using DES authentication, you need to also configure all necessary encryption keys before a node can join the cluster. See the keyserv(1M) and publickey(4) man pages for more information.

▼ How to Reset the Time of Day in a Cluster

Sun Cluster uses the Network Time Protocol (NTP) to maintain time synchronization between cluster nodes. Adjustments in the cluster occur automatically as needed when nodes synchronize their time. See the *Sun Cluster 3.0 U1 Concepts* document and the *Network Time Protocol User's Guide* for more information.

Caution – When using NTP, do not attempt to adjust the cluster time while the cluster is up and running. This includes using the date(1), rdate(1M), or xntpdate(1M) commands interactively or within cron(1M) scripts.

1. **Become superuser on any node in the cluster.**

2. **Shut down the cluster to the OBP prompt.**

   ```
   # scshutdown -g0 -y
   ```

3. **Boot each node into non-cluster node.**

   ```
   ok boot -x
   ```

4. **On a single node, set the time of day by running the** date**(1) command.**

   ```
   # date HHMMSS
   ```

5. **On the other machines, synchronize the time to that node by running the** rdate**(1M) command.**

   ```
   # rdate hostname
   ```

6. **Boot each node to restart the cluster.**

   ```
   # reboot
   ```

7. **Verify that the change took place on all cluster nodes.**

   On each node, run the date(1M) command.

   ```
   # date
   ```

▼ How to Enter the OpenBoot PROM (OBP) on a Node

Use this procedure if you need to configure or change OpenBoot PROM settings.

1. **Connect to the terminal concentrator port.**

   ```
   # telnet tc_name tc_port_number
   ```

tc_name	Specifies the name of the terminal concentrator.
tc_*port_number*	Specifies the port number on the terminal concentrator. Port numbers are configuration dependent. Typically, ports 2 and 3 (5002 and 5003) are used for the first cluster installed at a site.

2. **Shut down the cluster node gracefully by using the** scswitch(1M) **command to evacuate any resource or disk device groups and then** shutdown(1M) **to bring the node to the OBP prompt.**

```
# scswitch -S -h nodelist
# shutdown -g0 -y -i0
```

Caution – Do not use send brk on a cluster console to shut down a cluster node. If you use send brk and then type go at the OBP prompt to reboot, the node will panic. This functionality is not supported within a cluster.

3. **Execute the OBP commands.**

▼ How to Change the Private Hostname

Use this procedure to change the private hostname of a cluster node after installation has been completed.

Default private hostnames are assigned during initial cluster installation. The default private hostname takes the form clusternode<*nodeid*>-priv, for example: clusternode3-priv. You should only change a private hostname if the name is already in use in the domain.

Caution – Do not attempt to assign IP addresses to new private hostnames. The clustering software assigns them.

1. **Disable, on all nodes in the cluster, any Data Service resources or other applications that might cache private hostnames.**

```
# scswitch -n -j resource1, resource2
```

Include the following in the applications you disable.

- HA-DNS and HA-NFS services, if configured.
- Any application which has been custom configured to use the private hostname.

■ Any application which is being used by clients over the private interconnect.

See the scswitch(1M) man page and the *Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide* for information about using the scswitch command.

2. **Bring down the Network Time Protocol (NTP) daemon on each node of the cluster.**

See the xntpd man page for more information about the NTP daemon.

```
# /etc/init.d/xntpd stop
```

3. **Determine the name of the node on which you are changing the private hostname.**

```
# scconf -p | grep node
```

4. **Run the scsetup utility to change the private hostname.**

It is only necessary to do this from one of the nodes in the cluster.

Note – When selecting a new private hostname, be sure the name is unique to the cluster node.

---

5. **Select 5, Private Hostnames, from the Main Menu.**

6. **Select 1, Change a Private Hostname, from the Private Hostnames Menu.**

Answer the questions when prompted. You will be asked the name of the node whose private hostname is being changed (clusternode<*nodeid*>-priv), and the new private hostname.

7. **Flush the name service cache.**

Do this on each node in the cluster. This prevents the cluster applications and data services from trying to access the old private hostname.

```
# nscd -i hosts
```

8. **Edit the ntp.conf file on each node to change the private hostname to the new one.**

Use whatever editing tool you prefer.

If this is done at install time, also remember to remove names for nodes which are configured; the default template comes pre-configured with eight nodes. Typically, the ntp.conf file will be identical on each cluster node.

9. **Verify that you can successfully ping the new private hostname from all cluster nodes.**

10. **Restart the NTP daemon.**

    Do this on each node of the cluster.

    ```
    # /etc/init.d/xntpd start
    ```

11. **Enable all Data Service resources and other applications that were disabled in Step 1.**

    ```
    # scswitch -e -j resource1, resource2
    ```

    See the scswitch(1M) man page and the *Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide* for information about using the scswitch command.

Example—Changing the Private Hostname

The following example changes the private hostname from clusternode2-priv to clusternode4-priv, on node phys-schost-2.

```
[Disable all applications and data services as necessary.]
phys-schost-1# /etc/init.d/xntpd stop

phys-schost-1# scconf -p | grep node
 ...
 Cluster nodes:                              phys-schost-1 phys-schost-2 phys-
 schost-3
 Cluster node name:                               phys-schost-1
  Node private hostname:                          clusternode1-priv
 Cluster node name:                               phys-schost-2
  Node private hostname:                          clusternode2-priv
 Cluster node name:                               phys-schost-3
  Node private hostname:                          clusternode3-priv
 ...

phys-schost-1# scsetup

phys-schost-1# nscd -i hosts
```

```
phys-schost-1# vi /etc/inet/ntp.conf
 ...
 peer clusternode1-priv
 peer clusternode4-priv
 peer clusternode3-priv

phys-schost-1# ping clusternode4-priv

phys-schost-1# /etc/init.d/xntpd start

[Enable all applications and data services disabled at the beginning of the
procedure.]
```

▼ How to Put a Node Into Maintenance State

Put a cluster node into maintenance state when taking the node out of service for an extended period of time. This way, the node does not contribute to the quorum count while it is being serviced. To put a cluster node into maintenance state, the node must be brought down using scswitch(1M) and shutdown(1M).

Note – Use the Solaris shutdown command to shut down a single node. The scshutdown command should be used only when shutting down an entire cluster.

When a cluster node is brought down and put into maintenance state, all quorum devices that are configured with ports to the node have their quorum vote counts decremented by one. The node and quorum device vote counts are incremented by one when the node is taken out of maintenance mode and brought back online.

You need to use the scconf(1M) command to put a cluster node into maintenance state.The scsetup utility does not include the functionality for putting a quorum device into maintenance state.

1. **Become superuser on the node to be put into maintenance state.**

2. **Evacuate any resource groups and disk device groups from the node.**

   ```
   # scswitch -S -h nodelist
   ```

-S	Evacuates all device services and resource groups from the specified node.
–h *nodelist*	Specifies the node from which you are switching resource groups and devices groups.

3. **Bring the node you evacuated down to the OBP prompt and out of the cluster.**

```
# shutdown -g0 -y -i0
```

4. **Become superuser on another node in the cluster and put the node brought down in Step 3 into maintenance state.**

```
# scconf -c -q node=node,maintstate
```

-c	Specifies the change form of the scconf command.
-q	Manages the quorum options.
node=*node*	Specifies the node name or node ID of the node to change.
maintstate	Puts the node into maintenance state.

5. **Verify that the cluster node is now in maintenance state.**

```
# scconf -p | grep -i vote
```

Example—Putting a Cluster Node Into Maintenance State

The following example moves a cluster node into maintenance state and verifies the results.

```
[On the node to be put into maintenance state:]
phys-schost-1# scswitch -S -h phys-schost-1
phys-schost-1# shutdown -g0 -y -i0

[On another node in the cluster:]
phys-schost-2# scconf -c -q node=phys-schost-1,maintstate
phys-schost-2# scconf -p | grep -i quorum
 Node quorum vote count:                      1
Quorum devices:                              d20
Quorum device name:                          d20
  Quorum device votes:                       1
  Quorum device enabled:                     yes
  Quorum device path:                        /dev/did/rdsk/d5s2
  Quorum device hosts (disabled):            phys-schost-1
  Quorum device hosts (enabled):             phys-schost-2
```

Where to Go From Here

To bring a node back online, see .

▼ How to Bring a Node Out of Maintenance State

Use the following procedure to bring a node back online and reset the quorum vote count to the default. For cluster nodes, the default quorum count is one. For quorum devices, the default quorum count is N-1, where N is the number of nodes with non-zero vote counts that have ports to the quorum device.

When a node has been put into maintenance state, the node's quorum vote count is decremented by one. All quorum devices that are configured with ports to the node will also have their quorum vote counts decremented. When the quorum vote count is reset and a node is brought back out of maintenance state, both the node's quorum vote count and the quorum device vote count are incremented by one.

Run this procedure any time a node has been put into maintenance state and you are bringing it out of maintenance state.

Caution – If you do not specify either the `globaldev` or `node` options, the quorum count is reset for the entire cluster.

1. **Become superuser on any node of the cluster.**

2. **If using quorum, reset the cluster quorum count from a node other than the one in maintenance state.**

   You must reset the quorum count from a node other than the node in maintenance state before rebooting the node, or it might hang waiting for quorum.

   ```
   # scconf -c -q node=node,reset
   ```

-c	Specifies the change form of the scconf command.
-q	Manages the quorum options.
node=*node*	Specifies the name of the node to be reset, for example, phys-schost-1.
reset	The change flag that resets quorum.

3. **Reboot the node.**

4. **Verify the quorum vote count.**

   ```
   # scconf -p | grep -i vote
   ```

Example—Bringing a Cluster Node Out of Maintenance State and Resetting the Quorum Vote Count

The following example resets the quorum count for a cluster node and its quorum devices back to their defaults and verifies the result.

```
# scconf -c -q node=phys-schost-1,reset
# scconf -pv | grep -i vote
 Node quorum vote count:              1
 Node quorum vote count:              1
 Node quorum vote count:              1
 (d20) Quorum device votes:                    1
 (d21) Quorum device votes:                    1
```

Adding and Removing a Cluster Node

The following table lists the tasks to perform when adding a node to an existing cluster. To complete the procedure correctly, these tasks must be performed in the order shown.

TABLE 6-2   Task Map: Adding a Cluster Node

Task	For Instructions, Go To
Install the host adapter on the node and verify that the existing cluster interconnects can support the new node	*Sun Cluster 3.0 U1 Hardware Guide*
Add shared storage	*Sun Cluster 3.0 U1 Hardware Guide*
Add the node to the authorized node list - Use `scsetup`.	"How to Add a Cluster Node to the Authorized Node List" on page 119
Install and configure the software on the new cluster node - Install the Solaris Operating Environment and Sun Cluster software - Configure the node as part of the cluster	*Sun Cluster 3.0 U1 Installation Guide*: See the section on installing and configuring Sun Cluster software.

The following table lists the tasks to perform when removing a node from an existing cluster. To complete the procedure correctly, the tasks must be performed in the order shown.

TABLE 6-3   Task Map: Removing a Cluster Node

Task	For Instructions, Go To
Place node being removed into maintenance state - Use `shutdown` and `scconf`	"How to Put a Node Into Maintenance State" on page 114
Remove node from all resource groups - Use `scrgadm`	*Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide*: see the procedure for how to remove a node from an existing resource group.
Remove node from all device groups of which the node is a member - Use volume manager commands	"How to Remove a Node From a Disk Device Group (Solstice DiskSuite)" on page 40 or "How to Remove a Node From a Disk Device Group (VERITAS Volume Manager)" on page 53

TABLE 6-3   Task Map: Removing a Cluster Node

Task	For Instructions, Go To
Remove all logical transport connections to the node being removed - Use `scsetup`	"How to Remove Cluster Transport Cables, Transport Adapters, and Transport Junctions" on page 86 To remove the physical hardware from the node, see the *Sun Cluster 3.0 U1 Hardware Guide* section on installing and maintaining cluster interconnect and public network hardware.
Remove all quorum devices shared with the node being removed - Use `scsetup`	"How to Remove a Quorum Device" on page 73
Remove node from the cluster software configuration - Use `scconf`	"How to Remove a Node From the Cluster Software Configuration" on page 121
Remove required shared storage from the node and cluster - Follow the procedures in your volume manager documentation and hardware guide	SDS or VxVM administration guide *Sun Cluster 3.0 U1 Hardware Guide*

▼ How to Add a Cluster Node to the Authorized Node List

Before adding a machine to an existing cluster, be sure the node has all of the necessary hardware correctly installed and configured, including a good physical connection to the private cluster interconnect.

For hardware installation information, refer to the *Sun Cluster 3.0 U1 Hardware Guide* or the hardware documentation that shipped with your server.

This procedure permits a machine to install itself into a cluster by adding its node name to the list of authorized nodes for that cluster.

You must be superuser on a current cluster member node to complete this procedure.

1. **Be sure you have correctly completed all prerequisite hardware installation and configuration tasks listed in the task map for "Adding and Removing a Cluster Node" on page 118.**

2. **Execute the** `scsetup`**(1M) utility.**

```
# scsetup
```

The Main Menu is displayed.

3. **To modify the authorized list, type** 3 **at the New Nodes Menu, Specify the name of a machine which may add itself.**

Follow the prompts to add the node's name to the cluster. You will be asked for the name of the node to be added.

4. **To access the New Nodes Menu, type** 6 **at the Main Menu.**

5. **Verify that the task has been performed successfully.**

The scsetup utility prints a "Command completed successfully" message if it completes the task without error.

6. **To prevent any new machines from being added to the cluster, type** 1 **at the New Nodes Menu.**

Follow the `scsetup` prompts. This option tells the cluster to ignore all requests coming in over the public network from any new machine trying to add itself to the cluster.

7. **Quit the** `scsetup` **utility.**

8. **Install and configure the software on the new cluster node.**

Use either `scinstall` or JumpStart to complete the installation and configuration of the new node, as described in the *Sun Cluster 3.0 U1 Installation Guide*.

Example—Adding a Cluster Node to the Authorized Node List

The following example shows how to add a node named `phys-schost-3` to the authorized node list in an existing cluster.

```
[Become superuser and execute the scsetup utility.]
# scsetup

Select New nodes>Specify the name of a machine which may add itself.

Answer the questions when prompted.
```

```
Verify that the scconf command completed successfully.

scconf -a -T node=phys-schost-3

    Command completed successfully.

Select Prevent any new machines from being added to the cluster.

Quit the scsetup New Nodes Menu and Main Menu.

[Install the cluster software.]
```

Where to Go From Here

For an overall list of tasks for adding a cluster node, see

To add a node to an existing resource group, see the *Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide*.

▼ How to Remove a Node From the Cluster Software Configuration

This is the last software configuration procedure that needs to be accomplished in the process for removing a node from a cluster. You must be superuser on a node in the cluster to perform this procedure.

1. **Be sure you have correctly completed all prerequisite tasks listed in the task map for "Adding and Removing a Cluster Node" on page 118.**

   Note – Be sure you have placed the node in maintenance state and removed it from all resource groups, device groups, and quorum device configurations before continuing with this procedure.

2. **Determine the local disks in the cluster and their associated raw disk device group names, for example** dsk/d4**.**

   ```
   # scconf -pvv | grep Local_Disk
   ```

3. **Identify which local disks and raw disk device groups in the cluster are connected to the node being removed.**

```
# scconf -pvv | grep node-name | grep Device
```

4. **Disable the localonly property for each local disk identified in Step 3.**

See the scconf_dg_rawdisk(1M) man page for more information about the localonly property.

```
# scconf -c -D name=rawdisk-device-group,localonly=false
```

5. **Remove the node from all raw disk device groups, of which the node is a member.**

This step must be completed for each raw disk device group that is connected to the node being removed.

```
# scconf -r -D name=rawdisk-device-group,nodelist=node
```

6. **Remove the node from the cluster.**

```
# scconf -r -h node=node
```

7. **Verify the node removal using** scstat.

```
# scstat -n
```

8. **To physically remove the node from the cluster, remove the hardware connections as described in the** *Sun Cluster 3.0 U1 Hardware Guide*.

Note – After the device has been removed from the cluster, you must reinstall the Solaris operating environment on the removed host before it can be placed back into service in any capacity.

Example—Removing a Node From the Cluster Software Configuration

This example shows how to remove a node (phys-schost-2) from a cluster.

```
[Become superuser on any node.]

[Identify all local disks and their raw disk device group names:]
# scconf -pvv | grep Local_Disk
    (dsk/d4) Device group type:        Local_Disk
    (dsk/d8) Device group type:        Local_Disk

[Identify the local disks and raw disk device groups connected to
the node being removed:]
# scconf -pvv | grep phys-schost-2 | grep Device
    (dsk/d4) Device group node list:  phys-schost-2
    (dsk/d2) Device group node list:  phys-schost-1, phys-schost-2
    (dsk/d1) Device group node list:  phys-schost-1, phys-schost-2

[Remove the localonly flag for each local disk on the node:]
# scconf -c -D name=dsk/d4,localonly=false

[Remove the node from all raw disk device groups:]
# scconf -r -D name=dsk/d4,nodelist=phys-schost-2
# scconf -r -D name=dsk/d2,nodelist=phys-schost-2
# scconf -r -D name=dsk/d1,nodelist=phys-schost-2

[Remove the node from the cluster:]
# scconf -r -h node=phys-schost-2

[Verify node removal:]
# scstat -n

-- Cluster Nodes --
                    Node name          Status
                    ---------          ------
  Cluster node:     phys-schost-1      Online
```

Where to Go From Here

For hardware procedures, see the *Sun Cluster 3.0 U1 Hardware Guide*.

For an overall list of tasks for removing a cluster node, see TABLE 6-3, "Task Map: Removing a Cluster Node."

To add a node to an existing cluster, see .

Patching Sun Cluster Software and Firmware

This chapter provides the procedures for adding and removing patches for a Sun Cluster configuration.

This is a list of the procedures in this chapter.

Patching Sun Cluster Overview

Due to the nature of a cluster, all cluster member nodes must be at the same patch level for proper cluster operation. When patching a node with a Sun Cluster-specific patch, you might need to temporarily remove a node from cluster membership or stop the entire cluster before installing the patch. This section describes these steps.

Before applying a Sun Cluster patch, check the Sun Cluster web page for any special instructions; for the current URL, see the *Sun Cluster 3.0 U1 Release Notes* or contact Enterprise Services. If there aren't any special instructions, check the patch's README file.

Note – For Sun Cluster-specific patches, always defer to the patch's README file for instructions that supersede procedures in this chapter.

Patch installation on all cluster nodes falls into one of the following scenarios:

- **Rebooting patch (node)**—A node must be booted to single-user mode, using the command `boot -sx`, before the patch can be applied, then rebooted to join the cluster. In doing so, you need to put the node into a "quiet" state by first switching any resource groups or disk device groups from the node to be patched to another cluster member. Also, apply the patch to one cluster node at a time to avoid bringing down the entire cluster.

  The cluster itself remains available during this type of patch application, even though individual nodes are temporarily shut down. A patched node is able to rejoin a cluster as a member node even though other nodes are not yet at the same patch level.

- **Rebooting patch (cluster and firmware)**—The cluster must be shut down and each node must be booted to single-user mode, using the command `boot -sx`, to apply the software or firmware patch. Then, reboot the nodes to rejoin the cluster. For this type of patch, the cluster is unavailable during patch application.

- **Non-rebooting patch**—A node does not have to be in a "quiet" state (it can still be mastering resource groups or device groups), nor does it have to be shut down or rebooted when applying the patch. However, you should still apply the patch to one node at a time and verify that the patch works before patching another node.

Note – Underlying cluster protocols do not change due to a patch.

You use the `patchadd` command to apply a patch to the cluster, and `patchrm` to remove a patch (when possible).

Sun Cluster Patch Tips

Use the following tips to help you administer Sun Cluster patches more efficiently:

- Refer to the Sun Cluster website for any special instructions associated with the patch or firmware update. For the current URL, see the *Sun Cluster 3.0 U1 Release Notes* or contact Enterprise Services.
- Always read the patch README file before applying the patch.
- Apply all patches (required and recommended) before running the cluster in a production environment.
- Check the hardware firmware levels and install any required firmware updates that may be needed.
- All nodes acting as cluster members must have the same patches.
- Keep cluster subsystem patches up to date. This includes volume management, storage device firmware, cluster transport, and so forth.
- Review patch reports regularly, such as once a quarter, and patch a Sun Cluster configuration using the recommended patch suite.

- Apply selective patches as recommended by Enterprise Services.
- Test failover after major patch updates; be prepared to back out the patch if cluster operation is degraded or impaired.

Patching Sun Cluster

TABLE 7-1   Task Map: Patching the Cluster

Task	For Instructions, Go To...
Apply a non-rebooting Sun Cluster patch to one node at a time without having to stop the node	"How to Apply a Non-Rebooting Sun Cluster Patch" on page 133
Apply a rebooting Sun Cluster patch after taking the cluster member to non-cluster mode	"How to Apply a Rebooting Patch (Node)" on page 127 "How to Apply a Rebooting Patch (Cluster and Firmware)" on page 131
Remove a Sun Cluster patch - You can back out the patch if necessary	"How to Remove a Sun Cluster Patch" on page 134

▼ How to Apply a Rebooting Patch (Node)

Apply the patch to one node in the cluster at a time to keep the cluster itself operational during the patch process. With this procedure, you must first shut down the node and boot it to single-user mode using the `boot -sx` command, before applying the patch.

1. **Before applying the patch, check the Sun Cluster product web page for any special pre- or post-installation instructions.**

2. **Become superuser on the node to which you are applying the patch.**

3. **List the resource groups and device groups on the node being patched.**

```
# scrgadm -pv
# scstat
```

4. **Switch all resource groups, resources, and device groups from the node being patched to other cluster members.**

```
# scswitch -S -h nodelist
```

-S   Evacuates all device groups and resource groups from the specified node.

-h *nodelist*   Specifies the node from which you are switching the resource groups and device groups.

5. **Shut down the node.**

```
# shutdown -g0 [-y] [-i0]
```

6. **Boot the node in non-cluster single-user mode.**

```
ok boot -sx
```

7. **Apply the patch.**

```
# patchadd -M patch-dir patch-id
```

patch-dir   Specifies the directory location of the patch.

patch-id   Specifies the patch number of a given patch.

Note – Always defer to the instructions in the patch directory that supersede procedures in this chapter.

---

8. **Reboot the node into the cluster.**

```
# reboot
```

9. **Verify that the patch has been installed successfully.**

```
# showrev -p | grep patch-id
```

10. **Verify that the patch works, and that the node and cluster are operating normally.**

11. **Repeat Step 2 through Step 10 for all remaining cluster nodes.**

12. **Switch resource groups, resources, and device groups as needed.**

    After rebooting all the nodes, the last node rebooted will not have the resource groups and device groups online.

    ```
    # scswitch -z -h nodelist -D device-group
    # scswitch -z -h nodelist -g resource-group
    ```

-z	Specifies the change in mastery of a resource group or device group.
-h *nodelist*	Specifies the nodes to which you are switching the resource groups and device groups.
-D	Switches the specified device groups to the nodes identified by the -h option.
-g	Switches the specified resource groups to the nodes identified by the -h option. If -h is not specified, the resource groups are taken offline.

Example—Applying a Rebooting Patch (Node)

The following example shows the application of a rebooting Sun Cluster patch to a node.

```
# scrgadm -pv
...
RG Name: schost-sa-1
...
# scstat
...
Device Group Name:                         dg-schost-1
...
# scswitch -S -h phys-schost-2
# shutdown -g0 -y -i0
...
ok boot -sx
...
# patchadd -M /tmp/patches 234567-05
...
# reboot
...
# showrev -p | grep 234567-05
# scswitch -z -h phys-schost-1 -D dg-schost-1
# scswitch -z -h phys-schost-1 -g schost-sa-1
```

Where to Go From Here

If you need to back out a patch, see "How to Remove a Sun Cluster Patch" on page 134.

▼ How to Apply a Rebooting Patch (Cluster and Firmware)

With this procedure, you must first shut down the cluster and boot each node to single-user mode using the `boot -sx` command, before applying the patch.

1. **Before applying the patch, check the Sun Cluster product web page for any special pre- or post-installation instructions.**

2. **Become superuser on any node in the cluster.**

3. **Shut down the cluster.**

```
# scshutdown -y -g grace-period "message"
```

`-y`	Specifies to answer *yes* to the confirmation prompt.
`-g` *grace-period*	Specifies, in seconds, the amount of time to wait before shutting down. Default grace period is 60 seconds.
message	Specifies the warning message to broadcast. Use quotes if *message* contains multiple words.

4. **Boot each node into non-cluster single-user mode.**

   On the console of each node, run the following command.

```
ok boot -sx
```

5. **Apply the software or firmware patch.**

   On one node at a time, run the following command.

```
# patchadd -M patch-dir patch-id
```

patch-dir	Specifies the directory location of the patch.
patch-id	Specifies the patch number of a given patch.

Note – Always defer to the instructions in the patch directory that supersede procedures in this chapter.

---

6. **Verify that the patch has been installed successfully on each node.**

```
# showrev -p | grep patch-id
```

7. **After applying the patch to all nodes, reboot the nodes into the cluster.**

   On each node, run the following command.

```
# reboot
```

8. **Verify that the patch works, and that the nodes and cluster are operating normally.**

Example—Applying a Rebooting Patch (Cluster)

The following example shows the application of a rebooting Sun Cluster patch to a cluster.

```
# scshutdown -g 0 -y
...
ok boot -sx
...
# patchadd -M /tmp/patches 234567-05
(Apply patch to other cluster nodes)
...
# showrev -p | grep 234567-05
# reboot
```

Where to Go From Here

If you need to back out a patch, see "How to Remove a Sun Cluster Patch" on page 134.

▼ How to Apply a Non-Rebooting Sun Cluster Patch

Apply the patch to one node in the cluster at a time. When applying a non-rebooting patch, you do not need to first shut down the node receiving the patch.

1. **Before applying the patch, check the Sun Cluster product web page for any special pre- or post-installation instructions.**

2. **Apply the patch on a single node.**

   ```
   # patchadd -M patch-dir patch-id
   ```

   patch-dir   Specifies the directory location of the patch.

   patch-id   Specifies the patch number of a given patch.

3. **Verify that the patch has been installed successfully.**

   ```
   # showrev -p | grep patch-id
   ```

4. **Verify that the patch works, and that the node and cluster are operating normally.**

5. **Repeat Step 2 through Step 4 for the remaining cluster nodes.**

Example—Applying a Non-Rebooting Sun Cluster Patch

```
# patchadd -M /tmp/patches 234567-05
...
# showrev -p | grep 234567-05
```

Where to Go From Here

If you need to back out a patch, see "How to Remove a Sun Cluster Patch" on page 134.

▼ How to Remove a Sun Cluster Patch

If necessary, you can back out (remove) a Sun Cluster patch.

1. **Become superuser on the node from which you are removing the patch.**

2. **List the resource groups and device groups on the node having the patch removed.**

   ```
   # scrgadm -pv
   # scstat
   ```

3. **Switch all resource groups, resources, and device groups from the node having the patch removed to other cluster members.**

   ```
   # scswitch -S -h nodelist
   ```

   -S   Evacuates all device services and resource groups from the
                       specified node.

   -h *nodelist*   Specifies the nodes from which you are switching the resource
                       groups and device groups.

4. **Shut down the node.**

   ```
   # shutdown -g0 -y -i0 "message"
   ```

   -g0   Specifies, in seconds, the amount of time to wait before shutting
                       down. Default grace period is 60 seconds.

   -y   Specifies to answer *yes* to the confirmation prompt.

   -i0   Specifies init state of 0. This brings the node down to the OBP
                       prompt.

   message   Specifies the warning message to broadcast. Use quotes if
                       message contains multiple words.

5. **Boot the node in non-cluster single-user mode.**

   ```
   ok boot -sx
   ```

6. **Remove the patch.**

```
# patchrm patch-id
```

patch-id   Specifies the patch number of a given patch.

7. **Reboot the node.**

```
# reboot
```

8. **Verify that the patch has been removed successfully.**

```
# showrev -p | grep patch-id
```

9. **Verify that the node and cluster are operating normally.**

10. **Repeat Step 1 through Step 9 for the remaining cluster nodes.**

11. **Switch resource groups, resources, and device groups as needed (optional).**

    After rebooting all the nodes, the last node rebooted will not have the resource groups and device groups online.

```
# scswitch -z -h nodelist -D device-group
# scswitch -z -h nodelist -g resource-group
```

-z   Specifies the change in mastery of a resource group or device group.

-h *nodelist*   Specifies the nodes to which you are switching the resource groups and device groups.

-D   Switches the specified device groups to the nodes identified by the -h option.

-g   Switches the specified resource groups to the nodes identified by the -h option. If -h is not specified, the resource groups are taken offline.

Example—Removing a Sun Cluster Patch

The following example shows the removal of a Sun Cluster patch.

```
# scrgadm -pv
...
RG Name: schost-sa-1
...
# scstat
...
Device Group Name:                              dg-schost-1
...
# scswitch -S -h phys-schost-2
# shutdown -g0 -y -i0 "Rebooting down node for maintenance"
...
ok boot -x
...
# patchrm 234567-05
...
# reboot
...
# pkgchk -v 234567-05
...
# scswitch -z -h phys-schost-1 -D dg-schost-1
# scswitch -z -h phys-schost-1 -g schost-sa-1
```

Backing Up and Restoring a Cluster

This is a list of step-by-step instructions in this chapter.

Backing Up a Cluster

TABLE 8-1   Task Map: Backing Up Cluster Files

Task	For Instructions, Go To…
Find the names of the file systems you want to back up	
Calculate how many tapes you will need to contain a full backup	
Back up the root file system	
Perform online backup for mirrored or plexed file systems	

▼ How to Find File System Names to Back Up

Use this procedure to determine the names of the file systems you want to back up.

1. **Display the contents of the** `/etc/vfstab` **file.**

   You do not need to be superuser to run this command.

   ```
   % more /etc/vfstab
   ```

2. **Look in the mount point column for the name of the file system you want to back up.**

   Use this name when you back up the file system.

   ```
   % more /etc/vfstab
   ```

Example—Finding File System Names to Back Up

In the following example, the names of available file systems listed in the `/etc/vfstab` file are displayed.

```
% more /etc/vfstab
#device                device                mount  FS fsck  mount  mount
#to mount              to fsck               point  type  pass  at boot  options
#
#/dev/dsk/c1d0s2       /dev/rdsk/c1d0s2      /usr   ufs    1     yes      -
 f                     -                     /dev/fd fd    -     no       -
 /proc                 -                     /proc  proc   -     no       -
 /dev/dsk/c1t6d0s1     -                     -      swap   -     no       -
 /dev/dsk/c1t6d0s0     /dev/rdsk/c1t6d0s0    /      ufs    1     no       -
 /dev/dsk/c1t6d0s3     /dev/rdsk/c1t6d0s3    /cache ufs    2     yes      -
 swap                  -                     /tmp   tmpfs  -     yes      -
```

▼ How to Determine the Number of Tapes Needed for a Full Backup

Use this procedure to calculate the number of tapes you will need to back up a file system.

1. **Become superuser on the cluster node you want to back up.**

2. **Estimate the size of the backup in bytes.**

   ```
   # ufsdump S filesystem
   ```

   S   Displays the estimated number of bytes needed to perform the backup.

   filesystem   Specifies the name of the file system you want to back up.

3. **Divide the estimated size by the capacity of the tape to see how many tapes you need.**

Example—Determining the Number of Tapes Needed

In the following example, the file system size of 905,881,620 bytes will easily fit on a 4 GB tape (905,881,620 ÷ 4,000,000,000).

```
# ufsdump S /global/phys-schost-1
905881620
```

▼ How to Back Up the root (/) File System

Use this procedure to back up the root (/) file system of a cluster node. Be sure the cluster is running problem-free before performing the backup procedure.

1. **Become superuser on the cluster node you want to back up.**

2. **Switch each running data service from the node to be backed up to another node in the cluster.**

```
# scswitch -z -D disk-device-group -h nodelist
```

-z	Performs the switch.
-D *disk-device-group*	Name of the disk device group to be switched.
-h *nodelist*	Name of the cluster node to switch the disk device group to. This node becomes the new primary.

3. **Stop the node.**

```
# shutdown -g0 -y -i0
```

4. **At the** ok **prompt, reboot in non-cluster mode.**

```
ok boot -x
```

5. **Back up the root (/) file system.**

   - If the root disk is not encapsulated, use the following command.

```
# ufsdump 0ucf dump-device /
```

■ If the root disk is encapsulated, use the following command.

```
# ufsdump 0ucf dump-device /dev/vx/rdsk/rootvol
```

Refer to the ufsdump(1M) man page for more information.

6. **Reboot the node in cluster mode.**

```
# init 6
```

Example—Backing Up the root (/) File System

In the following example, the root (/) file system is backed up onto tape device /dev/rmt/0.

```
# ufsdump 0ucf /dev/rmt/0 /
  DUMP: Writing 63 Kilobyte records
  DUMP: Date of this level 0 dump: Tue Apr 18 18:06:15 2000
  DUMP: Date of last level 0 dump: the epoch
  DUMP: Dumping /dev/rdsk/c0t0d0s0 (phys-schost-1:/) to /dev/rmt/0
  DUMP: Mapping (Pass I) [regular files]
  DUMP: Mapping (Pass II) [directories]
  DUMP: Estimated 859086 blocks (419.48MB).
  DUMP: Dumping (Pass III) [directories]
  DUMP: Dumping (Pass IV) [regular files]
  DUMP: 859066 blocks (419.47MB) on 1 volume at 2495 KB/sec
  DUMP: DUMP IS DONE
  DUMP: Level 0 dump on Tue Apr 18 18:06:15 2000
```

▼ How to Perform Online Backups for Mirrors (Solstice DiskSuite)

A mirrored metadevice can be backed up without unmounting it or taking the entire mirror offline. One of the submirrors must be taken offline temporarily, thus losing mirroring, but it can be placed online and resynced as soon as the backup is complete, without halting the system or denying user access to the data. Using mirrors to perform online backups creates a backup that is a "snapshot" of an active file system.

A problem might occur if a program writes data onto the volume immediately
before the `lockfs` command is run. To prevent this problem, temporarily stop all
the services running on this node. Also, be sure the cluster is running problem-free
before performing the backup procedure.

1. **Become superuser on the cluster node you want to back up.**

2. **Use the `metaset`(1M) command to determine which node has the ownership on
   the backed up volume.**

   ```
   # metaset -s setname
   ```

   `-s` *setname*   Specifies the diskset name.

3. **Use the `lockfs`(1M) command with the `-w` option to lock the file system from
   writes.**

   ```
   # lockfs -w mount-point
   ```

   Note – You must lock the file system only if a UFS file system resides on the mirror.
   For example, if the metadevice is set up as a raw device for database management
   software or some other specific application, it would not be necessary to use the
   `lockfs` command. You may, however, want to run the appropriate vender-
   dependent utility to flush any buffers and lock access.

   ---

4. **Use the `metastat`(1M) command to determine the names of the submirrors.**

   ```
   # metastat -s setname -p
   ```

   `-p`   Displays the status in a format similar to the `md.tab` file.

5. **Use the `metadetach`(1M) command to take one submirror offline from the mirror.**

   ```
   # metadetach -s setname mirror submirror
   ```

Note – Reads will continue to be made from the other submirrors. However, the offline submirror will be out of sync as soon as the first write is made to the mirror. This inconsistency is corrected when the offline submirror is brought back online. You don't need to run fsck.

6. **Unlock the file systems and allow writes to continue, using the** lockfs **command with the** -u **option.**

   ```
   # lockfs -u mount-point
   ```

7. **Perform a file system check.**

   ```
   # fsck /dev/md/diskset/rdsk/submirror
   ```

8. **Back up the offline submirror to tape or another medium.**

   Use the ufsdump(1M) command or whatever other backup utility you normally use.

   ```
   # ufsdump 0ucf dump-device submirror
   ```

   Note – Use the raw device (/rdsk) name for the submirror, rather than the block device (/dsk) name.

9. **Use the** metattach**(1M) command to place the metadevice back online.**

   ```
   # metattach -s setname mirror submirror
   ```

   When the metadevice is placed online, it is automatically resynced with the mirror.

10. **Use the** metastat **command to verify that the submirror is resyncing.**

    ```
    # metastat -s setname mirror
    ```

Example—Performing Online Backups for Mirrors (Solstice DiskSuite)

In the following example, the cluster node `phys-schost-1` is the owner of the metaset `schost-1`, therefore the backup procedure is performed from `phys-schost-1`. The mirror `/dev/md/schost-1/dsk/d0` consists of the submirrors `d10`, `d20`, and `d30`.

```
[Determine the owner of the metaset:]
# metaset -s schost-1
Set name = schost-1, Set number = 1
Host                    Owner
  phys-schost-1      Yes
...

[Lock the file system from writes:]
# lockfs -w /global/schost-1

[List the submirrors:]
# metastat -s schost-1 -p
schost-1/d0 -m schost-1/d10 schost-1/d20 schost-1/d30 1
schost-1/d10 1 1 d4s0
schost-1/d20 1 1 d6s0
schost-1/d30 1 1 d8s0

[Take a submirror offline:]
# metadetach -s schost-1 d0 d30

[Unlock the file system:]
# lockfs -u /

[Check the file system:]
# fsck /dev/md/schost-1/rdsk/d30

[Copy the submirror to the backup device:]
# ufsdump 0ucf /dev/rmt/0 /dev/md/schost-1/rdsk/d30
  DUMP: Writing 63 Kilobyte records
  DUMP: Date of this level 0 dump: Tue Apr 25 16:15:51 2000
  DUMP: Date of last level 0 dump: the epoch
  DUMP: Dumping /dev/md/schost-1/rdsk/d30 to /dev/rdsk/c1t9d0s0.
  ...
  DUMP: DUMP IS DONE
```

```
[Bring the submirror back online:]
# metattach -s schost-1 d0 d30
schost-1/d0: submirror schost-1/d30 is attached

[Resync the submirror:]
# metastat -s schost-1 d0
schost-1/d0: Mirror
    Submirror 0: schost-0/d10
      State: Okay
    Submirror 1: schost-0/d20
      State: Okay
    Submirror 2: schost-0/d30
      State: Resyncing
    Resync in progress: 42% done
    Pass: 1
    Read option: roundrobin (default)
...
```

▼ How to Perform Online Backups for Volumes (VERITAS Volume Manager)

VERITAS Volume Manager identifies a mirrored volume as a plex. A plex can be backed up without unmounting it or taking the entire volume offline. This is done by creating a snapshot copy of the volume and backing up this temporary volume without halting the system or denying user access to the data.

Be sure the cluster is running problem-free before performing the backup procedure.

1. **Log on to any node in the cluster, and become superuser on the current primary node for the disk group on the cluster.**

2. **List the disk group information.**

   ```
   # vxprint -g diskgroup
   ```

3. **Run the scstat(1M) command to see which node has the disk group currently imported, indicating it is the primary node for the disk group.**

   ```
   # scstat -D
   ```

   -D   Shows the status for all disk device groups.

4. **Create a snapshot of the volume using the** vxassist**(1M) command.**

```
# vxassist -g diskgroup snapstart volume
```

Note – Creating a snapshot can take a long time depending on the size of your volume.

---

5. **Verify the new volume was created.**

```
# vxprint -g diskgroup
```

When the snapshot is complete, a status of Snapdone displays in the State field for the selected disk group.

6. **Stop any data services that are accessing the file system.**

```
# scswitch -z -g resource-group -h ""
```

Note – Stop all data services to ensure that the data file system is properly backed up. If no data services are running, you do not need to perform Step 6 and Step 8.

---

7. **Create a backup volume named** bkup-vol **and attach the snapshot volume to it using the** vxassist **command.**

```
# vxassist -g diskgroup snapshot volume bkup-vol
```

8. **Restart any data services that were stopped in Step 6, using the** scswitch **command.**

```
# scswitch -z -g resource-group -h nodelist
```

9. **Verify the volume is now attached to the new volume** bkup-vol **using the** vxprint **command.**

```
# vxprint -g diskgroup
```

10. **Register the disk group configuration change.**

   ```
   # scconf -c -D name=diskgroup,sync
   ```

11. **Check the backup volume using the** `fsck` **command.**

   ```
   # fsck -y /dev/vx/rdsk/diskgroup/bkup-vol
   ```

12. **Perform a backup to copy the volume** `bkup-vol` **to tape or another medium.**

   Use the `ufsdump`(1M) command or the backup utility you normally use.

   ```
   # ufsdump 0ucf dump-device /dev/vx/dsk/diskgroup/bkup-vol
   ```

13. **Remove the temporary volume using** `vxedit`**(1M).**

   ```
   # vxedit -rf rm bkup-vol
   ```

14. **Register the disk group configuration changes using the** `scconf` **command.**

   ```
   # scconf -c -D name=diskgroup,sync
   ```

Example—Performing Online Backups for Volumes (VERITAS Volume Manager)

In the following example, the cluster node `phys-schost-2` is the primary owner of the metaset disk group `schost-1`, therefore the backup procedure is performed from `phys-schost-2`. The volume `/vo101` is copied and then associated with a new volume, `bkup-vol`.

```
[Become superuser on the primary node.]
```

```
[Identify the current primary node for the disk group:]
# scstat -D
-- Device Group Servers --
                        Device Group      Primary           Secondary
                        ------------      -------           ---------
 Device group servers:  rmt/1             -                 -
 Device group servers:  schost-1          phys-schost-2     phys-schost-1

-- Device Group Status --
                        Device Group      Status
                        ------------      ------
 Device group status:   rmt/1             Offline
 Device group status:   schost-1          Online

[List the disk group information:]
# vxprint -g schost-1
TY NAME            ASSOC       KSTATE    LENGTH    PLOFFS STATE    TUTIL0   PUTIL0
dg schost-1        schost-1    -         -         -      -        -        -

dm schost-101      c1t1d0s2    -         17678493  -      -        -        -
dm schost-102      c1t2d0s2    -         17678493  -      -        -        -
dm schost-103      c2t1d0s2    -         8378640   -      -        -        -
dm schost-104      c2t2d0s2    -         17678493  -      -        -        -
dm schost-105      c1t3d0s2    -         17678493  -      -        -        -
dm schost-106      c2t3d0s2    -         17678493  -      -        -        -

v  vol01           gen         ENABLED   204800    -      ACTIVE   -        -
pl vol01-01        vol01       ENABLED   208331    -      ACTIVE   -        -
sd schost-101-01   vol01-01    ENABLED   104139    0      -        -        -
sd schost-102-01   vol01-01    ENABLED   104139    0      -        -        -
pl vol01-02        vol01       ENABLED   208331    -      ACTIVE   -        -
sd schost-103-01   vol01-02    ENABLED   103680    0      -        -        -
sd schost-104-01   vol01-02    ENABLED   104139    0      -        -        -
pl vol01-03        vol01       ENABLED   LOGONLY   -      ACTIVE   -        -
sd schost-103-02   vol01-03    ENABLED   5         LOG    -        -        -

[Start the snapshot operation:]
# vxassist -g schost-1 snapstart vol01
```

```
[Verify the new volume was created:]
# vxprint -g schost-1
TY NAME            ASSOC      KSTATE     LENGTH    PLOFFS STATE     TUTIL0   PUTIL0
dg schost-1        schost-1   -          -         -      -         -        -

dm schost-101      c1t1d0s2   -          17678493  -      -         -        -
dm schost-102      c1t2d0s2   -          17678493  -      -         -        -
dm schost-103      c2t1d0s2   -          8378640   -      -         -        -
dm schost-104      c2t2d0s2   -          17678493  -      -         -        -
dm schost-105      c1t3d0s2   -          17678493  -      -         -        -
dm schost-106      c2t3d0s2   -          17678493  -      -         -        -

v  vol01           gen        ENABLED    204800    -      ACTIVE    -        -
pl vol01-01        vol01      ENABLED    208331    -      ACTIVE    -        -
sd schost-101-01   vol01-01   ENABLED    104139    0      -         -        -
sd schost-102-01   vol01-01   ENABLED    104139    0      -         -        -
pl vol01-02        vol01      ENABLED    208331    -      ACTIVE    -        -
sd schost-103-01   vol01-02   ENABLED    103680    0      -         -        -
sd schost-104-01   vol01-02   ENABLED    104139    0      -         -        -
pl vol01-03        vol01      ENABLED    LOGONLY   -      ACTIVE    -        -
sd schost-103-02   vol01-03   ENABLED    5         LOG    -         -        -
pl vol01-04        vol01      ENABLED    208331    -      SNAPDONE  -        -
sd schost-105-01   vol01-04   ENABLED    104139    0      -         -        -
sd schost-106-01   vol01-04   ENABLED    104139    0      -         -        -


[Stop data services, if necessary:]
# scswitch -z -g nfs-rg -h ""


[Create a copy of the volume:]
# vxassist -g schost-1 snapshot vol01 bkup-vol


[Restart data services, if necessary:]
# scswitch -z -g nfs-rg -h phys-schost-1
```

```
[Verify bkup-vol was created:]
# vxprint -g schost-1
TY NAME            ASSOC         KSTATE    LENGTH    PLOFFS STATE    TUTIL0  PUTIL0
dg schost-1        schost-1      -         -         -      -        -       -

dm schost-101      c1t1d0s2      -         17678493  -      -        -       -
...

v  bkup-vol        gen           ENABLED   204800    -      ACTIVE   -       -
pl bkup-vol-01     bkup-vol      ENABLED   208331    -      ACTIVE   -       -
sd schost-105-01   bkup-vol-01   ENABLED   104139    0      -        -       -
sd schost-106-01   bkup-vol-01   ENABLED   104139    0      -        -       -

v  vol01           gen           ENABLED   204800    -      ACTIVE   -       -
pl vol01-01        vol01         ENABLED   208331    -      ACTIVE   -       -
sd schost-101-01   vol01-01      ENABLED   104139    0      -        -       -
sd schost-102-01   vol01-01      ENABLED   104139    0      -        -       -
pl vol01-02        vol01         ENABLED   208331    -      ACTIVE   -       -
sd schost-103-01   vol01-02      ENABLED   103680    0      -        -       -
sd schost-104-01   vol01-02      ENABLED   104139    0      -        -       -
pl vol01-03        vol01         ENABLED   LOGONLY   -      ACTIVE   -       -
sd schost-103-02   vol01-03      ENABLED   5         LOG    -        -       -


[Synchronize the disk group with cluster framework:]
# scconf -c -D name=schost-1,sync


[Check the file systems:]
# fsck -y /dev/vx/rdsk/schost-1/bkup-vol


[Copy bkup-vol to the backup device:]
# ufsdump 0ucf /dev/rmt/0 /dev/vx/rdsk/schost-1/bkup-vol
  DUMP: Writing 63 Kilobyte records
  DUMP: Date of this level 0 dump: Tue Apr 25 16:15:51 2000
  DUMP: Date of last level 0 dump: the epoch
  DUMP: Dumping /dev/vx/dsk/schost-2/bkup-vol to /dev/rmt/0.
  ...
  DUMP: DUMP IS DONE


[Remove the bkup-volume:]
# vxedit -rf rm bkup-vol


[Synchronize the disk group:]
# scconf -c -D name=schost-1,sync
```

Restoring Cluster Files Overview

The `ufsrestore` command copies files to disk, relative to the current working directory, from backups created using the `ufsdump` command. You can use `ufsrestore` to reload an entire file system hierarchy from a level 0 dump and incremental dumps that follow it, or to restore one or more single files from any dump tape. If `ufsrestore` is run as superuser, files are restored with their original owner, last modification time, and mode (permissions).

Before you start to restore files or file systems, you need to know the following information.

- Which tapes you need
- The raw device name on which you want to restore the file system
- The type of tape drive you will use
- The device name (local or remote) for the tape drive
- The partition scheme of any failed disk, because the partitions and file systems must be exactly duplicated on the replacement disk

Restoring Cluster Files

TABLE 8-2   Task Map: Restoring Cluster Files

Task	For Instructions, Go To…
For Solstice DiskSuite, restore files interactively following Solaris restore procedures	"How to Restore Individual Files Interactively (Solstice DiskSuite)" on page 152
For Solstice DiskSuite, restore the root (/) file system	"How to Restore the root (/) File System (Solstice DiskSuite)" on page 153
	"How to Restore a root (/) File System That Was on a Metadevice (Solstice DiskSuite)" on page 157
For VERITAS Volume Manager, restore a non-encapsulated root (/) file system	"How to Restore a Non-Encapsulated root (/) File System (VERITAS Volume Manager)" on page 162
For VERITAS Volume Manager, restore an encapsulated root (/) file system	"How to Restore an Encapsulated root (/) File System (VERITAS Volume Manager)" on page 166

▼ How to Restore Individual Files Interactively (Solstice DiskSuite)

Use this procedure to restore one or more individual files. Be sure the cluster is running problem-free before performing the restore procedure.

1. Become superuser on the cluster node you want to restore.

2. Stop all the data services that are using the files to be restored.

```
# scswitch -z -g resource-group -h ""
```

3. Restore the files using the `ufsrestore` **command.**

▼ How to Restore the root (/) File System (Solstice DiskSuite)

Use this procedure to restore the root (/) file systems to a new disk, such as after replacing a bad root disk. The node being restored should not be booted. Be sure the cluster is running problem-free before performing the restore procedure.

Note – Since you must partition the new disk using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and recreate file systems as appropriate.

---

1. **Become superuser on a cluster node with access to the metaset,** *other than* **the node you want to restore.**

2. **Remove from all metasets the hostname of the node being restored.**

   Run this command from a node in the metaset other than the node you are removing.

   ```
   # metaset -s setname -f -d -h nodelist
   ```

-s setname	Specifies the diskset name.
-f	Force.
-d	Deletes from the diskset.
-h *nodelist*	Specifies the name of the node to delete from the diskset.

3. **Replace the failed disk on the node on which the root (/) file system will be restored.**

   Refer to disk replacement procedures in the documentation that came with your server.

4. **Boot the node being restored.**

   - If using the Solaris CD-ROM, run the following command:

   ```
   ok boot cdrom -s
   ```

   - If using a JumpStart™ server, run the following command:

   ```
   ok boot net -s
   ```

5. **Create all the partitions and swap on the root disk using the** `format`**(1M) command.**

   Recreate the original partitioning scheme that was on the failed disk.

6. **Create the root (/) file system and other file systems as appropriate, using the** `newfs`**(1M) command.**

   Recreate the original file systems that were on the failed disk.

   Note – Be sure to create the `/global/.devices/node@nodeid` file system.

   ---

7. **Mount the root (/) file system on a temporary mount point.**

   ```
   # mount device temp-mount-point
   ```

8. **Use the following commands to restore the root (/) file system.**

   ```
   # cd temp-mount-point
   # ufsrestore rvf dump-device
   # rm restoresymtable
   # cd /
   # umount temp-mount-point
   # fsck raw-disk-device
   ```

   The file system is now restored.

9. **Install a new boot block on the new disk.**

   ```
   # /usr/sbin/installboot /usr/platform/'uname -i'/lib/fs/ufs/bootblk raw-disk-device
   ```

10. **Reboot the node in single-user mode.**

    ```
    # reboot -- "-s"
    ```

11. **Replace the disk ID using the** `scdidadm` **command.**

    ```
    # scdidadm -R rootdisk
    ```

12. **Use the** `metadb`**(1M) command to recreate the state database replicas.**

```
# metadb -c copies -af raw-disk-device
```

-c *copies*   Specifies the number of replicas to create.

-f *raw-disk-device* Raw disk device on which to create replicas.

-a   Adds replicas.

13. **Reboot the node in cluster mode.**

 a. Start the reboot.

```
# reboot
```

 During this boot you might see an error or warning message, ending with the
 following instruction:

```
Type control-d to proceed with normal startup,
(or give root password for system maintenance):
```

 b. Press CTRL-d to boot into multiuser mode.

14. **From a cluster node other than the restored node, use the** `metaset`**(1M) command
to add the restored node to all metasets.**

```
phys-schost-2# metaset -s setname -a -h nodelist
```

-a   Creates and adds the host to the diskset.

The node is rebooted into cluster mode. The cluster is ready to use.

Example—Restoring the root (/) File System (Solstice DiskSuite)

The following example shows the root (/) file system restored to the node phys-schost-1 from the tape device /dev/rmt/0. The metaset command is run from another node in the cluster, phys-schost-2, to remove and later add back node

phys-schost-1 to the diskset schost-1. All other commands are run from phys-schost-1. A new boot block is created on /dev/rdsk/c0t0d0s0, and three state database replicas are recreated on /dev/rdsk/c0t0d0s4.

```
[Become superuser on a cluster node other than the node to be
restored.]

[Remove the node from the metaset:]
phys-schost-2# metaset -s schost-1 -f -d -h phys-schost-1

[Replace the failed disk and boot the node:]
ok boot cdrom -s

[Use format and newfs to recreate partitions and file systems.]

[Mount the root file system on a temporary mount point:]
# mount /dev/dsk/c0t0d0s0 /a

[Restore the root file system:]
# cd /a
# ufsrestore rvf /dev/rmt/0
# rm restoresymtable
# cd /
# umount /a
# fsck /dev/rdsk/c0t0d0s0

[Install a new boot block:]
# /usr/sbin/installboot /usr/platform/`uname \
-i`/lib/fs/ufs/bootblk /dev/rdsk/c0t0d0s0

[Reboot in single-user mode:]
# reboot -- "-s"

[Replace the disk ID:]
# scdidadm -R /dev/dsk/c0t0d0

[Recreate state database replicas:]
# metadb -c 3 -af /dev/rdsk/c0t0d0s4

# reboot
Press CTL-d to boot into multiuser mode.

[Add the node back to the metaset:]
phys-schost-2# metaset -s schost-1 -a -h phys-schost-1
```

▼ How to Restore a root (/) File System That Was on a Metadevice (Solstice DiskSuite)

Use this procedure to restore a root (/) file system that was on a metadevice when the backups were performed. Perform this procedure under circumstances such as when a root disk is corrupted and replaced with a new disk. The node being restored should not be booted. Be sure the cluster is running problem-free before performing the restore procedure.

Note – Since you must partition the new disk using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and recreate file systems as appropriate.

---

1. **Become superuser on a cluster node with access to the metaset,** *other than* **the node you want to restore.**

2. **Remove from all metasets the hostname of the node being restored.**

   ```
   # metaset -s setname -f -d -h nodelist
   ```

-s *setname*	Specifies the metaset name.
-f	Force.
-d	Deletes from the metaset.
-h *nodelist*	Specifies the name of the node to delete from the metaset.

3. **Replace the failed disk on the node on which the root (/) file system will be restored.**

   Refer to disk replacement procedures in the documentation that came with your server.

4. **Boot the node being restored.**

   - If using the Solaris CD-ROM, run the following command:

   ```
   ok boot cdrom -s
   ```

   - If using a JumpStart server, run the following command:

   ```
   ok boot net -s
   ```

5. **Create all the partitions and swap on the root disk using the** `format`**(1M) command.**

   Recreate the original partitioning scheme that was on the failed disk.

6. **Create the root (/) file system and other file systems as appropriate, using the** `newfs`**(1M) command**

   Recreate the original file systems that were on the failed disk.

   Note – Be sure to create the `/global/.devices/node@nodeid` file system.

   ---

7. **Mount the root (/) file system on a temporary mount point.**

   ```
   # mount device temp-mount-point
   ```

8. **Use the following commands to restore the root (/) file system.**

   ```
   # cd temp-mount-point
   # ufsrestore rvf dump-device
   # rm restoresymtable
   ```

9. **Install a new boot block on the new disk.**

   ```
   # /usr/sbin/installboot /usr/platform/`uname -i`/lib/fs/ufs/bootblk raw-disk-device
   ```

10. **Remove the lines in the** */temp-mount-point*`/etc/system` **file for MDD root information.**

    ```
    * Begin MDD root info (do not edit)
    forceload: misc/md_trans
    forceload: misc/md_raid
    forceload: misc/md_mirror
    forceload: misc/md_hotspares
    forceload: misc/md_stripe
    forceload: drv/pcipsy
    forceload: drv/glm
    forceload: drv/sd
    rootdev:/pseudo/md@0:0,10,blk
    * End MDD root info (do not edit)
    ```

11. **Edit the** `/temp-mount-point/etc/vfstab` **file to change the root entry from a metadevice to a corresponding normal slice for each file system on the root disk that is part of the metadevice.**

```
Example:
Change from—
/dev/md/dsk/d10   /dev/md/rdsk/d10    /      ufs   1    no       -

Change to—
/dev/dsk/c0t0d0s0 /dev/rdsk/c0t0d0s0  /usr   ufs   1    no       -
```

12. **Unmount the temporary file system, and check the raw disk device.**

```
# cd /
# umount  temp-mount-point
# fsck  raw-disk-device
```

13. **Reboot the node in single-user mode.**

```
# reboot -- "-s"
```

14. **Replace the disk ID using the** `scdidadm` **command.**

```
# scdidadm -R  rootdisk
```

15. **Use the** `metadb`**(1M) command to recreate the state database replicas.**

```
# metadb -c  copies  -af  raw-disk-device
```

-c *copies*	Specifies the number of replicas to create.
-af *raw-disk-device*	Creates initial state database replicas on the named raw disk device.

16. **Reboot the node in cluster mode.**

a. **Start the reboot.**

```
# reboot
```

During this boot you will see error or warning messages, ending with the following instruction:

```
Type control-d to proceed with normal startup,
(or give root password for system maintenance):
```

b. **Press CTRL-d to boot into multiuser mode.**

17. **From a cluster node other than the restored node, use the** metaset**(1M) command to add the restored node to all metasets.**

```
phys-schost-2# metaset -s setname -a -h nodelist
```

-a   Adds (creates) the metaset.

Set up the metadevice/mirror for root (/) according to the Solstice DiskSuite documentation.

The node is rebooted into cluster mode. The cluster is ready to use.

Example—Restoring a root (/) File System That Was on a Metadevice (Solstice DiskSuite)

The following example shows the root (/) file system restored to the node phys-schost-1 from the tape device /dev/rmt/0. The metaset command is run from another node in the cluster, phys-schost-2, to remove and later add back node phys-schost-1 to the metaset schost-1. All other commands are run from phys-schost-1. A new boot block is created on /dev/rdsk/c0t0d0s0, and three state database replicas are recreated on /dev/rdsk/c0t0d0s4.

```
[Become superuser on a cluster node with access to the metaset,
other than the node to be restored.]

[Remove the node from the metaset:]
phys-schost-2# metaset -s schost-1 -f -d -h phys-schost-1
```

```
[Replace the failed disk and boot the node:]
ok boot cdrom -s

[Use format and newfs to recreate partitions and file systems.]

[Mount the root file system on a temporary mount point:]
# mount /dev/dsk/c0t0d0s0 /a

[Restore the root file system:]
# cd /a
# ufsrestore rvf /dev/rmt/0
# rm restoresymtable

[Install a new boot block:]
# /usr/sbin/installboot /usr/platform/'uname \
-i'/lib/fs/ufs/bootblk /dev/rdsk/c0t0d0s0

[Remove the lines in /temp-mount-point/etc/system file for MDD root
information:]
* Begin MDD root info (do not edit)
forceload: misc/md_trans
forceload: misc/md_raid
forceload: misc/md_mirror
forceload: misc/md_hotspares
forceload: misc/md_stripe
forceload: drv/pcipsy
forceload: drv/glm
forceload: drv/sd
rootdev:/pseudo/md@0:0,10,blk
* End MDD root info (do not edit)

[Edit the /temp-mount-point/etc/vfstab file]
Example:
Change from—
/dev/md/dsk/d10   /dev/md/rdsk/d10   /      ufs   1     no      -

Change to—
/dev/dsk/c0t0d0s0 /dev/rdsk/c0t0d0s0 /usr  ufs   1     no      -

[Unmount the temporary file system and check the raw disk device:]
# cd /
# umount /a
# fsck /dev/rdsk/c0t0d0s0
```

```
[Reboot in single-user mode:]
# reboot -- "-s"

[Replace the disk ID:]
# scdidadm -R /dev/dsk/c0t0d0

[Recreate state database replicas:]
# metadb -c 3 -af /dev/rdsk/c0t0d0s4

# reboot
Type CTRL-d to boot into multiuser mode.

[Add the node back to the metaset:]
phys-schost-2# metaset -s schost-1 -a -h phys-schost-1
```

▼ How to Restore a Non-Encapsulated root (/) File System (VERITAS Volume Manager)

Use this procedure to restore a non-encapsulated root (/) file system to a node. The node being restored should not be booted. Be sure the cluster is running problem-free before performing the restore procedure.

Note – Since you must partition the new disk using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and recreate file systems as appropriate.

---

1. **Replace the failed disk on the node where the root file system will be restored.**

   Refer to disk replacement procedures in the documentation that came with your server.

2. **Boot the node being restored.**

   - If using the Solaris CD-ROM, run the following command:

     ```
     ok boot cdrom -s
     ```

   - If using a JumpStart server, run the following command:

     ```
     ok boot net -s
     ```

3. **Create all the partitions and swap on the root disk using the** `format`**(1M) command.**

   Recreate the original partitioning scheme that was on the failed disk.

4. **Create the root (/) file system and other file systems as appropriate, using the** `newfs`**(1M) command.**

   Recreate the original file systems that were on the failed disk.

Note – Be sure to create the `/global/.devices/node@nodeid` file system.

---

5. **Mount the root (/) file system on a temporary mount point.**

   ```
   # mount device temp-mount-point
   ```

6. **Restore the root (/) file system from backup, and unmount and check the file system.**

   ```
   # cd temp-mount-point
   # ufsrestore rvf dump-device
   # rm restoresymtable
   # cd /
   # umount temp-mount-point
   # fsck raw-disk-device
   ```

   The file system is now restored.

7. **Install a new boot block on the new disk.**

   ```
   # /usr/sbin/installboot /usr/platform/'uname -i'/lib/fs/ufs/bootblk raw-disk-device
   ```

8. **Reboot the node into single-user mode.**

a. Start the reboot.

```
# reboot -- "-s"
```

During this boot you will see error or warning messages, ending with the following instruction:

```
Type control-d to proceed with normal startup,
(or give root password for system maintenance):
```

b. Type the root password.

9. **Determine if the root disk group is on a single slice on the root disk.**

   ■ If yes, create and set up the root disk group.

```
# vxdctl init
# vxdg init rootdg
# vxdctl add disk diskslice type=simple
# vxdisk -f init diskslice type=simple
# vxdg adddisk diskslice
# vxdctl enable
```

   ■ If no, proceed to Step 10.

10. **Update the disk ID using the** scdidadm **command.**

```
# scdidadm -R /dev/rdsk/disk-device
```

11. **Press CTRL-d to resume in multiuser mode.**

   The node reboots into cluster mode. The cluster is ready to use.

Example—Restoring a Non-Encapsulated root (/) File System (VERITAS Volume Manager)

The following example shows a non-encapsulated root (/) file system restored to the node phys-schost-1 from the tape device /dev/rmt/0.

```
[Replace the failed disk and boot the node:]
ok boot cdrom -s

[Use format and newfs to create partitions and file systems]

[Mount the root file system on a temporary mount point:]
# mount /dev/dsk/c0t0d0s0 /a

[Restore the root file system:]
# cd /a
# ufsrestore rvf /dev/rmt/0
# rm restoresymtable
# cd /
# umount /a
# fsck /dev/rdsk/c0t0d0s0

[Install a new boot block:]
# /usr/sbin/installboot /usr/platform/'uname \
-i'/lib/fs/ufs/bootblk /dev/rdsk/c0t0d0s0

[Reboot in single-user mode:]
# reboot -- "-s"

[If the root disk group is on a single slice on the root disk,
create the new root disk group:]
# vxdctl init
# vxdg init rootdg
# vxdctl add disk c0t0d0s4 type=simple
# vxdisk -f init c0t0d0s4 type=simple
# vxdg adddisk c0t0d0s4
# vxdctl enable

[Update the disk ID:]
# scdidadm -R /dev/rdsk/c0t0d0

[Press CTRL-d to resume in multiuser mode]
```

▼ How to Restore an Encapsulated root (/) File System (VERITAS Volume Manager)

Use this procedure to restore an encapsulated root (/) file system to a node. The node being restored should not be booted. Be sure the cluster is running problem-free before performing the restore procedure.

Note – Since you must partition the new disk using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and recreate file systems as appropriate.

---

1. **Replace the failed disk on the node where the root file system will be restored.**

   Refer to disk replacement procedures in the documentation that came with your server.

2. **Boot the node being restored.**

   ■ If using the Solaris CD-ROM, run the following command:

   ```
   ok boot cdrom -s
   ```

   ■ If using a JumpStart server, run the following command:

   ```
   ok boot net -s
   ```

3. **Create all the partitions and swap on the root disk using the** format**(1M) command.**

   Recreate the original partitioning scheme that was on the failed disk.

4. **Create the root (/) file system and other file systems as appropriate, using the** newfs**(1M) command.**

   Recreate the original file systems that were on the failed disk.

Note – Be sure to create the /global/.devices/*node@nodeid* file system.

---

5. **Mount the root (/) file system on a temporary mount point.**

   ```
   # mount device temp-mount-point
   ```

6. **Restore the root (/) file system from backup.**

```
# cd temp-mount-point
# ufsrestore rvf dump-device
# rm restoresymtable
```

7. **Create an empty** `install-db` **file.**

   This puts the node in VxVM install mode at the next reboot.

```
# touch /temp-mount-point/etc/vx/reconfig.d/state.d/install-db
```

8. **Remove or comment out the following entries from the** /temp-mount-point/etc/system **file.**

```
* rootdev:/pseudo/vxio@0:0
* set vxio:vol_rootdev_is_volume=1
```

9. **Edit the** /temp-mount-point/etc/vfstab **file and replace all VxVM mount points with the standard disk devices for the root disk, such as** /dev/dsk/c0t0d0s0**.**

```
Example:
Change from—
/dev/vx/dsk/rootdg/rootvol /dev/md/rdsk/rootdg/rootvol /     ufs    1     no -

Change to—
/dev/dsk/c0t0d0s0 /dev/rdsk/c0t0d0s0  / ufs   1     no        -
```

10. **Unmount the temporary file system and check the file system.**

```
# cd /
# umount temp-mount-point
# fsck raw-disk-device
```

11. **Install the boot block on the new disk.**

```
# /usr/sbin/installboot /usr/platform/`uname -i`/lib/fs/ufs/bootblk raw-disk-device
```

12. **Reboot the node in single-user mode.**

```
# reboot -- "-s"
```

13. **Update the disk ID using** scdidadm**(1M).**

```
# scdidadm -R /dev/rdsk/c0t0d0
```

14. **Run** vxinstall **to encapsulate the disk and reboot.**

```
# vxinstall
```

15. **If there is a conflict in minor number with any other system, unmount the global devices and reminor the disk group.**

   ■ Unmount the global devices file system on the cluster node.

```
# umount /global/.devices/node@nodeid
```

   ■ Reminor the rootdg disk group on the cluster node.

```
# vxdg reminor rootdg 100
```

16. **Shut down and reboot the node in cluster mode.**

```
# shutdown -g0 -i6 -y
```

Example—Restoring an Encapsulated root (/) File System (VERITAS Volume Manager)

The following example shows an encapsulated root (/) file system restored to the node phys-schost-1 from the tape device /dev/rmt/0.

```
[Replace the failed disk and boot the node:]
ok boot cdrom -s

[Use format and newfs to create partitions and file systems]
```

```
[Mount the root file system on a temporary mount point:]
# mount /dev/dsk/c0t0d0s0 /a

[Restore the root file system:]
# cd /a
# ufsrestore rvf /dev/rmt/0
# rm restoresymtable

[Create an empty install-db file:]
# touch /a/etc/vx/reconfig.d/state.d/install-db

[Edit /etc/system on the temporary file system and remove or
comment out the following entries:]
    # rootdev:/pseudo/vxio@0:0
    # set vxio:vol_rootdev_is_volume=1

[Edit /etc/vfstab on the temporary file system:]
Example:
Change from—
/dev/vx/dsk/rootdg/rootvol /dev/md/rdsk/rootdg/rootvol / ufs 1 no-

Change to—
/dev/dsk/c0t0d0s0 /dev/rdsk/c0t0d0s0  / ufs   1     no        -

[Unmount the temporary file system, then check the file system:]
# cd /
# umount /a
# fsck /dev/rdsk/c0t0d0s0

[Install a new boot block:]
# /usr/sbin/installboot /usr/platform/'uname \
-i'/lib/fs/ufs/bootblk /dev/rdsk/c0t0d0s0

[Reboot in single-user mode:]
# reboot -- "-s"

[Update the disk ID:]
# scdidadm -R /dev/rdsk/c0t0d0
```

```
[Run vxinstall:]
# vxinstall
Choose to encapsulate the root disk.

[If there is a conflict in minor number, reminor the rootdg disk
group:]
# umount /global/.devices/node@nodeid
# vxdg reminor rootdg 100
# shutdown -g0 -i6 -y
```

Where to Go From Here

For instructions about how to mirror the encapsulated root disk, see the *Sun Cluster 3.0 U1 Installation Guide*. Restoring Cluster Files

Administering Sun Cluster With the Graphical User Interfaces

This chapter provides descriptions of SunPlex Manager and Sun Management Center graphical user interface (GUI) tools, which you can use to administer some aspects of a cluster. It also contains procedures to configure and launch SunPlex Manager. The online help included with each GUI provides instructions for how to accomplish various administrative tasks using the GUI.

This is a list of the procedures in this chapter.

Sun Management Center Overview

The Sun Cluster-supplied module for Sun Management Center™ (formerly Sun Enterprise SyMON™) GUI Console enables you to graphically display cluster resources, resource types, and resource groups. It also enables you to monitor configuration changes and check the status of cluster components. However, the Sun Cluster-supplied module for Sun Management Center currently cannot perform all Sun Cluster administrative tasks. You must use the command-line interface for other operations. See "Command Line Interface" in Chapter 1 for more information.

For information on installing and starting the Sun Cluster module for Sun Management Center, and for viewing the cluster-specific online help supplied with the Sun Cluster module, see the *Sun Cluster 3.0 U1 Installation Guide*.

The Sun Cluster module of Sun Management Center is Simple Network Management Protocol (SNMP) compliant. Sun Cluster has created a Management Information Base (MIB) that can be used as the data definition by third-party management stations based on SNMP.

The Sun Cluster MIB file is located in `$BASEDIR/SUNWsymon/modules/cfg` on any cluster node.

The Sun Cluster MIB file is an ASN.1 specification of the Sun Cluster data that is modeled. This is the same specification used by all Sun Management Center MIBs. To use the Sun Cluster MIB, refer to the instructions for using other Sun Management Center MIBs in the *Sun Management Center 3.0 Software User's Guide*, Appendix B, "SNMP MIBs for Sun Management Center Modules."

SunPlex Manager Overview

SunPlex Manager is a GUI that enables you to graphically display cluster information, monitor configuration changes, and check the status of cluster components. It also allows you to perform some administrative tasks, including installing and configuring some data service applications. However, SunPlex Manager currently cannot perform all Sun Cluster administrative tasks. You must use the command-line interface for some operations. See for more information.

Information about installing and using SunPlex Manager can be found in the following locations.

- **Installing and starting SunPlex Manager:** See the *Sun Cluster 3.0 U1 Installation Guide*.
- **Configuring port numbers, server addresses, security certificates, and users:** See the section on configuring SunPlex Manager in the *Sun Cluster 3.0 U1 System Administration Guide*.
- **Installing and administering aspects of your cluster using SunPlex Manager:** See the online help supplied with SunPlex Manager.

Using SunPlex Manager Accessibility Features

SunPlex Manager supports third-party accessibility software when run through an accessible browser, such as Internet Explorer 5. This section describes some of these accessibility features.

- By default, the SunPlex Manager menu frame uses a JavaScript-enabled menu. Selecting an image or link in the menu frame expands or collapses any sub-items in the menu tree. Selecting an item in the menu frame also updates any information that is displayed in the content frame, that is related to the selected item.

  In addition to the Java-enabled menu, SunPlex Manager also provides a basic text-based menu that is always expanded and may interact better with accessibility software. To use that menu, start SunPlex Manager with the URL `https://`*nodename*`:3000/cgi-bin/index.pl?menu=basic` where *nodename* is replaced with the appropriate name of a node on which SunPlex Manager is installed. SunPlex Manager displays the text-based menu in the menu frame.

- SunPlex Manager uses combo box action menus to update and access a number of cluster components. When using keyboard control, select action menu items by opening the combo box pull-down menu before moving to the desired menu item. If you use the down arrow to move through the action menu items within the combo box, each menu item you move to is automatically selected and updated by using JavaScript. This could result in updating the wrong menu item.

  The following example shows how to access a combo box pull-down menu and select an item from that menu. The example assumes you are using keyboard commands with the Internet Explorer 5 browser.

  1. Tab to the desired combo box action menu.

  2. Press Alt-Down Arrow to display the pull-down menu.

  3. Press the down arrow key to move the cursor to the desired menu item.

  4. Press Return to select the menu item.

- SunPlex Manager provides several graphical topological views through Java applets. This information is also available in tabular status tables, since the Java applets may not be accessible.

Configuring SunPlex Manager

SunPlex Manager is a GUI that you can use to administer and view the status of some aspects of quorum devices, NAFO groups, interconnect components, and global devices. You can use it in place of many of the Sun Cluster CLI commands.

The procedure for installing SunPlex Manager on your cluster is included in the *Sun Cluster 3.0 U1 Installation Guide*. The SunPlex Manager online help contains instructions for completing various tasks using the GUI.

This section contains the following procedures for reconfiguring SunPlex Manager after initial installation.

▼ How to Change the Port Number for SunPlex Manager

If the default port number (3000) conflicts with another running process, change the port number of SunPlex Manager on each node of the cluster.

Note – The port number must be identical on all nodes of the cluster.

---

1. **Open the** /opt/SUNWscvw/conf/httpd.conf **configuration file using a text editor.**

2. **Change the Port number entry.**

   The Port entry is located under Section 2, 'Main' server configuration.

3. **Edit the VirtualHost entry to reflect the new port number.**

   The <VirtualHost _default_:3000> entry is located in the section titled "SSL Virtual Host Context".

4. **Save the configuration file and exit the editor.**

5. **Restart SunPlex Manager.**

   ```
   # /opt/SUNWscvw/bin/apachectl restart
   ```

6. **Repeat this procedure on each node of the cluster.**

▼ How to Change the Server Address for SunPlex Manager

If you change the hostname of a cluster node, you must change the address from which SunPlex Manager runs. Because the default security certificate is generated based on the node's hostname at the time SunPlex Manager is installed, you must remove one of the SunPlex Manager installation packages and reinstall it. You must complete this procedure on any node that has had its host name changed.

1. **Make the Sun Cluster 3.0 U1 CD-ROM image available to the node.**

2. **Remove the SUNWscvw package.**

```
# pkgrm SUNWscvw
```

3. **Re-install the SUNWscvw package.**

```
# cd <path to CD-ROM image>/SunCluster_3.0/Packages
# pkgadd -d . SUNWscvw
```

▼ How to Configure a New Security Certificate

You can generate your own security certificate to enable secure administration of your cluster, and then configure SunPlex Manager to use that certificate instead of the one generated by default. This procedure is an example of how to configure SunPlex Manager to use a security certificate generated by a particular security package. The actual tasks you must complete depend on the security package you use.

Note – You must generate an unencrypted certificate to allow the server to start on its own during boot up. Once you have generated a new certificate for each node of your cluster, configure SunPlex Manager to use those certificates. Each node must have its own security certificate.

---

1. **Copy the appropriate certificate to the node.**

2. **Open the** /opt/SUNWscvw/conf/httpd.conf **configuration file for editing.**

3. **Edit the following entry to enable SunPlex Manager to use the new certificate.**

```
SSLCertificateFile <path to certificate file>
```

4. **If the server private key is not combined with the certificate, edit the SSLCertificateKeyFile entry.**

```
SSLCertificateKeyFile <path to server key>
```

5. **Save the file and exit the editor.**

6. **Restart SunPlex Manager.**

```
# /opt/SUNWscvw/bin/apachectl restart
```

7. **Repeat this procedure for each node in the cluster.**

Example—Configuring SunPlex Manager to Use a New Security Certificate

The following example shows how to edit the SunPlex Manager configuration file to use a new security certificate.

```
[Copy the appropriate security certificates to each node.]

[Edit the configuration file.]
# vi /opt/SUNWscvw/conf/httpd.conf

[Edit the appropriate entries.]
SSLCertificateFile /opt/SUNWscvw/conf/ssl/phys-schost-1.crt
SSLCertificateKeyFile /opt/SUNWscvw/conf/ssl/phys-schost-1.key

[Save the file and exit the editor.]

[Restart SunPlex Manager.]
# /opt/SUNWscvw/bin/apachectl restart
```

Launching the SunPlex Manager Software

The SunPlex Manager graphical user interface (GUI) provides an easy way to administer some aspects of the Sun Cluster software. See the SunPlex Manager online help for more information.

▼ How to Launch SunPlex Manager

Follow this procedure to start SunPlex Manager on your cluster.

1. **Do you intend to access SunPlex Manager by using the cluster node** `root` **user name and password rather than set up a different user name and password?**
   - If yes, go to Step 5.
   - If no, go to Step 3 to set up SunPlex Manager user accounts.

2. **Become superuser on a cluster node.**

3. **Create a user account to access the cluster through SunPlex Manager.**

   You must set up at least one user account to access SunPlex Manager if you do not use the `root` system account.

   Note – SunPlex Manager user accounts are used only by SunPlex Manager. They do not correspond to any Solaris system user accounts.

   ---

   a. **Create a user ID name.**

   ```
   # /opt/SUNWscvw/bin/htpasswd -c /opt/SUNWscvw/conf/users userid
   ```

   b. **When prompted, type a password for that user ID.**

   c. **When prompted, type the password again for confirmation.**

   d. **Repeat Step a through Step c on each node of the cluster.**

   Use the same user ID name and password on all nodes.

   Note – Users who do not have a user account set up on a particular node cannot access the cluster through SunPlex Manager from that node, nor can users manage that node through another cluster node to which the users do have access.

   ---

4. **(Optional) Repeat Step 3 to set up additional user accounts.**

5. **From the administrative console or any other machine outside the cluster, launch a browser.**

6. **Disable the browser's Web proxy.**

   SunPlex Manager functionality is incompatible with Web proxies.

7. **Ensure that the browser's disk and memory cache sizes are set to greater than 0.**

8. **From the browser, connect to the SunPlex Manager port on one node of the cluster.**

   The default port number is 3000.

   ```
   https://node:3000/
   ```

Adding Non-Root Users to SunPlex Manager

To enable a user other than root to log in to SunPlex Manager, you must either create a new user with the "solaris.cluster.admin" Role Based Access Control (RBAC) authorization, or add that authorization to an existing user. Once given authorization to access the full SunPlex Manager functionality, the user can log in with his regular system password.

Note – Assigning the "solaris.cluster.admin" RBAC authorization to a non-root user allows that user to perform administrative actions usually performed only by root.

For more information, see "Role-Based Access Control" in the Solaris 8 System Administration Guide, Volume 2.

▼ How to Add RBAC Authorization to an Existing User Account

Add RBAC authorization to an existing user account to enable the user to log in to SunPlex Manager by using a regular system password.

Note – If you assign RBAC authorization to a non-`root` user account, that user account can perform a set of administrative actions usually performed only by `root`.

1. **Become superuser on any node of the cluster.**

2. **Add the following entry to the** `/etc/user_attr` **file on all nodes of the cluster.**

```
# vi /etc/user_attr
username:::type=normal;auths=solaris.cluster.admin
```

3. **If you are logged in to SunPlex Manager as the root user, exit the browser entirely, then restart it.**

4. **Connect to one node of the cluster.**

5. **Enter the login (username) and password to access SunPlex Manager.**

▼ How to Create a New User Account With RBAC Authorization

To provide a new non-root user account root access to the entire cluster through SunPlex Manager, create the account on all nodes of the cluster.

Note – If you assign RBAC authorization to a non-`root` user account, that user account can perform a set of administrative actions usually performed only by `root`.

---

1. **Become superuser on any node of the cluster.**

2. **Create the new user account.**

```
# useradd -d dir -A solaris.cluster.admin login
```

-d *dir*	Specifies the home directory of the new user
-A solaris.cluster.admin	Assigns the solaris.cluster.admin authorization to the new user account
login	Name of the new user account (username)

Note – The user name must be unique and must not already exist either on the local machine or in the network name service.

See the `useradd(1M)` man page for more information about creating user accounts.

3. Set the password on each node of the cluster.

```
# passwd login
```

Note – The password for this account must be the same on all nodes of the cluster.

4. If you are logged in to SunPlex Manager as the root user, exit the browser entirely, then restart it.

5. Connect to one node of the cluster.

6. Enter the new login (username) and password to access SunPlex Manager.