# Sun Cluster Software Installation Guide for Solaris OS

Adobe PostScript™

# Contents

# Preface

The *Sun Cluster Software Installation Guide for Solaris OS* contains guidelines for planning a Sun™ Cluster configuration, and provides procedures for installing, configuring, and upgrading the Sun Cluster software on both SPARC® based systems and x86 based systems. This book also provides a detailed example of how to use Sun StorEdge™ Availability Suite software to configure data replication between clusters.

**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. Do not use this document as a presales guide. You should have already determined your system requirements and purchased the appropriate equipment and software before reading this document.

The instructions in this book assume knowledge of the Solaris™ Operating System (Solaris OS) and expertise with the volume-manager software that is used with Sun Cluster software.

**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 about commands that are used to install, configure, or upgrade a Sun Cluster configuration. This document might not contain complete information about basic UNIX® commands and procedures such as shutting down the system, booting the system, and configuring devices.

See one or more of the following sources for this information.

- Online documentation for the Solaris OS
- Other software documentation that you received with your system
- Solaris OS 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.<br><br>Use `ls -a` to list all files.<br><br>`machine_name% you have mail.` |
| **AaBbCc123** | What you type, contrasted with onscreen computer output | `machine_name% `**`su`**<br><br>`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*.<br><br>Perform a *patch analysis*.<br><br>Do *not* save the file.<br><br>[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* |

| Topic | Documentation |
|---|---|
| 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.

# Related Third-Party Web Site References

Sun is not responsible for the availability of third-party web sites mentioned in this document. Sun does not endorse and is not responsible or liable for any content, advertising, products, or other materials that are available on or through such sites or resources. Sun will not be responsible or liable for any actual or alleged damage or loss caused or alleged to be caused by or in connection with use of or reliance on any such content, goods, or services that are available on or through such sites or resources.

# 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

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

- Your name and email address (if available)
- Your company name, address, and phone number
- The model number and serial number of your systems
- The release number of the Solaris OS (for example, Solaris 8)
- The release number of Sun Cluster (for example, Sun Cluster 3.1 8/05)

Use the following commands to gather information about 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.

# Planning the Sun Cluster Configuration

This chapter provides planning information and guidelines for installing a Sun Cluster configuration.

The following overview information is in this chapter:

- "Finding Sun Cluster Installation Tasks" on page 15
- "Planning the Solaris OS" on page 16
- "Planning the Sun Cluster Environment" on page 21
- "Planning the Global Devices and Cluster File Systems" on page 32
- "Planning Volume Management" on page 35

## Finding Sun Cluster Installation Tasks

The following table shows where to find instructions for various installation tasks for Sun Cluster software installation and the order in which you should perform the tasks.

**TABLE 1–1** Sun Cluster Software Installation Task Information

| Task | Instructions |
| --- | --- |
| Set up cluster hardware. | *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*<br><br>Documentation that shipped with your server and storage devices |
| Plan cluster software installation. | Chapter 1<br><br>"Installation and Configuration Worksheets" on page 286 |

**TABLE 1–1** Sun Cluster Software Installation Task Information    *(Continued)*

| Task | Instructions |
|---|---|
| Install software packages. Optionally, install and configure Sun StorEdge QFS software. | "Installing the Software" on page 45<br><br>*Sun StorEdge QFS and Sun StorEdge SAM-FS Software Installation and Configuration Guide* |
| Establish a new cluster or a new cluster node. | "Establishing the Cluster" on page 63 |
| Install and configure Solstice DiskSuite™ or Solaris Volume Manager software. | "Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software" on page 141<br><br>Solstice DiskSuite or Solaris Volume Manager documentation |
| SPARC: Install and configure VERITAS Volume Manager (VxVM) software. | "SPARC: Installing and Configuring VxVM Software" on page 177<br><br>VxVM documentation |
| Configure cluster file systems and other cluster components. | "Configuring the Cluster" on page 118 |
| *(Optional)* SPARC: Install and configure the Sun Cluster module to Sun Management Center. | "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 130<br><br>Sun Management Center documentation |
| Plan, install, and configure resource groups and data services. | *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* |
| Develop custom data services. | *Sun Cluster Data Services Developer's Guide for Solaris OS* |
| Upgrade to Sun Cluster 3.1 8/05 software. | Chapter 5<br><br>"Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software" on page 141 or "SPARC: Installing and Configuring VxVM Software" on page 177<br><br>Volume manager documentation |

# Planning the Solaris OS

This section provides guidelines for planning Solaris software installation in a cluster configuration. For more information about Solaris software, see your Solaris installation documentation.

# Guidelines for Selecting Your Solaris Installation Method

You can install Solaris software from a local CD-ROM or from a network installation server by using the JumpStart™ installation method. In addition, Sun Cluster software provides a custom method for installing both the Solaris OS and Sun Cluster software by using the JumpStart installation method. If you are installing several cluster nodes, consider a network installation.

See "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 72 for details about the `scinstall` JumpStart installation method. See your Solaris installation documentation for details about standard Solaris installation methods.

## Solaris OS Feature Restrictions

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

- Sun Cluster 3.1 8/05 software does not support non-global zones. All Sun Cluster software and software that is managed by the cluster must be installed only on the global zone of the node. Do not install cluster-related software on a non-global zone. In addition, all cluster-related software must be installed in a way that prevents propagation to a non-global zone that is later created on a cluster node. For more information, see "Adding a Package to the Global Zone Only" in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

- Solaris interface groups are not supported in a Sun Cluster configuration. The Solaris interface groups feature is disabled by default during Solaris software installation. Do not re-enable Solaris interface groups. See the `ifconfig`(1M) man page for more information about Solaris interface groups.

- Automatic power-saving shutdown is not supported in Sun Cluster configurations and should not be enabled. See the `pmconfig`(1M) and `power.conf`(4) man pages for more information.

- Sun Cluster software does not support Extensible Firmware Interface (EFI) disk labels.

- Sun Cluster software does not support filtering with Solaris IP Filter. The use of the STREAMS `autopush`(1M) mechanism by Solaris IP Filter conflicts with Sun Cluster software's use of the mechanism.

## Solaris Software Group Considerations

Sun Cluster 3.1 8/05 software requires at least the End User Solaris Software Group. However, other components of your cluster configuration might have their own Solaris software requirements as well. Consider the following information when you decide which Solaris software group you are installing.

- Check your server documentation for any Solaris software requirements. For example , Sun Enterprise™ 10000 servers require the Entire Solaris Software Group Plus OEM Support.

- If you intend to use SCI-PCI adapters, which are available for use in SPARC based clusters only, or the Remote Shared Memory Application Programming Interface (RSMAPI), ensure that you install the RSMAPI software packages (SUNWrsm and SUNWrsmo, and also SUNWrsmx and SUNWrsmox for the Solaris 8 or Solaris 9 OS). The RSMAPI software packages are included only in some Solaris software groups. For example, the Developer Solaris Software Group includes the RSMAPI software packages but the End User Solaris Software Group does not.

  If the software group that you install does not include the RSMAPI software packages, install the RSMAPI software packages manually before you install Sun Cluster software. Use the pkgadd(1M) command to manually install the software packages. See the Solaris 8 Section (3RSM) man pages for information about using the RSMAPI.

- You might need to install other Solaris software packages that are not part of the End User Solaris Software Group. The Apache HTTP server packages are one example. Third-party software, such as ORACLE®, might also require additional Solaris software packages. See your third-party documentation for any Solaris software requirements.

---

**Tip –** To avoid the need to manually install Solaris software packages, install the Entire Solaris Software Group Plus OEM Support.

---

## System Disk Partitions

Add this information to the appropriate .

When you install the Solaris OS, ensure that you create the required Sun Cluster partitions and that all partitions meet minimum space requirements.

- **swap** – The combined amount of swap space that is allocated for Solaris and Sun Cluster software must be no less than 750 Mbytes. For best results, add at least 512 Mbytes for Sun Cluster software to the amount that is required by the Solaris OS. In addition, allocate any additional swap amount that is required by applications that are to run on the cluster node.

---

**Note –** If you intend to create an additional swap file, do not create the swap file on a global device. Only use a local disk as a swap device for the node.

---

- /globaldevices – Create a 512-Mbyte file system that is to be used by the scinstall(1M) utility for global devices.

- **Volume manager** – Create a 20-Mbyte partition on a slice at the end of the disk (slice 7) for volume manager use. If your cluster uses VERITAS Volume Manager (VxVM) and you intend to encapsulate the root disk, you need to have two unused slices available for use by VxVM.

To meet these requirements, you must customize the partitioning if you are performing interactive installation of the Solaris OS.

See the following guidelines for additional partition planning information:

- "Guidelines for the Root (/) File System" on page 19
- "Guidelines for the /globaldevices File System" on page 20
- "Volume Manager Requirements" on page 20

## Guidelines for the Root (/) File System

As with any other system running the Solaris OS, you can configure the root (/), /var, /usr, and /opt directories as separate file systems. Or, you can include all the directories in the root (/) file system. The following describes the software contents of the root (/), /var, /usr, and /opt directories in a Sun Cluster configuration. Consider this information when you plan your partitioning scheme.

- root (/) – The Sun Cluster software itself occupies less than 40 Mbytes of space in the root (/) file system. Solstice DiskSuite or Solaris Volume Manager software requires less than 5 Mbytes, and VxVM software requires less than 15 Mbytes. To configure ample additional space and inode capacity, add at least 100 Mbytes to the amount of space you would normally allocate for your root ( /) file system. This space is used for the creation of both block special devices and character special devices used by the volume management software. You especially need to allocate this extra space if a large number of shared disks are in the cluster.

- /var – The Sun Cluster software occupies a negligible amount of space in the /var file system at installation time. However, you need to set aside ample space for log files. Also, more messages might be logged on a clustered node than would be found on a typical standalone server. Therefore, allow at least 100 Mbytes for the /var file system.

- /usr – Sun Cluster software occupies less than 25 Mbytes of space in the /usr file system. Solstice DiskSuite or Solaris Volume Manager and VxVM software each require less than 15 Mbytes.

- /opt – Sun Cluster framework software uses less than 2 Mbytes in the /opt file system. However, each Sun Cluster data service might use between 1 Mbyte and 5 Mbytes. Solstice DiskSuite or Solaris Volume Manager software does not use any space in the /opt file system. VxVM software can use over 40 Mbytes if all of its packages and tools are installed.

  In addition, most database and applications software is installed in the /opt file system.

  SPARC: If you use Sun Management Center software to monitor the cluster, you need an additional 25 Mbytes of space on each node to support the Sun Management Center agent and Sun Cluster module packages.

## Guidelines for the `/globaldevices` File System

Sun Cluster software requires you to set aside a special file system on one of the local disks for use in managing global devices. This file system is later mounted as a cluster file system. Name this file system `/globaldevices`, which is the default name that is recognized by the `scinstall`(1M) command.

The `scinstall` command later renames the file system `/global/.devices/node@`*nodeid*, where *nodeid* represents the number that is assigned to a node when it becomes a cluster member. The original `/globaldevices` mount point is removed.

The `/globaldevices` file system must have ample space and ample inode capacity for creating both block special devices and character special devices. This guideline is especially important if a large number of disks are in the cluster. A file system size of 512 Mbytes should suffice for most cluster configurations.

## Volume Manager Requirements

If you use Solstice DiskSuite or Solaris Volume Manager software, you must set aside a slice on the root disk for use in creating the state database replica. Specifically, set aside a slice for this purpose on each local disk. But, if you only have one local disk on a node, you might need to create three state database replicas in the same slice for Solstice DiskSuite or Solaris Volume Manager software to function properly. See your Solstice DiskSuite or Solaris Volume Manager documentation for more information.

SPARC: If you use VERITAS Volume Manager (VxVM) and you intend to encapsulate the root disk, you need to have two unused slices that are available for use by VxVM. Additionally, you need to have some additional unassigned free space at either the beginning or the end of the disk. See your VxVM documentation for more information about root disk encapsulation.

## Example—Sample File-System Allocations

Table 1–2 shows a partitioning scheme for a cluster node that has less than 750 Mbytes of physical memory. This scheme is to be installed with the End User Solaris Software Group, Sun Cluster software, and the Sun Cluster HA for NFS data service. The last slice on the disk, slice 7, is allocated with a small amount of space for volume-manager use.

This layout allows for the use of either Solstice DiskSuite or Solaris Volume Manager software or VxVM software. If you use Solstice DiskSuite or Solaris Volume Manager software, you use slice 7 for the state database replica. If you use VxVM, you later free slice 7 by assigning the slice a zero length. This layout provides the necessary two free slices, 4 and 7, as well as provides for unused space at the end of the disk.

**TABLE 1–2** Example File-System Allocation

| Slice | Contents | Size Allocation | Description |
|---|---|---|---|
| 0 | / | 6.75GB | Remaining free space on the disk after allocating space to slices 1 through 7. Used for the Solaris OS, Sun Cluster software, data-services software, volume-manager software, Sun Management Center agent and Sun Cluster module agent packages, root file systems, and database and application software. |
| 1 | swap | 1GB | 512 Mbytes for the Solaris OS.<br><br>512 Mbytes for Sun Cluster software. |
| 2 | overlap | 8.43GB | The entire disk. |
| 3 | /globaldevices | 512MB | The Sun Cluster software later assigns this slice a different mount point and mounts the slice as a cluster file system. |
| 4 | unused | - | Available as a free slice for encapsulating the root disk under VxVM. |
| 5 | unused | - | - |
| 6 | unused | - | - |
| 7 | volume manager | 20MB | Used by Solstice DiskSuite or Solaris Volume Manager software for the state database replica, or used by VxVM for installation after you free the slice. |

# Planning the Sun Cluster Environment

This section provides guidelines for planning and preparing the following components for Sun Cluster software installation and configuration:

- "Licensing" on page 22
- "Software Patches" on page 22
- "IP Addresses" on page 22
- "Console-Access Devices" on page 23
- "Logical Addresses" on page 23
- "Public Networks" on page 24
- "IP Network Multipathing Groups" on page 25
- "Guidelines for NFS" on page 26
- "Service Restrictions" on page 27
- "Sun Cluster Configurable Components" on page 28

For detailed information about Sun Cluster components, see the *Sun Cluster Overview for Solaris OS* and the *Sun Cluster Concepts Guide for Solaris OS*.

## Licensing

Ensure that you have available all necessary license certificates before you begin software installation. Sun Cluster software does not require a license certificate, but each node installed with Sun Cluster software must be covered under your Sun Cluster software license agreement.

For licensing requirements for volume-manager software and applications software, see the installation documentation for those products.

## Software Patches

After installing each software product, you must also install any required patches.

- For information about current required patches, see "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* or consult your Sun service provider.
- For general guidelines and procedures for applying patches, see Chapter 8, "Patching Sun Cluster Software and Firmware," in *Sun Cluster System Administration Guide for Solaris OS*.

## IP Addresses

You must set up a number of IP addresses for various Sun Cluster components, depending on your cluster configuration. Each node in the cluster configuration must have at least one public-network connection to the same set of public subnets.

The following table lists the components that need IP addresses assigned. Add these IP addresses to the following locations:

- Any naming services that are used
- The local `/etc/inet/hosts` file on each cluster node, after you install Solaris software
- For Solaris 10, the local `/etc/inet/iphosts` file on each cluster node, after you install Solaris software

**TABLE 1–3** Sun Cluster Components That Use IP Addresses

| Component | Number of IP Addresses Needed |
| --- | --- |
| Administrative console | 1 per subnet. |

**TABLE 1–3** Sun Cluster Components That Use IP Addresses     *(Continued)*

| Component | Number of IP Addresses Needed |
|---|---|
| IP Network Multipathing groups | ■ Single-adapter groups – 1 primary IP address. For Solaris 8, also 1 test IP address for each adapter in the group.<br>■ Multiple-adapter groups – 1 primary IP address plus 1 test IP address for each adapter in the group. |
| Cluster nodes | 1 per node, per subnet. |
| Domain console network interface (Sun Fire™ 15000) | 1 per domain. |
| Console-access device | 1. |
| Logical addresses | 1 per logical host resource, per subnet. |

For more information about planning IP addresses, see *System Administration Guide, Volume 3* (Solaris 8) or System Administration Guide: IP Services (Solaris 9 or Solaris 10).

For more information about test IP addresses to support IP Network Multipathing, see *IP Network Multipathing Administration Guide*.

# Console-Access Devices

You must have console access to all cluster nodes. If you install Cluster Control Panel software on your administrative console, you must provide the hostname of the console-access device that is used to communicate with the cluster nodes.

■ A terminal concentrator is used to communicate between the administrative console and the cluster node consoles.

■ A Sun Enterprise 10000 server uses a System Service Processor (SSP) instead of a terminal concentrator.

■ A Sun Fire™ server uses a system controller instead of a terminal concentrator.

For more information about console access, see the *Sun Cluster Concepts Guide for Solaris OS*.

# Logical Addresses

Consider the following points when you plan your logical addresses:

■ Each data-service resource group that uses a logical address must have a hostname specified for each public network from which the logical address can be accessed.

- The IP address must be on the same subnet as the test IP address that is used by the IP Network Multipathing group that hosts the logical address.

For more information, see the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*. For additional information about data services and resources, also see the *Sun Cluster Overview for Solaris OS* and the *Sun Cluster Concepts Guide for Solaris OS*.

## Public Networks

Public networks communicate outside the cluster. Consider the following points when you plan your public-network configuration:

- Public networks and the private network (cluster interconnect) must use separate adapters, or you must configure tagged VLAN on tagged-VLAN capable adapters and VLAN-capable switches to use the same adapter for both the private interconnect and the public network.

- You must have at least one public network that is connected to all cluster nodes.

- You can have as many additional public-network connections as your hardware configuration allows.

- Sun Cluster software supports IPv4 addresses on the public network.

- Sun Cluster software supports IPv6 addresses on the public network under the following conditions or restrictions:

    - Sun Cluster software does not support IPv6 addresses on the public network if the private interconnect uses SCI adapters.

    - On Solaris 9 OS and Solaris 10 OS, Sun Cluster software supports IPv6 addresses for both failover and scalable data services.

    - On Solaris 8 OS, Sun Cluster software supports IPv6 addresses for failover data services only.

- Each public network adapter must belong to an Internet Protocol (IP) Network Multipathing (IP Network Multipathing) group. See "IP Network Multipathing Groups" on page 25 for additional guidelines.

- All public network adapters must use network interface cards (NICs) that support local MAC address assignment. Local MAC address assignment is a requirement of IP Network Multipathing.

- The `local-mac-address?` variable must use the default value `true` for Ethernet adapters. Sun Cluster software does not support a `local-mac-address?` value of `false` for Ethernet adapters. This requirement is a change from Sun Cluster 3.0, which did require a `local-mac-address?` value of `false`.

- During Sun Cluster installation on the Solaris 9 or Solaris 10 OS, the `scinstall` utility automatically configures a single-adapter IP Network Multipathing group for each public-network adapter. To modify these backup groups after installation, follow the procedures in "Administering IPMP (Tasks)" in *System Administration*

*Guide: IP Services* (Solaris 9 or Solaris 10).

- Sun Cluster configurations do not support filtering with Solaris IP Filter.

See "IP Network Multipathing Groups" on page 25 for guidelines on planning public-network-adapter backup groups. For more information about public-network interfaces, see *Sun Cluster Concepts Guide for Solaris OS*.

## IP Network Multipathing Groups

Add this planning information to the "Public Networks Worksheet" on page 290.

Internet Protocol (IP) Network Multipathing groups, which replace Network Adapter Failover (NAFO) groups, provide public-network adapter monitoring and failover, and are the foundation for a network-address resource. A multipathing group provides high availability when the multipathing group is configured with two or more adapters. If one adapter fails, all of the addresses on the failed adapter fail over to another adapter in the multipathing group. In this way, the multipathing-group adapters maintain public-network connectivity to the subnet to which the adapters in the multipathing group connect.

The following describes the circumstances when you must manually configure IP Network Multipathing groups during a Sun Cluster software installation:

- For Sun Cluster software installations on the Solaris 8 OS, you must manually configure all public network adapters in IP Network Multipathing groups, with test IP addresses.
- If you use SunPlex Installer to install Sun Cluster software on the Solaris 9 or Solaris 10 OS, some but not all public network adapters might need to be manually configured in IP Network Multipathing groups.

For Sun Cluster software installations on the Solaris 9 or Solaris 10 OS, except when using SunPlex Installer, the `scinstall` utility automatically configures all public network adapters as single-adapter IP Network Multipathing groups.

Consider the following points when you plan your multipathing groups.

- Each public network adapter must belong to a multipathing group.
- In the following kinds of multipathing groups, you must configure a test IP address for each adapter in the group:
  - On the Solaris 8 OS, all multipathing groups require a test IP address for each adapter.
  - On the Solaris 9 or Solaris 10 OS, multipathing groups that contain two or more adapters require test IP addresses. If a multipathing group contains only one adapter, you do not need to configure a test IP address.
- 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 the test IP addresses are not highly available.

- In the `/etc/default/mpathd` file, the value of `TRACK_INTERFACES_ONLY_WITH_GROUPS` must be `yes`.

- The name of a multipathing group has no requirements or restrictions.

Most procedures, guidelines, and restrictions that are identified in the Solaris documentation for IP Network Multipathing are the same for both cluster and noncluster environments. Therefore, see the appropriate Solaris document for additional information about IP Network Multipathing:

- For the Solaris 8 OS, see Deploying Network Multipathing in *IP Network Multipathing Administration Guide*.

- For the Solaris 9 OS, see Chapter 28, "Administering Network Multipathing (Task)," in *System Administration Guide: IP Services*.

- For the Solaris 10 OS, see Chapter 30, "Administering IPMP (Tasks)," in *System Administration Guide: IP Services*.

Also see "IP Network Multipathing Groups" in *Sun Cluster Overview for Solaris OS* and *Sun Cluster Concepts Guide for Solaris OS*.

## Guidelines for NFS

Consider the following points when you plan the use of Network File System (NFS) in a Sun Cluster configuration.

- No Sun Cluster node can be an NFS client of a Sun Cluster HA for NFS-exported file system being mastered on a node in the same cluster. Such cross-mounting of Sun Cluster HA for NFS is prohibited. Use the cluster file system to share files among cluster nodes.

- Applications that run locally on the cluster must not lock files on a file system that is exported through NFS. Otherwise, local blocking (for example, `flock`(3UCB) or `fcntl`(2)) might interfere with the ability to restart the lock manager ( `lockd`(1M)). During restart, a blocked local process might be granted a lock which might be intended to be reclaimed by a remote client. This would cause unpredictable behavior.

- Sun Cluster software does not support the following options of the `share_nfs`(1M) command:

  - `secure`
  - `sec=dh`

However, Sun Cluster software does support the following security features for NFS:

- The use of secure ports for NFS. You enable secure ports for NFS by adding the entry set `nfssrv:nfs_portmon=1` to the `/etc/system` file on cluster nodes.
- The use of Kerberos with NFS. For more information, see "Securing Sun Cluster HA for NFS With Kerberos V5" in *Sun Cluster Data Service for NFS Guide for Solaris OS*.

## Service Restrictions

Observe the following service restrictions for Sun Cluster configurations:

- Do not configure cluster nodes as routers (gateways). If the system goes down, the clients cannot find an alternate router and cannot recover.
- Do not configure cluster nodes as NIS or NIS+ servers. There is no data service available for NIS or NIS+. However, cluster nodes can be NIS or NIS+ clients.
- Do not use a Sun Cluster configuration to provide a highly available boot or installation service on client systems.
- Do not use a Sun Cluster configuration to provide an `rarpd` service.
- If you install an RPC service on the cluster, the service must not use any of the following program numbers:

  - 100141
  - 100142
  - 100248

  These numbers are reserved for the Sun Cluster daemons `rgmd_receptionist`, `fed`, and `pmfd`, respectively.

  If the RPC service that you install also uses one of these program numbers, you must change that RPC service to use a different program number.

- Sun Cluster software does not support the running of high-priority process scheduling classes on cluster nodes. Do not run either of the following types of processes on cluster nodes:

  - Processes that run in the time-sharing scheduling class with a high priority
  - Processes that run in the real-time scheduling class

  Sun Cluster software relies on kernel threads that do not run in the real-time scheduling class. Other time-sharing processes that run at higher-than-normal priority or real-time processes can prevent the Sun Cluster kernel threads from acquiring needed CPU cycles.

# Sun Cluster Configurable Components

This section provides guidelines for the following Sun Cluster components that you configure:

-
-
-
-
-
-

Add this information to the appropriate configuration planning worksheet.

## Cluster Name

Specify a name for the cluster during Sun Cluster configuration. The cluster name should be unique throughout the enterprise.

## Node Names

The node name is the name that you assign to a machine when you install the Solaris OS. During Sun Cluster configuration, you specify the names of all nodes that you are installing as a cluster. In single-node cluster installations, the default cluster name is the node name.

## Private Network

---

**Note –** You do not need to configure a private network for a single-node cluster.

---

Sun Cluster software uses the private network for internal communication between nodes. A Sun Cluster configuration requires at least two connections to the cluster interconnect on the private network. You specify the private-network address and netmask when you configure Sun Cluster software on the first node of the cluster. You can either accept the default private-network address (`172.16.0.0`) and netmask (`255.255.0.0`) or type different choices.

**Note –** After the installation utility (`scinstall`, SunPlex Installer, or JumpStart) has finished processing and the cluster is established, you cannot change the private-network address and netmask. You must uninstall and reinstall the cluster software to use a different private-network address or netmask.

If you specify a private-network address other than the default, the address must meet the following requirements:

- The address must use zeroes for the last two octets of the address, as in the default address `172.16.0.0`. Sun Cluster software requires the last 16 bits of the address space for its own use.

- The address must be included in the block of addresses that RFC 1918 reserves for use in private networks. You can contact the InterNIC to obtain copies of RFCs or view RFCs online at http://www.rfcs.org.

- You can use the same private network address in more than one cluster. Private IP network addresses are not accessible from outside the cluster.

- Sun Cluster software does not support IPv6 addresses for the private interconnect. The system does configure IPv6 addresses on the private network adapters to support scalable services that use IPv6 addresses. But internode communication on the private network does not use these IPv6 addresses.

Although the `scinstall` utility lets you specify an alternate netmask, best practice is to accept the default netmask, `255.255.0.0`. There is no benefit if you specify a netmask that represents a larger network. And the `scinstall` utility does not accept a netmask that represents a smaller network.

See "Planning Your TCP/IP Network" in *System Administration Guide, Volume 3* (Solaris 8) or "Planning Your TCP/IP Network (Tasks)," in *System Administration Guide: IP Services* (Solaris 9 or Solaris 10) for more information about private networks.

## Private Hostnames

The private hostname is the name that is used for internode communication over the private-network interface. Private hostnames are automatically created during Sun Cluster configuration. These private hostnames follow the naming convention `clusternode`*nodeid* `-priv`, where *nodeid* is the numeral of the internal node ID. During Sun Cluster configuration, the node ID number is automatically assigned to each node when the node becomes a cluster member. After the cluster is configured, you can rename private hostnames by using the `scsetup`(1M) utility.

# Cluster Interconnect

---

**Note –** You do not need to configure a cluster interconnect for a single-node cluster. However, if you anticipate eventually adding nodes to a single-node cluster configuration, you might want to configure the cluster interconnect for future use.

---

The cluster interconnects provide the hardware pathways for private-network communication between cluster nodes. Each interconnect consists of a cable that is connected in one of the following ways:

- Between two transport adapters
- Between a transport adapter and a transport junction
- Between two transport junctions

During Sun Cluster configuration, you specify configuration information for two cluster interconnects. You can configure additional private-network connections after the cluster is established by using the scsetup(1M) utility.

For guidelines about cluster interconnect hardware, see "Interconnect Requirements and Restrictions" in *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*.For general information about the cluster interconnect, see "Cluster Interconnect" in *Sun Cluster Overview for Solaris OS* and *Sun Cluster Concepts Guide for Solaris OS*.

## *Transport Adapters*

For the transport adapters, such as ports on network interfaces, specify the transport adapter names and transport type. If your configuration is a two-node cluster, you also specify whether your interconnect is direct connected (adapter to adapter) or uses a transport junction.

Consider the following guidelines and restrictions:

- **IPv6** - Sun Cluster software does not support IPv6 communications over the private interconnects.

- **Local MAC address assignment** - All private network adapters must use network interface cards (NICs) that support local MAC address assignment. Link-local IPv6 addresses, which are required on private network adapters to support IPv6 public network addresses, are derived from the local MAC addresses.

- **Tagged VLAN adapters** – Sun Cluster software supports tagged Virtual Local Area Networks (VLANs) to share an adapter between the private interconnect and the public network. To configure a tagged VLAN adapter for the private interconnect, specify the adapter name and its VLAN ID (VID) in one of the following ways:

    - Specify the usual adapter name, which is the device name plus the instance number or physical point of attachment (PPA). For example, the name of instance 2 of a Cassini Gigabit Ethernet adapter would be ce2. If the

`scinstall` utility asks whether the adapter is part of a shared virtual LAN, answer **yes** and specify the adapter's VID number.

- Specify the adapter by its VLAN virtual device name. This name is composed of the adapter name plus the VLAN instance number. The VLAN instance number is derived from the formula $(1000*V)+N$, where $V$ is the VID number and $N$ is the PPA.

  As an example, for VID 73 on adapter `ce2`, the VLAN instance number would be calculated as (1000*73)+2. You would therefore specify the adapter name as `ce73002` to indicate that it is part of a shared virtual LAN.

  For more information about VLAN, see "Configuring VLANs" in *Solaris 9 9/04 Sun Hardware Platform Guide*.

- **SBus SCI adapters** – The SBus Scalable Coherent Interface (SCI) is not supported as a cluster interconnect. However, the SCI–PCI interface is supported.

- **Logical network interfaces** - Logical network interfaces are reserved for use by Sun Cluster software.

See the `scconf_trans_adap_*`(1M) family of man pages for information about a specific transport adapter.

### *Transport Junctions*

If you use transport junctions, such as a network switch, specify a transport junction name for each interconnect. You can use the default name `switch`*N*, where *N* is a number that is automatically assigned during configuration, or create another name. The exception is the Sun Fire Link adapter, which requires the junction name `sw-rsm`*N*. The `scinstall` utility automatically uses this junction name after you specify a Sun Fire Link adapter (`wrsm`*N*).

Also specify the junction port name or accept the default name. 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-PCI.

---

**Note –** Clusters with three or more nodes *must* use transport junctions. Direct connection between cluster nodes is supported only for two-node clusters.

---

If your two-node cluster is direct connected, you can still specify a transport junction for the interconnect.

---

**Tip –** If you specify a transport junction, you can more easily add another node to the cluster in the future.

---

## Quorum Devices

Sun Cluster configurations use quorum devices to maintain data and resource integrity. If the cluster temporarily loses connection to a node, the quorum device prevents amnesia or split-brain problems when the cluster node attempts to rejoin the cluster. During Sun Cluster installation of a two-node cluster, the scinstall utility automatically configures a quorum device. The quorum device is chosen from the available shared storage disks. The scinstall utility assumes that all available shared storage disks are supported to be quorum devices. After installation, you can also configure additional quorum devices by using the scsetup(1M) utility.

---

**Note –** You do not need to configure quorum devices for a single-node cluster.

---

If your cluster configuration includes third-party shared storage devices that are not supported for use as quorum devices, you must use the scsetup utility to configure quorum manually.

Consider the following points when you plan quorum devices.

- **Minimum** – A two-node cluster must have at least one quorum device, which can be a shared disk or a Network Appliance NAS device. For other topologies, quorum devices are optional.

- **Odd-number rule** – If more than one quorum device is configured in a two-node cluster, or in a pair of nodes directly connected to the quorum device, configure an odd number of quorum devices. This configuration ensures that the quorum devices have completely independent failure pathways.

- **Connection** – You must connect a quorum device to at least two nodes.

For more information about quorum devices, see "Quorum and Quorum Devices" in *Sun Cluster Concepts Guide for Solaris OS* and "Quorum Devices" in *Sun Cluster Overview for Solaris OS*.

---

# Planning the Global Devices and Cluster File Systems

This section provides the following guidelines for planning global devices and for planning cluster file systems:

- "Guidelines for Highly Available Global Devices and Cluster File Systems" on page 33
- "Disk Device Groups" on page 34

- "Mount Information for Cluster File Systems" on page 34

For more information about global devices and about cluster files systems, see *Sun Cluster Overview for Solaris OS* and *Sun Cluster Concepts Guide for Solaris OS*.

## Guidelines for Highly Available Global Devices and Cluster File Systems

Sun Cluster software does not require any specific disk layout or file system size. Consider the following points when you plan your layout for global devices and for cluster file systems.

- **Mirroring** – You must mirror all global devices for the global device to be considered highly available. You do not need to use software mirroring if the storage device provides hardware RAID as well as redundant paths to disks.
- **Disks** – When you mirror, lay out file systems so that the file systems are mirrored across disk arrays.
- **Availability** – You must physically connect a global device to more than one node in the cluster for the global device to be considered highly available. A global device with multiple physical connections can tolerate a single-node failure. A global device with only one physical connection is supported, but the global device becomes inaccessible from other nodes if the node with the connection is down.
- **Swap devices** - Do not create a swap file on a global device.

## Cluster File Systems

Consider the following points when you plan cluster file systems.

- **Quotas** - Quotas are not supported on cluster file systems.
- **Loopback file system (LOFS)** - Do *not* use the loopback file system (LOFS) if both conditions in the following list are met:

  - Sun Cluster HA for NFS is configured on a highly available local file system.
  - The `automountd` daemon is running.

  If both of these conditions are met, LOFS must be disabled to avoid switchover problems or other failures. If only one of these conditions is met, it is safe to enable LOFS.

  If you require both LOFS and the `automountd` daemon to be enabled, exclude from the automounter map all files that are part of the highly available local file system that is exported by Sun Cluster HA for NFS.

- **Process accounting log files** - Do not locate process accounting log files on a cluster file system or on a highly available local file system. A switchover would be blocked by writes to the log file, which would cause the node to hang. Use only a local file system to contain process accounting log files.

- **Communication endpoints** - The cluster file system does not support any of the file-system features of Solaris software by which one would put a communication endpoint in the file-system namespace.

    - Although you can create a UNIX domain socket whose name is a path name into the cluster file system, the socket would not survive a node failover.

    - Any FIFOs or named pipes that you create on a cluster file system would not be globally accessible.

    Therefore, do not attempt to use the `fattach` command from any node other than the local node.

## Disk Device Groups

Add this planning information to the "Disk Device Group Configurations Worksheet" on page 294.

You must configure all volume-manager disk groups as Sun Cluster disk device groups. This configuration enables a secondary node to host multihost disks if the primary node fails. Consider the following points when you plan disk device groups.

- **Failover** – You can configure multihost disks and properly configured volume-manager devices as failover devices. Proper configuration of a volume-manager device includes multihost disks and correct setup of the volume manager itself. This configuration ensures that multiple nodes can host the exported device. You cannot configure tape drives, CD-ROMs, or single-ported devices as failover devices.

- **Mirroring** – You must mirror the disks to protect the data from disk failure. See "Mirroring Guidelines" on page 41 for additional guidelines. See "Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software" on page 141 or "SPARC: Installing and Configuring VxVM Software" on page 177 and your volume-manager documentation for instructions about mirroring.

For more information about disk device groups, see "Devices" in *Sun Cluster Overview for Solaris OS* and *Sun Cluster Concepts Guide for Solaris OS*.

## Mount Information for Cluster File Systems

Consider the following points when you plan mount points for cluster file systems.

- **Mount-point location** – Create mount points for cluster file systems in the `/global` directory, unless you are prohibited by other software products. By using the `/global` directory, you can more easily distinguish cluster file systems, which are globally available, from local file systems.

- **SPARC: VxFS mount requirement** – If you use VERITAS File System (VxFS), globally mount and unmount a VxFS file system from the primary node. The primary node is the node that masters the disk on which the VxFS file system

resides. This method ensures that the mount or unmount operation succeeds. A VxFS file-system mount or unmount operation that is performed from a secondary node might fail.

- 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.

- **Nesting mount points** – Normally, you should not nest the mount points for cluster file systems. For example, do not set up one file system that is mounted on `/global/a` and another file system that is mounted on `/global/a/b`. To ignore this rule can cause availability and node boot-order problems. These problems would occur if the parent mount point is not present when the system attempts to mount a child of that file system. The only exception to this rule is if the devices for the two file systems have the same physical node connectivity. An example is different slices on the same disk.

- `forcedirectio` - Sun Cluster software does not support the execution of binaries off cluster file systems that are mounted by using the `forcedirectio` mount option.

# Planning Volume Management

Add this planning information to the "Disk Device Group Configurations Worksheet" on page 294 and the "Volume-Manager Configurations Worksheet" on page 296. For Solstice DiskSuite or Solaris Volume Manager, also add this planning information to the "Metadevices Worksheet (Solstice DiskSuite or Solaris Volume Manager)" on page 298.

This section provides the following guidelines for planning volume management of your cluster configuration:

Sun Cluster software uses volume-manager software to group disks into disk device groups which can then be administered as one unit. Sun Cluster software supports Solstice DiskSuite or Solaris Volume Manager software and VERITAS Volume Manager (VxVM) software that you install or use in the following ways.

**TABLE 1–4** Supported Use of Volume Managers With Sun Cluster Software

| Volume-Manager Software | Requirements |
|---|---|
| Solstice DiskSuite or Solaris Volume Manager | You must install Solstice DiskSuite or Solaris Volume Manager software on all nodes of the cluster, regardless of whether you use VxVM on some nodes to manage disks. |
| SPARC: VxVM with the cluster feature | You must install and license VxVM with the cluster feature on all nodes of the cluster. |
| SPARC: VxVM without the cluster feature | You are only required to install and license VxVM on those nodes that are attached to storage devices which VxVM manages. |
| SPARC: Both Solstice DiskSuite or Solaris Volume Manager and VxVM | If you install both volume managers on the same node, you must use Solstice DiskSuite or Solaris Volume Manager software to manage disks that are local to each node. Local disks include the root disk. Use VxVM to manage all shared disks. |

See your volume-manager documentation and "Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software" on page 141 or "SPARC: Installing and Configuring VxVM Software" on page 177 for instructions about how to install and configure the volume-manager software. For more information about volume management in a cluster configuration, see the *Sun Cluster Concepts Guide for Solaris OS*.

## Guidelines for Volume-Manager Software

Consider the following general guidelines when you configure your disks with volume-manager software:

- **Software RAID** – Sun Cluster software does not support software RAID 5.

- **Mirrored multihost disks** – You must mirror all multihost disks across disk expansion units. See "Guidelines for Mirroring Multihost Disks" on page 42 for guidelines on mirroring multihost disks. You do not need to use software mirroring

if the storage device provides hardware RAID as well as redundant paths to devices.

- **Mirrored root** – Mirroring the root disk ensures high availability, but such mirroring is not required. See "Mirroring Guidelines" on page 41 for guidelines about deciding whether to mirror the root disk.

- **Unique naming** – You might have local Solstice DiskSuite metadevices, local Solaris Volume Manager volumes, or VxVM volumes that are used as devices on which the /global/.devices/node@*nodeid* file systems are mounted. If so, the name of each local metadevice or local volume on which a /global/.devices/node@*nodeid* file system is to be mounted must be unique throughout the cluster.

- **Node lists** – To ensure high availability of a disk device group, make its node lists of potential masters and its failback policy identical to any associated resource group. Or, if a scalable resource group uses more nodes than its associated disk device group, make the scalable resource group's node list a superset of the disk device group's node list. See the resource group planning information in the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* for information about node lists.

- **Multihost disks** – You must connect, or port, all devices that are used to construct a device group to all of the nodes that are configured in the node list for that device group. Solstice DiskSuite or Solaris Volume Manager software can automatically check for this connection at the time that devices are added to a disk set. However, configured VxVM disk groups do not have an association to any particular set of nodes.

- **Hot spare disks** – You can use hot spare disks to increase availability, but hot spare disks are not required.

See your volume-manager documentation for disk layout recommendations and any additional restrictions.

## Guidelines for Solstice DiskSuite or Solaris Volume Manager Software

Consider the following points when you plan Solstice DiskSuite or Solaris Volume Manager configurations:

- **Local metadevice names or volume names** – The name of each local Solstice DiskSuite metadevice or Solaris Volume Manager volume on which a global–devices file system, /global/.devices/node@*nodeid*, is mounted must be unique throughout the cluster. Also, the name cannot be the same as any device-ID name.

- **Dual-string mediators** – Each disk set configured with exactly two disk strings and mastered by exactly two nodes must have Solstice DiskSuite or Solaris Volume Manager mediators configured for the disk set. A disk string consists of a disk enclosure, its physical disks, cables from the enclosure to the node(s), and the

interface adapter cards. Observe the following rules to configure dual-string mediators:

- You must configure each disk set with exactly two nodes that act as mediator hosts.
- You must use the same two nodes for all disk sets that require mediators. Those two nodes must master those disk sets.
- Mediators cannot be configured for disk sets that do not meet the two-string and two-host requirements.

See the mediator(7D) man page for details.

- /kernel/drv/md.conf **settings** – All Solstice DiskSuite metadevices or Solaris 9 Solaris Volume Manager volumes used by each disk set are created in advance, at reconfiguration boot time. This reconfiguration is based on the configuration parameters that exist in the /kernel/drv/md.conf file.

---

**Note –** With the Solaris 10 release, Solaris Volume Manager has been enhanced to configure volumes dynamically. You no longer need to edit the nmd and the md_nsets parameters in the /kernel/drv/md.conf file. New volumes are dynamically created, as needed.

---

You must modify the nmd and md_nsets fields as follows to support a Sun Cluster configuration on the Solaris 8 or Solaris 9 OS:

---

**Caution –** All cluster nodes must have identical /kernel/drv/md.conf files, regardless of the number of disk sets that are served by each node. Failure to follow this guideline can result in serious Solstice DiskSuite or Solaris Volume Manager errors and possible loss of data.

---

- md_nsets – The md_nsets field defines the total number of disk sets that can be created for a system to meet the needs of the entire cluster. Set the value of md_nsets to the expected number of disk sets in the cluster plus one additional disk set. Solstice DiskSuite or Solaris Volume Manager software uses the additional disk set to manage the private disks on the local host.

  The maximum number of disk sets that are allowed per cluster is 32. This number allows for 31 disk sets for general use plus one disk set for private disk management. The default value of md_nsets is 4.

- nmd – The nmd field defines the highest predicted value of any metadevice or volume name that will exist in the cluster. For example, if the highest value of the metadevice or volume names that are used in the first 15 disk sets of a cluster is 10, but the highest value of the metadevice or volume in the 16th disk set is 1000, set the value of nmd to at least 1000. Also, the value of nmd must be large enough to ensure that enough numbers exist for each device–ID name. The number must also be large enough to ensure that each local metadevice

name or local volume name can be unique throughout the cluster.

The highest allowed value of a metadevice or volume name per disk set is 8192. The default value of `nmd` is 128.

Set these fields at installation time to allow for all predicted future expansion of the cluster. To increase the value of these fields after the cluster is in production is time consuming. The value change requires a reconfiguration reboot for each node. To raise these values later also increases the possibility of inadequate space allocation in the root (/) file system to create all of the requested devices.

At the same time, keep the value of the `nmd`field and the `md_nsets` field as low as possible. Memory structures exist for all possible devices as determined by `nmd`and `md_nsets`, even if you have not created those devices. For optimal performance, keep the value of `nmd` and `md_nsets` only slightly higher than the number of metadevices or volumes you plan to use.

See "System and Startup Files" in *Solstice DiskSuite 4.2.1 Reference Guide* (Solaris 8) or "System Files and Startup Files" in *Solaris Volume Manager Administration Guide* (Solaris 9 or Solaris 10) for more information about the `md.conf` file.

# SPARC: Guidelines for VERITAS Volume Manager Software

Consider the following points when you plan VERITAS Volume Manager (VxVM) configurations.

- **Enclosure-Based Naming** – If you use Enclosure-Based Naming of devices, ensure that you use consistent device names on all cluster nodes that share the same storage. VxVM does not coordinate these names, so the administrator must ensure that VxVM assigns the same names to the same devices from different nodes. Failure to assign consistent names does not interfere with correct cluster behavior. However, inconsistent names greatly complicate cluster administration and greatly increase the possibility of configuration errors, potentially leading to loss of data.

- **Root disk group** – As of VxVM 4.0, the creation of a root disk group is optional.

  A root disk group can be created on the following disks:

  - The root disk, which must be encapsulated
  - One or more local nonroot disks, which you can encapsulate or initialize
  - A combination of root and local nonroot disks

  The root disk group must be local to the node.

- **Simple root disk groups** – Simple root disk groups (`rootdg` created on a single slice of the root disk) are not supported as disk types with VxVM on Sun Cluster software. This is a general VxVM software restriction.

- **Encapsulation** – Disks to be encapsulated must have two disk-slice table entries free.

- **Number of volumes** – Estimate the maximum number of volumes any given disk device group can use at the time the disk device group is created.

  - If the number of volumes is less than 1000, you can use default minor numbering.
  - If the number of volumes is 1000 or greater, you must carefully plan the way in which minor numbers are assigned to disk device group volumes. No two disk device groups can have overlapping minor number assignments.

- **Dirty Region Logging** – The use of Dirty Region Logging (DRL) decreases volume recovery time after a node failure. Using DRL might decrease I/O throughput.

- **Dynamic Multipathing (DMP)** – The use of DMP alone to manage multiple I/O paths per node to the shared storage is not supported. The use of DMP is supported only in the following configurations:

  - A single I/O path per node to the cluster's shared storage.
  - A supported multipathing solution, such as Sun Traffic Manager, EMC PowerPath, or Hitachi HDLM, that manages multiple I/O paths per node to the shared cluster storage.

See your VxVM installation documentation for additional information.

## File-System Logging

Logging is required for UFS and VxFS cluster file systems. This requirement does not apply to QFS shared file systems. Sun Cluster software supports the following choices of file-system logging:

- Solaris UFS logging – See the `mount_ufs`(1M) man page for more information.
- Solstice DiskSuite trans-metadevice logging or Solaris Volume Manager transactional-volume logging – See Chapter 2, "Creating DiskSuite Objects," in *Solstice DiskSuite 4.2.1 User's Guide* or "Transactional Volumes (Overview)" in *Solaris Volume Manager Administration Guide* for more information. Transactional volumes are no longer valid as of the Solaris 10 release of Solaris Volume Manager.
- SPARC: VERITAS File System (VxFS) logging – See the `mount_vxfs` man page provided with VxFS software for more information.

The following table lists the file-system logging supported by each volume manager.

**TABLE 1–5** Supported File System Logging Matrix

| Volume Manager | Supported File System Logging |
|---|---|
| Solstice DiskSuite or Solaris Volume Manager | ■ Solaris UFS logging<br>■ Solstice DiskSuite trans-metadevice logging<br>■ Solaris Volume Manager transactional-volume logging<br>■ VxFS logging |
| SPARC: VERITAS Volume Manager | ■ Solaris UFS logging<br>■ VxFS logging |

Consider the following points when you choose between Solaris UFS logging and Solstice DiskSuite trans-metadevice logging or Solaris Volume Manager transactional-volume logging for UFS cluster file systems:

■ Solaris Volume Manager transactional-volume logging (formerly Solstice DiskSuite trans-metadevice logging) is scheduled to be removed from the Solaris OS in an upcoming Solaris release. Solaris UFS logging provides the same capabilities but superior performance, as well as lower system administration requirements and overhead.

■ **Solaris UFS log size** – Solaris UFS logging always allocates the log by using free space on the UFS file system, and depending on the size of the file system.

  ■ On file systems less than 1 Gbyte, the log occupies 1 Mbyte.

  ■ On file systems 1 Gbyte or greater, the log occupies 1 Mbyte per Gbyte on the file system, to a maximum of 64 Mbytes.

■ **Log metadevice/transactional volume** – A Solstice DiskSuite trans metadevice or Solaris Volume Manager transactional volume manages UFS logging. The logging device component of a trans metadevice or transactional volume is a metadevice or volume that you can mirror and stripe. You can create a maximum 1-Gbyte log size, although 64 Mbytes is sufficient for most file systems. The minimum log size is 1 Mbyte.

## Mirroring Guidelines

This section provides the following guidelines for planning the mirroring of your cluster configuration:

## Guidelines for Mirroring Multihost Disks

To mirror all multihost disks in a Sun Cluster configuration enables the configuration to tolerate single-device failures. Sun Cluster software requires that you mirror all multihost disks across expansion units. You do not need to use software mirroring if the storage device provides hardware RAID as well as redundant paths to devices.

Consider the following points when you mirror multihost disks:

- **Separate disk expansion units** – Each submirror of a given mirror or plex should reside in a different multihost expansion unit.
- **Disk space** – Mirroring doubles the amount of necessary disk space.
- **Three-way mirroring** – Solstice DiskSuite or Solaris Volume Manager software and VERITAS Volume Manager (VxVM) software support three-way mirroring. However, Sun Cluster software requires only two-way mirroring.
- **Differing device sizes** – If you mirror to a device of a different size, your mirror capacity is limited to the size of the smallest submirror or plex.

For more information about multihost disks, see "Multihost Disk Storage" in *Sun Cluster Overview for Solaris OS* and *Sun Cluster Concepts Guide for Solaris OS*.

## Guidelines for Mirroring the Root Disk

Add this planning information to the "Local File System Layout Worksheet" on page 288.

For maximum availability, mirror root (/), /usr, /var, /opt, and swap on the local disks. Under VxVM, you encapsulate the root disk and mirror the generated subdisks. However, Sun Cluster software does not require that you mirror the root disk.

Before you decide whether to mirror the root disk, consider the risks, complexity, cost, and service time for the various alternatives that concern the root disk. No single mirroring strategy works for all configurations. You might want to consider your local Sun service representative's preferred solution when you decide whether to mirror root.

See your volume-manager documentation and "Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software" on page 141 or "SPARC: Installing and Configuring VxVM Software" on page 177 for instructions about how to mirror the root disk.

Consider the following points when you decide whether to mirror the root disk.

- **Boot disk** – You can set up the mirror to be a bootable root disk. You can then boot from the mirror if the primary boot disk fails.
- **Complexity** – To mirror the root disk adds complexity to system administration. To mirror the root disk also complicates booting in single-user mode.

- **Backups** – Regardless of whether you mirror the root disk, you also should perform regular backups of root. Mirroring alone does not protect against administrative errors. Only a backup plan enables you to restore files that have been accidentally altered or deleted.

- **Quorum devices** – Do not use a disk that was configured as a quorum device to mirror a root disk.

- **Quorum** – Under Solstice DiskSuite or Solaris Volume Manager software, in failure scenarios in which state database quorum is lost, you cannot reboot the system until maintenance is performed. See your Solstice DiskSuite or Solaris Volume Manager documentation for information about the state database and state database replicas.

- **Separate controllers** – Highest availability includes mirroring the root disk on a separate controller.

- **Secondary root disk** – With a mirrored root disk, the primary root disk can fail but work can continue on the secondary (mirror) root disk. Later, the primary root disk might return to service, for example, after a power cycle or transient I/O errors. Subsequent boots are then performed by using the primary root disk that is specified for the eeprom(1M) boot-device parameter. In this situation, no manual repair task occurs, but the drive starts working well enough to boot. With Solstice DiskSuite or Solaris Volume Manager software, a resync does occur. A resync requires a manual step when the drive is returned to service.

  If changes were made to any files on the secondary (mirror) root disk, they would not be reflected on the primary root disk during boot time. This condition would cause a stale submirror. For example, changes to the /etc/system file would be lost. With Solstice DiskSuite or Solaris Volume Manager software, some administrative commands might have changed the /etc/system file while the primary root disk was out of service.

  The boot program does not check whether the system is booting from a mirror or from an underlying physical device. The mirroring becomes active partway through the boot process, after the metadevices or volumes are loaded. Before this point, the system is therefore vulnerable to stale submirror problems.

# Installing and Configuring Sun Cluster Software

This chapter provides procedures for how to install and configure your cluster. You can also use these procedures to add a new node to an existing cluster. This chapter also provides procedures to uninstall certain cluster software.

The following sections are in this chapter.

# Installing the Software

This section provides information and procedures to install software on the cluster nodes.

The following task map lists the tasks that you perform to install software on multiple-node or single-node clusters. Complete the procedures in the order that is indicated.

**TABLE 2–1** Task Map: Installing the Software

| Task | Instructions |
|---|---|
| 1. Plan the layout of your cluster configuration and prepare to install software. | "How to Prepare for Cluster Software Installation" on page 46 |

**TABLE 2–1** Task Map: Installing the Software     *(Continued)*

| Task | Instructions |
|---|---|
| 2. (*Optional*) Install Cluster Control Panel (CCP) software on the administrative console. | "How to Install Cluster Control Panel Software on an Administrative Console" on page 48 |
| 3. Install the Solaris OS on all nodes. | "How to Install Solaris Software" on page 52 |
| 4. (*Optional*) SPARC: Install Sun StorEdge Traffic Manager software. | "SPARC: How to Install Sun Multipathing Software" on page 56 |
| 5. (*Optional*) SPARC: Install VERITAS File System software. | "SPARC: How to Install VERITAS File System Software" on page 59 |
| 6. Install Sun Cluster software packages and any Sun Java System data services for the Solaris 8 or Solaris 9 OS that you will use. | "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES `installer`)" on page 59 |
| 7. Set up directory paths. | "How to Set Up the Root Environment" on page 63 |
| 8. Establish the cluster or additional cluster nodes. | "Establishing the Cluster" on page 63 |

## ▼ How to Prepare for Cluster Software Installation

Before you begin to install software, make the following preparations.

**Steps** 1. **Ensure that the hardware and software that you choose for your cluster configuration are supported for this release of Sun Cluster software.**

Contact your Sun sales representative for the most current information about supported cluster configurations.

2. **Read the following manuals for information that can help you plan your cluster configuration and prepare your installation strategy.**

- *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* – Restrictions, bug workarounds, and other late-breaking information.

- *Sun Cluster 3.x Release Notes Supplement* – Post-release documentation about additional restrictions, bug workarounds, new features, and other late-breaking information. This document is regularly updated and published online at the following Web site.

  http://docs.sun.com

- *Sun Cluster Overview for Solaris OS* and *Sun Cluster Concepts Guide for Solaris OS* – Overviews of the Sun Cluster product.

- *Sun Cluster Software Installation Guide for Solaris OS* (this manual) – Planning guidelines and procedures for installing and configuring Solaris, Sun Cluster, and volume-manager software.

- *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* – Planning guidelines and procedures to install and configure data services.

3. **Have available all related documentation, including third-party documents.**

   The following is a partial list of products whose documentation you might need to reference during cluster installation:

   - Solaris OS
   - Solstice DiskSuite or Solaris Volume Manager software
   - Sun StorEdge QFS software
   - SPARC: VERITAS Volume Manager
   - SPARC: Sun Management Center
   - Third-party applications

4. **Plan your cluster configuration.**

   **Caution –** Plan your cluster installation completely. Identify requirements for all data services and third-party products **before** you begin Solaris and Sun Cluster software installation. Failure to do so might result in installation errors that require that you completely reinstall the Solaris and Sun Cluster software.

   For example, the Oracle Real Application Clusters Guard option of Oracle Real Application Clusters has special requirements for the hostnames that you use in the cluster. Another example with special requirements is Sun Cluster HA for SAP. You must accommodate these requirements before you install Sun Cluster software because you cannot change hostnames after you install Sun Cluster software.

   Also note that both Oracle Real Application Clusters and Sun Cluster HA for SAP are not supported for use in x86 based clusters.

   - Use the planning guidelines in Chapter 1 and in the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* to determine how to install and configure your cluster.
   - Fill out the cluster framework and data-services configuration worksheets that are referenced in the planning guidelines. Use your completed worksheets for reference during the installation and configuration tasks.

5. **Obtain all necessary patches for your cluster configuration.**

   See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

   a. **Copy the patches that are required for Sun Cluster into a single directory.**

      The directory must be on a file system that is accessible by all nodes. The default patch directory is /var/cluster/patches/.

      **Tip –** After you install Solaris software on a node, you can view the /etc/release file to see the exact version of Solaris software that is installed.

   b. **(Optional) If you are using SunPlex Installer, you can create a patch list file.**

If you specify a patch list file, SunPlex Installer only installs the patches that are listed in the patch list file. For information about creating a patch-list file, refer to the patchadd(1M) man page.

**c. Record the path to the patch directory.**

**Next Steps** If you want to use Cluster Control Panel software to connect from an administrative console to your cluster nodes, go to "How to Install Cluster Control Panel Software on an Administrative Console" on page 48.

Otherwise, choose the Solaris installation procedure to use.

- If you intend to install Sun Cluster software by using either the scinstall(1M) utility (text-based method) or SunPlex Installer (GUI-based method), go to "How to Install Solaris Software" on page 52 to first install Solaris software.

- If you intend to install Solaris and Sun Cluster software in the same operation (JumpStart method), go to "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 72.

## ▼ How to Install Cluster Control Panel Software on an Administrative Console

---

**Note –** You are not required to use an administrative console. If you do not use an administrative console, perform administrative tasks from one designated node in the cluster.

---

This procedure describes how to install the Cluster Control Panel (CCP) software on an administrative console. The CCP provides a single interface from which to start the cconsole(1M), ctelnet(1M), and crlogin(1M) tools. Each of these tools provides a multiple-window connection to a set of nodes, as well as a common window. You can use the common window to send input to all nodes at one time.

You can use any desktop machine that runs the Solaris 8 or Solaris 9 OS as an administrative console. In addition, you can also use the administrative console as a documentation server. If you are using Sun Cluster on a SPARC based system, you can use the administrative console as a Sun Management Center console or server as well. See Sun Management Center documentation for information about how to install Sun Management Center software. See the *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for additional information about how to install Sun Cluster documentation.

**Before You Begin** Ensure that a supported version of the Solaris OS and any Solaris patches are installed on the administrative console. All platforms require at least the End User Solaris Software Group.

**Steps** 1. **Become superuser on the administrative console.**

2. **Insert the Sun Cluster 2 of 2 CD-ROM in the CD-ROM drive of the administrative console.**

   If the volume management daemon vold(1M) is running and is configured to manage CD-ROM devices, the daemon automatically mounts the CD-ROM on the /cdrom/cdrom0/ directory.

3. **Change to the Solaris_*arch*/Product/sun_cluster/Solaris_*ver*/Packages/ directory, where *arch* is sparc or x86, and where *ver* is 8 for Solaris 8, 9 for Solaris 9, or 10 for Solaris 10 .**

   # **cd /cdrom/cdrom0/Solaris_*arch*/Product/sun_cluster/Solaris_*ver*/Packages/**

4. **Install the SUNWccon package.**

   # **pkgadd -d . SUNWccon**

5. **(Optional) Install the SUNWscman package.**

   # **pkgadd -d . SUNWscman**

   When you install the SUNWscman package on the administrative console, you can view Sun Cluster man pages from the administrative console before you install Sun Cluster software on the cluster nodes.

6. **(Optional) Install the Sun Cluster documentation packages.**

   ---

   **Note –** If you do not install the documentation on your administrative console, you can still view HTML or PDF documentation directly from the CD-ROM. Use a web browser to view the Solaris_*arch*/Product/sun_cluster/index.html file on the Sun Cluster 2 of 2 CD-ROM, where *arch* is sparc or x86.

   ---

   a. **Determine whether the SUNWsdocs package is already installed on the administrative console.**

   ```
   # pkginfo | grep SUNWsdocs
   application SUNWsdocs     Documentation Navigation for Solaris 9
   ```

   If the SUNWsdocs package is not yet installed, you must install it before you install the documentation packages.

   b. **Choose the Sun Cluster documentation packages to install.**

   The following documentation collections are available in both HTML and PDF format:

| Collection Title | HTML Package Name | PDF Package Name |
| --- | --- | --- |
| Sun Cluster 3.1 9/04 Software Collection for Solaris OS (SPARC Platform Edition) | SUNWscsdoc | SUNWpscsdoc |
| Sun Cluster 3.1 9/04 Software Collection for Solaris OS (x86 Platform Edition) | SUNWscxdoc | SUNWpscxdoc |
| Sun Cluster 3.x Hardware Collection for Solaris OS (SPARC Platform Edition) | SUNWschw | SUNWpschw |
| Sun Cluster 3.x Hardware Collection for Solaris OS (x86 Platform Edition) | SUNWscxhw | SUNWpscxhw |
| Sun Cluster 3.1 9/04 Reference Collection for Solaris OS | SUNWscref | SUNWpscref |

    **c. Install the `SUNWsdocs` package, if not already installed, and your choice of Sun Cluster documentation packages.**

---

**Note –** All documentation packages have a dependency on the SUNWsdocs package. The SUNWsdocs package must exist on the system before you can successfully install a documentation package on that system.

---

```
# pkgadd -d . SUNWsdocs pkg-list
```

**7. Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

```
# eject cdrom
```

**8. Create an `/etc/clusters` file on the administrative console.**

Add your cluster name and the physical node name of each cluster node to the file.

```
# vi /etc/clusters
```
*clustername node1 node2*

See the /opt/SUNWcluster/bin/clusters(4) man page for details.

**9. Create an `/etc/serialports` file.**

Add an entry for each node in the cluster to the file. Specify the physical node name, the hostname of the console-access device, and the port number. Examples of a console-access device are a terminal concentrator (TC), a System Service Processor (SSP), and a Sun Fire system controller.

```
# vi /etc/serialports
```
*node1 ca-dev-hostname port*
*node2 ca-dev-hostname port*

| | |
|---|---|
| *node1, node2* | Physical names of the cluster nodes |
| *ca-dev-hostname* | Hostname of the console-access device |
| *port* | Serial port number |

Note these special instructions to create an `/etc/serialports` file:

- For a Sun Fire 15000 system controller, use `telnet`(1) port number 23 for the serial port number of each entry.

- For all other console-access devices, use the `telnet` serial port number, not the physical port number. To determine the telnet serial port number, add 5000 to the physical port number. For example, if a physical port number is 6, the telnet serial port number is 5006.

- For Sun Enterprise 10000 servers, also see the `/opt/SUNWcluster/bin/serialports`(4) man page for details and special considerations.

10. **(Optional) For convenience, set the directory paths on the administrative console.**

   a. **Add the `/opt/SUNWcluster/bin/` directory to the `PATH`.**

   b. **Add the `/opt/SUNWcluster/man/` directory to the `MANPATH`.**

   c. **If you installed the `SUNWscman` package, also add the `/usr/cluster/man/` directory to the `MANPATH`.**

11. **Start the CCP utility.**

   `# /opt/SUNWcluster/bin/ccp &`

   Click the `cconsole`, `crlogin`, or `ctelnet` button in the CCP window to launch that tool. Alternately, you can start any of these tools directly. For example, to start `ctelnet`, type the following command:

   `# /opt/SUNWcluster/bin/ctelnet &`

   See the procedure "How to Remotely Log In to Sun Cluster" in "Beginning to Administer the Cluster" in *Sun Cluster System Administration Guide for Solaris OS* for additional information about how to use the CCP utility. Also see the `ccp`(1M) man page.

**Next Steps** Determine whether the Solaris OS is already installed to meet Sun Cluster software requirements.

- If the Solaris OS meets Sun Cluster requirements, go to "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES installer)" on page 59.

- If the Solaris OS does not meet Sun Cluster requirements, install, reconfigure, or reinstall the Solaris OS as needed. See "Planning the Solaris OS" on page 16 for information about Sun Cluster installation requirements for the Solaris OS. For installation procedures, go to "How to Install Solaris Software" on page 52.

## ▼ How to Install Solaris Software

Follow these procedures to install the Solaris OS on each node in the cluster or to install the Solaris OS on the master node that you will flash archive for a JumpStart installation. See "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 72 for more information about JumpStart installation of a cluster.

---

**Tip –** To speed installation, you can install the Solaris OS on each node at the same time.

---

If your nodes are already installed with the Solaris OS but do not meet Sun Cluster installation requirements, you might need to reinstall the Solaris software. Follow the steps in this procedure to ensure subsequent successful installation of Sun Cluster software. See "Planning the Solaris OS" on page 16 for information about required root-disk partitioning and other Sun Cluster installation requirements.

**Before You Begin** Perform the following tasks:

- Ensure that the hardware setup is complete and that connections are verified before you install Solaris software. See the *Sun Cluster Hardware Administration Collection* and your server and storage device documentation for details.

- Ensure that your cluster configuration planning is complete. See "How to Prepare for Cluster Software Installation" on page 46 for requirements and guidelines.

- Complete the "Local File System Layout Worksheet" on page 288.

- If you use a naming service, add address-to-name mappings for all public hostnames and logical addresses to any naming services that clients use for access to cluster services. See "IP Addresses" on page 22 for planning guidelines. See your Solaris system-administrator documentation for information about using Solaris naming services.

**Steps    1. If you are using a cluster administrative console, display a console screen for each node in the cluster.**

- **If Cluster Control Panel (CCP) software is installed and configured on your administrative console, use the `cconsole(1M)` utility to display the individual console screens.**

  Use the following command to start the `cconsole` utility:

  `# /opt/SUNWcluster/bin/cconsole` *clustername* `&`

  The `cconsole` utility also opens a master window from which you can send your input to all individual console windows at the same time.

- **If you do not use the `cconsole` utility, connect to the consoles of each node individually.**

**2. Install the Solaris OS as instructed in your Solaris installation documentation.**

---

**Note –** You must install all nodes in a cluster with the same version of the Solaris OS.

---

You can use any method that is normally used to install Solaris software. During Solaris software installation, perform the following steps:

**a. Install at least the End User Solaris Software Group.**

---

**Tip –** To avoid the need to manually install Solaris software packages, install the Entire Solaris Software Group Plus OEM Support.

---

See "Solaris Software Group Considerations" on page 17 for information about additional Solaris software requirements.

**b. Choose Manual Layout to set up the file systems.**

- **Create a file system of at least 512 Mbytes for use by the global-device subsystem.**

  If you intend to use SunPlex Installer to install Sun Cluster software, you must create the file system with a mount-point name of `/globaldevices`. The `/globaldevices` mount-point name is the default that is used by `scinstall`.

---

**Note –** Sun Cluster software requires a global-devices file system for installation to succeed.

---

- **Specify that slice 7 is at least 20 Mbytes in size.**

  If you intend to use SunPlex Installer to install Solstice DiskSuite software (Solaris 8) or configure Solaris Volume Manager software (Solaris 9 or Solaris 10), also make this file system mount on /sds.

  ---

  **Note –** If you intend to use SunPlex Installer to install Sun Cluster HA for NFS or Sun Cluster HA for Apache, SunPlex Installer must also install Solstice DiskSuite software (Solaris 8) or configure Solaris Volume Manager software (Solaris 9 or Solaris 10).

  ---

- **Create any other file-system partitions that you need, as described in "System Disk Partitions" on page 18.**

c. **For ease of administration, set the same root password on each node.**

3. **If you are adding a node to an existing cluster, prepare the cluster to accept the new node.**

   a. **On any active cluster member, start the scsetup(1M) utility.**

   ```
   # scsetup
   ```
   The Main Menu is displayed.

   b. **Choose the menu item, New nodes.**

   c. **Choose the menu item, Specify the name of a machine which may add itself.**

   d. **Follow the prompts to add the node's name to the list of recognized machines.**

   The scsetup utility prints the message Command completed successfully if the task is completed without error.

   e. **Quit the scsetup utility.**

   f. **From the active cluster node, display the names of all cluster file systems.**

   ```
   % mount | grep global | egrep -v node@ | awk '{print $1}'
   ```

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

   ```
   % mkdir -p mountpoint
   ```
   For example, if the mount command returned the file-system name /global/dg-schost-1, run mkdir -p /global/dg-schost-1 on the new node you are adding to the cluster.

4. **If you are adding a node and VxVM is installed on any node in the cluster, perform the following tasks.**

   a. **Ensure that the same `vxio` number is used on the VxVM-installed nodes.**

   ```
   # grep vxio /etc/name_to_major
   vxio NNN
   ```

   b. **Ensure that the `vxio` number is available for use on each of the nodes that do not have VxVM installed.**

   c. **If the `vxio` number is already in use on a node that does not have VxVM installed, change the `/etc/name_to_major` entry to use a different number.**

5. **If you installed the End User Solaris Software Group, use the `pkgadd` command to manually install any additional Solaris software packages that you might need.**

   The following Solaris packages are required to support some Sun Cluster functionality.

   ---

   **Note –** Install packages in the order in which they are listed in the following table.

   ---

   | Feature | Mandatory Solaris Software Packages |
   |---------|-------------------------------------|
   | RSMAPI, RSMRDT drivers, or SCI-PCI adapters (SPARC based clusters only) | *Solaris 8 or Solaris 9*: SUNWrsm SUNWrsmx SUNWrsmo SUNWrsmox |
   | | *Solaris 10*: SUNWrsm SUNWrsmo |
   | SunPlex Manager | SUNWapchr SUNWapchu |

   - **For the Solaris 8 or Solaris 9 OS, use the following command:**

     ```
     # pkgadd -d . packages
     ```

   - **For the Solaris 10 OS, use the following command:**

     ```
     # pkgadd -G -d . packages
     ```

     You must add these packages only to the global zone. The `-G` option adds packages to the current zone only. This option also specifies that the packages are *not* propagated to any existing non-global zone or to any non-global zone that is created later.

6. **Install any required Solaris OS patches and hardware-related firmware and patches, including those for storage-array support. Also download any needed firmware that is contained in the hardware patches.**

   See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

7. **x86: Set the default boot file to `kadb`.**

   ```
   # eeprom boot-file=kadb
   ```

   The setting of this value enables you to reboot the node if you are unable to access a login prompt.

8. **Update the `/etc/inet/hosts` file on each node with all IP addresses that are used in the cluster.**

   Perform this step regardless of whether you are using a naming service. See "IP Addresses" on page 22 for a listing of Sun Cluster components whose IP addresses you must add.

9. **If you will use `ce` adapters for the cluster interconnect, add the following entry to the `/etc/system` file.**

   ```
   set ce:ce_taskq_disable=1
   ```

   This entry becomes effective after the next system reboot.

10. **(Optional) On Sun Enterprise 10000 servers, configure the `/etc/system` file to use dynamic reconfiguration.**

    Add the following entry to the /etc/system file on each node of the cluster:

    ```
    set kernel_cage_enable=1
    ```

    This entry becomes effective after the next system reboot. See your server documentation for more information about dynamic reconfiguration.

**Next Steps**  If you intend to use Sun multipathing software, go to "SPARC: How to Install Sun Multipathing Software" on page 56.

If you intend to install VxFS, go to "SPARC: How to Install VERITAS File System Software" on page 59.

Otherwise, install the Sun Cluster software packages. Go to "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES `installer`)" on page 59.

**See Also**  See the *Sun Cluster System Administration Guide for Solaris OS* for procedures to perform dynamic reconfiguration tasks in a Sun Cluster configuration.

## ▼ SPARC: How to Install Sun Multipathing Software

Perform this procedure on each node of the cluster to install and configure Sun multipathing software for fiber channel (FC) storage. Multipathing software manages multiple I/O paths to the shared cluster storage.

- For the Solaris 8 or Solaris 9 OS, you install and configure Sun StorEdge Traffic Manager software.

■ For the Solaris 10 OS, you enable the Solaris multipathing feature, which is installed by default as part of the Solaris 10 software.

**Before You Begin**   Perform the following tasks:

■ Ensure that the Solaris OS is installed to support Sun Cluster software.

  If Solaris software is already installed on the node, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 52 for more information about installing Solaris software to meet Sun Cluster software requirements.

■ For the Solaris 8 or Solaris 9 OS, have available your software packages, patches, and documentation for Sun StorEdge Traffic Manager software and Sun StorEdge SAN Foundation software. See `http://www.sun.com/products-n-solutions/hardware/docs/` for links to documentation.

■ For the Solaris 10 OS, have available the *Solaris Fibre Channel Storage Configuration and Multipathing Administration Guide* at `http://docs.sun.com/source/819-0139/`.

**Steps**   **1. Become superuser.**

**2. For the Solaris 8 or Solaris 9 OS, install on each node Sun StorEdge Traffic Manager software and any necessary patches.**

  ■ For the procedure about how to install Sun StorEdge Traffic Manager software, see the *Sun StorEdge Traffic Manager Installation and Configuration Guide* at `http://www.sun.com/products-n-solutions/hardware/docs/`.

  ■ For a list of required patches for Sun StorEdge Traffic Manager software, see the *Sun StorEdge Traffic Manager Software Release Notes* at `http://www.sun.com/storage/san/`.

**3. Enable multipathing functionality.**

  ■ **For the Solaris 8 or 9 OS, change the value of the `mpxio-disable` parameter to `no`.**

    Modify this entry in the /kernel/drv/scsi_vhci.conf file on each node.

    ```
    set mpxio-disable=no
    ```

  ■ **For the Solaris 10 OS, issue the following command on each node:**

⚠ **Caution –** If Sun Cluster software is already installed, do not issue this command. Running the `stmsboot` command on an active cluster node might cause Solaris services to go into the maintenance state. Instead, follow instructions in the `stmsboot`(1M) man page for using the `stmsboot` command in a Sun Cluster environment.

```
# /usr/sbin/stmsboot -e
```

-e      Enables Solaris I/O multipathing

See the `stmsboot`(1M) man page for more information.

4. **For the Solaris 8 or Solaris 9 OS, determine whether your version of Sun StorEdge SAN Foundation software includes built-in support for your storage array.**

   If the software does not include built-in support for your storage array, edit the `/kernel/drv/scsi_vhci.conf` file on each node to include the necessary entries. For more information, see the release notes for your storage device.

5. **For the Solaris 8 or Solaris 9 OS, shut down each node and perform a reconfiguration boot.**

   The reconfiguration boot creates the new Solaris device files and links.

```
# shutdown -y -g0 -i0
ok boot -r
```

6. **After the reconfiguration reboot is finished on all nodes, perform any additional tasks that are necessary to complete the configuration of your storage array.**

   See installation instructions for your storage array in the *Sun Cluster Hardware Administration Collection* for details.

**Troubleshooting** If you installed Sun multipathing software after Sun Cluster software was installed on the cluster, DID mappings might require updating. Issue the following commands on each node of the cluster to regenerate the DID namespace.

```
# scdidadm -C
# scdidadm -r
(Solaris 8 or 9 only) # cfgadm -c configure
# scgdevs
```

See the scdidadm(1M), scgdevs(1M)man pages for more information.

**Next Steps** If you intend to install VxFS, go to .

Otherwise, install the Sun Cluster software packages. Go to "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES `installer`)" on page 59.

## ▼ SPARC: How to Install VERITAS File System Software

Perform this procedure on each node of the cluster.

**Steps**   1. **Follow the procedures in your VxFS installation documentation to install VxFS software on each node of the cluster.**

2. **Install any Sun Cluster patches that are required to support VxFS.**

   See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

3. **In the `/etc/system` file on each node, set the following values.**

   ```
   set rpcmod:svc_default_stksize=0x8000
   set lwp_default_stksize=0x6000
   ```

   These changes become effective at the next system reboot.

   - Sun Cluster software requires a minimum `rpcmod:svc_default_stksize` setting of `0x8000`. Because VxFS installation sets the value of the `rpcmod:svc_default_stksize` variable to `0x4000`, you must manually set the value to `0x8000` after VxFS installation is complete.

   - You must set the `lwp_default_stksize` variable in the `/etc/system` file to override the VxFS default value of `0x4000`.

**Next Steps**   Install the Sun Cluster software packages. Go to "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES `installer`)" on page 59.

## ▼ How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES `installer`)

Follow this procedure to use the Sun Java™ Enterprise System (Java ES) `installer` program to perform one or more of the following installation tasks:

- To install the Sun Cluster framework software packages on each node in the cluster.

- To install Sun Cluster framework software on the master node that you will flash archive for a JumpStart installation. See "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 72 for more information about JumpStart installation of a cluster.

- To install Sun Java System data services for the Solaris 8 or Solaris 9 OS from the Sun Cluster 2 of 2 CD-ROM.

---

**Note –** Do not use this procedure to install the following kinds of data service packages:

- Data services for the Solaris 10 OS from the Sun Cluster 2 of 2 CD-ROM - Instead, follow procedures in "How to Install Data-Service Software Packages (pkgadd)" on page 106.

- Data services from the Sun Cluster Agents CD - Instead, follow procedures in "How to Install Data-Service Software Packages (scinstall)" on page 108.

  For data services for the Solaris 8 or Solaris 9 OS from the Sun Cluster Agents CD, you can alternatively follow procedures in "How to Install Data-Service Software Packages (Web Start installer)" on page 111.

---

**Before You Begin**    Perform the following tasks:

- Ensure that the Solaris OS is installed to support Sun Cluster software.

  If Solaris software is already installed on the node, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 52 for more information about installing Solaris software to meet Sun Cluster software requirements.

- Have available the Sun Cluster 1 of 2 CD-ROM and the Sun Cluster 2 of 2 CD-ROM.

**Steps**    1. **(Optional) To use the `installer` program with a GUI, ensure that the display environment of the cluster node to install is set to display the GUI.**

```
% xhost +
% setenv DISPLAY nodename:0.0
```

2. **Become superuser on the cluster node to install.**

3. **Insert the Sun Cluster 1 of 2 CD-ROM in the CD-ROM drive.**

4. **Change to the directory of the CD-ROM where the `installer` program resides.**

```
# cd /cdrom/cdrom0/Solaris_arch/
```

In the Solaris_arch/ directory, *arch* is sparc or x86.

5. **Start the Java ES `installer` program.**

   ```
   # ./installer
   ```

6. **Follow instructions on the screen to install Sun Cluster framework software and data services on the node.**

   When prompted whether to configure Sun Cluster framework software, choose Configure Later.

   After installation is finished, you can view any available installation log. See the *Sun Java Enterprise System 2005Q5 Installation Guide* for additional information about using the Java ES `installer` program.

7. **Install additional packages if you intend to use any of the following features.**

   - Remote Shared Memory Application Programming Interface (RSMAPI)
   - SCI-PCI adapters for the interconnect transport
   - RSMRDT drivers

   ---

   **Note –** Use of the RSMRDT driver is restricted to clusters that run an Oracle9i release 2 SCI configuration with RSM enabled. Refer to Oracle9i release 2 user documentation for detailed installation and configuration instructions.

   ---

   a. **Determine which packages you must install.**

      The following table lists the Sun Cluster 3.1 8/05 packages that each feature requires, in the order in which you must install each group of packages. The Java ES `installer` program does not automatically install these packages.

      ---

      **Note –** Install packages in the order in which they are listed in the following table.

      ---

      | Feature | Additional Sun Cluster 3.1 8/05 Packages to Install |
      |---|---|
      | RSMAPI | `SUNWscrif` |
      | SCI-PCI adapters | ■ *Solaris 8 and 9*: `SUNWsci SUNWscid SUNWscidx` <br> ■ *Solaris 10*: `SUNWscir SUNWsci SUNWscidr SUNWscid` |
      | RSMRDT drivers | `SUNWscrdt` |

   b. **Insert the Sun Cluster 2 of 2 CD-ROM, if it is not already inserted in the CD-ROM drive.**

**c. Change to the Solaris_*arch*/Product/sun_cluster/Solaris_*ver*/Packages/ directory, where *arch* is sparc or x86, and where *ver* is 8 for Solaris 8, 9 for Solaris 9, or 10 for Solaris 10 .**

```
# cd /cdrom/cdrom0/Solaris_arch/Product/sun_cluster/Solaris_ver/Packages/
```

**d. Install the additional packages.**

```
# pkgadd -d . packages
```

**8. Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

```
# eject cdrom
```

**9. Ensure that the /usr/java/ directory is a symbolic link to the minimum or latest version of Java software.**

Sun Cluster software requires at least version 1.4.2_03 of Java software.

**a. Determine what directory the /usr/java/ directory is symbolically linked to.**

```
# ls -l /usr/java
lrwxrwxrwx   1 root   other    9 Apr 19 14:05 /usr/java -> /usr/j2se/
```

**b. Determine what version or versions of Java software are installed.**

The following are examples of commands that you can use to display the version of their related releases of Java software.

```
# /usr/j2se/bin/java -version
# /usr/java1.2/bin/java -version
# /usr/jdk/jdk1.5.0_01/bin/java -version
```

**c. If the /usr/java/ directory is not symbolically linked to a supported version of Java software, recreate the symbolic link to link to a supported version of Java software.**

The following example shows the creation of a symbolic link to the /usr/j2se/ directory, which contains Java 1.4.2_03 software.

```
# rm /usr/java
# ln -s /usr/j2se /usr/java
```

**Next Steps**   If you want to install Sun StorEdge QFS file system software, follow the procedures for initial installation in the *Sun StorEdge QFS and Sun StorEdge SAM-FS Software Installation and Configuration Guide*.

Otherwise, to set up the root user environment, go to "How to Set Up the Root Environment" on page 63.

## ▼ How to Set Up the Root Environment

---

**Note –** In a Sun Cluster configuration, user initialization files for the various shells must verify that they are run from an interactive shell. The files must verify this before they attempt to output to the terminal. Otherwise, unexpected behavior or interference with data services might occur. See "Customizing a User's Work Environment" in *System Administration Guide, Volume 1* (Solaris 8) or in *System Administration Guide: Basic Administration* (Solaris 9 or Solaris 10) for more information.

---

Perform this procedure on each node in the cluster.

**Steps**   **1. Become superuser on a cluster node.**

**2. Modify `PATH` and `MANPATH` entries in the `.cshrc` or `.profile` file.**

   **a. Set the `PATH` to include `/usr/sbin/` and `/usr/cluster/bin/`.**

   **b. Set the `MANPATH` to include `/usr/cluster/man/`.**

   See your volume manager documentation and other application documentation for additional file paths to set.

**3. (Optional) For ease of administration, set the same root password on each node, if you have not already done so.**

**Next Steps**   Configure Sun Cluster software on the cluster nodes. Go to "Establishing the Cluster" on page 63.

---

## Establishing the Cluster

This section provides information and procedures to establish a new cluster or to add a node to an existing cluster. Before you start to perform these tasks, ensure that you installed software packages for the Solaris OS, Sun Cluster framework, and other products as described in "Installing the Software" on page 45.

The following task map lists the tasks to perform. Complete the procedures in the order that is indicated.

**TABLE 2–2** Task Map: Establish the Cluster

| Method | Instructions |
|---|---|
| 1. Use one of the following methods to establish a new cluster or add a node to an existing cluster: | |
| ■ (*New clusters only*) Use the scinstall utility to establish the cluster. | "How to Configure Sun Cluster Software on All Nodes (scinstall)" on page 65 |
| ■ (*New clusters or added nodes*) Set up a JumpStart installation server. Then create a flash archive of the installed system. Finally, use the scinstall JumpStart option to install the flash archive on each node and establish the cluster. | "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 72 |
| ■ (*New multiple-node clusters only*)Use SunPlex Installer to establish the cluster. Optionally, also configure Solstice DiskSuite or Solaris Volume Manager disk sets, scalable Sun Cluster HA for Apache data service, and Sun Cluster HA for NFS data service. | "Using SunPlex Installer to Configure Sun Cluster Software" on page 86 <br><br> "How to Configure Sun Cluster Software (SunPlex Installer)" on page 89 |
| ■ (*Added nodes only*) Configure Sun Cluster software on the new node by using the scinstall utility. | "How to Configure Sun Cluster Software on Additional Cluster Nodes (scinstall)" on page 96 |
| 2. (*Oracle Real Application Clusters only*) If you added a node to a two-node cluster that runs Sun Cluster Support for Oracle Real Application Clusters and that uses a shared SCSI disk as the quorum device, update the SCSI reservations. | "How to Update SCSI Reservations After Adding a Node" on page 104 |
| 3. Install data-service software packages. | "How to Install Data-Service Software Packages (pkgadd)" on page 106 <br><br> "How to Install Data-Service Software Packages (scinstall)" on page 108 <br><br> "How to Install Data-Service Software Packages (Web Start installer)" on page 111 |
| 4. Assign quorum votes and remove the cluster from installation mode, if this operation was not already performed. | "How to Configure Quorum Devices" on page 114 |
| 5. Validate the quorum configuration. | "How to Verify the Quorum Configuration and Installation Mode" on page 117 |
| 6. Configure the cluster. | "Configuring the Cluster" on page 118 |

▼ **How to Configure Sun Cluster Software on All Nodes (`scinstall`)**

Perform this procedure from one node of the cluster to configure Sun Cluster software on all nodes of the cluster.

**Before You Begin**   Perform the following tasks:

- Ensure that the Solaris OS is installed to support Sun Cluster software.

  If Solaris software is already installed on the node, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 52 for more information about installing Solaris software to meet Sun Cluster software requirements.

- Ensure that Sun Cluster software packages are installed on the node. See "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES installer)" on page 59.

- Determine which mode of the `scinstall` utility you will use, Typical or Custom. For the Typical installation of Sun Cluster software, `scinstall` automatically specifies the following configuration defaults.

| Component | Default Value |
|-----------|---------------|
| Private-network address | `172.16.0.0` |
| Private-network netmask | `255.255.0.0` |
| Cluster-transport junctions | `switch1` and `switch2` |
| Global-devices file-system name | `/globaldevices` |
| Installation security (DES) | Limited |
| Solaris and Sun Cluster patch directory | `/var/cluster/patches/` |

- Complete one of the following cluster configuration worksheets, depending on whether you run the `scinstall` utility in Typical mode or Custom mode.

  - **Typical Mode** - If you will use Typical mode and accept all defaults, complete the following worksheet.

| Component | Description/Example | Answer |
|-----------|---------------------|--------|
| Cluster Name | What is the name of the cluster that you want to establish? | |
| Cluster Nodes | What are the names of the other cluster nodes planned for the initial cluster configuration? | |

| Component | Description/Example | Answer | |
|---|---|---|---|
| Cluster-Transport Adapters and Cables | What are the names of the two cluster-transport adapters that attach the node to the private interconnect? | *First* | *Second* |
| | Will this be a dedicated cluster transport adapter? | Yes \| No | Yes \| No |
| | If no, what is the VLAN ID for this adapter? | | |
| Quorum Configuration<br><br>*(two-node cluster only)* | Do you want to disable automatic quorum device selection? *(Answer Yes if any shared storage is not qualified to be a quorum device or if you want to configure a Network Appliance NAS device as a quorum device.)* | Yes \| No | |
| Check | Do you want to interrupt installation for sccheck errors? *(sccheck verifies that preconfiguration requirements are met)* | Yes \| No | |

- **Custom Mode** - If you will use Custom mode and customize the configuration data, complete the following worksheet.

| Component | Description/Example | Answer | |
|---|---|---|---|
| Cluster Name | What is the name of the cluster that you want to establish? | | |
| Cluster Nodes | What are the names of the other cluster nodes planned for the initial cluster configuration? | | |
| DES Authentication<br><br>*(multiple-node cluster only)* | Do you need to use DES authentication? | No \| Yes | |
| Network Address for the Cluster Transport<br><br>*(multiple-node cluster only)* | Do you want to accept the default network address (`172.16.0.0`)? | Yes \| No | |
| | If no, supply your own network address: | \_\_\_\_ . \_\_\_\_.0.0 | |
| | Do you want to accept the default netmask (`255.255.0.0`)? | Yes \| No | |
| | If no, supply your own netmask: | `255.255.___.___` | |
| Point-to-Point Cables<br><br>*(multiple-node cluster only)* | If this is a two-node cluster, does this cluster use transport junctions? | Yes \| No | |
| Cluster-Transport Junctions<br><br>*(multiple-node cluster only)* | If used, what are the names of the two transport junctions?<br>    Defaults: `switch1` and `switch2` | *First* | *Second* |

| Component | Description/Example | Answer | |
|---|---|---|---|
| Cluster-Transport Adapters and Cables *(multiple-node cluster only)* | Node name (*the node from which you run scinstall*): | | |
| | Transport adapters: | *First* | *Second* |
| | Will this be a dedicated cluster transport adapter? | Yes \| No | Yes \| No |
| | If no, what is the VLAN ID for this adapter? | | |
| | Where does each transport adapter connect to (*a transport junction or another adapter*)?<br>    Junction defaults: `switch1` and `switch2` | | |
| | For transport junctions, do you want to use the default port name? | Yes \| No | Yes \| No |
| | If no, what is the name of the port that you want to use? | | |
| | Do you want to use autodiscovery to list the available adapters for the other nodes? | Yes \| No | |
| | If no, supply the following information for each additional node: | | |
| *Specify for each additional node* *(multiple-node cluster only)* | Node name: | | |
| | Transport adapters: | *First* | *Second* |
| | Will this be a dedicated cluster transport adapter? | Yes \| No | Yes \| No |
| | If no, what is the VLAN ID for this adapter? | | |
| | Where does each transport adapter connect to (*a transport junction or another adapter*)?<br>    Defaults: `switch1` and `switch2` | | |
| | For transport junctions, do you want to use the default port name? | Yes \| No | Yes \| No |
| | If no, what is the name of the port that you want to use? | | |
| Software Patch Installation | Do you want `scinstall` to install patches for you? | Yes \| No | |
| | If yes, what is the name of the patch directory? | | |
| | Do you want to use a patch list? | Yes \| No | |
| Quorum Configuration *(two-node cluster only)* | Do you want to disable automatic quorum device selection?<br>*(Answer Yes if any shared storage is not qualified to be a quorum device or if you want to configure a Network Appliance NAS device as a quorum device.)* | Yes \| No | Yes \| No |

| Component | Description/Example | Answer |
|---|---|---|
| Global-Devices File System<br><br>(*specify for each node*) | Do you want to use the default name of the global-devices file system (/globaldevices)? | Yes  \|  No |
| | If no, do you want to use an already-existing file system? | Yes  \|  No |
| | What is the name of the file system that you want to use? | |
| Check<br><br>(*multiple-node cluster only*) | Do you want to interrupt installation for sccheck errors? *(sccheck verifies that preconfiguration requirements are met)* | Yes  \|  No |
| (*single-node cluster only*) | Do you want to run the sccheck utility to validate the cluster? | Yes  \|  No |
| Automatic Reboot<br><br>(*single-node cluster only*) | Do you want scinstall to automatically reboot the node after installation? | Yes  \|  No |

Follow these guidelines to use the interactive scinstall utility in this procedure:

- Interactive scinstall enables you to type ahead. Therefore, do not press the Return key more than once if the next menu screen does not appear immediately.

- Unless otherwise noted, you can press Control-D to return to either the start of a series of related questions or to the Main Menu.

- Default answers or answers to previous sessions are displayed in brackets ([ ]) at the end of a question. Press Return to enter the response that is in brackets without typing it.

**Steps**  **1. If you disabled remote configuration during Sun Cluster software installation, re-enable remote configuration.**

Enable remote shell (rsh(1M)) or secure shell (ssh(1)) access for superuser to all cluster nodes.

**2. (Optional) To use the scinstall(1M) utility to install patches, download patches to a patch directory.**

- **If you use Typical mode to install the cluster, use a directory named either /var/cluster/patches/ or /var/patches/ to contain the patches to install.**

   In Typical mode, the scinstall command checks both those directories for patches.

   - If neither of those directories exist, then no patches are added.
   - If both directories exist, then only the patches in the /var/cluster/patches/ directory are added.

- **If you use Custom mode to install the cluster, you specify the path to the patch directory. Specifying the path ensures that you do not have to use the**

patch directories that `scinstall` checks for in Typical mode.

You can include a patch-list file in the patch directory. The default patch-list file name is `patchlist`. For information about creating a patch-list file, refer to the `patchadd`(1M) man page.

3. **Become superuser on the cluster node from which you intend to configure the cluster.**

4. **Start the `scinstall` utility.**

   ```
   # /usr/cluster/bin/scinstall
   ```

5. **From the Main Menu, choose the menu item, Install a cluster or cluster node.**

   ```
    *** Main Menu ***

       Please select from one of the following (*) options:

         * 1) Install a cluster or cluster node
           2) Configure a cluster to be JumpStarted from this install server
           3) Add support for new data services to this cluster node
           4) Upgrade this cluster node
         * 5) Print release information for this cluster node
         * ?) Help with menu options
         * q) Quit

       Option:  1
   ```

6. **From the Install Menu, choose the menu item, Install all nodes of a new cluster.**

7. **From the Type of Installation menu, choose either Typical or Custom.**

8. **Follow the menu prompts to supply your answers from the configuration planning worksheet.**

   The `scinstall` utility installs and configures all cluster nodes and reboots the cluster. The cluster is established when all nodes have successfully booted into the cluster. Sun Cluster installation output is logged in a `/var/cluster/logs/install/scinstall.log.`*N* file.

9. **For the Solaris 10 OS, verify on each node that multi-user services for the Service Management Facility (SMF) are online.**

   If services are not yet online for a node, wait until the state becomes online before you proceed to the next step.

   ```
   # svcs multi-user-server
   STATE          STIME    FMRI
   online         17:52:55 svc:/milestone/multi-user-server:default
   ```

10. **From one node, verify that all nodes have joined the cluster.**

   Run the `scstat`(1M) command to display a list of the cluster nodes. You do not need to be logged in as superuser to run this command.

   ```
   % scstat -n
   ```

Output resembles the following.

```
-- Cluster Nodes --
                                Node name       Status
                                ---------       ------
    Cluster node:               phys-schost-1   Online
    Cluster node:               phys-schost-2   Online
```

**11. Install any necessary patches to support Sun Cluster software, if you have not already done so.**

**12. To re-enable the loopback file system (LOFS), delete the following entry from the `/etc/system` file on each node of the cluster.**

```
exclude:lofs
```

The re-enabling of LOFS becomes effective after the next system reboot.

---

**Note –** You cannot have LOFS enabled if you use Sun Cluster HA for NFS on a highly available local file system *and* have `automountd` running. LOFS can cause switchover problems for Sun Cluster HA for NFS. If you enable LOFS and later choose to add Sun Cluster HA for NFS on a highly available local file system, you must do one of the following:

- Restore the `exclude:lofs` entry to the `/etc/system` file on each node of the cluster and reboot each node. This change disables LOFS.

- Disable the `automountd` daemon.

- Exclude from the automounter map all files that are part of the highly available local file system that is exported by Sun Cluster HA for NFS. This choice enables you to keep both LOFS and the `automountd` daemon enabled.

---

See "Types of File Systems" in *System Administration Guide, Volume 1* (Solaris 8) or "The Loopback File System" in *System Administration Guide: Devices and File Systems* (Solaris 9 or Solaris 10) for more information about loopback file systems.

**Example 2–1**   Configuring Sun Cluster Software on All Nodes

The following example shows the `scinstall` progress messages that are logged as `scinstall` completes configuration tasks on the two-node cluster, `schost`. The cluster is installed from `phys-schost-1` by using the `scinstall` Typical mode. The other cluster node is `phys-schost-2`. The adapter names are `qfe2` and `qfe3`. The automatic selection of a quorum device is enabled.

```
  Installation and Configuration

    Log file - /var/cluster/logs/install/scinstall.log.24747

    Testing for "/globaldevices" on "phys-schost-1" ... done
    Testing for "/globaldevices" on "phys-schost-2" ... done
```

```
Checking installation status ... done

The Sun Cluster software is already installed on "phys-schost-1".
The Sun Cluster software is already installed on "phys-schost-2".
Starting discovery of the cluster transport configuration.

The following connections were discovered:

    phys-schost-1:qfe2  switch1  phys-schost-2:qfe2
    phys-schost-1:qfe3  switch2  phys-schost-2:qfe3

Completed discovery of the cluster transport configuration.

Started sccheck on "phys-schost-1".
Started sccheck on "phys-schost-2".

sccheck completed with no errors or warnings for "phys-schost-1".
sccheck completed with no errors or warnings for "phys-schost-2".

Removing the downloaded files ... done

Configuring "phys-schost-2" ... done
Rebooting "phys-schost-2" ... done

Configuring "phys-schost-1" ... done
Rebooting "phys-schost-1" ...

Log file - /var/cluster/logs/install/scinstall.log.24747

Rebooting ...
```

**Next Steps**  If you intend to install data services, go to the appropriate procedure for the data service that you want to install and for your version of the Solaris OS:

| | Sun Cluster 2 of 2 CD-ROM (Sun Java System data services) | | Sun Cluster Agents CD (All other data services) | |
|---|---|---|---|---|
| Procedure | Solaris 8 or 9 | Solaris 10 | Solaris 8 or 9 | Solaris 10 |
| "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES `installer`)" on page 59 | X | | | |
| "How to Install Data-Service Software Packages (`pkgadd`)" on page 106 | | X | | |
| "How to Install Data-Service Software Packages (`scinstall`)" on page 108 | | | X | X |
| "How to Install Data-Service Software Packages (Web Start `installer`)" on page 111 | | | X | |

Otherwise, go to the next appropriate procedure:

- If you installed a single-node cluster, cluster establishment is complete. Go to "Configuring the Cluster" on page 118 to install volume management software and configure the cluster.

- If you installed a multiple-node cluster and chose automatic quorum configuration, postinstallation setup is complete. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

- If you installed a multiple-node cluster and declined automatic quorum configuration, perform postinstallation setup. Go to "How to Configure Quorum Devices" on page 114.

**Troubleshooting** You cannot change the private-network address and netmask after scinstall processing is finished. If you need to use a different private-network address or netmask and the node is still in installation mode, follow the procedures in "How to Uninstall Sun Cluster Software to Correct Installation Problems" on page 136. Then perform the procedures in "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES installer)" on page 59 and then perform this procedure to reinstall the software and configure the node with the correct information.

## ▼ How to Install Solaris and Sun Cluster Software (JumpStart)

This procedure describes how to set up and use the scinstall(1M) custom JumpStart installation method. This method installs both Solaris OS and Sun Cluster software on all cluster nodes in the same operation and establishes the cluster. You can also use this procedure to add new nodes to an existing cluster.

**Before You Begin** Perform the following tasks:

- Ensure that the hardware setup is complete and connections are verified before you install Solaris software. See the *Sun Cluster Hardware Administration Collection* and your server and storage device documentation for details on how to set up the hardware.

- Determine the Ethernet address of each cluster node.

- If you use a naming service, ensure that the following information is added to any naming services that clients use to access cluster services. See "IP Addresses" on page 22 for planning guidelines. See your Solaris system-administrator documentation for information about using Solaris naming services.

  - Address-to-name mappings for all public hostnames and logical addresses
  - The IP address and hostname of the JumpStart server

- Ensure that your cluster configuration planning is complete. See "How to Prepare for Cluster Software Installation" on page 46 for requirements and guidelines.

- On the server from which you will create the flash archive, ensure that all Solaris OS software, patches, and firmware that is necessary to support Sun Cluster software is installed.

  If Solaris software is already installed on the server, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 52 for more information about installing Solaris software to meet Sun Cluster software requirements.

- Ensure that Sun Cluster software packages and patches are installed on the server from which you will create the flash archive. See "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES installer)" on page 59.

- Determine which mode of the scinstall utility you will use, Typical or Custom. For the Typical installation of Sun Cluster software, scinstall automatically specifies the following configuration defaults.

| Component | Default Value |
|---|---|
| Private-network address | 172.16.0.0 |
| Private-network netmask | 255.255.0.0 |
| Cluster-transport junctions | switch1 and switch2 |
| Global-devices file-system name | /globaldevices |
| Installation security (DES) | Limited |
| Solaris and Sun Cluster patch directory | /var/cluster/patches |

- Complete the appropriate planning worksheet. See "Planning the Sun Cluster Environment" on page 21 for planning guidelines.

  - **Typical Mode** - If you will use Typical mode and accept all defaults, complete the following worksheet.

| Component | Description/Example | Answer |
|---|---|---|
| JumpStart Directory | What is the name of the JumpStart directory to use? | |
| Cluster Name | What is the name of the cluster that you want to establish? | |
| Cluster Nodes | What are the names of the cluster nodes that are planned for the initial cluster configuration? | |

| Component | Description/Example | Answer | |
|---|---|---|---|
| Cluster-Transport Adapters and Cables | First node name: | | |
| | *First* / *Second* — Transport adapters: | *First* | *Second* |
| | Will this be a dedicated cluster transport adapter? | Yes \| No | Yes \| No |
| | If no, what is the VLAN ID for this adapter? | | |
| *Specify for each additional node* | Node name: | | |
| | Transport adapters: | *First* | *Second* |
| Quorum Configuration <br> *(two-node cluster only)* | Do you want to disable automatic quorum device selection? *(Answer Yes if any shared storage is not qualified to be a quorum device or if you want to configure a Network Appliance NAS device as a quorum device.)* | Yes \| No | Yes \| No |

- **Custom Mode** - If you will use Custom mode and customize the configuration data, complete the following worksheet.

| Component | Description/Example | Answer |
|---|---|---|
| JumpStart Directory | What is the name of the JumpStart directory to use? | |
| Cluster Name | What is the name of the cluster that you want to establish? | |
| Cluster Nodes | What are the names of the cluster nodes that are planned for the initial cluster configuration? | |
| DES Authentication <br> *(multiple-node cluster only)* | Do you need to use DES authentication? | No \| Yes |
| Network Address for the Cluster Transport <br> *(multiple-node cluster only)* | Do you want to accept the default network address (`172.16.0.0`)? | Yes \| No |
| | If no, supply your own network address: | _____ . _____ . 0 . 0 |
| | Do you want to accept the default netmask (`255.255.0.0`)? | Yes \| No |
| | If no, supply your own netmask: | 255.255.___ . ___ |
| Point-to-Point Cables <br> *(two-node cluster only)* | Does this cluster use transport junctions? | Yes \| No |

| Component | Description/Example | Answer | |
|---|---|---|---|
| Cluster-Transport Junctions<br><br>*(multiple-node cluster only)* | If used, what are the names of the two transport junctions?<br>    Defaults: `switch1` and `switch2` | *First* | *Second* |
| Cluster-Transport Adapters and Cables<br><br>*(multiple-node cluster only)* | First node name: | | |
| | Transport adapters: | *First* | *Second* |
| | Will this be a dedicated cluster transport adapter? | Yes \| No | Yes \| No |
| | If no, what is the VLAN ID for this adapter? | | |
| | Where does each transport adapter connect to (*a transport junction or another adapter*)?<br>    Junction defaults: `switch1` and `switch2` | | |
| | For transport junctions, do you want to use the default port name? | Yes \| No | Yes \| No |
| | If no, what is the name of the port that you want to use? | | |
| *Specify for each additional node*<br><br>*(multiple-node cluster only)* | Node name: | | |
| | Transport adapters: | *First* | *Second* |
| | Where does each transport adapter connect to (*a transport junction or another adapter*)?<br>    Junction defaults: `switch1` and `switch2` | | |
| | For transport junctions, do you want to use the default port name? | Yes \| No | Yes \| No |
| | If no, what is the name of the port that you want to use? | | |
| Global-Devices File System<br><br>*Specify for each node* | Do you want to use the default name of the global-devices file system (`/globaldevices`)? | Yes \| No | |
| | If no, do you want to use an already-existing file system? | Yes \| No | |
| | If no, do you want to create a new file system on an unused partition? | Yes \| No | |
| | What is the name of the file system? | | |
| Software Patch Installation | Do you want `scinstall` to install patches for you? | Yes \| No | |
| | If yes, what is the name of the patch directory? | | |
| | Do you want to use a patch list? | Yes \| No | |

| Component | Description/Example | Answer | |
|---|---|---|---|
| Quorum Configuration *(two-node cluster only)* | Do you want to disable automatic quorum device selection? *(Answer Yes if any shared storage is not qualified to be a quorum device or if you want to configure a Network Appliance NAS device as a quorum device.)* | Yes \| No | Yes \| No |

Follow these guidelines to use the interactive `scinstall` utility in this procedure:

- Interactive `scinstall` enables you to type ahead. Therefore, do not press the Return key more than once if the next menu screen does not appear immediately.

- Unless otherwise noted, you can press Control-D to return to either the start of a series of related questions or to the Main Menu.

- Default answers or answers to previous sessions are displayed in brackets ([ ]) at the end of a question. Press Return to enter the response that is in brackets without typing it.

**Steps 1. Set up your JumpStart installation server.**

- Follow the appropriate instructions for your software platform.

| Solaris OS Platform | Instructions |
|---|---|
| SPARC | See one of the following manuals for instructions about how to set up a JumpStart installation server:<br>- "Creating a Profile Server for Networked Systems" in *Solaris 8 Advanced Installation Guide*<br>- "Creating a Profile Server for Networked Systems" in *Solaris 9 9/04 Installation Guide*<br>- "Creating a Profile Server for Networked Systems" in *Solaris 10 Installation Guide: Custom JumpStart and Advanced Installations*<br><br>See also the `setup_install_server`(1M) and `add_install_client`(1M) man pages. |
| x86 | See "Solaris 9 Software Installation From a PXE Server" in *Sun Fire V60x and Sun Fire V65x Server Solaris Operating Environment Installation Guide* for instructions about how to set up a JumpStart Dynamic Host Configuration Protocol (DHCP) server and a Solaris network for Preboot Execution Environment (PXE) installations. |

- Ensure that the JumpStart installation server meets the following requirements.

  - The installation server is on the same subnet as the cluster nodes, or on the Solaris boot server for the subnet that the cluster nodes use.

  - The installation server is not itself a cluster node.

  - The installation server installs a release of the Solaris OS that is supported by the Sun Cluster software.

- A custom JumpStart directory exists for JumpStart installation of Sun Cluster software. This *jumpstart-dir* directory must contain a copy of the check(1M) utility. The directory must also be NFS exported for reading by the JumpStart installation server.

- Each new cluster node is configured as a custom JumpStart installation client that uses the custom JumpStart directory that you set up for Sun Cluster installation.

2. **If you are installing a new node to an existing cluster, add the node to the list of authorized cluster nodes.**

   a. **Switch to another cluster node that is active and start the scsetup(1M) utility.**

   b. **Use the scsetup utility to add the new node's name to the list of authorized cluster nodes.**

   For more information, see "How to Add a Node to the Authorized Node List" in *Sun Cluster System Administration Guide for Solaris OS*.

3. **On a cluster node or another machine of the same server platform, install the Solaris OS, if you have not already done so.**

   Follow procedures in "How to Install Solaris Software" on page 52.

4. **On the installed system, install Sun Cluster software, if you have not done so already.**

   Follow procedures in "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES installer)" on page 59.

5. **Enable the common agent container daemon to start automatically during system boots.**

   # **cacaoadm enable**

6. **On the installed system, install any necessary patches to support Sun Cluster software.**

7. **On the installed system, update the /etc/inet/hosts file with all IP addresses that are used in the cluster.**

   Perform this step regardless of whether you are using a naming service. See "IP Addresses" on page 22 for a listing of Sun Cluster components whose IP addresses you must add.

8. **For Solaris 10, on the installed system, update the /etc/inet/ipnodes file with all IP addresses that are used in the cluster.**

   Perform this step regardless of whether you are using a naming service.

9. **Create the flash archive of the installed system.**

   # **flarcreate -n** *name archive*

| -n name | Name to give the flash archive. |
| --- | --- |
| *archive* | File name to give the flash archive, with the full path. By convention, the file name ends in `.flar`. |

Follow procedures in one of the following manuals:

- Chapter 18, "Creating Web Start Flash Archives," in *Solaris 8 Advanced Installation Guide*
- Chapter 21, "Creating Solaris Flash Archives (Tasks)," in *Solaris 9 9/04 Installation Guide*
- Chapter 3, "Creating Solaris Flash Archives (Tasks)," in *Solaris 10 Installation Guide: Solaris Flash Archives (Creation and Installation)*

**10. Ensure that the flash archive is NFS exported for reading by the JumpStart installation server.**

See "Solaris NFS Environment" in *System Administration Guide, Volume 3* (Solaris 8) or "Managing Network File Systems (Overview)," in *System Administration Guide: Network Services* (Solaris 9 or Solaris 10) for more information about automatic file sharing.

See also the share(1M) and dfstab(4) man pages.

**11. From the JumpStart installation server, start the `scinstall`(1M) utility.**

The path /export/suncluster/sc31/ is used here as an example of the installation directory that you created. In the CD-ROM path, replace *arch* with sparc or x86 and replace *ver* with 8 for Solaris 8, 9 for Solaris 9, or 10 for Solaris 10.

```
# cd /export/suncluster/sc31/Solaris_arch/Product/sun_cluster/ \
Solaris_ver/Tools/
# ./scinstall
```

**12. From the Main Menu, choose the menu item, Configure a cluster to be JumpStarted from this installation server.**

This option is used to configure custom JumpStart finish scripts. JumpStart uses these finish scripts to install the Sun Cluster software.

```
 *** Main Menu ***

    Please select from one of the following (*) options:

      * 1) Install a cluster or cluster node
      * 2) Configure a cluster to be JumpStarted from this install server
        3) Add support for new data services to this cluster node
        4) Upgrade this cluster node
      * 5) Print release information for this cluster node

      * ?) Help with menu options
      * q) Quit
```

```
Option:  2
```

13. **Follow the menu prompts to supply your answers from the configuration planning worksheet.**

The `scinstall` command stores your configuration information and copies the `autoscinstall.class` default `class` file in the *jumpstart-dir*`/autoscinstall.d/3.1/` directory. This file is similar to the following example.

```
install_type    initial_install
system_type     standalone
partitioning    explicit
filesys         rootdisk.s0 free /
filesys         rootdisk.s1 750  swap
filesys         rootdisk.s3 512  /globaldevices
filesys         rootdisk.s7 20
cluster         SUNWCuser       add
package         SUNWman         add
```

14. **Make adjustments to the `autoscinstall.class` file to configure JumpStart to install the flash archive.**

  a. **Modify entries as necessary to match configuration choices you made when you installed the Solaris OS on the flash archive machine or when you ran the `scinstall` utility.**

  For example, if you assigned slice 4 for the global-devices file system and specified to `scinstall` that the file-system name is `/gdevs`, you would change the `/globaldevices` entry of the `autoscinstall.class` file to the following:

```
filesys         rootdisk.s4 512  /gdevs
```

  b. **Change the following entries in the `autoscinstall.class` file.**

| Existing Entry to Replace | | New Entry to Add | |
|---|---|---|---|
| `install_type` | `initial_install` | `install_type` | `flash_install` |
| `system_type` | `standalone` | `archive_location` | *retrieval_type location* |

  See "archive_location Keyword" in *Solaris 8 Advanced Installation Guide*, *Solaris 9 9/04 Installation Guide*, or *Solaris 10 Installation Guide: Custom JumpStart and Advanced Installations* for information about valid values for *retrieval_type* and *location* when used with the `archive_location` keyword.

  c. **Remove all entries that would install a specific package, such as the following entries.**

```
cluster         SUNWCuser       add
package         SUNWman         add
```

**15. Set up Solaris patch directories, if you did not already install the patches on the flash-archived system.**

---

**Note –** If you specified a patch directory to the `scinstall` utility, patches that are located in Solaris patch directories are not installed.

---

a. **Create** *jumpstart-dir*`/autoscinstall.d/nodes/`*node*`/patches/` **directories that are NFS exported for reading by the JumpStart installation server.**

Create one directory for each node in the cluster, where *node* is the name of a cluster node. Alternately, use this naming convention to create symbolic links to a shared patch directory.

```
# mkdir jumpstart-dir/autoscinstall.d/nodes/node/patches/
```

b. **Place copies of any Solaris patches into each of these directories.**

c. **Place copies of any hardware-related patches that you must install after Solaris software is installed into each of these directories.**

**16. If you are using a cluster administrative console, display a console screen for each node in the cluster.**

■ **If Cluster Control Panel (CCP) software is installed and configured on your administrative console, use the `cconsole`(1M) utility to display the individual console screens.**

Use the following command to start the `cconsole` utility:

```
# /opt/SUNWcluster/bin/cconsole clustername &
```

The `cconsole` utility also opens a master window from which you can send your input to all individual console windows at the same time.

■ **If you do not use the `cconsole` utility, connect to the consoles of each node individually.**

**17. Shut down each node.**

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

**18. Boot each node to start the JumpStart installation.**

■ **On SPARC based systems, do the following:**

```
ok boot net - install
```

---

**Note –** Surround the dash (-) in the command with a space on each side.

---

■ **On x86 based systems, do the following:**

a. **When the BIOS information screen appears, press the Esc key.**

The Select Boot Device screen appears.

b. **On the Select Boot Device screen, choose the listed IBA that is connected to the same network as the JumpStart PXE installation server.**

The lowest number to the right of the IBA boot choices corresponds to the lower Ethernet port number. The higher number to the right of the IBA boot choices corresponds to the higher Ethernet port number.

The node reboots and the Device Configuration Assistant appears.

c. **On the Boot Solaris screen, choose Net.**

d. **At the following prompt, choose Custom JumpStart and press Enter:**

```
Select the type of installation you want to perform:

        1 Solaris Interactive
        2 Custom JumpStart

Enter the number of your choice followed by the <ENTER> key.

If you enter anything else, or if you wait for 30 seconds,
an interactive installation will be started.
```

e. **When prompted, answer the questions and follow the instructions on the screen.**

JumpStart installs the Solaris OS and Sun Cluster software on each node. When the installation is successfully completed, each node is fully installed as a new cluster node. Sun Cluster installation output is logged in a /var/cluster/logs/install/scinstall.log. *N* file.

19. **For the Solaris 10 OS, verify on each node that multi-user services for the Service Management Facility (SMF) are online.**

If services are not yet online for a node, wait until the state becomes online before you proceed to the next step.

```
# svcs multi-user-server
STATE          STIME   FMRI
online         17:52:55 svc:/milestone/multi-user-server:default
```

20. **If you are installing a new node to an existing cluster, create mount points on the new node for all existing cluster file systems.**

a. **From another cluster node that is active, display the names of all cluster file systems.**

```
% mount | grep global | egrep -v node@ | awk '{print $1}'
```

b. **On the node that you added to the cluster, create a mount point for each cluster file system in the cluster.**

```
% mkdir -p mountpoint
```

For example, if a file-system name that is returned by the mount command is
/global/dg-schost-1, run mkdir -p /global/dg-schost-1 on the node
that is being added to the cluster.

---

**Note –** The mount points become active after you reboot the cluster in .

---

    c. **If VERITAS Volume Manager (VxVM) is installed on any nodes that are already in the cluster, view the `vxio` number on each VxVM–installed node.**

```
# grep vxio /etc/name_to_major
vxio NNN
```

- Ensure that the same vxio number is used on each of the VxVM-installed nodes.
- Ensure that the vxio number is available for use on each of the nodes that do not have VxVM installed.
- If the vxio number is already in use on a node that does not have VxVM installed, free the number on that node. Change the /etc/name_to_major entry to use a different number.

**21. (Optional) To use dynamic reconfiguration on Sun Enterprise 10000 servers, add the following entry to the `/etc/system` file. Add this entry on each node in the cluster.**

```
set kernel_cage_enable=1
```

This entry becomes effective after the next system reboot. See the *Sun Cluster System Administration Guide for Solaris OS* for procedures to perform dynamic reconfiguration tasks in a Sun Cluster configuration. See your server documentation for more information about dynamic reconfiguration.

**22. To re-enable the loopback file system (LOFS), delete the following entry from the `/etc/system` file on each node of the cluster.**

```
exclude:lofs
```

The re-enabling of LOFS becomes effective after the next system reboot.

**Note –** You cannot have LOFS enabled if you use Sun Cluster HA for NFS on a highly available local file system *and* have `automountd` running. LOFS can cause switchover problems for Sun Cluster HA for NFS. If you enable LOFS and later choose to add Sun Cluster HA for NFS on a highly available local file system, you must do one of the following:

- Restore the `exclude:lofs` entry to the `/etc/system` file on each node of the cluster and reboot each node. This change disables LOFS.

- Disable the `automountd` daemon.

- Exclude from the automounter map all files that are part of the highly available local file system that is exported by Sun Cluster HA for NFS. This choice enables you to keep both LOFS and the `automountd` daemon enabled.

See "Types of File Systems" in *System Administration Guide, Volume 1* (Solaris 8) or "The Loopback File System" in *System Administration Guide: Devices and File Systems* (Solaris 9 or Solaris 10) for more information about loopback file systems.

23. **x86: Set the default boot file to `kadb`.**

    ```
    # eeprom boot-file=kadb
    ```

    The setting of this value enables you to reboot the node if you are unable to access a login prompt.

24. **If you performed a task that requires a cluster reboot, follow these steps to reboot the cluster.**

    The following are some of the tasks that require a reboot:

    - Adding a new node to an existing cluster
    - Installing patches that require a node or cluster reboot
    - Making configuration changes that require a reboot to become active

    a. **From one node, shut down the cluster.**

       ```
       # scshutdown
       ```

       **Note –** Do not reboot the first-installed node of the cluster until *after* the cluster is shut down. Until cluster installation mode is disabled, only the first-installed node, which established the cluster, has a quorum vote. In an established cluster that is still in installation mode, if the cluster is not shut down before the first-installed node is rebooted, the remaining cluster nodes cannot obtain quorum. The entire cluster then shuts down.

       Cluster nodes remain in installation mode until the first time that you run the `scsetup`(1M) command. You run this command during the procedure "How to Configure Quorum Devices" on page 114.

**b. Reboot each node in the cluster.**

▪ **On SPARC based systems, do the following:**

```
ok boot
```

▪ **On x86 based systems, do the following:**

```
                    <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci-ide@7,1/ata@1/cmdk@0,0:b
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
```

The scinstall utility installs and configures all cluster nodes and reboots the cluster. The cluster is established when all nodes have successfully booted into the cluster. Sun Cluster installation output is logged in a /var/cluster/logs/install/scinstall.log.*N* file.

**25. From one node, verify that all nodes have joined the cluster.**

Run the scstat(1M) command to display a list of the cluster nodes. You do not need to be logged in as superuser to run this command.

```
% scstat -n
```

Output resembles the following.

```
-- Cluster Nodes --

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

**Next Steps**  If you added a node to a two-node cluster, go to "How to Update SCSI Reservations After Adding a Node" on page 104.

If you intend to install data services, go to the appropriate procedure for the data service that you want to install and for your version of the Solaris OS:

| | Sun Cluster 2 of 2 CD-ROM (Sun Java System data services) | | Sun Cluster Agents CD (All other data services) | |
|---|---|---|---|---|
| Procedure | Solaris 8 or 9 | Solaris 10 | Solaris 8 or 9 | Solaris 10 |
| "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES installer)" on page 59 | X | | | |
| "How to Install Data-Service Software Packages (pkgadd)" on page 106 | | X | | |
| "How to Install Data-Service Software Packages (scinstall)" on page 108 | | | X | X |
| "How to Install Data-Service Software Packages (Web Start installer)" on page 111 | | | X | |

Otherwise, go to the next appropriate procedure:

- If you installed a single-node cluster, cluster establishment is complete. Go to "Configuring the Cluster" on page 118 to install volume management software and configure the cluster.

- If you added a new node to an existing cluster, verify the state of the cluster. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

- If you installed a multiple-node cluster and chose automatic quorum configuration, postinstallation setup is complete. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

- If you installed a multiple-node cluster and declined automatic quorum configuration, perform postinstallation setup. Go to "How to Configure Quorum Devices" on page 114.

- If you added a node to a cluster that had less or more than two nodes, go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

**Troubleshooting** **Disabled scinstall option** – If the JumpStart option of the scinstall command does not have an asterisk in front, the option is disabled. This condition indicates that JumpStart setup is not complete or that the setup has an error. To correct this condition, first quit the scinstall utility. Repeat Step 1 through Step 10 to correct JumpStart setup, then restart the scinstall utility.

**Error messages about nonexistent nodes** – Unless you have installed your own /etc/inet/ntp.conf file, the scinstall command installs a default ntp.conf file for you. The default file is shipped with references to the maximum number of nodes. Therefore, the xntpd(1M) daemon might issue error messages regarding some of these references at boot time. You can safely ignore these messages. See "How to Configure Network Time Protocol (NTP)" on page 127 for information about how to suppress these messages under otherwise normal cluster conditions.

**Changing the private-network address** – You cannot change the private-network address and netmask after `scinstall` processing has finished. If you need to use a different private-network address or netmask and the node is still in installation mode, follow the procedures in "How to Uninstall Sun Cluster Software to Correct Installation Problems" on page 136. Then repeat this procedure to reinstall and configure the node with the correct information.

## Using SunPlex Installer to Configure Sun Cluster Software

---

**Note –** Do not use this configuration method in the following circumstances:

- To configure a single-node cluster. Instead, follow procedures in "How to Configure Sun Cluster Software on All Nodes (`scinstall`)" on page 65.

- To use a different private-network IP address or netmask than the defaults. SunPlex Installer automatically specifies the default private-network address (`172.16.0.0`) and netmask (`255.255.0.0`). Instead, follow procedures in "How to Configure Sun Cluster Software on All Nodes (`scinstall`)" on page 65.

- To configure tagged-VLAN capable adapters or SCI-PCI adapters for the cluster transport. Instead, follow procedures in "How to Configure Sun Cluster Software on All Nodes (`scinstall`)" on page 65.

- To add a new node to an existing cluster. Instead, follow procedures in "How to Configure Sun Cluster Software on Additional Cluster Nodes (`scinstall`)" on page 96 or "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 72.

---

This section describes how to use SunPlex Installer, the installation module of SunPlex Manager, to establish a new cluster. You can also use SunPlex Installer to install or configure one or more of the following additional software products:

- (On Solaris 8 only) Solstice DiskSuite software – After it installs Solstice DiskSuite software, SunPlex Installer configures up to three metasets and associated metadevices. SunPlex Installer also creates and mounts cluster file systems for each metaset.

- (On Solaris 9 or Solaris 10 only) Solaris Volume Manager software – SunPlex Installer configures up to three Solaris Volume Manager volumes. SunPlex Installer also creates and mounts cluster file systems for each volume. Solaris Volume Manager software is already installed as part of Solaris software installation.

- Sun Cluster HA for NFS data service.

- Sun Cluster HA for Apache scalable data service.

## Installation Requirements

The following table lists SunPlex Installer installation requirements for these
additional software products.

**TABLE 2–3** Requirements to Use SunPlex Installer to Install Software

| Software Package | Installation Requirements |
| --- | --- |
| Solstice DiskSuite or Solaris Volume Manager | A partition that uses /sds as the mount–point name. The partition must be at least 20 Mbytes in size. |
| Sun Cluster HA for NFS data service | ■ At least two shared disks, of the same size, that are connected to the same set of nodes.<br>■ Solstice DiskSuite software installed, or Solaris Volume Manager software configured, by SunPlex Installer.<br>■ A logical hostname for use by Sun Cluster HA for NFS. The logical hostname must have a valid IP address that is accessible by all cluster nodes. The IP address must be on the same subnet as any of the adapters in the IP Network Multipathing group that hosts the logical address.<br>■ A test IP address for each node of the cluster. SunPlex Installer uses these test IP addresses to create Internet Protocol (IP) Network Multipathing (IP Network Multipathing) groups for use by Sun Cluster HA for NFS. |
| Sun Cluster HA for Apache scalable data service | ■ At least two shared disks of the same size that are connected to the same set of nodes.<br>■ Solstice DiskSuite software installed, or Solaris Volume Manager software configured, by SunPlex Installer.<br>■ A shared address for use by Sun Cluster HA for Apache. The shared address must have a valid IP address that is accessible by all cluster nodes. The IP address must be on the same subnet as any of the adapters in the IP Network Multipathing group that hosts the logical address.<br>■ A test IP address for each node of the cluster. SunPlex Installer uses these test IP addresses to create Internet Protocol (IP) Network Multipathing (IP Network Multipathing) groups for use by Sun Cluster HA for Apache. |

## Test IP Addresses

The test IP addresses that you supply must meet the following requirements:

■ 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 the test IP addresses are not highly available.

The following table lists each metaset name and cluster-file-system mount point that is created by SunPlex Installer. The number of metasets and mount points that SunPlex Installer creates depends on the number of shared disks that are connected to the node. For example, if a node is connected to four shared disks, SunPlex Installer creates the `mirror-1` and `mirror-2` metasets. However, SunPlex Installer does not create the `mirror-3` metaset, because the node does not have enough shared disks to create a third metaset.

**TABLE 2–4** Metasets Created by SunPlex Installer

| Shared Disks | Metaset Name | Cluster File System Mount Point | Purpose |
|---|---|---|---|
| First pair | `mirror-1` | `/global/mirror-1` | Sun Cluster HA for NFS or Sun Cluster HA for Apache scalable data service, or both |
| Second pair | `mirror-2` | `/global/mirror-2` | Unused |
| Third pair | `mirror-3` | `/global/mirror-3` | Unused |

**Note –** If the cluster does not meet the minimum shared-disk requirement, SunPlex Installer still installs the Solstice DiskSuite packages. However, without sufficient shared disks, SunPlex Installer cannot configure the metasets, metadevices, or volumes. SunPlex Installer then cannot configure the cluster file systems that are needed to create instances of the data service.

## Character-Set Limitations

SunPlex Installer recognizes a limited character set to increase security. Characters that are not a part of the set are silently filtered out when HTML forms are submitted to the SunPlex Installer server. The following characters are accepted by SunPlex Installer:

`()+,-./0-9:=@A-Z^_a-z{|}~`

This filter can cause problems in the following two areas:

- **Password entry for Sun Java™ System services –** If the password contains unusual characters, these characters are stripped out, resulting in one of the following problems:

  - The resulting password therefore fails because it has less than eight characters.
  - The application is configured with a different password than the user expects.

- **Localization –** Alternative character sets, such as accented characters or Asian characters, do not work for input.

## ▼ How to Configure Sun Cluster Software (SunPlex Installer)

Perform this procedure to use SunPlex Installer to configure Sun Cluster software and install patches on all nodes in the cluster in a single operation. In addition, you can use this procedure to install Solstice DiskSuite software and patches (Solaris 8) and to configure Solstice DiskSuite or Solaris Volume Manager mirrored disk sets.

---

**Note –** Do not use this configuration method in the following circumstances:

- To configure a single-node cluster. Instead, follow procedures in "How to Configure Sun Cluster Software on All Nodes (`scinstall`)" on page 65.

- To use a different private-network IP address or netmask than the defaults. SunPlex Installer automatically specifies the default private-network address (`172.16.0.0`) and netmask (`255.255.0.0`). Instead, follow procedures in "How to Configure Sun Cluster Software on All Nodes (`scinstall`)" on page 65.

- To configure tagged-VLAN capable adapters or SCI-PCI adapters for the cluster transport. Instead, follow procedures in "How to Configure Sun Cluster Software on All Nodes (`scinstall`)" on page 65.

- To add a new node to an existing cluster. Instead, follow procedures in "How to Configure Sun Cluster Software on Additional Cluster Nodes (`scinstall`)" on page 96 or "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 72.

---

The installation process might take from 30 minutes to two or more hours. The actual length of time depends on the number of nodes that are in the cluster, your choice of data services to install, and the number of disks that are in your cluster configuration.

**Before You Begin**   Perform the following tasks:

- Ensure that the cluster configuration meets the requirements to use SunPlex Installer to install software. See "Using SunPlex Installer to Configure Sun Cluster Software" on page 86 for installation requirements and restrictions.

- Ensure that the Solaris OS is installed to support Sun Cluster software.

  If Solaris software is already installed on the node, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 52 for more information about installing Solaris software to meet Sun Cluster software requirements.

- Ensure that Apache software packages and Apache software patches are installed on the node.

  ```
  # pkginfo SUNWapchr SUNWapchu SUNWapchd
  ```

If necessary, install any missing Apache software packages from the Solaris Software 2 of 2 CD-ROM.

- Ensure that Sun Cluster software packages are installed on the node. See "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES installer)" on page 59.

- If you intend to use the root password to access SunPlex Installer or SunPlex Manager, ensure that the root password is the same on every node of the cluster. If necessary, also use the chkey command to update the RPC key pair. See the chkey(1) man page.

- If you intend to install Sun Cluster HA for NFS or Sun Cluster HA for Apache, ensure that the cluster configuration meets all applicable requirements. See "Using SunPlex Installer to Configure Sun Cluster Software" on page 86.

- Complete the following configuration planning worksheet. See "Planning the Solaris OS" on page 16 and "Planning the Sun Cluster Environment" on page 21 for planning guidelines. See the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* for data-service planning guidelines.

| Component | Description/Example | Answer |
|---|---|---|
| Cluster Name | What is the name of the cluster that you want to establish? | |
| | How many nodes are you installing in the cluster? | |
| Node Names | What are the names of the cluster nodes? | |
| Cluster-Transport Adapters and Cables | What are the names of the two transport adapters to use, two adapters per node? | |
| Solstice DiskSuite or Solaris Volume Manager | ■ Solaris 8: Do you want to install Solstice DiskSuite?<br>■ Solaris 9 or Solaris 10: Do you want to configure Solaris Volume Manager? | Yes \| No |
| Sun Cluster HA for NFS<br><br>*Requires Solstice DiskSuite or Solaris Volume Manager* | Do you want to install Sun Cluster HA for NFS?<br><br>If yes, also specify the following: | Yes \| No |
| | What is the logical hostname that the data service is to use? | |
| | What are the test IP addresses to use?<br><br>*Supply one test IP address for each node in the cluster.* | |
| Sun Cluster HA for Apache (scalable)<br><br>*Requires Solstice DiskSuite or Solaris Volume Manager* | Do you want to install scalable Sun Cluster HA for Apache?<br><br>If yes, also specify the following: | Yes \| No |
| | What is the logical hostname that the data service is to use? | |
| | What are the test IP addresses to use?<br><br>*Supply one test IP address for each node in the cluster, if not already supplied for Sun Cluster HA for NFS.* | |

| Component | Description/Example | Answer |
|---|---|---|
| CD-ROM Paths<br><br>*The path to the Sun Cluster framework is always required, even though the packages are already installed.* | What is the path for each of the following components that you want to install?<br><br>*The CD-ROM path must end with the directory that contains the* `.cdtoc` *file. For Sun Cluster CDs, this is usually the media mount point.* | |
| | Solstice DiskSuite: | |
| | Sun Cluster (framework): | |
| | Sun Cluster data services (agents): | |
| | Patches: | |
| Validation Checks | Do you want to run the `sccheck` utility to validate the cluster? | Yes  &#124;  No |

**Steps**    **1. Prepare file-system paths to a CD-ROM image of each software product that you intend to install.**

Follow these guidelines to prepare the file-system paths:

- Provide each CD-ROM image in a location that is available to each node.
- Ensure that the CD-ROM images are accessible to all nodes of the cluster from the same file-system path. These paths can be one or more of the following locations:
  - CD-ROM drives that are exported to the network from machines outside the cluster.
  - Exported file systems on machines outside the cluster.
  - CD-ROM images that are copied to local file systems on each node of the cluster. The local file system must use the same name on each node.

**2. x86: Determine whether you are using the Netscape Navigator™ browser or the Microsoft Internet Explorer browser on your administrative console.**

- **If you are using Netscape Navigator, proceed to Step 3.**

- **If you are using Internet Explorer, skip to Step 4.**

**3. x86: Ensure that the Java plug-in is installed and working on your administrative console.**

   **a. Start the Netscape Navigator browser on the administrative console that you use to connect to the cluster.**

   **b. From the Help menu, choose About Plug-ins.**

   **c. Determine whether the Java plug-in is listed.**

- **If yes, skip to Step 5.**

- **If no, proceed to Step d.**

d. **Download the latest Java plug-in from http://java.sun.com/products/plugin.**

e. **Install the plug-in on your administrative console.**

f. **Create a symbolic link to the plug-in.**

```
% cd ~/.netscape/plugins/
% ln -s /usr/j2se/plugin/i386/ns4/javaplugin.so .
```

g. **Skip to Step 5.**

4. **x86: Ensure that Java 2 Platform, Standard Edition (J2SE) for Windows is installed and working on your administrative console.**

a. **On your Microsoft Windows desktop, click Start, point to Settings, and then select Control Panel.**

The Control Panel window appears.

b. **Determine whether the Java Plug-in is listed.**

- **If no, proceed to Step c.**

- **If yes, double-click the Java Plug-in control panel. When the control panel window opens, click the About tab.**

    - **If an earlier version is shown, proceed to Step c.**

    - **If version 1.4.1 or a later version is shown, skip to Step 5.**

c. **Download the latest version of J2SE for Windows from http://java.sun.com/j2se/downloads.html.**

d. **Install the J2SE for Windows software on your administrative console.**

e. **Restart the system on which your administrative console runs.**

The J2SE for Windows control panel is activated.

5. **If patches exist that are required to support Sun Cluster or Solstice DiskSuite software, determine how to install those patches.**

- **To manually install patches, use the `patchadd` command to install all patches *before* you use SunPlex Installer.**

- **To use SunPlex Installer to install patches, copy patches into a single directory.**

    Ensure that the patch directory meets the following requirements:

    - The patch directory resides on a file system that is available to each node.

- Only one version of each patch is present in this patch directory. If the patch directory contains multiple versions of the same patch, SunPlex Installer cannot determine the correct patch dependency order.

- The patches are uncompressed.

6. **From the administrative console or any other machine outside the cluster, launch a browser.**

7. **Disable the browser's Web proxy.**

   SunPlex Installer installation functionality is incompatible with Web proxies.

8. **Ensure that disk caching and memory caching is enabled.**

   The disk cache and memory cache size must be greater than 0.

9. **From the browser, connect to port 3000 on a node of the cluster.**

   `https://`*node*`:3000`

   The Sun Cluster Installation screen is displayed in the browser window.

   ---

   **Note –** If SunPlex Installer displays the data services installation screen instead of the Sun Cluster Installation screen, Sun Cluster framework software is already installed and configured on that node. Check that the name of the node in the URL is the correct name of the cluster node to install.

   ---

10. **If the browser displays a New Site Certification window, follow the onscreen instructions to accept the certificate.**

11. **Log in as superuser.**

12. **In the Sun Cluster Installation screen, verify that the cluster meets the listed requirements for using SunPlex Installer.**

    If you meet all listed requirements, click Next to continue to the next screen.

13. **Follow the menu prompts to supply your answers from the configuration planning worksheet.**

14. **Click Begin Installation to start the installation process.**

    Follow these guidelines to use SunPlex Installer:

    - Do *not* close the browser window or change the URL during the installation process.

    - If the browser displays a New Site Certification window, follow the onscreen instructions to accept the certificate.

    - If the browser prompts for login information, type the appropriate superuser ID and password for the node that you connect to.

SunPlex Installer installs and configures all cluster nodes and reboots the cluster. The cluster is established when all nodes have successfully booted into the cluster. Sun Cluster installation output is logged in a `/var/cluster/logs/install/scinstall.log.`*N* file.

During installation, the screen displays brief messages about the status of the cluster installation process. When installation and configuration is complete, the browser displays the cluster monitoring and administration GUI.

SunPlex Installer installation output is logged in the `/var/cluster/spm/messages` file. Sun Cluster installation output is logged in a `/var/cluster/logs/install/scinstall.log.` *N* file.

**15. From one node, verify that all nodes have joined the cluster.**

Run the `scstat`(1M) command to display a list of the cluster nodes. You do not need to be logged in as superuser to run this command.

```
% scstat -n
```
Output resembles the following.

```
-- Cluster Nodes --
                          Node name       Status
                          ---------       ------
  Cluster node:           phys-schost-1   Online
  Cluster node:           phys-schost-2   Online
```

**16. Verify the quorum assignments and modify those assignments, if necessary.**

For clusters with three or more nodes, the use of shared quorum devices is optional. SunPlex Installer might or might not have assigned quorum votes to any quorum devices, depending on whether appropriate shared disks were available. You can use SunPlex Manager to designate quorum devices and to reassign quorum votes in the cluster. See Chapter 5, "Administering Quorum," in *Sun Cluster System Administration Guide for Solaris OS* for more information.

**17. To re-enable the loopback file system (LOFS), delete the following entry from the `/etc/system` file on each node of the cluster.**

```
exclude:lofs
```
The re-enabling of LOFS becomes effective after the next system reboot.

**Note –** You cannot have LOFS enabled if you use Sun Cluster HA for NFS on a highly available local file system *and* have `automountd` running. LOFS can cause switchover problems for Sun Cluster HA for NFS. If you enable LOFS and later choose to add Sun Cluster HA for NFS on a highly available local file system, you must do one of the following:

- Restore the `exclude:lofs` entry to the `/etc/system` file on each node of the cluster and reboot each node. This change disables LOFS.

- Disable the `automountd` daemon.

- Exclude from the automounter map all files that are part of the highly available local file system that is exported by Sun Cluster HA for NFS. This choice enables you to keep both LOFS and the `automountd` daemon enabled.

See "Types of File Systems" in *System Administration Guide, Volume 1* (Solaris 8) or "The Loopback File System" in *System Administration Guide: Devices and File Systems* (Solaris 9 or Solaris 10) for more information about loopback file systems.

**Next Steps** If you intend to install data services, go to the appropriate procedure for the data service that you want to install and for your version of the Solaris OS:

| | Sun Cluster 2 of 2 CD-ROM (Sun Java System data services) | | Sun Cluster Agents CD (All other data services) | |
|---|---|---|---|---|
| **Procedure** | **Solaris 8 or 9** | **Solaris 10** | **Solaris 8 or 9** | **Solaris 10** |
| "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES `installer`)" on page 59 | X | | | |
| "How to Install Data-Service Software Packages (`pkgadd`)" on page 106 | | X | | |
| "How to Install Data-Service Software Packages (`scinstall`)" on page 108 | | | X | X |
| "How to Install Data-Service Software Packages (Web Start `installer`)" on page 111 | | | X | |

Otherwise, go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

## ▼ How to Configure Sun Cluster Software on Additional Cluster Nodes (`scinstall`)

Perform this procedure to add a new node to an existing cluster. To use JumpStart to add a new node, instead follow procedures in "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 72.

**Before You Begin**

Perform the following tasks:

- Ensure that all necessary hardware is installed.

  - Ensure that the host adapter is installed on the new node. See the *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*.

  - Verify that any existing cluster interconnects can support the new node. See the *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS*.

  - Ensure that any additional storage is installed. See the appropriate manual from the Sun Cluster 3.x Hardware Administration Collection.

- Ensure that the Solaris OS is installed to support Sun Cluster software.

  If Solaris software is already installed on the node, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 52 for more information about installing Solaris software to meet Sun Cluster software requirements.

- Ensure that Sun Cluster software packages are installed on the node. See "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES installer)" on page 59.

- Determine which mode of the `scinstall` utility you will use, Typical or Custom. For the Typical installation of Sun Cluster software, `scinstall` automatically specifies the following configuration defaults.

| Component | Default Value |
|---|---|
| Cluster-transport junctions | `switch1` and `switch2` |
| Global-devices file-system name | `/globaldevices` |
| Solaris and Sun Cluster patch directory | `/var/cluster/patches` |

- Complete one of the following configuration planning worksheets. See "Planning the Solaris OS" on page 16 and "Planning the Sun Cluster Environment" on page 21 for planning guidelines.

  - **Typical Mode** - If you will use Typical mode and accept all defaults, complete the following worksheet.

| Component | Description/Example | Answer | |
|-----------|---------------------|--------|---|
| Sponsoring Node | What is the name of the sponsoring node?<br><br>*Choose any node that is active in the cluster.* | | |
| Cluster Name | What is the name of the cluster that you want the node to join? | | |
| Check | Do you want to run the `sccheck` validation utility? | Yes \| No | |
| Autodiscovery of Cluster Transport | Do you want to use autodiscovery to configure the cluster transport?<br><br>If no, supply the following additional information: | Yes \| No | |
| Point-to-Point Cables | Does the node that you are adding to the cluster make this a two-node cluster? | Yes \| No | |
| | Does the cluster use transport junctions? | Yes \| No | |
| Cluster–Transport Junctions | If used, what are the names of the two transport junctions?<br>Defaults: `switch1` and `switch2` | *First* | *Second* |
| Cluster-Transport Adapters and Cables | What are the names of the two transport adapters? | *First* | *Second* |
| | Where does each transport adapter connect to (*a transport junction or another adapter*)?<br>Junction defaults: `switch1` and `switch2` | | |
| | For transport junctions, do you want to use the default port name? | Yes \| No | Yes \| No |
| | If no, what is the name of the port that you want to use? | | |
| Automatic Reboot | Do you want `scinstall` to automatically reboot the node after installation? | Yes \| No | |

  - **Custom Mode** - If you will use Custom mode and customize the configuration data, complete the following worksheet.

| Component | Description/Example | Answer |
|-----------|---------------------|--------|
| Software Patch Installation | Do you want `scinstall` to install patches for you? | Yes \| No |
| | If yes, what is the name of the patch directory? | |
| | Do you want to use a patch list? | Yes \| No |

| Component | Description/Example | Answer | |
|---|---|---|---|
| Sponsoring Node | What is the name of the sponsoring node? *Choose any node that is active in the cluster.* | | |
| Cluster Name | What is the name of the cluster that you want the node to join? | | |
| Check | Do you want to run the `sccheck` validation utility? | Yes │ No | |
| Autodiscovery of Cluster Transport | Do you want to use autodiscovery to configure the cluster transport? If no, supply the following additional information: | Yes │ No | |
| Point-to-Point Cables | Does the node that you are adding to the cluster make this a two-node cluster? | Yes │ No | |
| | Does the cluster use transport junctions? | Yes │ No | |
| Cluster-Transport Junctions | If used, what are the names of the two transport junctions? Defaults: `switch1` and `switch2` | *First* | *Second* |
| Cluster-Transport Adapters and Cables | What are the names of the two transport adapters? | *First* | *Second* |
| | Where does each transport adapter connect to (*a transport junction or another adapter*)? Junction defaults: `switch1` and `switch2` | | |
| | For transport junctions, do you want to use the default port name? | Yes │ No | Yes │ No |
| | If no, what is the name of the port that you want to use? | | |
| Global-Devices File System | What is the name of the global-devices file system? Default: `/globaldevices` | | |
| Automatic Reboot | Do you want `scinstall` to automatically reboot the node after installation? | Yes │ No | |

Follow these guidelines to use the interactive `scinstall` utility in this procedure:

- Interactive `scinstall` enables you to type ahead. Therefore, do not press the Return key more than once if the next menu screen does not appear immediately.

- Unless otherwise noted, you can press Control-D to return to either the start of a series of related questions or to the Main Menu.

- Default answers or answers to previous sessions are displayed in brackets ([ ]) at the end of a question. Press Return to enter the response that is in brackets without typing it.

**Steps**    1. **If you are adding this node to a single-node cluster, ensure that two cluster interconnects already exist by displaying the interconnect configuration.**

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

You must have at least two cables or two adapters configured before you can add a node.

- **If the output shows configuration information for two cables or for two adapters, proceed to Step 2.**

- **If the output shows no configuration information for either cables or adapters, or shows configuration information for only one cable or adapter, configure new cluster interconnects.**

    a. **On the existing cluster node, start the scsetup(1M) utility.**

    ```
    # scsetup
    ```

    b. **Choose the menu item, Cluster interconnect.**

    c. **Choose the menu item, Add a transport cable.**
       Follow the instructions to specify the name of the node to add to the cluster, the name of a transport adapter, and whether to use a transport junction.

    d. **If necessary, repeat Step c to configure a second cluster interconnect.**
       When finished, quit the scsetup utility.

    e. **Verify that the cluster now has two cluster interconnects configured.**

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

    The command output should show configuration information for at least two cluster interconnects.

2. **If you are adding this node to an existing cluster, add the new node to the cluster authorized–nodes list.**

    a. **On any active cluster member, start the scsetup(1M) utility.**

    ```
    # scsetup
    ```

    The Main Menu is displayed.

    b. **Choose the menu item, New nodes.**

    c. **Choose the menu item, Specify the name of a machine which may add itself.**

    d. **Follow the prompts to add the node's name to the list of recognized machines.**
       The scsetup utility prints the message Command completed successfully if the task is completed without error.

e.  Quit the `scsetup` utility.

3. **Become superuser on the cluster node to configure.**

4. **Start the `scinstall` utility.**

   ```
   # /usr/cluster/bin/scinstall
   ```

5. **From the Main Menu, choose the menu item, Install a cluster or cluster node.**

   ```
   *** Main Menu ***

     Please select from one of the following (*) options:

       * 1) Install a cluster or cluster node
         2) Configure a cluster to be JumpStarted from this install server
         3) Add support for new data services to this cluster node
         4) Upgrade this cluster node
       * 5) Print release information for this cluster node

       * ?) Help with menu options
       * q) Quit

     Option:  1
   ```

6. **From the Install Menu, choose the menu item, Add this machine as a node in an existing cluster.**

7. **Follow the menu prompts to supply your answers from the configuration planning worksheet.**

   The `scinstall` utility configures the node and boots the node into the cluster.

8. **Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

   ```
   # eject cdrom
   ```

9. **Install any necessary patches to support Sun Cluster software, if you have not already done so.**

10. **Repeat this procedure on any other node to add to the cluster until all additional nodes are fully configured.**

11. **For the Solaris 10 OS, verify on each node that multi-user services for the Service Management Facility (SMF) are online.**

    If services are not yet online for a node, wait until the state becomes online before you proceed to the next step.

    ```
    # svcs multi-user-server
    STATE          STIME    FMRI
    online         17:52:55 svc:/milestone/multi-user-server:default
    ```

12. **From an active cluster member, prevent any other nodes from joining the cluster.**

    ```
    # /usr/cluster/bin/scconf -a -T node=.
    ```

| -a | Specifies the add form of the command |
| -T | Specifies authentication options |
| node=. | Specifies the node name of dot (.) to add to the authentication list, to prevent any other node from adding itself to the cluster |

Alternately, you can use the scsetup(1M) utility. See "How to Add a Node to the Authorized Node List" in *Sun Cluster System Administration Guide for Solaris OS* for procedures.

13. **From one node, verify that all nodes have joined the cluster.**

    Run the scstat(1M) command to display a list of the cluster nodes. You do not need to be logged in as superuser to run this command.

    ```
    % scstat -n
    ```

    Output resembles the following.

    ```
    -- Cluster Nodes --

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

14. **To re-enable the loopback file system (LOFS), delete the following entry from the /etc/system file on each node of the cluster.**

    ```
    exclude:lofs
    ```

    The re-enabling of LOFS becomes effective after the next system reboot.

    ---

    **Note –** You cannot have LOFS enabled if you use Sun Cluster HA for NFS on a highly available local file system *and* have automountd running. LOFS can cause switchover problems for Sun Cluster HA for NFS. If you enable LOFS and later choose to add Sun Cluster HA for NFS on a highly available local file system, you must do one of the following:

    - Restore the exclude:lofs entry to the /etc/system file on each node of the cluster and reboot each node. This change disables LOFS.

    - Disable the automountd daemon.

    - Exclude from the automounter map all files that are part of the highly available local file system that is exported by Sun Cluster HA for NFS. This choice enables you to keep both LOFS and the automountd daemon enabled.

    ---

    See "Types of File Systems" in *System Administration Guide, Volume 1* (Solaris 8) or "The Loopback File System" in *System Administration Guide: Devices and File Systems* (Solaris 9 or Solaris 10) for more information about loopback file systems.

**Example 2–2** Configuring Sun Cluster Software on an Additional Node

The following example shows the node `phys-schost-3` added to the cluster schost. The sponsoring node is `phys-schost-1`.

```
*** Adding a Node to an Existing Cluster ***
Fri Feb  4 10:17:53 PST 2005


scinstall -ik -C schost -N phys-schost-1 -A trtype=dlpi,name=qfe2 -A trtype=dlpi,name=qfe3
-m endpoint=:qfe2,endpoint=switch1 -m endpoint=:qfe3,endpoint=switch2


Checking device to use for global devices file system ... done

Adding node "phys-schost-3" to the cluster configuration ... done
Adding adapter "qfe2" to the cluster configuration ... done
Adding adapter "qfe3" to the cluster configuration ... done
Adding cable to the cluster configuration ... done
Adding cable to the cluster configuration ... done

Copying the config from "phys-schost-1" ... done

Copying the postconfig file from "phys-schost-1" if it exists ... done
Copying the Common Agent Container keys from "phys-schost-1" ... done


Setting the node ID for "phys-schost-3" ... done (id=1)

Setting the major number for the "did" driver ...
Obtaining the major number for the "did" driver from "phys-schost-1" ... done
"did" driver major number set to 300

Checking for global devices global file system ... done
Updating vfstab ... done

Verifying that NTP is configured ... done
Initializing NTP configuration ... done

Updating nsswitch.conf ...
done

Adding clusternode entries to /etc/inet/hosts ... done


Configuring IP Multipathing groups in "/etc/hostname.<adapter>" files

Updating "/etc/hostname.hme0".

Verifying that power management is NOT configured ... done

Ensure that the EEPROM parameter "local-mac-address?" is set to "true" ... done
The "local-mac-address?" parameter setting has been changed to "true".
```

```
Ensure network routing is disabled ... done

Updating file ("ntp.conf.cluster") on node phys-schost-1 ... done
Updating file ("hosts") on node phys-schost-1 ... done

Rebooting ...
```

**Next Steps**   Determine your next step:

If you added a node to a two-node cluster, go to "How to Update SCSI Reservations After Adding a Node" on page 104.

If you intend to install data services, go to the appropriate procedure for the data service that you want to install and for your version of the Solaris OS:

| | Sun Cluster 2 of 2 CD-ROM (Sun Java System data services) | | Sun Cluster Agents CD (All other data services) | |
|---|---|---|---|---|
| Procedure | Solaris 8 or 9 | Solaris 10 | Solaris 8 or 9 | Solaris 10 |
| "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES `installer`)" on page 59 | X | | | |
| "How to Install Data-Service Software Packages (`pkgadd`)" on page 106 | | X | | |
| "How to Install Data-Service Software Packages (`scinstall`)" on page 108 | | | X | X |
| "How to Install Data-Service Software Packages (Web Start `installer`)" on page 111 | | | X | |

Otherwise, go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

**Troubleshooting** When you increase or decrease the number of node attachments to a quorum device, the cluster does not automatically recalculate the quorum vote count. To reestablish the correct quorum vote, use the `scsetup` utility to remove each quorum device and then add it back into the configuration. Do this on one quorum device at a time.

If the cluster has only one quorum device, configure a second quorum device before you remove and readd the original quorum device. Then remove the second quorum device to return the cluster to its original configuration.

## ▼ How to Update SCSI Reservations After Adding a Node

If you added a node to a two-node cluster that uses one or more shared SCSI disks as quorum devices, you must update the SCSI Persistent Group Reservations (PGR). To do this, you remove the quorum devices which have SCSI-2 reservations. If you want to add back quorum devices, the newly configured quorum devices will have SCSI-3 reservations.

**Before You Begin**    Ensure that you have completed installation of Sun Cluster software on the added node.

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

2. **View the current quorum configuration.**
   The following example output shows the status of quorum device d3.

   # **scstat -q**
   Note the name of each quorum device that is listed.

3. **Remove the original quorum device.**
   Perform this step for each quorum device that is configured.

   # **scconf -r -q globaldev=***devicename*

   -r                                     Removes

   -q globaldev=*devicename*    Specifies the name of the quorum device

4. **Verify that all original quorum devices are removed.**

   # **scstat -q**

5. **(Optional) Add a SCSI quorum device.**
   You can configure the same device that was originally configured as the quorum device or choose a new shared device to configure.

   a. **(Optional) If you want to choose a new shared device to configure as a quorum device, display all devices that the system checks.**
      Otherwise, skip to Step c.

      # **scdidadm -L**
      Output resembles the following:

      ```
      1        phys-schost-1:/dev/rdsk/c0t0d0 /dev/did/rdsk/d1
      2        phys-schost-1:/dev/rdsk/c1t1d0 /dev/did/rdsk/d2
      2        phys-schost-2:/dev/rdsk/c1t1d0 /dev/did/rdsk/d2
      3        phys-schost-1:/dev/rdsk/c1t2d0 /dev/did/rdsk/d3
      3        phys-schost-2:/dev/rdsk/c1t2d0 /dev/did/rdsk/d3
      ```

```
...
```

   **b. From the output, choose a shared device to configure as a quorum device.**

   **c. Configure the shared device as a quorum device.**

   ```
   # scconf -a -q globaldev=devicename
   ```

   -a    Adds

   **d. Repeat for each quorum device that you want to configure.**

6. **If you added any quorum devices, verify the new quorum configuration.**

   ```
   # scstat -q
   ```

   Each new quorum device should be Online and have an assigned vote.

**Example 2–3** Updating SCSI Reservations After Adding a Node

The following example identifies the original quorum device d2, removes that quorum device, lists the available shared devices, and configures d3 as a new quorum device.

```
    (List quorum devices)
# scstat -q
...
-- Quorum Votes by Device --

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

    (Remove the original quorum device)
# scconf -r -q globaldev=d2

    (Verify the removal of the original quorum device)
# scstat -q
...
-- Quorum Votes by Device --

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

    (List available devices)
# scdidadm -L
...
3      phys-schost-1:/dev/rdsk/c1t2d0 /dev/did/rdsk/d3
3      phys-schost-2:/dev/rdsk/c1t2d0 /dev/did/rdsk/d3
...

    (Add a quorum device)
# scconf -a -q globaldev=d3

    (Verify the addition of the new quorum device)
```

```
# scstat -q
...
-- Quorum Votes by Device --

                       Device Name         Present Possible Status
                       -----------         ------- -------- ------
      Device votes:    /dev/did/rdsk/d3s2 2    2        Online
```

**Next Steps** ■ If you intend to install data services, go to the appropriate procedure for the data service that you want to install and for your version of the Solaris OS:

| | Sun Cluster 2 of 2 CD-ROM (Sun Java System data services) | | Sun Cluster Agents CD (All other data services) | |
|---|---|---|---|---|
| **Procedure** | **Solaris 8 or 9** | **Solaris 10** | **Solaris 8 or 9** | **Solaris 10** |
| "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES installer)" on page 59 | X | | | |
| "How to Install Data-Service Software Packages (pkgadd)" on page 106 | | X | | |
| "How to Install Data-Service Software Packages (scinstall)" on page 108 | | | X | X |
| "How to Install Data-Service Software Packages (Web Start installer)" on page 111 | | | X | |

■ Otherwise, go to "How to Verify the Quorum Configuration and Installation Mode" on page 117

## ▼ How to Install Data-Service Software Packages (pkgadd)

Perform this procedure to install data services for the Solaris 10 OS from the Sun Cluster 2 of 2 CD-ROM. The Sun Cluster 2 of 2 CD-ROM contains the data services for Sun Java System applications. This procedure uses the pkgadd(1M) program to install the packages. Perform this procedure on each node in the cluster on which you want to run a chosen data service.

---

**Note –** Do not use this procedure for the following kinds of data-service packages:

- Data services for the Solaris 8 or Solaris 9 OS from the Sun Cluster 2 of 2 CD-ROM
  - Instead, follow installation procedures in "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES `installer`)" on page 59.

- Data services for the Solaris 10 OS from the Sun Cluster Agents CD - Instead, follow installation procedures in "How to Install Data-Service Software Packages (`scinstall`)" on page 108. The Web Start `installer` program on the Sun Cluster Agents CD is not compatible with the Solaris 10 OS.

---

**Steps** 1. **Become superuser on the cluster node.**

2. **Insert the Sun Cluster 2 of 2 CD-ROM in the CD-ROM drive.**

   If the volume management daemon `vold`(1M) is running and is configured to manage CD-ROM devices, the daemon automatically mounts the CD-ROM on the `/cdrom/cdrom0/` directory.

3. **Change to the Solaris_*arch*/Product/sun_cluster_agents/Solaris_10/Packages/ directory, where *arch* is sparc or x86.**

   ```
   # cd /cdrom/cdrom0/Solaris_arch/Product/sun_cluster_agents/ \
   Solaris_10/Packages/
   ```

4. **Install the data service packages on the global zone.**

   ```
   # pkgadd -G -d . [packages]
   ```

   -G          Adds packages to the current zone only. You must add Sun Cluster packages only to the global zone. This option also specifies that the packages are *not* propagated to any existing non-global zone or to any non-global zone that is created later.

   -d          Specifies the location of the packages to install.

   *packages*   *Optional.* Specifies the name of one or more packages to install. If no package name is specified, the `pkgadd` program displays a pick list of all packages that are available to install.

5. **Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

   ```
   # eject cdrom
   ```

6. **Install any patches for the data services that you installed.**

   See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

   You do not have to reboot after you install Sun Cluster data-service patches unless a reboot is specified by the patch special instructions. If a patch instruction requires

that you reboot, perform the following steps:

a. **From one node, shut down the cluster by using the `scshutdown`(1M) command.**

b. **Reboot each node in the cluster.**

---

**Note –** Until cluster installation mode is disabled, only the first-installed node, which established the cluster, has a quorum vote. In an established multiple-node cluster which is still in installation mode, if the cluster is not shut down before the first-installed node is rebooted, the remaining cluster nodes cannot obtain quorum. The entire cluster then shuts down.

If you chose automatic quorum configuration during Sun Cluster installation or used SunPlex Installer to install Sun Cluster software, the installation utility automatically assigns quorum votes and removes the cluster from installation mode during installation reboot. However, if you did not choose one of these methods, cluster nodes remain in installation mode until you run the `scsetup`(1M) command, during the procedure "How to Configure Quorum Devices" on page 114.

---

**Next Steps**
- If you installed a single-node cluster, cluster establishment is complete. Go to "Configuring the Cluster" on page 118 to install volume management software and configure the cluster.
- If you added a new node to an existing cluster, verify the state of the cluster. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.
- If you declined automatic quorum configuration during Sun Cluster software installation of a multiple-node cluster, perform postinstallation setup. Go to "How to Configure Quorum Devices" on page 114.
- If you chose automatic quorum configuration during Sun Cluster software installation of a multiple-node cluster, postinstallation setup is complete. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.
- If you used SunPlex Installer to install a multiple-node cluster, postinstallation setup is complete. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

## ▼ How to Install Data-Service Software Packages (`scinstall`)

Perform this procedure to install data services from the Sun Cluster Agents CD of the Sun Cluster 3.1 8/05 release. This procedure uses the interactive `scinstall` utility to install the packages. Perform this procedure on each node in the cluster on which you want to run a chosen data service.

**Note –** Do not use this procedure for the following kinds of data-service packages:

- Data services for the Solaris 10 OS from the Sun Cluster 2 of 2 CD-ROM - Instead, follow installation procedures in "How to Install Data-Service Software Packages (pkgadd)" on page 106.

- Data services for the Solaris 8 or Solaris 9 OS from the Sun Cluster 2 of 2 CD-ROM - Instead, follow installation procedures in "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES installer)" on page 59.

You do not need to perform this procedure if you used SunPlex Installer to install Sun Cluster HA for NFS or Sun Cluster HA for Apache or both and you do not intend to install any other data services. Instead, go to "How to Configure Quorum Devices" on page 114.

To install data services from the Sun Cluster 3.1 10/03 release or earlier, you can alternatively use the Web Start installer program to install the packages. See "How to Install Data-Service Software Packages (Web Start installer)" on page 111.

Follow these guidelines to use the interactive scinstall utility in this procedure:

- Interactive scinstall enables you to type ahead. Therefore, do not press the Return key more than once if the next menu screen does not appear immediately.

- Unless otherwise noted, you can press Control-D to return to either the start of a series of related questions or to the Main Menu.

- Default answers or answers to previous sessions are displayed in brackets ([ ]) at the end of a question. Press Return to enter the response that is in brackets without typing it.

**Steps** 1. **Become superuser on the cluster node.**

2. **Insert the Sun Cluster Agents CD in the CD-ROM drive on the node.**

   If the volume management daemon vold(1M) is running and is configured to manage CD-ROM devices, the daemon automatically mounts the CD-ROM on the /cdrom/cdrom0/ directory.

3. **Change to the directory where the CD-ROM is mounted.**

   ```
   # cd /cdrom/cdrom0/
   ```

4. **Start the scinstall(1M) utility.**

   ```
   # scinstall
   ```

5. **From the Main Menu, choose the menu item, Add support for new data services to this cluster node.**

6. **Follow the prompts to select the data services to install.**

   You must install the same set of data-service packages on each node. This requirement applies even if a node is not expected to host resources for an installed data service.

7. **After the data services are installed, quit the `scinstall` utility.**

8. **Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

   ```
   # eject cdrom
   ```

9. **Install any Sun Cluster data-service patches.**

   See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

   You do not have to reboot after you install Sun Cluster data-service patches unless a reboot is specified by the patch special instructions. If a patch instruction requires that you reboot, perform the following steps:

   a. **From one node, shut down the cluster by using the `scshutdown(1M)` command.**

   b. **Reboot each node in the cluster.**

   ---

   **Note –** Until cluster installation mode is disabled, only the first-installed node, which established the cluster, has a quorum vote. In an established multiple-node cluster which is still in installation mode, if the cluster is not shut down before the first-installed node is rebooted, the remaining cluster nodes cannot obtain quorum. This inability to obtain quorum causes the entire cluster to shut down.

   If you chose automatic quorum configuration during Sun Cluster installation or used SunPlex Installer to install Sun Cluster software, the installation utility automatically assigns quorum votes and removes the cluster from installation mode during installation reboot. However, if you did not choose one of these methods, cluster nodes remain in installation mode until you run the `scsetup`(1M) command, during the procedure "How to Configure Quorum Devices" on page 114.

   ---

**Next Steps**  ■ If you installed a single-node cluster, cluster establishment is complete. Go to "Configuring the Cluster" on page 118 to install volume management software and configure the cluster.

■ If you added a new node to an existing cluster, verify the state of the cluster. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

■ If you declined automatic quorum configuration during Sun Cluster software installation of a multiple-node cluster, perform postinstallation setup. Go to "How to Configure Quorum Devices" on page 114.

- If you chose automatic quorum configuration during Sun Cluster software installation of a multiple-node cluster, postinstallation setup is complete. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

- If you used SunPlex Installer to install a multiple-node cluster, postinstallation setup is complete. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

## ▼ How to Install Data-Service Software Packages (Web Start `installer`)

Perform this procedure to install data services for the Solaris 8 or Solaris 9 OS from the Sun Cluster Agents CD. This procedure uses the Web Start `installer` program on the CD-ROM to install the packages. Perform this procedure on each node in the cluster on which you want to run a chosen data service.

---

**Note –** Do not use this procedure for the following kinds of data-service packages:

- Data services for the Solaris 10 OS from the Sun Cluster Agents CD - Instead, follow installation procedures in "How to Install Data-Service Software Packages (`scinstall`)" on page 108. The Web Start `installer` program on the Sun Cluster Agents CD is not compatible with the Solaris 10 OS.

- Data services for the Solaris 10 OS from the Sun Cluster 2 of 2 CD-ROM - Instead, follow installation procedures in "How to Install Data-Service Software Packages (`pkgadd`)" on page 106.

- Data services for the Solaris 8 or Solaris 9 OS from the Sun Cluster 2 of 2 CD-ROM - Instead, follow installation procedures in "How to Install Sun Cluster Framework and Data-Service Software Packages (Java ES `installer`)" on page 59.

You do not need to perform this procedure if you used SunPlex Installer to install Sun Cluster HA for NFS or Sun Cluster HA for Apache or both and you do not intend to install any other data services. Instead, go to "How to Configure Quorum Devices" on page 114.

---

To install data services from the Sun Cluster 3.1 10/03 release or earlier, you can alternatively follow the procedures in "How to Install Data-Service Software Packages (`scinstall`)" on page 108.

You can run the `installer` program with a command-line interface (CLI) or with a graphical user interface (GUI). The content and sequence of instructions in the CLI and the GUI are similar. For more information about the `installer` program, see the `installer`(1M) man page.

**Before You Begin**  If you intend to use the installer program with a GUI, ensure that the DISPLAY environment variable is set.

**Steps**  1. **Become superuser on the cluster node.**

2. **Insert the Sun Cluster Agents CD in the CD-ROM drive.**

   If the volume management daemon vold(1M) is running and is configured to manage CD-ROM devices, the daemon automatically mounts the CD-ROM on the /cdrom/cdrom0/ directory.

3. **Change to the directory of the CD-ROM where the installer program resides.**

   # **cd /cdrom/cdrom0/Solaris_*arch*/**

   In the Solaris_*arch*/ directory, *arch* is sparc or x86.

4. **Start the Web Start installer program.**

   # **./installer**

5. **When you are prompted, select the type of installation.**

   See the Sun Cluster Release Notes for a listing of the locales that are available for each data service.

   - **To install all data services on the CD-ROM, select Typical.**

   - **To install only a subset of the data services on the CD-ROM, select Custom.**

6. **When you are prompted, select the locale to install.**

   - **To install only the C locale, select Typical.**

   - **To install other locales, select Custom.**

7. **Follow instructions on the screen to install the data-service packages on the node.**

   After the installation is finished, the installer program provides an installation summary. This summary enables you to view logs that the program created during the installation. These logs are located in the /var/sadm/install/logs/ directory.

8. **Quit the installer program.**

9. **Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

   # **eject cdrom**

10. **Install any Sun Cluster data-service patches.**

See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

You do not have to reboot after you install Sun Cluster data-service patches unless a reboot is specified by the patch special instructions. If a patch instruction requires that you reboot, perform the following steps:

a. **From one node, shut down the cluster by using the `scshutdown(1M)` command.**

b. **Reboot each node in the cluster.**

---

**Note –** Until cluster installation mode is disabled, only the first-installed node, which established the cluster, has a quorum vote. In an established multiple-node cluster which is still in installation mode, if the cluster is not shut down before the first-installed node is rebooted, the remaining cluster nodes cannot obtain quorum. The entire cluster then shuts down.

If you chose automatic quorum configuration during Sun Cluster installation or used SunPlex Installer to install Sun Cluster software, the installation utility automatically assigns quorum votes and removes the cluster from installation mode during installation reboot. However, if you did not choose one of these methods, cluster nodes remain in installation mode until you run the `scsetup(1M)` command, during the procedure "How to Configure Quorum Devices" on page 114.

---

**Next Steps** ■ If you installed a single-node cluster, cluster establishment is complete. Go to "Configuring the Cluster" on page 118 to install volume management software and configure the cluster.

■ If you added a new node to an existing cluster, verify the state of the cluster. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

■ If you declined automatic quorum configuration during Sun Cluster software installation of a multiple-node cluster, perform postinstallation setup. Go to "How to Configure Quorum Devices" on page 114.

■ If you chose automatic quorum configuration during Sun Cluster software installation of a multiple-node cluster, postinstallation setup is complete. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

■ If you used SunPlex Installer to install a multiple-node cluster, postinstallation setup is complete. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

# ▼ How to Configure Quorum Devices

---

**Note –** You do not need to configure quorum devices in the following circumstances:

- You chose automatic quorum configuration during Sun Cluster software configuration.
- You used SunPlex Installer to install the cluster. SunPlex Installer assigns quorum votes and removes the cluster from installation mode for you.
- You installed a single-node cluster.
- You added a node to an existing cluster and already have sufficient quorum votes assigned.

Instead, proceed to "How to Verify the Quorum Configuration and Installation Mode" on page 117.

---

Perform this procedure one time only, after the cluster is fully formed. Use this procedure to assign quorum votes and then to remove the cluster from installation mode.

**Before You Begin**  If you intend to configure a Network Appliance network-attached storage (NAS) device as a quorum device, do the following:

- Install the NAS device hardware and software. See Chapter 1, "Installing and Maintaining Network Appliance Network-Attached Storage Devices in a Sun Cluster Environment," in *Sun Cluster 3.1 With Network-Attached Storage Devices Manual for Solaris OS* and your device documentation for requirements and installation procedures for NAS hardware and software.

- Have available the following information:

  - The name of the NAS device
  - The LUN ID of the NAS device

  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*.

**Steps**   1. **If you want to use a shared SCSI disk as a quorum device, verify device connectivity to the cluster nodes and choose the device to configure.**

a. **From one node of the cluster, display a list of all the devices that the system checks.**

   You do not need to be logged in as superuser to run this command.

   ```
   % scdidadm -L
   ```

   Output resembles the following:

   ```
   1          phys-schost-1:/dev/rdsk/c0t0d0 /dev/did/rdsk/d1
   2          phys-schost-1:/dev/rdsk/c1t1d0 /dev/did/rdsk/d2
   2          phys-schost-2:/dev/rdsk/c1t1d0 /dev/did/rdsk/d2
   3          phys-schost-1:/dev/rdsk/c1t2d0 /dev/did/rdsk/d3
   3          phys-schost-2:/dev/rdsk/c1t2d0 /dev/did/rdsk/d3
   ...
   ```

b. **Ensure that the output shows all connections between cluster nodes and storage devices.**

c. **Determine the global device-ID name of each shared disk that you are configuring as a quorum device.**

   ---
   **Note –** Any shared disk that you choose must be qualified for use as a quorum device. See "Quorum Devices" on page 32 for further information about choosing quorum devices.

   ---

   Use the scdidadm output from Step a to identify the device–ID name of each shared disk that you are configuring as a quorum device. For example, the output in Step a shows that global device d2 is shared by phys-schost-1 and phys-schost-2.

2. **Become superuser on one node of the cluster.**

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

   ```
   # scsetup
   ```

   The Initial Cluster Setup screen is displayed.

> **Note –** If the Main Menu is displayed instead, initial cluster setup was already
> successfully performed. Skip to Step 8.

4. **Answer the prompt `Do you want to add any quorum disks?`.**

   - **If your cluster is a two-node cluster, you must configure at least one shared
     quorum device. Type `Yes` to configure one or more quorum devices.**

   - **If your cluster has three or more nodes, quorum device configuration is
     optional.**

     - **Type `No` if you do not want to configure additional quorum devices. Then
       skip to Step 7.**

     - **Type `Yes` to configure additional quorum devices. Then proceed to Step 5.**

5. **Specify what type of device you want to configure as a quorum device.**

   - **Choose `scsi` to configure a shared SCSI disk.**

   - **Choose `netapp_nas` to configure a Network Appliance NAS device.**

6. **Specify the name of the device to configure as a quorum device.**

   For a Network Appliance NAS device, also specify the following information:

   - The name of the NAS device
   - The LUN ID of the NAS device

7. **At the prompt `Is it okay to reset "installmode"?`, type `Yes`.**

   After the `scsetup` utility sets the quorum configurations and vote counts for the
   cluster, the message `Cluster initialization is complete` is displayed. The
   utility returns you to the Main Menu.

8. **Quit the `scsetup` utility.**

**Next Steps**   Verify the quorum configuration and that installation mode is disabled. Go to "How to
Verify the Quorum Configuration and Installation Mode" on page 117.

**Troubleshooting** **Interrupted scsetup processing** — If the quorum setup process is interrupted or fails
to be completed successfully, rerun `scsetup`.

**Changes to quorum vote count** — If you later 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 by removing each quorum
device and then add it back into the configuration, one quorum device at a time. 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. See the procedure "How to Modify a Quorum Device Node List" in Chapter 5, "Administering Quorum," in *Sun Cluster System Administration Guide for Solaris OS*.

## ▼ How to Verify the Quorum Configuration and Installation Mode

Perform this procedure to verify that quorum configuration was completed successfully and that cluster installation mode is disabled.

**Steps** 1. **From any node, verify the device and node quorum configurations.**

```
% scstat -q
```

2. **From any node, verify that cluster installation mode is disabled.**
   You do not need to be superuser to run this command.

```
% scconf -p | grep "install mode"
Cluster install mode:              disabled
```

Cluster installation is complete.

**Next Steps** Go to to install volume management software and perform other configuration tasks on the cluster or new cluster node.

---

**Note –** If you added a new node to a cluster that uses VxVM, you must perform steps in to do one of the following tasks:

- Install VxVM on that node.
- Modify that node's /etc/name_to_major file, to support coexistence with VxVM.

---

# Configuring the Cluster

This section provides information and procedures to configure the software that you installed on the cluster or new cluster node. Before you start to perform these tasks, ensure that you completed the following tasks:

- Installed software packages for the Solaris OS, Sun Cluster framework, and other products as described in "Installing the Software" on page 45
- Established the new cluster or cluster node as described in "Establishing the Cluster" on page 63

The following table lists the tasks to perform to configure your cluster. Complete the procedures in the order that is indicated.

---

**Note –** If you added a new node to a cluster that uses VxVM, you must perform steps in "SPARC: How to Install VERITAS Volume Manager Software" on page 179 to do one of the following tasks:

- Install VxVM on that node.
- Modify that node's /etc/name_to_major file, to support coexistence with VxVM.

---

**TABLE 2–5** Task Map: Configuring the Cluster

| Task | Instructions |
|---|---|
| 1. Install and configure volume management software: | |
| ■ Install and configure Solstice DiskSuite or Solaris Volume Manager software | Chapter 3<br><br>Solstice DiskSuite or Solaris Volume Manager documentation |
| ■ SPARC: Install and configure VERITAS Volume Manager software. | Chapter 4<br><br>VERITAS Volume Manager documentation |
| 2. Create and mount cluster file systems. | "How to Create Cluster File Systems" on page 119 |
| 3. *(Solaris 8 or SunPlex Installer installations)* Create Internet Protocol (IP) Network Multipathing groups for each public-network adapter that is not already configured in an IP Network Multipathing group. | "How to Configure Internet Protocol (IP) Network Multipathing Groups" on page 125 |

**TABLE 2–5** Task Map: Configuring the Cluster     *(Continued)*

| Task | Instructions |
|------|-------------|
| 4. *(Optional)* Change a node's private hostname. | "How to Change Private Hostnames" on page 126 |
| 5. Create or modify the NTP configuration file. | "How to Configure Network Time Protocol (NTP)" on page 127 |
| 6. *(Optional)* SPARC: Install the Sun Cluster module to Sun Management Center software. | "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 130 <br><br> Sun Management Center documentation |
| 7. Install third-party applications and configure the applications, data services, and resource groups. | *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* <br><br> Third-party application documentation |

## ▼ How to Create Cluster File Systems

Perform this procedure for each cluster file system that you want to create. Unlike a local file system, a cluster file system is accessible from any node in the cluster. If you used SunPlex Installer to install data services, SunPlex Installer might have already created one or more cluster file systems.

**Caution –** Any data on the disks is destroyed when you create a file system. Be sure that you specify the correct disk device name. If you specify the wrong device name, you might erase data that you did not intend to delete.

**Before You Begin**   Perform the following tasks:

- Ensure that volume-manager software is installed and configured. For volume-manager installation procedures, see "Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software" on page 141 or "SPARC: Installing and Configuring VxVM Software" on page 177.

- Determine the mount options to use for each cluster file system that you want to create. Observe the Sun Cluster mount-option requirements and restrictions that are described in the following tables:

  - Mount Options for UFS Cluster File Systems

| Mount Option | Description |
|--------------|-------------|
| `global` | **Required**. This option makes the file system globally visible to all nodes in the cluster. |

| Mount Option | Description |
| --- | --- |
| logging | **Required**. This option enables logging. |
| forcedirectio | **Required** for cluster file systems that will host Oracle Real Application Clusters RDBMS data files, log files, and control files. |
| | **Note –** Oracle Real Application Clusters is supported for use only in SPARC based clusters. |
| onerror=panic | **Required**. You do not have to explicitly specify the onerror=panic mount option in the /etc/vfstab file. This mount option is already the default value if no other onerror mount option is specified. |
| | **Note –** Only the onerror=panic mount option is supported by Sun Cluster software. Do not use the onerror=umount or onerror=lock mount options. These mount options are not supported on cluster file systems for the following reasons:<ul><li>Use of the onerror=umount or onerror=lock mount option might cause the cluster file system to lock or become inaccessible. This condition might occur if the cluster file system experiences file corruption.</li><li>The onerror=umount or onerror=lock mount option might cause the cluster file system to become unmountable. This condition might thereby cause applications that use the cluster file system to hang or prevent the applications from being killed.</li></ul> |
| | A node might require rebooting to recover from these states. |
| syncdir | **Optional**. If you specify syncdir, you are guaranteed POSIX-compliant file system behavior for the write() system call. If a write() succeeds, then this mount option ensures that sufficient space is on the disk. |
| | If you do not specify syncdir, the same behavior occurs 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 (ENOSPC) until you close a file. |
| | You see ENOSPC on close only during a very short time after a failover. With syncdir, as with POSIX behavior, the out-of-space condition would be discovered before the close. |

See the mount_ufs(1M) man page for more information about UFS mount options.

■ Mount Parameters for Sun StorEdge QFS Shared File Systems

| Mount Parameter | Description |
| --- | --- |
| shared | **Required**. This option specifies that this is a shared file system, therefore globally visible to all nodes in the cluster. |

**Caution –** Ensure that settings in the `/etc/vfstab` file do not conflict with settings in the `/etc/opt/SUNWsamfs/samfs.cmd` file. Settings in the `/etc/vfstab` file override settings in the `/etc/opt/SUNWsamfs/samfs.cmd` file.

See the `mount_samfs`(1M) man page for more information about QFS mount parameters.

Certain data services such as Sun Cluster Support for Oracle Real Application Clusters have additional requirements and guidelines for QFS mount parameters. See your data service manual for any additional requirements.

**Note –** Logging is not enabled by an `/etc/vfstab` mount parameter, nor does Sun Cluster software require logging for QFS shared file systems.

■ Mount Options for VxFS Cluster File Systems

| Mount Option | Description |
| --- | --- |
| global | **Required**. This option makes the file system globally visible to all nodes in the cluster. |
| log | **Required**. This option enables logging. |

See the VxFS `mount_vxfs` man page and "Administering Cluster File Systems Overview" in *Sun Cluster System Administration Guide for Solaris OS* for more information about VxFS mount options.

**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 create a file system.

2. **Create a file system.**

   ■ **For a UFS file system, use the newfs(1M) command.**

   # **newfs** *raw-disk-device*

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

| Volume Manager | Sample Disk Device Name | Description |
| --- | --- | --- |
| Solstice DiskSuite or Solaris Volume Manager | `/dev/md/nfs/rdsk/d1` | Raw disk device d1 within the nfs 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 d1s3 |

- **For a Sun StorEdge QFS file system, follow the procedures for defining the configuration in the** *Sun StorEdge QFS and Sun StorEdge SAM-FS Software Installation and Configuration Guide***.**

- **SPARC: For a VERITAS File System (VxFS) file system, follow the procedures that are provided in your VxFS documentation.**

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 is not accessed on that node.

   ---

   **Tip –** For ease of administration, create the mount point in the `/global/`*device-group*`/` directory. 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.**

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

   a. **In each entry, specify the required mount options for the type of file system that you use.**

> **Note –** Do *not* use the `logging` mount option for Solstice DiskSuite trans metadevices or Solaris Volume Manager transactional volumes. Trans metadevices and transactional volumes provide their own logging.
>
> In addition, Solaris Volume Manager transactional-volume logging (formerly Solstice DiskSuite trans-metadevice logging) is scheduled to be removed from the Solaris OS in an upcoming Solaris release. Solaris UFS logging provides the same capabilities but superior performance, as well as lower system administration requirements and overhead.

 

    **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/`.

**5. On any node in the cluster, run the `sccheck(1M)` utility.**

The `sccheck` utility verifies that the mount points exist. The utility also verifies that `/etc/vfstab` file entries are correct on all nodes of the cluster.

```
# sccheck
```

If no errors occur, nothing is returned.

**6. Mount the cluster file system.**

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

   ■  **For UFS and QFS, mount the cluster file system from any node in the cluster.**

   ■  **SPARC: For VxFS, mount the cluster file system from the current master of *device-group* to ensure that the file system mounts successfully. In addition, unmount a VxFS file system from the current master of *device-group* to ensure that the file system unmounts successfully.**

---

**Note –** 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.

---

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

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

**Example 2–4**   Creating a Cluster File System

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

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

     (on each node)
# mkdir -p /global/oracle/d1
# vi /etc/vfstab
#device          device          mount   FS      fsck    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 /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 2000
```

**Next Steps**   If you installed Sun Cluster software on the Solaris 8 OS or you used SunPlex Installer to install the cluster, go to "How to Configure Internet Protocol (IP) Network Multipathing Groups" on page 125.

If you want to change any private hostnames, go to "How to Change Private Hostnames" on page 126.

If you did not install your own /etc/inet/ntp.conf file before you installed Sun Cluster software, install or create the NTP configuration file. Go to "How to Configure Network Time Protocol (NTP)" on page 127.

SPARC: If you want to configure Sun Management Center to monitor the cluster, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 130.

Otherwise, install third-party applications, register resource types, set up resource groups, and configure data services. Follow procedures in the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* and in the documentation that is supplied with your application software.

## ▼ How to Configure Internet Protocol (IP) Network Multipathing Groups

Perform this task on each node of the cluster. If you used SunPlex Installer to install Sun Cluster HA for Apache or Sun Cluster HA for NFS, SunPlex Installer configured IP Network Multipathing groups for the public-network adapters those data services use. You must configure IP Network Multipathing groups for the remaining public-network adapters.

---

**Note –** All public-network adapters *must* belong to an IP Network Multipathing group.

---

**Before You Begin**   Have available your completed "Public Networks Worksheet" on page 290.

**Step**   ● **Configure IP Network Multipathing groups.**

  ■ Perform procedures in "Deploying Network Multipathing" in *IP Network Multipathing Administration Guide* (Solaris 8), "Configuring Multipathing Interface Groups" in *System Administration Guide: IP Services* (Solaris 9), or "Configuring IPMP Groups" in *System Administration Guide: IP Services* (Solaris 10).

  ■ Follow these additional requirements to configure IP Network Multipathing groups in a Sun Cluster configuration:

    ■ Each public network adapter must belong to a multipathing group.

    ■ In the following kinds of multipathing groups, you must configure a test IP address for each adapter in the group:

      ■ On the Solaris 8 OS, all multipathing groups require a test IP address for each adapter.

      ■ On the Solaris 9 or Solaris 10 OS, multipathing groups that contain two or more adapters require test IP addresses. If a multipathing group contains only one adapter, you do not need to configure a test IP address.

    ■ 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 the test IP addresses are not highly available.

- In the /etc/default/mpathd file, the value of TRACK_INTERFACES_ONLY_WITH_GROUPS must be yes.
- The name of a multipathing group has no requirements or restrictions.

**Next Steps**   If you want to change any private hostnames, go to .

If you did not install your own /etc/inet/ntp.conf file before you installed Sun Cluster software, install or create the NTP configuration file. Go to .

If you are using Sun Cluster on a SPARC based system and you want to use Sun Management Center to monitor the cluster, install the Sun Cluster module for Sun Management Center. Go to .

Otherwise, install third-party applications, register resource types, set up resource groups, and configure data services. Follow procedures in the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* and in the documentation that is supplied with your application software.

## ▼ How to Change Private Hostnames

Perform this task if you do not want to use the default private hostnames, clusternode*nodeid*-priv, that are assigned during Sun Cluster software installation.

---

**Note –** Do *not* perform this procedure after applications and data services have been configured and have been started. Otherwise, an application or data service might continue to use the old private hostname after the hostname is renamed, which would cause hostname conflicts. If any applications or data services are running, stop them before you perform this procedure.

---

Perform this procedure on one active node of the cluster.

**Steps**   1. **Become superuser on a node in the cluster.**

2. **Start the scsetup(1M) utility.**

   # **scsetup**

3. **From the Main Menu, choose the menu item, Private hostnames.**

4. **From the Private Hostname Menu, choose the menu item, Change a private hostname.**

5. **Follow the prompts to change the private hostname.**

   Repeat for each private hostname to change.

6. **Verify the new private hostnames.**

   ```
   # scconf -pv | grep "private hostname"
   (phys-schost-1) Node private hostname:        phys-schost-1-priv
   (phys-schost-3) Node private hostname:        phys-schost-3-priv
   (phys-schost-2) Node private hostname:        phys-schost-2-priv
   ```

**Next Steps**   If you did not install your own /etc/inet/ntp.conf file before you installed Sun Cluster software, install or create the NTP configuration file. Go to "How to Configure Network Time Protocol (NTP)" on page 127.

SPARC: If you want to configure Sun Management Center to monitor the cluster, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 130.

Otherwise, install third-party applications, register resource types, set up resource groups, and configure data services. See the documentation that is supplied with the application software and the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

## ▼ How to Configure Network Time Protocol (NTP)

---

**Note –** If you installed your own /etc/inet/ntp.conf file **before** you installed Sun Cluster software, you do not need to perform this procedure. Determine your next step:

- SPARC: If you want to configure Sun Management Center to monitor the cluster, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 130.
- Otherwise, install third-party applications, register resource types, set up resource groups, and configure data services. See the documentation that is supplied with the application software and the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

---

Perform this task to create or modify the NTP configuration file after you perform any of the following tasks:

- Install Sun Cluster software
- Add a node to an existing cluster

- Change the private hostname of a node in the cluster

If you added a node to a single-node cluster, you must ensure that the NTP configuration file that you use is copied to the original cluster node as well as to the new node.

The primary requirement when you configure NTP, or any time synchronization facility within the cluster, is that all cluster nodes must be synchronized to the same time. Consider accuracy of time on individual nodes to be of secondary importance to the synchronization of time among nodes. You are free to configure NTP as best meets your individual needs if this basic requirement for synchronization is met.

See the *Sun Cluster Concepts Guide for Solaris OS* for further information about cluster time. See the `/etc/inet/ntp.cluster` template file for additional guidelines on how to configure NTP for a Sun Cluster configuration.

**Steps**  1. **Become superuser on a cluster node.**

2. **If you have your own file, copy your file to each node of the cluster.**

3. **If you do not have your own `/etc/inet/ntp.conf` file to install, use the `/etc/inet/ntp.conf.cluster` file as your NTP configuration file.**

---

**Note –** Do not rename the `ntp.conf.cluster` file as `ntp.conf`.

---

If the `/etc/inet/ntp.conf.cluster` file does not exist on the node, you might have an `/etc/inet/ntp.conf` file from an earlier installation of Sun Cluster software. Sun Cluster software creates the `/etc/inet/ntp.conf.cluster` file as the NTP configuration file if an `/etc/inet/ntp.conf` file is not already present on the node. If so, perform the following edits instead on that `ntp.conf` file.

a. **Use your preferred text editor to open the `/etc/inet/ntp.conf.cluster` file on one node of the cluster for editing.**

b. **Ensure that an entry exists for the private hostname of each cluster node.**

   If you changed any node's private hostname, ensure that the NTP configuration file contains the new private hostname.

c. **If necessary, make other modifications to meet your NTP requirements.**

4. **Copy the NTP configuration file to all nodes in the cluster.**
   The contents of the NTP configuration file must be identical on all cluster nodes.

5. **Stop the NTP daemon on each node.**

Wait for the command to complete successfully on each node before you proceed to Step 6.

- For the Solaris 8 or Solaris 9 OS, use the following command:

  # **/etc/init.d/xntpd stop**

- For the Solaris 10 OS, use the following command:

  # **svcadm disable ntp**

6. **Restart the NTP daemon on each node.**

   - **If you use the `ntp.conf.cluster` file, run the following command:**

     # **/etc/init.d/xntpd.cluster start**

     The xntpd.cluster startup script first looks for the /etc/inet/ntp.conf file.

     - If the ntp.conf file exists, the script exits immediately without starting the NTP daemon.
     - If the ntp.conf file does not exist but the ntp.conf.cluster file does exist, the script starts the NTP daemon. In this case, the script uses the ntp.conf.cluster file as the NTP configuration file.

   - **If you use the `ntp.conf` file, run one of the following commands:**

     - For the Solaris 8 or Solaris 9 OS, use the following command:

       # **/etc/init.d/xntpd start**

     - For the Solaris 10 OS, use the following command:

       # **svcadm enable ntp**

**Next Steps**   SPARC: To configure Sun Management Center to monitor the cluster, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 130.

Otherwise, install third-party applications, register resource types, set up resource groups, and configure data services. See the documentation that is supplied with the application software and the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

# SPARC: Installing the Sun Cluster Module for Sun Management Center

This section provides information and procedures to install software for the Sun Cluster module to Sun Management Center.

The Sun Cluster module for Sun Management Center enables you to use Sun Management Center to monitor the cluster. The following table lists the tasks to perform to install the Sun Cluster–module software for Sun Management Center.

**TABLE 2–6** Task Map: Installing the Sun Cluster Module for Sun Management Center

| Task | Instructions |
|---|---|
| 1. Install Sun Management Center server, help-server, agent, and console packages. | Sun Management Center documentation<br><br>"SPARC: Installation Requirements for Sun Cluster Monitoring" on page 130 |
| 2. Install Sun Cluster–module packages. | "SPARC: How to Install the Sun Cluster Module for Sun Management Center" on page 131 |
| 3. Start Sun Management Center server, console, and agent processes. | "SPARC: How to Start Sun Management Center" on page 132 |
| 4. Add each cluster node as a Sun Management Center agent host object. | "SPARC: How to Add a Cluster Node as a Sun Management Center Agent Host Object" on page 133 |
| 5. Load the Sun Cluster module to begin to monitor the cluster. | "SPARC: How to Load the Sun Cluster Module" on page 134 |

## SPARC: Installation Requirements for Sun Cluster Monitoring

The Sun Cluster module for Sun Management Center is used to monitor a Sun Cluster configuration. Perform the following tasks before you install the Sun Cluster module packages.

■ **Space requirements** - Ensure that 25 Mbytes of space is available on each cluster node for Sun Cluster–module packages.

■ **Sun Management Center installation** - Follow procedures in your Sun Management Center installation documentation to install Sun Management Center software.

The following are additional requirements for a Sun Cluster configuration:

- Install the Sun Management Center agent package on each cluster node.
- When you install Sun Management Center on an agent machine (cluster node), choose whether to use the default of 161 for the agent (SNMP) communication port or another number. This port number enables the server to communicate with this agent. Record the port number that you choose for reference later when you configure the cluster nodes for monitoring.

  See your Sun Management Center installation documentation for information about choosing an SNMP port number.
- Install the Sun Management Center server, help–server, and console packages on noncluster nodes.
- If you have an administrative console or other dedicated machine, you can run the console process on the administrative console and the server process on a separate machine. This installation approach improves Sun Management Center performance.
- **Web browser** - Ensure that the web browser that you use to connect to Sun Management Center is supported by Sun Management Center. Certain features, such as online help, might not be available on unsupported web browsers. See your Sun Management Center documentation for information about supported web browsers and any configuration requirements.

## ▼ SPARC: How to Install the Sun Cluster Module for Sun Management Center

Perform this procedure to install the Sun Cluster–module server and help–server packages.

---

**Note –** The Sun Cluster–module agent packages, SUNWscsal and SUNWscsam, are already added to cluster nodes during Sun Cluster software installation.

---

**Before You Begin**   Ensure that all Sun Management Center core packages are installed on the appropriate machines. This task includes installing Sun Management Center agent packages on each cluster node. See your Sun Management Center documentation for installation instructions.

**Steps**   1. **On the server machine, install the Sun Cluster–module server package SUNWscssv.**

   a. **Become superuser.**

   b. **Insert the Sun Cluster 2 of 2 CD-ROM for the SPARC platform in the CD-ROM drive.**

If the volume management daemon vold(1M) is running and is configured to manage CD-ROM devices, the daemon automatically mounts the CD-ROM on the /cdrom/cdrom0/ directory.

**c. Change to the Solaris_sparc/Product/sun_cluster/Solaris_*ver*/Packages/ directory, where *ver* is 8 for Solaris 8, 9 for Solaris 9, or 10 for Solaris 10.**

```
# cd /cdrom/cdrom0/Solaris_sparc/Product/sun_cluster/Solaris_ver/Packages/
```

**d. Install the Sun Cluster–module server package.**

```
# pkgadd -d . SUNWscssv
```

**e. Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

```
# eject cdrom
```

**2. On the Sun Management Center 3.0 help-server machine or the Sun Management Center 3.5 server machine, install the Sun Cluster–module help–server package SUNWscshl.**

Use the same procedure as in the previous step.

**3. Install any Sun Cluster–module patches.**

See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

**Next Steps**  Start Sun Management Center. Go to "SPARC: How to Start Sun Management Center" on page 132.

## ▼ SPARC: How to Start Sun Management Center

Perform this procedure to start the Sun Management Center server, agent, and console processes.

**Steps**  **1. As superuser, on the Sun Management Center server machine, start the Sun Management Center server process.**

The *install-dir* is the directory on which you installed the Sun Management Center software. The default directory is /opt.

```
# /install-dir/SUNWsymon/sbin/es-start -S
```

**2. As superuser, on each Sun Management Center agent machine (cluster node), start the Sun Management Center agent process.**

```
# /install-dir/SUNWsymon/sbin/es-start -a
```

3.  **On each Sun Management Center agent machine (cluster node), ensure that the `scsymon_srv` daemon is running.**

    `# `**`ps -ef | grep scsymon_srv`**

    If any cluster node is not already running the scsymon_srv daemon, start the daemon on that node.

    `# `**`/usr/cluster/lib/scsymon/scsymon_srv`**

4.  **On the Sun Management Center console machine (administrative console), start the Sun Management Center console.**

    You do not need to be superuser to start the console process.

    `% /`*install-dir*`/`**`SUNWsymon/sbin/es-start -c`**

**Next Steps**  Add a cluster node as a monitored host object. Go to "SPARC: How to Add a Cluster Node as a Sun Management Center Agent Host Object" on page 133.

## ▼ SPARC: How to Add a Cluster Node as a Sun Management Center Agent Host Object

Perform this procedure to create a Sun Management Center agent host object for a cluster node.

**Steps**  1.  **Log in to Sun Management Center.**

    See your Sun Management Center documentation.

2.  **From the Sun Management Center main window, select a domain from the Sun Management Center Administrative Domains pull-down list.**

    This domain contains the Sun Management Center agent host object that you create. During Sun Management Center software installation, a Default Domain was automatically created for you. You can use this domain, select another existing domain, or create a new domain.

    See your Sun Management Center documentation for information about how to create Sun Management Center domains.

3.  **Choose Edit⇒Create an Object from the pull-down menu.**

4.  **Click the Node tab.**

5.  **From the Monitor Via pull-down list, select Sun Management Center Agent - Host.**

6.  **Fill in the name of the cluster node, for example, `phys-schost-1`, in the Node Label and Hostname text fields.**

    Leave the IP text field blank. The Description text field is optional.

7. **In the Port text field, type the port number that you chose when you installed the Sun Management Center agent machine.**

8. **Click OK.**

   A Sun Management Center agent host object is created in the domain.

**Next Steps**  Load the Sun Cluster module. Go to "SPARC: How to Load the Sun Cluster Module" on page 134.

**Troubleshooting**  You need only one cluster node host object to use Sun Cluster–module monitoring and configuration functions for the entire cluster. However, if that cluster node becomes unavailable, connection to the cluster through that host object also becomes unavailable. Then you need another cluster-node host object to reconnect to the cluster.

## ▼  SPARC: How to Load the Sun Cluster Module

Perform this procedure to start cluster monitoring.

**Steps**  1. **In the Sun Management Center main window, right click the icon of a cluster node.**

   The pull-down menu is displayed.

2. **Choose Load Module.**

   The Load Module window lists each available Sun Management Center module and whether the module is currently loaded.

3. **Choose Sun Cluster: Not Loaded and click OK.**

   The Module Loader window shows the current parameter information for the selected module.

4. **Click OK.**

   After a few moments, the module is loaded. A Sun Cluster icon is then displayed in the Details window.

5. **Verify that the Sun Cluster module is loaded.**

   Under the Operating System category, expand the Sun Cluster subtree in either of the following ways:

   - **In the tree hierarchy on the left side of the window, place the cursor over the Sun Cluster module icon and single-click the left mouse button.**

   - **In the topology view on the right side of the window, place the cursor over the Sun Cluster module icon and double-click the left mouse button.**

**See Also** See the Sun Cluster module online help for information about how to use Sun Cluster module features.

- To view online help for a specific Sun Cluster module item, place the cursor over the item. Then click the right mouse button and select Help from the pop-up menu.
- To access the home page for the Sun Cluster module online help, place the cursor over the Cluster Info icon. Then click the right mouse button and select Help from the pop-up menu.
- To directly access the home page for the Sun Cluster module online help, click the Sun Management Center Help button to launch the help browser. Then go to the following URL, where *install-dir* is the directory on which you installed the Sun Management Center software:

  `file:/`*install-dir*`/SUNWsymon/lib/locale/C/help/main.top.html`

---

**Note –** The Help button in the Sun Management Center browser accesses online help for Sun Management Center, not the topics specific to the Sun Cluster module.

---

See Sun Management Center online help and your Sun Management Center documentation for information about how to use Sun Management Center.

**Next Steps** Install third-party applications, register resource types, set up resource groups, and configure data services. See the documentation that is supplied with the application software and the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

---

# Uninstalling the Software

This section provides the following procedures to uninstall or remove Sun Cluster software:

- "How to Uninstall Sun Cluster Software to Correct Installation Problems" on page 136
- "How to Uninstall the `SUNWscrdt` Package" on page 137
- "How to Unload the RSMRDT Driver Manually" on page 138

# ▼ How to Uninstall Sun Cluster Software to Correct Installation Problems

Perform this procedure if the installed node cannot join the cluster or if you need to correct configuration information. For example, perform this procedure to reconfigure the transport adapters or the private-network address.

---

**Note –** If the node has already joined the cluster and is no longer in installation mode, as described in Step 2 of "How to Verify the Quorum Configuration and Installation Mode" on page 117, do not perform this procedure. Instead, go to "How to Uninstall Sun Cluster Software From a Cluster Node" in "Adding and Removing a Cluster Node" in *Sun Cluster System Administration Guide for Solaris OS*.

---

**Before You Begin**  Attempt to reinstall the node. You can correct certain failed installations by repeating Sun Cluster software installation on the node.

**Steps**  1. **Add to the cluster's node-authentication list the node that you intend to uninstall.**

   If you are uninstalling a single-node cluster, skip to Step 2.

   a. **Become superuser on an active cluster member other than the node that you are uninstalling.**

   b. **Specify the name of the node to add to the authentication list.**

      # **/usr/cluster/bin/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

      You can also use the scsetup(1M) utility to perform this task. See "How to Add a Node to the Authorized Node List" in *Sun Cluster System Administration Guide for Solaris OS* for procedures.

   2. **Become superuser on the node that you intend to uninstall.**

   3. **Shut down the node that you intend to uninstall.**

      # **shutdown -g0 -y -i0**

   4. **Reboot the node into noncluster mode.**

      ■ **On SPARC based systems, do the following:**

         ok **boot -x**

■ **On x86 based systems, do the following:**

```
                        <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci-ide@7,1/ata@1/cmdk@0,0:b
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. **Change to a directory, such as the root (/) directory, that does not contain any files that are delivered by the Sun Cluster packages.**

   `# cd /`

6. **Uninstall Sun Cluster software from the node.**

   `# /usr/cluster/bin/scinstall -r`

   See the scinstall(1M) man page for more information.

7. **Reinstall and reconfigure Sun Cluster software on the node.**

   Refer to Table 2–1 for the list of all installation tasks and the order in which to perform the tasks.

## ▼ How to Uninstall the SUNWscrdt Package

Perform this procedure on each node in the cluster.

**Before You Begin**   Verify that no applications are using the RSMRDT driver before performing this procedure.

**Steps**   1. **Become superuser on the node to which you want to uninstall the SUNWscrdt package.**

2. **Uninstall the SUNWscrdt package.**

   `# pkgrm SUNWscrdt`

## ▼ How to Unload the RSMRDT Driver Manually

If the driver remains loaded in memory after completing "How to Uninstall the SUNWscrdt Package" on page 137, perform this procedure to unload the driver manually.

**Steps**   **1.  Start the `adb` utility.**

```
# adb -kw
```

**2.  Set the kernel variable `clifrsmrdt_modunload_ok` to 1.**

physmem *NNNN*
**clifrsmrdt_modunload_ok/W 1**

**3.  Exit the `adb` utility by pressing Control-D.**

**4.  Find the `clif_rsmrdt` and `rsmrdt` module IDs.**

```
# modinfo | grep rdt
```

**5.  Unload the `clif_rsmrdt` module.**

You must unload the clif_rsmrdt module before you unload the rsmrdt module.

```
# modunload -i clif_rsmrdt_id
```

*clif_rsmrdt_id*       Specifies the numeric ID for the module being unloaded

**6.  Unload the `rsmrdt` module.**

```
# modunload -i rsmrdt_id
```

*rsmrdt_id*       Specifies the numeric ID for the module being unloaded

**7.  Verify that the module was successfully unloaded.**

```
# modinfo | grep rdt
```

**Example 2–5**   Unloading the RSMRDT Driver

The following example shows the console output after the RSMRDT driver is manually unloaded.

```
# adb -kw
physmem fc54
clifrsmrdt_modunload_ok/W 1
clifrsmrdt_modunload_ok: 0x0 = 0x1
^D
# modinfo | grep rsm
 88 f064a5cb 974 - 1 rsmops (RSMOPS module 1.1)
 93 f08e07d4 b95 - 1 clif_rsmrdt (CLUSTER-RSMRDT Interface module)
 94 f0d3d000 13db0 194 1 rsmrdt (Reliable Datagram Transport dri)
```

```
# modunload -i 93
# modunload -i 94
# modinfo | grep rsm
 88 f064a5cb 974 - 1 rsmops (RSMOPS module 1.1)
#
```

**Troubleshooting** If the `modunload` command fails, applications are probably still using the driver.
Terminate the applications before you run `modunload` again.

# Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software

Install and configure your local and multihost disks for Solstice DiskSuite or Solaris Volume Manager software by using the procedures in this chapter, along with the planning information in "Planning Volume Management" on page 35. See your Solstice DiskSuite or Solaris Volume Manager documentation for additional details.

---

**Note –** DiskSuite Tool (Solstice DiskSuite `metatool`) and the Enhanced Storage module of Solaris Management Console (Solaris Volume Manager) are not compatible with Sun Cluster software. Use the command-line interface or Sun Cluster utilities to configure Solstice DiskSuite or Solaris Volume Manager software.

---

The following sections are in this chapter:

---

# Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software

This section provides information and procedures to install and configure Solstice DiskSuite or Solaris Volume Manager software. You can skip certain procedures under the following conditions:

- If you used SunPlex Installer to install Solstice DiskSuite software (Solaris 8), the procedures "How to Install Solstice DiskSuite Software" on page 143 through "How to Create State Database Replicas" on page 147 are already completed. Go to "Mirroring the Root Disk" on page 148 or "Creating Disk Sets in a Cluster" on page 163 to continue to configure Solstice DiskSuite software.
- If you installed Solaris 9 or Solaris 10 software, Solaris Volume Manager is already installed. You can start configuration at "How to Set the Number of Metadevice or Volume Names and Disk Sets" on page 145.

The following table lists the tasks that you perform to install and configure Solstice DiskSuite or Solaris Volume Manager software for Sun Cluster configurations.

**TABLE 3–1** Task Map: Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software

| Task | Instructions |
|------|-------------|
| 1. Plan the layout of your Solstice DiskSuite or Solaris Volume Manager configuration. | "Planning Volume Management" on page 35 |
| 2. (*Solaris 8 only*) Install Solstice DiskSuite software. | "How to Install Solstice DiskSuite Software" on page 143 |
| 3. (*Solaris 8 and Solaris 9 only*) Calculate the number of metadevice names and disk sets needed for your configuration, and modify the /kernel/drv/md.conf file. | "How to Set the Number of Metadevice or Volume Names and Disk Sets" on page 145 |
| 4. Create state database replicas on the local disks. | "How to Create State Database Replicas" on page 147 |
| 5. (*Optional*) Mirror file systems on the root disk. | "Mirroring the Root Disk" on page 148 |

# ▼ How to Install Solstice DiskSuite Software

---

**Note –** Do not perform this procedure under the following circumstances:

- You installed Solaris 9 software. Solaris Volume Manager software is automatically installed with Solaris 9 software. Instead, go to "How to Set the Number of Metadevice or Volume Names and Disk Sets" on page 145.

- You installed Solaris 10 software. Instead, go to "How to Create State Database Replicas" on page 147.

- You used SunPlex Installer to install Solstice DiskSuite software. Instead, do one of the following:

  - If you plan to create additional disk sets, go to "How to Set the Number of Metadevice or Volume Names and Disk Sets" on page 145.
  - If you do not plan to create additional disk sets, go to "Mirroring the Root Disk" on page 148.

---

Perform this task on each node in the cluster.

**Before You Begin**    Perform the following tasks:

- Make mappings of your storage drives.

- Complete the following configuration planning worksheets. See "Planning Volume Management" on page 35 for planning guidelines.

  - "Local File System Layout Worksheet" on page 288
  - "Disk Device Group Configurations Worksheet" on page 294
  - "Volume-Manager Configurations Worksheet" on page 296
  - "Metadevices Worksheet (Solstice DiskSuite or Solaris Volume Manager)" on page 298

**Steps**    1. **Become superuser on the cluster node.**

2. **If you install from the CD-ROM, insert the Solaris 8 Software 2 of 2 CD-ROM in the CD-ROM drive on the node.**

   This step assumes that the Volume Management daemon `vold`(1M) is running and configured to manage CD-ROM devices.

3. **Install the Solstice DiskSuite software packages.**

   Install the packages in the order that is shown in the following example.

```
# cd /cdrom/sol_8_sparc_2/Solaris_8/EA/products/DiskSuite_4.2.1/sparc/Packages
# pkgadd -d . SUNWmdr SUNWmdu [SUNWmdx] optional-pkgs
```

   - The `SUNWmdr` and `SUNWmdu` packages are required for all Solstice DiskSuite installations.

- The `SUNWmdx` package is also required for the 64-bit Solstice DiskSuite installation.

- See your Solstice DiskSuite installation documentation for information about optional software packages.

---

**Note –** If you have Solstice DiskSuite software patches to install, do not reboot after you install the Solstice DiskSuite software.

---

**4. If you installed from a CD-ROM, eject the CD-ROM.**

**5. Install any Solstice DiskSuite patches.**

See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

**6. Repeat Step 1 through Step 5 on each of the other nodes of the cluster.**

**7. From one node of the cluster, manually populate the global-device namespace for Solstice DiskSuite.**

```
# scgdevs
```

**Next Steps**  If you used SunPlex Installer to install Solstice DiskSuite software, go to "Mirroring the Root Disk" on page 148.

If the cluster runs on the Solaris 10 OS, go to "How to Create State Database Replicas" on page 147.

Otherwise, go to "How to Set the Number of Metadevice or Volume Names and Disk Sets" on page 145.

**Troubleshooting** The scgdevs command might return a message similar to the message Could not open /dev/rdsk/c0t6d0s2 to verify device id, Device busy. If the listed device is a CD-ROM device, you can safely ignore the message.

## ▼ How to Set the Number of Metadevice or Volume Names and Disk Sets

---

**Note –** Do not perform this procedure. in the following circumstances:

- The cluster runs on the Solaris 10 OS. Instead, go to "How to Create State Database Replicas" on page 147.

  With the Solaris 10 release, Solaris Volume Manager has been enhanced to configure volumes dynamically. You no longer need to edit the nmd and the md_nsets parameters in the /kernel/drv/md.conf file. New volumes are dynamically created, as needed.

- You used SunPlex Installer to install Solstice DiskSuite software. Instead, go to "Mirroring the Root Disk" on page 148.

---

This procedure describes how to determine the number of Solstice DiskSuite metadevice or Solaris Volume Manager volume names and disk sets that are needed for your configuration. This procedure also describes how to modify the /kernel/drv/md.conf file to specify these numbers.

---

**Tip –** The default number of metadevice or volume names per disk set is 128, but many configurations need more than the default. Increase this number before you implement a configuration, to save administration time later.

At the same time, keep the value of the nmdfield and the md_nsets field as low as possible. Memory structures exist for all possible devices as determined by nmdand md_nsets, even if you have not created those devices. For optimal performance, keep the value of nmd and md_nsets only slightly higher than the number of metadevices or volumes you plan to use.

---

**Before You Begin**  Have available the completed "Disk Device Group Configurations Worksheet" on page 294.

**Steps**  **1. Calculate the total number of disk sets that you expect to need in the cluster, then add one more disk set for private disk management.**

The cluster can have a maximum of 32 disk sets, 31 disk sets for general use plus one disk set for private disk management. The default number of disk sets is 4. You supply this value for the md_nsets field in Step 3.

**2. Calculate the largest metadevice or volume name that you expect to need for any disk set in the cluster.**

Each disk set can have a maximum of 8192 metadevice or volume names. You supply this value for the nmd field in Step 3.

a. **Determine the quantity of metadevice or volume names that you expect to need for each disk set.**

If you use local metadevices or volumes, ensure that each local metadevice or volume name on which a global-devices file system, `/global/.devices/node@` *nodeid*, is mounted is unique throughout the cluster and does not use the same name as any device-ID name in the cluster.

---

**Tip –** Choose a range of numbers to use exclusively for device-ID names and a range for each node to use exclusively for its local metadevice or volume names. For example, device-ID names might use the range from `d1` to `d100`. Local metadevices or volumes on node 1 might use names in the range from `d100` to `d199`. And local metadevices or volumes on node 2 might use `d200` to `d299`.

---

b. **Calculate the highest of the metadevice or volume names that you expect to use in any disk set.**

The quantity of metadevice or volume names to set is based on the metadevice or volume name *value* rather than on the *actual quantity* . For example, if your metadevice or volume names range from `d950` to `d1000`, Solstice DiskSuite or Solaris Volume Manager software requires that you set the value at 1000 names, not 50.

3. **On each node, become superuser and edit the `/kernel/drv/md.conf` file.**

---

**Caution –** All cluster nodes (or cluster pairs in the cluster-pair topology) must have identical `/kernel/drv/md.conf` files, regardless of the number of disk sets served by each node. Failure to follow this guideline can result in serious Solstice DiskSuite or Solaris Volume Manager errors and possible loss of data.

---

a. **Set the `md_nsets` field to the value that you determined in Step 1.**

b. **Set the `nmd` field to the value that you determined in Step 2.**

4. **On each node, perform a reconfiguration reboot.**

```
# touch /reconfigure
# shutdown -g0 -y -i6
```

Changes to the /kernel/drv/md.conf file become operative after you perform a reconfiguration reboot.

**Next Steps** Create local state database replicas. Go to "How to Create State Database Replicas" on page 147.

## ▼ How to Create State Database Replicas

Perform this procedure on each node in the cluster.

**Steps**   1. **Become superuser on the cluster node.**

2. **Create state database replicas on one or more local devices for each cluster node.**
   Use the physical name (c*N*t*X*d*Y* s*Z*), not the device-ID name (d*N*), to specify the slices to use.

   # **metadb -af** *slice-1 slice-2 slice-3*

   **Tip –** To provide protection of state data, which is necessary to run Solstice DiskSuite or Solaris Volume Manager software, create at least three replicas for each node. Also, you can place replicas on more than one device to provide protection if one of the devices fails.

   See the metadb(1M) man page and your Solstice DiskSuite or Solaris Volume Manager documentation for details.

3. **Verify the replicas.**

   # **metadb**
   The metadb command displays the list of replicas.

**Example 3–1**   Creating State Database Replicas

The following example shows three Solstice DiskSuite state database replicas. Each replica is created on a different device. For Solaris Volume Manager, the replica size would be larger.

```
# metadb -af c0t0d0s7 c0t1d0s7 c1t0d0s7
# metadb
flags              first blk       block count
    a       u      16              1034           /dev/dsk/c0t0d0s7
    a       u      16              1034           /dev/dsk/c0t1d0s7
    a       u      16              1034           /dev/dsk/c1t0d0s7
```

**Next Steps**   To mirror file systems on the root disk, go to "Mirroring the Root Disk" on page 148.

Otherwise, go to "Creating Disk Sets in a Cluster" on page 163 to create Solstice DiskSuite or Solaris Volume Manager disk sets.

# Mirroring the Root Disk

Mirroring the root disk prevents the cluster node itself from shutting down because of a system disk failure. Four types of file systems can reside on the root disk. Each file-system type is mirrored by using a different method.

Use the following procedures to mirror each type of file system.

- "How to Mirror the Root (/) File System" on page 148
- "How to Mirror the Global Namespace" on page 152
- "How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted" on page 155
- "How to Mirror File Systems That Can Be Unmounted" on page 159

> **Caution –** For local disk mirroring, do not use /dev/global as the path when you specify the disk name. If you specify this path for anything other than cluster file systems, the system cannot boot.

## ▼ How to Mirror the Root (/) File System

Use this procedure to mirror the root (/) file system.

**Steps** 1. **Become superuser on the node.**

2. **Place the root slice in a single-slice (one-way) concatenation.**
   Specify the physical disk name of the root-disk slice (c*N*t*X*d*Y* s*Z*).

   # **metainit -f** *submirror1* **1 1** *root-disk-slice*

3. **Create a second concatenation.**

   # **metainit** *submirror2* **1 1** *submirror-disk-slice*

4. **Create a one-way mirror with one submirror.**

   # **metainit** *mirror* **-m** *submirror1*

   > **Note –** If the device is a local device to be used to mount a global-devices file system, /global/.devices/node@*nodeid*, the metadevice or volume name for the mirror *must* be unique throughout the cluster.

5. **Run the metaroot(1M) command.**
   This command edits the /etc/vfstab and /etc/system files so the system can be booted with the root (/) file system on a metadevice or volume.

   # **metaroot** *mirror*

6. **Run the `lockfs(1M)` command.**

   This command flushes all transactions out of the log and writes the transactions to the master file system on all mounted UFS file systems.

   ```
   # lockfs -fa
   ```

7. **Move any resource groups or device groups from the node.**

   ```
   # scswitch -S -h from-node
   ```

   | | |
   |---|---|
   | -S | Moves all resource groups and device groups |
   | -h from-node | Specifies the name of the node from which to move resource or device groups |

8. **Reboot the node.**

   This command remounts the newly mirrored root (/) file system.

   ```
   # shutdown -g0 -y -i6
   ```

9. **Use the `metattach(1M)` command to attach the second submirror to the mirror.**

   ```
   # metattach mirror submirror2
   ```

10. **If the disk that is used to mirror the root disk is physically connected to more than one node (multihosted), enable the `localonly` property.**

    Perform the following steps to enable the localonly property of the raw-disk device group for the disk that is used to mirror the root disk. You must enable the localonly property to prevent unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

    a. **If necessary, use the `scdidadm(1M) -L` command to display the full device-ID path name of the raw-disk device group.**

       In the following example, the raw-disk device-group name dsk/d2 is part of the third column of output, which is the full device-ID path name.

       ```
       # scdidadm -L
       ...
       1          phys-schost-3:/dev/rdsk/c1t1d0     /dev/did/rdsk/d2
       ```

    b. **View the node list of the raw-disk device group.**

       Output looks similar to the following:

       ```
       # scconf -pvv | grep dsk/d2
       Device group name:                      dsk/d2
       ...
         (dsk/d2) Device group node list:      phys-schost-1, phys-schost-3
       ...
       ```

    c. **If the node list contains more than one node name, remove all nodes from the node list except the node whose root disk you mirrored.**

Only the node whose root disk you mirrored should remain in the node list for the raw-disk device group.

```
# scconf -r -D name=dsk/dN,nodelist=node
```

-D name=dsk/dN        Specifies the cluster-unique name of the raw-disk device group

nodelist=node        Specifies the name of the node or nodes to remove from the node list

   **d. Use the `scconf(1M)` command to enable the `localonly` property.**

   When the `localonly` property is enabled, the raw-disk device group is used exclusively by the node in its node list. This usage prevents unintentional fencing of the node from its boot device if the boot device is connected to multiple nodes.

```
# scconf -c -D name=rawdisk-groupname,localonly=true
```

-D name=rawdisk-groupname        Specifies the name of the raw-disk device group

   For more information about the `localonly` property, see the `scconf_dg_rawdisk(1M)` man page.

**11. Record the alternate boot path for possible future use.**

If the primary boot device fails, you can then boot from this alternate boot device. See Chapter 7, "Troubleshooting the System," in *Solstice DiskSuite 4.2.1 User's Guide*, "Special Considerations for Mirroring root (/)" in *Solaris Volume Manager Administration Guide*, or "Creating a RAID-1 Volume" in *Solaris Volume Manager Administration Guide* for more information about alternate boot devices.

```
# ls -l /dev/rdsk/root-disk-slice
```

**12. Repeat Step 1 through Step 11 on each remaining node of the cluster.**

Ensure that each metadevice or volume name for a mirror on which a global-devices file system, `/global/.devices/node@nodeid`, is to be mounted is unique throughout the cluster.

**Example 3–2**    Mirroring the Root (/) File System

The following example shows the creation of mirror `d0` on the node `phys-schost-1`, which consists of submirror `d10` on partition `c0t0d0s0` and submirror `d20` on partition `c2t2d0s0`. Device `c2t2d0` is a multihost disk, so the `localonly` property is enabled.

```
    (Create the mirror)
# metainit -f d10 1 1 c0t0d0s0
d11: Concat/Stripe is setup
# metainit d20 1 1 c2t2d0s0
d12: Concat/Stripe is setup
```

```
# metainit d0 -m d10
d10: Mirror is setup
# metaroot d0
# lockfs -fa

    (Move resource groups and device groups from phys-schost-1)
# scswitch -S -h phys-schost-1

    (Reboot the node)
# shutdown -g0 -y -i6

    (Attach the second submirror)
# metattach d0 d20
d0: Submirror d20 is attached

    (Display the device-group node list)
# scconf -pvv | grep dsk/d2
Device group name:                        dsk/d2
...
  (dsk/d2) Device group node list:        phys-schost-1, phys-schost-3
...

    (Remove phys-schost-3 from the node list)
# scconf -r -D name=dsk/d2,nodelist=phys-schost-3

    (Enable the localonly property)
# scconf -c -D name=dsk/d2,localonly=true

    (Record the alternate boot path)
# ls -l /dev/rdsk/c2t2d0s0
lrwxrwxrwx  1 root     root          57 Apr 25 20:11 /dev/rdsk/c2t2d0s0
-> ../../devices/node@1/pci@1f,0/pci@1/scsi@3,1/disk@2,0:a,raw
```

**Next Steps**     To mirror the global namespace, /global/.devices/node@*nodeid*, go to "How to Mirror the Global Namespace" on page 152.

To mirror file systems than cannot be unmounted, go to "How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted" on page 155.

To mirror user-defined file systems, go to "How to Mirror File Systems That Can Be Unmounted" on page 159.

Otherwise, go to "Creating Disk Sets in a Cluster" on page 163 to create a disk set.

**Troubleshooting** Some of the steps in this mirroring procedure might cause an error message similar to metainit: dg-schost-1: d1s0: not a metadevice. Such an error message is harmless and can be ignored.

# ▼ How to Mirror the Global Namespace

Use this procedure to mirror the global namespace,
`/global/.devices/node@`*nodeid*`/`.

**Steps** 1. **Become superuser on a node of the cluster.**

2. **Place the global namespace slice in a single-slice (one-way) concatenation.**
   Use the physical disk name of the disk slice (c*N*t*X*d*Y* s*Z*).

   `# `**`metainit -f`** *submirror1* **`1 1`** *diskslice*

3. **Create a second concatenation.**

   `# `**`metainit`** *submirror2* **`1 1`** *submirror-diskslice*

4. **Create a one-way mirror with one submirror.**

   `# `**`metainit`** *mirror* **`-m`** *submirror1*

   ---

   **Note –** The metadevice or volume name for a mirror on which a global-devices file
   system, `/global/.devices/node@`*nodeid*, is to be mounted *must* be unique
   throughout the cluster.

   ---

5. **Attach the second submirror to the mirror.**
   This attachment starts a synchronization of the submirrors.

   `# `**`metattach`** *mirror* *submirror2*

6. **Edit the `/etc/vfstab` file entry for the `/global/.devices/node@`*nodeid* file
   system.**
   Replace the names in the `device to mount` and `device to fsck` columns with
   the mirror name.

```
#
vi /etc/vfstab
#device        device         mount    FS     fsck    mount     mount
#to mount      to fsck        point    type   pass    at boot   options
#
/dev/md/dsk/mirror /dev/md/rdsk/mirror /global/.devices/node@nodeid ufs 2 no global
```

7. **Repeat Step 1 through Step 6 on each remaining node of the cluster.**

8. **Wait for the synchronization of the mirrors, started in Step 5, to be completed.**
   Use the `metastat`(1M) command to view mirror status and to verify that mirror
   synchronization is complete.

   `# `**`metastat`** *mirror*

9. **If the disk that is used to mirror the global namespace is physically connected to more than one node (multihosted), enable the `localonly` property.**

Perform the following steps to enable the `localonly` property of the raw-disk device group for the disk that is used to mirror the global namespace. You must enable the `localonly` property to prevent unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

   a. **If necessary, use the `scdidadm(1M)` command to display the full device-ID path name of the raw-disk device group.**

   In the following example, the raw-disk device-group name dsk/d2 is part of the third column of output, which is the full device-ID path name.

   ```
   # scdidadm -L
   ...
   1          phys-schost-3:/dev/rdsk/c1t1d0     /dev/did/rdsk/d2
   ```

   b. **View the node list of the raw-disk device group.**

   Output looks similar to the following.

   ```
   # scconf -pvv | grep dsk/d2
   Device group name:                      dsk/d2
   ...
     (dsk/d2) Device group node list:      phys-schost-1, phys-schost-3
   ...
   ```

   c. **If the node list contains more than one node name, remove all nodes from the node list except the node whose disk is mirrored.**

   Only the node whose disk is mirrored should remain in the node list for the raw-disk device group.

   ```
   # scconf -r -D name=dsk/dN,nodelist=node
   ```

   | | |
   |---|---|
   | -D name=dsk/dN | Specifies the cluster-unique name of the raw-disk device group |
   | nodelist=node | Specifies the name of the node or nodes to remove from the node list |

   d. **Enable the `localonly` property.**

   When the `localonly` property is enabled, the raw-disk device group is used exclusively by the node in its node list. This usage prevents unintentional fencing of the node from its boot device if the boot device is connected to multiple nodes.

   ```
   # scconf -c -D name=rawdisk-groupname,localonly=true
   ```

   | | |
   |---|---|
   | -D name=rawdisk-groupname | Specifies the name of the raw-disk device group |

   For more information about the `localonly` property, see the `scconf_dg_rawdisk(1M)` man page.

**Example 3–3** Mirroring the Global Namespace

The following example shows creation of mirror d101, which consists of submirror d111 on partition c0t0d0s3 and submirror d121 on partition c2t2d0s3. The /etc/vfstab file entry for /global/.devices/node@1 is updated to use the mirror name d101. Device c2t2d0 is a multihost disk, so the localonly property is enabled.

```
    (Create the mirror)
# metainit -f d111 1 1 c0t0d0s3
d111: Concat/Stripe is setup
# metainit d121 1 1 c2t2d0s3
d121: Concat/Stripe is setup
# metainit d101 -m d111
d101: Mirror is setup
# metattach d101 d121
d101: Submirror d121 is attached

    (Edit the /etc/vfstab file)
# vi /etc/vfstab
#device         device          mount   FS      fsck    mount     mount
#to mount       to fsck         point   type    pass    at boot   options
#
/dev/md/dsk/d101 /dev/md/rdsk/d101 /global/.devices/node@1 ufs 2 no global

    (View the sync status)
# metastat d101
d101: Mirror
      Submirror 0: d111
         State: Okay
      Submirror 1: d121
         State: Resyncing
      Resync in progress: 15 % done
...

    (Identify the device-ID name of the mirrored disk's raw-disk device group)
# scdidadm -L
...
1          phys-schost-3:/dev/rdsk/c2t2d0     /dev/did/rdsk/d2

    (Display the device-group node list)
# scconf -pvv | grep dsk/d2
Device group name:                      dsk/d2
...
  (dsk/d2) Device group node list:      phys-schost-1, phys-schost-3
...

    (Remove phys-schost-3 from the node list)
# scconf -r -D name=dsk/d2,nodelist=phys-schost-3

    (Enable the localonly property)
# scconf -c -D name=dsk/d2,localonly=true
```

**Next Steps** To mirror file systems other than root (/) that cannot be unmounted, go to .

To mirror user-defined file systems, go to

Otherwise, go to to create a disk set.

**Troubleshooting** Some of the steps in this mirroring procedure might cause an error message similar to `metainit: dg-schost-1: d1s0: not a metadevice`. Such an error message is harmless and can be ignored.

## ▼ How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted

Use this procedure to mirror file systems other than root (/) that cannot be unmounted during normal system usage, such as /usr, /opt, or swap.

**Steps** 1. **Become superuser on a node of the cluster.**

2. **Place the slice on which an unmountable file system resides in a single-slice (one-way) concatenation.**

   Specify the physical disk name of the disk slice ($cNtX\ dYsZ$).

   # **metainit -f** *submirror1* **1 1** *diskslice*

3. **Create a second concatenation.**

   # **metainit** *submirror2* **1 1** *submirror-diskslice*

4. **Create a one-way mirror with one submirror.**

   # **metainit** *mirror* **-m** *submirror1*

   ---

   **Note –** The metadevice or volume name for this mirror does *not* need to be unique throughout the cluster.

   ---

5. **Repeat Step 1 through Step 4 for each remaining unmountable file system that you want to mirror.**

6. **On each node, edit the /etc/vfstab file entry for each unmountable file system you mirrored.**

Replace the names in the device to mount and device to fsck columns with the mirror name.

```
# vi /etc/vfstab
#device         device         mount    FS    fsck    mount    mount
#to mount       to fsck        point    type  pass    at boot  options
#
/dev/md/dsk/mirror /dev/md/rdsk/mirror /filesystem ufs 2 no global
```

**7. Move any resource groups or device groups from the node.**

```
# scswitch -S -h from-node
```

-S              Moves all resource groups and device groups

-h from-node    Specifies the name of the node from which to move resource or device groups

**8. Reboot the node.**

```
# shutdown -g0 -y -i6
```

**9. Attach the second submirror to each mirror.**

This attachment starts a synchronization of the submirrors.

```
# metattach mirror submirror2
```

**10. Wait for the synchronization of the mirrors, started in Step 9, to complete.**

Use the metastat(1M) command to view mirror status and to verify that mirror synchronization is complete.

```
# metastat mirror
```

**11. If the disk that is used to mirror the unmountable file system is physically connected to more than one node (is multihosted), enable the `localonly` property.**

Perform the following steps to enable the localonly property of the raw-disk device group for the disk that is used to mirror the unmountable file system. You must enable the localonly property to prevent unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

**a. If necessary, use the `scdidadm -L` command to display the full device-ID path name of the raw-disk device group.**

In the following example, the raw-disk device-group name dsk/d2 is part of the third column of output, which is the full device-ID path name.

```
# scdidadm -L
...
1              phys-schost-3:/dev/rdsk/c1t1d0    /dev/did/rdsk/d2
```

**b. View the node list of the raw-disk device group.**

Output looks similar to the following.

```
# scconf -pvv | grep dsk/d2
Device group name:                       dsk/d2
...
  (dsk/d2) Device group node list:       phys-schost-1, phys-schost-3
...
```

**c. If the node list contains more than one node name, remove all nodes from the node list except the node whose root disk is mirrored.**

Only the node whose root disk is mirrored should remain in the node list for the raw-disk device group.

```
# scconf -r -D name=dsk/dN,nodelist=node
```

-D name=dsk/dN    Specifies the cluster-unique name of the raw-disk device group

nodelist=node     Specifies the name of the node or nodes to remove from the node list

**d. Enable the `localonly` property.**

When the localonly property is enabled, the raw-disk device group is used exclusively by the node in its node list. This usage prevents unintentional fencing of the node from its boot device if the boot device is connected to multiple nodes.

```
# scconf -c -D name=rawdisk-groupname,localonly=true
```

-D name=rawdisk-groupname    Specifies the name of the raw-disk device group

For more information about the localonly property, see the scconf_dg_rawdisk(1M) man page.

**Example 3–4    Mirroring File Systems That Cannot Be Unmounted**

The following example shows the creation of mirror d1 on the node phys-schost-1 to mirror /usr, which resides on c0t0d0s1. Mirror d1 consists of submirror d11 on partition c0t0d0s1 and submirror d21 on partition c2t2d0s1. The /etc/vfstab file entry for /usr is updated to use the mirror name d1. Device c2t2d0 is a multihost disk, so the localonly property is enabled.

```
    (Create the mirror)
# metainit -f d11 1 1 c0t0d0s1
d11: Concat/Stripe is setup
# metainit d21 1 1 c2t2d0s1
d21: Concat/Stripe is setup
# metainit d1 -m d11
d1: Mirror is setup

    (Edit the /etc/vfstab file)
```

```
# vi /etc/vfstab
#device         device          mount    FS    fsck    mount    mount
#to mount       to fsck         point    type  pass    at boot  options
#
/dev/md/dsk/d1 /dev/md/rdsk/d1 /usr ufs  2       no global
```

*(Move resource groups and device groups from phys-schost-1)*
```
# scswitch -S -h phys-schost-1
```

*(Reboot the node)*
```
# shutdown -g0 -y -i6
```

*(Attach the second submirror)*
```
# metattach d1 d21
d1: Submirror d21 is attached
```

*(View the sync status)*
```
# metastat d1
d1: Mirror
    Submirror 0: d11
        State: Okay
    Submirror 1: d21
        State: Resyncing
    Resync in progress: 15 % done
...
```

*(Identify the device-ID name of the mirrored disk's raw-disk device group)*
```
# scdidadm -L
...
1         phys-schost-3:/dev/rdsk/c2t2d0     /dev/did/rdsk/d2
```

*(Display the device-group node list)*
```
# scconf -pvv | grep dsk/d2
Device group name:                    dsk/d2
...
  (dsk/d2) Device group node list:        phys-schost-1, phys-schost-3
...
```

*(Remove phys-schost-3 from the node list)*
```
# scconf -r -D name=dsk/d2,nodelist=phys-schost-3
```

*(Enable the localonly property)*
```
# scconf -c -D name=dsk/d2,localonly=true
```

**Next Steps**    To mirror user-defined file systems, go to "How to Mirror File Systems That Can Be Unmounted" on page 159.

Otherwise, go to "Creating Disk Sets in a Cluster" on page 163 to create a disk set.

**Troubleshooting** Some of the steps in this mirroring procedure might cause an error message similar to `metainit: dg-schost-1: d1s0: not a metadevice`. Such an error message is harmless and can be ignored.

## ▼ How to Mirror File Systems That Can Be Unmounted

Use this procedure to mirror user-defined file systems that can be unmounted. In this procedure, the nodes do not need to be rebooted.

**Steps** 1. **Become superuser on a node of the cluster.**

2. **Unmount the file system to mirror.**
   Ensure that no processes are running on the file system.

   # **umount /***mount-point*

   See the `umount`(1M) man page and Chapter 19, "Mounting and Unmounting File Systems (Tasks)," in *System Administration Guide: Devices and File Systems* for more information.

3. **Place in a single-slice (one-way) concatenation the slice that contains a user-defined file system that can be unmounted.**
   Specify the physical disk name of the disk slice (c*N*t*X* d*Y*s*Z*).

   # **metainit -f** *submirror1* **1 1** *diskslice*

4. **Create a second concatenation.**

   # **metainit** *submirror2* **1 1** *submirror-diskslice*

5. **Create a one-way mirror with one submirror.**

   # **metainit** *mirror* **-m** *submirror1*

   ---
   **Note –** The metadevice or volume name for this mirror does *not* need to be unique throughout the cluster.

   ---

6. **Repeat Step 1 through Step 5 for each mountable file system to be mirrored.**

7. **On each node, edit the /etc/vfstab file entry for each file system you mirrored.**

Replace the names in the device to mount and device to fsck columns with the mirror name.

```
# vi /etc/vfstab
#device         device          mount      FS      fsck    mount     mount
#to mount       to fsck         point      type    pass    at boot   options
#
/dev/md/dsk/mirror /dev/md/rdsk/mirror /filesystem ufs 2 no global
```

8. **Attach the second submirror to the mirror.**

   This attachment starts a synchronization of the submirrors.

   ```
   # metattach mirror submirror2
   ```

9. **Wait for the synchronization of the mirrors, started in Step 8, to be completed.**

   Use the metastat(1M) command to view mirror status.

   ```
   # metastat mirror
   ```

10. **If the disk that is used to mirror the user-defined file system is physically connected to more than one node (multihosted), enable the localonly property.**

    Perform the following steps to enable the localonly property of the raw-disk device group for the disk that is used to mirror the user-defined file system. You must enable the localonly property to prevent unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

    a. **If necessary, use the scdidadm -L command to display the full device-ID path name of the raw-disk device group.**

       In the following example, the raw-disk device-group name dsk/d4 is part of the third column of output, which is the full device-ID path name.

       ```
       # scdidadm -L
       ...
       1          phys-schost-3:/dev/rdsk/c1t1d0      /dev/did/rdsk/d2
       ```

    b. **View the node list of the raw-disk device group.**

       Output looks similar to the following.

       ```
       # scconf -pvv | grep dsk/d2
       Device group name:                    dsk/d2
       ...
         (dsk/d2) Device group node list:     phys-schost-1, phys-schost-3
       ...
       ```

    c. **If the node list contains more than one node name, remove all nodes from the node list except the node whose root disk you mirrored.**

       Only the node whose root disk you mirrored should remain in the node list for the raw-disk device group.

       ```
       # scconf -r -D name=dsk/dN,nodelist=node
       ```

| `-D name=dsk/dN` | Specifies the cluster-unique name of the raw-disk device group |
|---|---|
| `nodelist=node` | Specifies the name of the node or nodes to remove from the node list |

   **d. Enable the `localonly` property.**

When the `localonly` property is enabled, the raw-disk device group is used exclusively by the node in its node list. This usage prevents unintentional fencing of the node from its boot device if the boot device is connected to multiple nodes.

   # **scconf -c -D name=**_rawdisk-groupname_**,localonly=true**

| `-D name=rawdisk-groupname` | Specifies the name of the raw-disk device group |
|---|---|

For more information about the `localonly` property, see the `scconf_dg_rawdisk`(1M) man page.

**11. Mount the mirrored file system.**

   # **mount** _/mount-point_

See the `mount`(1M) man page and Chapter 19, "Mounting and Unmounting File Systems (Tasks)," in _System Administration Guide: Devices and File Systems_ for more information.

**Example 3–5**   Mirroring File Systems That Can Be Unmounted

The following example shows creation of mirror `d4` to mirror `/export`, which resides on `c0t0d0s4`. Mirror `d4` consists of submirror `d14` on partition `c0t0d0s4` and submirror `d24` on partition `c2t2d0s4`. The `/etc/vfstab` file entry for `/export` is updated to use the mirror name `d4`. Device `c2t2d0` is a multihost disk, so the `localonly` property is enabled.

```
       (Unmount the file system)
# umount /export

       (Create the mirror)
# metainit -f d14 1 1 c0t0d0s4
d14: Concat/Stripe is setup
# metainit d24 1 1 c2t2d0s4
d24: Concat/Stripe is setup
# metainit d4 -m d14
d4: Mirror is setup

       (Edit the /etc/vfstab file)
# vi /etc/vfstab
#device          device          mount     FS     fsck    mount     mount
#to mount        to fsck         point     type   pass    at boot   options
#
# /dev/md/dsk/d4 /dev/md/rdsk/d4 /export ufs 2 no    global
```

```
        (Attach the second submirror)
# metattach d4 d24
d4: Submirror d24 is attached

        (View the sync status)
# metastat d4
d4: Mirror
        Submirror 0: d14
            State: Okay
        Submirror 1: d24
            State: Resyncing
        Resync in progress: 15 % done
...

        (Identify the device-ID name of the mirrored disk's raw-disk device group)
# scdidadm -L
...
1           phys-schost-3:/dev/rdsk/c2t2d0      /dev/did/rdsk/d2

        (Display the device-group node list)
# scconf -pvv | grep dsk/d2
Device group name:                      dsk/d2
...
  (dsk/d2) Device group node list:      phys-schost-1, phys-schost-3
...

        (Remove phys-schost-3 from the node list)
# scconf -r -D name=dsk/d2,nodelist=phys-schost-3

        (Enable the localonly property)
# scconf -c -D name=dsk/d2,localonly=true

        (Mount the file system)
# mount /export
```

**Next Steps**  If you need to create disk sets, go to one of the following:

- To create a Solaris Volume Manager for Sun Cluster disk set for use by Oracle Real Application Clusters, go to "Creating a Multi-Owner Disk Set in Solaris Volume Manager for Sun Cluster for the Oracle Real Application Clusters Database" in *Sun Cluster Data Service for Oracle Real Application Clusters Guide for Solaris OS*.

- To create a disk set for any other application, go to .

- If you used SunPlex Installer to install Solstice DiskSuite, one to three disk sets might already exist. See for information about the metasets that were created by SunPlex Installer.

If you have sufficient disk sets for your needs, go to one of the following:

- If your cluster contains disk sets that are configured with exactly two disk enclosures and two nodes, you must add dual-string mediators. Go to "Configuring Dual-String Mediators" on page 172.

- If your cluster configuration does not require dual-string mediators, go to "How to Create Cluster File Systems" on page 119.

**Troubleshooting** Some of the steps in this mirroring procedure might cause an error message that is similar to `metainit: dg-schost-1: d1s0: not a metadevice`. Such an error message is harmless and can be ignored.

# Creating Disk Sets in a Cluster

This section describes how to create disk sets for a cluster configuration. You might not need to create disk sets under the following circumstances:

- If you used SunPlex Installer to install Solstice DiskSuite, one to three disk sets might already exist. See "Using SunPlex Installer to Configure Sun Cluster Software" on page 86 for information about the metasets that were created by SunPlex Installer.

- To create a Solaris Volume Manager for Sun Cluster disk set for use by Oracle Real Application Clusters, do not use these procedures. Instead, perform the procedures in "Creating a Multi-Owner Disk Set in Solaris Volume Manager for Sun Cluster for the Oracle Real Application Clusters Database" in *Sun Cluster Data Service for Oracle Real Application Clusters Guide for Solaris OS*.

The following table lists the tasks that you perform to create disk sets.

**TABLE 3–2** Task Map: Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software

| Task | Instructions |
|------|--------------|
| 1. Create disk sets by using the `metaset` command. | "How to Create a Disk Set" on page 164 |
| 2. Add drives to the disk sets. | "How to Add Drives to a Disk Set" on page 167 |
| 3. (*Optional*) Repartition drives in a disk set to allocate space to slices 1 through 6. | "How to Repartition Drives in a Disk Set" on page 168 |
| 4. List DID pseudo-driver mappings and define metadevices or volumes in the `/etc/lvm/md.tab` files. | "How to Create an `md.tab` File" on page 169 |

**TABLE 3–2** Task Map: Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software     *(Continued)*

| Task | Instructions |
|------|-------------|
| 5. Initialize the `md.tab` files. | "How to Activate Metadevices or Volumes" on page 171 |

## ▼ How to Create a Disk Set

Perform this procedure to create disk sets.

**Steps**  1. *(Solaris 8 or Solaris 9)* **Determine whether, after you create the new disk sets, the cluster will have more than three disk sets.**

   - If the cluster will have no more than three disk sets, skip to Step 9.

   - If the cluster will have four or more disk sets, proceed to Step 2 to prepare the cluster. You must perform this task whether you are installing disk sets for the first time or whether you are adding more disk sets to a fully configured cluster.

   - If the cluster runs on the Solaris 10 OS, Solaris Volume Manager automatically makes the necessary configuration changes. Skip to Step 9.

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

3. **If the total 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` to the desired value.**

   The maximum permissible number of disk sets is one less than the configured value of `md_nsets`. The maximum possible value of `md_nsets` is 32, therefore the maximum permissible number of disk sets that you can create is 31.

4. **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 or Solaris Volume Manager errors and possible loss of data.

5. **If you made changes to the `md.conf` file on any node, perform the following steps to make those changes active.**

   a. **From one node, shut down the cluster.**

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

   b. **Reboot each node of the cluster.**

      ```
      ok> boot
      ```

6. **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.

7. **From one node of the cluster, run the `scgdevs`(1M) command to update the global-devices namespace.**

8. **On each node, verify that the `scgdevs` command has completed processing 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
   ```

9. **Ensure that the disk set that you intend to create meets one of the following requirements.**

   - If the disk set is configured with exactly two disk strings, the disk set must connect to exactly two nodes and use exactly two mediator hosts. These mediator hosts must be the same two hosts used for the disk set. See "Configuring Dual-String Mediators" on page 172 for details on how to configure dual-string mediators.

   - If the disk set is configured with more than two disk strings, ensure that for any two disk strings S1 and S2, the sum of the number of drives on those strings exceeds the number of drives on the third string S3. Stated as a formula, the requirement is that count(S1) + count(S2) > count(S3).

10. **Ensure that the local state database replicas exist.**

    For instructions, see "How to Create State Database Replicas" on page 147.

11. **Become superuser on the cluster node that will master the disk set.**

12. **Create the disk set.**

    The following command creates the disk set and registers the disk set as a Sun Cluster disk device group.

    ```
    # metaset -s setname -a -h node1 node2
    ```

    | | |
    |---|---|
    | -s setname | Specifies the disk set name |
    | -a | Adds (creates) the disk set |
    | -h node1 | Specifies the name of the primary node to master the disk set |
    | node2 | Specifies the name of the secondary node to master the disk set |

> **Note –** When you run the `metaset` command to configure a Solstice DiskSuite or Solaris Volume Manager device group on a cluster, the command designates one secondary node by default. You can change the desired number of secondary nodes in the device group by using the `scsetup`(1M) utility after the device group is created. Refer to "Administering Disk Device Groups" in *Sun Cluster System Administration Guide for Solaris OS* for more information about how to change the `numsecondaries` property.

13. **Verify the status of the new disk set.**

    ```
    # metaset -s setname
    ```

**Example 3–6**    Creating a Disk Set

The following command creates two disk sets, dg-schost-1 and dg-schost-2, with the nodes phys-schost-1 and phys-schost-2 specified as the potential primaries.

```
# metaset -s dg-schost-1 -a -h phys-schost-1 phys-schost-2
# metaset -s dg-schost-2 -a -h phys-schost-1 phys-schost-2
```

**Next Steps**    Add drives to the disk set. Go to .

## Adding Drives to a Disk Set

When you add a drive to a disk set, the volume management software repartitions the drive as follows so that the state database for the disk set can be placed on the drive.

- A small portion of each drive is reserved in slice 7 for use by Solstice DiskSuite or Solaris Volume Manager software. The remainder of the space on each drive is placed into slice 0.

- Drives are repartitioned when they are added to the disk set only if slice 7 is not configured correctly.

- Any existing data on the drives is lost by the repartitioning.

- If slice 7 starts at cylinder 0, and the drive partition is large enough to contain a state database replica, the drive is not repartitioned.

## ▼ How to Add Drives to a Disk Set

**Before You Begin**  Ensure that the disk set has been created. For instructions, see

**Steps**  **1. Become superuser on the node.**

**2. List the DID mappings.**

```
# scdidadm -L
```

- Choose drives that are shared by the cluster nodes that will master or potentially master the disk set.

- Use the full device-ID path names when you add drives to a disk set.

The first column of output is the DID instance number, the second column is the full physical path name, and the third column is the full device-ID path name (pseudo path). A shared drive has more than one entry for the same DID instance number.

In the following example, the entries for DID instance number 2 indicate a drive that is shared by phys-schost-1 and phys-schost-2, and the full device-ID path name is /dev/did/rdsk/d2.

```
1        phys-schost-1:/dev/rdsk/c0t0d0 /dev/did/rdsk/d1
2        phys-schost-1:/dev/rdsk/c1t1d0 /dev/did/rdsk/d2
2        phys-schost-2:/dev/rdsk/c1t1d0 /dev/did/rdsk/d2
3        phys-schost-1:/dev/rdsk/c1t2d0 /dev/did/rdsk/d3
3        phys-schost-2:/dev/rdsk/c1t2d0 /dev/did/rdsk/d3
...
```

**3. Become owner of the disk set.**

```
# metaset -s setname -t
```

-s setname      Specifies the disk set name

-t              Takes ownership of the disk set

**4. Add the drives to the disk set.**

Use the full device-ID path name.

```
# metaset -s setname -a drivename
```

-a              Adds the drive to the disk set

drivename       Full device-ID path name of the shared drive

> **Note –** Do *not* use the lower-level device name (c*N*t*X*d*Y*) when you add a drive to a disk set. Because the lower-level device name is a local name and not unique throughout the cluster, using this name might prevent the metaset from being able to switch over.

5. **Verify the status of the disk set and drives.**

   ```
   # metaset -s setname
   ```

**Example 3–7**    Adding Drives to a Disk Set

The metaset command adds the drives /dev/did/rdsk/d1 and /dev/did/rdsk/d2 to the disk set dg-schost-1.

```
# metaset -s dg-schost-1 -a /dev/did/rdsk/d1 /dev/did/rdsk/d2
```

**Next Steps**    To repartition drives for use in metadevices or volumes, go to "How to Repartition Drives in a Disk Set" on page 168.

Otherwise, go to "How to Create an md.tab File" on page 169 to define metadevices or volumes by using an md.tab file.

## ▼ How to Repartition Drives in a Disk Set

The metaset(1M) command repartitions drives in a disk set so that a small portion of each drive is reserved in slice 7 for use by Solstice DiskSuite or Solaris Volume Manager software. The remainder of the space on each drive is placed into slice 0. To make more effective use of the drive, use this procedure to modify the disk layout. If you allocate space to slices 1 through 6, you can use these slices when you set up Solstice DiskSuite metadevices or Solaris Volume Manager volumes.

**Steps**    1. **Become superuser on the cluster node.**

2. **Use the `format` command to change the disk partitioning for each drive in the disk set.**

   When you repartition a drive, you must meet the following conditions to prevent the metaset(1M) command from repartitioning the drive.

   - Create slice 7 starting at cylinder 0, large enough to hold a state database replica. See your Solstice DiskSuite or Solaris Volume Manager administration guide to determine the size of a state database replica for your version of the volume-manager software.

- Set the `Flag` field in slice 7 to `wu` (read-write, unmountable). Do not set it to read-only.
- Do not allow slice 7 to overlap any other slice on the drive.

See the `format`(1M) man page for details.

**Next Steps**  Define metadevices or volumes by using an `md.tab` file. Go to "How to Create an `md.tab` File" on page 169.

## ▼ How to Create an `md.tab` File

Create an `/etc/lvm/md.tab` file on each node in the cluster. Use the `md.tab` file to define Solstice DiskSuite metadevices or Solaris Volume Manager volumes for the disk sets that you created.

---

**Note –** If you are using local metadevices or volumes, ensure that local metadevices or volumes names are distinct from the device-ID names used to form disk sets. For example, if the device-ID name `/dev/did/dsk/d3` is used in a disk set, do not use the name `/dev/md/dsk/d3` for a local metadevice or volume. This requirement does not apply to shared metadevices or volumes, which use the naming convention `/dev/md/`*setname*`/{r}dsk/d#`.

---

**Steps**  1. **Become superuser on the cluster node.**

2. **List the DID mappings for reference when you create your `md.tab` file.**

   Use the full device-ID path names in the `md.tab` file in place of the lower-level device names (c*N* t*X*d*Y*).

   ```
   # scdidadm -L
   ```

   In the following example, the first column of output is the DID instance number, the second column is the full physical path name, and the third column is the full device-ID path name (pseudo path).

   ```
   1        phys-schost-1:/dev/rdsk/c0t0d0 /dev/did/rdsk/d1
   2        phys-schost-1:/dev/rdsk/c1t1d0 /dev/did/rdsk/d2
   2        phys-schost-2:/dev/rdsk/c1t1d0 /dev/did/rdsk/d2
   3        phys-schost-1:/dev/rdsk/c1t2d0 /dev/did/rdsk/d3
   3        phys-schost-2:/dev/rdsk/c1t2d0 /dev/did/rdsk/d3
   ...
   ```

3. **Create an `/etc/lvm/md.tab` file and edit it by hand with your preferred text editor.**

> **Note –** If you have existing data on the drives that will be used for the submirrors, you must back up the data before metadevice or volume setup. Then restore the data onto the mirror.

To avoid possible confusion between local metadevices or volumes in a cluster environment, use a naming scheme that makes each local metadevice or volume name unique throughout the cluster. For example, for node 1 choose names from d100-d199. And for node 2 use d200-d299.

See your Solstice DiskSuite or Solaris Volume Manager documentation and the `md.tab`(4) man page for details about how to create an `md.tab` file.

**Example 3–8**   Sample `md.tab` File

The following sample `md.tab` file defines the disk set that is named `dg-schost-1`. The ordering of lines in the `md.tab` file is not important.

```
dg-schost-1/d0 -m dg-schost-1/d10 dg-schost-1/d20
    dg-schost-1/d10 1 1 /dev/did/rdsk/d1s0
    dg-schost-1/d20 1 1 /dev/did/rdsk/d2s0
```

The following example uses Solstice DiskSuite terminology. For Solaris Volume Manager, a trans metadevice is instead called a *transactional volume* and a metadevice is instead called a *volume*. Otherwise, the following process is valid for both volume managers.

The sample `md.tab` file is constructed as follows.

1. The first line defines the device `d0` as a mirror of metadevices `d10` and `d20`. The `-m` signifies that this device is a mirror device.

   ```
   dg-schost-1/d0 -m dg-schost-1/d0 dg-schost-1/d20
   ```

2. The second line defines metadevice `d10`, the first submirror of `d0`, as a one-way stripe.

   ```
   dg-schost-1/d10 1 1 /dev/did/rdsk/d1s0
   ```

3. The third line defines metadevice `d20`, the second submirror of `d0`, as a one-way stripe.

   ```
   dg-schost-1/d20 1 1 /dev/did/rdsk/d2s0
   ```

**Next Steps**   Activate the metadevices or volumes that are defined in the `md.tab` files. Go to .

# ▼ How to Activate Metadevices or Volumes

Perform this procedure to activate Solstice DiskSuite metadevices or Solaris Volume Manager volumes that are defined in md.tab files.

**Steps**
1. **Become superuser on the cluster node.**

2. **Ensure that md.tab files are located in the /etc/lvm directory.**

3. **Ensure that you have ownership of the disk set on the node where the command will be executed.**

4. **Take ownership of the disk set.**

   ```
   # scswitch -z setname -h node
   ```

   -z *setname*      Specifies the disk set name

   -h *node*          Specifies the node that takes ownership

5. **Activate the disk set's metadevices or volumes, which are defined in the md.tab file.**

   ```
   # metainit -s setname -a
   ```

   -s *setname*      Specifies the disk set name

   -a                Activates all metadevices in the md.tab file

6. **For each master and log device, attach the second submirror (*submirror2*).**

   When the metadevices or volumes in the md.tab file are activated, only the first submirror (*submirror1*) of the master and log devices is attached, so *submirror2* must be attached manually.

   ```
   # metattach mirror submirror2
   ```

7. **Repeat Step 3 through Step 6 for each disk set in the cluster.**

   If necessary, run the metainit(1M) command from another node that has connectivity to the drives. This step is required for cluster-pair topologies, where the drives are not accessible by all nodes.

8. **Check the status of the metadevices or volumes.**

   ```
   # metastat -s setname
   ```

   See the metastat(1M) man page for more information.

**Example 3–9**   Activating Metadevices or Volumes in the md.tab File

In the following example, all metadevices that are defined in the md.tab file for disk set dg-schost-1 are activated. Then the second submirrors of master device dg-schost-1/d1 and log device dg-schost-1/d4 are activated.

```
# metainit -s dg-schost-1 -a
# metattach dg-schost-1/d1 dg-schost-1/d3
# metattach dg-schost-1/d4 dg-schost-1/d6
```

**Next Steps**    If your cluster contains disk sets that are configured with exactly two disk enclosures and two nodes, add dual-string mediators. Go to "Configuring Dual-String Mediators" on page 172.

Otherwise, go to "How to Create Cluster File Systems" on page 119 to create a cluster file system.

# Configuring Dual-String Mediators

This section provides information and procedures to configure dual-string mediator hosts. Dual-string mediators are required for all Solstice DiskSuite or Solaris Volume Manager disk sets that are configured with exactly two disk strings and two cluster nodes. The use of mediators enables the Sun Cluster software to ensure that the most current data is presented in the instance of a single-string failure in a dual-string configuration.

A *dual-string mediator*, or mediator host, is a cluster node that stores mediator data. Mediator data provides information about the location of other mediators and contains a commit count that is identical to the commit count that is stored in the database replicas. This commit count is used to confirm that the mediator data is in sync with the data in the database replicas.

A *disk string* consists of a disk enclosure, its physical drives, cables from the enclosure to the node(s), and the interface adapter cards.

The following table lists the tasks that you perform to configure dual-string mediator hosts.

**TABLE 3–3** Task Map: Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software

| Task | Instructions |
|------|-------------|
| 1. Configure dual-string mediator hosts. | "Requirements for Dual-String Mediators" on page 173 |
| | "How to Add Mediator Hosts" on page 173 |
| 2. Check the status of mediator data. | "How to Check the Status of Mediator Data" on page 174 |

**TABLE 3–3** Task Map: Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software      *(Continued)*

| Task | Instructions |
|---|---|
| 3. If necessary, fix bad mediator data. | "How to Fix Bad Mediator Data" on page 174 |

# Requirements for Dual-String Mediators

The following rules apply to dual-string configurations that use mediators.

- Disk sets must be configured with exactly two mediator hosts. Those two mediator hosts must be the same two cluster nodes that are used for the disk set.
- A disk set cannot have more than two mediator hosts.
- Mediators cannot be configured for disk sets that do not meet the two-string and two-host criteria.

These rules do not require that the entire cluster must have exactly two nodes. Rather, only those disk sets that have two disk strings must be connected to exactly two nodes. An N+1 cluster and many other topologies are permitted under these rules.

# ▼ How to Add Mediator Hosts

Perform this procedure if your configuration requires dual-string mediators.

**Steps 1. Become superuser on the node that currently masters the disk set to which you intend to add mediator hosts.**

**2. Add each node with connectivity to the disk set as a mediator host for that disk set.**

```
# metaset -s setname -a -m mediator-host-list
```

| -s setname | Specifies the disk set name |
|---|---|
| -a | Adds to the disk set |
| -m mediator-host-list | Specifies the name of the node to add as a mediator host for the disk set |

See the mediator(7D) man page for details about mediator-specific options to the metaset command.

**Example 3–10** Adding Mediator Hosts

The following example adds the nodes phys-schost-1 and phys-schost-2 as mediator hosts for the disk set dg-schost-1. Both commands are run from the node phys-schost-1.

```
# metaset -s dg-schost-1 -a -m phys-schost-1
# metaset -s dg-schost-1 -a -m phys-schost-2
```

**Next Steps**    Check the status of mediator data. Go to "How to Check the Status of Mediator Data" on page 174.

## ▼ How to Check the Status of Mediator Data

**Before You Begin**    Ensure that you have added mediator hosts as described in "How to Add Mediator Hosts" on page 173.

**Steps**    1. **Display the status of the mediator data.**

```
# medstat -s setname
```

-s setname         Specifies the disk set name

See the medstat(1M) man page for more information.

2. **If Bad is the value in the Status field of the medstat output, repair the affected mediator host.**
   Go to "How to Fix Bad Mediator Data" on page 174.

**Next Steps**    Go to "How to Create Cluster File Systems" on page 119 to create a cluster file system.

## ▼ How to Fix Bad Mediator Data

Perform this procedure to repair bad mediator data.

**Steps**    1. **Identify all mediator hosts with bad mediator data as described in the procedure "How to Check the Status of Mediator Data" on page 174.**

2. **Become superuser on the node that owns the affected disk set.**

3. **Remove all mediator hosts with bad mediator data from all affected disk sets.**

```
# metaset -s setname -d -m mediator-host-list
```

| | |
|---|---|
| -s setname | Specifies the disk set name |
| -d | Deletes from the disk set |
| -m mediator-host-list | Specifies the name of the node to remove as a mediator host for the disk set |

4.  **Restore each mediator host that you removed in Step 3.**

    # **metaset -s** *setname* **-a -m** *mediator-host-list*

    -a                              Adds to the disk set

    -m mediator-host-list    Specifies the name of the node to add as a
                                    mediator host for the disk set

    See the mediator(7D) man page for details about mediator-specific options to the
    metaset command.

**Next Steps**    Create cluster file systems. Go to "How to Create Cluster File Systems" on page 119.

# SPARC: Installing and Configuring VERITAS Volume Manager

Install and configure your local and multihost disks for VERITAS Volume Manager (VxVM) by using the procedures in this chapter, along with the planning information in "Planning Volume Management" on page 35. See your VxVM documentation for additional details.

The following sections are in this chapter:

- "SPARC: Installing and Configuring VxVM Software" on page 177
- "SPARC: Creating Disk Groups in a Cluster" on page 186
- "SPARC: Unencapsulating the Root Disk" on page 189

# SPARC: Installing and Configuring VxVM Software

This section provides information and procedures to install and configure VxVM software on a Sun Cluster configuration.

The following table lists the tasks to perform to install and configure VxVM software for Sun Cluster configurations.

TABLE 4–1 SPARC: Task Map: Installing and Configuring VxVM Software

| Task | Instructions |
|---|---|
| 1. Plan the layout of your VxVM configuration. | "Planning Volume Management" on page 35 |

| Task | Instructions |
|---|---|
| 2. Determine how you will create the root disk group on each node. As of VxVM 4.0, the creation of a root disk group is optional. | "SPARC: Setting Up a Root Disk Group Overview" on page 178 |
| 3. Install VxVM software. | "SPARC: How to Install VERITAS Volume Manager Software" on page 179<br><br>VxVM installation documentation |
| 4. If necessary, create a root disk group. You can either encapsulate the root disk or create the root disk group on local, nonroot disks. | "SPARC: How to Encapsulate the Root Disk" on page 181<br><br>"SPARC: How to Create a Root Disk Group on a Nonroot Disk" on page 182 |
| 5. *(Optional)* Mirror the encapsulated root disk. | "SPARC: How to Mirror the Encapsulated Root Disk" on page 184 |
| 6. Create disk groups. | "SPARC: Creating Disk Groups in a Cluster" on page 186 |

# SPARC: Setting Up a Root Disk Group Overview

As of VxVM 4.0, the creation of a root disk group is optional. If you do not intend to create a root disk group, proceed to "SPARC: How to Install VERITAS Volume Manager Software" on page 179.

For VxVM 3.5, each cluster node requires the creation of a root disk group after VxVM is installed. This root disk group is used by VxVM to store configuration information, and has the following restrictions.

- Access to a node's root disk group must be restricted to only that node.

- Remote nodes must never access data stored in another node's root disk group.

- Do not use the scconf(1M) command to register the root disk group as a disk device group.

- Whenever possible, configure the root disk group for each node on a nonshared disk.

Sun Cluster software supports the following methods to configure the root disk group.

- **Encapsulate the node's root disk** – This method enables the root disk to be mirrored, which provides a boot alternative if the root disk is corrupted or damaged. To encapsulate the root disk you need two free disk slices as well as free cylinders, preferably at the beginning or the end of the disk.

- **Use local nonroot disks** – This method provides an alternative to encapsulating the root disk. If a node's root disk is encapsulated, certain tasks you might later perform, such as upgrade the Solaris OS or perform disaster recovery procedures,

could be more complicated than if the root disk is not encapsulated. To avoid this potential added complexity, you can instead initialize or encapsulate local nonroot disks for use as root disk groups.

A root disk group that is created on local nonroot disks is local to that node, neither globally accessible nor highly available. As with the root disk, to encapsulate a nonroot disk you need two free disk slices as well as free cylinders at the beginning or the end of the disk.

See your VxVM installation documentation for more information.

## ▼ SPARC: How to Install VERITAS Volume Manager Software

Perform this procedure to install VERITAS Volume Manager (VxVM) software on each node that you want to install with VxVM. You can install VxVM on all nodes of the cluster, or install VxVM just on the nodes that are physically connected to the storage devices that VxVM will manage.

**Before You Begin**  Perform the following tasks:
- Ensure that all nodes in the cluster are running in cluster mode.
- Obtain any VERITAS Volume Manager (VxVM) license keys that you need to install.
- Have available your VxVM installation documentation.

**Steps**  1. **Become superuser on a cluster node that you intend to install with VxVM.**

2. **Insert the VxVM CD-ROM in the CD-ROM drive on the node.**

3. **For VxVM 4.1, follow procedures in your VxVM installation guide to install and configure VxVM software and licenses.**

---

**Note –** For VxVM 4.1, the `scvxinstall` command no longer performs installation of VxVM packages and licenses, but does perform necessary postinstallation tasks.

---

4. **Run the `scvxinstall` utility in noninteractive mode.**

   - **For VxVM 4.0 and earlier, use the following command:**

     # **scvxinstall -i -L** {*license* | **none**}

     -i                            Installs VxVM but does not encapsulate the root disk

     -L {license | none}           Installs the specified *license*. The none argument specifies that no additional license key is being

added.

- **For VxVM 4.1, use the following command:**

  ```
  # scvxinstall -i
  ```

  -i      For VxVM 4.1, verifies that VxVM is installed but does not encapsulate the root disk

The scvxinstall utility also selects and configures a cluster-wide vxio driver major number. See the scvxinstall(1M) man page for more information.

5. **If you intend to enable the VxVM cluster feature, supply the cluster feature license key, if you did not already do so.**

   See your VxVM documentation for information about how to add a license.

6. **(Optional) Install the VxVM GUI.**

   See your VxVM documentation for information about installing the VxVM GUI.

7. **Eject the CD-ROM.**

8. **Install any VxVM patches.**

   See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

9. **(Optional) For VxVM 4.0 and earlier, if you prefer not to have VxVM man pages reside on the cluster node, remove the man-page package.**

   ```
   # pkgrm VRTSvmman
   ```

10. **Repeat Step 1 through Step 9 to install VxVM on any additional nodes.**

    ---

    **Note –** If you intend to enable the VxVM cluster feature, you *must* install VxVM on all nodes of the cluster.

    ---

11. **If you do *not* install one or more nodes with VxVM, modify the /etc/name_to_major file on each non-VxVM node.**

    a. **On a node that is installed with VxVM, determine the vxio major number setting.**

       ```
       # grep vxio /etc/name_to_major
       ```

    b. **Become superuser on a node that you do *not* intend to install with VxVM.**

c. **Edit the /etc/name_to_major file and add an entry to set the vxio major number to** *NNN*, **the number derived in Step a.**

```
# vi /etc/name_to_major
vxio NNN
```

d. **Initialize the vxio entry.**

```
# drvconfig -b -i vxio -m NNN
```

e. **Repeat Step a through Step d on all other nodes that you do *not* intend to install with VxVM.**

When you finish, each node of the cluster should have the same vxio entry in its /etc/name_to_major file.

12. **To create a root disk group, go to "SPARC: How to Encapsulate the Root Disk" on page 181 or "SPARC: How to Create a Root Disk Group on a Nonroot Disk" on page 182.**

Otherwise, proceed to Step 13.

---

**Note –** VxVM 3.5 requires that you create a root disk group. For VxVM 4.0 and later, a root disk group is optional.

---

13. **Reboot each node on which you installed VxVM.**

```
# shutdown -g0 -y -i6
```

**Next Steps**   To create a root disk group, go to "SPARC: How to Encapsulate the Root Disk" on page 181 or "SPARC: How to Create a Root Disk Group on a Nonroot Disk" on page 182.

Otherwise, create disk groups. Go to "SPARC: Creating Disk Groups in a Cluster" on page 186.

## ▼ SPARC: How to Encapsulate the Root Disk

Perform this procedure to create a root disk group by encapsulating the root disk. Root disk groups are required for VxVM 3.5. For VxVM 4.0 and later, root disk groups are optional. See your VxVM documentation for more information.

> **Note –** If you want to create the root disk group on nonroot disks, instead perform procedures in "SPARC: How to Create a Root Disk Group on a Nonroot Disk" on page 182.

**Before You Begin**   Ensure that you have installed VxVM as described in "SPARC: How to Install VERITAS Volume Manager Software" on page 179.

**Steps**   **1. Become superuser on a node that you installed with VxVM.**

     **2. Encapsulate the root disk.**

```
# scvxinstall -e
```

     -e      Encapsulates the root disk

     See the scvxinstall(1M) for more information.

     **3. Repeat for any other node on which you installed VxVM.**

**Next Steps**   To mirror the encapsulated root disk, go to "SPARC: How to Mirror the Encapsulated Root Disk" on page 184.

Otherwise, go to "SPARC: Creating Disk Groups in a Cluster" on page 186.

## ▼ SPARC: How to Create a Root Disk Group on a Nonroot Disk

Use this procedure to create a root disk group by encapsulating or initializing local disks other than the root disk. As of VxVM 4.0, the creation of a root disk group is optional.

> **Note –** If you want to create a root disk group on the root disk, instead perform procedures in "SPARC: How to Encapsulate the Root Disk" on page 181.

**Before You Begin**   If the disks are to be encapsulated, ensure that each disk has at least two slices with 0 cylinders. If necessary, use the format(1M) command to assign 0 cylinders to each VxVM slice.

**Steps**   **1. Become superuser on the node.**

2. **Start the `vxinstall` utility.**

   # **vxinstall**

   When prompted, make the following choices or entries.

   - If you intend to enable the VxVM cluster feature, supply the cluster feature license key.
   - Choose Custom Installation.
   - Do not encapsulate the boot disk.
   - Choose any disks to add to the root disk group.
   - Do not accept automatic reboot.

3. **If the root disk group that you created contains one or more disks that connect to more than one node, enable the `localonly` property.**

   Use the following command to enable the `localonly` property of the raw-disk device group for each shared disk in the root disk group.

   # **scconf -c -D name=dsk/d**$N$**,localonly=true**

   When the `localonly` property is enabled, the raw-disk device group is used exclusively by the node in its node list. This usage prevents unintentional fencing of the node from the disk that is used by the root disk group if that disk is connected to multiple nodes.

   For more information about the `localonly` property, see the `scconf_dg_rawdisk`(1M) man page.

4. **Move any resource groups or device groups from the node.**

   # **scswitch -S -h** *from-node*

   | | |
   |---|---|
   | -S | Moves all resource groups and device groups |
   | -h *from-node* | Specifies the name of the node from which to move resource or device groups |

5. **Reboot the node.**

   # **shutdown -g0 -y -i6**

6. **Use the `vxdiskadm` command to add multiple disks to the root disk group.**

   The root disk group becomes tolerant of a disk failure when it contains multiple disks. See VxVM documentation for procedures.

**Next Steps**   Create disk groups. Go to "SPARC: Creating Disk Groups in a Cluster" on page 186.

# ▼ SPARC: How to Mirror the Encapsulated Root Disk

After you install VxVM and encapsulate the root disk, perform this procedure on each node on which you mirror the encapsulated root disk.

**Before You Begin**  Ensure that you have encapsulated the root disk as described in "SPARC: How to Encapsulate the Root Disk" on page 181.

**Steps**  **1. Mirror the encapsulated root disk.**

Follow the procedures in your VxVM documentation. For maximum availability and simplified administration, use a local disk for the mirror. See "Guidelines for Mirroring the Root Disk" on page 42 for additional guidelines.

---

**Caution –** Do not use a quorum device to mirror a root disk. Using a quorum device to mirror a root disk might prevent the node from booting from the root-disk mirror under certain circumstances.

---

**2. Display the DID mappings.**

```
# scdidadm -L
```

**3. From the DID mappings, locate the disk that is used to mirror the root disk.**

**4. Extract the raw-disk device-group name from the device-ID name of the root-disk mirror.**

The name of the raw-disk device group follows the convention dsk/d*N*, where *N* is a number. In the following output, the portion of a scdidadm output line from which you extract the raw-disk device-group name is highlighted in bold.

*N*        *node*:/dev/rdsk/c*N*t*X*d*Y*        /dev/did/r**dsk/d***N*

**5. View the node list of the raw-disk device group.**

Output looks similar to the following.

```
# scconf -pvv | grep dsk/dN
Device group name:                      dsk/dN
...
 (dsk/dN) Device group node list:       phys-schost-1, phys-schost-3
...
```

**6. If the node list contains more than one node name, remove from the node list all nodes except the node whose root disk you mirrored.**

Only the node whose root disk you mirrored should remain in the node list for the raw-disk device group.

```
# scconf -r -D name=dsk/dN,nodelist=node
```

<table>
<tr><td>-D name=dsk/d<em>N</em></td><td>Specifies the cluster-unique name of the raw-disk device group</td></tr>
<tr><td>nodelist=<em>node</em></td><td>Specifies the name of the node or nodes to remove from the node list</td></tr>
</table>

7. **Enable the `localonly` property of the raw-disk device group.**

   When the localonly property is enabled, the raw-disk device group is used exclusively by the node in its node list. This usage prevents unintentional fencing of the node from its boot device if the boot device is connected to multiple nodes.

   ```
   # scconf -c -D name=dsk/dN,localonly=true
   ```

   For more information about the localonly property, see the scconf_dg_rawdisk(1M) man page.

8. **Repeat this procedure for each node in the cluster whose encapsulated root disk you want to mirror.**

**Example 4–1**    SPARC: Mirroring the Encapsulated Root Disk

The following example shows a mirror created of the root disk for the node phys-schost-1. The mirror is created on the disk c1t1d0, whose raw-disk device-group name is dsk/d2. Disk c1t1d0 is a multihost disk, so the node phys-schost-3 is removed from the disk's node list and the localonly property is enabled.

```
    (Display the DID mappings)
# scdidadm -L
...
2        phys-schost-1:/dev/rdsk/c1t1d0 /dev/did/rdsk/d2
2        phys-schost-3:/dev/rdsk/c1t1d0 /dev/did/rdsk/d2
...

    (Display the node list of the mirror disk's raw-disk device group)
# scconf -pvv | grep dsk/d2
Device group name:                       dsk/d2
...
  (dsk/d2) Device group node list:       phys-schost-1, phys-schost-3
...

    (Remove phys-schost-3 from the node list)
# scconf -r -D name=dsk/d2,nodelist=phys-schost-3

    (Enable the localonly property)
# scconf -c -D name=dsk/d2,localonly=true
```

**Next Steps**    Create disk groups. Go to "SPARC: Creating Disk Groups in a Cluster" on page 186.

# SPARC: Creating Disk Groups in a Cluster

This section describes how to create VxVM disk groups in a cluster.

The following table lists the tasks to perform to create VxVM disk groups for Sun Cluster configurations.

**TABLE 4–2** SPARC: Task Map: Creating VxVM Disk Groups

| Task | Instructions |
|---|---|
| 1. Create disk groups and volumes. | "SPARC: How to Create and Register a Disk Group" on page 186 |
| 2. If necessary, resolve any minor-number conflicts between disk device groups by assigning a new minor number. | "SPARC: How to Assign a New Minor Number to a Disk Device Group" on page 188 |
| 3. Verify the disk groups and volumes. | "SPARC: How to Verify the Disk Group Configuration" on page 189 |

## ▼ SPARC: How to Create and Register a Disk Group

Use this procedure to create your VxVM disk groups and volumes.

---

**Note –** After a disk group is registered with the cluster as a disk device group, you should never import or deport a VxVM disk group by using VxVM commands. The Sun Cluster software can handle all cases where disk groups need to be imported or deported. See "Administering Disk Device Groups" in *Sun Cluster System Administration Guide for Solaris OS* for procedures about how to manage Sun Cluster disk device groups.

---

Perform this procedure from a node that is physically connected to the disks that make the disk group that you add.

**Before You Begin** Perform the following tasks:

- Make mappings of your storage disk drives. See the appropriate manual in the *Sun Cluster Hardware Administration Collection* to perform an initial installation of your storage device.
- Complete the following configuration planning worksheets.

- "Local File System Layout Worksheet" on page 288
- "Disk Device Group Configurations Worksheet" on page 294
- "Volume-Manager Configurations Worksheet" on page 296

See "Planning Volume Management" on page 35 for planning guidelines.

- If you did not create root disk groups, ensure that you have rebooted each node on which you installed VxVM, as instructed in Step 13 of "SPARC: How to Install VERITAS Volume Manager Software" on page 179.

**Steps** 1. **Become superuser on the node that will own the disk group.**

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

   If you are installing Oracle Real Application Clusters, create shared VxVM disk groups by using the cluster feature of VxVM as described in the *VERITAS Volume Manager Administrator's Reference Guide*. Otherwise, create VxVM disk groups by using the standard procedures that are documented in the VxVM documentation.

   ---

   **Note –** You can use Dirty Region Logging (DRL) to decrease volume recovery time if a node failure occurs. However, DRL might decrease I/O throughput.

   ---

3. **If the VxVM cluster feature is *not* enabled, register the disk group as a Sun Cluster disk device group.**

   If the VxVM cluster feature is enabled, do *not* register a shared disk group as a Sun Cluster disk device group. Instead, go to "SPARC: How to Verify the Disk Group Configuration" on page 189.

   a. **Start the `scsetup(1M)` utility.**

      ```
      # scsetup
      ```

   b. **Choose the menu item, Device groups and volumes.**

   c. **Choose the menu item, Register a VxVM disk group.**

   d. **Follow the instructions to specify the VxVM disk group that you want to register as a Sun Cluster disk device group.**

   e. **When finished, quit the `scsetup` utility.**

   f. **Verify that the disk device group is registered.**

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

      ```
      # scstat -D
      ```

**Next Steps** Go to "SPARC: How to Verify the Disk Group Configuration" on page 189.

**Troubleshooting** **Failure to register the device group** – If when you attempt to register the disk device group you encounter the error message scconf: Failed to add device group - in use, reminor the disk device group. Use the procedure "SPARC: How to Assign a New Minor Number to a Disk Device Group" on page 188. This procedure enables you to assign a new minor number that does not conflict with a minor number that is used by existing disk device groups.

**Stack overflow** – If a stack overflows when the disk device group is brought online, the default value of the thread stack size might be insufficient. On each node, add the entry set cl_comm:rm_thread_stacksize=0x*size* to the /etc/system file, where *size* is a number greater than 8000, which is the default setting.

**Configuration changes** – If you change any configuration information for a VxVM disk group or volume, you must register the configuration changes by using the scsetup utility. Configuration changes you must register include adding or removing volumes and changing the group, owner, or permissions of existing volumes. See "Administering Disk Device Groups" in *Sun Cluster System Administration Guide for Solaris OS* for procedures to register configuration changes to a disk device group.

## ▼ SPARC: How to Assign a New Minor Number to a Disk Device Group

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. Perform this procedure to reminor a disk group.

**Steps** 1. **Become superuser on a node of the cluster.**

2. **Determine the minor numbers in use.**

   # **ls -l /global/.devices/node@1/dev/vx/dsk/***

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

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

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

**Example 4–2** 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 reminors the new disk device group to use the base minor number 5000.

```
# ls -l /global/.devices/node@1/dev/vx/dsk/*
/global/.devices/node@1/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@1/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
```

**Next Steps**    Register the disk group as a Sun Cluster disk device group. Go to .

## ▼ SPARC: How to Verify the Disk Group Configuration

Perform this procedure on each node of the cluster.

**Steps**    1. **Verify that only the local disks are included in the root disk group, and disk groups are imported on the current primary node only.**

```
# vxdisk list
```

2. **Verify that all volumes have been started.**

```
# vxprint
```

3. **Verify that all disk groups have been registered as Sun Cluster disk device groups and are online.**

```
# scstat -D
```

**Next Steps**    Go to .

---

# SPARC: Unencapsulating the Root Disk

This section describes how to unencapsulate the root disk in a Sun Cluster configuration.

## ▼ SPARC: How to Unencapsulate the Root Disk

Perform this procedure to unencapsulate the root disk.

**Before You Begin**    Perform the following tasks:

- Ensure that only Solaris root file systems are present on the root disk. The Solaris root file systems are root (/), swap, the global devices namespace, /usr, /var, /opt, and /home.

- Back up and remove from the root disk any file systems other than Solaris root file systems that reside on the root disk.

**Steps**  **1. Become superuser on the node that you intend to unencapsulate.**

**2. Move all resource groups and device groups from the node.**

```
# scswitch -S -h from-node
```

-S                Moves all resource groups and device groups

-h *from-node*     Specifies the name of the node from which to move resource or device groups

**3. Determine the node-ID number of the node.**

```
# clinfo -n
```

**4. Unmount the global-devices file system for this node, where *N* is the node ID number that is returned in Step 3.**

```
# umount /global/.devices/node@N
```

**5. View the /etc/vfstab file and determine which VxVM volume corresponds to the global-devices file system.**

```
# vi /etc/vfstab
#device         device          mount      FS      fsck    mount     mount
#to mount       to fsck         point      type    pass    at boot   options
#
#NOTE: volume rootdiskxNvol (/global/.devices/node@N) encapsulated
#partition cNtXdYsZ
```

**6. Remove from the root disk group the VxVM volume that corresponds to the global-devices file system.**

```
# vxedit -g rootdiskgroup -rf rm rootdiskxNvol
```

**Caution –** Do not store data other than device entries for global devices in the global-devices file system. All data in the global-devices file system is destroyed when you remove the VxVM volume. Only data that is related to global devices entries is restored after the root disk is unencapsulated.

**7. Unencapsulate the root disk.**

---

**Note** – Do **not** accept the shutdown request from the command.

---

```
# /etc/vx/bin/vxunroot
```

See your VxVM documentation for details.

8. **Use the `format(1M)` command to add a 512-Mbyte partition to the root disk to use for the global-devices file system.**

---

**Tip** – Use the same slice that was allocated to the global-devices file system before the root disk was encapsulated, as specified in the /etc/vfstab file.

---

9. **Set up a file system on the partition that you created in Step 8.**

   ```
   # newfs /dev/rdsk/cNtXdYsZ
   ```

10. **Determine the DID name of the root disk.**

    ```
    # scdidadm -l cNtXdY
    1           phys-schost-1:/dev/rdsk/cNtXdY    /dev/did/rdsk/dN
    ```

11. **In the `/etc/vfstab` file, replace the path names in the global-devices file system entry with the DID path that you identified in Step 10.**

    The original entry would look similar to the following.

```
# vi /etc/vfstab
/dev/vx/dsk/rootdiskxNvol /dev/vx/rdsk/rootdiskxNvol /global/.devices/node@N ufs 2 no global
```

    The revised entry that uses the DID path would look similar to the following.

```
/dev/did/dsk/dNsX /dev/did/rdsk/dNsX /global/.devices/node@N ufs 2 no global
```

12. **Mount the global-devices file system.**

    ```
    # mount /global/.devices/node@N
    ```

13. **From one node of the cluster, repopulate the global-devices file system with device nodes for any raw-disk devices and Solstice DiskSuite or Solaris Volume Manager devices.**

    ```
    # scgdevs
    ```
    VxVM devices are recreated during the next reboot.

14. **Reboot the node.**

    ```
    # reboot
    ```

**15. Repeat this procedure on each node of the cluster to unencapsulate the root disk on those nodes.**

# Upgrading Sun Cluster Software

This chapter provides the following information and procedures to upgrade a Sun Cluster 3.x configuration to Sun Cluster 3.1 8/05 software:

- "Overview of Upgrading a Sun Cluster Configuration" on page 193
- "Performing a Nonrolling Upgrade" on page 195
- "Performing a Rolling Upgrade" on page 220
- "Recovering From Storage Configuration Changes During Upgrade" on page 240
- "SPARC: Upgrading Sun Management Center Software" on page 242

# Overview of Upgrading a Sun Cluster Configuration

This section provides the following guidelines to upgrade a Sun Cluster configuration:

- "Upgrade Requirements and Software Support Guidelines" on page 193
- "Choosing a Sun Cluster Upgrade Method" on page 194

## Upgrade Requirements and Software Support Guidelines

Observe the following requirements and software-support guidelines when you upgrade to Sun Cluster 3.1 8/05 software:

- **Supported hardware** - The cluster hardware must be a supported configuration for Sun Cluster 3.1 8/05 software. Contact your Sun representative for information about current supported Sun Cluster configurations.

- **Architecture changes during upgrade** - Sun Cluster 3.1 8/05 software does not support upgrade between architectures.

- **Minimum Solaris OS** - The cluster must run on or be upgraded to at least Solaris 8 2/02 software, including the most current required patches.

- **Restriction on upgrade to the March 2005 distribution of the Solaris 10 OS** - Sun Cluster 3.1 8/05 software does not support upgrade to the original release of the Solaris 10 OS, which was distributed in March 2005. You must upgrade to at least Solaris 10 10/05 software or compatible.

- **Upgrading between Solaris major releases** - Sun Cluster 3.1 8/05 software supports only nonrolling upgrade from Solaris 8 software to Solaris 9 software or from Solaris 9 software to Solaris 10 10/05 software or compatible.

- **Upgrading to compatible versions** - You must upgrade all software to a version that is supported by Sun Cluster 3.1 8/05 software. For example, if a data service is supported on Sun Cluster 3.0 software but is not supported on Sun Cluster 3.1 8/05 software, you must upgrade that data service to the version of that data service that is supported on Sun Cluster 3.1 8/05 software. See "Supported Products" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for support information about specific data services.

  If the related application of a data service is not supported on Sun Cluster 3.1 8/05 software, you must upgrade that application to a supported release.

- **Minimum Sun Cluster software version** - Sun Cluster 3.1 8/05 software supports direct upgrade only from Sun Cluster 3.x software.

- **Converting from NAFO to IPMP groups** - For upgrade from a Sun Cluster 3.0 release, have available the test IP addresses to use with your public-network adapters when NAFO groups are converted to Internet Protocol (IP) Network Multipathing groups. The `scinstall` upgrade utility prompts you for a test IP address for each public-network adapter in the cluster. A test IP address must be on the same subnet as the primary IP address for the adapter.

  See the *IP Network Multipathing Administration Guide* (Solaris 8) or "IPMP," in *System Administration Guide: IP Services* (Solaris 9 or Solaris 10) for information about test IP addresses for IP Network Multipathing groups.

- **Downgrade** - Sun Cluster 3.1 8/05 software does not support any downgrade of Sun Cluster software.

- **Limitation of scinstall for data-service upgrades** - The `scinstall` upgrade utility only upgrades those data services that are provided with Sun Cluster 3.1 8/05 software. You must manually upgrade any custom or third-party data services.

## Choosing a Sun Cluster Upgrade Method

Choose from the following methods to upgrade your cluster to Sun Cluster 3.1 8/05 software:

- **Nonrolling upgrade –** In a *nonrolling upgrade*, you shut down the cluster before you upgrade the cluster nodes. You return the cluster to production after all nodes are fully upgraded. You **must** use the nonrolling-upgrade method if one or more of the following conditions apply:

- You are upgrading from Sun Cluster 3.0 software.
- You are upgrading from Solaris 8 software to Solaris 9 software or from Solaris 9 software to Solaris 10 10/05 software or compatible.
- Any software products that you are upgrading, such as applications or databases, require that the same version of the software is running on all cluster nodes at the same time.
- You are upgrading the Sun Cluster module software for Sun Management Center.
- You are also upgrading VxVM or VxFS.

- **Rolling upgrade –** In a *rolling upgrade*, you upgrade one node of the cluster at a time. The cluster remains in production with services running on the other nodes. You can use the rolling-upgrade method only if **all** of the following conditions apply:
  - You are upgrading from Sun Cluster 3.1 software.
  - You are upgrading the Solaris operating system only to a Solaris update, if at all.
  - For any applications or databases you must upgrade, the current version of the software can coexist in a running cluster with the upgrade version of that software.

  If your cluster configuration meets the requirements to perform a rolling upgrade, you can still choose to perform a nonrolling upgrade instead. A nonrolling upgrade might be preferable to a rolling upgrade if you wanted to use the Cluster Control Panel to issue commands to all cluster nodes at the same time and you could tolerate the cluster downtime.

For overview information about planning your Sun Cluster 3.1 8/05 configuration, see Chapter 1.

# Performing a Nonrolling Upgrade

Follow the tasks in this section to perform a nonrolling upgrade from Sun Cluster 3.x software to Sun Cluster 3.1 8/05 software. In a nonrolling upgrade, you shut down the entire cluster before you upgrade the cluster nodes. This procedure also enables you to upgrade the cluster from Solaris 8 software to Solaris 9 software or from Solaris 9 software to Solaris 10 10/05 software or compatible.

---

**Note –** To perform a rolling upgrade to Sun Cluster 3.1 8/05 software, instead follow the procedures in "Performing a Rolling Upgrade" on page 220.

---

**TABLE 5–1** Task Map: Performing a Nonrolling Upgrade to Sun Cluster 3.1 8/05 Software

| Task | Instructions |
|------|-------------|
| 1. Read the upgrade requirements and restrictions. | "Upgrade Requirements and Software Support Guidelines" on page 193 |
| 2. Remove the cluster from production and back up shared data. If the cluster uses dual-string mediators for Solstice DiskSuite or Solaris Volume Manager software, unconfigure the mediators. | "How to Prepare the Cluster for a Nonrolling Upgrade" on page 196 |
| 3. Upgrade the Solaris software, if necessary, to a supported Solaris update. Optionally, upgrade VERITAS Volume Manager (VxVM). | "How to Perform a Nonrolling Upgrade of the Solaris OS" on page 201 |
| 4. Install or upgrade software on which Sun Cluster 3.1 8/05 software has a dependency. | "How to Upgrade Dependency Software Before a Nonrolling Upgrade" on page 205 |
| 5. Upgrade to Sun Cluster 3.1 8/05 framework and data-service software. If necessary, upgrade applications. If the cluster uses dual-string mediators, reconfigure the mediators. SPARC: If you upgraded VxVM, upgrade disk groups. | "How to Perform a Nonrolling Upgrade of Sun Cluster 3.1 8/05 Software" on page 210 |
| 6. Enable resources and bring resource groups online. Optionally, migrate existing resources to new resource types. | "How to Finish a Nonrolling Upgrade to Sun Cluster 3.1 8/05 Software" on page 217 |
| 7. (*Optional*) SPARC: Upgrade the Sun Cluster module for Sun Management Center, if needed. | "SPARC: How to Upgrade Sun Cluster Module Software for Sun Management Center" on page 242 |

## ▼ How to Prepare the Cluster for a Nonrolling Upgrade

Perform this procedure to remove the cluster from production.

**Before You Begin**   Perform the following tasks:

- Ensure that the configuration meets requirements for upgrade. See "Upgrade Requirements and Software Support Guidelines" on page 193.

- Have available the CD-ROMs, documentation, and patches for all software products you are upgrading, including the following software:

  - Solaris OS
  - Sun Cluster 3.1 8/05 framework
  - Sun Cluster 3.1 8/05 data services (agents)
  - Applications that are managed by Sun Cluster 3.1 8/05 data-service agents
  - SPARC: VERITAS Volume Manager, if applicable

See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

■ If you are upgrading from Sun Cluster 3.0 software, have available your list of test IP addresses. Each public-network adapter in the cluster must have at least one test IP address. This requirement applies regardless of whether the adapter is the active adapter or the backup adapter in the group. The test IP addresses are used to reconfigure the adapters to use IP Network Multipathing.

---

**Note –** Each test IP address must be on the same subnet as the existing IP address that is used by the public-network adapter.

---

To list the public-network adapters on a node, run the following command:

```
% pnmstat
```

See one of the following manuals for more information about test IP addresses for IP Network Multipathing:

■ *IP Network Multipathing Administration Guide* (Solaris 8)

■ "Configuring Test Addresses" in "Administering Multipathing Groups With Multiple Physical Interfaces" in *System Administration Guide: IP Services* (Solaris 9)

■ "Test Addresses" in *System Administration Guide: IP Services* (Solaris 10)

**Steps** **1. Ensure that the cluster is functioning normally.**

■ To view the current status of the cluster, run the following command from any node:

```
% scstat
```

See the scstat(1M) man page for more information.

■ Search the /var/adm/messages log on the same node for unresolved error messages or warning messages.

■ Check the volume-manager status.

**2. (Optional) Install Sun Cluster 3.1 8/05 documentation.**

Install the documentation packages on your preferred location, such as an administrative console or a documentation server. See the Solaris_*arch*/Product/sun_cluster/index.html file on the Sun Cluster 2 of 2 CD-ROM, where *arch* is sparc or x86, to access installation instructions.

**3. Notify users that cluster services will be unavailable during the upgrade.**

**4. Become superuser on a node of the cluster.**

5. **Start the `scsetup(1m)` utility.**

   `# scsetup`

   The Main Menu is displayed.

6. **Switch each resource group offline.**

   a. **From the `scsetup` Main Menu, choose the menu item, Resource groups.**

   b. **From the Resource Group Menu, choose the menu item, Online/Offline or Switchover a resource group.**

   c. **Follow the prompts to take offline all resource groups and to put them in the unmanaged state.**

   d. **When all resource groups are offline, type `q` to return to the Resource Group Menu.**

7. **Disable all resources in the cluster.**

   The disabling of resources before upgrade prevents the cluster from bringing the resources online automatically if a node is mistakenly rebooted into cluster mode.

   a. **From the Resource Group Menu, choose the menu item, Enable/Disable a resource.**

   b. **Choose a resource to disable and follow the prompts.**

   c. **Repeat Step b for each resource.**

   d. **When all resources are disabled, type `q` to return to the Resource Group Menu.**

8. **Exit the `scsetup` utility.**

   Type `q` to back out of each submenu or press `Ctrl-C`.

9. **Verify that all resources on all nodes are `Offline` and that all resource groups are in the `Unmanaged` state.**

   `# scstat -g`

10. **If your cluster uses dual-string mediators for Solstice DiskSuite or Solaris Volume Manager software, unconfigure your mediators.**

    See "Configuring Dual-String Mediators" on page 172 for more information.

    a. **Run the following command to verify that no mediator data problems exist.**

       `# medstat -s` *setname*

       `-s setname`    Specifies the disk set name

       If the value in the Status field is `Bad`, repair the affected mediator host. Follow the procedure "How to Fix Bad Mediator Data" on page 174.

b. **List all mediators.**

Save this information for when you restore the mediators during the procedure "How to Finish a Nonrolling Upgrade to Sun Cluster 3.1 8/05 Software" on page 217.

c. **For a disk set that uses mediators, take ownership of the disk set if no node already has ownership.**

# **scswitch -z -D** *setname* **-h** *node*

-z        Changes mastery

-D        Specifies the name of the disk set

-h *node*   Specifies the name of the node to become primary of the disk set

d. **Unconfigure all mediators for the disk set.**

# **metaset -s** *setname* **-d -m** *mediator-host-list*

| -s setname | Specifies the disk set name |
|---|---|
| -d | Deletes from the disk set |
| -m mediator-host-list | Specifies the name of the node to remove as a mediator host for the disk set |

See the mediator(7D) man page for further information about mediator-specific options to the metaset command.

e. **Repeat Step c through Step d for each remaining disk set that uses mediators.**

11. **For a two-node cluster that uses Sun StorEdge Availability Suite software, ensure that the configuration data for availability services resides on the quorum disk.**

The configuration data must reside on a quorum disk to ensure the proper functioning of Sun StorEdge Availability Suite after you upgrade the cluster software.

a. **Become superuser on a node of the cluster that runs Sun StorEdge Availability Suite software.**

b. **Identify the device ID and the slice that is used by the Sun StorEdge Availability Suite configuration file.**

# **/usr/opt/SUNWscm/sbin/dscfg**
/dev/did/rdsk/d*N*s*S*

In this example output, *N* is the device ID and *S* the slice of device *N*.

c. **Identify the existing quorum device.**

# **scstat -q**
-- Quorum Votes by Device --
                    Device Name        Present Possible Status

```
                   -----------       ------- -------- ------
    Device votes:    /dev/did/rdsk/dQsS  1       1        Online
```
In this example output, d*QsS* is the existing quorum device.

**d. If the quorum device is not the same as the Sun StorEdge Availability Suite configuration-data device, move the configuration data to an available slice on the quorum device.**

```
# dd if='/usr/opt/SUNWesm/sbin/dscfg' of=/dev/did/rdsk/dQsS
```

---

**Note –** You must use the name of the raw DID device, `/dev/did/rdsk/`, not the block DID device, `/dev/did/dsk/`.

---

**e. If you moved the configuration data, configure Sun StorEdge Availability Suite software to use the new location.**

As superuser, issue the following command on each node that runs Sun StorEdge Availability Suite software.

```
# /usr/opt/SUNWesm/sbin/dscfg -s /dev/did/rdsk/dQsS
```

**12. Stop all applications that are running on each node of the cluster.**

**13. Ensure that all shared data is backed up.**

**14. From one node, shut down the cluster.**

```
# scshutdown -g0 -y
```
See the scshutdown(1M) man page for more information.

**15. Boot each node into noncluster mode.**

- **On SPARC based systems, perform the following command:**

    ```
    ok boot -x
    ```

- **On x86 based systems, perform the following commands:**

    ```
    ...
                        <<< Current Boot Parameters >>>
    Boot path: /pci@0,0/pci-ide@7,1/ata@1/cmdk@0,0:b
    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
    ```

**16. Ensure that each system disk is backed up.**

To upgrade Solaris software before you perform Sun Cluster software upgrade, go to "How to Perform a Nonrolling Upgrade of the Solaris OS" on page 201.

- If Sun Cluster 3.1 8/05 software does not support the release of the Solaris OS that you currently run on your cluster, you must upgrade the Solaris software to a supported release. See "Supported Products" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for more information.

- If Sun Cluster 3.1 8/05 software supports the release of the Solaris OS that you currently run on your cluster, further Solaris software upgrade is optional.

Otherwise, upgrade dependency software. Go to "How to Upgrade Dependency Software Before a Nonrolling Upgrade" on page 205.

## ▼ How to Perform a Nonrolling Upgrade of the Solaris OS

Perform this procedure on each node in the cluster to upgrade the Solaris OS. If the cluster already runs on a version of the Solaris OS that supports Sun Cluster 3.1 8/05 software, further upgrade of the Solaris OS is optional. If you do not intend to upgrade the Solaris OS, proceed to "How to Perform a Nonrolling Upgrade of Sun Cluster 3.1 8/05 Software" on page 210.

---

**⚠ Caution –** Sun Cluster 3.1 8/05 software does not support upgrade from the Solaris 9 OS to the original release of the Solaris 10 OS, which was distributed in March 2005. You must upgrade to at least the Solaris 10 10/05 release or compatible.

---

**Before You Begin** Perform the following tasks:

- Ensure that the cluster runs at least the minimum required level of the Solaris OS to support Sun Cluster 3.1 8/05 software. See "Supported Products" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for more information.

- Ensure that all steps in "How to Prepare the Cluster for a Nonrolling Upgrade" on page 196 are completed.

**Steps** 1. **Become superuser on the cluster node to upgrade.**

2. **(Optional) SPARC: Upgrade VxFS.**

   Follow procedures that are provided in your VxFS documentation.

3. **Determine whether the following Apache run control scripts exist and are enabled or disabled:**

```
/etc/rc0.d/K16apache
/etc/rc1.d/K16apache
/etc/rc2.d/K16apache
/etc/rc3.d/S50apache
/etc/rcS.d/K16apache
```

Some applications, such as Sun Cluster HA for Apache, require that Apache run control scripts be disabled.

- If these scripts exist and contain an uppercase K or S in the file name, the scripts are enabled. No further action is necessary for these scripts.
- If these scripts do not exist, in Step 8 you must ensure that any Apache run control scripts that are installed during the Solaris OS upgrade are disabled.
- If these scripts exist but the file names contain a lowercase k or s, the scripts are disabled. In Step 8 you must ensure that any Apache run control scripts that are installed during the Solaris OS upgrade are disabled.

4. **Comment out all entries for globally mounted file systems in the node's `/etc/vfstab` file.**

   a. **For later reference, make a record of all entries that are already commented out.**

   b. **Temporarily comment out all entries for globally mounted file systems in the `/etc/vfstab` file.**

   Entries for globally mounted file systems contain the `global` mount option. Comment out these entries to prevent the Solaris upgrade from attempting to mount the global devices.

5. **Determine which procedure to follow to upgrade the Solaris OS.**

| Volume Manager | Procedure | Location of Instructions |
|---|---|---|
| Solstice DiskSuite or Solaris Volume Manager | Any Solaris upgrade method *except* the Live Upgrade method | Solaris installation documentation |
| SPARC: VERITAS Volume Manager | "Upgrading VxVM and Solaris" | VERITAS Volume Manager installation documentation |

**Note –** If your cluster has VxVM installed, you must reinstall the existing VxVM software or upgrade to the Solaris 9 version of VxVM software as part of the Solaris upgrade process.

6. **Upgrade the Solaris software, following the procedure that you selected in Step 5.**

   Make the following changes to the procedures that you use:

   - **When you are instructed to reboot a node during the upgrade process, always reboot into noncluster mode.**

     - **For the `boot` and `reboot` commands, add the `-x` option to the command.**

       The `-x` option ensures that the node reboots into noncluster mode. For example, either of the following two commands boot a node into single-user noncluster mode:

       - On SPARC based systems, perform either of the following commands:

         ```
         # reboot -- -xs
              or
         ok boot -xs
         ```

       - On x86 based systems, perform either of the following commands:

         ```
         # reboot -- -xs
              or
         ...
                            <<< Current Boot Parameters >>>
         Boot path: /pci@0,0/pci-ide@7,1/ata@1/cmdk@0,0:b
         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 -xs
         ```

     - **If the instruction says to run the `init S` command, use the `reboot -- -xs` command instead.**

   - *Do not* **perform the final reboot instruction in the Solaris software upgrade. Instead, do the following:**

     a. **Return to this procedure to perform Step 7 and Step 8.**

     b. **Reboot into noncluster mode in Step 9 to complete Solaris software upgrade.**

7. **In the `/a/etc/vfstab` file, uncomment those entries for globally mounted file systems that you commented out in Step 4.**

8. **If Apache run control scripts were disabled or did not exist before you upgraded the Solaris OS, ensure that any scripts that were installed during Solaris upgrade are disabled.**

To disable Apache run control scripts, use the following commands to rename the files with a lowercase k or s.

```
# mv /a/etc/rc0.d/K16apache /a/etc/rc0.d/k16apache
# mv /a/etc/rc1.d/K16apache /a/etc/rc1.d/k16apache
# mv /a/etc/rc2.d/K16apache /a/etc/rc2.d/k16apache
# mv /a/etc/rc3.d/S50apache /a/etc/rc3.d/s50apache
# mv /a/etc/rcS.d/K16apache /a/etc/rcS.d/k16apache
```

Alternatively, you can rename the scripts to be consistent with your normal administration practices.

9. **Reboot the node into noncluster mode.**

   Include the double dashes (--) in the following command:

   ```
   # reboot -- -x
   ```

10. **SPARC: If your cluster runs VxVM, perform the remaining steps in the procedure "Upgrading VxVM and Solaris" to reinstall or upgrade VxVM.**

    Make the following changes to the procedure:

    - **After VxVM upgrade is complete but before you reboot, verify the entries in the `/etc/vfstab` file.**

      If any of the entries that you uncommented in Step 7 were commented out, make those entries uncommented again.

    - **When the VxVM procedures instruct you to perform a final reconfiguration reboot,** *do not* **use the `-r` option alone. Instead, reboot into noncluster mode by using the `-rx` options.**

      ```
      # reboot -- -rx
      ```

    ---

    **Note –** If you see a message similar to the following, type the root password to continue upgrade processing. Do *not* run the `fsck` command nor type Ctrl-D.

    ```
    WARNING - Unable to repair the /global/.devices/node@1 filesystem.
    Run fsck manually (fsck -F ufs /dev/vx/rdsk/rootdisk_13vol). Exit the
    shell when done to continue the boot process.

    Type control-d to proceed with normal startup,
    (or give root password for system maintenance):      Type the root password
    ```

    ---

11. **Install any required Solaris software patches and hardware-related patches, and download any needed firmware that is contained in the hardware patches.**

    For Solstice DiskSuite software (Solaris 8), also install any Solstice DiskSuite software patches.

---

**Note –** Do not reboot after you add patches. Wait to reboot the node until after you upgrade the Sun Cluster software.

---

See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

**Next Steps**  Upgrade dependency software. Go to "How to Upgrade Dependency Software Before a Nonrolling Upgrade" on page 205.

---

**Note –** To complete the upgrade from Solaris 8 to Solaris 9 software or from Solaris 9 to Solaris 10 10/05 software or compatible, you must also upgrade to the Solaris 9 or Solaris 10 version of Sun Cluster 3.1 8/05 software, including dependency software. You must perform this task even if the cluster already runs on Sun Cluster 3.1 8/05 software for another version of Solaris software.

---

## ▼ How to Upgrade Dependency Software Before a Nonrolling Upgrade

Perform this procedure on each cluster node to install or upgrade software on which Sun Cluster 3.1 8/05 software has a dependency. The cluster remains in production during this procedure.

If you are running SunPlex Manager, status on a node will not be reported during the period that the node's security file agent is stopped. Status reporting resumes when the security file agent is restarted, after the common agent container software is upgraded.

**Before You Begin**  Perform the following tasks:

- Ensure that all steps in "How to Prepare the Cluster for a Nonrolling Upgrade" on page 196 are completed.
- If you upgraded from Solaris 8 to Solaris 9 software or from Solaris 9 to Solaris 10 10/05 software or compatible, ensure that all steps in "How to Perform a Nonrolling Upgrade of the Solaris OS" on page 201 are completed.
- Ensure that you have installed all required Solaris software patches and hardware-related patches.
- If the cluster runs Solstice DiskSuite software (Solaris 8), ensure that you have installed all required Solstice DiskSuite software patches.

**Steps**  1. **Become superuser on the cluster node.**

2. **For the Solaris 8 and Solaris 9 OS, ensure that the Apache Tomcat package is at the required patch level, if the package is installed.**

   a. **Determine whether the `SUNWtcatu` package is installed.**

   ```
   # pkginfo SUNWtcatu
   SUNWtcatu        Tomcat Servlet/JSP Container
   ```

   b. **If the Apache Tomcat package is installed, determine whether the required patch for the platform is installed.**

   - SPARC based platforms require at least 114016-01
   - x86 based platforms require at least 114017-01

   ```
   # patchadd -p | grep 114016
   Patch: 114016-01 Obsoletes: Requires: Incompatibles: Packages: SUNWtcatu
   ```

   c. **If the required patch is not installed, remove the Apache Tomcat package.**

   ```
   # pkgrm SUNWtcatu
   ```

3. **Insert the Sun Cluster 1 of 2 CD-ROM.**

4. **Change to the `/cdrom/cdrom0/Solaris_arch/Product/shared_components/Packages/` directory, where *arch* is `sparc` or `x86` .**

   ```
   # cd /cdrom/cdrom0/Solaris_arch/Product/shared_components/Packages/
   ```

5. **Ensure that at least version 4.3.1 of the Explorer packages is installed.**

   These packages are required by Sun Cluster software for use by the sccheck utility.

   a. **Determine whether the Explorer packages are installed and, if so, what version.**

   ```
   # pkginfo -l SUNWexplo | grep SUNW_PRODVERS
   SUNW_PRODVERS=4.3.1
   ```

   b. **If a version earlier than 4.3.1 is installed, remove the existing Explorer packages.**

   ```
   # pkgrm SUNWexplo SUNWexplu SUNWexplj
   ```

   c. **If you removed Explorer packages or none were installed, install the latest Explorer packages from the Sun Cluster 1 of 2 CD-ROM.**

   - **For the Solaris 8 or Solaris 9 OS, use the following command:**

     ```
     # pkgadd -d . SUNWexpl*
     ```

   - **For the Solaris 10 OS, use the following command:**

     ```
     # pkgadd -G -d . SUNWexpl*
     ```

The `-G` option adds packages to the current zone only. You must add these packages only to the global zone. Therefore, this option also specifies that the packages are *not* propagated to any existing non-global zone or to any non-global zone that is created later.

6. **Ensure that at least version 5.1,REV=34 of the Java Dynamic Management Kit (JDMK) packages is installed.**

   a. **Determine whether JDMK packages are installed and, if so, what version.**

   ```
   # pkginfo -l SUNWjdmk-runtime | grep VERSION
   VERSION=5.1,REV=34
   ```

   b. **If a version earlier than 5.1,REV=34 is installed, remove the existing JDMK packages.**

   ```
   # pkgrm SUNWjdmk-runtime SUNWjdmk-runtime-jmx
   ```

   c. **If you removed JDMK packages or none were installed, install the latest JDMK packages from the Sun Cluster 1 of 2 CD-ROM.**

      ■ **For the Solaris 8 or Solaris 9 OS, use the following command:**

      ```
      # pkgadd -d . SUNWjdmk*
      ```

      ■ **For the Solaris 10 OS, use the following command:**

      ```
      # pkgadd -G -d . SUNWjdmk*
      ```

7. **Change to the `Solaris_`*arch*`/Product/shared_components/Solaris_`*ver*`/Packages/` directory, where** *arch* **is** `sparc` **or** `x86` **and where** *ver* **is** `8` **for Solaris 8,** `9` **for Solaris 9, or** `10` **for Solaris 10.**

   ```
   # cd ../Solaris_ver/Packages
   ```

8. **Ensure that at least version 4.5.0 of the Netscape Portable Runtime (NSPR) packages is installed.**

   a. **Determine whether NSPR packages are installed and, if so, what version.**

   ```
   # cat /var/sadm/pkg/SUNWpr/pkginfo | grep SUNW_PRODVERS
   SUNW_PRODVERS=4.5.0
   ```

   b. **If a version earlier than 4.5.0 is installed, remove the existing NSPR packages.**

   ```
   # pkgrm packages
   ```

   The following table lists the applicable packages for each hardware platform.

   ---

   **Note –** Install packages in the order in which they are listed in the following table.

   ---

| Hardware Platform | NSPR Package Names |
|---|---|
| SPARC | SUNWpr SUNWprx |
| x86 | SUNWpr |

    **c. If you removed NSPR packages or none were installed, install the latest NSPR packages.**

      ■ **For the Solaris 8 or Solaris 9 OS, use the following command:**

        # **pkgadd -d .** *packages*

      ■ **For the Solaris 10 OS, use the following command:**

        # **pkgadd -G -d .** *packages*

**9. Ensure that at least version 3.9.4 of the Network Security Services (NSS) packages is installed.**

    **a. Determine whether NSS packages are installed and, if so, what version.**

    # **cat /var/sadm/pkg/SUNWtls/pkginfo | grep SUNW_PRODVERS**
    SUNW_PRODVERS=3.9.4

    **b. If a version earlier than 3.9.4 is installed, remove the existing NSS packages.**

    # **pkgrm** *packages*

    The following table lists the applicable packages for each hardware platform.

    **Note –** Install packages in the order in which they are listed in the following table.

| Hardware Platform | NSS Package Names |
|---|---|
| SPARC | SUNWtls SUNWtlsu SUNWtlsx |
| x86 | SUNWtls SUNWtlsu |

    **c. If you removed NSS packages or none were installed, install the latest NSS packages from the Sun Cluster 1 of 2 CD-ROM.**

      ■ **For the Solaris 8 or Solaris 9 OS, use the following command:**

        # **pkgadd -d .** *packages*

      ■ **For the Solaris 10 OS, use the following command:**

        # **pkgadd -G -d .** *packages*

10. **Change back to the Solaris_*arch*/Product/shared_components//Packages/ directory.**

    ```
    # cd ../../Packages
    ```

11. **Ensure that at least version 1.0,REV=25 of the common agent container packages is installed.**

    a. **Determine whether the common agent container packages are installed and, if so, what version.**

    ```
    # pkginfo -l SUNWcacao | grep VERSION
    VERSION=1.0,REV=25
    ```

    b. **If a version earlier than 1.0,REV=25 is installed, stop the security file agent for the common agent container on each cluster node.**

    ```
    # /opt/SUNWcacao/bin/cacaoadm stop
    ```

    c. **If a version earlier than 1.0,REV=25 is installed, remove the existing common agent container packages.**

    ```
    # pkgrm SUNWcacao SUNWcacaocfg
    ```

    d. **If you removed the common agent container packages or none were installed, install the latest common agent container packages from the Sun Cluster 1 of 2 CD-ROM.**

    - **For the Solaris 8 or Solaris 9 OS, use the following command:**

      ```
      # pkgadd -d . SUNWcacao*
      ```

    - **For the Solaris 10 OS, use the following command:**

      ```
      # pkgadd -G -d . SUNWcacao*
      ```

12. **Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

    ```
    # eject cdrom
    ```

13. **Insert the Sun Cluster 2 of 2 CD-ROM.**

14. **For upgrade from Solaris 8 to Solaris 9 OS, install or upgrade Sun Java Web Console packages.**

    a. **Change to the Solaris_*arch*/Product/sunwebconsole/ directory, where *arch* is sparc or x86.**

    b. **Install the Sun Java Web Console packages.**

    ```
    # ./setup
    ```

    The setup command installs or upgrades all packages to support Sun Java Web Console.

**15. Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

```
# eject cdrom
```

**16. Ensure that the `/usr/java/` directory is a symbolic link to the minimum or latest version of Java software.**

Sun Cluster software requires at least version 1.4.2_03 of Java software.

**a. Determine what directory the `/usr/java/` directory is symbolically linked to.**

```
# ls -l /usr/java
lrwxrwxrwx   1 root   other    9 Apr 19 14:05 /usr/java -> /usr/j2se/
```

**b. Determine what version or versions of Java software are installed.**

The following are examples of commands that you can use to display the version of their related releases of Java software.

```
# /usr/j2se/bin/java -version
# /usr/java1.2/bin/java -version
# /usr/jdk/jdk1.5.0_01/bin/java -version
```

**c. If the `/usr/java/` directory is not symbolically linked to a supported version of Java software, recreate the symbolic link to link to a supported version of Java software.**

The following example shows the creation of a symbolic link to the /usr/j2se/ directory, which contains Java 1.4.2_03 software.

```
# rm /usr/java
# ln -s /usr/j2se /usr/java
```

**Next Steps**  Upgrade to Sun Cluster 3.1 8/05 software. Go to

## ▼ How to Perform a Nonrolling Upgrade of Sun Cluster 3.1 8/05 Software

Perform this procedure to upgrade each node of the cluster to Sun Cluster 3.1 8/05 software. You must also perform this procedure to complete cluster upgrade from Solaris 8 to Solaris 9 software or from Solaris 9 to Solaris 10 10/05 software or compatible.

---

**Tip –** You can perform this procedure on more than one node at the same time.

---

**Before You Begin**  Ensure that dependency software is installed or upgraded. See

**Steps** 1. **Become superuser on a node of the cluster.**

2. **Insert the Sun Cluster 2 of 2 CD-ROM in the CD-ROM drive on the node.**

   If the volume management daemon `vold`(1M) is running and is configured to manage CD-ROM devices, the daemon automatically mounts the CD-ROM on the `/cdrom/cdrom0/` directory.

3. **Change to the `Solaris_`*arch*`/Product/sun_cluster/Solaris_`*ver*`/Tools/` directory, where** *arch* **is `sparc` or `x86` and where** *ver* **is 8 for Solaris 8, 9 for Solaris 9, or `10` for Solaris 10 .**

   ```
   # cd /cdrom/cdrom0/Solaris_arch/Product/sun_cluster/Solaris_ver/Tools
   ```

4. **Start the `scinstall` utility.**

   ```
   # ./scinstall
   ```

   ---
   **Note –** Do not use the `/usr/cluster/bin/scinstall` command that is already installed on the node. You must use the `scinstall` command on the Sun Cluster 2 of 2 CD-ROM.

   ---

5. **From the Main Menu, choose the menu item, Upgrade this cluster node.**

   ```
      *** Main Menu ***

     Please select from one of the following (*) options:

       * 1) Install a cluster or cluster node
         2) Configure a cluster to be JumpStarted from this install server
       * 3) Add support for new data services to this cluster node
       * 4) Upgrade this cluster node
       * 5) Print release information for this cluster node

       * ?) Help with menu options
       * q) Quit

     Option:  4
   ```

6. **From the Upgrade Menu, choose the menu item, Upgrade Sun Cluster framework on this node.**

7. **Follow the menu prompts to upgrade the cluster framework.**

   During the Sun Cluster upgrade, `scinstall` might make one or more of the following configuration changes:

   - Convert NAFO groups to IP Network Multipathing groups but keep the original NAFO-group name.

See one of the following manuals for information about test addresses for IP Network Multipathing:

- *IP Network Multipathing Administration Guide* (Solaris 8)
- "Configuring Test Addresses" in "Administering Multipathing Groups With Multiple Physical Interfaces" in *System Administration Guide: IP Services* (Solaris 9)
- "Test Addresses" in *System Administration Guide: IP Services* (Solaris 10)

See the scinstall(1M) man page for more information about the conversion of NAFO groups to IP Network Multipathing during Sun Cluster software upgrade.

- Rename the ntp.conf file to ntp.conf.cluster, if ntp.conf.cluster does not already exist on the node.
- Set the local-mac-address? variable to true, if the variable is not already set to that value.

Upgrade processing is finished when the system displays the message Completed Sun Cluster framework upgrade and prompts you to press Enter to continue.

8. **Press Enter.**

The Upgrade Menu is displayed.

9. **(Optional) Upgrade Java Enterprise System data services from the Sun Cluster 2 of 2 CD-ROM.**

   a. **From the Upgrade Menu of the scinstall utility, choose the menu item, Upgrade Sun Cluster data service agents on this node.**

   b. **Follow the menu prompts to upgrade Sun Cluster data service agents that are installed on the node.**

   You can choose from the list of data services that are available to upgrade or choose to upgrade all installed data services.

   Upgrade processing is finished when the system displays the message Completed upgrade of Sun Cluster data services agents and prompts you to press Enter to continue.

   c. **Press Enter.**

   The Upgrade Menu is displayed.

10. **Quit the scinstall utility.**

11. **Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

    # **eject cdrom**

12. **Upgrade Sun Cluster data services from the Sun Cluster 2 of 2 CD-ROM.**

- If you are using the Sun Cluster HA for NFS data service and you upgrade to the Solaris 10 OS, you must upgrade the data service and migrate the resource type to the new version. See "Upgrading the SUNW.nfs Resource Type" in *Sun Cluster Data Service for NFS Guide for Solaris OS* for more information.

- If you are using the Sun Cluster HA for Oracle 3.0 64-bit for Solaris 9 data service, you must upgrade to the Sun Cluster 3.1 8/05 version.

- The upgrade of any other data services to the Sun Cluster 3.1 8/05 version is optional. You can continue to use any other Sun Cluster 3.x data services after you upgrade the cluster to Sun Cluster 3.1 8/05 software.

Only those data services that are delivered on the Sun Cluster Agents CD are automatically upgraded by the `scinstall`(1M) utility. You must manually upgrade any custom or third-party data services. Follow the procedures that are provided with those data services.

a. **Insert the Sun Cluster Agents CD in the CD-ROM drive on the node.**

b. **Start the `scinstall` utility.**

   For data-service upgrades, you can use the `/usr/cluster/bin/scinstall` command that is already installed on the node.

   ```
   # scinstall
   ```

c. **From the Main Menu, choose the menu item, Upgrade this cluster node.**

d. **From the Upgrade Menu, choose the menu item, Upgrade Sun Cluster data service agents on this node.**

e. **Follow the menu prompts to upgrade Sun Cluster data service agents that are installed on the node.**

   You can choose from the list of data services that are available to upgrade or choose to upgrade all installed data services.

   Upgrade processing is finished when the system displays the message `Completed upgrade of Sun Cluster data services agents` and prompts you to press Enter to continue.

f. **Press Enter.**

   The Upgrade Menu is displayed.

g. **Quit the `scinstall` utility.**

h. **Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

   ```
   # eject cdrom
   ```

13. **As needed, manually upgrade any custom data services that are not supplied on the product media.**

14. **Verify that each data-service update is installed successfully.**

View the upgrade log file that is referenced at the end of the upgrade output messages.

15. **Install any Sun Cluster 3.1 8/05 software patches, if you did not already install them by using the `scinstall` command.**

16. **Install any Sun Cluster 3.1 8/05 data-service software patches.**

    See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

17. **Upgrade software applications that are installed on the cluster.**

    Ensure that application levels are compatible with the current versions of Sun Cluster and Solaris software. See your application documentation for installation instructions.

18. **After all nodes are upgraded, reboot each node into the cluster.**

    ```
    # reboot
    ```

19. **Copy the security files for the common agent container to all cluster nodes.**

    This step ensures that security files for the common agent container are identical on all cluster nodes and that the copied files retain the correct file permissions.

    a. **On each node, stop the Sun Java Web Console agent.**

       ```
       # /usr/sbin/smcwebserver stop
       ```

    b. **On each node, stop the security file agent.**

       ```
       # /opt/SUNWcacao/bin/cacaoadm stop
       ```

    c. **On one node, change to the `/etc/opt/SUNWcacao/` directory.**

       ```
       phys-schost-1# cd /etc/opt/SUNWcacao/
       ```

    d. **Create a tar file of the `/etc/opt/SUNWcacao/security/` directory.**

       ```
       phys-schost-1# tar cf /tmp/SECURITY.tar security
       ```

    e. **Copy the `/tmp/SECURITY.tar` file to each of the other cluster nodes.**

    f. **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
       ```

    g. **Delete the `/tmp/SECURITY.tar` file from each node in the cluster.**

You must delete each copy of the tar file to avoid security risks.

```
phys-schost-1# rm /tmp/SECURITY.tar
phys-schost-2# rm /tmp/SECURITY.tar
```

**h. On each node, start the security file agent.**

```
phys-schost-1# /opt/SUNWcacao/bin/cacaoadm start
phys-schost-2# /opt/SUNWcacao/bin/cacaoadm start
```

**i. On each node, start the Sun Java Web Console agent.**

```
phys-schost-1# /usr/sbin/smcwebserver start
phys-schost-2# /usr/sbin/smcwebserver start
```

**Next Steps**    Go to "How to Verify a Nonrolling Upgrade of Sun Cluster 3.1 8/05 Software" on page 215

## ▼ How to Verify a Nonrolling Upgrade of Sun Cluster 3.1 8/05 Software

Perform this procedure to verify that the cluster is successfully upgraded to Sun Cluster 3.1 8/05 software.

**Before You Begin**    Ensure that all upgrade procedures are completed for all cluster nodes that you are upgrading.

**Steps**    **1. On each upgraded node, view the installed levels of Sun Cluster software.**

```
# scinstall -pv
```

The first line of output states which version of Sun Cluster software the node is running. This version should match the version that you just upgraded to.

**2. From any node, verify that all upgraded cluster nodes are running in cluster mode (Online).**

```
# scstat -n
```

See the scstat(1M) man page for more information about displaying cluster status.

**3. If you upgraded from Solaris 8 to Solaris 9 software, verify the consistency of the storage configuration.**

**a. On each node, run the following command to verify the consistency of the storage configuration.**

```
# scdidadm -c
```

-c      Performs a consistency check

> ⚠ **Caution –** Do not proceed to Step b until your configuration passes this consistency check. Failure to pass this check might result in errors in device identification and cause data corruption.

The following table lists the possible output from the scdidadm -c command and the action you must take, if any.

| Example Message | Action |
|---|---|
| device id for 'phys-schost-1:/dev/rdsk/c1t3d0' does not match physical device's id, device may have been replaced | Go to "Recovering From Storage Configuration Changes During Upgrade" on page 240 and perform the appropriate repair procedure. |
| device id for 'phys-schost-1:/dev/rdsk/c0t0d0' needs to be updated, run scdidadm –R to update | None. You update this device ID in Step b. |
| No output message | None. |

See the scdidadm(1M) man page for more information.

**b. On each node, migrate the Sun Cluster storage database to Solaris 9 device IDs.**

   ```
   # scdidadm -R all
   ```

   -R      Performs repair procedures

   all     Specifies all devices

**c. On each node, run the following command to verify that storage database migration to Solaris 9 device IDs is successful.**

   ```
   # scdidadm -c
   ```

   ■ If the scdidadm command displays a message, return to Step a to make further corrections to the storage configuration or the storage database.

   ■ If the scdidadm command displays no messages, the device-ID migration is successful. When device-ID migration is verified on all cluster nodes, proceed to "How to Finish a Nonrolling Upgrade to Sun Cluster 3.1 8/05 Software" on page 217.

**Example 5–1**   Verifying a Nonrolling Upgrade From Sun Cluster 3.0 to Sun Cluster 3.1 8/05 Software

The following example shows the commands used to verify a nonrolling upgrade of a two-node cluster from Sun Cluster 3.0 to Sun Cluster 3.1 8/05 software on the Solaris 8 OS. The cluster node names are phys-schost-1 and phys-schost-2.

```
        (Verify that software versions are the same on all nodes)
# scinstall -pv

        (Verify cluster membership)
# scstat -n
-- Cluster Nodes --
                        Node name      Status
                        ---------      ------
  Cluster node:     phys-schost-1  Online
  Cluster node:     phys-schost-2  Online
```

**Next Steps**  Go to "How to Finish a Nonrolling Upgrade to Sun Cluster 3.1 8/05 Software" on page 217.

## ▼ How to Finish a Nonrolling Upgrade to Sun Cluster 3.1 8/05 Software

Perform this procedure to finish Sun Cluster upgrade. First, reregister all resource types that received a new version from the upgrade. Second, modify eligible resources to use the new version of the resource type that the resource uses. Third, re-enable resources. Finally, bring resource groups back online.

**Before You Begin**  Ensure that all steps in "How to Verify a Nonrolling Upgrade of Sun Cluster 3.1 8/05 Software" on page 215 are completed.

**Steps**  1. **If you upgraded any data services that are not supplied on the product media, register the new resource types for those data services.**

Follow the documentation that accompanies the data services.

2. **If you upgraded Sun Cluster HA for SAP liveCache from the version for Sun Cluster 3.0 to the version for Sun Cluster 3.1, modify the /opt/SUNWsclc/livecache/bin/lccluster configuration file.**

   a. **Become superuser on a node that will host the liveCache resource.**

   b. **Copy the new /opt/SUNWsclc/livecache/bin/lccluster file to the /sapdb/*LC_NAME*/db/sap/ directory.**

   Overwrite the lccluster file that already exists from the previous configuration of the data service.

   c. **Configure this /sapdb/*LC_NAME*/db/sap/lccluster file as documented in "How to Register and Configure Sun Cluster HA for SAP liveCache" in** *Sun Cluster Data Service for SAP liveCache Guide for Solaris OS.*

3. **If your configuration uses dual-string mediators for Solstice DiskSuite or Solaris Volume Manager software, restore the mediator configurations.**

   a. **Determine which node has ownership of a disk set to which you will add the mediator hosts.**

   `# `**`metaset -s`** *setname*

   -s *setname*     Specifies the disk set name

   b. **If no node has ownership, take ownership of the disk set.**

   `# `**`scswitch -z -D`** *setname* **`-h`** *node*

   -z             Changes mastery

   -D *setname*     Specifies the name of the disk set

   -h *node*        Specifies the name of the node to become primary of the disk set

   c. **Re-create the mediators.**

   `# `**`metaset -s`** *setname* **`-a -m`** *mediator-host-list*

   -a                           Adds to the disk set

   -m mediator-host-list     Specifies the names of the nodes to add as mediator hosts for the disk set

   d. **Repeat these steps for each disk set in the cluster that uses mediators.**

4. **SPARC: If you upgraded VxVM, upgrade all disk groups.**

   a. **Bring online and take ownership of a disk group to upgrade.**

   `# `**`scswitch -z -D`** *setname* **`-h`** *thisnode*

   b. **Run the following command to upgrade a disk group to the highest version supported by the VxVM release you installed.**

   `# `**`vxdg upgrade`** *dgname*

   See your VxVM administration documentation for more information about upgrading disk groups.

   c. **Repeat for each remaining VxVM disk group in the cluster.**

5. **Migrate resources to new resource type versions.**

**Note –** If you upgrade to the Sun Cluster HA for NFS data service for the Solaris 10 OS, you must migrate to the new resource type version. See "Upgrading the SUNW.nfs Resource Type" in *Sun Cluster Data Service for NFS Guide for Solaris OS* for more information.

For all other data services, this step is optional.

See "Upgrading a Resource Type" in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*, which contains procedures which use the command line. Alternatively, you can perform the same tasks by using the Resource Group menu of the `scsetup` utility. The process involves performing the following tasks:

- Registration of the new resource type
- Migration of the eligible resource to the new version of its resource type
- Modification of the extension properties of the resource type as specified in the manual for the related data service

6. **From any node, start the `scsetup`(1M) utility.**

   ```
   # scsetup
   ```

7. **Re-enable all disabled resources.**

   a. **From the Resource Group Menu, choose the menu item, Enable/Disable a resource.**

   b. **Choose a resource to enable and follow the prompts.**

   c. **Repeat Step b for each disabled resource.**

   d. **When all resources are re-enabled, type `q` to return to the Resource Group Menu.**

8. **Bring each resource group back online.**

   a. **From the Resource Group Menu, choose the menu item, Online/Offline or Switchover a resource group.**

   b. **Follow the prompts to put each resource group into the managed state and then bring the resource group online.**

9. **When all resource groups are back online, exit the `scsetup` utility.**

   Type `q` to back out of each submenu, or press Ctrl-C.

**Next Steps**   If you have a SPARC based system and use Sun Management Center to monitor the cluster, go to .

Otherwise, the cluster upgrade is complete.

**See Also**   To upgrade future versions of resource types, see "Upgrading a Resource Type" in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

# Performing a Rolling Upgrade

This section provides procedures to perform a rolling upgrade from Sun Cluster 3.1 software to Sun Cluster 3.1 8/05 software. In a rolling upgrade, you upgrade one cluster node at a time, while the other cluster nodes remain in production. After all nodes are upgraded and have rejoined the cluster, you must commit the cluster to the new software version before you can use any new features.

To upgrade from Sun Cluster 3.0 software, follow instead the procedures in .

---

**Note –** Sun Cluster 3.1 8/05 software does not support rolling upgrade from Solaris 8 software to Solaris 9 software or from Solaris 9 software to Solaris 10 10/05 software. You can only upgrade Solaris software to an update release during Sun Cluster rolling upgrade. To upgrade a Sun Cluster configuration from Solaris 8 software to Solaris 9 software or from Solaris 9 software to Solaris 10 10/05 software or compatible, perform instead the procedures in .

---

**TABLE 5–2** Task Map: Performing a Rolling Upgrade to Sun Cluster 3.1 8/05 Software

| Task | Instructions |
|---|---|
| 1. Read the upgrade requirements and restrictions. | "Upgrade Requirements and Software Support Guidelines" on page 193 |
| 2. On one node of the cluster, move resource groups and device groups to another cluster node, and ensure that shared data and system disks are backed up. If the cluster uses dual-string mediators for Solstice DiskSuite or Solaris Volume Manager software, unconfigure the mediators. Then reboot the node into noncluster mode. | "How to Prepare a Cluster Node for a Rolling Upgrade" on page 221 |

| Task | Instructions |
|---|---|
| 3. Upgrade the Solaris OS on the cluster node, if necessary, to a supported Solaris update release. SPARC: Optionally, upgrade VERITAS File System (VxFS) and VERITAS Volume Manager (VxVM). | "How to Perform a Rolling Upgrade of a Solaris Maintenance Update" on page 225 |
| 4. On all cluster nodes, install or upgrade software on which Sun Cluster 3.1 8/05 software has a dependency. | "How to Upgrade Dependency Software Before a Rolling Upgrade" on page 226 |
| 5. Upgrade the cluster node to Sun Cluster 3.1 8/05 framework and data-service software. If necessary, upgrade applications. SPARC: If you upgraded VxVM, upgrade disk groups. Then reboot the node back into the cluster. | "How to Perform a Rolling Upgrade of Sun Cluster 3.1 8/05 Software" on page 232 |
| 6. Repeat Tasks 3 through 5 on each remaining node to upgrade. | |
| 7. Use the `scversions` command to commit the cluster to the upgrade. If the cluster uses dual-string mediators, reconfigure the mediators. Optionally, migrate existing resources to new resource types. | "How to Finish a Rolling Upgrade to Sun Cluster 3.1 8/05 Software" on page 237 |
| 8. (*Optional*) SPARC: Upgrade the Sun Cluster module to Sun Management Center. | "SPARC: How to Upgrade Sun Cluster Module Software for Sun Management Center" on page 242 |

## ▼ How to Prepare a Cluster Node for a Rolling Upgrade

Perform this procedure on one node at a time. You will take the upgraded node out of the cluster while the remaining nodes continue to function as active cluster members.

**Before You Begin**    Perform the following tasks:

- Ensure that the configuration meets requirements for upgrade. See "Upgrade Requirements and Software Support Guidelines" on page 193.

- Have available the CD-ROMs, documentation, and patches for all the software products you are upgrading before you begin to upgrade the cluster, including the following software:

  - Solaris OS
  - Sun Cluster 3.1 8/05 framework
  - Sun Cluster 3.1 8/05 data services (agents)
  - Applications that are managed by Sun Cluster 3.1 8/05 data-service agents

  See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

Observe the following guidelines when you perform a rolling upgrade:

- Do *not* make any changes to the cluster configuration during a rolling upgrade. For example, do not add to or change the cluster interconnect or quorum devices. If you need to make such a change, do so before you start the rolling upgrade procedure or wait until after all nodes are upgraded and the cluster is committed to the new software version.

- Limit the amount of time that you take to complete a rolling upgrade of all cluster nodes. After a node is upgraded, begin the upgrade of the next cluster node as soon as possible. You can experience performance penalties and other penalties when you run a mixed-version cluster for an extended period of time.

- Avoid installing new data services or issuing any administrative configuration commands during the upgrade.

- Until all nodes of the cluster are successfully upgraded and the upgrade is committed, new features that are introduced by the new release might not be available.

**Steps** 1. **(Optional) Install Sun Cluster 3.1 8/05 documentation.**

Install the documentation packages on your preferred location, such as an administrative console or a documentation server. See the `Solaris_arch/Product/sun_cluster/index.html` file on the Sun Cluster 2 of 2 CD-ROM, where *arch* is `sparc` or `x86`, to access installation instructions.

2. **If you are upgrading from the Sun Cluster 3.1 9/04 release, ensure that the latest Sun Cluster 3.1 Core Patch is installed.**

This Core Patch contains the code fix for 6210440, which is necessary to enable rolling upgrade from Sun Cluster 3.1 9/04 software to Sun Cluster 3.1 8/05 software.

3. **Become superuser on one node of the cluster to upgrade.**

4. **For a two-node cluster that uses Sun StorEdge Availability Suite software, ensure that the configuration data for availability services resides on the quorum disk.**

The configuration data must reside on a quorum disk to ensure the proper functioning of Sun StorEdge Availability Suite after you upgrade the cluster software.

   a. **Become superuser on a node of the cluster that runs Sun StorEdge Availability Suite software.**

   b. **Identify the device ID and the slice that is used by the Sun StorEdge Availability Suite configuration file.**

   ```
   # /usr/opt/SUNWscm/sbin/dscfg
   /dev/did/rdsk/dNsS
   ```

   In this example output, *N* is the device ID and *S* the slice of device *N*.

c. **Identify the existing quorum device.**

```
# scstat -q
-- Quorum Votes by Device --
                      Device Name        Present Possible Status
                      -----------        ------- -------- ------
    Device votes:     /dev/did/rdsk/dQsS  1       1        Online
```

In this example output, d*Qs*S is the existing quorum device.

d. **If the quorum device is not the same as the Sun StorEdge Availability Suite configuration-data device, move the configuration data to an available slice on the quorum device.**

```
# dd if=`/usr/opt/SUNWesm/sbin/dscfg` of=/dev/did/rdsk/dQsS
```

---

**Note –** You must use the name of the raw DID device, `/dev/did/rdsk/`, not the block DID device, `/dev/did/dsk/`.

---

e. **If you moved the configuration data, configure Sun StorEdge Availability Suite software to use the new location.**

As superuser, issue the following command on each node that runs Sun StorEdge Availability Suite software.

```
# /usr/opt/SUNWesm/sbin/dscfg -s /dev/did/rdsk/dQsS
```

5. **From any node, view the current status of the cluster.**

Save the output as a baseline for later comparison.

```
% scstat
% scrgadm -pv[v]
```

See the scstat(1M) and scrgadm(1M) man pages for more information.

6. **Move all resource groups and device groups that are running on the node to upgrade.**

```
# scswitch -S -h from-node
```

-S              Moves all resource groups and device groups

-h *from-node*   Specifies the name of the node from which to move resource groups and device groups

See the scswitch(1M) man page for more information.

7. **Verify that the move was completed successfully.**

```
# scstat -g -D
```

-g    Shows status for all resource groups

-D     Shows status for all disk device groups

8. **Ensure that the system disk, applications, and all data are backed up.**

9. **If your cluster uses dual-string mediators for Solstice DiskSuite or Solaris Volume Manager software, unconfigure your mediators.**

   See "Configuring Dual-String Mediators" on page 172 for more information.

   a. **Run the following command to verify that no mediator data problems exist.**

      ```
      # medstat -s setname
      ```

      -s setname     Specifies the disk set name

      If the value in the Status field is Bad, repair the affected mediator host. Follow the procedure "How to Fix Bad Mediator Data" on page 174.

   b. **List all mediators.**

      Save this information for when you restore the mediators during the procedure "How to Finish a Rolling Upgrade to Sun Cluster 3.1 8/05 Software" on page 237.

   c. **For a disk set that uses mediators, take ownership of the disk set if no node already has ownership.**

      ```
      # scswitch -z -D setname -h node
      ```

      -z           Changes mastery

      -D           Specifies the name of the disk set

      -h *node*    Specifies the name of the node to become primary of the disk set

   d. **Unconfigure all mediators for the disk set.**

      ```
      # metaset -s setname -d -m mediator-host-list
      ```

      | | |
      |---|---|
      | -s setname | Specifies the disk-set name |
      | -d | Deletes from the disk set |
      | -m mediator-host-list | Specifies the name of the node to remove as a mediator host for the disk set |

      See the mediator(7D) man page for further information about mediator-specific options to the metaset command.

   e. **Repeat these steps for each remaining disk set that uses mediators.**

10. **Shut down the node that you want to upgrade and boot it into noncluster mode.**

   ■ **On SPARC based systems, perform the following commands:**

      ```
      # shutdown -y -g0
      ok boot -x
      ```

- **On x86 based systems, perform the following commands:**

```
# shutdown -y -g0
...
                     <<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci-ide@7,1/ata@1/cmdk@0,0:b
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
```

The other nodes of the cluster continue to function as active cluster members.

**Next Steps**   To upgrade the Solaris software to a Maintenance Update release, go to "How to Perform a Rolling Upgrade of a Solaris Maintenance Update" on page 225.

---

**Note –** The cluster must already run on, or be upgraded to, at least the minimum required level of the Solaris OS to support Sun Cluster 3.1 8/05 software. See the *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for information about supported releases of the Solaris OS.

---

If you do not intend to upgrade the Solaris OS, go to "How to Upgrade Dependency Software Before a Rolling Upgrade" on page 226.

## ▼ How to Perform a Rolling Upgrade of a Solaris Maintenance Update

Perform this procedure to upgrade the Solaris OS to a supported Maintenance Update release.

---

**Note –** To upgrade a cluster from Solaris 8 to Solaris 9 software or from Solaris 9 to Solaris 10 10/05 software or compatible, with or without upgrading Sun Cluster software as well, you must instead perform a nonrolling upgrade. Go to "Performing a Nonrolling Upgrade" on page 195.

---

**Before You Begin**   Ensure that all steps in "How to Prepare a Cluster Node for a Rolling Upgrade" on page 221 are completed.

**Steps**  **1. Temporarily comment out all entries for globally mounted file systems in the node's `/etc/vfstab` file.**

Perform this step to prevent the Solaris upgrade from attempting to mount the global devices.

**2. Follow the instructions in the Solaris maintenance update installation guide to install the Maintenance Update release.**

---
**Note –** Do not reboot the node when prompted to reboot at the end of installation processing.

---

**3. Uncomment all entries in the `/a/etc/vfstab` file for globally mounted file systems that you commented out in Step 1.**

**4. Install any required Solaris software patches and hardware-related patches, and download any needed firmware that is contained in the hardware patches.**

---
**Note –** Do not reboot the node until Step 5.

---

**5. Reboot the node into noncluster mode.**

Include the double dashes (--) in the following command:

```
# reboot -- -x
```

**Next Steps**   Upgrade dependency software. Go to "How to Upgrade Dependency Software Before a Rolling Upgrade" on page 226.

## ▼ How to Upgrade Dependency Software Before a Rolling Upgrade

Perform this procedure on each cluster node to install or upgrade software on which Sun Cluster 3.1 8/05 software has a dependency. The cluster remains in production during this procedure. If you are running SunPlex Manager, status on a node will not be reported during the period that the node's security file agent is stopped. Status reporting resumes when the security file agent is restarted, after the common agent container software is upgraded.

**Before You Begin**   Perform the following tasks:

- Ensure that all steps in "How to Prepare a Cluster Node for a Rolling Upgrade" on page 221 are completed.
- If you upgraded the Solaris OS to a Maintenance Update release, ensure that all steps in "How to Perform a Rolling Upgrade of a Solaris Maintenance Update" on page 225 are completed.
- Ensure that you have installed all required Solaris software patches and hardware-related patches.
- If the cluster runs Solstice DiskSuite software (Solaris 8), ensure that you have installed all required Solstice DiskSuite software patches.

**Steps**  1.  **Become superuser on the cluster node.**

2.  **For the Solaris 8 and Solaris 9 OS, ensure that the Apache Tomcat package is at the required patch level, if the package is installed.**

   a.  **Determine whether the `SUNWtcatu` package is installed.**

   ```
   # pkginfo SUNWtcatu
   SUNWtcatu        Tomcat Servlet/JSP Container
   ```

   b.  **If the Apache Tomcat package is installed, determine whether the required patch for the platform is installed.**

   - SPARC based platforms require at least 114016-01
   - x86 based platforms require at least 114017-01

   ```
   # patchadd -p | grep 114016
   Patch: 114016-01 Obsoletes: Requires: Incompatibles: Packages: SUNWtcatu
   ```

   c.  **If the required patch is not installed, remove the Apache Tomcat package.**

   ```
   # pkgrm SUNWtcatu
   ```

3.  **Insert the Sun Cluster 1 of 2 CD-ROM.**

4.  **Change to the `/cdrom/cdrom0/Solaris_arch/Product/shared_components/Packages/` directory, where *arch* is `sparc` or `x86` .**

   ```
   # cd Solaris_arch/Product/shared_components/Packages/
   ```

5.  **Ensure that at least version 4.3.1 of the Explorer packages is installed.**

   These packages are required by Sun Cluster software for use by the sccheck utility.

   a.  **Determine whether the Explorer packages are installed and, if so, what version.**

   ```
   # pkginfo -l SUNWexplo | grep SUNW_PRODVERS
   SUNW_PRODVERS=4.3.1
   ```

b. **If a version earlier than 4.3.1 is installed, remove the existing Explorer packages.**

```
# pkgrm SUNWexplo SUNWexplu SUNWexplj
```

c. **If you removed Explorer packages or none were installed, install the latest Explorer packages from the Sun Cluster 1 of 2 CD-ROM.**

- **For the Solaris 8 or Solaris 9 OS, use the following command:**

```
# pkgadd -d . SUNWexpl*
```

- **For the Solaris 10 OS, use the following command:**

```
# pkgadd -G -d . SUNWexpl*
```

The `-G` option adds packages to the current zone only. You must add these packages only to the global zone. Therefore, this option also specifies that the packages are *not* propagated to any existing non-global zone or to any non-global zone that is created later.

6. **Ensure that at least version 5.1,REV=34 of the Java Dynamic Management Kit (JDMK) packages is installed.**

a. **Determine whether JDMK packages are installed and, if so, what version.**

```
# pkginfo -l SUNWjdmk-runtime | grep VERSION
VERSION=5.1,REV=34
```

b. **If a version earlier than 5.1,REV=34 is installed, remove the existing JDMK packages.**

```
# pkgrm SUNWjdmk-runtime SUNWjdmk-runtime-jmx
```

c. **If you removed JDMK packages or none were installed, install the latest JDMK packages from the Sun Cluster 1 of 2 CD-ROM.**

- **For the Solaris 8 or Solaris 9 OS, use the following command:**

```
# pkgadd -d . SUNWjdmk*
```

- **For the Solaris 10 OS, use the following command:**

```
# pkgadd -G -d . SUNWjdmk*
```

7. **Change to the `Solaris_arch/Product/shared_components/Solaris_ver/Packages/` directory, where *arch* is `sparc` or `x86` and where *ver* is 8 for Solaris 8, 9 for Solaris 9, or 10 for Solaris 10.**

```
# cd ../Solaris_ver/Packages
```

8. **Ensure that at least version 4.5.0 of the Netscape Portable Runtime (NSPR) packages is installed.**

a.  **Determine whether NSPR packages are installed and, if so, what version.**

    ```
    # cat /var/sadm/pkg/SUNWpr/pkginfo | grep SUNW_PRODVERS
    SUNW_PRODVERS=4.5.0
    ```

b.  **If a version earlier than 4.5.0 is installed, remove the existing NSPR packages.**

    ```
    # pkgrm packages
    ```
    The following table lists the applicable packages for each hardware platform.

    ---

    **Note –** Install packages in the order in which they are listed in the following table.

    ---

    | Hardware Platform | NSPR Package Names |
    | --- | --- |
    | SPARC | SUNWpr SUNWprx |
    | x86 | SUNWpr |

c.  **If you removed NSPR packages or none were installed, install the latest NSPR packages.**

    ■ **For the Solaris 8 or Solaris 9 OS, use the following command:**

    ```
    # pkgadd -d . packages
    ```

    ■ **For the Solaris 10 OS, use the following command:**

    ```
    # pkgadd -G -d . packages
    ```

9.  **Ensure that at least version 3.9.4 of the Network Security Services (NSS) packages is installed.**

    a.  **Determine whether NSS packages are installed and, if so, what version.**

        ```
        # cat /var/sadm/pkg/SUNWtls/pkginfo | grep SUNW_PRODVERS
        SUNW_PRODVERS=3.9.4
        ```

    b.  **If a version earlier than 3.9.4 is installed, remove the existing NSS packages.**

        ```
        # pkgrm packages
        ```
        The following table lists the applicable packages for each hardware platform.

        ---

        **Note –** Install packages in the order in which they are listed in the following table.

        ---

| Hardware Platform | NSS Package Names |
| --- | --- |
| SPARC | `SUNWtls SUNWtlsu SUNWtlsx` |
| x86 | `SUNWtls SUNWtlsu` |

   c. **If you removed NSS packages or none were installed, install the latest NSS packages from the Sun Cluster 1 of 2 CD-ROM.**

     ■ **For the Solaris 8 or Solaris 9 OS, use the following command:**

       `# `**`pkgadd -d .`** *packages*

     ■ **For the Solaris 10 OS, use the following command:**

       `# `**`pkgadd -G -d .`** *packages*

10. **Change back to the `Solaris_`*arch*`/Product/shared_components/Packages/` directory.**

    `# `**`cd ../../Packages`**

11. **Ensure that at least version 1.0,REV=25 of the common agent container packages is installed.**

   a. **Determine whether the common agent container packages are installed and, if so, what version.**

    `# `**`pkginfo -l SUNWcacao | grep VERSION`**
    `VERSION=1.0,REV=25`

   b. **If a version earlier than 1.0,REV=25 is installed, stop the security file agent for the common agent container on each cluster node.**

    `# `**`/opt/SUNWcacao/bin/cacaoadm stop`**

   c. **If a version earlier than 1.0,REV=25 is installed, remove the existing common agent container packages.**

    `# `**`pkgrm SUNWcacao SUNWcacaocfg`**

   d. **If you removed the common agent container packages or none were installed, install the latest common agent container packages from the Sun Cluster 1 of 2 CD-ROM.**

     ■ **For the Solaris 8 or Solaris 9 OS, use the following command:**

       `# `**`pkgadd -d . SUNWcacao*`**

     ■ **For the Solaris 10 OS, use the following command:**

       `# `**`pkgadd -G -d . SUNWcacao*`**

**12. Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

```
# eject cdrom
```

**13. Insert the Sun Cluster 2 of 2 CD-ROM.**

**14. Install or upgrade Sun Java Web Console packages.**

    **a. Change to the `Solaris_`*arch*`/Product/sunwebconsole/` directory, where *arch* is `sparc` or `x86`.**

    **b. Install the Sun Java Web Console packages.**

       
```
# ./setup
```

       The setup command installs or upgrades all packages to support Sun Java Web Console.

**15. Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

```
# eject cdrom
```

**16. Ensure that the `/usr/java/` directory is a symbolic link to the minimum or latest version of Java software.**

Sun Cluster software requires at least version 1.4.2_03 of Java software.

    **a. Determine what directory the `/usr/java/` directory is symbolically linked to.**

```
# ls -l /usr/java
lrwxrwxrwx   1 root   other    9 Apr 19 14:05 /usr/java -> /usr/j2se/
```

    **b. Determine what version or versions of Java software are installed.**

       The following are examples of commands that you can use to display the version of their related releases of Java software.

```
# /usr/j2se/bin/java -version
# /usr/java1.2/bin/java -version
# /usr/jdk/jdk1.5.0_01/bin/java -version
```

    **c. If the `/usr/java/` directory is not symbolically linked to a supported version of Java software, recreate the symbolic link to link to a supported version of Java software.**

       The following example shows the creation of a symbolic link to the /usr/j2se/ directory, which contains Java 1.4.2_03 software.

```
# rm /usr/java
# ln -s /usr/j2se /usr/java
```

**Next Steps**    Upgrade Sun Cluster software. Go to "How to Perform a Rolling Upgrade of Sun Cluster 3.1 8/05 Software" on page 232

## ▼ How to Perform a Rolling Upgrade of Sun Cluster 3.1 8/05 Software

Perform this procedure to upgrade a node to Sun Cluster 3.1 8/05 software while the remaining cluster nodes are in cluster mode.

---

**Note –** Until all nodes of the cluster are upgraded and the upgrade is committed, new features that are introduced by the new release might not be available.

---

**Before You Begin** Ensure that dependency software is installed or upgraded. See "How to Upgrade Dependency Software Before a Rolling Upgrade" on page 226.

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

2. **Insert the Sun Cluster 2 of 2 CD-ROM in the CD-ROM drive on the node.**

   If the volume management daemon `vold`(1M) is running and is configured to manage CD-ROM devices, the daemon automatically mounts the CD-ROM on the `/cdrom/cdrom0/` directory.

3. **Change to the `Solaris_arch/Product/sun_cluster/Solaris_ver/Tools/` directory, where *arch* is `sparc` or `x86` and where *ver* is 8 for Solaris 8, 9 for Solaris 9, or 10 for Solaris 10 .**

   ```
   # cd /cdrom/cdrom0/Solaris_arch/Product/sun_cluster/Solaris_ver/Tools
   ```

4. **Start the `scinstall` utility.**

   ```
   # ./scinstall
   ```

---

**Note –** Do not use the `/usr/cluster/bin/scinstall` command that is already installed on the node. You must use the `scinstall` command on the Sun Cluster 2 of 2 CD-ROM.

---

5. **From the Main Menu, choose the menu item, Upgrade this cluster node.**

   ```
       *** Main Menu ***

     Please select from one of the following (*) options:

         * 1) Install a cluster or cluster node
           2) Configure a cluster to be JumpStarted from this install server
         * 3) Add support for new data services to this cluster node
         * 4) Upgrade this cluster node
         * 5) Print release information for this cluster node
   ```

```
* ?) Help with menu options
* q) Quit

Option:  4
```

6. **From the Upgrade Menu, choose the menu item, Upgrade Sun Cluster framework on this node.**

7. **Follow the menu prompts to upgrade the cluster framework.**

   During Sun Cluster upgrade, scinstall might make one or more of the following configuration changes:

   ■ Convert NAFO groups to IP Network Multipathing groups but keep the original NAFO-group name.

   See one of the following manuals for information about test addresses for IP Network Multipathing:

   ■ *IP Network Multipathing Administration Guide* (Solaris 8)

   ■ "Configuring Test Addresses" in "Administering Multipathing Groups With Multiple Physical Interfaces" in *System Administration Guide: IP Services* (Solaris 9)

   ■ "Test Addresses" in *System Administration Guide: IP Services* (Solaris 10)

   See the scinstall(1M) man page for more information about the conversion of NAFO groups to IP Network Multipathing during Sun Cluster software upgrade.

   ■ Rename the ntp.conf file to ntp.conf.cluster, if ntp.conf.cluster does not already exist on the node.

   ■ Set the local-mac-address? variable to true, if the variable is not already set to that value.

   Upgrade processing is finished when the system displays the message Completed Sun Cluster framework upgrade and prompts you to press Enter to continue.

8. **Press Enter.**

   The Upgrade Menu is displayed.

9. **(Optional) Upgrade Java Enterprise System data services from the Sun Cluster 2 of 2 CD-ROM.**

   a. **From the Upgrade Menu of the scinstall utility, choose the menu item, Upgrade Sun Cluster data service agents on this node.**

   b. **Follow the menu prompts to upgrade Sun Cluster data service agents that are installed on the node.**

   You can choose from the list of data services that are available to upgrade or choose to upgrade all installed data services.

   Upgrade processing is finished when the system displays the message Completed upgrade of Sun Cluster data services agents and

prompts you to press Enter to continue.

**c. Press Enter.**

The Upgrade Menu is displayed.

**10. Quit the `scinstall` utility.**

**11. Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

# `eject cdrom`

**12. Upgrade Sun Cluster data services from the Sun Cluster Agents CD.**

- If you are using the Sun Cluster HA for NFS data service and you upgrade to the Solaris 10 OS, you must upgrade the data service and migrate the resource type to the new version. See "Upgrading the SUNW.nfs Resource Type" in *Sun Cluster Data Service for NFS Guide for Solaris OS* for more information.

- If you are using the Sun Cluster HA for Oracle 3.0 64-bit for Solaris 9 data service, you must upgrade to the Sun Cluster 3.1 8/05 version.

- The upgrade of any other data services to the Sun Cluster 3.1 8/05 version is optional. You can continue to use any other Sun Cluster 3.x data services after you upgrade the cluster to Sun Cluster 3.1 8/05 software.

**a. Insert the Sun Cluster Agents CD in the CD-ROM drive on the node.**

**b. Start the `scinstall` utility.**

For data-service upgrades, you can use the `/usr/cluster/bin/scinstall` command that is already installed on the node.

# `scinstall`

**c. From the Main Menu, choose the menu item, Upgrade this cluster node.**

**d. From the Upgrade Menu, choose the menu item, Upgrade Sun Cluster data service agents on this node.**

**e. Follow the menu prompts to upgrade Sun Cluster data service agents that are installed on the node.**

You can choose from the list of data services that are available to upgrade or choose to upgrade all installed data services.

Upgrade processing is finished when the system displays the message `Completed upgrade of Sun Cluster data services agents` and prompts you to press Enter to continue.

**f. Press Enter.**

The Upgrade Menu is displayed.

**g. Quit the `scinstall` utility.**

**h. Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

```
# eject cdrom
```

**13. As needed, manually upgrade any custom data services that are not supplied on the product media.**

**14. Verify that each data-service update is installed successfully.**

View the upgrade log file that is referenced at the end of the upgrade output messages.

**15. Install any Sun Cluster 3.1 8/05 software patches, if you did not already install them by using the `scinstall` command.**

**16. Install any Sun Cluster 3.1 8/05 data-service software patches.**

See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

**17. Upgrade software applications that are installed on the cluster.**

Ensure that application levels are compatible with the current versions of Sun Cluster and Solaris software. See your application documentation for installation instructions. In addition, follow these guidelines to upgrade applications in a Sun Cluster 3.1 8/05 configuration:

- If the applications are stored on shared disks, you must master the relevant disk groups and manually mount the relevant file systems before you upgrade the application.

- If you are instructed to reboot a node during the upgrade process, always add the -x option to the command.

  The -x option ensures that the node reboots into noncluster mode. For example, either of the following two commands boot a node into single-user noncluster mode:

  - On SPARC based systems, perform the following commands:

    ```
    # reboot -- -xs
    ok boot -xs
    ```

  - On x86 based systems, perform the following commands:

    ```
    # reboot -- -xs
    ...
                        <<< Current Boot Parameters >>>
    Boot path: /pci@0,0/pci-ide@7,1/ata@1/cmdk@0,0:b
    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 -xs
```

---

**Note –** Do not upgrade an application if the newer version of the application cannot coexist in the cluster with the older version of the application.

---

18. **Reboot the node into the cluster.**

    ```
    # reboot
    ```

19. **Run the following command on the upgraded node to verify that Sun Cluster 3.1 8/05 software was installed successfully.**

    ```
    # scinstall -pv
    ```

    The first line of output states which version of Sun Cluster software the node is running. This version should match the version you just upgraded to.

20. **From any node, verify the status of the cluster configuration.**

    ```
    % scstat
    % scrgadm -pv[v]
    ```

    Output should be the same as for Step 5 in "How to Prepare a Cluster Node for a Rolling Upgrade" on page 221.

21. **If you have another node to upgrade, return to "How to Prepare a Cluster Node for a Rolling Upgrade" on page 221 and repeat all upgrade procedures on the next node to upgrade.**

**Example 5–2**    Rolling Upgrade From Sun Cluster 3.1 to Sun Cluster 3.1 8/05 Software

The following example shows the process of a rolling upgrade of a cluster node from Sun Cluster 3.1 to Sun Cluster 3.1 8/05 software on the Solaris 8 OS. The example includes the upgrade of all installed data services that have new versions on the Sun Cluster Agents CD. The cluster node name is phys-schost-1.

```
(Upgrade framework software from the Sun Cluster 2 of 2 CD-ROM)
phys-schost-1# cd /cdrom/cdrom0/Solaris_sparc/Product/sun_cluster/Solaris_8/Tools/
phys-schost-1# ./scinstall

(Upgrade data services from the Sun Cluster Agents CD)
phys-schost-1# scinstall

(Reboot the node into the cluster)
phys-schost-1# reboot

(Verify that software upgrade succeeded)
# scinstall -pv

(Verify cluster status)
```

```
# scstat
# scrgadm -pv
```

## ▼ How to Finish a Rolling Upgrade to Sun Cluster 3.1 8/05 Software

**Before You Begin**  Ensure that all upgrade procedures are completed for all cluster nodes that you are upgrading.

**Steps**  1. **From one node, check the upgrade status of the cluster.**

```
# scversions
```

2. **From the following table, perform the action that is listed for the output message from Step 1.**

| Output Message | Action |
|---|---|
| `Upgrade commit is needed.` | Proceed to Step 4. |
| `Upgrade commit is NOT needed. All versions match.` | Skip to Step 6. |
| `Upgrade commit cannot be performed until all cluster nodes are upgraded. Please run scinstall(1m) on cluster nodes to identify older versions.` | Return to "How to Perform a Rolling Upgrade of Sun Cluster 3.1 8/05 Software" on page 232 to upgrade the remaining cluster nodes. |
| `Check upgrade cannot be performed until all cluster nodes are upgraded. Please run scinstall(1m) on cluster nodes to identify older versions.` | Return to "How to Perform a Rolling Upgrade of Sun Cluster 3.1 8/05 Software" on page 232 to upgrade the remaining cluster nodes. |

3. **After all nodes have rejoined the cluster, from one node commit the cluster to the upgrade.**

```
# scversions -c
```

Committing the upgrade enables the cluster to utilize all features in the newer software. New features are available only after you perform the upgrade commitment.

4. **From one node, verify that the cluster upgrade commitment has succeeded.**

```
# scversions
Upgrade commit is NOT needed. All versions match.
```

5. **Copy the security files for the common agent container to all cluster nodes.**

   This step ensures that security files for the common agent container are identical on all cluster nodes and that the copied files retain the correct file permissions.

   a. **On each node, stop the Sun Java Web Console agent.**

      ```
      # /usr/sbin/smcwebserver stop
      ```

   b. **On each node, stop the security file agent.**

      ```
      # /opt/SUNWcacao/bin/cacaoadm stop
      ```

   c. **On one node, change to the `/etc/opt/SUNWcacao/` directory.**

      ```
      phys-schost-1# cd /etc/opt/SUNWcacao/
      ```

   d. **Create a tar file of the `/etc/opt/SUNWcacao/security/` directory.**

      ```
      phys-schost-1# tar cf /tmp/SECURITY.tar security
      ```

   e. **Copy the `/tmp/SECURITY.tar` file to each of the other cluster nodes.**

   f. **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
      ```

   g. **Delete the `/tmp/SECURITY.tar` file from each node in the cluster.**

      You must delete each copy of the tar file to avoid security risks.

      ```
      phys-schost-1# rm /tmp/SECURITY.tar
      phys-schost-2# rm /tmp/SECURITY.tar
      ```

   h. **On each node, start the security file agent.**

      ```
      phys-schost-1# /opt/SUNWcacao/bin/cacaoadm start
      phys-schost-2# /opt/SUNWcacao/bin/cacaoadm start
      ```

   i. **On each node, start the Sun Java Web Console agent.**

      ```
      phys-schost-1# /usr/sbin/smcwebserver start
      phys-schost-2# /usr/sbin/smcwebserver start
      ```

6. **If your configuration uses dual-string mediators for Solstice DiskSuite or Solaris Volume Manager software, restore the mediator configurations.**

   a. **Determine which node has ownership of a disk set to which you are adding the mediator hosts.**

      ```
      # metaset -s setname
      ```

      -s setname      Specifies the disk-set name

b. **If no node has ownership, take ownership of the disk set.**

   # **scswitch -z -D** *setname* **-h** *node*

   -z          Changes mastery

   -D          Specifies the name of the disk set

   -h *node*     Specifies the name of the node to become primary of the disk set

c. **Re-create the mediators.**

   # **metaset -s** *setname* **-a -m** *mediator-host-list*

   -a                         Adds to the disk set

   -m mediator-host-list      Specifies the names of the nodes to add as
                              mediator hosts for the disk set

d. **Repeat Step a through Step c for each disk set in the cluster that uses
   mediators.**

7. **If you upgraded any data services that are not supplied on the product media,
   register the new resource types for those data services.**

   Follow the documentation that accompanies the data services.

8. **(Optional) Switch each resource group and device group back to its original
   node.**

   # **scswitch -z -g** *resource-group* **-h** *node*
   # **scswitch -z -D** *disk-device-group* **-h** *node*

   -z                         Performs the switch

   -g *resource-group*          Specifies the resource group to switch

   -h *node*                    Specifies the name of the node to switch to

   -D *disk-device-group*       Specifies the device group to switch

9. **Restart any applications.**

   Follow the instructions that are provided in your vendor documentation.

10. **Migrate resources to new resource type versions.**

---

**Note –** If you upgrade to the Sun Cluster HA for NFS data service for the Solaris 10
OS, you must migrate to the new resource type version. See "Upgrading the
SUNW.nfs Resource Type" in *Sun Cluster Data Service for NFS Guide for Solaris OS* for
more information.

For all other data services, this step is optional.

---

See "Upgrading a Resource Type" in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*, which contains procedures which use the command line. Alternatively, you can perform the same tasks by using the Resource Group menu of the `scsetup` utility. The process involves performing the following tasks:

- Registration of the new resource type
- Migration of the eligible resource to the new version of its resource type
- Modification of the extension properties of the resource type as specified in the manual for the related data service

**Next Steps**    If you have a SPARC based system and use Sun Management Center to monitor the cluster, go to "SPARC: How to Upgrade Sun Cluster Module Software for Sun Management Center" on page 242.

Otherwise, the cluster upgrade is complete.

# Recovering From Storage Configuration Changes During Upgrade

This section provides the following repair procedures to follow if changes were inadvertently made to the storage configuration during upgrade:

- "How to Handle Storage Reconfiguration During an Upgrade" on page 240
- "How to Resolve Mistaken Storage Changes During an Upgrade" on page 241

## ▼ How to Handle Storage Reconfiguration During an Upgrade

Any changes to the storage topology, including running Sun Cluster commands, should be completed before you upgrade the cluster to Solaris 9 software. If, however, changes were made to the storage topology during the upgrade, perform the following procedure. This procedure ensures that the new storage configuration is correct and that existing storage that was not reconfigured is not mistakenly altered.

**Before You Begin**    Ensure that the storage topology is correct. Check whether the devices that were flagged as possibly being replaced map to devices that actually were replaced. If the devices were not replaced, check for and correct possible accidental configuration changes, such as incorrect cabling.

**Steps**    1. **Become superuser on a node that is attached to the unverified device.**

2. **Manually update the unverified device.**

   ```
   # scdidadm -R device
   ```

   -R device     Performs repair procedures on the specified *device*

   See the scdidadm(1M) man page for more information.

3. **Update the DID driver.**

   ```
   # scdidadm -ui
   # scdidadm -r
   ```

   -u     Loads the device-ID configuration table into the kernel

   -i     Initializes the DID driver

   -r     Reconfigures the database

4. **Repeat Step 2 through Step 3 on all other nodes that are attached to the unverified device.**

**Next Steps**    Return to the remaining upgrade tasks.

- For a nonrolling upgrade, go to Step 3 in "How to Perform a Nonrolling Upgrade of Sun Cluster 3.1 8/05 Software" on page 210.
- For a rolling upgrade, go to Step 4 in "How to Perform a Rolling Upgrade of Sun Cluster 3.1 8/05 Software" on page 232.

## ▼ How to Resolve Mistaken Storage Changes During an Upgrade

If accidental changes are made to the storage cabling during the upgrade, perform the following procedure to return the storage configuration to the correct state.

---

**Note –** This procedure assumes that no physical storage was actually changed. If physical or logical storage devices were changed or replaced, instead follow the procedures in "How to Handle Storage Reconfiguration During an Upgrade" on page 240.

---

**Before You Begin**  Return the storage topology to its original configuration. Check the configuration of the devices that were flagged as possibly being replaced, including the cabling.

**Steps**  **1. As superuser, update the DID driver on each node of the cluster.**

```
# scdidadm -ui
# scdidadm -r
```

-u    Loads the device–ID configuration table into the kernel

-i    Initializes the DID driver

-r    Reconfigures the database

See the scdidadm(1M) man page for more information.

**2. If the `scdidadm` command returned any error messages in Step 1, make further modifications as needed to correct the storage configuration, then repeat Step 1.**

**Next Steps**  Return to the remaining upgrade tasks.

- For a nonrolling upgrade, go to Step 3 in "How to Perform a Nonrolling Upgrade of Sun Cluster 3.1 8/05 Software" on page 210.
- For a rolling upgrade, go to Step 4 in "How to Perform a Rolling Upgrade of Sun Cluster 3.1 8/05 Software" on page 232.

# SPARC: Upgrading Sun Management Center Software

This section provides the following procedures to upgrade the Sun Cluster module for Sun Management Center:

- "SPARC: How to Upgrade Sun Cluster Module Software for Sun Management Center" on page 242
- "SPARC: How to Upgrade Sun Management Center Software" on page 244

## ▼ SPARC: How to Upgrade Sun Cluster Module Software for Sun Management Center

Perform the following steps to upgrade Sun Cluster module software on the Sun Management Center server machine, help-server machine, and console machine.

---

**Note –** If you intend to upgrade the Sun Management Center software itself, do not perform this procedure. Instead, go to "SPARC: How to Upgrade Sun Management Center Software" on page 244 to upgrade the Sun Management Center software and the Sun Cluster module.

---

**Before You Begin**    Have available the Sun Cluster 2 of 2 CD-ROM for the SPARC platform or the path to the CD-ROM image.

**Steps**    1. **As superuser, remove the existing Sun Cluster module packages from each machine.**

Use the pkgrm(1M) command to remove all Sun Cluster module packages from all locations that are listed in the following table.

| Location | Module Package to Remove |
|---|---|
| Sun Management Center console machine | SUNWscscn |
| Sun Management Center server machine | SUNWscssv |
| Sun Management Center 3.0 help-server machine or Sun Management Center 3.5 server machine | SUNWscshl |

# **pkgrm** *module-package*

---

**Note –** Sun Cluster module software on the cluster nodes was already upgraded during the cluster-framework upgrade.

---

2. **As superuser, reinstall Sun Cluster module packages on each machine.**

   a. **Insert the Sun Cluster 2 of 2 CD-ROM for the SPARC platform into the CD-ROM drive of the machine.**

   b. **Change to the Solaris_sparc/Product/sun_cluster/Solaris_*ver*/Packages/ directory, where *ver* is 8 for Solaris 8, 9 for Solaris 9, or 10 for Solaris 10.**

      # **cd Solaris_sparc/Product/sun_cluster/Solaris_*ver*/Packages/**

   c. **Install the appropriate module packages, as listed in the following table.**

| Location | Module Package to Install |
|---|---|
| Sun Management Center console machine | SUNWscshl |
| Sun Management Center server machine | SUNWscssv |
| Sun Management Center 3.0 help-server machine or Sun Management Center 3.5 server machine | SUNWscshl |

Note that you install the help-server package SUNWscshl on both the console machine and the Sun Management Center 3.0 help-server machine or the Sun Management Center 3.5 server machine. Also, you do not upgrade to a new SUNWscscn package on the console machine.

```
# pkgadd -d . module-package
```

   **d. Change to a directory that does *not* reside on the CD-ROM and eject the CD-ROM.**

```
# eject cdrom
```

## ▼ SPARC: How to Upgrade Sun Management Center Software

Perform the following steps to upgrade from Sun Management Center 2.1.1 to either Sun Management Center 3.0 software or Sun Management Center 3.5 software.

**Before You Begin**   Have available the following items:

- Sun Cluster 2 of 2 CD-ROM for the SPARC platform and, if applicable, for the x86 platform, or the paths to the CD-ROM images. You use the CD-ROM to reinstall the Sun Cluster 3.1 8/05 version of the Sun Cluster module packages after you upgrade Sun Management Center software.

   **Note –** The agent packages to install on the cluster nodes are available for both SPARC based systems and x86 based systems. The packages for the console, server, and help-server machines are available for SPARC based systems only.

- Sun Management Center documentation.
- Sun Management Center patches and Sun Cluster module patches, if any.

   See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 8/05 Release Notes for Solaris OS* for the location of patches and installation instructions.

**Steps**   **1. Stop any Sun Management Center processes.**

a. **If the Sun Management Center console is running, exit the console.**

In the console window, choose File⇒Exit.

b. **On each Sun Management Center agent machine (cluster node), stop the Sun Management Center agent process.**

```
# /opt/SUNWsymon/sbin/es-stop -a
```

c. **On the Sun Management Center server machine, stop the Sun Management Center server process.**

```
# /opt/SUNWsymon/sbin/es-stop -S
```

2. **As superuser, remove Sun Cluster–module packages.**

Use the pkgrm(1M) command to remove all Sun Cluster module packages from all locations that are listed in the following table.

| Location | Module Package to Remove |
|---|---|
| Each cluster node | SUNWscsam, SUNWscsal |
| Sun Management Center console machine | SUNWscscn |
| Sun Management Center server machine | SUNWscssv |
| Sun Management Center 3.0 help-server machine or Sun Management Center 3.5 server machine | SUNWscshl |

```
# pkgrm module-package
```

If you do not remove the listed packages, the Sun Management Center software upgrade might fail because of package dependency problems. You reinstall these packages in Step 4, after you upgrade Sun Management Center software.

3. **Upgrade the Sun Management Center software.**

Follow the upgrade procedures in your Sun Management Center documentation.

4. **As superuser, reinstall Sun Cluster module packages from the to the locations that are listed in the following table.**

| Location | Module Package to Install |
|---|---|
| Each cluster node | SUNWscsam, SUNWscsal |
| Sun Management Center server machine | SUNWscssv |
| Sun Management Center console machine | SUNWscshl |

| Location | Module Package to Install |
|---|---|
| Sun Management Center 3.0 help-server machine or Sun Management Center 3.5 server machine | SUNWscshl |

You install the help-server package SUNWscshl on both the console machine and the Sun Management Center 3.0 help-server machine or the Sun Management Center 3.5 server machine.

a. **Insert the Sun Cluster 2 of 2 CD-ROM for the appropriate platform in the CD-ROM drive of the machine.**

b. **Change to the Solaris_*arch*/Product/sun_cluster/Solaris_*ver*/Packages/ directory, where *arch* is sparc or x86, and *ver* is 8 for Solaris 8, 9 for Solaris 9, or 10 for Solaris 10.**

   # **cd /cdrom/cdrom0/Solaris_*arch*/Product/sun_cluster/Solaris_*ver*/Packages/**

   ---
   **Note –** The agent packages to install on the cluster nodes are available for both SPARC based systems and x86 based systems. The packages for the console, server, and help-server machines are available for SPARC based systems only.

   ---

c. **Install the appropriate module package on the machine.**

   ■ **For cluster nodes that run the Solaris 10 OS, use the following command:**

      # **pkgadd -G -d .** *module-package*

      The -G option adds packages to the current zone only. You must add these packages only to the global zone. Therefore, this option also specifies that the packages are *not* propagated to any existing non-global zone or to any non-global zone that is created later.

   ■ **For cluster nodes that run the Solaris 8 or Solaris 9 OS and for the console, server, and help-server machines, use the following command:**

      # **pkgadd -d .** *module-package*

5. **Apply any Sun Management Center patches and any Sun Cluster module patches to each node of the cluster.**

6. **Restart Sun Management Center agent, server, and console processes.**
   Follow procedures in "SPARC: How to Start Sun Management Center" on page 132.

7. **Load the Sun Cluster module.**
   Follow procedures in "SPARC: How to Load the Sun Cluster Module" on page 134.

If the Sun Cluster module was previously loaded, unload the module and then reload it to clear all cached alarm definitions on the server. To unload the module, choose Unload Module from the Module menu on the console's Details window.

# Configuring Data Replication With Sun StorEdge Availability Suite Software

This chapter provides guidelines for configuring data replication between clusters by using Sun StorEdge Availability Suite 3.1 or 3.2 software.

This chapter also contains an example of how data replication was configured for an NFS application by using Sun StorEdge Availability Suite software. This example uses a specific cluster configuration and provides detailed information about how individual tasks can be performed. It does not include all of the steps that are required by other applications or other cluster configurations.

This chapter contains the following sections:

# Introduction to Data Replication

This section introduces disaster tolerance and describes the data replication methods that Sun StorEdge Availability Suite software uses.

# What Is Disaster Tolerance?

Disaster tolerance is the ability of a system to restore an application on an alternate cluster when the primary cluster fails. Disaster tolerance is based on *data replication* and *failover*.

Data replication is the copying of data from a primary cluster to a backup or secondary cluster. Through data replication, the secondary cluster has an up-to-date copy of the data on the primary cluster. The secondary cluster can be located far away from the primary cluster.

Failover is the automatic relocation of a resource group or device group from a primary cluster to a secondary cluster. If the primary cluster fails, the application and the data are immediately available on the secondary cluster.

# Data Replication Methods Used by Sun StorEdge Availability Suite Software

This section describes the remote mirror replication method and the point-in-time snapshot method used by Sun StorEdge Availability Suite software. This software uses the sndradm(1RPC) and iiadm(1II) commands to replicate data. For more information about these commands, see one of the following manuals:

- Sun StorEdge Availability Suite 3.1 software - *Sun Cluster 3.0 and Sun StorEdge Software Integration Guide*
- Sun StorEdge Availability Suite 3.2 software - *Sun Cluster 3.0/3.1 and Sun StorEdge Availability Suite 3.2 Software Integration Guide*

## Remote Mirror Replication

Remote mirror replication is illustrated in Figure 6–1. Data from the master volume of the primary disk is replicated to the master volume of the secondary disk through a TCP/IP connection. A remote mirror bitmap tracks differences between the master volume on the primary disk and the master volume on the secondary disk.

**Primary disk**                    **Secondary disk**

Updated datablocks are periodically copied to the secondary disk

Master volume

Remote mirror bitmap volume

Differences between the master volume on the primary disk and the master volume on the secondary disk are tracked on the remote mirror bitmap volume
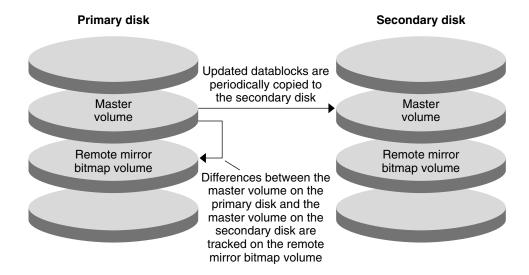
**FIGURE 6–1** Remote Mirror Replication

Remote mirror replication can be performed synchronously in real time, or asynchronously. Each volume set in each cluster can be configured individually, for synchronous replication or asynchronous replication.

- In synchronous data replication, a write operation is not confirmed as complete until the remote volume has been updated.

- In asynchronous data replication, a write operation is confirmed as complete before the remote volume is updated. Asynchronous data replication provides greater flexibility over long distances and low bandwidth.

## Point-in-Time Snapshot

Point-in-time snapshot is illustrated in Figure 6–2. Data from the master volume of each disk is copied to the shadow volume on the same disk. The point-in-time bitmap tracks differences between the master volume and the shadow volume. When data is copied to the shadow volume, the point-in-time bitmap is reset.

**Primary disk**                                      **Secondary disk**

Master volume · Updated datablocks are periodically copied to the shadow volume · Master volume

Shadow volume · Shadow volume

Point-in-time bitmap volume · Differences between the master volume and shadow volume are tracked on the point-in-time bitmap volume · Point-in-time bitmap volume
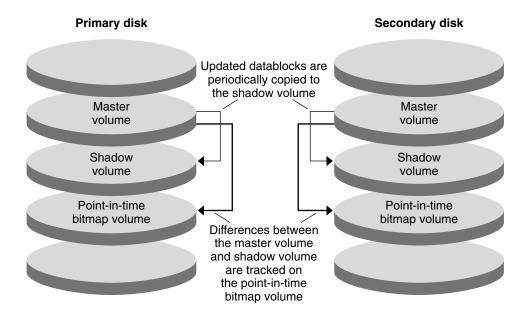
**FIGURE 6–2** Point-in-Time Snapshot

## Replication in the Example Configuration

The following figure illustrates how remote mirror replication and point-in-time snapshot are used in this example configuration.

**Primary disk**                                        **Secondary disk**



| Master volume | Remote mirror replication → | Master volume |
| Shadow volume | Point-in-time replication | Shadow volume |
| Point-in-time bitmap volume | | Point-in-time bitmap volume |
| Remote mirror bitmap volume | | Remote mirror bitmap volume |
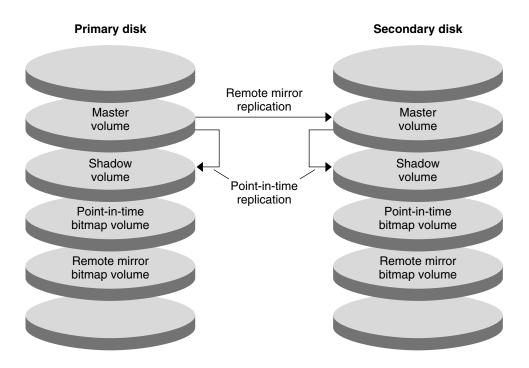
**FIGURE 6–3** Replication in the Example Configuration

# Guidelines for Configuring Data Replication

This section provides guidelines for configuring data replication between clusters. This section also contains tips for configuring replication resource groups and application resource groups. Use these guidelines when you are configuring data replication for your cluster.

This section discusses the following topics:

- "Configuring Replication Resource Groups" on page 254
- "Configuring Application Resource Groups" on page 254
    - "Configuring Resource Groups for a Failover Application" on page 255
    - "Configuring Resource Groups for a Scalable Application" on page 256
- "Guidelines for Managing a Failover or Switchover" on page 257

# Configuring Replication Resource Groups

Replication resource groups colocate the device group under Sun StorEdge Availability Suite software control with the logical hostname resource. A replication resource group must have the following characteristics:

- Be a failover resource group

  A failover resource can run on only one node at a time. When a failover occurs, failover resources take part in the failover.

- Have a logical hostname resource

  The logical hostname must be hosted by the primary cluster. After a failover or switchover, the logical hostname must be hosted by the secondary cluster. The Domain Name System (DNS) is used to associate the logical hostname with a cluster.

- Have an HAStoragePlus resource

  The HAStoragePlus resource enforces the switchover of the device group when the replication resource group is switched over or failed over. Sun Cluster software also enforces the switchover of the replication resource group when the device group is switched over. In this way, the replication resource group and the device group are always colocated, or mastered by the same node.

  The following extension properties must be defined in the HAStoragePlus resource:

  - *GlobalDevicePaths*. This extension property defines the device group to which a volume belongs.

  - *AffinityOn property* = `True`. This extension property causes the device group to switch over or fail over when the replication resource group switches over or fails over. This feature is called an *affinity switchover*.

  For more information about HAStoragePlus, see the `SUNW.HAStoragePlus`(5) man page.

- Be named after the device group with which it is colocated, followed by `-stor-rg`

  For example, *devicegroup*`-stor-rg`.

- Be online on both the primary cluster and the secondary cluster

# Configuring Application Resource Groups

To be highly available, an application must be managed as a resource in an application resource group. An application resource group can be configured for a failover application or a scalable application.

Application resources and application resource groups configured on the primary cluster must also be configured on the secondary cluster. Also, the data accessed by the application resource must be replicated to the secondary cluster.

This section provides guidelines for configuring the following application resource groups:

## Configuring Resource Groups for a Failover Application

In a failover application, an application runs on one node at a time. If that node fails, the application fails over to another node in the same cluster. A resource group for a failover application must have the following characteristics:

- Have an HAStoragePlus resource to enforce the switchover of the device group when the application resource group is switched over or failed over

  The device group is colocated with the replication resource group and the application resource group. Therefore, the switchover of the application resource group enforces the switchover of the device group and replication resource group. The application resource group, the replication resource group, and the device group are mastered by the same node.

  Note, however, that a switchover or failover of the device group or the replication resource group does not cause a switchover or failover of the application resource group.

    - If the application data is globally mounted, the presence of an HAStoragePlus resource in the application resource group is not required but is advised.

    - If the application data is mounted locally, the presence of an HAStoragePlus resource in the application resource group is required.

      Without an HAStoragePlus resource, the switchover or failover of the application resource group would not trigger the switchover or failover of the replication resource group and device group. After a switchover or failover, the application resource group, replication resource group, and device group would not be mastered by the same node.

  For more information about HAStoragePlus, see the `SUNW.HAStoragePlus`(5) man page.

- Must be online on the primary cluster and offline on the secondary cluster

  The application resource group must be brought online on the secondary cluster when the secondary cluster takes over as the primary cluster.

The following figure illustrates the configuration of an application resource group and a replication resource group in a failover application.

**Primary cluster**                    **Secondary cluster**
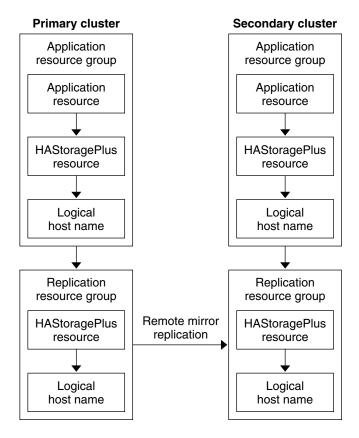
**FIGURE 6–4** Configuration of Resource Groups in a Failover Application

## Configuring Resource Groups for a Scalable Application

In a scalable application, an application runs on several nodes to create a single, logical service. If a node that is running a scalable application fails, failover does not occur. The application continues to run on the other nodes.

When a scalable application is managed as a resource in an application resource group, it is not necessary to colocate the application resource group with the device group. Therefore, it is not necessary to create an HAStoragePlus resource for the application resource group.

A resource group for a scalable application must have the following characteristics:

- Have a dependency on the shared address resource group

   The shared address is used by the nodes that are running the scalable application, to distribute incoming data.

- Be online on the primary cluster and offline on the secondary cluster

The following figure illustrates the configuration of resource groups in a scalable application.
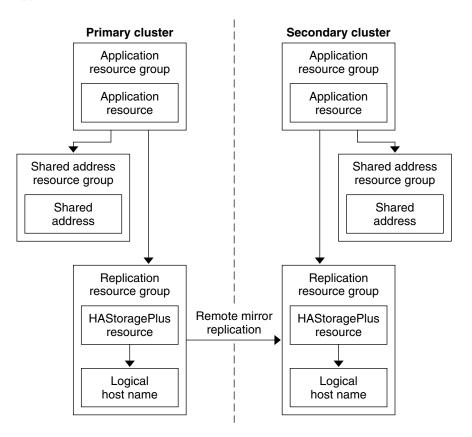


**FIGURE 6–5** Configuration of Resource Groups in a Scalable Application

## Guidelines for Managing a Failover or Switchover

If the primary cluster fails, the application must be switched over to the secondary cluster as soon as possible. To enable the secondary cluster to take over, the DNS must be updated.

The DNS associates a client with the logical hostname of an application. After a failover or switchover, the DNS mapping to the primary cluster must be removed, and a DNS mapping to the secondary cluster must be created. The following figure shows how the DNS maps a client to a cluster.
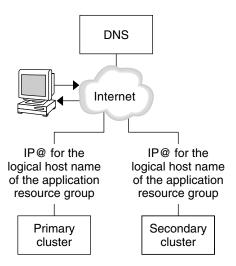
**FIGURE 6–6** DNS Mapping of a Client to a Cluster

To update the DNS, use the `nsupdate` command. For information, see the `nsupdate`(1M) man page. For an example of how to manage a failover or switchover, see "Example of How to Manage a Failover or Switchover" on page 282.

After repair, the primary cluster can be brought back online. To switch back to the original primary cluster, perform the following tasks:

1. Synchronize the primary cluster with the secondary cluster to ensure that the primary volume is up-to-date.

2. Update the DNS so that clients can access the application on the primary cluster.

# Task Map: Example of a Data–Replication Configuration

The following task map lists the tasks in this example of how data replication was configured for an NFS application by using Sun StorEdge Availability Suite software.

**TABLE 6–1** Task Map: Example of a Data Replication Configuration

| Task | Instructions |
|---|---|
| 1. Connect and install the clusters. | "Connecting and Installing the Clusters" on page 259 |
| 2. Configure disk device groups, file systems for the NFS application, and resource groups on the primary cluster and on the secondary cluster. | "Example of How to Configure Device Groups and Resource Groups" on page 261 |
| 3. Enable data replication on the primary cluster and on the secondary cluster. | "How to Enable Replication on the Primary Cluster" on page 274<br><br>"How to Enable Replication on the Secondary Cluster" on page 276 |
| 4. Perform data replication. | "How to Perform a Remote Mirror Replication" on page 277<br><br>"How to Perform a Point-in-Time Snapshot" on page 278 |
| 5. Verify the data replication configuration. | "How to Verify That Replication Is Configured Correctly" on page 279 |

# Connecting and Installing the Clusters

Figure 6–7 illustrates the cluster configuration used in the example configuration. The secondary cluster in the example configuration contains one node, but other cluster configurations can be used.
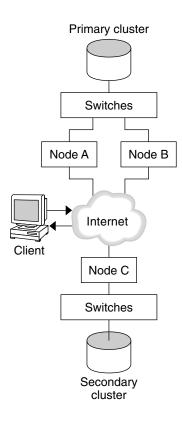
**FIGURE 6–7** Example Cluster Configuration

Table 6–2 summarizes the hardware and software required by the example configuration. The Solaris OS, Sun Cluster software, and volume manager software must be installed on the cluster nodes *before* you install Sun StorEdge Availability Suite software and patches.

**TABLE 6–2** Required Hardware and Software

| Hardware or Software | Requirement |
| --- | --- |
| Node hardware | Sun StorEdge Availability Suite software is supported on all servers using the Solaris OS. |
|  | For information about which hardware to use, see the *Sun Cluster 3.0-3.1 Hardware Administration Manual for Solaris OS* |
| Disk space | Approximately 15 Mbytes. |

**TABLE 6–2** Required Hardware and Software     *(Continued)*

| Hardware or Software | Requirement |
| --- | --- |
| Solaris OS | Solaris OS releases that are supported by Sun Cluster software. |
| | All nodes must use the same version of the Solaris OS. |
| | For information about installation, see "Installing the Software" on page 45. |
| Sun Cluster software | Sun Cluster 3.1 8/05 software. |
| | For information about installation, see Chapter 2. |
| Volume manager software | Solstice DiskSuite or Solaris Volume Manager software or VERITAS Volume Manager (VxVM) software. |
| | All nodes must use the same version of volume manager software. |
| | Information about installation is in "Installing and Configuring Solstice DiskSuite or Solaris Volume Manager Software" on page 141 and "SPARC: Installing and Configuring VxVM Software" on page 177. |
| Sun StorEdge Availability Suite software | For information about how to install the software, see the installation manuals for your release of Sun StorEdge Availability Suite software:<br>■ Sun StorEdge Availability Suite 3.1 - *Sun StorEdge Availability Suite 3.1 Point-in-Time Copy Software Installation Guide* and *Sun StorEdge Availability Suite 3.1 Remote Mirror Software Installation Guide*<br>■ Sun StorEdge Availability Suite 3.2 - *Sun StorEdge Availability Suite 3.2 Software Installation Guide* |
| Sun StorEdge Availability Suite software patches | For information about the latest patches, see `http://www.sunsolve.com`. |

# Example of How to Configure Device Groups and Resource Groups

This section describes how disk device groups and resource groups are configured for an NFS application. For additional information, see "Configuring Replication Resource Groups" on page 254 and "Configuring Application Resource Groups" on page 254.

This section contains the following procedures:

■ "How to Configure a Disk Device Group on the Primary Cluster" on page 263
■ "How to Configure a Disk Device Group on the Secondary Cluster" on page 264

The following table lists the names of the groups and resources that are created for the example configuration.

**TABLE 6–3** Summary of the Groups and Resources in the Example Configuration

| Group or Resource | Name | Description |
|---|---|---|
| Disk device group | *devicegroup* | The disk device group |
| Replication resource group and resources | *devicegroup*-stor-rg | The replication resource group |
| | *lhost-reprg-prim*, *lhost-reprg-sec* | The logical hostnames for the replication resource group on the primary cluster and the secondary cluster |
| | *devicegroup-stor* | The HAStoragePlus resource for the replication resource group |
| Application resource group and resources | *nfs-rg* | The application resource group |
| | *lhost-nfsrg-prim*, *lhost-nfsrg-sec* | The logical hostnames for the application resource group on the primary cluster and the secondary cluster |
| | *nfs-dg-rs* | The HAStoragePlus resource for the application |
| | *nfs-rs* | The NFS resource |

With the exception of *devicegroup*-stor-rg, the names of the groups and resources are example names that can be changed as required. The replication resource group must have a name with the format *devicegroup*-stor-rg .

This example configuration uses VxVM software. For information about Solstice DiskSuite or Solaris Volume Manager software, see Chapter 3.

The following figure illustrates the volumes that are created in the disk device group.
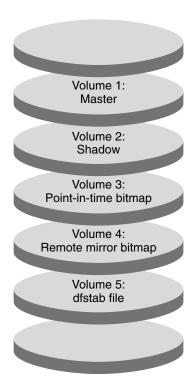
**FIGURE 6–8** Volumes for the Disk Device Group

---

**Note –** The volumes defined in this procedure must not include disk-label private areas, for example, cylinder 0. The VxVM software manages this constraint automatically.

---

## ▼ How to Configure a Disk Device Group on the Primary Cluster

**Before You Begin**

Ensure that you have completed the following tasks:

- Read the guidelines and requirements in the following sections:

  - "Introduction to Data Replication" on page 249
  - "Guidelines for Configuring Data Replication" on page 253

- Set up the primary and secondary clusters as described in "Connecting and Installing the Clusters" on page 259.

**Steps** 1. **Access `nodeA` as superuser.**

nodeA is the first node of the primary cluster. For a reminder of which node is nodeA, see Figure 6–7.

2. **Create a disk group on `nodeA` that contains four volumes: volume 1, `vol01` through volume 4, `vol04`.**

For information about configuring a disk group by using the VxVM software, see Chapter 4.

3. **Configure the disk group to create a disk device group.**

*nodeA*# **/usr/cluster/bin/scconf -a** \
**-D type=vxvm,name=***devicegroup***,nodelist=***nodeA*:*nodeB*

The disk device group is called *devicegroup*.

4. **Create the file system for the disk device group.**

*nodeA*# **/usr/sbin/newfs /dev/vx/rdsk/***devicegroup***/vol01 < /dev/null**
*nodeA*# **/usr/sbin/newfs /dev/vx/rdsk/***devicegroup***/vol02 < /dev/null**

No file system is needed for vol03 or vol04, which are instead used as raw volumes.

**Next Steps** Go to "How to Configure a Disk Device Group on the Secondary Cluster" on page 264.

## ▼ How to Configure a Disk Device Group on the Secondary Cluster

**Before You Begin** Ensure that you completed steps in "How to Configure a Disk Device Group on the Primary Cluster" on page 263.

**Steps** 1. **Access `nodeC` as superuser.**

2. **Create a disk group on `nodeC` that contains four volumes: volume 1, `vol01` through volume 4, `vol04`.**

3. **Configure the disk group to create a disk device group.**

*nodeC*# **/usr/cluster/bin/scconf -a** \
**-D type=vxvm,name=***devicegroup***,nodelist=***nodeC*

The disk device group is called *devicegroup*.

4. **Create the file system for the disk device group.**

*nodeC*# **/usr/sbin/newfs /dev/vx/rdsk/***devicegroup***/vol01 < /dev/null**
*nodeC*# **/usr/sbin/newfs /dev/vx/rdsk/***devicegroup***/vol02 < /dev/null**

No file system is needed for vol03 or vol04, which are instead used as raw volumes.

## ▼ How to Configure the File System on the Primary Cluster for the NFS Application

**Before You Begin** Ensure that you completed steps in "How to Configure a Disk Device Group on the Secondary Cluster" on page 264.

**Steps** 1. **On nodeA and nodeB, create a mount point directory for the NFS file system.**
For example:

*nodeA*# **mkdir /global/***mountpoint*

2. **On nodeA and nodeB, configure the master volume to be mounted automatically on the mount point.**
Add or replace the following text to the /etc/vfstab file on nodeA and nodeB. The text must be on a single line.

/dev/vx/dsk/*devicegroup*/vol01 /dev/vx/rdsk/*devicegroup*/vol01 \
/global/*mountpoint* ufs 3 no global,logging

For a reminder of the volumes names and volume numbers used in the disk device group, see Figure 6–8.

3. **On nodeA, create a volume for the file system information that is used by the Sun Cluster HA for NFS data service.**

*nodeA*# **/usr/sbin/vxassist -g** *devicegroup* **make vol05 120m** *disk1*

Volume 5, vol05, contains the file system information that is used by the Sun Cluster HA for NFS data service.

4. **On nodeA, resynchronize the device group with the Sun Cluster software.**

*nodeA*# **/usr/cluster/bin/scconf -c -D name=***devicegroup***,sync**

5. **On nodeA, create the file system for vol05.**

*nodeA*# **/usr/sbin/newfs /dev/vx/rdsk/***devicegroup***/vol05**

6. **On nodeA and nodeB, create a mount point for vol05.**
For example:

*nodeA*# **mkdir /global/etc**

7. **On nodeA and nodeB, configure vol05 to be mounted automatically on the mount point.**

Add or replace the following text to the /etc/vfstab file on nodeA and nodeB. The text must be on a single line.

```
/dev/vx/dsk/devicegroup/vol05 /dev/vx/rdsk/devicegroup/vol05 \
/global/etc ufs 3 yes global,logging
```

8. **Mount vol05 on nodeA.**

   *nodeA#* **mount /global/etc**

9. **Make vol05 accessible to remote systems.**

   a. **Create a directory called /global/etc/SUNW.nfs on nodeA.**

      *nodeA#* **mkdir -p /global/etc/SUNW.nfs**

   b. **Create the file /global/etc/SUNW.nfs/dfstab.***nfs-rs* **on nodeA.**

      *nodeA#* **touch /global/etc/SUNW.nfs/dfstab.***nfs-rs*

   c. **Add the following line to the /global/etc/SUNW.nfs/dfstab.***nfs-rs* **file on nodeA:**

      ```
      share -F nfs -o rw -d "HA NFS" /global/mountpoint
      ```

**Next Steps**   Go to

## ▼ How to Configure the File System on the Secondary Cluster for the NFS Application

**Before You Begin**   Ensure that you completed steps in

**Steps**   1. **On nodeC, create a mount point directory for the NFS file system.**
   For example:

   *nodeC#* **mkdir /global/***mountpoint*

2. **On nodeC, configure the master volume to be mounted automatically on the mount point.**
   Add or replace the following text to the /etc/vfstab file on nodeC. The text must be on a single line.

   ```
   /dev/vx/dsk/devicegroup/vol01 /dev/vx/rdsk/devicegroup/vol01 \
   /global/mountpoint ufs 3 no global,logging
   ```

3. **On nodeC, create a volume for the file system information that is used by the Sun Cluster HA for NFS data service.**

   *nodeC#* **/usr/sbin/vxassist -g** *devicegroup* **make vol05 120m** *disk1*

Volume 5, `vol05`, contains the file system information that is used by the Sun Cluster HA for NFS data service.

4. **On `nodeC`, resynchronize the device group with the Sun Cluster software.**

   *nodeC#* `/usr/cluster/bin/scconf -c -D name=`*devicegroup*`,sync`

5. **On `nodeC`, create the file system for `vol05`.**

   *nodeC#* `/usr/sbin/newfs /dev/vx/rdsk/`*devicegroup*`/vol05`

6. **On `nodeC`, create a mount point for `vol05`.**

   For example:

   *nodeC#* `mkdir /global/etc`

7. **On `nodeC`, configure `vol05` to be mounted automatically on the mount point.**

   Add or replace the following text to the `/etc/vfstab` file on `nodeC`. The text must be on a single line.

   `/dev/vx/dsk/`*devicegroup*`/vol05 /dev/vx/rdsk/`*devicegroup*`/vol05 \`
   `/global/etc ufs 3 yes global,logging`

8. **Mount `vol05` on `nodeC`.**

   *nodeC#* `mount /global/etc`

9. **Make `vol05` accessible to remote systems.**

   a. **Create a directory called `/global/etc/SUNW.nfs` on `nodeC`.**

      *nodeC#* `mkdir -p /global/etc/SUNW.nfs`

   b. **Create the file `/global/etc/SUNW.nfs/dfstab.`*nfs-rs* on `nodeC`.**

      *nodeC#* `touch /global/etc/SUNW.nfs/dfstab.`*nfs-rs*

   c. **Add the following line to the `/global/etc/SUNW.nfs/dfstab.`*nfs-rs* file on `nodeC`:**

      `share -F nfs -o rw -d "HA NFS" /global/`*mountpoint*

**Next Steps**   Go to

## ▼ How to Create a Replication Resource Group on the Primary Cluster

**Before You Begin**   Ensure that you completed steps in

**Steps**   1.   **Access `nodeA` as superuser.**

2.   **Register `SUNW.HAStoragePlus` as a resource type.**

  *nodeA#* `/usr/cluster/bin/scrgadm -a -t SUNW.HAStoragePlus`

3.   **Create a replication resource group for the disk device group.**

  *nodeA#* `/usr/cluster/bin/scrgadm -a -g` *devicegroup*`-stor-rg -h` *nodeA,nodeB*

  | | |
  |---|---|
  | *devicegroup* | The name of the disk device group |
  | *devicegroup*`-stor-rg` | The name of the replication resource group |
  | `-h` *nodeA, nodeB* | Specifies the cluster nodes that can master the replication resource group |

4.   **Add a `SUNW.HAStoragePlus` resource to the replication resource group.**

  *nodeA#* `/usr/cluster/bin/scrgadm -a -j` *devicegroup-stor* `\`
  `-g` *devicegroup*`-stor-rg -t SUNW.HAStoragePlus \`
  `-x GlobalDevicePaths=`*devicegroup* `\`
  `-x AffinityOn=True`

  | | |
  |---|---|
  | *devicegroup-stor* | The HAStoragePlus resource for the replication resource group. |
  | `-x GlobalDevicePaths=` | Specifies the extension property that Sun StorEdge Availability Suite software relies on. |
  | `-x AffinityOn=True` | Specifies that the `SUNW.HAStoragePlus` resource must perform an affinity switchover for the global devices and cluster file systems defined by `-x GlobalDevicePaths=`. Therefore, when the replication resource group fails over or is switched over, the associated device group is switched over. |

  For more information about these extension properties, see the `SUNW.HAStoragePlus`(5) man page.

5.   **Add a logical hostname resource to the replication resource group.**

  *nodeA#* `/usr/cluster/bin/scrgadm -a -L -j` *lhost-reprg-prim* `\`
  `-g` *devicegroup*`-stor-rg -l` *lhost-reprg-prim*

  *lhost-reprg-prim* is the logical hostname for the replication resource group on the primary cluster.

6.   **Enable the resources, manage the resource group, and bring the resource group online.**

  *nodeA#* `/usr/cluster/bin/scswitch -Z -g` *devicegroup*`-stor-rg`
  *nodeA#* `/usr/cluster/bin/scswitch -z -g` *devicegroup*`-stor-rg -h` *nodeA*

7. **Verify that the resource group is online.**

*nodeA#* **/usr/cluster/bin/scstat -g**

Examine the resource group state field to confirm that the replication resource group is online on nodeA.

**Next Steps**  Go to

## ▼ How to Create a Replication Resource Group on the Secondary Cluster

**Before You Begin**  Ensure that you completed steps in

**Steps**  1. **Access nodeC as superuser.**

2. **Register SUNW.HAStoragePlus as a resource type.**

   *nodeC#* **/usr/cluster/bin/scrgadm -a -t SUNW.HAStoragePlus**

3. **Create a replication resource group for the disk device group.**

   *nodeC#* **/usr/cluster/bin/scrgadm -a -g** *devicegroup***-stor-rg -h** *nodeC*

   | *devicegroup* | The name of the disk device group |
   | *devicegroup***-stor-rg** | The name of the replication resource group |
   | **-h** *nodeC* | Specifies the cluster node that can master the replication resource group |

4. **Add a SUNW.HAStoragePlus resource to the replication resource group.**

   *nodeC#* **/usr/cluster/bin/scrgadm -a -j** *devicegroup-stor* **\**
   **-g** *devicegroup-stor-rg* **-t SUNW.HAStoragePlus \**
   **-x GlobalDevicePaths=***devicegroup* **\**
   **-x AffinityOn=True**

   | *devicegroup-stor* | The HAStoragePlus resource for the replication resource group. |
   | **-x GlobalDevicePaths=** | Specifies the extension property that Sun StorEdge Availability Suite software relies on. |
   | **-x AffinityOn=True** | Specifies that the SUNW.HAStoragePlus resource must perform an affinity switchover for the global devices and cluster file systems defined by -x GlobalDevicePaths=. Therefore, when the |

replication resource group fails over or is switched over, the associated device group is switched over.

For more information about these extension properties, see the SUNW.HAStoragePlus(5) man page.

5. **Add a logical hostname resource to the replication resource group.**

   *nodeC#* **/usr/cluster/bin/scrgadm -a -L -j** *lhost-reprg-sec* **\**
   **-g** *devicegroup***-stor-rg -l** *lhost-reprg-sec*

   *lhost-reprg-sec* is the logical hostname for the replication resource group on the primary cluster.

6. **Enable the resources, manage the resource group, and bring the resource group online.**

   *nodeC#* **/usr/cluster/bin/scswitch -Z -g** *devicegroup***-stor-rg**

7. **Verify that the resource group is online.**

   *nodeC#* **/usr/cluster/bin/scstat -g**

   Examine the resource group state field to confirm that the replication resource group is online on nodeC.

**Next Steps**    Go to .

## ▼ How to Create an NFS Application Resource Group on the Primary Cluster

This procedure describes how application resource groups are created for NFS. This procedure is specific to this application and cannot be used for another type of application.

**Before You Begin**    Ensure that you completed steps in .

**Steps**    1. **Access nodeA as superuser.**

2. **Register SUNW.nfs as a resource type.**

   *nodeA#* **scrgadm -a -t SUNW.nfs**

3. **If SUNW.HAStoragePlus has not been registered as a resource type, register it.**

   *nodeA#* **scrgadm -a -t SUNW.HAStoragePlus**

4. **Create an application resource group for the** *devicegroup*.

   *nodeA#* `scrgadm -a -g` *nfs-rg* `\`
   `-y Pathprefix=/global/etc \`
   `-y Auto_start_on_new_cluster=False \`
   `-y RG_dependencies=`*devicegroup*`-stor-rg`

   *nfs-rg*
   > The name of the application resource group.

   **Pathprefix=/global/etc**
   > Specifies a directory into which the resources in the group can write administrative files.

   **Auto_start_on_new_cluster=False**
   > Specifies that the application resource group is not started automatically.

   **RG_dependencies=**_devicegroup_**-stor-rg**
   > Specifies the resource groups that the application resource group depends on. In this example, the application resource group depends on the replication resource group.
   >
   > If the application resource group is switched over to a new primary node, the replication resource group is automatically switched over. However, if the replication resource group is switched over to a new primary node, the application resource group must be manually switched over.

5. **Add a `SUNW.HAStoragePlus` resource to the application resource group.**

   *nodeA#* `scrgadm -a -j` *nfs-dg-rs* `-g` *nfs-rg* `\`
   `-t SUNW.HAStoragePlus \`
   `-x FileSystemMountPoints=/global/`*mountpoint* `\`
   `-x AffinityOn=True`

   *nfs-dg-rs*
   > Is the name of the HAStoragePlus resource for the NFS application.

   **-x FileSystemMountPoints=/global/**
   > Specifies that the mount point for the file system is global.

   **-t SUNW.HAStoragePlus**
   > Specifies that the resource is of the type SUNW.HAStoragePlus.

   **-x AffinityOn=True**
   > Specifies that the application resource must perform an affinity switchover for the global devices and cluster file systems defined by -x GlobalDevicePaths=. Therefore, when the application resource group fails over or is switched over, the associated device group is switched over.

   For more information about these extension properties, see the SUNW.HAStoragePlus(5) man page.

6. **Add a logical hostname resource to the application resource group.**

   *nodeA#* `/usr/cluster/bin/scrgadm -a -L -j` *lhost-nfsrg-prim* `-g` *nfs-rg* `\`
   `-l` *lhost-nfsrg-prim*

*lhost-nfsrg-prim* is the logical hostname of the application resource group on the primary cluster.

7. **Enable the resources, manage the application resource group, and bring the application resource group online.**

   a. **Bring the HAStoragePlus resource for the NFS application online.**

      *nodeA#* `/usr/cluster/bin/scrgadm -a -g` *nfs-rg* `\`
      `-j` *nfs-rs* `-t SUNW.nfs -y Resource_dependencies=`*nfs-dg-rs*

   b. **Bring the application resource group online on `nodeA`.**

      *nodeA#* `/usr/cluster/bin/scswitch -Z -g` *nfs-rg*
      *nodeA#* `/usr/cluster/bin/scswitch -z -g` *nfs-rg* `-h` *nodeA*

8. **Verify that the application resource group is online.**

   *nodeA#* `/usr/cluster/bin/scstat -g`

   Examine the resource group state field to determine whether the application resource group is online for `nodeA` and `nodeB`.

## ▼ How to Create an NFS Application Resource Group on the Secondary Cluster

**Steps**   1. **Access `nodeC` as superuser.**

2. **Register `SUNW.nfs` as a resource type.**

   *nodeC#* `scrgadm -a -t SUNW.nfs`

3. **If `SUNW.HAStoragePlus` has not been registered as a resource type, register it.**

   *nodeC#* `scrgadm -a -t SUNW.HAStoragePlus`

4. **Create an application resource group for the *devicegroup*.**

   *nodeC#* `scrgadm -a -g` *nfs-rg* `\`
   `-y Pathprefix=/global/etc \`
   `-y Auto_start_on_new_cluster=False \`
   `-y RG_dependencies=`*devicegroup*`-stor-rg`

   *nfs-rg*
   　　The name of the application resource group.

**Pathprefix=/global/etc**
Specifies a directory into which the resources in the group can write administrative files.

**Auto_start_on_new_cluster=False**
Specifies that the application resource group is not started automatically.

**RG_dependencies=***devicegroup***-stor-rg**
Specifies the resource groups that the application resource group depends on. In this example, the application resource group depends on the replication resource group.

If the application resource group is switched over to a new primary node, the replication resource group is automatically switched over. However, if the replication resource group is switched over to a new primary node, the application resource group must be manually switched over.

5. **Add a SUNW.HAStoragePlus resource to the application resource group.**

*nodeC#* **scrgadm -a -j** *nfs-dg-rs* **-g** *nfs-rg* \
**-t SUNW.HAStoragePlus** \
**-x FileSystemMountPoints=/global/***mountpoint* \
**-x AffinityOn=True**

*nfs-dg-rs*
Is the name of the HAStoragePlus resource for the NFS application.

**-x FileSystemMountPoints=/global/**
Specifies that the mount point for the file system is global.

**-t SUNW.HAStoragePlus**
Specifies that the resource is of the type SUNW.HAStoragePlus.

**-x AffinityOn=True**
Specifies that the application resource must perform an affinity switchover for the global devices and cluster file systems defined by -x GlobalDevicePaths=. Therefore, when the application resource group fails over or is switched over, the associated device group is switched over.

For more information about these extension properties, see the SUNW.HAStoragePlus(5) man page.

6. **Add a logical hostname resource to the application resource group.**

*nodeC#* **/usr/cluster/bin/scrgadm -a -L -j** *lhost-nfsrg-sec* **-g** *nfs-rg* \
**-l** *lhost-nfsrg-sec*

*lhost-nfsrg-sec* is the logical hostname of the application resource group on the secondary cluster.

7. **Add an NFS resource to the application resource group.**

*nodeC#* **/usr/cluster/bin/scrgadm -a -g** *nfs-rg* \
**-j** *nfs-rs* **-t SUNW.nfs -y Resource_dependencies=***nfs-dg-rs*

8. **Ensure that the application resource group does not come online on `nodeC`.**

   *nodeC#* **/usr/cluster/bin/scswitch -n -j** *nfs-rs*
   *nodeC#* **/usr/cluster/bin/scswitch -n -j** *nfs-dg-rs*
   *nodeC#* **/usr/cluster/bin/scswitch -n -j** *lhost-nfsrg-sec*
   *nodeC#* **/usr/cluster/bin/scswitch -z -g** *nfs-rg* **-h ""**

   The resource group remains offline after a reboot, because
   *Auto_start_on_new_cluster=False*.

9. **If the global volume is mounted on the primary cluster, unmount the global volume from the secondary cluster.**

   *nodeC#* **umount /global/***mountpoint*

   If the volume is mounted on a secondary cluster, the synchronization fails.

**Next Steps**   Go to

# Example of How to Enable Data Replication

This section describes how data replication is enabled for the example configuration. This section uses the Sun StorEdge Availability Suite software commands `sndradm` and `iiadm`. For more information about these commands, see the *Sun Cluster 3.0 and Sun StorEdge Software Integration Guide*.

This section contains the following procedures:

- "How to Enable Replication on the Primary Cluster" on page 274
- "How to Enable Replication on the Secondary Cluster" on page 276

## ▼ How to Enable Replication on the Primary Cluster

**Steps**   1. **Access `nodeA` as superuser.**

2. **Flush all transactions.**

   *nodeA#* **/usr/sbin/lockfs -a -f**

3. **Confirm that the logical hostnames** *lhost-reprg-prim* **and** *lhost-reprg-sec* **are online.**

   *nodeA#* **/usr/cluster/bin/scstat -g**
   *nodeC#* **/usr/cluster/bin/scstat -g**

Examine the state field of the resource group.

4. **Enable remote mirror replication from the primary cluster to the secondary cluster.**

   This step enables replication from the master volume on the primary cluster to the master volume on the secondary cluster. In addition, this step enables replication to the remote mirror bitmap on vol04.

   ■ If the primary cluster and secondary cluster are unsynchronized, run this command:

   *nodeA#* **/usr/opt/SUNWesm/sbin/sndradm -n -e** *lhost-reprg-prim* \
   **/dev/vx/rdsk/***devicegroup***/vol01** \
   **/dev/vx/rdsk/***devicegroup***/vol04** *lhost-reprg-sec* \
   **/dev/vx/rdsk/***devicegroup***/vol01** \
   **/dev/vx/rdsk/***devicegroup***/vol04 ip sync**

   ■ If the primary cluster and secondary cluster are synchronized, run this command:

   *nodeA#* **/usr/opt/SUNWesm/sbin/sndradm -n -E** *lhost-reprg-prim* \
   **/dev/vx/rdsk/***devicegroup***/vol01** \
   **/dev/vx/rdsk/***devicegroup***/vol04** *lhost-reprg-sec* \
   **/dev/vx/rdsk/***devicegroup***/vol01** \
   **/dev/vx/rdsk/***devicegroup***/vol04 ip sync**

5. **Enable autosynchronization.**

   *nodeA#* **/usr/opt/SUNWesm/sbin/sndradm -n -a on** *lhost-reprg-prim* \
   **/dev/vx/rdsk/***devicegroup***/vol01** \
   **/dev/vx/rdsk/***devicegroup***/vol04** *lhost-reprg-sec* \
   **/dev/vx/rdsk/***devicegroup***/vol01** \
   **/dev/vx/rdsk/***devicegroup***/vol04 ip sync**

   This step enables autosynchronization. When the active state of autosynchronization is set to on, the volume sets are resynchronized if the system reboots or a failure occurs.

6. **Verify that the cluster is in logging mode.**

   *nodeA#* **/usr/opt/SUNWesm/sbin/sndradm -P**

   The output should resemble the following:

   /dev/vx/rdsk/*devicegroup*/vol01 ->
   *lhost-reprg-sec*:/dev/vx/rdsk/*devicegroup*/vol01
   autosync: off, max q writes:4194304, max q fbas:16384, mode:sync,ctag:
   *devicegroup*, state: logging

   In logging mode, the state is logging, and the active state of autosynchronization is off. When the data volume on the disk is written to, the bitmap file on the same disk is updated.

7. **Enable point-in-time snapshot.**

   *nodeA#* **/usr/opt/SUNWesm/sbin/iiadm -e ind** \
   **/dev/vx/rdsk/***devicegroup***/vol01** \

   Chapter 6 • Configuring Data Replication With Sun StorEdge Availability Suite Software **275**

```
/dev/vx/rdsk/devicegroup/vol02 \
/dev/vx/rdsk/devicegroup/vol03
nodeA# /usr/opt/SUNWesm/sbin/iiadm -w \
/dev/vx/rdsk/devicegroup/vol02
```

This step enables the master volume on the primary cluster to be copied to the shadow volume on the same cluster. The master volume, shadow volume, and point-in-time bitmap volume must be in the same device group. In this example, the master volume is vol01, the shadow volume is vol02, and the point-in-time bitmap volume is vol03.

8.  **Attach the point-in-time snapshot to the remote mirror set.**

```
nodeA# /usr/opt/SUNWesm/sbin/sndradm -I a \
/dev/vx/rdsk/devicegroup/vol01 \
/dev/vx/rdsk/devicegroup/vol02 \
/dev/vx/rdsk/devicegroup/vol03
```

This step associates the point-in-time snapshot with the remote mirror volume set. Sun StorEdge Availability Suite software ensures that a point-in-time snapshot is taken before remote mirror replication can occur.

**Next Steps** Go to

## ▼ How to Enable Replication on the Secondary Cluster

**Before You Begin** Ensure that you completed steps in

**Steps**
1.  **Access nodeC as superuser.**

2.  **Flush all transactions.**

    ```
    nodeC# /usr/sbin/lockfs -a -f
    ```

3.  **Enable remote mirror replication from the primary cluster to the secondary cluster.**

    ```
    nodeC# /usr/opt/SUNWesm/sbin/sndradm -n -e lhost-reprg-prim \
    /dev/vx/rdsk/devicegroup/vol01 \
    /dev/vx/rdsk/devicegroup/vol04 lhost-reprg-sec \
    /dev/vx/rdsk/devicegroup/vol01 \
    /dev/vx/rdsk/devicegroup/vol04 ip sync
    ```

    The primary cluster detects the presence of the secondary cluster and starts synchronization. Refer to the system log file /var/opt/SUNWesm/ds.log for information about the status of the clusters.

4.  **Enable independent point-in-time snapshot.**

    ```
    nodeC# /usr/opt/SUNWesm/sbin/iiadm -e ind \
    /dev/vx/rdsk/devicegroup/vol01 \
    ```

```
/dev/vx/rdsk/devicegroup/vol02 \
/dev/vx/rdsk/devicegroup/vol03
nodeC# /usr/opt/SUNWesm/sbin/iiadm -w \
/dev/vx/rdsk/devicegroup/vol02
```

**5. Attach the point-in-time snapshot to the remote mirror set.**

```
nodeC# /usr/opt/SUNWesm/sbin/sndradm -I a \
/dev/vx/rdsk/devicegroup/vol01 \
/dev/vx/rdsk/devicegroup/vol02 \
/dev/vx/rdsk/devicegroup/vol03
```

**Next Steps**   Go to

# Example of How to Perform Data Replication

This section describes how data replication is performed for the example configuration. This section uses the Sun StorEdge Availability Suite software commands sndradm and iiadm. For more information about these commands, see the *Sun Cluster 3.0 and Sun StorEdge Software Integration Guide*.

This section contains the following procedures:

-
-
-

## ▼ How to Perform a Remote Mirror Replication

In this procedure, the master volume of the primary disk is replicated to the master volume on the secondary disk. The master volume is vol01 and the remote mirror bitmap volume is vol04.

**Steps**   **1. Access nodeA as superuser.**

**2. Verify that the cluster is in logging mode.**

```
nodeA# /usr/opt/SUNWesm/sbin/sndradm -P
```

The output should resemble the following:

```
/dev/vx/rdsk/devicegroup/vol01 ->
lhost-reprg-sec:/dev/vx/rdsk/devicegroup/vol01
autosync: off, max q writes:4194304, max q fbas:16384, mode:sync,ctag:
```

*devicegroup*, state: logging

In logging mode, the state is logging, and the active state of autosynchronization is off. When the data volume on the disk is written to, the bitmap file on the same disk is updated.

3. **Flush all transactions.**

   *nodeA*# **/usr/sbin/lockfs -a -f**

4. **Repeat Step 1 through Step 3 on nodeC.**

5. **Copy the master volume of nodeA to the master volume of nodeC.**

   *nodeA*# **/usr/opt/SUNWesm/sbin/sndradm -n -m** *lhost-reprg-prim* \
   **/dev/vx/rdsk/***devicegroup***/vol01** \
   **/dev/vx/rdsk/***devicegroup***/vol04** *lhost-reprg-sec* \
   **/dev/vx/rdsk/***devicegroup***/vol01** \
   **/dev/vx/rdsk/***devicegroup***/vol04 ip sync**

6. **Wait until the replication is complete and the volumes are synchronized.**

   *nodeA*# **/usr/opt/SUNWesm/sbin/sndradm -n -w** *lhost-reprg-prim* \
   **/dev/vx/rdsk/***devicegroup***/vol01** \
   **/dev/vx/rdsk/***devicegroup***/vol04** *lhost-reprg-sec* \
   **/dev/vx/rdsk/***devicegroup***/vol01** \
   **/dev/vx/rdsk/***devicegroup***/vol04 ip sync**

7. **Confirm that the cluster is in replicating mode.**

   *nodeA*# **/usr/opt/SUNWesm/sbin/sndradm -P**

   The output should resemble the following:

   ```
   /dev/vx/rdsk/devicegroup/vol01 ->
   lhost-reprg-sec:/dev/vx/rdsk/devicegroup/vol01
   autosync: on, max q writes:4194304, max q fbas:16384, mode:sync,ctag:
   devicegroup, state: replicating
   ```

   In replicating mode, the state is replicating, and the active state of autosynchronization is on. When the primary volume is written to, the secondary volume is updated by Sun StorEdge Availability Suite software.

**Next Steps**     Go to "How to Perform a Point-in-Time Snapshot" on page 278.

## ▼ How to Perform a Point-in-Time Snapshot

In this procedure, point-in-time snapshot is used to synchronize the shadow volume of the primary cluster to the master volume of the primary cluster. The master volume is vol01, the bitmap volume is vol04, and the shadow volume is vol02.

**Before You Begin**     Ensure that you completed steps in "How to Perform a Remote Mirror Replication" on page 277.

**Steps** 1. **Access nodeA as superuser.**

2. **Disable the resource that is running on nodeA.**

   *nodeA#* `/usr/cluster/bin/scswitch -n -j` *nfs-rs*

3. **Change the primary cluster to logging mode.**

   *nodeA#* `/usr/opt/SUNWesm/sbin/sndradm -n -l` *lhost-reprg-prim* `\`
   `/dev/vx/rdsk/`*devicegroup*`/vol01 \`
   `/dev/vx/rdsk/`*devicegroup*`/vol04` *lhost-reprg-sec* `\`
   `/dev/vx/rdsk/`*devicegroup*`/vol01 \`
   `/dev/vx/rdsk/`*devicegroup*`/vol04 ip sync`

   When the data volume on the disk is written to, the bitmap file on the same disk is updated. No replication occurs.

4. **Synchronize the shadow volume of the primary cluster to the master volume of the primary cluster.**

   *nodeA#* `/usr/opt/SUNWesm/sbin/iiadm -u s /dev/vx/rdsk/`*devicegroup*`/vol02`
   *nodeA#* `/usr/opt/SUNWesm/sbin/iiadm -w /dev/vx/rdsk/`*devicegroup*`/vol02`

5. **Synchronize the shadow volume of the secondary cluster to the master volume of the secondary cluster.**

   *nodeC#* `/usr/opt/SUNWesm/sbin/iiadm -u s /dev/vx/rdsk/`*devicegroup*`/vol02`
   *nodeC#* `/usr/opt/SUNWesm/sbin/iiadm -w /dev/vx/rdsk/`*devicegroup*`/vol02`

6. **Restart the application on nodeA.**

   *nodeA#* `/usr/cluster/bin/scswitch -e -j` *nfs-rs*

7. **Resynchronize the secondary volume with the primary volume.**

   *nodeA#* `/usr/opt/SUNWesm/sbin/sndradm -n -u` *lhost-reprg-prim* `\`
   `/dev/vx/rdsk/`*devicegroup*`/vol01 \`
   `/dev/vx/rdsk/`*devicegroup*`/vol04` *lhost-reprg-sec* `\`
   `/dev/vx/rdsk/`*devicegroup*`/vol01 \`
   `/dev/vx/rdsk/`*devicegroup*`/vol04 ip sync`

**Next Steps** Go to

## ▼ How to Verify That Replication Is Configured Correctly

**Before You Begin** Ensure that you completed steps in .

**Steps** 1. **Verify that the primary cluster is in replicating mode, with autosynchronization on.**

   *nodeA#* `/usr/opt/SUNWesm/sbin/sndradm -P`

The output should resemble the following:

```
/dev/vx/rdsk/devicegroup/vol01 ->
lhost-reprg-sec:/dev/vx/rdsk/devicegroup/vol01
autosync: on, max q writes:4194304, max q fbas:16384, mode:sync,ctag:
devicegroup, state: replicating
```

In replicating mode, the state is `replicating`, and the active state of autosynchronization is `on`. When the primary volume is written to, the secondary volume is updated by Sun StorEdge Availability Suite software.

**2. If the primary cluster is not in replicating mode, put it into replicating mode.**

*nodeA#* **/usr/opt/SUNWesm/sbin/sndradm -n -u** *lhost-reprg-prim* \
**/dev/vx/rdsk/***devicegroup***/vol01** \
**/dev/vx/rdsk/***devicegroup***/vol04** *lhost-reprg-sec* \
**/dev/vx/rdsk/***devicegroup***/vol01** \
**/dev/vx/rdsk/***devicegroup***/vol04 ip sync**

**3. Create a directory on a client machine.**

**a. Log in to a client machine as superuser.**

You see a prompt that resembles the following:

*client-machine#*

**b. Create a directory on the client machine.**

*client-machine#* **mkdir /***dir*

**4. Mount the directory to the application on the primary cluster, and display the mounted directory.**

**a. Mount the directory to the application on the primary cluster.**

*client-machine#* **mount -o rw** *lhost-nfsrg-prim***:/global/***mountpoint* **/***dir*

**b. Display the mounted directory.**

*client-machine#* **ls /***dir*

**5. Mount the directory to the application on the secondary cluster, and display the mounted directory.**

**a. Unmount the directory to the application on the primary cluster.**

*client-machine#* **umount /***dir*

**b. Take the application resource group offline on the primary cluster.**

*nodeA#* **/usr/cluster/bin/scswitch -n -j** *nfs-rs*
*nodeA#* **/usr/cluster/bin/scswitch -n -j** *nfs-dg-rs*
*nodeA#* **/usr/cluster/bin/scswitch -n -j** *lhost-nfsrg-prim*
*nodeA#* **/usr/cluster/bin/scswitch -z -g** *nfs-rg* **-h ""**

c. **Change the primary cluster to logging mode.**

*nodeA#* **/usr/opt/SUNWesm/sbin/sndradm -n -l** *lhost-reprg-prim* **\**
**/dev/vx/rdsk/***devicegroup***/vol01 \**
**/dev/vx/rdsk/***devicegroup***/vol04** *lhost-reprg-sec* **\**
**/dev/vx/rdsk/***devicegroup***/vol01 \**
**/dev/vx/rdsk/***devicegroup***/vol04 ip sync**

When the data volume on the disk is written to, the bitmap file on the same disk is updated. No replication occurs.

d. **Ensure that the `PathPrefix` directory is available.**

*nodeC#* **mount | grep /global/etc**

e. **Bring the application resource group online on the secondary cluster.**

*nodeC#* **/usr/cluster/bin/scswitch -Z -g** *nfs-rg*

f. **Access the client machine as superuser.**

You see a prompt that resembles the following:

*client-machine#*

g. **Mount the directory that was created in Step 3 to the application on the secondary cluster.**

*client-machine#* **mount -o rw** *lhost-nfsrg-sec***:/global/***mountpoint* **/***dir*

h. **Display the mounted directory.**

*client-machine#* **ls /***dir*

6. **Ensure that the directory displayed in Step 4 is the same as that displayed in Step 5.**

7. **Return the application on the primary cluster to the mounted directory.**

a. **Take the application resource group offline on the secondary cluster.**

*nodeC#* **/usr/cluster/bin/scswitch -n -j** *nfs-rs*
*nodeC#* **/usr/cluster/bin/scswitch -n -j** *nfs-dg-rs*
*nodeC#* **/usr/cluster/bin/scswitch -n -j** *lhost-nfsrg-sec*
*nodeC#* **/usr/cluster/bin/scswitch -z -g** *nfs-rg* **-h ""**

b. **Ensure that the global volume is unmounted from the secondary cluster.**

*nodeC#* **umount /global/***mountpoint*

c. **Bring the application resource group online on the primary cluster.**

*nodeA#* **/usr/cluster/bin/scswitch -Z -g** *nfs-rg*

d. **Change the primary cluster to replicating mode.**

*nodeA#* **/usr/opt/SUNWesm/sbin/sndradm -n -u** *lhost-reprg-prim* **\**
**/dev/vx/rdsk/***devicegroup***/vol01 \**

Chapter 6 • Configuring Data Replication With Sun StorEdge Availability Suite Software    **281**

```
/dev/vx/rdsk/devicegroup/vol04 lhost-reprg-sec \
/dev/vx/rdsk/devicegroup/vol01 \
/dev/vx/rdsk/devicegroup/vol04 ip sync
```

When the primary volume is written to, the secondary volume is updated by
Sun StorEdge Availability Suite software.

**See Also**

---

# Example of How to Manage a Failover or Switchover

This section describes how to provoke a switchover and how the application is
transferred to the secondary cluster. After a switchover or failover, update the DNS
entries. For additional information, see "Guidelines for Managing a Failover or
Switchover" on page 257.

This section contains the following procedures:

- "How to Provoke a Switchover" on page 282
- "How to Update the DNS Entry" on page 283

## ▼ How to Provoke a Switchover

**Steps**   1. **Change the primary cluster to logging mode.**

*nodeA#* `/usr/opt/SUNWesm/sbin/sndradm -n -l` *lhost-reprg-prim* `\`
`/dev/vx/rdsk/`*devicegroup*`/vol01 \`
`/dev/vx/rdsk/`*devicegroup*`/vol04` *lhost-reprg-sec* `\`
`/dev/vx/rdsk/`*devicegroup*`/vol01 \`
`/dev/vx/rdsk/`*devicegroup*`/vol04 ip sync`

When the data volume on the disk is written to, the bitmap volume on the same
device group is updated. No replication occurs.

2. **Confirm that the primary cluster and the secondary cluster are in logging mode,
   with autosynchronization off.**

   a. **On `nodeA`, confirm the mode and setting:**

   *nodeA#* `/usr/opt/SUNWesm/sbin/sndradm -P`

   The output should resemble the following:

   `/dev/vx/rdsk/`*devicegroup*`/vol01 ->`
   *lhost-reprg-sec*`:/dev/vx/rdsk/`*devicegroup*`/vol01`

```
autosync:off, max q writes:4194304,max q fbas:16384,mode:sync,ctag:
```
*devicegroup*, state: logging

**b. On nodeC, confirm the mode and setting:**

*nodeC#* **/usr/opt/SUNWesm/sbin/sndradm -P**

The output should resemble the following:

```
/dev/vx/rdsk/devicegroup/vol01 <-
lhost-reprg-prim:/dev/vx/rdsk/devicegroup/vol01
autosync:off, max q writes:4194304,max q fbas:16384,mode:sync,ctag:
```
*devicegroup*, state: logging

For nodeA and nodeC, the state should be logging, and the active state of
autosynchronization should be off.

**3. Confirm that the secondary cluster is ready to take over from the primary cluster.**

*nodeC#* **/usr/sbin/fsck -y /dev/vx/rdsk/***devicegroup***/vol01**

**4. Switch over to the secondary cluster.**

*nodeC#* **scswitch -Z -g** *nfs-rg*

**Next Steps**  Go to "How to Update the DNS Entry" on page 283.


## ▼ How to Update the DNS Entry

For an illustration of how DNS maps a client to a cluster, see Figure 6–6.

**Before You Begin**  Ensure that you completed all steps in "How to Provoke a Switchover" on page 282.

**Steps**  **1. Start the nsupdate command.**
For information, see the nsupdate(1M) man page.

**2. Remove the current DNS mapping between the logical hostname of the application resource group and the cluster IP address, for both clusters.**

```
> update delete lhost-nfsrg-prim A
> update delete lhost-nfsrg-sec A
> update delete ipaddress1rev.in-addr.arpa ttl PTR lhost-nfsrg-prim
> update delete ipaddress2rev.in-addr.arpa ttl PTR lhost-nfsrg-sec
```

*ipaddress1rev*    The IP address of the primary cluster, in reverse order.

*ipaddress2rev*    The IP address of the secondary cluster, in reverse order.

*ttl*              The time to live, in seconds. A typical value is 3600.

**3. Create a new DNS mapping between the logical hostname of the application resource group and the cluster IP address, for both clusters.**

Map the primary logical hostname to the IP address of the secondary cluster and map the secondary logical hostname to the IP address of the primary cluster.

```
> update add lhost-nfsrg-prim  ttl  A  ipaddress2fwd
> update add lhost-nfsrg-sec  ttl  A  ipaddress1fwd
> update add ipaddress2rev.in-addr.arpa  ttl  PTR  lhost-nfsrg-prim
> update add ipaddress1rev.in-addr.arpa  ttl  PTR  lhost-nfsrg-sec
```

*ipaddress2fwd*    The IP address of the secondary cluster, in forward order.

*ipaddress1fwd*    The IP address of the primary cluster, in forward order.

*ipaddress2rev*    The IP address of the secondary cluster, in reverse order.

*ipaddress1rev*    The IP address of the primary cluster, in reverse order.

# Sun Cluster Installation and Configuration Worksheets

This appendix provides worksheets to plan various components of your cluster configuration and examples of completed worksheets for your reference. See "Installation and Configuration Worksheets" in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* for configuration worksheets for resources, resource types, and resource groups.

# Installation and Configuration Worksheets

If necessary, make additional copies of a worksheet to accommodate all the components in your cluster configuration. Follow planning guidelines in Chapter 1 to complete these worksheets. Then refer to your completed worksheets during cluster installation and configuration.

---

**Note –** The data used in the worksheet examples is intended as a guide only. The examples do not represent a complete configuration of a functional cluster.

---

The following table lists the planning worksheets and examples provided in this appendix, as well as the titles of sections in Chapter 1 that contain related planning guidelines.

**TABLE A–1** Cluster Installation Worksheets and Related Planning Guidelines

| Worksheet | Example | Section Titles of Related Planning Guidelines |
|---|---|---|
| "Local File System Layout Worksheet" on page 288 | "Example: Local File System Layout Worksheets, With and Without Mirrored Root" on page 289 | "System Disk Partitions" on page 18<br><br>"Guidelines for Mirroring the Root Disk" on page 42 |
| "Public Networks Worksheet" on page 290 | "Example: Public Networks Worksheet" on page 291 | "Public Networks" on page 24 |
| "Local Devices Worksheets" on page 292 | "Example: Local Devices Worksheets" on page 293 | --- |
| "Disk Device Group Configurations Worksheet" on page 294 | "Example: Disk Device Group Configurations Worksheet" on page 295 | "Disk Device Groups" on page 34<br><br>"Planning Volume Management" on page 35 |
| "Volume-Manager Configurations Worksheet" on page 296 | "Example: Volume-Manager Configurations Worksheet" on page 297 | "Planning Volume Management" on page 35<br><br>Your volume manager documentation |

**TABLE A–1** Cluster Installation Worksheets and Related Planning Guidelines      *(Continued)*

| Worksheet | Example | Section Titles of Related Planning Guidelines |
|---|---|---|
| "Metadevices Worksheet (Solstice DiskSuite or Solaris Volume Manager)" on page 298 | "Example: Metadevices Worksheet (Solstice DiskSuite or Solaris Volume Manager)" on page 299 | "Planning Volume Management" on page 35<br><br>*Solstice DiskSuite 4.2.1 Installation and Product Notes* or *Solaris Volume Manager Administration Guide* (Solaris 9 or Solaris 10) |

# Local File System Layout Worksheet

Node name: _____

**TABLE A–2** Local File Systems With Mirrored Root Worksheet

| Volume Name | Component | Component | File System | Size |
|---|---|---|---|---|
| | | | / | |
| | | | swap | |
| | | | /globaldevices | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

**TABLE A–3** Local File Systems With Nonmirrored Root Worksheet

| Device Name | File System | Size |
|---|---|---|
| | / | |
| | swap | |
| | /globaldevices | |
| | | |
| | | |
| | | |
| | | |

## Example: Local File System Layout Worksheets, With and Without Mirrored Root

Node name: **phys-schost-1**

**TABLE A–4** Example: Local File Systems With Mirrored Root Worksheet

| Volume Name | Component | Component | File System | Size |
|---|---|---|---|---|
| **d1** | **c0t0d0s0** | **c1t0d0s0** | / | **6.75 GB** |
| **d2** | **c0t0d0s1** | **c1t0d0s1** | swap | **750 MB** |
| **d3** | **c0t0d0s3** | **c1t0d0s3** | /globaldevices | **512 MB** |
| | | | | |
| | | | | |
| | | | | |
| **d7** | **c0t0d0s7** | **c1t0d0s7** | **SDS replica** | **20 MB** |

**TABLE A–5** Example: Local File Systems With Nonmirrored Root Worksheet

| Device Name | File System | Size |
|---|---|---|
| **c0t0d0s0** | / | **6.75 GB** |
| **c0t0d0s1** | swap | **750 MB** |
| **c0t0d0s3** | /globaldevices | **512 MB** |
| | | |
| | | |
| | | |
| **c0t0d0s7** | **SDS replica** | **20 MB** |

# Public Networks Worksheet

**TABLE A–6** Public Networks Worksheet

| Component | Name |
|---|---|
| **Node name** | |
| **Primary hostname** | |
| IP Network Multipathing group | |
| Adapter name | |
| Backup adapter(s) (optional) | |
| Network name | |
| **Secondary hostname** | |
| IP Network Multipathing group | |
| Adapter name | |
| Backup adapter(s) (optional) | |
| Network name | |
| **Secondary hostname** | |
| IP Network Multipathing group | |
| Adapter name | |
| Backup adapter(s) (optional) | |
| Network name | |
| **Secondary hostname** | |
| IP Network Multipathing group | |
| Adapter name | |
| Backup adapter(s) (optional) | |
| Network name | |

# Example: Public Networks Worksheet

TABLE A–7 Example: Public Networks Worksheet

| Component | Name |
|---|---|
| **Node name** | `phys-schost-1` |
| **Primary hostname** | `phys-schost-1` |
| IP Network Multipathing group | `ipmp0` |
| Adapter name | `qfe0` |
| Backup adapter(s) (optional) | `qfe4` |
| Network name | `net-85` |
| **Secondary hostname** | `phys-schost-1-86` |
| IP Network Multipathing group | `ipmp1` |
| Adapter name | `qfe1` |
| Backup adapter(s) (optional) | `qfe5` |
| Network name | `net-86` |
| **Secondary hostname** | |
| IP Network Multipathing group | |
| Adapter name | |
| Backup adapter(s) (optional) | |
| Network name | |
| **Secondary hostname** | |
| IP Network Multipathing group | |
| Adapter name | |
| Backup adapter(s) (optional) | |
| Network name | |

# Local Devices Worksheets

Node name:_____

**TABLE A–8** Local Disks Worksheet

| Local Disk Name | Size |
|---|---|
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

**TABLE A–9** Other Local Devices Worksheet

| Device Type | Name |
|---|---|
|  |  |
|  |  |
|  |  |
|  |  |

## Example: Local Devices Worksheets

Node name: **phys-schost-1**

**TABLE A–10** Example: Local Disks Worksheet

| Local Disk Name | Size |
|---|---|
| c0t0d0 | 2G |
| c0t1d0 | 2G |
| c1t0d0 | 2G |
| c1t1d0 | 2G |
| | |
| | |
| | |
| | |

**TABLE A–11** Example: Other Local Devices Worksheet

| Device Type | Name |
|---|---|
| tape | /dev/rmt/0 |
| | |
| | |
| | |

# Disk Device Group Configurations Worksheet

Volume manager (circle one):

Solstice DiskSuite | Solaris Volume Manager | VxVM

**TABLE A–12** Disk Device Groups Worksheet

| Disk Group/<br>Disk Set Name | Node Names<br>(indicate priority if ordered list) | Ordered priority?<br>(circle one) | Failback?<br>(circle one) |
|---|---|---|---|
| | | Yes | No | Yes | No |
| | | Yes | No | Yes | No |
| | | Yes | No | Yes | No |
| | | Yes | No | Yes | No |
| | | Yes | No | Yes | No |
| | | Yes | No | Yes | No |
| | | Yes | No | Yes | No |
| | | Yes | No | Yes | No |
| | | Yes | No | Yes | No |
| | | Yes | No | Yes | No |
| | | Yes | No | Yes | No |
| | | Yes | No | Yes | No |

# Example: Disk Device Group Configurations Worksheet

Volume manager (circle one):

Solstice DiskSuite

**TABLE A–13** Example: Disk Device Groups Configurations Worksheet

| Disk Group/<br>Disk Set Name | Node Names<br>(indicate priority if ordered list) | Ordered priority?<br>(circle one) | Failback?<br>(circle one) |
|---|---|---|---|
| **dg-schost-1** | **1) phys-schost-1,**<br>**2) phys-schost-2** | **Yes** | **Yes** |
| | | Yes \| No | Yes \| No |
| | | Yes \| No | Yes \| No |
| | | Yes \| No | Yes \| No |
| | | Yes \| No | Yes \| No |
| | | Yes \| No | Yes \| No |
| | | Yes \| No | Yes \| No |
| | | Yes \| No | Yes \| No |
| | | Yes \| No | Yes \| No |
| | | Yes \| No | Yes \| No |
| | | Yes \| No | Yes \| No |
| | | Yes \| No | Yes \| No |

# Volume-Manager Configurations Worksheet

Volume manager (circle one):

Solstice DiskSuite | Solaris Volume Manager | VxVM

**TABLE A–14** Volume-Manager Configurations Worksheet

| Name | Type | Component | Component |
|------|------|-----------|-----------|
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

# Example: Volume-Manager Configurations Worksheet

Volume manager (circle one):

Solstice DiskSuite

**TABLE A–15** Example: Volume-Manager Configurations Worksheet

| Name | Type | Component | Component |
|------|------|-----------|-----------|
| dg-schost-1/d0 | trans | dg-schost-1/d1 | dg-schost-1/d4 |
| dg-schost-1/d1 | mirror | c0t0d0s4 | c4t4d0s4 |
| dg-schost-1/d4 | mirror | c0t0d2s5 | d4t4d2s5 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

# Metadevices Worksheet (Solstice DiskSuite or Solaris Volume Manager)

**TABLE A–16** Metadevices Worksheet (Solstice DiskSuite or Solaris Volume Manager)

| File System | Metatrans | Metamirrors | | Submirrors | | Hot-Spare Pool | Physical Device | |
|---|---|---|---|---|---|---|---|---|
| | | (Data) | (Log) | (Data) | (Log) | | (Data) | (Log) |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |

# Example: Metadevices Worksheet (Solstice DiskSuite or Solaris Volume Manager)

**TABLE A–17** Example: Metadevices Worksheet (Solstice DiskSuite or Solaris Volume Manager)

| File System | Metatrans | Metamirrors | | Submirrors | | Hot-Spare Pool | Physical Device | |
|---|---|---|---|---|---|---|---|---|
| | | (Data) | (Log) | (Data) | (Log) | | (Data) | (Log) |
| /A | d10 | d11 | | d12, d13 | | hsp000 | c1t0d0s0, c2t0d1s0 | |
| | | | d14 | | d15 | hsp006 | | c1t0d1s6, c2t1d1s6 |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |

# Index