Sun Cluster System Administration Guide for Solaris OS

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95054
U.S.A.

Part No: 819–0580–10
August 2005, Revision A

050505@11223

Contents

Preface

The *Sun Cluster System Administration Guide for Solaris OS* provides procedures for administering a Sun™ Cluster configuration on both SPARC® and x86 based systems.

Note – In this document, the term "x86" refers to the Intel 32-bit family of microprocessor chips and compatible microprocessor chips made by AMD.

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.

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

Note – Sun Cluster software runs on two platforms, SPARC and x86. The information in this document pertains to both platforms unless otherwise specified in a special chapter, section, note, bulleted item, figure, table, or example.

Using UNIX Commands

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

See one or more of the following for this information:

- Online documentation for the Solaris software

- Other software documentation that you received with your system
- Solaris operating system man pages

Typographic Conventions

The following table describes the typographic changes that are used in this book.

TABLE P–1 Typographic Conventions

Typeface or Symbol	Meaning	Example
`AaBbCc123`	The names of commands, files, and directories, and onscreen computer output	Edit your `.login` file. Use `ls -a` to list all files. `machine_name% you have mail.`
`AaBbCc123`	What you type, contrasted with onscreen computer output	`machine_name%` **`su`** `Password:`
AaBbCc123	Command-line placeholder: replace with a real name or value	The command to remove a file is `rm` *filename*.
AaBbCc123	Book titles, new terms, and terms to be emphasized	Read Chapter 6 in the *User's Guide*. Perform a *patch analysis*. Do *not* save the file. [Note that some emphasized items appear bold online.]

Shell Prompts in Command Examples

The following table shows the default system prompt and superuser prompt for the C shell, Bourne shell, and Korn shell.

TABLE P–2 Shell Prompts

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

Related Documentation

Information about related Sun Cluster topics is available in the documentation that is listed in the following table. All Sun Cluster documentation is available at http://docs.sun.com.

Topic	Documentation
Overview	*Sun Cluster Overview for Solaris OS*
Concepts	*Sun Cluster Concepts Guide for Solaris OS*
Hardware installation and administration	*Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*
	Individual hardware administration guides
Software installation	*Sun Cluster Software Installation Guide for Solaris OS*
Data service installation and administration	*Sun Cluster Data Services Planning and Administration Guide for Solaris OS*
	Individual data service guides
Data service development	*Sun Cluster Data Services Developer's Guide for Solaris OS*
System administration	*Sun Cluster System Administration Guide for Solaris OS*
Error messages	*Sun Cluster Error Messages Guide for Solaris OS*
Command and function references	*Sun Cluster Reference Manual for Solaris OS*

For a complete list of Sun Cluster documentation, see the release notes for your release of Sun Cluster software at http://docs.sun.com.

For a complete list of Sun Cluster documentation, see the release notes for your release of Sun Cluster at http://docs.sun.com.

Documentation, Support, and Training

Sun Function	URL	Description
Documentation	`http://www.sun.com/documentation/`	Download PDF and HTML documents, and order printed documents
Support and Training	`http://www.sun.com/supportraining/`	Obtain technical support, download patches, and learn about Sun courses

Getting Help

Contact your service provider if you have problems installing or using Sun Cluster. Provide the following information to your service provider.

- Your name and email address
- 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 3.1 4/04

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
SPARC: `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.

CHAPTER 1

Introduction to Administering Sun Cluster

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

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.

Familiarize yourself with the planning information in the *Sun Cluster Software Installation Guide for Solaris OS* and the *Sun Cluster Concepts Guide for Solaris OS* before beginning administration tasks. 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 Data Services Planning and Administration Guide for Solaris OS*.

- Service tasks, such as adding or repairing storage or network hardware. These tasks are described in the *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*.

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, to impose minimal impact on the system. If you plan to take down the cluster or a cluster node, notify users ahead of time.

Solaris OS Feature Restrictions

The following Solaris OS features are not supported in a Sun Cluster configuration:

- Do not enable or disable the following Sun Cluster services using the Solaris 10 Service Management Facility (SMF) management interface.

Sun Cluster Service	FMRI
pnm	svc:/system/cluster/pnm:default
cl_event	svc:/system/cluster/cl_event:default
cl_eventlog	svc:/system/cluster/cl_eventlog:default
rpc_pmf	svc:/system/cluster/rpc_pmf:default
rpc_fed	svc:/system/cluster/rpc_fed:default
rgm	svc:/system/cluster/rgm:default
scdpm	svc:/system/cluster/scdpm:default
cl_ccra	svc:/system/cluster/cl_ccra:default
scsymon_srv	svc:/system/cluster/scsymon_srv:default
spm	svc:/system/cluster/spm:default
cl_svc_cluster_milestone	svc:/system/cluster/cl_svc_cluster_milestone:default
cl_svc_enable	svc:/system/cluster/cl_svc_enable:default
network-multipathing	svc:/system/cluster/network-multipathing

Administration Tools

You can perform administrative tasks on Sun Cluster by using a Graphical User Interface (GUI) or by using the command-line. The following section provides an overview of the GUI and command-line tools.

Graphical User Interface

Sun Cluster supports Graphical User Interface (GUI) tools that you can use to perform various administrative tasks on your cluster. These GUI tools are SunPlex™ Manager and, if you are using Sun Cluster on a SPARC based system, Sun Management Center. See Chapter 10 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 Main Menu items through the `scsetup` utility.

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

You can administer the following Resource Group Menu items through the `scsetup` utility.

- Create a resource group
- Add a network resource to a resource group
- Add a data service resource to a resource group
- Online/Offline or Switchover a resource group
- Enable/Disable a resource
- Change properties of a resource group
- Change properties of a resource
- Remove a resource from a resource group
- Remove a resource group

- Clear the stop_failed error flag from a resource

Table 1-1 lists other commands that 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.
`if_mpadm`(1M)	Use to switch IP addresses from one adapter to another in an IP Network Multipathing group.
`sccheck`(1M)	Checks and validates the Sun Cluster configuration to ensure the very basic configuration for a cluster to be functional.
`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. The command 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. **Note –** Resource type, resource group, and resource property names are case insensitive when executing `scrgadm`.
`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 that affect node mastery and states for resource groups and disk device groups.

In addition, 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™, VERITAS Volume Manager, or Solaris 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

Document the hardware aspects that are unique to your site as your Sun Cluster configuration is scaled. Refer to your hardware documentation when you change or upgrade the cluster to save administration labor. Labeling cables and connections between the various cluster components can also make administration easier.

Reduce the time required by a third-party service provider when servicing your cluster by keeping records of your original cluster configuration, and subsequent changes.

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 Log In to Sun Cluster Remotely" on page 20. For instructions on installing the Cluster Control Panel module for Sun Management Center and SunPlex Manager GUI tools, see the *Sun Cluster Software Installation Guide for Solaris OS*.

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 SPARC cluster consists of a Sun Enterprise™ 10000 server, you must log in from the administrative console to the System Service Processor (SSP). 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, you can use `netcon` in "exclusive" mode by setting the `-f` option. You can also send `~*` during a normal `netcon` session. Either of the previous solutions give you the option of toggling to the serial interface if the network becomes unreachable.

Sun Cluster does not require a dedicated administrative console, but using a console 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

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, Sun Cluster is not a replacement for regular backups. Sun Cluster can survive multiple failures, but 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/Solaris Volume Manager 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 3.1 4/04 Administration Tools

Task	Tool	Documentation
Log in to the Cluster Remotely	Use the `ccp` command to launch the Cluster Control Panel (CCP). Then select one of the following icons: `cconsole`(1M), `crlogin`(1M), or `ctelnet`(1M).	"How to Log In to Sun Cluster Remotely" on page 20
Configure the Cluster Interactively	Launch the`scsetup`(1M) utility.	"How to Access the `scsetup` Utility" on page 20
Display Sun Cluster Release Number and Version Information	Use the `scinstall`(1M) command with either the `-p` or `-pv` options.	"How to Display Sun Cluster Release and Version Information" on page 21

TABLE 1–2 Sun Cluster 3.1 4/04 Administration Tools *(Continued)*

Task	Tool	Documentation
Display Installed Resources, Resource Groups, and Resource Types **Note –** Resource type, resource group, and resource property names are case insensitive when executing `scrgadm`.	Use the `scrgadm`(1M) `-p` command.	"How to Display Configured Resource Types, Resource Groups, and Resources" on page 22
Monitor Cluster Components Graphically	Use SunPlex Manager.	SunPlex Manager online help
Administer Some Cluster Components Graphically	Use SunPlex Manager or the Sun Cluster module for Sun Management Center (which is available with Sun Cluster on SPARC based systems only).	SunPlex Manager or Sun Cluster module for Sun Management Center online help
Check the Status of Cluster Components	Use the `scstat`(1M) command.	"How to Check the Status of Cluster Components" on page 23
Check the Status of IP Network Multipathing Groups on the Public Network	Use the`scstat`(1M) command with the `-i` option.	"How to Check the Status of the Public Network" on page 25
View the Cluster Configuration	Use the `scconf`(1M) `-p` command.	"How to View the Cluster Configuration" on page 25
Check Global Mount Points	Use the `sccheck`(1M) command.	"How to Validate a Basic Cluster Configuration" on page 27
Look at Sun Cluster System Messages	Examine the`/var/adm/messages` file.	"Viewing System Messages" in *System Administration Guide: Advanced Administration* in *System Administration:Advanced Administration*
Monitor the Status of Solstice DiskSuite	Use the `metastat` commands.	Solstice DiskSuite/Solaris Volume Manager documentation
Monitor the Status of VERITAS Volume Manager if running Solaris 8	Use the `vxstat` or `vxva` commands.	VERITAS Volume Manager documentation
Monitor the Status of Solaris Volume Manager if running Solaris 9	Use the `svmstat`command	*Solaris Volume Manager Administration Guide*

▼ How to Log In to Sun Cluster Remotely

The Cluster Control Panel (CCP) provides a launch pad for `cconsole`(1M), `crlogin`(1M), and `ctelnet`(1M) 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 to the common window is sent to each of the 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.

Steps

1. **Verify that the following prerequisites are met before starting the CCP.**
   - Install the `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. The files 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 10000 server platform.**
   - If no, proceed to Step 3.
   - If yes, log into the System Service Processor (SSP) and connect by using the `netcon` command. After the connection is made, 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 either the cconsole, crlogin, or ctelnet icon in the CCP launch pad.**

See Also You can also start `cconsole`, `crlogin`, or `ctelnet` sessions from the command line.

▼ 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.

Steps

1. **Become superuser on any node in the cluster.**
2. **Enter the `scsetup` utility.**

   ```
   # scsetup
   ```

   The Main Menu is displayed.
3. **Make your configuration selection from the menu. Follow the onscreen instructions to complete a task.**

   See the `scsetup` online help for more information.

▼ How to Display Sun Cluster Patch Information

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

Step

- **Type the following command.**

  ```
  % showrev -p
  ```

  Sun Cluster update releases are identified by the main product patch number plus the update version.

Example 1–1 Displaying Sun Cluster Patch Information

The following example displays information about patch 110648-05.

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

▼ How to Display Sun Cluster Release and Version Information

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

Step

- **Type the following command.**

  ```
  % scinstall -pv
  ```

  This command displays Sun Cluster release number and version strings for all Sun Cluster packages.

Example 1–2 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 3.1
SUNWscr:       3.1.0,REV=2000.10.01.01.00
SUNWscdev:     3.1.0,REV=2000.10.01.01.00
SUNWscu:       3.1.0,REV=2000.10.01.01.00
SUNWscman:     3.1.0,REV=2000.10.01.01.00
SUNWscsal:     3.1.0,REV=2000.10.01.01.00
SUNWscsam:     3.1.0,REV=2000.10.01.01.00
SUNWscvm:      3.1.0,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. Refer to Chapter 10. See the SunPlex Manager online help for more information.

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

Step

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

  ```
  % scrgadm -p
  ```

Example 1–3 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.

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

```
% scstat -p
```

Example 1–4 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

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

-- IPMP Groups --
```

```
              Node Name         Group        Status         Adapter   Status
              ---------         -----        ------         -------   ------
  IPMP Group: phys-schost-1     sc_ipmp0   Online         qfe1      Online

  IPMP Group: phys-schost-2     sc_ipmp0   Online         qfe1      Online

------------------------------------------------------------------
```

▼ How to Check the Status of the Public Network

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.

To check the status of the IP Network Multipathing groups, use the `scstat`(1M) command.

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

```
% scstat -i
```

Example 1–5 Checking the Public Network Status

The following example provides a sample of status information for cluster components returned by `scstat -i`.

```
% scstat -i
------------------------------------------------------------------

-- IPMP Groups --

              Node Name         Group        Status         Adapter   Status
              ---------         -----        ------         -------   ------
  IPMP Group: phys-schost-1     sc_ipmp1      Online         qfe2      Online

  IPMP Group: phys-schost-1     sc_ipmp0      Online         qfe1      Online

  IPMP Group: phys-schost-2     sc_ipmp1      Online         qfe2      Online

  IPMP Group: phys-schost-2     sc_ipmp0      Online         qfe1      Online

------------------------------------------------------------------
```

▼ 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.

Step ● **View the cluster configuration**

```
% scconf -p
```

To display more information using the `scconf` command, use the verbose options. See the `scconf`(1M) man page for details.

Example 1–6 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:            192.168.0.0
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:                                 SVM
  Device group failback enabled:                     no
  Device group node list:                            phys-schost-3, phys-schost-4
  Diskset name:                                      schost-3
```

▼ How to Validate a Basic Cluster Configuration

The `sccheck`(1M) command runs a set of checks to validate the basic configuration required for a cluster to function properly. If no checks fail, `sccheck` returns to the shell prompt. If a check fails, `sccheck` produces reports in either the specified or the default output directory. If you run `sccheck` against more than one node, `sccheck` will produce a report for each node and a report for multi-node checks.

The `sccheck` command runs in two steps: data collection and analysis. Data collection can be time consuming, depending on the system configuration. You can invoke `sccheck` in verbose mode with the `-v1` flag to print progress messages, or you can use the `-v2` flag to run `sccheck` in highly verbose mode which prints more detailed progress messages, especially during data collection.

Note – Run `sccheck` after performing an administration procedure that might result in changes to devices, volume management components, or the Sun Cluster configuration.

Steps

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

   ```
   % su
   ```

2. **Verify the cluster configuration.**

   ```
   # sccheck
   ```

Example 1–7 Checking the Cluster Configuration With All Checks Passing

The following example shows `sccheck` being run in verbose mode against nodes `phys-schost-1` and `phys-schost-2` with all checks passing.

```
# sccheck -v1 -h phys-schost-1,phys-schost-2

sccheck: Requesting explorer data and node report from phys-schost-1.
sccheck: Requesting explorer data and node report from phys-schost-2.
sccheck: phys-schost-1: Explorer finished.
sccheck: phys-schost-1: Starting single-node checks.
sccheck: phys-schost-1: Single-node checks finished.
sccheck: phys-schost-2: Explorer finished.
sccheck: phys-schost-2: Starting single-node checks.
sccheck: phys-schost-2: Single-node checks finished.
sccheck: Starting multi-node checks.
sccheck: Multi-node checks finished
#
```

Example 1–8 Checking the Cluster Configuration With a Failed Check

The following example shows the node `phys-schost-2` in the cluster `suncluster` missing the mount point `/global/phys-schost-1`. Reports are created in the output directory `/var/cluster/sccheck/myReports/`.

```
# sccheck -v1 -h phys-schost-1,phys-schost-2 -o /var/cluster/sccheck/myReports

sccheck: Requesting explorer data and node report from phys-schost-1.
sccheck: Requesting explorer data and node report from phys-schost-2.
sccheck: phys-schost-1: Explorer finished.
sccheck: phys-schost-1: Starting single-node checks.
sccheck: phys-schost-1: Single-node checks finished.
sccheck: phys-schost-2: Explorer finished.
sccheck: phys-schost-2: Starting single-node checks.
sccheck: phys-schost-2: Single-node checks finished.
sccheck: Starting multi-node checks.
sccheck: Multi-node checks finished.
sccheck: One or more checks failed.
sccheck: The greatest severity of all check failures was 3 (HIGH).
sccheck: Reports are in /var/cluster/sccheck/myReports.
#
# cat /var/cluster/sccheck/myReports/sccheck-results.suncluster.txt
...
================================================
= ANALYSIS DETAILS =
================================================
------------------------------------
CHECK ID : 3065
SEVERITY : HIGH
FAILURE  : Global filesystem /etc/vfstab entries are not consistent across
all Sun Cluster 3.x nodes.
ANALYSIS : The global filesystem /etc/vfstab entries are not consistent across
all nodes in this cluster.
Analysis indicates:
```

```
FileSystem '/global/phys-schost-1' is on 'phys-schost-1' but missing from 'phys-schost-2'.
RECOMMEND: Ensure each node has the correct /etc/vfstab entry for the
filesystem(s) in question.
...
 #
```

▼ How to Check the Global Mount Points

The `sccheck`(1M) command includes checks which examine the `/etc/vfstab` file for configuration errors with the cluster file system and its global mount points.

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

Steps

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

   ```
   % su
   ```

2. **Verify the cluster configuration.**

   ```
   # sccheck
   ```

Example 1–9 Checking the Global Mount Points

The following example shows the node `phys-schost-2` of the cluster `suncluster` missing the mount point `/global/schost-1`. Reports are being sent to the output directory `/var/cluster/sccheck/myReports/`

```
# sccheck -v1 -h phys-schost-1,phys-schost-2 -o /var/cluster/sccheck/myReports

sccheck: Requesting explorer data and node report from phys-schost-1.
sccheck: Requesting explorer data and node report from phys-schost-2.
sccheck: phys-schost-1: Explorer finished.
sccheck: phys-schost-1: Starting single-node checks.
sccheck: phys-schost-1: Single-node checks finished.
sccheck: phys-schost-2: Explorer finished.
sccheck: phys-schost-2: Starting single-node checks.
sccheck: phys-schost-2: Single-node checks finished.
sccheck: Starting multi-node checks.
sccheck: Multi-node checks finished.
sccheck: One or more checks failed.
sccheck: The greatest severity of all check failures was 3 (HIGH).
sccheck: Reports are in /var/cluster/sccheck/myReports.
#
# cat /var/cluster/sccheck/myReports/sccheck-results.suncluster.txt

...
================================================
```

```
= ANALYSIS DETAILS =
================================================
------------------------------------
CHECK ID : 3065
SEVERITY : HIGH
FAILURE  : Global filesystem /etc/vfstab entries are not consistent across
all Sun Cluster 3.x nodes.
ANALYSIS : The global filesystem /etc/vfstab entries are not consistent across
all nodes in this cluster.
Analysis indicates:
FileSystem '/global/phys-schost-1' is on 'phys-schost-1' but missing from 'phys-schost-2'.
RECOMMEND: Ensure each node has the correct /etc/vfstab entry for the
filesystem(s) in question.
...
#
# cat /var/cluster/sccheck/myReports/sccheck-results.phys-schost-1.txt

...
================================================
= ANALYSIS DETAILS =
================================================
------------------------------------
CHECK ID : 1398
SEVERITY : HIGH
FAILURE  : An unsupported server is being used as a Sun Cluster 3.x node.
ANALYSIS : This server may not been qualified to be used as a Sun Cluster 3.x node.
Only servers that have been qualified with Sun Cluster 3.x are supported as
Sun Cluster 3.x nodes.
RECOMMEND: Because the list of supported servers is always being updated, check with
your Sun Microsystems representative to get the latest information on what servers
are currently supported and only use a server that is supported with Sun Cluster 3.x.
...
#
```

CHAPTER 2

Sun Cluster and RBAC

This chapter describes RBAC (Role-Based Access Control) in relation to Sun Cluster. Topics covered include:

Setting Up and Using RBAC With Sun Cluster

Use the following table to determine the documentation to consult about setting up and using RBAC. Specific steps that you follow to set up and use RBAC with Sun Cluster are presented later in this chapter.

To	Refer to
Learn more about RBAC	Chapter 8, "Using Roles and Privileges (Overview)," in *System Administration Guide: Security Services*
Set up, manage elements of, and use RBAC	Chapter 9, "Using Role-Based Access Control (Tasks)," in *System Administration Guide: Security Services*
Learn more about RBAC elements and tools	Chapter 10, "Role-Based Access Control (Reference)," in *System Administration Guide: Security Services*

Sun Cluster RBAC Rights Profiles

SunPlex Manager and selected Sun Cluster commands and options that you issue on the command line use RBAC for authorization. Several RBAC rights profiles are included in Sun Cluster. You can assign these rights profiles to users or to roles to give them different levels of access to Sun Cluster. Sun provides the following rights profiles with Sun Cluster software.

Rights Profile	Includes Authorizations	This Authorization Permits the Role Identity to
Sun Cluster Commands	None, but includes a list of Sun Cluster commands that run with `euid=0`	Execute selected Sun Cluster commands that you use to configure and manage a cluster, including:
		`scgdevs`(1M)
		`scswitch`(1M) (selected options)
		`scha_control`(1HA)
		`scha_resource_get`(1HA)
		`scha_resource_setstatus`(1HA)
		`scha_resourcegroup_get`(1HA)
		`scha_resourcetype_get`(1HA)
Basic Solaris User	This existing Solaris rights profile contains Solaris authorizations, as well as:	Perform the same operations that the Basic Solaris User role identity can perform, as well as:
	`solaris.cluster.device.read`	Read information about device groups
	`solaris.cluster.gui`	Access SunPlex Manager
	`solaris.cluster.network.read`	Read information about IP Network Multipathing
	`solaris.cluster.node.read`	Read information about attributes of nodes
	`solaris.cluster.quorum.read`	Read information about quorum devices and the quorum state
	`solaris.cluster.resource.read`	Read information about resources and resource groups
	`solaris.cluster.system.read`	Read the status of the cluster
	`solaris.cluster.transport.read`	Read information about transports
Cluster Operation	`solaris.cluster.appinstall`	Install clustered applications
	`solaris.cluster.device.admin`	Perform administrative tasks on device group attributes

Rights Profile	Includes Authorizations	This Authorization Permits the Role Identity to
	`solaris.cluster.device.read`	Read information about device groups
	`solaris.cluster.gui`	Access SunPlex Manager
	`solaris.cluster.install`	Install clustering software
	`solaris.cluster.network.admin`	Perform administrative tasks on IP Network Multipathing attributes
	`solaris.cluster.network.read`	Read information about IP Network Multipathing
	`solaris.cluster.node.admin`	Perform administrative tasks on node attributes
	`solaris.cluster.node.read`	Read information about attributes of nodes
	`solaris.cluster.quorum.admin`	Perform administrative tasks on quorum devices and quorum state attributes
	`solaris.cluster.quorum.read`	Read information about quorum devices and the quorum state
	`solaris.cluster.resource.admin`	Perform administrative tasks on resource attributes and resource group attributes
	`solaris.cluster.resource.read`	Read information about resources and resource groups
	`solaris.cluster.system.admin`	Administer the system
	`solaris.cluster.system.read`	Read the status of the cluster
	`solaris.cluster.transport.admin`	Perform administrative tasks on transport attributes
	`solaris.cluster.transport.read`	Read information about transports
System Administrator	This existing Solaris rights profile contains the same authorizations that the Cluster Management profile contains.	Perform the same operations that the Cluster Management role identity can perform, in addition to other system administration operations.
Cluster Management	This rights profile contains the same authorizations that the Cluster Operation profile contains, as well as the following authorizations:	Perform the same operations that the Cluster Operation role identity can perform, as well as:
	`solaris.cluster.device.modify`	Modify device group attributes
	`solaris.cluster.gui`	Access SunPlex Manager
	`solaris.cluster.network.modify`	Modify IP Network Multipathing attributes
	`solaris.cluster.node.modify`	Modify node attributes

Rights Profile	Includes Authorizations	This Authorization Permits the Role Identity to
	`solaris.cluster.quorum.modify`	Modify quorum devices and quorum state attributes
	`solaris.cluster.resource.modify`	Modify resource attributes and resource group attributes
	`solaris.cluster.system.modify`	Modify system attributes
	`solaris.cluster.transport.modify`	Modify transport attributes

Creating and Assigning an RBAC Role With a Sun Cluster Management Rights Profile

To create a role, you must either assume a role that has the Primary Administrator rights profile assigned to it or run as `root` user.

TABLE 2–1 Add Administrative Role Wizard: Dialog Boxes and Fields

Dialog Box	Fields	Field Description
Step 1: Enter a role name	Role Name	Short name of the role.
	Full Name	Long version of the name.
	Description	Description of the role.
	Role ID Number	UID for the role, automatically incremented.
	Role Shell	The profile shells that are available to roles: Administrator's C, Administrator's Bourne, or Administrator's Korn shell.
	Create a role mailing list	Makes a mailing list for users who are assigned to this role.
Step 2: Enter a role password	Role Password	********
	Confirm Password	********

TABLE 2–1 Add Administrative Role Wizard: Dialog Boxes and Fields *(Continued)*

Dialog Box	Fields	Field Description
Step 3: Select role rights	Available Rights / Granted Rights	Assigns or removes a role's rights profiles. Note that the system does not prevent you from typing multiple occurrences of the same command. The attributes that are assigned to the first occurrence of a command in a rights profile have precedence and all subsequent occurrences are ignored. Use the Up and Down arrows to change the order.
Step 4: Select a home directory	Server	Server for the home directory.
	Path	Home directory path.
Step 5: Assign users to this role	Add	Adds users who can assume this role. Must be in the same scope.
	Delete	Deletes users who are assigned to this role.

▼ How to Create a Role by Using the Administrative Roles Tool

Steps

1. **Start the Administrative Roles tool.**

   Run the Administrative Roles tool, start the Solaris Management Console, as described in "How to Assume a Role in the Solaris Management Console" in *System Administration Guide: Security Services* in System Administration Guide: Security Services. Then, open the User Tool Collection, and click the Administrative Roles icon.

2. **Start the Add Administrative Role wizard.**

   Select Add Administrative Role from the Action menu to start the Add Administrative Role wizard for configuring roles.

3. **Set up a role to which the Cluster Management rights profile is assigned.**

   Use the Next and Back buttons to navigate between dialog boxes. Note that the Next button does not become active until you have filled in all required fields. The last dialog box enables you to review the entered data, at which point you can go back to change entries or click Finish to save the new role. Table 2–1 summarizes the dialog boxes.

Note – You need to place this profile first in the list of profiles that are assigned to the role.

4. **Add users who need to use the SunPlex Manager features or Sun Cluster commands to the newly created role.**

   You use the `useradd`(1M) command to add a user account to the system. The `-P` option assigns a role to a user's account.

5. **Click Finish when you are done.**

6. **Open a terminal window, become `root`, and start and stop the name service cache daemon.**

   The new role does not take effect until the name service cache daemon is restarted. After becoming `root`, type as follows:

   ```
   # /etc/init.d/nscd stop
   # /etc/init.d/nscd start
   ```

▼ How to Create a Role From the Command Line

Steps

1. **Become superuser or assume a role that is capable of creating other roles.**

2. **Select a method for creating a role:**

   - For roles in the local scope, use the `roleadd`(1M) command to specify a new local role and its attributes.
   - Alternatively, for roles in the local scope, edit the `user_attr`(4) file to add a user with `type=role`.

     This method is recommended for emergencies only, as it is easy to make mistakes while you are typing.
   - For roles in a name service, use the `smrole`(1M) command to specify the new role and its attributes.

     This command requires authentication by superuser or a role that is capable of creating other roles. You can apply the `smrole` to all name services. This command runs as a client of the Solaris Management Console server.

3. **Start and stop the name service cache daemon.**

   New roles do not take effect until the name service cache daemon is restarted. As `root`, type as follows:

   ```
   # /etc/init.d/nscd stop
   # /etc/init.d/nscd start
   ```

Example 2–1 Creating a Custom Operator Role by Using the `smrole` Command

The following sequence demonstrates how a role is created with the `smrole` command. In this example, a new version of the Operator role is created that has assigned to it the standard Operator rights profile and the Media Restore rights profile.

```
% su primaryadmin
# /usr/sadm/bin/smrole add -H myHost -- -c "Custom Operator" -n oper2 -a johnDoe \
-d /export/home/oper2 -F "Backup/Restore Operator" -p "Operator" -p "Media Restore"

Authenticating as user: primaryadmin

Type /? for help, pressing <enter> accepts the default denoted by [ ]
Please enter a string value for: password ::     <type primaryadmin password>

Loading Tool: com.sun.admin.usermgr.cli.role.UserMgrRoleCli from myHost
Login to myHost as user primaryadmin was successful.
Download of com.sun.admin.usermgr.cli.role.UserMgrRoleCli from myHost was successful.

Type /? for help, pressing <enter> accepts the default denoted by [ ]
Please enter a string value for: password ::    <type oper2 password>

# /etc/init.d/nscd stop
# /etc/init.d/nscd start
```

To view the newly created role (and any other roles), use `smrole` with the `list` option, as follows:

```
# /usr/sadm/bin/smrole list --
Authenticating as user: primaryadmin

Type /? for help, pressing <enter> accepts the default denoted by [ ]
Please enter a string value for: password ::     <type primaryadmin password>

Loading Tool: com.sun.admin.usermgr.cli.role.UserMgrRoleCli from myHost
Login to myHost as user primaryadmin was successful.
Download of com.sun.admin.usermgr.cli.role.UserMgrRoleCli from myHost was successful.
root                    0               Super-User
primaryadmin            100             Most powerful role
sysadmin                101             Performs non-security admin tasks
oper2                   102             Custom Operator
```

Modifying a User's RBAC Properties

To modify a user's properties, you must either be running the User Tool Collection as `root` user or assume a role that has the Primary Administrator rights profile assigned to it.

▼ How to Modify a User's RBAC Properties by Using the User Accounts Tool

Steps

1. **Start the User Accounts tool.**

   To run the User Accounts tool, you need to start the Solaris Management Console, as described in "How to Assume a Role in the Solaris Management Console" in *System Administration Guide: Security Services* in *System Administration Guide: Security Services*. Then, open the User Tool Collection, and click the User Accounts icon.

   After the User Accounts tool starts, the icons for the existing user accounts are displayed in the view pane.

2. **Click the user account icon to be changed and select Properties from the Action menu (or simply double-click the user account icon).**

3. **Click the appropriate tab in the dialog box for the property to be changed, as follows:**

   - To change the roles that are assigned to the user, click the Roles tab and move the role assignment to be changed to the appropriate column: Available Roles or Assigned Roles.
   - To change the rights profiles that are assigned to the user, click the Rights tab and move it to the appropriate column: Available Rights or Assigned Rights.

   Note – It is not good practice to assign rights profiles directly to users. The preferred approach is to force users to assume roles in order to perform privileged applications. This strategy avoids the possibility of normal users abusing privileges.

▼ How to Modify a User's RBAC Properties From the Command Line

Steps

1. **Become superuser or assume a role that can modify user files.**

2. **Use the appropriate command:**

   - To change the authorizations, roles, or rights profiles that are assigned to a user who is defined in the local scope, use the `usermod`(1M) command.
   - Alternatively, to change the authorizations, roles, or rights profiles that are assigned to a user who is defined in the local scope, edit the `user_attr` file.

     This method is recommended for emergencies only, as it is easy to make a mistake while you are typing.

- To change the authorizations, roles, or rights profiles that are assigned to a user who is defined in a name service, use the `smuser`(1M) command.

  This command requires authentication as superuser or as a role that is capable of changing user files. You can apply `smuser` to all name services. `smuser` runs as a client of the Solaris Management Console server.

CHAPTER 3

Shutting Down and Booting a Cluster

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

For a high-level description of the related procedures in this chapter, see Table 3–1 and Table 3–2.

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 use the `scshutdown` command when moving the location of a cluster. You can also use the command to shut down the cluster if you have data corruption caused by an application error.

Note – Use the `scshutdown` command instead of the `shutdown` or `halt` commands to ensure proper shutdown of the entire cluster. The Solaris `shutdown` command is used with the `scswitch`(1M) command to shut down individual nodes. See "How to Shut Down a Cluster" on page 43 or "Shutting Down and Booting a Single Cluster Node" on page 53 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 OpenBoot™ PROM `ok` prompt on a SPARC based system or to a boot subsystem on an x86 based system. Boot subsystems are described in more detail in "Boot Subsystems" in *System Administration Guide: Basic Administration*.

Note – If necessary, you can boot a node in non-cluster mode so that the node does not participate in cluster membership. Non-cluster mode 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 64 for more information.

TABLE 3–1 Task List: Shutting Down and Booting a Cluster

Task	For Instructions
Stop the cluster -Use `scshutdown`(1M)	See "How to Shut Down a Cluster" on page 43
Start the cluster by booting all nodes. The nodes must have a working connection to the cluster interconnect to attain cluster membership.	See "How to Boot a Cluster" on page 44
Reboot the cluster - Use `scshutdown` At the `ok` prompt or the `Select (b)oot or (i)nterpreter` prompt on the Current Boot Parameters screen, boot each node individually with the `boot`(1M) or the `b` command. The nodes must have a working connection to the cluster interconnect to attain cluster membership.	See "How to Reboot a Cluster" on page 48

▼ How to Shut Down a Cluster

Caution – Do not use `send brk` on a cluster console to shut down a cluster node. The command is not supported within a cluster.

Steps

1. **SPARC: If your cluster is running Oracle Parallel Server or Real Application Clusters, shut down all instances of the database.**

   Refer to the Oracle Parallel Server or Oracle Real Application Clusters product documentation for shutdown procedures.

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

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

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

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

4. **Verify that all nodes are showing the `ok` prompt on a SPARC based system or a Boot Subsystem on an x86 based system.**

   Do not power off any nodes until all cluster nodes are at the `ok` prompt on a SPARC based system or in a Boot Subsystem on an x86 based system.

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

Example 3–1 SPARC: Shutting Down a Cluster

The following example shows the console output when stopping normal cluster operation and bringing down all nodes so that the `ok` prompt is shown. 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
Wed Mar 10 13:47:32 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
```

Example 3–2 x86: Shutting Down a Cluster

The following example shows the console output when stopping normal cluster operation and bringing down all nodes. 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 10:32:57 phys-schost-1 cl_runtime:
WARNING: CMM: Monitoring disabled.
root@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)
failfasts already disabled on node 1
Print services already stopped.
May  2 10:33:13 phys-schost-1 syslogd: going down on signal 15
The system is down.
syncing file systems... done
Type any key to continue
```

See Also See "How to Boot a Cluster" on page 44 to restart a cluster that has been shut down.

▼ How to Boot a Cluster

Steps 1. **To start a cluster whose nodes have been shut down and are at the `ok` prompt or at the `Select (b)oot or (i)nterpreter` prompt on the Current Boot Parameters screen, `boot`(1M) 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.

- SPARC:

```
ok boot
```

- x86:

```
                        <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>  to boot with options
or      i <ENTER>                           to enter boot interpreter
or      <ENTER>                             to boot with defaults

                    <<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b
```

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 68.

Example 3–3 SPARC: 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 phys-schost-1 with votecount = 1 added.
NOTICE: Node phys-schost-2 with votecount = 1 added.
NOTICE: Node phys-schost-3 with votecount = 1 added.
...
NOTICE: Node phys-schost-1: attempting to join cluster
...
NOTICE: Node phys-schost-2 (incarnation # 937690106) has become reachable.
NOTICE: Node phys-schost-3 (incarnation # 937690290) has become reachable.
NOTICE: cluster has reached quorum.
NOTICE: node phys-schost-1 is up; new incarnation number = 937846227.
NOTICE: node phys-schost-2 is up; new incarnation number = 937690106.
NOTICE: node phys-schost-3 is up; new incarnation number = 937690290.
NOTICE: Cluster members: phys-schost-1 phys-schost-2 phys-schost-3.
...
```

Example 3–4 x86: 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.

```
ATI RAGE SDRAM BIOS P/N GR-xlint.007-4.330
*                                       BIOS Lan-Console 2.0
```

```
Copyright (C) 1999-2001  Intel Corporation
MAC ADDR: 00 02 47 31 38 3C
AMIBIOS (C)1985-2002 American Megatrends Inc.,
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064
2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

Press <F2> to enter SETUP, <F12> Network

Adaptec AIC-7899 SCSI BIOS v2.57S4
(c) 2000 Adaptec, Inc. All Rights Reserved.
    Press <Ctrl><A> for SCSISelect(TM) Utility!

Ch B,  SCSI ID: 0 SEAGATE  ST336605LC        160
       SCSI ID: 1 SEAGATE  ST336605LC        160
       SCSI ID: 6 ESG-SHV  SCA HSBP M18      ASYN
Ch A,  SCSI ID: 2 SUN      StorEdge 3310     160
       SCSI ID: 3 SUN      StorEdge 3310     160

AMIBIOS (C)1985-2002 American Megatrends Inc.,
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064

2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

SunOS - Intel Platform Edition             Primary Boot Subsystem, vsn 2.0

                        Current Disk Partition Information

                  Part#   Status    Type      Start       Length
                 ================================================
                    1     Active   X86 BOOT     2428       21852
                    2              SOLARIS     24280     71662420
                    3              <unused>
                    4              <unused>
               Please select the partition you wish to boot: *       *

Solaris DCB

                    loading /solaris/boot.bin

SunOS Secondary Boot version 3.00
```

```
                  Solaris Intel Platform Edition Booting System

Autobooting from bootpath: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/
pci8086,341a@7,1/sd@0,0:a

If the system hardware has changed, or to boot from a different
device, interrupt the autoboot process by pressing ESC.
Press ESCape to interrupt autoboot in 2 seconds.
Initializing system
Please wait...
Warning: Resource Conflict - both devices are added

NON-ACPI device: ISY0050
     Port: 3F0-3F5, 3F7; IRQ: 6; DMA: 2
ACPI device: ISY0050
     Port: 3F2-3F3, 3F4-3F5, 3F7; IRQ: 6; DMA: 2

                     <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>  to boot with options
or      i <ENTER>                           to enter boot interpreter
or      <ENTER>                             to boot with defaults

                 <<< timeout in 5 seconds >>>

Select (b)oot or (i)nterpreter:
Size: 275683 + 22092 + 150244 Bytes
/platform/i86pc/kernel/unix loaded - 0xac000 bytes used
SunOS Release 5.9 Version Generic_112234-07 32-bit
Copyright 1983-2003 Sun Microsystems, Inc.  All rights reserved.
Use is subject to license terms.
configuring IPv4 interfaces: e1000g2.
Hostname: phys-schost-1
Booting as part of a cluster
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) with votecount = 1 added.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) with votecount = 1 added.
NOTICE: CMM: Quorum device 1 (/dev/did/rdsk/d1s2) added; votecount = 1, bitmask
of nodes with configured paths = 0x3.
NOTICE: clcomm: Adapter e1000g3 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being initiated
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 online
NOTICE: clcomm: Adapter e1000g0 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being constructed
NOTICE: CMM: Node phys-schost-1: attempting to join cluster.
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being initiated
NOTICE: CMM: Quorum device /dev/did/rdsk/d1s2: owner set to node 1.
NOTICE: CMM: Cluster has reached quorum.
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) is up; new incarnation number = 1068496374.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) is up; new incarnation number = 1068496374.
NOTICE: CMM: Cluster members: phys-schost-1 phys-schost-2.
NOTICE: CMM: node reconfiguration #1 completed.
```

```
NOTICE: CMM: Node phys-schost-1: joined cluster.
```

▼ How to Reboot a Cluster

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

Steps

1. **SPARC: If your cluster is running Oracle Parallel Server or Oracle Real Application Clusters, shut down all instances of the database.**

   Refer to the Oracle Parallel Server or Oracle Real Application Clusters product documentation for shutdown procedures.

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

3. **Shut down the cluster.**

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

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

   Each node is shut down.

   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. If you make configuration changes between shutdowns, start the node with the most current configuration first.

   - SPARC:

     ```
     ok boot
     ```

   - x86:

     ```
                          <<< Current Boot Parameters >>>
     Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
     sd@0,0:a
     Boot args:

     Type    b [file-name] [boot-flags] <ENTER>  to boot with options
     or      i <ENTER>                           to enter boot interpreter
     or      <ENTER>                             to boot with defaults

                       <<< timeout in 5 seconds >>>
     Select (b)oot or (i)nterpreter: b
     ```

   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 68.

Example 3–5 SPARC: 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
Wed Mar 10 13:47:32 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 phys-schost-1: attempting to join cluster
...
NOTICE: Node phys-schost-2 (incarnation # 937690106) has become reachable.
NOTICE: Node phys-schost-3 (incarnation # 937690290) has become reachable.
NOTICE: cluster has reached quorum.
...
NOTICE: Cluster members: phys-schost-1 phys-schost-2 phys-schost-3.
...
NOTICE: Node phys-schost-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 phys-schost-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:
```

Example 3–6 x86: Rebooting a Cluster

The following example shows the console output when stopping normal cluster operation, bringing down all nodes, 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 10:32:57 phys-schost-1 cl_runtime:
WARNING: CMM: Monitoring disabled.
root@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)
failfasts already disabled on node 1
Print services already stopped.
May  2 10:33:13 phys-schost-1 syslogd: going down on signal 15
The system is down.
syncing file systems... done
Type any key to continue

ATI RAGE SDRAM BIOS P/N GR-xlint.007-4.330
*                                       BIOS Lan-Console 2.0
Copyright (C) 1999-2001  Intel Corporation
MAC ADDR: 00 02 47 31 38 3C
AMIBIOS (C)1985-2002 American Megatrends Inc.,
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064
2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

Press <F2> to enter SETUP, <F12> Network

Adaptec AIC-7899 SCSI BIOS v2.57S4
(c) 2000 Adaptec, Inc. All Rights Reserved.
    Press <Ctrl><A> for SCSISelect(TM) Utility!
```

```
Ch B,  SCSI ID: 0 SEAGATE  ST336605LC        160
       SCSI ID: 1 SEAGATE  ST336605LC        160
       SCSI ID: 6 ESG-SHV  SCA HSBP M18      ASYN
Ch A,  SCSI ID: 2 SUN      StorEdge 3310     160
       SCSI ID: 3 SUN      StorEdge 3310     160

AMIBIOS (C)1985-2002 American Megatrends Inc.,
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064

2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

SunOS - Intel Platform Edition             Primary Boot Subsystem, vsn 2.0

                        Current Disk Partition Information

                  Part#   Status    Type      Start       Length
                 ================================================
                    1     Active   X86 BOOT     2428       21852
                    2              SOLARIS     24280     71662420
                    3              <unused>
                    4              <unused>
               Please select the partition you wish to boot: *       *

Solaris DCB

                    loading /solaris/boot.bin

SunOS Secondary Boot version 3.00

                   Solaris Intel Platform Edition Booting System

Autobooting from bootpath: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/
pci8086,341a@7,1/sd@0,0:a

If the system hardware has changed, or to boot from a different
device, interrupt the autoboot process by pressing ESC.
Press ESCape to interrupt autoboot in 2 seconds.
Initializing system
Please wait...
Warning: Resource Conflict - both devices are added

NON-ACPI device: ISY0050
     Port: 3F0-3F5, 3F7; IRQ: 6; DMA: 2
ACPI device: ISY0050
     Port: 3F2-3F3, 3F4-3F5, 3F7; IRQ: 6; DMA: 2

                     <<< Current Boot Parameters >>>
```

```
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>  to boot with options
or      i <ENTER>                           to enter boot interpreter
or      <ENTER>                             to boot with defaults

                  <<< timeout in 5 seconds >>>

Select (b)oot or (i)nterpreter: b
Size: 275683 + 22092 + 150244 Bytes
/platform/i86pc/kernel/unix loaded - 0xac000 bytes used
SunOS Release 5.9 Version Generic_112234-07 32-bit
Copyright 1983-2003 Sun Microsystems, Inc.  All rights reserved.
Use is subject to license terms.
configuring IPv4 interfaces: e1000g2.
Hostname: phys-schost-1
Booting as part of a cluster
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) with votecount = 1 added.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) with votecount = 1 added.
NOTICE: CMM: Quorum device 1 (/dev/did/rdsk/d1s2) added; votecount = 1, bitmask
of nodes with configured paths = 0x3.
NOTICE: clcomm: Adapter e1000g3 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being initiated
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 online
NOTICE: clcomm: Adapter e1000g0 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being constructed
NOTICE: CMM: Node phys-schost-1: attempting to join cluster.
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being initiated
NOTICE: CMM: Quorum device /dev/did/rdsk/d1s2: owner set to node 1.
NOTICE: CMM: Cluster has reached quorum.
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) is up; new incarnation number = 1068496374.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) is up; new incarnation number = 1068496374.
NOTICE: CMM: Cluster members: phys-schost-1 phys-schost-2.
NOTICE: CMM: node reconfiguration #1 completed.
NOTICE: CMM: Node phys-schost-1: joined cluster.
WARNING: mod_installdrv: no major number for rsmrdt
ip: joining multicasts failed (18) on clprivnet0 - will use link layer
broadcasts for multicast
The system is coming up.  Please wait.
checking ufs filesystems
/dev/rdsk/c1t0d0s5: is clean.
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 online
NIS domain name is dev.eng.mycompany.com
starting rpc services: rpcbind keyserv ypbind done.
Setting netmask of e1000g2 to 192.168.255.0
Setting netmask of e1000g3 to 192.168.255.128
Setting netmask of e1000g0 to 192.168.255.128
Setting netmask of clprivnet0 to 192.168.255.0
Setting default IPv4 interface for multicast: add net 224.0/4: gateway phys-schost-1
syslog service starting.
obtaining access to all attached disks
```

```
********************************************************************************
*
* The X-server can not be started on display :0...
*
********************************************************************************
volume management starting.
Starting Fault Injection Server...
The system is ready.

phys-schost-1 console login:
```

Shutting Down and Booting a Single Cluster Node

Note – Use the `scswitch`(1M) command in conjunction with the Solaris `shutdown`(1M) command to shut down an individual node. Use the `scshutdown` command only when shutting down an entire cluster.

TABLE 3–2 Task Map: Shutting Down and Booting a Cluster Node

Task	For Instructions
Stop a cluster node - Use `scswitch`(1M) and `shutdown`(1M)	"How to Shut Down a Cluster Node" on page 54
Start a node The node must have a working connection to the cluster interconnect to attain cluster membership.	"How to Boot a Cluster Node" on page 56
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 59

TABLE 3–2 Task Map: Shutting Down and Booting a Cluster Node *(Continued)*

Task	For Instructions
Boot a node so that the node does not participate in cluster membership - Use `scswitch` and `shutdown`, then `boot -x` or `b -x`	"How to Boot a Cluster Node in Non-Cluster Mode" on page 64

▼ How to Shut Down a Cluster Node

Caution – Do not use `send brk` on a cluster console to shut down a cluster node. The command is not supported within a cluster.

Steps

1. **SPARC: If your cluster is running Oracle Parallel Server or Oracle Real Application Clusters, shut down all instances of the database.**
   Refer to the Oracle Parallel Server or Oracle Real Application Clusters 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 node
   ```

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

   -h *node* Specifies the node from which you are switching resource groups and device groups.
4. **Shut down the cluster node.**
   On the node to be shut down, type the following command.

   ```
   # shutdown -g0 -y -i0
   ```
5. **Verify that the cluster node is showing the `ok` prompt or the `Select (b)oot or (i)nterpreter` prompt on the Current Boot Parameters screen.**
6. **If necessary, power off the node.**

Example 3–7 SPARC: 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
Wed Mar 10 13:47:32 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
```

Example 3–8 x86: 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
Shutdown started.    Wed Mar 10 13:47:32 PST 2004

Changing to init state 0 - please wait
Broadcast Message from root (console) on phys-schost-1 Wed Mar 10 13:47:32...
THE SYSTEM phys-schost-1 IS BEING SHUT DOWN NOW ! ! !
Log off now or risk your files being damaged

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)
failfasts disabled on node 1
Print services already stopped.
Mar 10 13:47:44 phys-schost-1 syslogd: going down on signal 15
umount: /global/.devices/node@2 busy
umount: /global/.devices/node@1 busy
```

```
The system is down.
syncing file systems... done
WARNING: CMM: Node being shut down.
Type any key to continue
```

See Also See "How to Boot a Cluster Node" on page 56 to restart a cluster node that has been shut down.

▼ How to Boot a Cluster Node

If you intend to shut down or reboot other, active nodes in the cluster, wait until the node you are booting has at least reached the login prompt. Otherwise, the node will not be available to take over services from other nodes in the cluster that you shut down or reboot.

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. You should have one quorum count for each node and one quorum count 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. The required cluster quorum count (two) must be present.

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

- SPARC:

```
ok boot
```

- x86:

```
                        <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>  to boot with options
or      i <ENTER>                           to enter boot interpreter
or      <ENTER>                             to boot with defaults

                    <<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b
```

Messages appear on the booted nodes' 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` 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 68.

Example 3–9 SPARC: 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 phys-schost-1: attempting to join cluster
...
NOTICE: Node phys-schost-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:
```

Example 3–10 x86: Booting a Cluster Node

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

```
                <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>   to boot with options
```

```
or      i <ENTER>                           to enter boot interpreter
or      <ENTER>                             to boot with defaults

                 <<< timeout in 5 seconds >>>

Select (b)oot or (i)nterpreter: Size: 276915 + 22156 + 150372 Bytes
/platform/i86pc/kernel/unix loaded - 0xac000 bytes used
SunOS Release 5.9 Version on81-feature-patch:08/30/2003 32-bit
Copyright 1983-2003 Sun Microsystems, Inc.  All rights reserved.
Use is subject to license terms.
configuring IPv4 interfaces: e1000g2.
Hostname: phys-schost-1
Booting as part of a cluster
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) with votecount = 1 added.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) with votecount = 1 added.
NOTICE: CMM: Quorum device 1 (/dev/did/rdsk/d1s2) added; votecount = 1, bitmask
of nodes with configured paths = 0x3.
WARNING: CMM: Initialization for quorum device /dev/did/rdsk/d1s2 failed with
error EACCES. Will retry later.
NOTICE: clcomm: Adapter e1000g3 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being initiated
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 online
NOTICE: clcomm: Adapter e1000g0 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being constructed
NOTICE: CMM: Node phys-schost-1: attempting to join cluster.
WARNING: CMM: Reading reservation keys from quorum device /dev/did/rdsk/d1s2
failed with error 2.
NOTICE: CMM: Cluster has reached quorum.
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) is up; new incarnation number =
1068503958.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) is up; new incarnation number =
1068496374.
NOTICE: CMM: Cluster members: phys-schost-1 phys-schost-2.
NOTICE: CMM: node reconfiguration #3 completed.
NOTICE: CMM: Node phys-schost-1: joined cluster.
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being initiated
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 online
NOTICE: CMM: Retry of initialization for quorum device /dev/did/rdsk/d1s2 was
successful.
WARNING: mod_installdrv: no major number for rsmrdt
ip: joining multicasts failed (18) on clprivnet0 - will use link layer
broadcasts for multicast
The system is coming up.  Please wait.
checking ufs filesystems
/dev/rdsk/c1t0d0s5: is clean.
NIS domain name is dev.eng.mycompany.com
starting rpc services: rpcbind keyserv ypbind done.
Setting netmask of e1000g2 to 192.168.255.0
Setting netmask of e1000g3 to 192.168.255.128
Setting netmask of e1000g0 to 192.168.255.128
Setting netmask of clprivnet0 to 192.168.255.0
Setting default IPv4 interface for multicast: add net 224.0/4: gateway phys-schost-1
syslog service starting.
```

```
obtaining access to all attached disks


********************************************************************************
*
* The X-server can not be started on display :0...
*
********************************************************************************
volume management starting.
Starting Fault Injection Server...
The system is ready.

phys-schost-1 console login:
```

▼ How to Reboot a Cluster Node

If you intend to shut down or reboot other, active nodes in the cluster, wait until the node you are rebooting has at least reached the login prompt. Otherwise, the node will not be available to take over services from other nodes in the cluster that you shut down or reboot.

Steps

1. **SPARC: If the cluster node is running Oracle Parallel Server or Oracle Real Application Clusters, shut down all instances of the database.**

   Refer to the Oracle Parallel Server or Oracle Real Application Clusters 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 the node shuts down.

   ```
   # scswitch -S -h node
   # 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 3–11 SPARC: Rebooting a Cluster Node

The following example shows the console output when rebooting node phys-schost-1. Messages for this node, such as shutdown and startup notification, appear on the consoles of other nodes in the cluster.

```
# scswitch -S -h phys-schost-1
# shutdown -g0 -y -i6
Shutdown started.    Wed Mar 10 13:47:32 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 phys-schost-1: attempting to join cluster
...
NOTICE: Node phys-schost-1: joined cluster
...
The system is coming up.  Please wait.
The system is ready.
phys-schost-1 console login:
```

Example 3–12 x86: Rebooting a Cluster Node

The following example shows the console output when rebooting node phys-schost-1. Messages for this node, such as shutdown and startup notification, appear on the consoles of other nodes in the cluster.

```
# scswitch -S -h phys-schost-1
# shutdown -g0 -y -i6
Shutdown started.    Wed Mar 10 13:47:32 PST 2004

Changing to init state 6 - please wait
```

```
Broadcast Message from root (console) on phys-schost-1 Wed Mar 10 13:47:32...
THE SYSTEM phys-schost-1 IS BEING SHUT DOWN NOW ! ! !
Log off now or risk your files being damaged

phys-schost-1#
INIT: New run level: 6
The system is coming down.  Please wait.
System services are now being stopped.
/etc/rc0.d/K05initrgm: Calling scswitch -S (evacuate)
Print services already stopped.
Mar 10 13:47:44 phys-schost-1 syslogd: going down on signal 15

umount: /global/.devices/node@2 busy
umount: /global/.devices/node@1 busy
The system is down.
syncing file systems... done
WARNING: CMM: Node being shut down.
rebooting...

ATI RAGE SDRAM BIOS P/N GR-xlint.007-4.330
*                                       BIOS Lan-Console 2.0
Copyright (C) 1999-2001  Intel Corporation
MAC ADDR: 00 02 47 31 38 3C
AMIBIOS (C)1985-2002 American Megatrends Inc.,
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064
2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

Press <F2> to enter SETUP, <F12> Network

Adaptec AIC-7899 SCSI BIOS v2.57S4
(c) 2000 Adaptec, Inc. All Rights Reserved.
    Press <Ctrl><A> for SCSISelect(TM) Utility!

Ch B,  SCSI ID: 0 SEAGATE  ST336605LC        160
       SCSI ID: 1 SEAGATE  ST336605LC        160
       SCSI ID: 6 ESG-SHV  SCA HSBP M18      ASYN
Ch A,  SCSI ID: 2 SUN      StorEdge 3310     160
       SCSI ID: 3 SUN      StorEdge 3310     160

AMIBIOS (C)1985-2002 American Megatrends Inc.,
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064

2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
```

```
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

SunOS - Intel Platform Edition             Primary Boot Subsystem, vsn 2.0

                        Current Disk Partition Information

                 Part#   Status    Type      Start       Length
                ================================================
                   1     Active   X86 BOOT     2428       21852
                   2              SOLARIS     24280     71662420
                   3              <unused>
                   4              <unused>
               Please select the partition you wish to boot: *       *

Solaris DCB

                   loading /solaris/boot.bin

SunOS Secondary Boot version 3.00

                  Solaris Intel Platform Edition Booting System

Autobooting from bootpath: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/
pci8086,341a@7,1/sd@0,0:a

If the system hardware has changed, or to boot from a different
device, interrupt the autoboot process by pressing ESC.
Press ESCape to interrupt autoboot in 2 seconds.
Initializing system
Please wait...
Warning: Resource Conflict - both devices are added

NON-ACPI device: ISY0050
     Port: 3F0-3F5, 3F7; IRQ: 6; DMA: 2
ACPI device: ISY0050
     Port: 3F2-3F3, 3F4-3F5, 3F7; IRQ: 6; DMA: 2

                     <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>   to boot with options
or      i <ENTER>                            to enter boot interpreter
or      <ENTER>                              to boot with defaults

                  <<< timeout in 5 seconds >>>

Select (b)oot or (i)nterpreter: Size: 276915 + 22156 + 150372 Bytes
/platform/i86pc/kernel/unix loaded - 0xac000 bytes used
SunOS Release 5.9 Version on81-feature-patch:08/30/2003 32-bit
Copyright 1983-2003 Sun Microsystems, Inc.  All rights reserved.
```

```
Use is subject to license terms.
configuring IPv4 interfaces: e1000g2.
Hostname: phys-schost-1
Booting as part of a cluster
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) with votecount = 1 added.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) with votecount = 1 added.
NOTICE: CMM: Quorum device 1 (/dev/did/rdsk/d1s2) added; votecount = 1, bitmask
of nodes with configured paths = 0x3.
WARNING: CMM: Initialization for quorum device /dev/did/rdsk/d1s2 failed with
error EACCES. Will retry later.
NOTICE: clcomm: Adapter e1000g3 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being initiated
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 online
NOTICE: clcomm: Adapter e1000g0 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being constructed
NOTICE: CMM: Node phys-schost-1: attempting to join cluster.
WARNING: CMM: Reading reservation keys from quorum device /dev/did/rdsk/d1s2
failed with error 2.
NOTICE: CMM: Cluster has reached quorum.
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) is up; new incarnation number =
1068503958.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) is up; new incarnation number =
1068496374.
NOTICE: CMM: Cluster members: phys-schost-1 phys-schost-2.
NOTICE: CMM: node reconfiguration #3 completed.
NOTICE: CMM: Node phys-schost-1: joined cluster.
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being initiated
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 online
NOTICE: CMM: Retry of initialization for quorum device /dev/did/rdsk/d1s2 was
successful.
WARNING: mod_installdrv: no major number for rsmrdt
ip: joining multicasts failed (18) on clprivnet0 - will use link layer
broadcasts for multicast
The system is coming up.  Please wait.
checking ufs filesystems
/dev/rdsk/c1t0d0s5: is clean.
NIS domain name is dev.eng.mycompany.com
starting rpc services: rpcbind keyserv ypbind done.
Setting netmask of e1000g2 to 192.168.255.0
Setting netmask of e1000g3 to 192.168.255.128
Setting netmask of e1000g0 to 192.168.255.128
Setting netmask of clprivnet0 to 192.168.255.0
Setting default IPv4 interface for multicast: add net 224.0/4: gateway phys-schost-1
syslog service starting.
obtaining access to all attached disks


********************************************************************************
*
* The X-server can not be started on display :0...
*
********************************************************************************
volume management starting.
```

```
Starting Fault Injection Server...
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 the node does not participate in the cluster membership, that is, in non-cluster mode. Non-cluster mode is useful when installing the cluster software or performing certain administrative procedures, such as patching a node.

Steps

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

3. **Verify that the node is showing the `ok` prompt or the `Select (b)oot or (i)nterpreter` prompt on the Current Boot Parameters screen.**
4. **Boot the node in non-cluster mode by using the `boot`(1M) or `b` command with the `-x` option.**

   - SPARC:

     ```
     ok boot -x
     ```

   - x86:

     ```
                          <<< Current Boot Parameters >>>
     Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
     sd@0,0:a
     Boot args:

     Type    b [file-name] [boot-flags] <ENTER>  to boot with options
     or      i <ENTER>                           to enter boot interpreter
     or      <ENTER>                             to boot with defaults

                      <<< timeout in 5 seconds >>>
     Select (b)oot or (i)nterpreter: b -x
     ```

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

Example 3–13 SPARC: Booting a Cluster Node in Non-Cluster Mode

The following example shows the console output when shutting down node `phys-schost-1` then restarting the node 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
Shutdown started.    Wed Mar 10 13:47:32 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 phys-schost-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:
```

Example 3–14 x86: Booting a Cluster Node in Non-Cluster Mode

The following example shows the console output when shutting down node `phys-schost-1` then restarting the node 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
Shutdown started.    Wed Mar 10 13:47:32 PST 2004

phys-schost-1#
INIT: New run level: 0
The system is coming down.  Please wait.
```

```
System services are now being stopped.
Print services already stopped.
Mar 10 13:47:44 phys-schost-1 syslogd: going down on signal 15
...
The system is down.
syncing file systems... done
WARNING: CMM: Node being shut down.
Type any key to continue

ATI RAGE SDRAM BIOS P/N GR-xlint.007-4.330
*                                       BIOS Lan-Console 2.0
Copyright (C) 1999-2001  Intel Corporation
MAC ADDR: 00 02 47 31 38 3C
AMIBIOS (C)1985-2002 American Megatrends Inc.,
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064
2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

Press <F2> to enter SETUP, <F12> Network

Adaptec AIC-7899 SCSI BIOS v2.57S4
(c) 2000 Adaptec, Inc. All Rights Reserved.
    Press <Ctrl><A> for SCSISelect(TM) Utility!

Ch B,  SCSI ID: 0 SEAGATE  ST336605LC        160
       SCSI ID: 1 SEAGATE  ST336605LC        160
       SCSI ID: 6 ESG-SHV  SCA HSBP M18      ASYN
Ch A,  SCSI ID: 2 SUN      StorEdge 3310     160
       SCSI ID: 3 SUN      StorEdge 3310     160

AMIBIOS (C)1985-2002 American Megatrends Inc.,
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064

2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

SunOS - Intel Platform Edition             Primary Boot Subsystem, vsn 2.0

                        Current Disk Partition Information

                 Part#   Status    Type      Start       Length
                ================================================
```

```
                    1     Active   X86 BOOT     2428       21852
                    2              SOLARIS     24280     71662420
                    3              <unused>
                    4              <unused>
               Please select the partition you wish to boot: *       *

Solaris DCB

                    loading /solaris/boot.bin

SunOS Secondary Boot version 3.00

                  Solaris Intel Platform Edition Booting System

Autobooting from bootpath: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/
pci8086,341a@7,1/sd@0,0:a

If the system hardware has changed, or to boot from a different
device, interrupt the autoboot process by pressing ESC.
Press ESCape to interrupt autoboot in 2 seconds.
Initializing system
Please wait...
Warning: Resource Conflict - both devices are added

NON-ACPI device: ISY0050
     Port: 3F0-3F5, 3F7; IRQ: 6; DMA: 2
ACPI device: ISY0050
     Port: 3F2-3F3, 3F4-3F5, 3F7; IRQ: 6; DMA: 2

                     <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>  to boot with options
or      i <ENTER>                           to enter boot interpreter
or      <ENTER>                             to boot with defaults

                  <<< timeout in 5 seconds >>>

Select (b)oot or (i)nterpreter: b -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. Refer to "Viewing System Messages" in *System Administration Guide: Advanced Administration* in *System Administration Guide: Advanced Administration* for more information.

Steps

1. **Become superuser on the cluster node with the full `/var` file system.**
2. **Clear the full file system.**
   For example, delete nonessential files that are contained in the file system.

CHAPTER 4

Administering Global Devices, Disk-Path Monitoring, and Cluster File Systems

This chapter provides the procedures for administering global devices, disk-path monitoring and cluster file systems.

Following is a list of the procedures in this chapter.

For a high-level description of the related procedures in this chapter, see Table 4–2.

See the *Sun Cluster Concepts Guide for Solaris OS* for conceptual information related to global devices, the global namespace, disk device groups, disk-path monitoring 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 that is installed on the cluster. Solstice DiskSuite/Solaris Volume Manager is "cluster-aware," so you add, register, and remove disk device groups by using the Solstice DiskSuite/Solaris Volume Manager `metaset`(1M) command. If you are using VERITAS Volume Manager (VxVM), you create disk groups by using VxVM commands. You register the disk groups as Sun Cluster disk device groups with the `scsetup`(1M) utility. When removing VxVM disk device groups, you use both the `scsetup` command and VxVM commands.

Sun Cluster software automatically creates a rawdisk device group for each disk and tape device in the cluster. However, cluster device groups remain in an offline state until you access the groups as global devices. 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. The global namespace is automatically set up during installation and automatically updated during Solaris operating system reboots. However, if the global namespace needs to be updated, you can run the `scgdevs`(1M) command from any cluster node. This command 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/Solaris Volume Manager

Changes made to global device permissions are not automatically propagated to all the nodes in the cluster for Solstice DiskSuite/Solaris Volume Manager 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.

Dynamic Reconfiguration With Global Devices

Following are issues the you must consider when completing dynamic reconfiguration (DR) operations on disk and tape devices in a cluster.

- All of the requirements, procedures, and restrictions that are documented for the Solaris DR feature also apply to Sun Cluster DR support. The only exception is for the operating environment quiescence operation. Therefore, review the documentation for the Solaris DR feature *before* using the DR feature with Sun Cluster software. You should review in particular the issues that affect non-network IO devices during a DR detach operation.
- Sun Cluster rejects DR remove-board operations on active devices in the primary node. DR operations can be performed on non-active devices in the primary node and on any devices in the secondary nodes.
- After the DR operation, cluster data access continues as before.
- Sun Cluster rejects DR operations that impact the availability of quorum devices. See "Dynamic Reconfiguration With Quorum Devices" on page 121 for more information.

Caution – If the current primary node fails while you are performing the DR operation on a secondary node, cluster availability is impacted. The primary node will have no place to fail over until a new secondary node is provided.

To perform DR operations on global devices, complete the following steps in the order indicated.

TABLE 4–1 Task Map: Dynamic Reconfiguration with Disk and Tape Devices

Task	For Instructions
1. If a DR operation that affects an active device group must be performed on the current primary node, switch the primary and secondary nodes before performing the DR remove operation on the device.	"How to Switch the Primary for a Device Group" on page 103
2. Perform the DR removal operation on the device being removed.	*Sun Enterprise 10000 DR Configuration Guide* and the *Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual* in the *Solaris 8 on Sun Hardware* and *Solaris 9 on Sun Hardware* collections.

SPARC: 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. Changing the volume permissions, owner, or group ID also impacts the namespace.

  Note – Never import or deport VxVM disk groups by using VxVM commands once the disk group has been registered with the cluster as a Sun Cluster disk device group. The Sun Cluster software handles all cases where disk groups need to be imported or be 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, the minor number is sufficient to

guarantee uniqueness. The minor number for a newly-created disk group might conflict with the minor number of a pre-existing disk group that was imported on a different node. In this case, attempting to register the Sun Cluster disk device group fails. 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 after a node failure. Use of DRL is strongly recommended, although use of DRL could decrease I/O throughput.
- VxVM does not support the `chmod` command. To change global device permissions in VxVM, consult the VxVM administrator's guide.
- Sun Cluster 3.1 4/04 software does not support VxVM Dynamic Multipathing (DMP) to manage multiple paths from the same node.
- If you use VxVM to set up shared disk groups for Oracle Parallel Server or Oracle Real Application Clusters, use the cluster functionality of VxVM as described in the *VERITAS Volume Manager Administrator's Reference Guide*. There are differences between creating shared disk groups for Oracle Parallel Server or Oracle Real Application Clusters and creating other disk groups. You must import the Oracle Parallel Server or Oracle Real Application Clusters shared disk groups by using `vxdg -s`. You do not register the Oracle Parallel Server or Oracle Real Application Clusters shared disk groups with the cluster framework. To create other VxVM disk groups, see "SPARC: How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)" on page 84.

Administering Cluster File Systems Overview

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

Note – When the cluster file system reads files, the file system does not update the access time on those files.

Cluster File System Restrictions

The following restrictions apply to the cluster file system administration:

- The command `unlink`(1M) is not supported on non-empty directories.
- The command `lockfs -d` is not supported. Use `lockfs -n` as a workaround.
- You cannot remount a cluster file system with the `directio` mount option added at remount time.
- You cannot set the `directio` mount option on a single file by using the `directio` ioctl.

SPARC: Guidelines to Support VxFS

The following VxFS features are not supported in a Sun Cluster 3.1 cluster file system. They are, however, supported in a local file system.

- Quick I/O
- Snapshots
- Storage checkpoints
- VxFS-specific mount options:
  - `convosync` (Convert O_SYNC)
  - `mincache`
  - `qlog`, `delaylog`, `tmplog`
- VERITAS cluster file system (requires VxVM cluster feature & VERITAS Cluster Server)

Cache advisories can be used, but the effect is observed on the given node only.

All other VxFS features and options that are supported in a cluster file system are supported by Sun Cluster 3.1 software. See VxFS documentation for details about VxFS options that are supported in a cluster configuration.

The following guidelines for how to use VxFS to create highly available cluster file systems are specific to a Sun Cluster 3.1 4/04 configuration.

- Create a VxFS file system by following procedures in VxFS documentation.
- Mount and unmount a VxFS file system from the primary node. The primary node masters the disk on which the VxFS file system resides. A VxFS file system mount or unmount operation that is performed from a secondary node might fail.
- Perform all VxFS administration commands from the primary node of the VxFS cluster file system.

The following guidelines for how to administer VxFS cluster file systems are not specific to Sun Cluster 3.1 4/04 software. However, the guidelines are different from the way you administer UFS cluster file systems.

- You can administer files on a VxFS cluster file system from any node in the cluster. The exception is ioctls, which you must issue only from the primary node. If you do not know whether an administration command involves ioctls, issue the command from the primary node.
- If a VxFS cluster file system fails over to a secondary node, all standard-system-call operations that were in progress during failover are re-issued transparently on the new primary. However, any ioctl-related operation in progress during the failover will fail. After a VxFS cluster file system failover, check the state of the cluster file system. Administrative commands that were issued on the old primary before failover might require corrective measures. See VxFS documentation for more information.

Administering Disk Device Groups

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

Note – Sun Cluster software automatically creates a raw disk device group for each disk and tape device in the cluster. However, cluster device groups remain in an offline state until you access the groups as global devices.

TABLE 4–2 Task List: Administering Disk Device Groups

Task	For Instructions, Go To…
Update the global device namespace without a reconfiguration reboot - Use `scgdevs`(1M)	"How to Update the Global Device Namespace" on page 77
Add Solstice DiskSuite/Solaris Volume Manager disk sets and register them as disk device groups - Use `metaset`(1M)	"How to Add and Register a Disk Device Group (Solstice DiskSuite/Solaris Volume Manager)" on page 78
Remove Solstice DiskSuite/Solaris Volume Manager disk device groups from the configuration - Use `metaset` and `metaclear`(1M)	"How to Remove and Unregister a Disk Device Group (Solstice DiskSuite/Solaris Volume Manager)" on page 79

TABLE 4–2 Task List: Administering Disk Device Groups *(Continued)*

Task	For Instructions, Go To…
Remove a node from all disk device groups - Use `scconf`, `metaset`, and `scsetup`	"How to Remove a Node From All Disk Device Groups" on page 79
Remove a node from a Solstice DiskSuite/Solaris Volume Manager disk device group - Use `metaset`	"How to Remove a Node From a Disk Device Group (Solstice DiskSuite/Solaris Volume Manager)" on page 80
SPARC: Add VERITAS Volume Manager disk groups as disk device groups	"SPARC: How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)" on page 84
- Use VxVM commands and `scsetup`(1M)	"SPARC: How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)" on page 85
	"SPARC: How to Add a New Volume to an Existing Disk Device Group (VERITAS Volume Manager)" on page 86
	"SPARC: How to Make an Existing Disk Group Into a Disk Device Group (VERITAS Volume Manager)" on page 87
	"SPARC: How to Assign a New Minor Number to a Disk Device Group (VERITAS Volume Manager)" on page 87
	"SPARC: How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)" on page 88
	"SPARC: How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 91
SPARC: Remove VERITAS Volume Manager disk device groups from the configuration	"SPARC: How to Remove a Volume From a Disk Device Group (VERITAS Volume Manager)" on page 93
- Use `scsetup` (to generate scconf)	"SPARC: How to Remove and Unregister a Disk Device Group (VERITAS Volume Manager)" on page 94
SPARC: Add a node to a VERITAS Volume Manager disk device group - Use `scsetup` to generate `scconf`	"SPARC: How to Add a Node to a Disk Device Group (VERITAS Volume Manager)" on page 95

TABLE 4–2 Task List: Administering Disk Device Groups *(Continued)*

Task	For Instructions, Go To…
SPARC: Remove a node from a VERITAS Volume Manager disk device group - Use `scsetup` to generate `scconf`	"SPARC: How to Remove a Node From a Disk Device Group (VERITAS Volume Manager)" on page 96
Remove a node from a raw disk device group - Use `scconf`(1M)	"SPARC: How to Remove a Node From a Raw Disk Device Group" on page 98
Change disk device group properties - Use `scsetup` to generate `scconf`	"How to Change Disk Device Properties" on page 99
Display disk device groups and properties - Use `scconf`	"How to List a Disk Device Group Configuration" on page 102
Change the desired number of secondaries for a device group - Use `scsetup` to generate `scconf`	"How to Change the Desired Number of Secondaries for a Device Group" on page 100
Switch the primary for a disk device group - Use `scswitch`(1M)	"How to Switch the Primary for a Device Group" on page 103
Put a disk device group in maintenance state - Use `metaset` or `vxdg`	"How to Put a Disk Device Group in Maintenance State" on page 104

▼ 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 that is running the command is not currently a cluster member. The command also has no effect if the `/global/.devices/node@` *nodeID* file system is not mounted.

Steps

1. **Become superuser on any node of the cluster.**
2. **Use the `scgdevs` command to reconfigure the namespace.**

```
# scgdevs
```

Example 4–1 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/Solaris Volume Manager)

Use the `metaset` command to create a Solstice DiskSuite/Solaris Volume Manager disk set and register the disk set as a Sun Cluster disk device group. When you register the disk set, the name that you assigned to the disk set is automatically assigned to the disk device group.

Steps

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

2. **Calculate the number of names for Solstice DiskSuite metadevices or Solaris Volume Manager volumes that you need for your configuration, and modify the `/kernel/drv/md.conf` file on each node.**

   See "How to Set the Number of Metadevice or Volume Names and Disk Sets " in the *Sun Cluster Software Installation Guide for Solaris OS*.

3. **Use the `metaset`(1M)command to add the Solstice DiskSuite/Solaris Volume Manager disk set and register it as a disk device group with Sun Cluster. To create a multi-owner disk group, use the `-M` option.**

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

`-s` *diskset*	Specifies the disk set to be created.
`-a -h` *nodelist*	Adds the list of nodes that can master the disk set.
`-M`	Designates the disk group as multi-owner.

   Note – Running the `metaset` command to set up a Solstice DiskSuite/Solaris Volume Manager device group on a cluster results in one secondary by default, regardless of the number of nodes that are included in that device group. You can change the desired number of secondary nodes by using the `scsetup`(1M) utility after the device group has been created. Refer to "How to Change the Desired Number of Secondaries for a Device Group" on page 100 for more information about disk failover.

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

The disk device group name matches the disk set name that is specified with `metaset`.

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

Example 4–2 Adding a Solstice DiskSuite/Solaris Volume Manager Disk Device Group

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

```
# metaset -s dg-schost-1 -a -h phys-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/Solaris Volume Manager)

Disk device groups are Solstice DiskSuite/Solaris Volume Manager disk sets that have been registered with Sun Cluster. To remove a Solstice DiskSuite/Solaris Volume Manager disk device group, use the `metaclear` and `metaset` commands. These commands remove the disk device group with the same name and unregister the disk group as a Sun Cluster disk device group.

Refer to the Solstice DiskSuite/Solaris Volume Manager documentation for the steps to remove a disk set.

▼ How to Remove a Node From All Disk Device Groups

Use this procedure to remove a cluster node from all disk device groups that list the node in their lists of potential primaries.

Steps

1. **Become superuser on the node you want to remove as a potential primary of all disk device groups.**
2. **Determine the disk device group(s) of which the node to be removed is a member.**

Look for the node name in the `Device group node list` for each disk device group.

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

3. **Are any of the disk device groups identified in Step 2 of the device group type `SDS/SVM`?**
   - If yes, perform the procedures in "How to Remove a Node From a Disk Device Group (Solstice DiskSuite/Solaris Volume Manager)" on page 80.
   - If no, go to Step 4.
4. **Are any of the disk device groups identified in Step 2 of the device group type `VxVM`?**
   - If yes, perform the procedures in "SPARC: How to Remove a Node From a Disk Device Group (VERITAS Volume Manager)" on page 96.
   - If no, go to Step 5.
5. **Determine the raw disk device groups of which the node to be removed is a member.**

   Note that the following command contains two "v"s in `-pvv`. The second "v" is needed to display raw disk device groups.

   ```
   # scconf -pvv | grep "Device group"
   ```
6. **Are any of the disk device groups listed in Step 5 of the device group types `Disk`, `Local_Disk`, or both?**
   - If yes, perform the procedures in "SPARC: How to Remove a Node From a Raw Disk Device Group" on page 98.
   - If no, go to Step 7.
7. **Verify that the node has been removed from the potential primaries list of all disk device groups.**

   The command returns nothing if the node is no longer listed as a potential primary of any disk device group.

   ```
   # scconf -pvv | grep "Device group" | grep nodename
   ```

▼ How to Remove a Node From a Disk Device Group (Solstice DiskSuite/Solaris Volume Manager)

Use this procedure to remove a cluster node from the list of potential primaries of a Solstice DiskSuite/Solaris Volume Manager disk device group. Repeat the `metaset` command for each disk device group from which you want to remove the node.

Steps

1. **Verify that the node is still a member of the disk device group and that the disk device group is a Solstice DiskSuite/Solaris Volume Manager disk device group.**

   Device group type `SDS/SVM` indicates a Solstice DiskSuite/Solaris Volume Manager disk device group.

```
phys-schost-1% scconf -pv | grep '(global-galileo)'
  (global-galileo) Device group type:              SDS/SVM
  (global-galileo) Device group failback enabled:  no
  (global-galileo) Device group node list:         phys-schost-1, phys-schost-2
  (global-galileo) Diskset name:                   global-galileo
phys-schost-1%
```

2. **Determine which node is the current primary for the device group.**

```
# scstat -D
```

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

4. **Delete the node's hostname from the disk device group.**

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

`-s` *setname*	Specifies the disk device group name.
`-d`	Deletes from the disk device group the nodes identified with `-h`.
`-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.

   If the command fails, add the `-f` (Force) option to the command.

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

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

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

   The disk device group name matches the disk set name that is specified with `metaset`.

```
phys-schost-1% scconf -pv | grep  devicegroup
Device group node list:  phys-schost-1, phys-schost-2
```

Example 4–3 Removing a Node From a Disk Device Group (Solstice DiskSuite/Solaris Volume Manager)

The following example shows the removal of the host name `phys-schost-2` from a disk device group configuration. This example 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. Check that the removed node is no longer displayed in the screen text.

```
[Determine the Solstice DiskSuite/Solaris Volume Manager
disk device group(2) for the node:]
# scconf -pv | grep Device
  Device group name:                dg-schost-1
    Device group type:              SDS/SVM
    Device group failback enabled:  no
    Device group node list:         phys-schost-1, phys-schost-2
    Device group ordered node list: yes
    Device group diskset name:        dg-schost-1
[Determine which node is the current primary for the disk device group:]
# scstat -D
  -- Device Group Servers --
                       Device Group  Primary        Secondary
                       ------------  -------        ---------
  Device group servers: dg-schost-1  phys-schost-1  phys-schost-2
[Become superuser on the node that currently owns the disk device group.]
[Remove the hostname from the disk device group:]
# metaset -s dg-schost-1 -d -h phys-schost-2
[Verify removal of the node:]
phys-schost-1% scconf -pv | grep dg-schost-1
  -- Device Group Servers --
                       Device Group  Primary        Secondary
                       ------------  -------        ---------
  Device group servers: dg-schost-1  phys-schost-1
```

▼ How to Create More Than Three Disk Sets in a Cluster

If you intend to create more than three disk sets in the cluster, perform the following steps before you create the disk sets. Follow these steps if you are installing disk sets for the first time or if you are adding more disk sets to a fully configured cluster.

Steps **1. Ensure that the value of the `md_nsets` variable is high enough. The value should accommodate the total number of disk sets you intend to create in the cluster.**

a. On any node of the cluster, check the value of the `md_nsets` variable in the `/kernel/drv/md.conf` file.

b. If the number of disk sets in the cluster will be greater than the existing value of `md_nsets` minus one, increase the value of `md_nsets` on each node.

The maximum permissible number of disk sets is the value of `md_nsets` minus one. The maximum possible value of `md_nsets` is 32.

c. Ensure that the `/kernel/drv/md.conf` file is identical on each node of the cluster.

Caution – Failure to follow this guideline can result in serious Solstice DiskSuite/Solaris Volume Manager errors and possible loss of data.

d. From one node, shut down the cluster.

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

e. Reboot each node in the cluster.

- SPARC:

```
ok boot
```

- x86:

```
                      <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@
7,1/sd@0,0:a
Boot args:

Type b [file-name] [boot-flags] <ENTER> to boot with options
or   i <ENTER>                          to enter boot interpreter
or   <ENTER>                            to boot with defaults

                  <<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b
```

2. On each node in the cluster, run the `devfsadm`(1M) command.

You can run this command on all nodes in the cluster at the same time.

3. From one node of the cluster, run the `scgdevs`(1M) command.

4. On each node, verify that the `scgdevs` command has completed before you attempt to create any disk sets.

The `scgdevs` command calls itself remotely on all nodes, even when the command is run from just one node. To determine whether the `scgdevs` command has completed processing, run the following command on each node of the cluster.

```
% ps -ef | grep scgdevs
```

▼ SPARC: 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 "SPARC: How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)" on page 85.

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

If you use VxVM to set up shared disk groups for Oracle Parallel Server or Oracle Real Application Clusters, use the cluster functionality of VxVM as described in the *VERITAS Volume Manager Administrator's Reference Guide*.

Steps

1. **Become superuser on any cluster node that is *physically connected* to 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 after a node failure. However, DRL might decrease I/O throughput.

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

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

   See "SPARC: How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)" on page 88.

   Do not register the Oracle Parallel Server or Oracle Real Application Clusters shared disk groups with the cluster framework.

▼ SPARC: 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 "SPARC: How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)" on page 84.

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

Disk encapsulation is only supported during initial creation of a VxVM disk group. After 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.

If you use VxVM to set up shared disk groups for Oracle Parallel Server or Oracle Real Application Clusters, use the cluster functionality of VxVM as described in the *VERITAS Volume Manager Administrator's Reference Guide*.

Steps

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`.**

   Set back to `yes` once the disk is 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 two must also be set to the entire disk. See the `vxdiskadm` 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. Use `shutdown` to shut down and restart the node.

   ```
   # scswitch -S -h node[,...]
   # 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 -D disk-device-group -h node[,...]
# scswitch -z -g resource-group -h node[,...]
```

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

   See "SPARC: How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)" on page 88.

   Do not register the Oracle Parallel Server or Oracle Real Application Clusters shared disk groups with the cluster framework.

7. **If you set the `mount at boot` option to `no` in Step 2, set it back to `yes`.**

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

When you add a new volume to an existing VxVM disk device group, perform the procedure from the primary node of the online disk device group.

Note – After adding the volume, you need to register the configuration change by using the procedure "SPARC: How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 91.

Steps

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

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

   ```
   # scstat -D
   ```

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

   ```
   # scswitch -z -D disk-device-group -h node[,...]
   ```

`-z -D` *disk-device-group*	Switches the specified device group.
`-h` *node*	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 to update the global namespace.**

   DPM

   "SPARC: How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 91.

▼ SPARC: 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 importing the disk group onto the current node, then registering the disk group as a Sun Cluster disk device group.

Steps

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 "SPARC: How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)" on page 88.

▼ SPARC: 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.

Steps

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 another multiple of 1000 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 "SPARC: How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)" on page 88.

Example 4–4 SPARC: 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
```

▼ SPARC: 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 – After a disk device group has been registered with the cluster, never import or deport a VxVM disk group by using VxVM commands. If you make a change to the VxVM disk group or volume, use the procedure "SPARC: How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 91 to register the disk device group configuration changes. This procedure ensures 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.
- A desired number of secondary nodes for 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.

Non-primary cluster nodes (spares) transition to secondary according to the node preference order. The default number of secondaries for a device group is normally set to one. This default setting minimizes performance degradation caused by primary checkpointing of multiple secondary nodes during normal operation. For example, in a four node cluster, the default behavior configures one primary, one secondary, and two spare nodes. See also "SPARC: How to Set the Desired Number of Secondaries (VERITAS Volume Manager)" on page 91.

Steps

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.

   If you use VxVM to set up shared disk groups for Oracle Parallel Server/Real Application Clusters, you do not register the shared disk groups with the cluster framework. Use the cluster functionality of VxVM as described in the *VERITAS Volume Manager Administrator's Reference Guide*.

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 "SPARC: How to Assign a New Minor Number to a Disk Device Group (VERITAS Volume Manager)" on page 87. 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 synchronize the disk device group by using `scsetup`(1M). 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 77.

Example 4–5 SPARC: Registering a VERITAS Volume Manager Disk Device Group

The following example shows the `scconf` command generated by `scsetup` when registering 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
```

See Also To create a cluster file system on the VxVM disk device group, see "How to Add a Cluster File System" on page 106.

If there are problems with the minor number, see "SPARC: How to Assign a New Minor Number to a Disk Device Group (VERITAS Volume Manager)" on page 87.

▼ SPARC: 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. Registration ensures that the global namespace is in the correct state.

Steps

1. **Become superuser on any node in the cluster.**
2. **Run 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 4–6 SPARC: Registering VERITAS Volume Manager Disk Group Configuration Changes

The following example shows the `scconf` command generated by `scsetup` when registering 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
```

▼ SPARC: How to Set the Desired Number of Secondaries (VERITAS Volume Manager)

The `numsecondaries` property specifies the number of nodes within a device group that can master the group if the primary node fails. The default number of secondaries for device services is one. The value can be set to any integer between one and the number of operational non-primary provider nodes in the device group.

This setting is an important factor in balancing cluster performance and availability. For example, increasing the desired number of secondaries increases the device group's opportunity to survive multiple failures that occur simultaneously within a cluster. Increasing the number of secondaries also decreases performance regularly during normal operation. A smaller number of secondaries typically results in better performance, but reduces availability. However, a larger number of secondaries does not always result in greater availability of the file system or device group in question. Refer to Chapter 3, "Key Concepts for System Administrators and Application Developers," in *Sun Cluster Concepts Guide for Solaris OS* for more information.

Steps

1. **Become superuser on any node of the cluster.**
2. **Run 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 change key properties of a device group, type `6` (Change key properties of a device group).**

   The Change Key Properties Menu is displayed.
5. **To change the desired number of secondaries, type `2` (Change the `numsecondaries` property).**

   Follow the instructions and type the desired number of secondaries to be configured for the disk device group. After an appropriate value has been typed, the corresponding `scconf` command is executed. Following, a log is printed, and the user is returned to the previous menu.
6. **Validate the device group configuration by using the `scconf -p` command.**

```
# scconf -p | grep Device
Device group name:                         dg-schost-1
Device group type:                         VxVM
Device group failback enabled:             yes
Device group node list:                    phys-schost-1,phys-schost-2, phys-schost-3
Device group ordered node list:            yes
Device group desired number of secondaries: 1
Device group diskset name:                 dg-schost-1
```

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 77.

7. **Verify the primary node and status for the disk device group.**

```
# scstat -D
```

Example 4–7 SPARC: Setting the Desired Number of Secondaries (VERITAS Volume Manager)

The following example shows the `scconf` command that is generated by `scsetup` when it configures the desired number of secondaries for a device group (`diskgrp1`). See "How to Change the Desired Number of Secondaries for a Device Group" on page 100 for information about changing the desired number of secondaries after a device group is created.

```
# scconf -a -D type=vxvm,name=diskgrp1,nodelist=host1:host2:host3\
,preferenced=true,failback=enabled,numsecondaries=2
```

▼ SPARC: 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 "SPARC: How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 91.

Steps

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 node
```

`-z`	Performs the switch.
`-D` *disk-device-group*	Specifies the device group to switch.
`-h` *node*	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. The `-r` option makes the operation recursive. The `-f` option is required to remove an enabled volume.

5. **Register the disk device group configuration changes to update the global namespace, using `scsetup`(1M).**

   See "SPARC: How to Register Disk Group Configuration Changes (VERITAS Volume Manager)" on page 91.

▼ SPARC: 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.

Steps

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 4–8 SPARC: 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` command generated by `scsetup` when it removes and unregisters the disk device group.

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

   scconf -r -D name=dg1
```

▼ SPARC: 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

Steps

1. **Become superuser on any node of the cluster.**
2. **At the prompt, type the `scsetup` command.**

   ```
   # 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 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 4–9 SPARC: 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
```

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

Use this procedure to remove a cluster node from the list of potential primaries of a VERITAS Volume Manager (VxVM) disk device group (disk group).

Steps

1. **Verify that the node is still a member of the group and that the group is an VxVM device group.**

   Device group type `VxVM` indicates a VxVM disk device group.

```
phys-schost-1% scconf -pv | grep '(global-galileo)'
  (global-galileo) Device group type:              VxVM
  (global-galileo) Device group failback enabled:  no
  (global-galileo) Device group node list:         phys-schost-1, phys-schost-2
  (global-galileo) Diskset name:                   global-galileo
phys-schost-1%
```

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

3. **Execute the `scsetup`(1M) command.**

   ```
   # 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(s).**

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

Example 4–10 SPARC: Removing a Node From a Disk Device Group (VxVM)

This example shows removal of the node named `phys-schost-1` from the `dg1` VxVM disk device group.

```
[Determine the VxVM disk device group for the node:]
# scconf -p | grep Device
  Device group name:                 dg1
    Device group type:               VxVM
    Device group failback enabled:   no
    Device group node list:          phys-schost-1, phys-schost-2
    Device group diskset name:        dg1
[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-1

    Command completed successfully.
Quit the scsetupDevice 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-2
    Device group diskset name:        dg1
```

▼ SPARC: How to Remove a Node From a Raw Disk Device Group

Use this procedure to remove a cluster node from the list of potential primaries of a VERITAS Volume Manager (VxVM) disk device group (disk group).

Use this procedure to remove a cluster node from the list of potential primaries of a raw disk device group.

Steps

1. **Become superuser on a node in the cluster other than the node to remove.**
2. **Identify the disk device groups that are connected to the node being removed.**
   Look for the node name in the `Device group node list` entry.

   ```
   # scconf -pvv | grep nodename | grep "Device group node list"
   ```

3. **Determine which disk device groups identified in Step 2 are raw disk device groups.**
   Raw disk device groups are of the `Disk` or `Local_Disk` device group type.

   ```
   # scconf -pvv | grep group-type
   ```

4. **Disable the `localonly` property of each `Local_Disk` raw disk device group.**

   ```
   # scconf -c -D name=rawdisk-device-group,localonly=false
   ```

   See the `scconf_dg_rawdisk`(1M) man page for more information about the `localonly` property.
5. **Verify that you have disabled the `localonly` property of all raw disk device groups that are connected to the node being removed.**
   The `Disk` device group type indicates that the `localonly` property is disabled for that raw disk device group.

   ```
   # scconf -pvv | grep group-type
   ```

6. **Remove the node from all raw disk device groups identified in Step 3.**
   You must complete this step for each raw disk device group that is connected to the node being removed.

   ```
   # scconf -r -D name=rawdisk-device-group,nodelist=nodename
   ```

Example 4–11 SPARC: Removing a Node From a Raw Disk Device Group

This example shows how to remove a node (`phys-schost-2`) from a raw disk device group. All commands are run from another node of the cluster (`phys-schost-1`).

```
[Identify the disk device groups connected to the node being removed:]
phys-schost-1# scconf -pvv | grep phys-schost-2 | grep "Device group node list"
    (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
[Identify the raw disk device groups:]
phys-schost-1# scconf -pvv | grep Disk
    (dsk/d4) Device group type:          Local_Disk
    (dsk/d8) Device group type:          Local_Disk
[Disable the localonly flag for each local disk on the node:]
phys-schost-1# scconf -c -D name=dsk/d4,localonly=false
[Verify that the localonly flag is disabled:]
phys-schost-1# scconf -pvv | grep Disk
    (dsk/d4) Device group type:          Disk
    (dsk/d8) Device group type:          Local_Disk
[Remove the node from all raw disk device groups:]
phys-schost-1# scconf -r -D name=dsk/d4,nodelist=phys-schost-2

phys-schost-1# scconf -r -D name=dsk/d2,nodelist=phys-schost-2

phys-schost-1# scconf -r -D name=dsk/d1,nodelist=phys-schost-2
```

▼ 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/Solaris Volume Manager 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.

Steps

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

2. **Run the `scsetup` command.**

   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 key properties of a device group, type `6` (Change key properties of a VxVM or Solstice DiskSuite/Solaris Volume Manager device group).**

   The Change Key Properties Menu is displayed

5. **To change a device group property, type `1` (Change the preferenced and/or failback properties).**

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

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

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

   ```
   # scconf -p
   ```

Example 4–12 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 -Dname=dg-schost-1,nodelist=phys-schost-1:phys-schost-2, \
preferenced=true,failback=enabled,numsecondaries=1

# 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 desired number of secondaries: 1
   Device group diskset name:                   dg-schost-1
```

▼ How to Change the Desired Number of Secondaries for a Device Group

The default number of secondary nodes for a device group is set to one. This setting specifies the number of nodes within a device group that can become primary owner of the group if the primary node fails. The desired number of secondaries value can be set to any integer between one and the number of non-primary provider nodes in the device group.

If the `numsecondaries` property is changed, secondary nodes are added or removed from the device group if the change causes a mismatch between the actual number of secondaries and the desired number.

This procedure uses `scsetup`(1M) to set or unset the `numsecondaries` property for Solstice DiskSuite/Solaris Volume Manager or VxVM disk device groups. Refer to `scconf_dg_rawdisk`(1M), `scconf_dg_sds`(1M), `scconf_dg_svm`(1M), and `scconf_dg_vxvm`(1M) for information about disk device group options when configuring any device group.

Steps

1. **Become superuser on any node of the cluster.**
2. **Run the `scsetup` utility.**

   ```
   # scsetup
   ```

   The Main Menu is displayed.
3. **To work with disk device groups, type `5` (Device groups and volumes).**

   The Device Groups Menu is displayed.
4. **To change key properties of a device group, type `6` (Change key properties of a device group).**

   The Change Key Properties Menu is displayed.
5. **To change the desired number of secondaries, type `2` (Change the `numsecondaries` property).**

   Follow the instructions and type the desired number of secondaries to be configured for the disk device group. After an appropriate value has been entered, the corresponding `scconf` command is executed, a log is printed, and the user returns to the previous menu.
6. **Verify that the disk device group attribute has been changed.**

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

   ```
   # scconf -p
   ```

Example 4–13 Changing the Desired Number of Secondaries

The following example shows the `scconf` command that is generated by `scsetup` when it configures the desired number of secondaries for a device group (`dg-schost-1`). This example assumes that the disk group and volume were created previously.

```
# scconf -c -D name=phys-host-1,nodelist=phys-schost-1:phys-schost-2,phys-schost-3 \
preferenced=true,failback=enabled,numsecondaries=1

# scconf -p | grep Device
Device group name:                              dg-schost-1
   Device group type:                           SDS/SVM
   Device group failback enabled:               yes
   Device group node list:                      phys-schost-1, phys-scost-2, phys-schost-3
```

```
   Device group ordered node list:             yes
   Device group desired number of secondaries: 1
   Device group diskset name:                  dg-schost-1
```

The following example shows use of a null string value to configure the default number of secondaries. The device group will be configured to use the default value, even if the default value changes.

```
# scconf -c -D name=diskgrp1, nodelist=host1:host2:host3, \
preferenced=false,failback=enabled,numsecondaries=
 # scconf -p | grep Device
Device group name:                          dg-schost-1
   Device group type:                       SDS/SVM
   Device group failback enabled:           yes
   Device group node list:                  phys-schost-1, phost-2, phys-schost-3
   Device group ordered node list:          yes
   Device group desired number of secondaries: 1
   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.

Step

- **Use one method from the following list.**
  - **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 4–14 Listing the Disk Device Group Configuration With `scstat`

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

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

```
  Device group status:        phys-schost-2          Offline
  Device group status:        phys-schost-1          Online
  Device group status:        phys-schost-3          Offline
```

Example 4–15 Listing the Disk Device Group Configuration With `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/SVM
    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.

Steps

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 node
   ```

`-z`	Performs the switch.
`-D` *disk-device-group*	Specifies the device group to switch.
`-h` *node*	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 4–16 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:    dg-schost-1               phys-schost-1       phys-schost-2

-- Device Group Status --
                              Device Group             Status
                             ------------             ------
  Device group status:        dg-schost-1              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 acquiesced 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 disk set or disk group is being repaired on another node.

Note – Before a device group can be placed in maintenance state, all access to its devices must be stopped, and all dependent file systems must be unmounted.

Steps

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 disk set or disk group, manually import that disk set or disk group.**

   For Solstice DiskSuite/Solaris Volume Manager:

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

Caution – If you are taking ownership of a Solstice DiskSuite/Solaris Volume Manager disk set, 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 disk set or disk group.**

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

- For Solstice DiskSuite/Solaris Volume Manager:

  ```
  # 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 node
   ```

Example 4–17 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 disk set or disk group.]
For Solstice DiskSuite/Solaris Volume Manager:
  # 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/Solaris Volume Manager:
  # 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

The cluster file system is a globally available file system that can be read and accessed from any node of the cluster.

TABLE 4–3 Task Map: Administering Cluster File Systems

Task	For Instructions, Go To…
Add cluster file systems after the initial Sun Cluster installation - Use `newfs`(1M) and `mkdir`	"How to Add a Cluster File System" on page 106
Remove a cluster file system - Use `fuser`(1M) and `umount`(1M)	"How to Remove a Cluster File System" on page 110
Check global mount points in a cluster for consistency across nodes - Use `sccheck`(1M)	"How to Check Global Mounts in a Cluster" on page 112

▼ 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.
- Volume manager software must be installed and configured on the cluster.

- A device group (Solstice DiskSuite/Solaris Volume Manager device group or a VxVM device group) or block disk slice must exist, upon which to create the cluster file system.

If you used SunPlex Manger to install data services, one or more cluster file systems already exist if there were sufficient shared disks on which to create the cluster file systems.

Steps

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` command.**

   Note – The `newfs` command is only valid for creating new UFS file systems. To create a new VxFS file system, follow procedures provided in your VxFS documentation

   ```
   # 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.

If Your Volume Manager Is …	A Disk Device Name Might Be …	Description
Solstice DiskSuite/Solaris Volume Manager	`/dev/md/oracle/rdsk/d1`	Raw disk device `d1` within the `oracle` disk set.
SPARC: 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.

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.

```
# mkdir -p /global/device-group mountpoint
```

device-group	Name of the directory that corresponds to the name of the device group that contains the device.
mountpoint	Name of the directory on which to mount the cluster file system.

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/Solaris Volume Manager trans metadevice or transactional volume –** Use the `global` mount option (do not use the `logging` mount option). See your Solstice DiskSuite/Solaris Volume Manager documentation for information about setting up trans metadevices and transactional volumes.

Note – Transactional volumes are scheduled to be removed from the Solaris operating system in an upcoming Solaris release. Solaris UFS logging, available since the Solaris 8 release, provides the same capabilities but superior performance, as well as lower system administration requirements and overhead.

- **VxFS logging –** Use the `global` and `log` mount options. See the `mount_vxfs` man page that is provided with VxFS software for more information.

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 mountpoint
```

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

You can use either the `df` or `mount` command to list mounted file systems.

To manage a VxFS cluster file system in a Sun Cluster environment, run administrative commands only from the primary node on which the VxFS cluster file system is mounted.

Example 4–18 Adding a Cluster File System

The following example creates a UFS cluster file system on the Solstice DiskSuite/Solaris Volume Manager metadevice or volume `/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  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]

[on one node:]
# sccheck
# mount /dev/md/oracle/dsk/d1 /global/oracle/d1
# mount
...
/global/oracle/d1 on /dev/md/oracle/dsk/d1 read/write/setuid/global/logging/largefiles
on Sun Oct 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.

Steps

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 ] mountpoint
   ```

`-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.
mountpoint	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 mountpoint
   ```

   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 mountpoint
   ```

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

   ```
   # umount mountpoint
   ```

mountpoint	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/volume/plex`.**

   See your volume manager documentation for more information.

Example 4–19 Removing a Cluster File System

The following example removes a UFS cluster file system mounted on the Solstice DiskSuite/Solaris Volume Manager metadevice or volume/dev/md/oracle/rdsk/d1.

```
# mount -v
...
/global/oracle/d1 on /dev/md/oracle/dsk/d1 read/write/setuid/global/logging/largefiles
# 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.]
```

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.

Steps

1. **Become superuser on any node in the cluster.**
2. **Check the cluster global mounts.**

   ```
   # sccheck
   ```

Administering Disk-Path Monitoring

Disk path monitoring (DPM) administration commands enable you to receive notification of secondary disk-path failure. Use the procedures in this section to perform administrative tasks that are associated with monitoring disk paths. Refer to Chapter 3, "Key Concepts for System Administrators and Application Developers," in *Sun Cluster Concepts Guide for Solaris OS* for conceptual information about the disk-path monitoring daemon. Refer to the `scdpm`(1M) man page for a description of the `scdpm` command options and related commands. Refer to the `syslogd`(1M) man page for logged errors that are reported by the daemon.

Note – Disk paths are automatically added to the monitoring list monitored when I/O devices are added to a node by using the `scgdevs` or `scdidadm` commands. Disk paths are also automatically unmonitored when devices are removed from a node by using Sun Cluster commands.

TABLE 4–4 Task Map: Administering Disk-Path Monitoring

Task	For Instructions
Monitor a disk path by using the `scdpm` command	"How to Monitor a Disk Path" on page 114
Unmonitor a disk path by using the `scdpm` command	"How to Unmonitor a Disk Path" on page 115
Print the status of faulted disk paths for a node by using `scdpm`	"How to Print Faulted Disk Paths" on page 116
Monitor or unmonitor disk paths from a file by using `scdpm -f`	"How to Monitor Disk Paths From a File" on page 117

The procedures in the following section issue the `scdpm` command with the disk-path argument. The disk-path argument is always constituted of a node name and a disk name. The node name is not required and defaults to `all` if none is specified. The following table describes the conventions that are used for naming the disk path.

Note – Use of the global disk-path name is strongly recommended, because the global disk-path name is consistent throughout the cluster. The UNIX disk-path name is not consistent throughout the cluster. The UNIX disk path for one disk can differ from cluster node to cluster node. The disk path could be `c1t0d0` on one node and `c2t0d0` on another node. If you use UNIX disk-path names, use the `scdidadm -L` command to map the UNIX disk-path name to the global disk-path name before issuing DPM commands. See the `scdidadm`(1M) man page.

TABLE 4–5 Sample Disk-Path Names

Name Type	Sample Disk-Path Name	Description
Global disk path	`phys-schost-1:/dev/did/dsk/d1`	Disk path `d1` on the `phys-schost-1` node
	`all:d1`	Disk path `d1` on all nodes in the cluster
UNIX disk path	`phys-schost-1:/dev/rdsk/c0t0d0s0`	Disk path `c0t0d0s0` on the `phys-schost-1` node
	`phys-schost-1:all`	All disk paths on all nodes of the cluster

▼ How to Monitor a Disk Path

Perform this task to monitor disk paths in your cluster.

Caution – DPM is not supported on nodes that run versions that were released prior to Sun Cluster 3.1 10/03 software. Do not use DPM commands while a rolling upgrade is in progress. After all nodes are upgraded, the nodes must be online to use DPM commands.

Steps

1. **Become superuser on any node in the cluster.**
2. **Monitor a disk path by using the `scdpm` command.**

   # **`scdpm -m`** *node:disk path*

   Refer to for naming conventions for the *node:disk path* argument.
3. **Verify that the disk path is monitored.**

   # **`scdpm -p`** *node*`:all`

Example 4–20 Monitoring a Disk Path on a Single Node

The following example monitors the `schost-1:/dev/did/rdsk/d1` disk path from a single node. Only the DPM daemon on the node `schost-1` monitors the path to the disk `/dev/did/dsk/d1` .

```
# scdpm -m schost-1:d1
# scdpm -p schost-1:d1
         schost-1:/dev/did/dsk/d1   Ok
```

Example 4–21 Monitoring a Disk Path on All Nodes

The following example monitors the `schost-1:/dev/did/dsk/d1` disk path from all nodes. DPM starts on all nodes for which `/dev/did/dsk/d1` is a valid path.

```
# scdpm -m all:/dev/did/dsk/d1
# scdpm -p schost-1:d1
         schost-1:/dev/did/dsk/d1   Ok
```

Example 4–22 Rereading the Disk Configuration From the CCR

The following example forces the daemon to reread the disk configuration from the CCR and prints the monitored disk paths with status.

```
# scdpm -m all:all
# scdpm -p all:all
         schost-1:/dev/did/dsk/d4   Ok
         schost-1:/dev/did/dsk/d3   Ok
         schost-2:/dev/did/dsk/d4   Fail
         schost-2:/dev/did/dsk/d3   Ok
         schost-2:/dev/did/dsk/d5   Unmonitored
         schost-2:/dev/did/dsk/d6   Ok
```

▼ How to Unmonitor a Disk Path

Use this procedure to unmonitor a disk path.

Caution – DPM is not supported on nodes that run versions that were released prior to Sun Cluster 3.1 10/03 software. Do not use DPM commands while a rolling upgrade is in progress. After all nodes are upgraded, the nodes must be online to use DPM commands.

Steps

1. **Become superuser on any node in the cluster.**
2. **Determine the state of the disk path to unmonitor.**

   # **scdpm -p [all:]** *disk path*

`-p`	Prints a detailed listing of the current status of a specified disk path
`[:all]`	Displays all monitored and unmonitored disk paths

3. **On each node, unmonitor the appropriate disk paths.**

   `#` **`scdpm -u`** *node:disk path*

   Refer to Table 4–5 for naming conventions for the *node:disk path* argument.

Example 4–23 Unmonitoring a Disk Path

The following example unmonitors the `schost-2:/dev/did/rdsk/d1` disk path and prints disk paths with status for the entire cluster.

```
# scdpm -u schost-2:/dev/did/rdsk/d1
# scdpm -p all:all
         schost-1:/dev/did/dsk/d4   Ok
         schost-1:/dev/did/dsk/d3   Ok
         schost-2:/dev/did/dsk/d4   Fail
         schost-2:/dev/did/dsk/d3   Ok
         schost-2:/dev/did/dsk/d1   Unmonitored
         schost-2:/dev/did/dsk/d6   Ok
```

▼ How to Print Faulted Disk Paths

Use the following procedure to print the faulted disk paths for a cluster.

Caution – DPM is not supported on nodes that run versions that were released prior to Sun Cluster 3.1 10/03 software. Do not use DPM commands while a rolling upgrade is in progress. After all nodes are upgraded, the nodes must be online to use DPM commands.

Steps

1. **Become superuser on any node in the cluster.**
2. **Print the faulted disk paths throughout the cluster.**

   `#` **`scdpm -p -F`** *node:disk path*

   Refer to Table 4–5 for naming conventions for the *node:disk path* argument.

Example 4–24 Printing Faulted Disk Paths

The following example prints faulted disk paths for the entire cluster.

```
# scdpm -p -F [all:]all
         schost-1:/dev/did/dsk/d4   Fail
         schost-1:/dev/did/dsk/d3   Fail
         schost-2:/dev/did/dsk/d4   Fail
         schost-2:/dev/did/dsk/d3   Fail
         schost-2:/dev/did/dsk/d5   Fail
         schost-2:/dev/did/dsk/d6   Fail
```

▼ How to Monitor Disk Paths From a File

Use the following procedure to monitor or unmonitor disk paths from a file. The file must list the commands to monitor or unmonitor, the node names, and the disk-path names. Each field of the file should be separated by a column. Format each listing by following the example.

```
syntax in command file:
[u,m] [node|all]:<[/dev/did/rdsk/]d- | [/dev/rdsk/]c-t-d- | all>


command file entry
u schost-1:/dev/did/rdsk/d5
m schost-2:all
```

Caution – DPM is not supported on nodes that run versions that were released prior to Sun Cluster 3.1 10/03 software. Do not use DPM commands while a rolling upgrade is in progress. After all nodes are upgraded, the nodes must be online to use DPM commands.

Steps

1. **Become superuser on any node in the cluster.**
2. **Monitor the disk paths from a file.**

   # **scdpm -f** *filename*
3. **Verify the disk paths with status for the cluster.**

   # **scdpm -p all:all**

Example 4–25 Monitor or Unmonitor Disk Paths From a File

The following example monitors or unmonitors disk paths from a file.

```
# scdpm -f schost_config
# scdpm -p all:all
         schost-1:/dev/did/dsk/d4   Ok
         schost-1:/dev/did/dsk/d3   Ok
```

```
schost-2:/dev/did/dsk/d4   Fail
schost-2:/dev/did/dsk/d3   Ok
schost-2:/dev/did/dsk/d5   Unmonitored
schost-2:/dev/did/dsk/d6   Ok
```

CHAPTER 5

Administering Quorum

This chapter provides the procedures for administering quorum within Sun Cluster. For information on quorum concepts, see "Quorum and Quorum Devices" in *Sun Cluster Concepts Guide for Solaris OS*

This is a list of the procedures in this chapter.

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

See the *Sun Cluster Concepts Guide for Solaris OS* 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` command is an interactive interface to the `scconf` command. When `scsetup` runs, the command 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 5–1 Task List: Administering Quorum

Task	For Instructions, Go To...
Add a quorum device to a cluster - Use `scsetup`(1M)	"Adding a Quorum Device" on page 121
Remove a quorum device from a cluster - Use `scsetup` (to generate scconf)	"How to Remove a Quorum Device" on page 126
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 128
Replace a quorum device in a cluster - Use the add and remove procedures	"How to Replace a Quorum Device" on page 129
Modify a Quorum Device List - Use the add and remove procedures	"How to Modify a Quorum Device Node List" on page 130
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 132
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 133

TABLE 5–1 Task List: Administering Quorum *(Continued)*

Task	For Instructions, Go To...
List the quorum devices and vote counts - Use `scconf`(1M)	"How to List the Quorum Configuration" on page 135

Dynamic Reconfiguration With Quorum Devices

There are a few issues you must consider when completing dynamic reconfiguration (DR) operations on quorum devices in a cluster.

- All of the requirements, procedures, and restrictions that are documented for the Solaris DR feature also apply to Sun Cluster DR support (except for the operating environment quiescence operation). Therefore, review the documentation for the Solaris DR feature *before* using the DR feature with Sun Cluster software. You should review in particular the issues that affect non-network IO devices during a DR detach operation.
- Sun Cluster rejects DR remove-board operations that are performed when an interface is present that is configured for a quorum device.
- If the DR operation would pertain to an active device, Sun Cluster rejects the operation and identifies the devices that would be affected by the operation.

To remove a quorum device, you must complete the following steps, in the order indicated.

TABLE 5–2 Task Map: Dynamic Reconfiguration with Quorum Devices

Task	For Instructions, Go To...
1. Enable a new quorum device to replace the one being removed.	"Adding a Quorum Device" on page 121
2. Disable the quorum device to be removed.	"How to Remove a Quorum Device" on page 126
3. Perform the DR remove operation on the device being removed.	*Sun Enterprise 10000 DR Configuration Guide* and the *Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual* (from the *Solaris 8 on Sun Hardware* and *Solaris 9 on Sun Hardware* collections.)

Adding a Quorum Device

This section provides procedures to add a quorum device. For information on determining the number of quorum vote counts necessary for your cluster, recommended quorum configurations, and failure fencing, see "Quorum and Quorum Devices" in *Sun Cluster Concepts Guide for Solaris OS*.

Sun Cluster supports two types of quorum devices: SCSI and Network Appliance (NetApp) NAS. Procedures for adding these devices are provided in the following sections:

- "How to Add a SCSI Quorum Device" on page 122
- "How to Add a Network Appliance Network-Attached Storage (NAS) Quorum Device" on page 123

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

See the `scsetup`(1M) and `scconf`(1M) man pages for information on the commands used in the following procedures.

▼ How to Add a SCSI Quorum Device

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

Steps

1. **Become superuser on any node of the cluster.**
2. **Start the `scsetup` utility.**

   ```
   # scsetup
   ```

   The `scsetup` Main Menu is displayed.
3. **Type the number that corresponds to the option for Quorum.**

   The Quorum Menu is displayed.
4. **Type the number that corresponds to the option for adding a quorum device, then type `yes` when the scsetup utility asks you to confirm that you want to add a quorum device.**

   The `scsetup` utility asks what type of quorum device you want to add.
5. **Type the number that corresponds to the option for a SCSI quorum device.**

   The `scsetup` utility asks which global device you want to use .
6. **Type the global device you are using.**

   The `scsetup` utility asks you to confirm that the new quorum device should be added the global device you specified.
7. **Type `yes` to continue adding the new quorum device.**

   If the new quorum device is added successfully, the `scsetup` utility displays a message to that affect.
8. **Verify that the quorum device has been added.**

   ```
   # scstat -q
   ```

9. **Repeat Step 3 through Step 8 for each group of nodes that shares a storage enclosure.**

Example 5–1 Adding a SCSI Quorum Device

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

Become superuser on any cluster node.

[*Execute the scsetup utility:*]
scsetup
Select Quorum>Add a quorum device .
Select SCSI device.
Type the global device your are using.
d20
Type yes to continue adding the new quorum device.

yes
[*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 added:*]
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/d20s1  1        1       Online
  Device votes:  /dev/did/rdsk/d20s2  1        1       Online
```

▼ How to Add a Network Appliance Network-Attached Storage (NAS) Quorum Device

When you use a Network Appliance (NetApp) network-attached storage (NAS) device as a quorum device, the following are required:

- You must install the iSCSI license from NetApp.
- You must configure an iSCSI LUN on the clustered filer for use as the quorum device.
- You must configure the NetApp NAS unit to use NTP for synchronizing time.
- At least one of the NTP servers selected for the clustered filer must be an NTP server for the Sun Cluster nodes.
- When booting the cluster, always boot the NAS device before you boot the cluster nodes.

  If you boot devices in the wrong order, your nodes cannot find the quorum device. If a node should fail in this situation, your cluster might be unable to remain in service. If this happens, you must either reboot the entire cluster or remove the NetApp NAS quorum device and add it again.
- A cluster can use a NAS device for only a single quorum device.

  You can configure other shared storage if you need additional quorum devices. Additional clusters using the same NAS device can use separate LUNs on that device as their quorum devices.

See the following Network Appliance NAS documentation for information about creating and setting up a Network Appliance NAS device and LUN. You can access the following documents at `http://now.netapp.com`.

- Setting up a NAS device

  System Administration File Access Management Guide
- Setting up a LUN

  Host Cluster Tool for Unix Installation Guide
- Installing ONTAP software

  Software Setup Guide, *Upgrade Guide*
- Exporting volumes for the cluster

  Data ONTAP Storage Management Guide
- Installing NAS support software packages on cluster nodes

  Log in to http://now.netapp.com. From the Software Download page, download the *Host Cluster Tool for Unix Installation Guide*.

See the following Sun Cluster documentation for information on installing a NetApp NAS storage device in a Sun Cluster environment: *Sun Cluster 3.1 With Network-Attached Storage Devices Manual for Solaris OS*.

Steps

1. **Make sure that all Sun Cluster nodes are online and can communicate with the NetApp clustered filer.**
2. **Become superuser on any node of the cluster.**

3. **Start the `scsetup` utility.**

   ```
   # scsetup
   ```

   The `scsetup` Main Menu is displayed.

4. **Type the number that corresponds to the option for Quorum.**

   The Quorum Menu is displayed.

5. **Type the number that corresponds to the option for adding a quorum device, then type `yes` when the `scsetup` utility asks you to confirm that you want to add a quorum device.**

   The `scsetup` utility asks what type of quorum device you want to add.

6. **Type the number that corresponds to the option for a netapp_nas quorum device, then type `yes` when the `scsetup` utility asks you to confirm that you want to add a netapp_nas quorum device.**

   The `scsetup` utility asks you provide the name of the new quorum device.

7. **Type the name of the quorum device you are adding.**

   The quorum device name can be any name you choose. The name is only used to process future administrative commands.

   The `scsetup` utility asks you to provide the name of the filer for the new quorum device.

8. **Type the name of the filer of the new quorum device.**

   This name is the network accessible name or address of the filer.

   The `scsetup` utility asks you to provide the LUN ID for the filer.

9. **Type the ID of the quorum device LUN on the filer.**

   The `scsetup` utility asks you to confirm that the new quorum device should be added on the filer.

10. **Type `yes` to continue adding the new quorum device.**

    If the new quorum device is added successfully, the `scsetup` utility displays a message to that affect.

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

    ```
    # scstat -q
    ```

12. **Repeat Step 4 through Step 11 for each group of nodes that shares a storage enclosure.**

Example 5–2 Adding a NetApp NAS Quorum Device

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

Make sure that all Sun Cluster nodes are online and can communicate with the NetApp clustered filer.

Become superuser on any cluster node.
[*Execute the scsetup utility:*]
scsetup
Select Quorum>Add a quorum device .
Select Netapp_nas quorum device.
Type the name of the quorum device you are adding.

qd1
Type the name of the filer of the new quorum device.

nas1.sun.com
Type the ID of the quorum device LUN on the filer.
0
Type yes to continue adding the new quorum device.

yes
[*Verify that the scconf command completed successfully:*]

```
scconf -a -q name=qd1,type-=netapp_nas,filer=nas1.sun.com,lun_id=0

    Command completed successfully.
```

Quit the scsetup Quorum Menu and Main Menu.
[*Verify the quorum device is added:*]
scstat -q

```
-- Quorum Summary --

  Quorum votes possible: 5
  Quorum votes needed:   3
  Quorum votes present:  5

-- 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:     qd1                 1        1       Online
 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 device 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 128.

Steps

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. **Type the number that corresponds to the option for Quorum.**
5. **Type the number that corresponds with the option to remove a quorum device.**
   Answer the questions displayed during the removal process.
6. **Quit `scsetup`.**
7. **Verify that the quorum device is removed.**

   ```
   # scstat -q
   ```

Example 5–3 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 device .
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 126.

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.

Steps

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 161.

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

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

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

   The `scsetup`(1M) cluster-administration menu options are not available while the cluster is in install mode.

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

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

   ```
   # scstat -q
   ```

Example 5–4 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 name=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

Use this procedure to replace an existing quorum device with another quorum device. You can replace a quorum device with a similar device type, such as replacing a NAS device with another NAS device, or you can replace the device with a dissimilar device, such as replacing a NAS device with a shared disk.

Steps

1. **Configure a new quorum device.**

   You need to first add a new quorum device to the configuration to take the place of the old device. See "Adding a Quorum Device" on page 121 to add a new quorum device to the cluster.

2. **Remove the device that you are replacing as a quorum device.**

   See "How to Remove a Quorum Device" on page 126 to remove the old quorum device from the configuration.

3. **If the quorum device is a failed disk, replace the disk.**

   Refer to the hardware procedures for your disk enclosure in the *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*.

▼ How to Modify a Quorum Device Node List

You can use the `scsetup`(1M) utility to add a node to or remove a node from the node list of an existing quorum device. To modify a quorum device's node list, you must remove the quorum device, modify the physical connections of nodes to the quorum device you removed, then add the quorum device to the cluster configuration again. When a quorum device is added, `scconf`(1M) automatically configures the node-to-disk paths for all nodes attached to the disk.

Steps

1. **Become superuser on any node of the cluster.**
2. **Determine the name of the quorum device you are modifying.**

   ```
   # scconf -p | grep Quorum
   ```

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

   ```
   # scsetup
   ```

   The Main Menu is displayed.
4. **Type the number that corresponds to the Quorum option.**

   The Quorum Menu is displayed.
5. **Type the number that corresponds to the option to remove a quorum device.**

   Follow the instructions. You will be asked the name of the disk to be removed.
6. **Add or delete the physical node connections to the quorum device.**
7. **Type the number that corresponds to the option to add a quorum device.**

   Follow the instructions. You will be asked the name of the disk to be used as the quorum device.
8. **Verify that the quorum device has been added.**

   ```
   # scstat -q
   ```

Example 5–5 Modifying a Quorum Device Node List

The following example shows how to use the `scsetup` utility to add nodes to or delete nodes from a quorum device node list. In this example, the quorum device name is d2, and the final result of the procedures adds another node to the quorum device node list.

```
[Become superuser on any node in the cluster.
]
[Determine the quorum device name:]
# scconf -p | grep Quorum
Quorum devices:                                 d2
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):
```
[*Execute the utility:*]
```
# scsetup
```
Type the number that corresponds with the quorum option.
.
Type the number that corresponds with the option to remove a quorum device.
.
Answer the questions when prompted.
You Will Need:
Example:
```
   quorum device name     d2
```
[*Verify that the scconf command completed successfully:*]
```
scconf -r -q globaldev=d2
Command completed successfully.
```
Type the number that corresponds with the Quorum option.
.
Type the number that corresponds with the option to add a quorum device.
.Answer the questions when prompted.
You Will Need: *Example:*
```
   quorum device name     d2

[Verify that the scconf command completed successfully:]

scconf -a -q globaldev=d2
Command completed successfully.
```
Quit the scsetup utility.
```
[Verify that the correct nodes have paths to the quorum device.
In this example, note that phys-schost-3 has been added to the
enabled hosts list.]
# scconf -p | grep Quorum
Quorum devices:                                   d2
Quorum device name:                               d2
  Quorum device votes:                            2
  Quorum device enabled:                          yes
  Quorum device name:                             /dev/did/rdsk/d2s2
  Quorum device hosts (enabled):                  phys-schost-1 phys-schost-2
                                                  phys-schost-3
  Quorum device hosts (disabled):
```
[*Verify that the modified quorum device is online.*]
```
# scstat -q
-- Quorum Votes by Device --
                    Device Name         Present Possible Status
                    -----------         ------- -------- ------
  Device votes:     /dev/did/rdsk/d2s2  1        1       Online
```
[*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 Put a Quorum Device Into Maintenance State

Use the `scconf`(1M) command to put a quorum device into maintenance state. The `scsetup`(1M) utility does not currently have this capability. 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 quorum device's quorum vote count is set to zero and does not contribute to the quorum count while the device is being serviced. While in maintenance state, the quorum device's configuration information is preserved.

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 161.

Steps

1. **Become superuser on any node of the cluster.**
2. **Put the quorum device into the maintenance state.**

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

   `-c` Specifies the change form of the `scconf` command.

`-q`	Manages the quorum options.
`name=`*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 5–6 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 name=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):
```

See Also To re-enable the quorum device, see "How to Bring a Quorum Device Out of Maintenance State" on page 133.

To put a node into maintenance state, see "How to Put a Node Into Maintenance State" on page 161.

▼ How to Bring a Quorum Device Out of Maintenance State

Run this procedure each time a quorum device is in maintenance state and you want to bring the quorum device out of maintenance state and reset the quorum vote count to the default.

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

When you configure a quorum device, Sun Cluster software assigns the quorum device a vote count of *N*-1 where *N* is the number of connected votes to the quorum device. For example, a quorum device that is connected to two nodes with nonzero vote counts has a quorum count of one (two minus one).

- 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 163.
- To learn more about quorum vote counts, see "About Quorum Vote Counts" in *Sun Cluster Concepts Guide for Solaris OS*.

Steps

1. **Become superuser on any node of the cluster.**
2. **Reset the quorum count.**

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

`-c`	Specifies the change form of the `scconf` command.
`-q`	Manages the quorum options.
`name=`*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 5–7 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 name=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.

Note – When you increase or decrease the number of node attachments to a quorum device, the quorum vote count is not automatically recalculated. You can reestablish the correct quorum vote if you remove all quorum devices and then add them back into the configuration. For a two-node cluster, temporarily add a new quorum device before you remove and add back the original quorum device. Then remove the temporary quorum device.

Step

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

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

Example 5–8 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):
```

▼ How to Repair a Quorum Device

Use this procedure to replace a malfunctioning quorum device.

Steps

1. **Remove the disk device that you are replacing as a quorum device.**

Note – If the device you intend to remove is the last quorum device, you might want to first add another disk as a new quorum device. This will ensure that there will be a valid quorum device in case there is a failure during the replacement procedure. See "Adding a Quorum Device" on page 121 to add a new quorum device.

See "How to Remove a Quorum Device" on page 126 to remove a disk device as a quorum device .

2. **Replace the disk device.**

   To replace the disk device, see the hardware procedures for the disk enclosure in the *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*.

3. **Add the replaced disk as a new quorum device.**

   See "Adding a Quorum Device" on page 121 to add a disk as a new quorum device.

Note – If you added an additional quorum device in Step 1, it is now safe to remove it. See "How to Remove a Quorum Device" on page 126 to remove the quorum device.

CHAPTER 6

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 Internet Protocol (IP) Network Multipathing groups, when you initially install and configure the cluster. If you later need to alter a cluster interconnect network configuration, you can use the software procedures in this chapter. For information about configuring IP Network Multipathing groups in a cluster, see the section "Administering the Public Network" on page 148.

This is a list of the procedures in this chapter.

For a high-level description of the related procedures in this chapter, see Table 6–1 and Table 6–3.

Refer to the *Sun Cluster Concepts Guide for Solaris OS* 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`(1M) utility to administer the cluster transport for the cluster interconnects. See the `scsetup` man page for more information.

For cluster software installation procedures, see the *Sun Cluster Software Installation Guide for Solaris OS*. For procedures about servicing cluster hardware components, see the *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*.

Note – You can usually choose to use the default port name, where appropriate, during cluster interconnect procedures. The default port name is the same as the internal node ID number of the node that hosts the adapter end of the cable. However, you cannot use the default port name for certain adapter types, such as SCI.

TABLE 6–1 Task List: Administering the Cluster Interconnect

Task	For Instructions, Go To...
Administer the cluster transport - Use `scsetup`(1M)	"How to Access the `scsetup` Utility" on page 20
Check the status of the cluster interconnect - Use `scstat`	"How to Check the Status of the Cluster Interconnect" on page 140
Add a cluster transport cable, transport adapter, or transport junction - Use `scstat`(1M)	"How to Add Cluster Transport Cables, Transport Adapters, or Transport Junctions" on page 141
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 142
Enable a cluster transport cable - Use `scsetup`	"How to Enable a Cluster Transport Cable" on page 145

TABLE 6–1 Task List: Administering the Cluster Interconnect *(Continued)*

Task	For Instructions, Go To...
Disable a cluster transport cable - Use `scsetup`	"How to Disable a Cluster Transport Cable" on page 146
Determining an transport adapter's instance number	"How to Determine a Transport Adapter's Instance Number" on page 147

Dynamic Reconfiguration With Cluster Interconnects

There are a few issues you must consider when completing dynamic reconfiguration (DR) operations on cluster interconnects.

- All of the requirements, procedures, and restrictions that are documented for the Solaris DR feature also apply to Sun Cluster DR support (except for the operating environment quiescence operation). Therefore, review the documentation for the Solaris DR feature *before* using the DR feature with Sun Cluster software. You should review in particular the issues that affect non-network IO devices during a DR detach operation.
- Sun Cluster rejects DR remove-board operations performed on active private interconnect interfaces.
- If the DR remove-board operation pertains to an active private interconnect interface, Sun Cluster rejects the operation and identifies the interface that would be affected by the operation.

Caution – Sun Cluster requires that each cluster node has at least one functioning path to every other cluster node. Do not disable a private interconnect interface that supports the last path to any cluster node.

Complete the following procedures in the order indicated when performing DR operations on public network interfaces.

TABLE 6–2 Task Map: Dynamic Reconfiguration with Public Network Interfaces

Task	For Instructions, Go To...
1. Disable and remove the interface from the active interconnect	"Dynamic Reconfiguration With Public Network Interfaces" on page 150

TABLE 6–2 Task Map: Dynamic Reconfiguration with Public Network Interfaces *(Continued)*

Task	For Instructions, Go To...
2. Perform the DR operation on the public network interface.	<ulink type="text" href="http://docs.sun.com/app/docs/doc/816-3630">Sun Enterprise 10000 DR Configuration Guide</ulink> and the *Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual* (from the *Solaris 8 on Sun Hardware* and *Solaris 9 on Sun Hardware* collections)

▼ 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.

Steps

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 6–1 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

For information about the requirements for the cluster private transport, see "Interconnect Requirements and Restrictions" in *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*.

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

Steps

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

   For the procedure on installing a cluster transport cable, see the *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*.

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 `4` (Cluster interconnect).**

   Note – If your configuration uses SCI adapters, do not accept the default when you are prompted for the adapter connections (the port name) during the "Add" portion of this procedure. Instead, provide the port name (0, 1, 2, or 3) found on the Dolphin switch, to which the node is **physically** cabled.

   ---

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 6–2 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.]
# 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 – Each cluster node needs at least one functioning transport path to every other node in the cluster. No two nodes should be isolated from one another. 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.

Steps

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 `4` (Cluster interconnect).**

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

   Follow the instructions and enter the requested information. You will need to know the applicable node names, adapter names, and junction names.

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. You will need to know the applicable node names, adapter names, and junction names.

Note – If you are removing a physical adapter from a node, see the *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS* for hardware service procedures.

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

   Follow the instructions and enter the requested information. You will need to know the applicable node names, adapter names, and junction names.

   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 6–3 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 utility:]
# scsetup
Type 4 (Cluster interconnect).
Select either Remove a transport cable,
Remove a transport adapter to a node,
or Remove 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.

Steps

1. **Become superuser on any node in the cluster.**
2. **Enter the `scsetup`(1M) utility.**

   ```
   # scsetup
   ```

   The Main Menu is displayed.
3. **Access the Cluster Interconnect Menu by typing `4` (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 6–4 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.

Caution – Each cluster node needs at least one functioning transport path to every other node in the cluster. No two nodes should be isolated from one another. 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.

Steps

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`(1M) utility.**

```
# scsetup
```

The Main Menu is displayed.

4. **Access the Cluster Interconnect Menu by typing `4` (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 6–5 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
```

▼ How to Determine a Transport Adapter's Instance Number

You need to determine a transport adapter's instance number to ensure that you add and remove the correct transport adapter through the `scsetup` command. The adapter name is a combination of the type of the adapter and the adapter's instance number. This procedure uses an SCI-PCI adapter as an example.

Steps 1. **Based on the slot number, find the adapter's name.**

The following screen is an example and might not reflect your hardware.

```
# prtdiag
...
======================== IO Cards =========================
                           Bus  Max
 IO  Port Bus      Freq Bus  Dev,
Type  ID  Side Slot MHz  Freq Func State Name Model
---- ---- ---- ---- ---- ---- ---- ----- --------------------------------
PCI   8    B    2    33   33  2,0  ok    pci11c8,0-pci11c8,d665.11c8.0.0
PCI   8    B    3    33   33  3,0  ok    pci11c8,0-pci11c8,d665.11c8.0.0
...
```

2. **Using the adapter's name and slot number, find the adapter's instance number.**

The following screen is an example and might not reflect your hardware.

```
# prtconf
...
 pci, instance #0
             pci11c8,0, instance #0
            pci11c8,0, instance #1
...
```

Administering the Public Network

Sun Cluster 3.1 4/04 supports the Solaris implementation of Internet Protocol (IP) Network Multipathing for public networks. Basic IP Network Multipathing administration is the same for both cluster and non-cluster environments. Multipathing administration is covered in the appropriate Solaris documentation. However, review the guidelines that follow before administering IP Network Multipathing in a Sun Cluster environment.

How to Administer IP Network Multipathing Groups in a Cluster

Before performing IP Network Multipathing procedures on a cluster, consider the guidelines below.

- Each public network adapter must belong to a multipathing group.
- The `local-mac-address?` variable must have a value of `true` for Ethernet adapters.
- You must configure a test IP address for each adapter in the following kinds of multipathing groups:

- All multipathing groups in a cluster that runs on the Solaris 8 OS.
- All multiple-adapter multipathing groups in a cluster that runs on the Solaris 9 or Solaris 10 OS. Single-adapter multipathing groups on the Solaris 9 or Solaris 10 OS do not require test IP addresses.

- Test IP addresses for all adapters in the same multipathing group must belong to a single IP subnet.
- Test IP addresses must not be used by normal applications because they are not highly available.
- There are no restrictions on multipathing group naming. However, when configuring a resource group, the `netiflist` naming convention is any multipathing name followed by either the nodeID number or the node name. For example, given a multipathing group named `sc_ipmp0` , the `netiflist` naming could be either `sc_ipmp0@1` or `sc_ipmp0@phys-schost-1`, where the adapter is on the node `phys-schost-1` which has the nodeID of `1`.
- Avoid unconfiguring (unplumbing) or bringing down an adapter of an IP Network Multipathing group without first switching over the IP addresses from the adapter to be removed to an alternate adapter in the group, using the `if_mpadm`(1M) command.
- Avoid rewiring adapters to different subnets without first removing them from their respective multipathing groups.
- Logical adapter operations can be done on an adapter even if monitoring is on for the multipathing 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.
- To view the status of IP Network Multipathing groups on a cluster, use the command `scstat`(1M) with the `-i` option.

For more information about IP Network Multipathing, see the appropriate documentation in the Solaris system administration documentation set.

TABLE 6–3 Task Map: Administering the Public Network

Solaris Operating System Release	For Instructions, Go To...
Solaris 8 operating system	*IP Network Multipathing Administration Guide*
Solaris 9 operating system	"IP Network Multipathing Topics" in*System Administration Guide: IP Series*

For cluster software installation procedures, see the *Sun Cluster Software Installation Guide for Solaris OS*. For procedures about servicing public networking hardware components, see the *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*.

Dynamic Reconfiguration With Public Network Interfaces

There are a few issues you must consider when completing dynamic reconfiguration (DR) operations on public network interfaces in a cluster.

- All of the requirements, procedures, and restrictions that are documented for the Solaris DR feature also apply to Sun Cluster DR support (except for the operating environment quiescence operation). Therefore, review the documentation for the Solaris DR feature *before* using the DR feature with Sun Cluster software. You should review in particular the issues that affect non-network IO devices during a DR detach operation.
- DR remove-board operations can succeed only when public network interfaces are not active. Before removing an active public network interface, switch the IP addresses from the adapter to be removed to another adapter in the multipathing group, using the `if_mpadm`(1M) command.
- If you try to remove a public network interface card without having properly disabled it as an active network interface, Sun Cluster rejects the operation and identifies the interface that would be affected by the operation.

Caution – For multipathing groups with two adapters, if the remaining network adapter fails while you are performing the DR remove operation on the disabled network adapter, availability is impacted. The remaining adapter has no place to fail over for the duration of the DR operation.

Complete the following procedures in the order indicated when performing DR operations on public network interfaces.

TABLE 6–4 Task Map: Dynamic Reconfiguration with Public Network Interfaces

Task	For Instructions, Go To...
1. Switch the IP addresses from the adapter to be removed to another adapter in the multipathing group, using the `if_mpadm`.	`if_mpadm`(1M) man page. The appropriate Solaris documentation: Solaris 8: *IP Network Multipathing Administration Guide* Solaris 9: "IP Network Multipathing Topics" in *System Administration Guide: IP Services*

TABLE 6–4 Task Map: Dynamic Reconfiguration with Public Network Interfaces *(Continued)*

Task	For Instructions, Go To...
2. Remove the adapter from the multipathing group using the `ifconfig` command.	The appropriate Solaris documentation: Solaris 8: *IP Network Multipathing Administration Guide* Solaris 9: "IP Network Multipathing Topics" in *System Administration Guide: IP Services* `ifconfig`(1M) man page
3. Perform the DR operation on the public network interface.	<ulink type="text" href="http://docs.sun.com/app/docs/doc/816-3630">Sun Enterprise 10000 DR Configuration Guide</ulink> and the *Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual* (from the *Solaris 8 on Sun Hardware* and *Solaris 9 on Sun Hardware* collections)

CHAPTER 7

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 7–1 Task List: Administering the Cluster

Task	For Instructions, Go To
Change the name of the cluster	"How to Change the Cluster Name" on page 154

TABLE 7–1 Task List: Administering the Cluster *(Continued)*

Task	For Instructions, Go To
List node IDs and their corresponding node names	"How to Map Node ID to Node Name" on page 155
Permit or deny new nodes to add themselves to the cluster	"How to Work With New Cluster Node Authentication" on page 155
Change the time for a cluster using the Network Time Protocol (NTP)	"How to Reset the Time of Day in a Cluster" on page 156
Bring down a node to the OpenBoot PROM `ok` prompt on a SPARC based system or to a Boot Subsystem on an x86 based system	"SPARC: How to Enter the OpenBoot PROM (OBP) on a Node" on page 158
Change the private hostname	"How to Change the Private Hostname" on page 158
Put a cluster node in maintenance state	"How to Put a Node Into Maintenance State" on page 161
Bring a cluster node out of maintenance state	"How to Bring a Node Out of Maintenance State" on page 163
Add a node to a cluster	"How to Add a Node to the Authorized Node List" on page 167
Remove a node from a cluster	"How to Remove a Node From the Cluster Software Configuration" on page 169

▼ How to Change the Cluster Name

If necessary, you can change the cluster name after initial installation.

Steps

1. **Become superuser on any node in the cluster.**
2. **Type the `scsetup` command.**

   ```
   # scsetup
   ```

   The Main Menu is displayed.
3. **To change the cluster name, type `8` (Other cluster properties).**

   The Other Cluster Properties menu is displayed.
4. **Make your selection from the menu and follow the onscreen instructions.**

Example 7–1 Changing the Cluster Name

The following example shows the `scconf`(1M) command generated from the `scsetup`(1M) 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.

Step

- **Use the `scconf`(1M) command to list the cluster configuration information.**

```
% scconf -pv | grep "Node ID"
```

Example 7–2 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.

Steps

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 `7` (New nodes).**

The New Nodes menu is displayed.

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

Example 7–3 Preventing New Machines From Being Added to the Cluster

The following example shows the `scconf` command generated from the `scsetup` utility that would prevent new machines from being added to the cluster.

```
# scconf -a -T node=.
```

Example 7–4 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
```

Example 7–5 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
```

Example 7–6 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
```

Example 7–7 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
```

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 Concepts Guide for Solaris OS* 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), `xntpd`(1M), or `svcadm`(1M) commands interactively or within `cron`(1M) scripts.

Steps

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

2. **Shut down the cluster.**

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

3. **Verify that the node is showing the `ok` prompt or the `Select (b)oot or (i)nterpreter` prompt on the Current Boot Parameters screen.**

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

   - SPARC:

     ```
     ok boot -x
     ```

   - x86:

     ```
                            <<< Current Boot Parameters >>>
     Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
     sd@0,0:a
     Boot args:

     Type    b [file-name] [boot-flags] <ENTER>  to boot with options
     or      i <ENTER>                           to enter boot interpreter
     or      <ENTER>                             to boot with defaults

                       <<< timeout in 5 seconds >>>
     Select (b)oot or (i)nterpreter: b -x
     ```

5. **On a single node, set the time of day by running the `date` command.**

   ```
   # date HHMM.SS
   ```

6. **On the other machines, synchronize the time to that node by running the `rdate`(1M) command.**

   ```
   # rdate hostname
   ```

7. **Boot each node to restart the cluster.**

   ```
   # reboot
   ```

8. **Verify that the change took place on all cluster nodes.**

   On each node, run the `date` command.

   ```
   # date
   ```

▼ SPARC: How to Enter the OpenBoot PROM (OBP) on a Node

Use this procedure if you need to configure or change OpenBoot PROM settings.

Steps

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`.**

   ```
   # scswitch -S -h node[,...]
   # shutdown -g0 -y -i0
   ```

Caution – Do not use `send brk` on a cluster console to shut down a cluster node.

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.

Steps

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 Data Services Planning and Administration Guide for Solaris OS* for information about using the `scswitch` command.

2. **Bring down the Network Time Protocol (NTP) daemon on each node of the cluster.**

   - If you are using Solaris 8 or Solaris 9, use the `xntpd` command to bring down the Network Time Protocol (NTP) daemon. See the `xntpd`(1M) man page for more information about the NTP daemon.

     ```
     # /etc/init.d/xntpd.cluster stop
     ```

   - If you are using Solaris 10, use the `svcadm` command to bring down the Network Time Protocol (NTP) daemon. See the `svcadm`(1M) man page for more information about the NTP daemon.

     ```
     # svcadm disable ntp
     ```

3. **Run the `scsetup`(1M) utility to change the private hostname of the appropriate node.**

   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.

   ---

4. **Select `5`, `Private Hostnames`, from the Main Menu.**

5. **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.

6. **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
   ```

7. **Edit the `ntp.conf.cluster` 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 sixteen nodes. Typically, the `ntp.conf.cluster` file will be identical on each cluster node.

8. **Verify that you can successfully ping the new private hostname from all cluster nodes.**

9. **Restart the NTP daemon.**

   Do this on each node of the cluster.

   - If you are using Solaris 8 or Solaris 9, use the `xntpd` command to restart the NTP daemon.

     ```
     # /etc/init.d/xntpd.cluster start
     ```

   - If you are using Solaris 10, use the `svcadm` command to restart the NTP daemon.

     ```
     # svcadm enable ntp
     ```

10. **Enable all Data Service resources and other applications that were disabled in Step 1.**

    ```
    # scswitch -e -j resource1, resource2
    ```

    See the `scswitch` man page and the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* for information about using the `scswitch` command.

Example 7–8 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`(1M) utility does not include the functionality for putting a quorum device into maintenance state.

Steps

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 node[,...]
   ```

`-S`	Evacuates all device services and resource groups from the specified node.
`-h` *node[,...]*	Specifies the node from which you are switching resource groups and devices groups.

3. **Shut down the node that you evacuated.**

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

4. **Become superuser on another node in the cluster and put the node that you shut 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.**

   # **scstat -q**

   The node you put into maintenance state should have a `Status` of `offline` and `0` (zero) for `Present` and `Possible` quorum votes.

Example 7–9 Putting a Cluster Node Into Maintenance State

The following example moves a cluster node into maintenance state and verifies the results. The `scstat -q` output shows the `Node votes` for phys-schost-1 to be `0` (zero) and the status to be `Offline`. The `Quorum Summary` should also show reduced vote counts. Depending on your configuration, the `Quorum Votes by Device` output might indicate that some quorum disk devices are offline as well.

```
[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# 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       0        0       Offline
  Node votes:       phys-schost-2       1        1       Online
  Node votes:       phys-schost-3       1        1       Online

-- Quorum Votes by Device --
                    Device Name         Present Possible Status
                    -----------         ------- -------- ------
  Device votes:     /dev/did/rdsk/d3s2  0        0       Offline
  Device votes:     /dev/did/rdsk/d17s2 0        0       Offline
```

```
Device votes:     /dev/did/rdsk/d31s2 1        1       Online
```

See Also To bring a node back online, see "How to Bring a Node Out of Maintenance State" on page 163.

▼ 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.

Steps

1. **Become superuser on any node of the cluster, other than the one in maintenance state.**

2. **Are you bringing a node out of maintenance state in a two-node cluster?**
   - If yes, go to Step 4.
   - If no, go to Step 3.

3. **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.

4. **Boot the node that you want to bring out of maintenance state.**

5. **Verify the quorum vote count.**

   ```
   # scstat -q
   ```

   The node you brought out of maintenance state should have a status of `online` and show the appropriate vote count for `Present` and `Possible` quorum votes.

Example 7–10 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 to their defaults and verifies the result. The `scstat -q` output shows the `Node votes` for phys-schost-1 to be `1` and the status to be `online`. The `Quorum Summary` should also show an increase in vote counts.

```
phys-schost-2# scconf -c -q node=phys-schost-1,reset
```

On phys-schost-1:

- SPARC:

  ```
  ok boot -x
  ```

- x86:

  ```
                         <<< Current Boot Parameters >>>
  Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
  sd@0,0:a
  Boot args:

  Type    b [file-name] [boot-flags] <ENTER>  to boot with options
  or      i <ENTER>                           to enter boot interpreter
  or      <ENTER>                             to boot with defaults

                     <<< timeout in 5 seconds >>>
  Select (b)oot or (i)nterpreter: b -x
  ```

```
phys-schost-1# scstat -q

-- 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
  Node votes:       phys-schost-3       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/d17s2 1        1       Online
  Device votes:     /dev/did/rdsk/d31s2 1        1       Online
```

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 7–2 Task Map: Adding a Cluster Node to an Existing Cluster

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 3.0-3.1 Hardware Administration Manual for Solaris OS*
Add shared storage	*Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*
Add the node to the authorized node list - Use `scsetup`.	"How to Add a Node to the Authorized Node List" on page 167
Install and configure the software on the new cluster node - Install the Solaris operating system and Sun Cluster software - Configure the node as part of the cluster	Chapter 2, "Installing and Configuring Sun Cluster Software," in *Sun Cluster Software Installation Guide for Solaris OS*

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.

Caution – Do not use this procedure if your cluster is running an OPS configuration. At this time, removing a node in an OPS configuration might cause nodes to panic at reboot.

TABLE 7–3 Task Map: Removing a Cluster Node (5/02)

Task	For Instructions, Go To
Move all resource groups and disk device groups off of the node to be removed. - Use `scswitch`(1M)	`# scswitch -S -h` *from-node*
Remove the node from all resource groups. - Use `scrgadm`(1M)	*Sun Cluster Data Services Planning and Administration Guide for Solaris OS*
Remove node from all disk device groups - Use `scconf`(1M), `metaset`(1M), and `scsetup`(1M)	"How to Remove a Node From a Disk Device Group (Solstice DiskSuite/Solaris Volume Manager)" on page 80 "SPARC: How to Remove a Node From a Disk Device Group (VERITAS Volume Manager)" on page 96 "SPARC: How to Remove a Node From a Raw Disk Device Group" on page 98 **Caution: If the number of desired secondaries is configured as 2 or more, it must be decreased to 1.**
Remove all fully connected quorum devices. - Use `scsetup`.	**Caution: Do not remove the quorum device if you are removing a node from a two-node cluster.** "How to Remove a Quorum Device" on page 126 Note that although you must remove the quorum device before you remove the storage device in the next step, you can add the quorum device back immediately afterward.
Remove all fully connected storage devices from the node. - Use `devfsadm`(1M), `scdidadm`(1M).	**Caution: Do not remove the quorum device if you are removing a node from a two-node cluster.** "How to Remove Connectivity Between an Array and a Single Node, in a Cluster With Greater Than Two-Node Connectivity" on page 170

TABLE 7–3 Task Map: Removing a Cluster Node (5/02) *(Continued)*

Task	For Instructions, Go To
Add back the quorum devices (to only the nodes that are intended to remain in the cluster). - Use `scconf -a -q globaldev=d[n],node=node1,node=node2`	`scconf`(1M)
Place the node being removed into maintenance state. - Use `scswitch`(1M), `shutdown`(1M), and `scconf`(1M).	"How to Put a Node Into Maintenance State" on page 161
Remove all logical transport connections (transport cables and adapters) to the node being removed. - Use `scsetup`.	"How to Remove Cluster Transport Cables, Transport Adapters, and Transport Junctions" on page 142
Remove all quorum devices connected to the node being removed. - Use `scsetup`, `scconf(1M)`.	"How to Remove the Last Quorum Device From a Cluster" on page 128
Remove node from the cluster software configuration. - Use `scconf(1M)`.	"How to Remove a Node From the Cluster Software Configuration" on page 169

▼ How to Add a 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 3.0-3.1 Hardware Administration Manual for Solaris OS* 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 to complete this procedure.

Steps

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 165.**

2. **Type the `scsetup` command.**

   ```
   # scsetup
   ```

   The Main Menu is displayed.

3. **To access the New Nodes Menu, type `7` at the Main Menu.**

4. **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.

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. **Quit the `scsetup` utility.**

7. **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 Software Installation Guide for Solaris OS*.

8. **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.

Example 7–11 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.]
```

See Also For an overall list of tasks for adding a cluster node, see Table 7–2, "Task Map: Adding a Cluster Node."

To add a node to an existing resource group, see the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

▼ How to Remove a Node From the Cluster Software Configuration

Perform this procedure to remove a node from the cluster.

Steps

1. **Be sure you have correctly completed all prerequisite tasks listed in the "Removing a Cluster Node" task map in "Adding and Removing a Cluster Node" on page 165.**

   Note – Be sure you have removed the node from all resource groups, disk device groups, and quorum device configurations and placed it in maintenance state before you continue with this procedure.

2. **Become superuser on a node in the cluster other than the node to remove.**
3. **Remove the node from the cluster.**

   ```
   # scconf -r -h node=node-name
   ```

4. **Verify the node removal by using `scstat`(1M).**

   ```
   # scstat -n
   ```

5. **Do you intend to uninstall Sun Cluster software from the removed node?**
   - If yes, go to "How to Uninstall Sun Cluster Software From a Cluster Node" on page 173. Or instead, you can reinstall Solaris software on the node.
   - If no, to physically remove the node from the cluster, remove the hardware connections as described in the *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*.

Example 7–12 Removing a Node From the Cluster Software Configuration

This example shows how to remove a node (`phys-schost-2`) from a cluster. All commands are run from another node of the cluster (`phys-schost-1`).

```
[Remove the node from the cluster:]
phys-schost-1# scconf -r -h node=phys-schost-2
[Verify node removal:]
phys-schost-1# scstat -n
-- Cluster Nodes --
                    Node name           Status
```

```
                    ---------           ------
Cluster node:       phys-schost-1       Online
```

See Also To uninstall Sun Cluster software from the removed node, see "How to Uninstall Sun Cluster Software From a Cluster Node" on page 173

For hardware procedures, see the *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*.

For an overall list of tasks for removing a cluster node, see Table 7–3.

To add a node to an existing cluster, see "How to Add a Node to the Authorized Node List" on page 167.

▼ How to Remove Connectivity Between an Array and a Single Node, in a Cluster With Greater Than Two-Node Connectivity

Use this procedure to detach a storage array from a single cluster node, in a cluster that has three- or four-node connectivity.

Steps

1. **Back up all database tables, data services, and volumes that are associated with the storage array that you are removing.**

2. **Determine the resource groups and device groups that are running on the node to be disconnected.**

   ```
   # scstat
   ```

3. **If necessary, move all resource groups and device groups off the node to be disconnected.**

Caution (SPARC only) – If your cluster is running Oracle Parallel Server/Real Application Clusters software, shut down the Oracle Parallel Server/Real Application Clusters database instance that is running on the node before you move the groups off the node. For instructions see the *Oracle Database Administration Guide*.

```
# scswitch -S -h from-node
```

4. **Put the device groups into maintenance state.**

   For the procedure on acquiescing I/O activity to Veritas shared disk groups, see your VxVM documentation.

   For the procedure on putting a device group in maintenance state, see the Chapter 7.

5. **Remove the node from the device groups.**
   - If you use VxVM or raw disk, use the `scconf`(1M) command to remove the device groups.
   - If you use Solstice DiskSuite, use the `metaset` command to remove the device groups.
6. **If the cluster is running HAStorage or HAStoragePlus, remove the node from the resource group's nodelist.**

   `# scrgadm -a -g` *resource-group* `-h` *nodelist*

   See the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* for more information on changing a resource group's nodelist.

   Note – Resource type, resource group, and resource property names are case insensitive when executing `scrgadm`.

   ---

7. **If the storage array you are removing is the last storage array that is connected to the node, disconnect the fiber-optic cable between the node and the hub or switch that is connected to this storage array (otherwise, skip this step).**
8. **Do you want to remove the host adapter from the node you are disconnecting?**
   - If yes, shut down and power off the node.
   - If no, skip to Step 11.
9. **Remove the host adapter from the node.**

   For the procedure on removing host adapters, see the documentation that shipped with your node.
10. **Without allowing the node to boot, power on the node.**
11. **Boot the node into non-cluster mode.**
    - SPARC:

      ```
      ok boot -x
      ```

    - x86:

      ```
                            <<< Current Boot Parameters >>>
      Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
      sd@0,0:a
      Boot args:

      Type    b [file-name] [boot-flags] <ENTER>  to boot with options
      or      i <ENTER>                           to enter boot interpreter
      or      <ENTER>                             to boot with defaults

                        <<< timeout in 5 seconds >>>
      Select (b)oot or (i)nterpreter: b -x
      ```

Caution (SPARC only) – The node must be in non-cluster mode before you remove Oracle Parallel Server/Real Application Clusters software in the next step or the node panics and potentially causes a loss of data availability.

12. SPARC: If Oracle Parallel Server/Real Application Clusters software has been installed, remove the Oracle Parallel Server/Real Application Clusters software package from the node that you are disconnecting.

```
# pkgrm SUNWscucm
```

Caution (SPARC only) – If you do not remove the Oracle Parallel Server/Real Application Clusters software from the node you disconnected, the node will panic when the node is reintroduced to the cluster and potentially cause a loss of data availability.

13. Boot the node into cluster mode.

- SPARC:

```
ok boot
```

- x86:

```
                      <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>  to boot with options
or      i <ENTER>                           to enter boot interpreter
or      <ENTER>                             to boot with defaults

                  <<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b
```

14. On the node, update the device namespace by updating the /devices and /dev entries.

```
# devfsadm -C
# scdidadm -C
```

15. Bring the device groups back online.

For procedures about bringing a VERITAS shared disk group online, see your VERITAS Volume Manager documentation.

For the procedure on bringing a device group online, see the procedure on putting a device group into maintenance state.

▼ How to Uninstall Sun Cluster Software From a Cluster Node

Perform this procedure to uninstall Sun Cluster software from a cluster node before you disconnect it from a fully established cluster configuration. You can use this procedure to uninstall software from the last remaining node of a cluster.

Note – To uninstall Sun Cluster software from a node that has not yet joined the cluster or is still in install mode, do not perform this procedure. Instead, go to "How to Uninstall Sun Cluster Software to Correct Installation Problems" in the *Sun Cluster Software Installation Guide for Solaris OS*.

Steps

1. **Be sure you have correctly completed all prerequisite tasks listed in the task map for removing a cluster node.**

   See "Adding and Removing a Cluster Node" on page 165.

   Note – Be sure you have removed the node from all resource groups, device groups, and quorum device configurations, placed it in maintenance state, and removed it from the cluster before you continue with this procedure.

   ---

2. **Become superuser on an active cluster member other than the node you will uninstall.**

3. **From the active cluster member, add the node you intend to uninstall to the cluster's node authentication list.**

   ```
   # scconf -a -T node=nodename
   ```

`-a`	Add
`-T`	Specifies authentication options
`node=`*nodename*	Specifies the name of the node to add to the authentication list

   Alternately, you can use the `scsetup`(1M) utility. See "How to Add a Node to the Authorized Node List" on page 167 for procedures.

4. **Become superuser on the node to uninstall.**

5. **Reboot the node into non-cluster mode.**

   - SPARC:

     ```
     # shutdown -g0 -y -i0ok boot -x
     ```

   - x86:

```
# shutdown -g0 -y -i0
...
                      <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>  to boot with options
or      i <ENTER>                           to enter boot interpreter
or      <ENTER>                             to boot with defaults

                  <<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b -x
```

6. **In the `/etc/vfstab` file, remove all globally mounted file system entries except the `/global/.devices` global mounts.**

7. **Uninstall Sun Cluster software from the node.**

   Run the command from a directory that is not associated with any Sun Cluster packages.

   ```
   # cd /
   # scinstall -r
   ```

   See the `scinstall`(1M) man page for more information. If `scinstall` returns error messages, see "Unremoved Cluster File System Entries" on page 175.

8. **Disconnect the transport cables and the transport junction, if any, from the other cluster devices.**

   a. **If the uninstalled node is connected to a storage device that uses a parallel SCSI interface, install a SCSI terminator to the open SCSI connector of the storage device after you disconnect the transport cables.**

      If the uninstalled node is connected to a storage device that uses Fibre Channel interfaces, no termination is necessary.

   b. **Follow the documentation that shipped with your host adapter and server for disconnection procedures.**

▼ How to Correct Error Messages

To correct the error messages in the previous sections, perform this procedure.

Steps

1. **Attempt to rejoin the node to the cluster.**

   ```
   # boot
   ```

2. **Did the node successfully rejoin the cluster?**

- If no, proceed to Step 3.
- If yes, perform the following steps to remove the node from disk device groups.

a. If the node successfully rejoins the cluster, remove the node from the remaining disk device group(s).

Follow procedures in "How to Remove a Node From All Disk Device Groups" on page 79.

b. After you remove the node from all disk device groups, return to "How to Uninstall Sun Cluster Software From a Cluster Node" on page 173 and repeat the procedure.

3. If the node could not rejoin the cluster, rename the node's `/etc/cluster/ccr` file to any other name you choose, for example, `ccr.old`.

```
# mv /etc/cluster/ccr /etc/cluster/ccr.old
```

4. Return to "How to Uninstall Sun Cluster Software From a Cluster Node" on page 173 and repeat the procedure.

Troubleshooting a Node Uninstallation

This section describes error messages you might receive when you run the `scinstall -r` command and the corrective actions to take.

Unremoved Cluster File System Entries

The following error messages indicate that the node you removed still has cluster file systems referenced in its `vfstab` file.

```
Verifying that no unexpected global mounts remain in /etc/vfstab ... failed
scinstall:  global-mount1 is still configured as a global mount.
scinstall:  global-mount1 is still configured as a global mount.
scinstall:  /global/dg1 is still configured as a global mount.

scinstall:  It is not safe to uninstall with these outstanding errors.
scinstall:  Refer to the documentation for complete uninstall instructions.
scinstall:  Uninstall failed.
```

To correct this error, return to "How to Uninstall Sun Cluster Software From a Cluster Node" on page 173 and repeat the procedure. Ensure that you successfully complete Step 6 in the procedure before you rerun the `scinstall -r` command.

Unremoved Listing in Disk Device Groups

The following error messages indicate that the node you removed is still listed with a disk device group.

```
Verifying that no device services still reference this node ... failed
scinstall:  This node is still configured to host device service "
service".
scinstall:  This node is still configured to host device service "
service2".
scinstall:  This node is still configured to host device service "
service3".
scinstall:  This node is still configured to host device service "
dg1".

scinstall:  It is not safe to uninstall with these outstanding errors.
scinstall:  Refer to the documentation for complete uninstall instructions.
scinstall:  Uninstall failed.
```

CHAPTER 8

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. Occasionally, when patching a node with a Sun Cluster 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 3.1 8/05 Release Notes for Solaris OS* or contact Enterprise Services. If there aren't any special instructions, check the patch's `README` file.

Note – For Sun Cluster patches, always defer to the patch's `README` file and to SunSolve 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` or `b -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` or `b -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 3.1 8/05 Release Notes for Solaris OS* 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 8–1 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 185
Apply a rebooting Sun Cluster patch after taking the cluster member to non-cluster mode	"How to Apply a Rebooting Patch (Node)" on page 179 "How to Apply a Rebooting Patch (Cluster and Firmware)" on page 182
Remove a Sun Cluster patch - You can back out the patch if necessary	"How to Remove a Sun Cluster Patch" on page 185

▼ 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` or `b -sx`command, before applying the patch.

Steps

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 node[,...]
   ```

`-S`	Evacuates all device groups and resource groups from the specified node.
`-h` *node[,...]*	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.**

   - SPARC:

```
ok boot -sx
```

   - x86:

```
                      <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>  to boot with options
or      i <ENTER>                           to enter boot interpreter
or      <ENTER>                             to boot with defaults

                  <<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b -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. **Verify that the patch has been installed successfully.**

```
# showrev -p | grep patch-id
```

9. **Reboot the node into the cluster.**

```
# reboot
```

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 -D  device-group[,...]
    -h node[,...]
    # scswitch -z -g  resource-group[,...]
    -h nod[,...]e
    ```

`-z`	Specifies the change in mastery of a resource group or device group.
`-h` *node[,...]*	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.

13. **Check to see if you need to commit the patch software by using the `scversions` command.**

    ```
    # /usr/cluster/bin/scversions
    ```

    You will see one of the following results:

    ```
    Upgrade commit is needed.

    Upgrade commit is NOT needed. All versions match.
    ```

14. **If a commit is needed, commit the patch software.**

    ```
    #/usr/cluster/bin/scversions -c
    ```

    The `-c` options causes cluster to commit to and run the new patched software.

Note – Running `scversions(1m)` will cause one or more CMM reconfigurations, depending on the situation.

Example 8–1 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
...
```

Boot the node in non-cluster, single-user mode.

- SPARC:

```
ok boot -sx
```

- x86:

```
                      <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>  to boot with options
or      i <ENTER>                           to enter boot interpreter
or      <ENTER>                             to boot with defaults

                  <<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b -sx
```

```
# patchadd -M /var/tmp/patches 234567-05
...
# showrev -p | grep 234567-05

...
# reboot
...
# scswitch -z -D dg-schost-1 -h phys-schost-1
# scswitch -z -g schost-sa-1 -h phys-schost-1
# scversions
Upgrade commit is needed.
# scversions -c
```

See Also If you need to back out a patch, see "How to Remove a Sun Cluster Patch" on page 185.

▼ 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` or `b -sx` command, before applying the patch.

Steps

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.

   - SPARC:

     ok **boot -sx**

   - x86:

```
                      <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>  to boot with options
or      i <ENTER>                           to enter boot interpreter
or      <ENTER>                             to boot with defaults

                  <<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b -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 8–2 Applying a Rebooting Patch (Cluster)

The following example shows the application of a rebooting Sun Cluster patch to a cluster.

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

Boot the cluster in non-cluster, single-user mode.

- SPARC:

  ```
  ok boot -sx
  ```

- x86:

  ```
                         <<< Current Boot Parameters >>>
  Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
  sd@0,0:a
  Boot args:

  Type    b [file-name] [boot-flags] <ENTER>  to boot with options
  or      i <ENTER>                           to enter boot interpreter
  or      <ENTER>                             to boot with defaults

                    <<< timeout in 5 seconds >>>
  Select (b)oot or (i)nterpreter: b -sx
  ```

```
...
# patchadd -M /var/tmp/patches 234567-05
(Apply patch to other cluster nodes)
...
# showrev -p | grep 234567-05
# reboot
```

See Also If you need to back out a patch, see "How to Remove a Sun Cluster Patch" on page 185.

▼ 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.

Steps

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 8–3 Applying a Non-Rebooting Sun Cluster Patch

```
# patchadd -M /tmp/patches 234567-05
...
# showrev -p | grep 234567-05
```

See Also If you need to back out a patch, see "How to Remove a Sun Cluster Patch" on page 185.

▼ How to Remove a Sun Cluster Patch

If necessary, you can back out (remove) a Sun Cluster patch.

Steps

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** *node[,...]*

`-S`	Evacuates all device services and resource groups from the specified node.
`-h` *node[,...]*	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. Using this option brings down a node to the OpenBoot PROM `ok` prompt on a SPARC based system or to the Boot Subsystems on an x86 based system.
message	Specifies the warning message to broadcast. Use quotes if *message* contains multiple words.

5. **Boot the node in non-cluster, single-user mode.**

   - SPARC:

     ```
     ok boot -sx
     ```

   - x86:

     ```
                          <<< Current Boot Parameters >>>
     Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
     sd@0,0:a
     Boot args:

     Type    b [file-name] [boot-flags] <ENTER>  to boot with options
     or      i <ENTER>                           to enter boot interpreter
     or      <ENTER>                             to boot with defaults

                      <<< timeout in 5 seconds >>>
     Select (b)oot or (i)nterpreter: b -sx
     ```

6. **Remove the patch.**

   # **patchrm** *patch-id*

patch-id	Specifies the patch number of a given patch.

7. **Verify that the patch has been removed successfully.**

   # **showrev -p** | **grep** *patch-id*

8. **Reboot the node.**

   ```
   # reboot
   ```

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 -D device-group[,...]  -h node
    # scswitch -z -g resource-group[,...] -h node
    ```

`-z`	Specifies the change in mastery of a resource group or device group.
`-h` *node[,...]*	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 8–4 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"
...
```

Boot the node in non-cluster mode.

- SPARC:

  ```
  ok boot -x
  ```

- x86:

```
                        <<< Current Boot Parameters >>>
   Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
   sd@0,0:a
   Boot args:

   Type    b [file-name] [boot-flags] <ENTER>  to boot with options
   or      i <ENTER>                           to enter boot interpreter
   or      <ENTER>                             to boot with defaults

                    <<< timeout in 5 seconds >>>
   Select (b)oot or (i)nterpreter: b -x

...
# patchrm 234567-05
...
# showrev -p | grep 234567-05
...
# reboot
...
# scswitch -z -D dg-schost-1 -h phys-schost-1
# scswitch -z -g schost-sa-1 -h phys-schost-1
```

CHAPTER 9

Backing Up and Restoring a Cluster

This is a list of step-by-step instructions in this chapter.

Backing Up a Cluster

TABLE 9–1 Task Map: Backing Up Cluster Files

Task	For Instructions, Go To…
Find the names of the file systems you want to back up	"How to Find File System Names to Back Up" on page 190

TABLE 9–1 Task Map: Backing Up Cluster Files *(Continued)*

Task	For Instructions, Go To…
Calculate how many tapes you will need to contain a full backup	"How to Determine the Number of Tapes Needed for a Full Backup" on page 191
Back up the root file system	"How to Back Up the root (/) File System" on page 191
Perform online backup for mirrored or plexed file systems	"How to Perform Online Backups for Mirrors (Solstice DiskSuite/Solaris Volume Manager)" on page 193 "SPARC: How to Perform Online Backups for Volumes (VERITAS Volume Manager)" on page 196

▼ 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.

Steps

1. **Display the contents of the `/etc/vfstab` file.**

   You do not need to be superuser or assume an equivalent role 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 9–1 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.

Steps

1. **Become superuser or assume an equivalent role 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 9–2 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.

Steps

1. **Become superuser or assume an equivalent role 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** *node[,...]*

`-z`	Performs the switch.
`-D` *disk-device-group*	Name of the disk device group to be switched.
`-h` *node*	Name of the cluster node to switch the disk device group to. This node becomes the new primary.

3. **Shut down the node.**

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

4. **Reboot the node in non-cluster mode.**

   - SPARC:

```
ok boot -x
```

   - x86:

```
                      <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type  b [file-name] [boot-flags] <ENTER> to boot with options
or    i <ENTER>                         to enter boot interpreter
or    <ENTER>                           to boot with defaults

                  <<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b -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 9–3 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/Solaris Volume Manager)

A mirrored Solstice DiskSuite metadevice or Solaris Volume Manager volume 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.

Steps

1. **Become superuser or assume an equivalent role 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 disk set name.

3. **Use the `lockfs`(1M) command with the `-w` option to lock the file system from writes.**

   # **`lockfs -w`** *mountpoint*

   Note – You must lock the file system only if a UFS file system resides on the mirror. For example, if the Solstice DiskSuite metadevice or Solaris Volume Manager volume 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`** *mountpoint*

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 the backup utility that you usually 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 or volume back online.**

   # **`metattach -s`** *setname mirror submirror*

   When the metadevice or volume 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 9–4 Performing Online Backups for Mirrors (Solstice DiskSuite/Solaris Volume Manager)

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)
...
```

▼ SPARC: 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.

Steps

1. **Log on to any node in the cluster, and become superuser or assume an equivalent role 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` 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`(1M) command.**

   # **scswitch -z -g** *resource-group[,...]* **-h** *node[,...]*

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`.**

    # **vxedit -rf rm** *bkup-vol*

14. **Register the disk group configuration changes using the `scconf`(1M) command.**

    # **scconf -c -D name=***diskgroup***,sync**

Example 9–5 SPARC: Performing Online Backups for Volumes (VERITAS Volume Manager)

In the following example, the cluster node `phys-schost-2` is the primary owner of the disk device group `schost-1`, therefore the backup procedure is performed from `phys-schost-2`. The volume `/vol01` is copied and then associated with a new volume, `bkup-vol`.

```
[Become superuser or assume an equivalent role on the primary node.]
[Identify the current primary node for the disk device 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 device 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`(1M) command copies files to disk, relative to the current working directory, from backups created using the `ufsdump`(1M) 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 or assume an equivalent role, 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 9–2 Task Map: Restoring Cluster Files

Task	For Instructions, Go To…
For Solstice DiskSuite/Solaris Volume Manager, restore files interactively following Solaris restore procedures	"How to Restore Individual Files Interactively (Solstice DiskSuite/Solaris Volume Manager)" on page 201
For Solstice DiskSuite/Solaris Volume Manager, restore the root (/) file system	"How to Restore the root (/) File System (Solstice DiskSuite/Solaris Volume Manager)" on page 201
	"How to Restore a root (/) File System That Was on a Solstice DiskSuite Metadevice or Solaris Volume Manager Volume" on page 203
For VERITAS Volume Manager, restore a non-encapsulated root (/) file system	"SPARC: How to Restore a Non-Encapsulated root (/) File System (VERITAS Volume Manager)" on page 208

TABLE 9–2 Task Map: Restoring Cluster Files *(Continued)*

Task	For Instructions, Go To…
For VERITAS Volume Manager, restore an encapsulated root (/) file system	"SPARC: How to Restore an Encapsulated root (/) File System (VERITAS Volume Manager)" on page 210

▼ How to Restore Individual Files Interactively (Solstice DiskSuite/Solaris Volume Manager)

Use this procedure to restore one or more individual files. Be sure the cluster is running problem-free before performing the restore procedure.

Steps

1. **Become superuser or assume an equivalent role 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/Solaris Volume Manager)

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 – Because 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.

Steps

1. **Become superuser or assume an equivalent role on a cluster node with access to the disksets to which the node to be restored is also attached.**

   Become superuser or assume an equivalent role on a node *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 disk set name.
`-f`	Force.
`-d`	Deletes from the disk set.
`-h` *nodelist*	Specifies the name of the node to delete from the disk set.

3. **Restore the `root` (`/`) and `/usr` file systems.**

   To restore the `root` and `/usr` file systems, following the procedure in Chapter 27, "Restoring Files and File Systems (Tasks)," in *System Administration Guide: Devices and File Systems*. Omit the step in the Solaris procedure to reboot the system.

Note – Be sure to create the `/global/.devices/node@nodeid` file system.

---

4. **Reboot the node in multiuser mode.**

   ```
   # reboot
   ```

5. **Replace the disk ID using the `scdidadm`(1M) command.**

   ```
   # scdidadm -R rootdisk
   ```

6. **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.

7. **From a cluster node other than the restored node, use the `metaset` command to add the restored node to all disksets.**

   ```
   phys-schost-2# metaset -s setname -a -h nodelist
   ```

`-a`	Creates and adds the host to the disk set.

   The node is rebooted into cluster mode. The cluster is ready to use.

Example 9–6 Restoring the root (/) File System (Solstice DiskSuite/Solaris Volume Manager)

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 disk set `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 or assume an equivalent role 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:]
Restore the root (/) and /usr file system using the procedure in the Solaris system administration documentation
 [Reboot:]
# reboot
[Replace the disk ID:]
# scdidadm -R /dev/dsk/c0t0d0
[Recreate state database replicas:]
# metadb -c 3 -af /dev/rdsk/c0t0d0s4
[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 Solstice DiskSuite Metadevice or Solaris Volume Manager Volume

Use this procedure to restore a root (/) file system that was on a Solstice DiskSuite metadevice or Solaris Volume Manager volume 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.

Steps

1. **Become superuser or assume an equivalent role on a cluster node with access to the diskset,** *other than* **the node you want to restore.**

2. **Remove from all disksets 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 that you want to restore.**
   - If you are using the Solaris CD:
     - SPARC: At the OpenBoot PROM `ok` prompt, type the following command:

       ```
       ok boot cdrom -s
       ```

     - x86: Insert the CD into the system's CD drive and boot the system by shutting it down and then turning it off and on. On the Current Boot Parameters screen, type the following command:

       ```
                            <<< Current Boot Parameters >>>
       Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@
       7,1/sd@0,0:a
       Boot args:

       Type b [file-name] [boot-flags] <ENTER> to boot with options
       or   i <ENTER>                          to enter boot interpreter
       or   <ENTER>                            to boot with defaults

                        <<< timeout in 5 seconds >>>
       Select (b)oot or (i)nterpreter: b -s
       ```

   - If you are using a Solaris JumpStart™ server:
     - SPARC: At the OpenBoot PROM `ok` prompt, type the following command:

       ```
       ok boot net -s
       ```

     - x86: Boot the system by shutting it down and then turning it off and on. On the Current Boot Parameters screen, type the following command:

       ```
                            <<< Current Boot Parameters >>>
       Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@
       7,1/sd@0,0:a
       Boot args:

       Type b [file-name] [boot-flags] <ENTER> to boot with options
       or   i <ENTER>                          to enter boot interpreter
       or   <ENTER>                            to boot with defaults

                        <<< timeout in 5 seconds >>>
       Select (b)oot or (i)nterpreter: b -s
       ```

5. **Create all the partitions and swap on the root disk using the `format` 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` 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-mountpoint*

8. **Use the following commands to restore the root (/) file system.**

   # **cd** *temp-mountpoint*
   # **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-mountpoint***`/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-mountpoint/etc/vfstab` file to change the root entry from a Solstice DiskSuite metadevice or a Solaris Volume Manager volume to a corresponding normal slice for each file system on the root disk that is part of the metadevice or volume.**

```
Example:
Change from—
/dev/md/dsk/d10   /dev/md/rdsk/d10    /      ufs   1     no       -

Change to—
/dev/dsk/c0t0d0s0 /dev/rdsk/c0t0d0s0  /      ufs   1     no       -
```

12. **Unmount the temporary file system, and check the raw disk device.**

    # **cd /**
    # **umount** *temp-mountpoint*
    # **fsck** *raw-disk-device*

13. Reboot the node in multiuser mode.

```
# reboot
```

14. Replace the disk ID using the `scdidadm` command.

```
# scdidadm -R rootdisk
```

15. Use the `metadb` 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. From a cluster node other than the restored node, use the `metaset` command to add the restored node to all disksets.

```
phys-schost-2# metaset -s setname -a -h nodelist
```

`-a`	Adds (creates) the metaset.

Set up the metadevice or volume/mirror for root (`/`) according to the Solstice DiskSuite documentation.

The node is rebooted into cluster mode. The cluster is ready to use.

Example 9–7 Restoring a root (/) File System That Was on a Solstice DiskSuite Metadevice or Solaris Volume Manager Volume

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 or assume an equivalent role 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:]
```

Boot the node from the Solaris CD:

- SPARC: At the OpenBoot PROM `ok` prompt, type the following command:

  ```
  ok boot cdrom -s
  ```

- x86: Insert the CD into the system's CD drive and boot the system by shutting it down and then turning it off and on. On the Current Boot Parameters screen, type the following command:

```
                        <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type  b [file-name] [boot-flags] <ENTER> to boot with options
or    i <ENTER>                          to enter boot interpreter
or    <ENTER>                            to boot with defaults

                    <<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b -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-mountpoint/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-mountpoint/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:]

```
# reboot
```

```
[Replace the disk ID:]
# scdidadm -R /dev/rdsk/c0t0d0
[Recreate state database replicas:]
# metadb -c 3 -af /dev/rdsk/c0t0d0s4
[Add the node back to the metaset:]
phys-schost-2# metaset -s schost-1 -a -h phys-schost-1
```

▼ SPARC: 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.

Steps

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 that you want to restore.**

   - If you are using the Solaris CD, at the OpenBoot PROM `ok` prompt, type the following command:

     `ok` **`boot cdrom -s`**

   - If you are using a Solaris JumpStart™ server, at the OpenBoot PROM `ok` prompt, type the following command:

     `ok` **`boot net -s`**

3. **Create all the partitions and swap on the root disk using the `format` 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` 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-mountpoint*

6. **Restore the root (/) file system from backup, and unmount and check the file system.**

   # **cd** *temp-mountpoint*
   # **ufsrestore rvf** *dump-device*
   # **rm restoresymtable**
   # **cd /**
   # **umount** *temp-mountpoint*
   # **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 in multiuser mode.**

   # **reboot**

9. **Update the disk ID using the `scdidadm` command.**

   # **scdidadm -R /dev/rdsk/***disk-device*

10. **Press Control-d to resume in multiuser mode.**

    The node reboots into cluster mode. The cluster is ready to use.

Example 9–8 SPARC: 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*:]

Boot the node from the Solaris CD. At the OpenBoot PROM `ok` prompt, type the following command:

```
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:]
# reboot
[Update the disk ID:]
# scdidadm -R /dev/rdsk/c0t0d0
```

▼ SPARC: 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.

Steps

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 that you want to restore.**

   - If you are using the Solaris CD, at the OpenBoot PROM `ok` prompt, type the following command:

     ```
     ok boot cdrom -s
     ```

   - If you are using a Solaris JumpStart™ server, at the OpenBoot PROM `ok` prompt, type the following command:

     ```
     ok boot net -s
     ```

3. **Create all the partitions and swap on the root disk using the `format` 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` 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-mountpoint*

6. **Restore the root (/) file system from backup.**

   # **cd** *temp-mountpoint*
   # **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-mountpoint*/**etc/vx/reconfig.d/state.d/install-db**

8. **Remove the following entries from the /** *temp-mountpoint*/**etc/system file.**

   ```
   * rootdev:/pseudo/vxio@0:0
   * set vxio:vol_rootdev_is_volume=1
   ```

9. **Edit the /***temp-mountpoint* **/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/vx/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-mountpoint*
    # **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 multiuser mode.**

    # **reboot**

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 9–9 SPARC: 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:]

Boot the node from the Solaris CD. At the OpenBoot PROM ok prompt, type the
following command:

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/vx/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:]
# reboot
[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
```

See Also For instructions about how to mirror the encapsulated root disk, see the *Sun Cluster Software Installation Guide for Solaris OS*.

CHAPTER 10

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 many 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.

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. SunPlex Manager also allows you to perform many administrative tasks for the following Sun Cluster components. However, SunPlex Manager currently cannot perform all Sun Cluster administrative tasks. You must use the command-line interface for some operations.

- Adapters
- Cables
- Data Services
- Global Devices

- Interconnects
- Junctions
- Nodes
- Quorum Devices
- Resource Groups
- Resources

SunPlex Installer, an installation module of SunPlex Manager, can be used to install certain Sun Cluster data services. You can SunPlex Installer once you've launched SunPlex Manager. SunPlex Installer is located at the following port.

`https://`*node*`:6789/`

Information about installing and using SunPlex Manager can be found in the following locations.

- **Installing and starting SunPlex Manager:** See the *Sun Cluster Software Installation Guide for Solaris OS*.
- **Configuring port numbers, server addresses, security certificates, and users:** See "Configuring SunPlex Manager" on page 217.
- **Installing and administering aspects of your cluster using SunPlex Manager:** See the online help supplied with SunPlex Manager.
- **Regenerating SunPlex Manager security keys:** See "How to Regenerate Common Agent Container Security Keys" on page 220.

SPARC: Sun Management Center Overview

The Sun Cluster 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 module for Sun Management Center cannot perform Sun Cluster configuration tasks. You must use the command-line interface for configuration 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 Software Installation Guide for Solaris OS*.

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 at `/opt/SUNWsymon/modules/cfg/sun-cluster-mib.mib` 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 "SNMP MIBs for Sun Management Center Modules" in *Sun Management Center 3.5 Update 2 User's Guide* in "SNMP MIBs for Sun Management Center Modules" in *Sun Management Center 3.5 Update 2 User's Guide*.

Configuring SunPlex Manager

SunPlex Manager is a GUI that you can use to administer and view the status all aspects of quorum devices, IPMP 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 Software Installation Guide for Solaris OS*. 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.

Setting up RBAC Roles

The SunPlex Manager uses RBAC to determine who has rights to administer the cluster. Several RBAC rights profiles are included in the Sun Cluster software. You can assign these rights profiles to users or to roles to give users different levels of access to Sun Cluster. For more information about how to set up and manage RBAC for Sun Cluster, see Sun Cluster and RBAC in the *Sun Cluster Systems Administration Guide*.

▼ How to Use the Common Agent Container to Change the Port Numbers for Services or Management Agents

If the default port numbers for your common agent container services conflict with other running processes, you can use the `cacaoadm` command to change the port number of the conflicting service or management agent on each node of the cluster.

Steps

1. **On all cluster nodes, stop the common agent container management daemon.**

   ```
   # /opt/SUNWcacao/bin/cacaoadm stop
   ```

2. **Stop Sun Java Web Console.**

   ```
   # /usr/sbin/sunmcwebserver stop
   ```

3. **If you do not know the port number currently used by the common agent container service for which you want to change the port number, use the `cacaoadm` command with the `get-param` subcommand to retrieve the port number.**

   ```
   # /opt/SUNWcacao/bin/cacaoadm get-param parameterName
   ```

   You can use the `cacaoadm` command to change the port numbers for the following common agent container services. The following list provides some examples of services and agents that can be managed by the common agent container, along with corresponding parameter names.

JMX connector port	`jmxmp-connector-port`
SNMP port	`snmp-adaptor-port`
SNMP trap port	`snmp-adaptor-trap-port`
Command stream port	`commandstream-adaptor-port`

4. **To change a port number, use the `cacaoadm` command with the `setparam` subcommand and the parameter name.**

   ```
   # /opt/SUNWcacao/bin/cacaoadm set-param parameterName=parameterValue
   =parameterValue
   ```

5. **Repeat Step 4 on each node of the cluster.**

6. **Restart Sun Java Web Console.**

   ```
   # /usr/sbin/sunmcwebserver start
   ```

7. **Restart the common agent container management daemon on all cluster nodes.**

   ```
   # /opt/SUNWcacao/bin/cacaoadm start
   ```

▼ 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. The default security certificate is generated based on the node's hostname at the time SunPlex Manager is installed. To reset the node's hostname, delete the certificate file, `keystore` and restart SunPlex Manager. SunPlex Manager will automatically create a new certificate file with the new hostname. You must complete this procedure on any node that has had its hostname changed.

Steps

1. **Remove the certificate file, `keystore`, located in `/etc/opt/webconsole`.**

   ```
   # cd /etc/opt/webconsole
   # pkgrm keystore
   ```

2. **Restart SunPlex Manager.**

   ```
   # /usr/sbin/smcwebserver restart
   ```

▼ 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 booting. 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.

Steps

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.**

```
# /usr/sbin/smcwebserver restart
```

7. **Repeat this procedure for each node in the cluster.**

Example 10–1 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.]
# /usr/sbin/smcwebserver restart
```

▼ How to Regenerate Common Agent Container Security Keys

SunPlex Manager uses strong encryption techniques to ensure secure communication between the SunPlex Manager web server and each cluster node.

The keys used by the SunPlex Manager are stored under the `/etc/opt/SUNWcacao/security` directory on each node. They should be identical across all cluster nodes.

Under normal operation, these keys can be left in their default configuration. If you change the hostname of a cluster node, you must regenerate the common agent container security keys. You may also need to regenerate the keys due to a possible key compromise (for example, root compromise on the machine). To regenerate the security keys, using the following procedure.

Steps 1. **On all cluster nodes, stop the common agent container management daemon.**

```
# /opt/SUNWcacao/bin/cacaoadm stop
```

2. **On one node of the cluster, regenerate the security keys.**

```
phys-schost-1# /opt/SUNWcacao/bin/cacaoadm create-keys --force
```

3. **Restart the common agent container management daemon on the node on which you regenerated the security keys.**

```
phys-schost-1# /opt/SUNWcacao/bin/cacaoadm start
```

4. **Create a tarfile of the `/etc/opt/SUNWcacao/security` directory.**

```
phys-schost-1# tar cf /tmp/SECURITY.tar security
```

5. **Copy the `/tmp/Security.tar` file to each of the cluster nodes.**

6. **On each node to which you copied the `/tmp/SECURITY.tar` file, extract the security files.**

   Any security files that already exist in the `/etc/opt/SUNWcacao/` directory are overwritten.

```
phys-schost-2# cd /etc/opt/SUNWcacao
phys-schost-2# tar xf /tmp/SECURITY.tar
```

7. **Delete the `/tmp/SECURITY.tar` file from each node in the cluster.**

   You must delete each copy of the tarfile to avoid security risks.

```
phys-schost-1# rm /tmp/SECURITY.tar
phys-schost-2# rm /tmp/SECURITY.tar
```

8. **On all nodes, restart the common agent container management daemon.**

```
phys-schost-1# /opt/SUNWcacao/bin/cacaoadm start
```

9. **Restart SunPlex Manager.**

```
# /usr/sbin/smcwebserver 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.

Steps

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 use the `useradd`(1M) command to add a user account to the system. You must set up at least one user account to access SunPlex Manager if you do not use the `root` system account. SunPlex Manager user accounts are used only by SunPlex Manager. They do not correspond to any Solaris system user accounts. Creating and assigning an RBAC role to a user account is described in more detail in "Creating and Assigning an RBAC Role With a Sun Cluster Management Rights Profile" on page 34.

   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. **Ensure that the browser's disk and memory cache sizes are set to a value that is greater than 0.**
7. **From the browser, connect to the SunPlex Manager port on one node of the cluster.**

   The default port number is 6789.

   `https://`*node*`:6789/`

▼ SPARC: How to Launch SunPlex Manager From the Sun Management Center Web Console

---

Note – You must possess the `solaris.cluster.gui` Role-Based Access Control (RBAC) authorization to log into SunPlex Manager. You can learn more about RBAC authorizations in Chapter 8, "Using Roles and Privileges (Overview)," in *System Administration Guide: Security Services*, Chapter 10, "Role-Based Access Control (Reference)," in *System Administration Guide: Security Services*, and in Chapter 2.

Steps

1. **Log in to the Sun Management Center Web Console.**

   The default port number is 6789.

   `https://`*`node`***`:6789/`**

2. **Choose the SunPlex Manager link**

   If you selected the "Start Each Application in a New Window" option after you logged in, SunPlex will display in a new browser window. Otherwise, SunPlex Manager will display in an existing browser window.

3. **To exit SunPlex Manager, click Log Out at the top, right corner of the SunPlex Manager workspace page.**

   SunPlex Manager exits.

Index

D

E

F

G

I

J

K

L

M

N

O

P

Q

R

S

T

U

V