Sun Cluster Software Installation Guide for Solaris OS

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

Part No: 817–4229–10
April 2004, Revision A

040325@7940

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

Typographic Conventions

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

TABLE P–1 Typographic Conventions

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

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
Concepts	*Sun Cluster Concepts Guide for Solaris OS*
Overview	*Sun Cluster Overview for Solaris OS*
Hardware administration	*Sun Cluster Hardware Administration Manual for Solaris OS*
	Individual hardware administration guides
Software installation	*Sun Cluster Software Installation Guide for Solaris OS*
Data service administration	*Sun Cluster Data Service Planning and Administration Guide for Solaris OS*
	Individual data service guides
Data service development	*Sun Cluster Data Services Developer's Guide for Solaris OS*
System administration	*Sun Cluster System Administration Guide for Solaris OS*
Error messages	*Sun Cluster Error Messages Guide for Solaris OS*

Topic	Documentation
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`.

Accessing Sun Documentation Online

The docs.sun.com^{SM} Web site enables you to access Sun technical documentation online. You can browse the docs.sun.com archive or search for a specific book title or subject. The URL is `http://docs.sun.com`.

Ordering Sun Documentation

Sun Microsystems offers select product documentation in print. For a list of documents and how to order them, see "Buy printed documentation" at http://docs.sun.com.

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 operating environment (for example, Solaris 8)
- The release number of Sun Cluster (for example, Sun Cluster 3.1 4/04)

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

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

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

CHAPTER 1

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:

Where to Find 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.x Hardware Administration Manual* ■ Documentation that shipped with your server and storage devices
Plan cluster software installation.	■ Chapter 1 ■ "Installation and Configuration Worksheets" on page 236

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

Task	Instructions
Install a new cluster or add nodes to an existing cluster.	"Installing the Software" on page 40
Install and configure Solstice DiskSuite™/Solaris Volume Manager software.	■ "Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software" on page 116 ■ Solstice DiskSuite/Solaris Volume Manager documentation
SPARC: Install and configure VERITAS Volume Manager (VxVM) software.	■ "SPARC: Installing and Configuring VxVM Software" on page 149 ■ VxVM documentation
Configure cluster framework software and optionally install and configure the Sun Cluster module to Sun Management Center (which is available on SPARC based systems only).	"Configuring the Cluster" on page 100
Plan, install, and configure resource groups and data services.	*Sun Cluster Data Service 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 4/04 software.	■ Chapter 5 ■ "Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software" on page 116 or "SPARC: Installing and Configuring VxVM Software" on page 149 ■ Volume manager documentation

Planning the Solaris Operating Environment

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 operating environment 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 68 for details about the `scinstall` JumpStart installation method. See your Solaris installation documentation for details about standard Solaris installation methods.

Solaris Operating-Environment Feature Restrictions

The following Solaris operating-environment features are not supported in a Sun Cluster configuration:

- 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.

Solaris Software Group Considerations

Sun Cluster 3.1 4/04 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`, `SUNWrsmx`, `SUNWrsmo`, and `SUNWrsmox`). The RSMAPI software packages are included in only 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 "Local File System Layout Worksheet" on page 238.

When you install the Solaris operating environment, 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 operating environment. 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 operating environment.

See the following guidelines for additional partition planning information:

Guidelines for the Root (/) File System

As with any other system running the Solaris operating environment, 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/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 either Solstice DiskSuite/Solaris Volume Manager or VxVM 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/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/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/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/Solaris Volume Manager software to function properly. See your Solstice DiskSuite/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 Solaris operating environment 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/Solaris Volume Manager software or VxVM. If you use Solstice DiskSuite/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	Allocation (in Mbytes)	Description
0	/	6.75GB	Remaining free space on the disk after allocating space to slices 1 through 7. Used for Solaris operating environment software, 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 Solaris operating environment software. 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/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:

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 Release Notes for Solaris OS* or consult your Sun service provider.
- For general guidelines and procedures for applying patches, see "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 any naming services that are used. Also add these IP addresses to the local `/etc/inet/hosts` file on each cluster node after you install Solaris software.

- For more information about IP addresses, see *System Administration Guide, Volume 3* (Solaris 8) or *System Administration Guide: IP Services* (Solaris 9).
- For more information about test IP addresses to support IP Network Multipathing, see *IP Network Multipathing Administration Guide*.

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 ■ 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

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

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.

- For more information, see the *Sun Cluster Data Service 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.
- 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.
- 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, the `scinstall` utility configures a single-adapter IP Network Multipathing group for each public-network adapter. To modify these backup groups after installation, follow the procedures in "Deploying Network Multipathing" in *IP Network Multipathing Administration Guide* (Solaris 8) or"Administering Network Multipathing (Task)" in *System Administration Guide: IP Services* (Solaris 9).

See "IP Network Multipathing Groups" on page 27 for guidelines on planning public-network-adapter backup groups. For more information about public-network interfaces, see *Sun Cluster Concepts Guide for Solaris OS*.

Sun Cluster Configurable Components

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

Add this information to the appropriate configuration worksheet.

TABLE 1–4 Worksheets for Sun Cluster Configuration

Configuration Worksheet	Location
Table 2–2 or Table 2–3	"How to Configure Sun Cluster Software on All Nodes (`scinstall`)" on page 52
Table 2–6	"How to Install and Configure Sun Cluster Software (SunPlex Manager)" on page 62

TABLE 1–4 Worksheets for Sun Cluster Configuration *(Continued)*

Configuration Worksheet	Location
Table 2–7	"How to Install Solaris and Sun Cluster Software (JumpStart)" on page 68
Table 2–9	"How to Configure Sun Cluster Software on Additional Cluster Nodes (`scinstall`)" on page 83

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 operating environment. 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 node name is the same as the cluster 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 if the default network address is already in use elsewhere in the enterprise.

Note – After the installation utility (`scinstall`, SunPlex Manager, 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:

- Use zeroes for the last two octets of the address.
- Follow the guidelines in RFC 1597 for network address assignments.

  You can contact the InterNIC to obtain copies of RFCs. See "Planning Your TCP/IP Network" in *System Administration Guide, Volume 3* (Solaris 8) or "Planning Your TCP/IP Network (Task)" in *System Administration Guide: IP Services* (Solaris 9) for instructions.

If you specify a netmask other than the default, the netmask must minimally mask all bits that are given in the private-network address.

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 the following information for two cluster interconnects:

- **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. 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.

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 Firelink adapter, which requires the junction name `sw-rsm`*N*. The `scinstall` utility automatically uses this junction name after you specify a Sun Firelink 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.

You can configure additional private-network connections after the cluster is established by using the `scsetup`(1M) utility.

For more information about the cluster interconnect, see "Cluster Interconnect" in *Sun Cluster Overview for Solaris OS* and *Sun Cluster Concepts Guide for Solaris OS*.

IP Network Multipathing Groups

Add this planning information to the "Public Networks Worksheet" on page 240.

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.

Consider the following points when you plan your multipathing groups.

- Each public network adapter must belong to a multipathing group.

- For multipathing groups that contain two or more adapters, you must configure a test IP address for each adapter in the group. 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, do no change the value of `TRACK_INTERFACES_ONLY_WITH_GROUPS` from `yes` to `no`.
- The name of a multipathing group has no requirements or restrictions.

For more information about IP Network Multipathing, see "Deploying Network Multipathing" in *IP Network Multipathing Administration Guide* (Solaris 8) or"Administering Network Multipathing (Task)" in *System Administration Guide: IP Services* (Solaris 9). Also see "IP Network Multipathing Groups" in *Sun Cluster Overview for Solaris OS* and *Sun Cluster Concepts Guide for Solaris OS*.

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. You assign quorum devices by using the `scsetup`(1M) utility.

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

Consider the following points when you plan quorum devices.

- **Minimum –** A two-node cluster must have at least one shared disk assigned as a quorum 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 Devices" in *Sun Cluster Overview for Solaris OS* and *Sun Cluster Concepts Guide 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:

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.

Disk Device Groups

Add this planning information to the "Disk Device Group Configurations Worksheet" on page 244.

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 multiported disks and properly configured volume-manager devices as failover devices. Proper configuration of a volume-manager device includes multiported 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 disks as failover devices.
- **Mirroring –** You must mirror the disks to protect the data from disk failure. See "Mirroring Guidelines" on page 36 for additional guidelines. See "Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software" on page 116 or "SPARC: Installing and Configuring VxVM Software" on page 149 and your volume-manager documentation for instructions on 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: The following VxFS features are not supported in a Sun Cluster 3.1 configuration.
  - Quick I/O
  - Snapshots
  - Storage checkpoints
  - VxFS-specific mount options:
    - `convosync` (Convert O_SYNC)
    - `mincache`
    - `qlog`, `delaylog`, `tmplog`
  - VERITAS CFS requires VERITAS cluster feature & VCS

  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 configuration are supported by Sun Cluster 3.1 software. See VxFS documentation for details about VxFS options that are supported in a cluster configuration.
- **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.

- **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.

Planning Volume Management

Add this planning information to the "Disk Device Group Configurations Worksheet" on page 244 and the "Volume-Manager Configurations Worksheet" on page 246. For Solstice DiskSuite/Solaris Volume Manager, also add this planning information to the "Metadevices Worksheet (Solstice DiskSuite/Solaris Volume Manager)" on page 248.

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/Solaris Volume Manager software and VERITAS Volume Manager (VxVM) software that you install or use in the following ways.

TABLE 1–5 Supported Use of Volume Managers with Sun Cluster Software

Volume-Manager Software	Requirements
Solstice DiskSuite/Solaris Volume Manager	You must install Solstice DiskSuite/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.

TABLE 1–5 Supported Use of Volume Managers with Sun Cluster Software *(Continued)*

Volume-Manager Software	Requirements
SPARC: Both Solstice DiskSuite/Solaris Volume Manager and VxVM	If you install both volume managers on the same node, you must use Solstice DiskSuite/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/Solaris Volume Manager Software" on page 116 or "SPARC: Installing and Configuring VxVM Software" on page 149 for instructions on 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:

- **Mirrored multihost disks –** You must mirror all multihost disks across disk expansion units. See "Guidelines for Mirroring Multihost Disks" on page 37 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 disks.
- **Mirrored root –** Mirroring the root disk ensures high availability, but such mirroring is not required. See "Mirroring Guidelines" on page 36 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 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 Service Planning and Administration Guide for Solaris OS* for information about node lists.
- **Multiported disks –** You must connect, or port, all disks used to construct a device group within the cluster to all of the nodes that are configured in the node list for that device group. Solstice DiskSuite/Solaris Volume Manager software can automatically check for this connection at the time that disks are added to a diskset. 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/Solaris Volume Manager Software

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

- **Local metadevice names or volume names –** The name of each local Solstice DiskSuite metadevice or Solaris Volume Manager volume must be unique throughout the cluster. Also, the name cannot be the same as any device-ID name.
- **Dual-string mediators –** Each diskset configured with exactly two disk strings and mastered by exactly two nodes must have Solstice DiskSuite/Solaris Volume Manager mediators configured for the diskset. 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 diskset with exactly two nodes that act as mediator hosts.
  - You must use the same two nodes for all disksets that require mediators. Those two nodes must master those disksets.
  - Mediators cannot be configured for disksets 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 Volume Manager volumes used by each diskset 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.

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

You must modify the `nmd` and `md_nsets` fields as follows to support a Sun Cluster configuration:

– `md_nsets` – The `md_nsets` field defines the total number of disksets 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 disksets in the cluster plus one additional diskset. Solstice DiskSuite/Solaris Volume Manager software uses the additional diskset to manage the private disks on the local host. The private disks are those metadevices or volumes that are not in the local diskset.

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

– `nmd` – The `nmd` field defines the number of metadevices or volumes that are created for each diskset. Set the value of `nmd` to the predicted highest value of metadevice or volume name that is used by any one of the disksets in the cluster. For example, if a cluster uses 10 metadevices or volumes in its first 15 disksets, but 1000 metadevices or volumes in the 16th diskset, 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 diskset 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* or "System Files and Startup Files" in *Solaris Volume Manager Administration Guide* 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 –** Enclosure-Based Naming is a feature that was introduced in VxVM version 3.2. 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 –** You must create a default root disk group on each node. The 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.
- **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 –** Using 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.

File-System Logging

Logging is required for cluster 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 "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.

- 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 1–6 Supported File System Logging Matrix

Volume Manager	Supported File System Logging
Solstice DiskSuite/Solaris Volume Manager	Solaris UFS logging, Solstice DiskSuite trans-metadevice logging or Solaris Volume Manager transactional-volume logging, VxFS logging
SPARC: VERITAS Volume Manager	Solaris UFS logging, VxFS logging

Consider the following points when you choose between Solaris UFS logging and Solstice DiskSuite trans-metadevice logging/Solaris Volume Manager transactional-volume logging:

- Solaris Volume Manager transactional-volume logging (formerly Solstice DiskSuite trans-metadevice logging) is scheduled to be removed from the Solaris operating environment 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" on page 37
- "Guidelines for Mirroring the Root Disk" on page 37

Guidelines for Mirroring Multihost Disks

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

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 disk-expansion unit.
- **Disk space –** Mirroring doubles the amount of necessary disk space.
- **Three-way mirroring –** Solstice DiskSuite/Solaris Volume Manager software and VERITAS Volume Manager (VxVM) support three-way mirroring. However, Sun Cluster software requires only two-way mirroring.
- **Number of metadevices or volumes –** Under Solstice DiskSuite/Solaris Volume Manager software, mirrors consist of other Solstice DiskSuite metadevices or Solaris Volume Manager volumes such as concatenations or stripes. Large configurations might contain a large number of metadevices or volumes.
- **Differing disk sizes –** If you mirror to a disk 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 238.

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/Solaris Volume Manager Software" on page 116 or "SPARC: Installing and Configuring VxVM Software" on page 149 for instructions on 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/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/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/Solaris Volume Manager, 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/Solaris Volume Manager, 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.

CHAPTER 2

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.

The following information and procedures are in this chapter.

Installing the Software

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

Task Map: Installing the Software

The following task map lists the tasks that you perform to install software on multinode 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 42
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 44

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

Task	Instructions
3. Install the Solaris operating environment and Sun Cluster software. Choose one of the following methods:	
■ **Method 1** – (*New clusters only*) Install Solaris software. Install Sun Cluster software on all nodes. Then use the `scinstall` utility to establish the cluster.	1. "How to Install Solaris Software" on page 47 2. "How to Install Sun Cluster Software Packages" on page 51 3. "How to Configure Sun Cluster Software on All Nodes (`scinstall`)" on page 52
■ **Method 2** – (*New clusters only*) Install Solaris software. Then install SunPlex™ Manager and use it to install Sun Cluster software.	1. "How to Install Solaris Software" on page 47 2. "Using SunPlex Manager to Install Sun Cluster Software" on page 57
■ **Method 3** – (*New clusters or added nodes*) Install Solaris software and Sun Cluster software in one operation by using the `scinstall` utility's custom JumpStart option.	"How to Install Solaris and Sun Cluster Software (JumpStart)" on page 68
■ **Method 4** – (*New single-node clusters*) Install Solaris software and then install Sun Cluster software by using the `scinstall -iFo` command.	1. "How to Install Solaris Software" on page 47 2. "How to Install Sun Cluster Software on a Single-Node Cluster" on page 79
■ **Method 5** – (*Added nodes only*) Install Solaris software on the new nodes. Install Sun Cluster software on the new node. Then configure Sun Cluster software on the new node by using the `scinstall` utility.	1. "How to Install Solaris Software" on page 47 2. "How to Install Sun Cluster Software Packages" on page 51 3. "How to Configure Sun Cluster Software on Additional Cluster Nodes (`scinstall`)" on page 83
4. (*Optional*) SPARC: Install VERITAS File System software.	"SPARC: How to Install VERITAS File System Software" on page 88
5. Configure the name-service look-up order.	"How to Configure the Name-Service Switch" on page 88
6. Set up directory paths.	"How to Set Up the Root Environment" on page 90
7. Install data-service software packages.	"How to Install Data-Service Software Packages (`installer`)" on page 91 or "How to Install Data-Service Software Packages (`scinstall`)" on page 93
8. Assign quorum votes and remove the cluster from installation mode, if not already performed during Sun Cluster installation.	"How to Perform Postinstallation Setup and Configure Quorum Devices" on page 95
9. Validate the quorum configuration.	"How to Verify the Quorum Configuration and Installation Mode" on page 98
10. Install and configure volume-manager software:	

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

Task	Instructions
■ Install and configure Solstice DiskSuite/Solaris Volume Manager software.	■ "Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software" on page 116 ■ Solstice DiskSuite/Solaris Volume Manager documentation
■ SPARC: Install and configure VERITAS Volume Manager software.	■ "SPARC: Installing and Configuring VxVM Software" on page 149 ■ VERITAS Volume Manager documentation
11. Configure the cluster.	"Configuring the Cluster" on page 100

▼ How to Prepare for Cluster Software Installation

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

1. **Read the following manuals for information that can help you plan your cluster configuration and prepare your installation strategy.**
   - *Sun Cluster 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 Service Planning and Administration Guide for Solaris OS* – Planning guidelines and procedures to install and configure data services.
   - Documentation for all third-party software products.
2. **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 software
   - Solstice DiskSuite/Solaris Volume Manager software
   - SPARC: VERITAS Volume Manager
   - SPARC: Sun Management Center

- Third-party applications

3. **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 Parallel Fail Safe/Real Application Clusters Guard option of Oracle Parallel Server/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 Parallel Server/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 Service 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.

4. **Get all necessary patches for your cluster configuration.**

   See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 Release Notes* 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 not using SunPlex Manager, you can create a patch list file.**

      If you specify a patch list file, SunPlex Manager 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.**

5. **Do you intend to use Cluster Control Panel software to connect from an administrative console to your cluster nodes?**
   - If yes, go to "How to Install Cluster Control Panel Software on an Administrative Console" on page 44.
   - If no, go to one of the following procedures.
     - To install Sun Cluster software by using either the `scinstall`(1M) utility (text-based method) or SunPlex Manager (GUI-based method), go to "How to Install Solaris Software" on page 47 to first install Solaris software.
     - To install Solaris and Sun Cluster software in the same operation (JumpStart based method), go to "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 68.

▼ 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 launchpad for 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 operating environment 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 on how to install Sun Management Center software. See the *Sun Cluster Release Notes for Solaris OS* for additional information on how to install Sun Cluster documentation.

1. **Become superuser on the administrative console.**
2. **Ensure that a supported version of the Solaris operating environment and any Solaris patches are installed on the administrative console.**

   All platforms require at least the End User Solaris Software Group.
3. **Insert the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM into the CD-ROM drive of the administrative console.**

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

4. **From the `/cdrom/cdrom0` directory, 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) or `9` (for Solaris 9) .**

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

5. **Install the `SUNWccon` package.**

   ```
   # pkgadd -d . SUNWccon
   ```

6. **(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.

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

   If you do not install the documentation on your administrative console, you can still view HTML or PDF documentation directly from the CD-ROM.

   a. **Start the `pkgadd` utility in interactive mode.**

      ```
      # pkgadd -d .
      ```

   b. **Select the Documentation Navigation for Solaris 9 package, if it has not already been installed on the administrative console.**

   c. **Select the Sun Cluster documentation packages to install.**

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

      - Sun Cluster 3.1 4/04 Software Collection for Solaris OS (SPARC Platform Edition)
      - Sun Cluster 3.1 4/04 Software Collection for Solaris OS (x86 Platform Edition)
      - Sun Cluster 3.x Hardware Collection for Solaris OS (SPARC Platform Edition)
      - Sun Cluster 3.x Hardware Collection for Solaris OS (x86 Platform Edition)
      - Sun Cluster 3.1 4/04 Reference Collection for Solaris OS

   d. **Follow onscreen instructions to continue package installation.**

8. **Unload the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM from the CD-ROM drive.**

a. To ensure that the CD-ROM is not being used, change to a directory that does *not* reside on the CD-ROM.

b. Eject the CD-ROM.

```
# eject cdrom
```

9. 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.

10. 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.

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

- Add the `/opt/SUNWcluster/bin` directory to the `PATH`.
- Add the `/opt/SUNWcluster/man` directory to the `MANPATH`.
- If you installed the `SUNWscman` package, also add the `/usr/cluster/man` directory to the `MANPATH`.

12. 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.

13. **Is the Solaris operating environment already installed on each cluster node to meet Sun Cluster software requirements?**

    - If yes, go to "How to Install Sun Cluster Software Packages" on page 51.
    - If no, install, reconfigure, or reinstall the Solaris operating environment as needed to meet Sun Cluster software requirements, such as required partitioning. See "Planning the Solaris Operating Environment" on page 16 for information about Sun Cluster installation requirements for the Solaris operating environment.
      - To install only Solaris software, go to "How to Install Solaris Software" on page 47.
      - To install both Solaris and Sun Cluster software by using the `scinstall` JumpStart option, go to "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 68.

▼ How to Install Solaris Software

If you do not use the `scinstall`(1M) custom JumpStart installation method to install software, perform this task. Follow these procedures to install the Solaris operating environment on each node in the cluster.

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

If your nodes are already installed with the Solaris operating environment 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 Operating Environment" on page 16 for information about required root-disk partitioning and other Sun Cluster installation requirements.

1. **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.

2. **Ensure that your cluster configuration planning is complete.**

   See "How to Prepare for Cluster Software Installation" on page 42 for requirements and guidelines.

3. **Have available your completed "Local File System Layout Worksheet" on page 238.**

4. **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. You set up local hostname information in Step 15.**

   See "IP Addresses" on page 22 for planning guidelines. See your Solaris system-administrator documentation for information about using Solaris naming services.

5. **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, you can use the `cconsole`(1M) utility to display the individual console screens. The `cconsole` utility also opens a master window from which you can send your input to all individual console windows at the same time. Use the following command to start `cconsole`:

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

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

6. **Install the Solaris operating environment as instructed in your Solaris installation documentation.**

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

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.**

   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 Manager 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 Manager to install Solstice DiskSuite software (Solaris 8) or configure Solaris Volume Manager software (Solaris 9), also make this file system mounted on `/sds`.
- Create any other file-system partitions that you need, as described in "System Disk Partitions" on page 18.

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

---

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

7. Are you installing a new node to an existing cluster?

- If no, skip to Step 12.
- If yes, proceed to Step 8.

8. Have you added the new node to the cluster's authorized-node list?

- If yes, proceed to Step 9.
- If no, add the new node's name to the list of authorized cluster nodes.

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

```
# scsetup
```

The Main Menu is displayed.

b. Select `New nodes`.

c. Select `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.

e. Verify that the task has succeeded.

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

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

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

10. **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.

11. **Is VERITAS Volume Manager (VxVM) installed on any nodes that are already in the cluster?**

    - If no, proceed to Step 12.
    - If yes, ensure that the same `vxio` number is used on the VxVM-installed nodes. Also ensure that the `vxio` number is available for use on each of the nodes that do not have VxVM installed.

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

      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.

12. **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.**

Feature	Required Solaris Software Packages (shown in installation order)
RSMAPI SCI-PCI adapters (SPARC based clusters only)	`SUNWrsm SUNWrsmx SUNWrsmo SUNWrsmox`
SunPlex Manager	`SUNWapchr SUNWapchu`

13. **Install any hardware-related patches. Also download any needed firmware that is contained in the hardware patches.**

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

14. **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.

15. **Update the `/etc/inet/hosts` file on each node with all public hostnames and logical addresses for the cluster.**

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

16. **Do you intend to use dynamic reconfiguration on Sun Enterprise 10000 servers?**

    - If no, proceed to Step 17.
    - If yes, 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 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.

17. **Install Sun Cluster software packages.**

    Go to "How to Install Sun Cluster Software Packages" on page 51.

▼ How to Install Sun Cluster Software Packages

Perform this procedure to install the Sun Cluster software packages.

1. **Ensure that the Solaris operating environment 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 47 for more information about installing Solaris software to meet Sun Cluster software requirements.

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

3. **Install Sun Cluster framework software packages and Sun Java System data service packages.**

   Follow procedures in the *Sun Java Enterprise System Installation Guide*.

4. **Repeat Step 1 through Step 3 on each remaining cluster node to install.**

5. **Configure Sun Cluster software on the cluster nodes.**

- To establish a new cluster, go to "How to Configure Sun Cluster Software on All Nodes (`scinstall`)" on page 52.
- To add a new node to an existing cluster, go to "How to Configure Sun Cluster Software on Additional Cluster Nodes (`scinstall`)" on page 83.

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

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

1. **Ensure that the Solaris operating environment 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 47 for more information about installing Solaris software to meet Sun Cluster software requirements.

2. **Did you disable remote configuration during Sun Cluster software installation?**

   - If no, proceed to Step 3.
   - If yes, enable remote shell (`rsh`(1M)) or secure shell (`ssh`(1)) access for superuser to all cluster nodes.

3. **Do you intend to use the `scinstall` utility to install patches?**

   - If no, proceed to Step 4.
   - If yes, download patches to a patch directory.

     The `scinstall` command checks for patches in the `/var/cluster/patches` directory or the `/var/patches` directory. If neither of the directories exist, then no patches are added. If both directories exist, then only the patches in the `/var/cluster/patches` directory are added.

     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.

4. **Complete one of the following configuration worksheets:**

   - Table 2–2
   - Table 2–3

   See "Planning the Sun Cluster Environment" on page 21 for planning guidelines.

TABLE 2–2 Interactive `scinstall` Configuration Worksheet (Typical)

Component	Description/Example	Enter Answers Here
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?	
Cluster-Transport Adapters and Cables	What are the names of the two cluster-transport adapters that attach the node to the private interconnect?	
Validation Checks	Do you want to interrupt installation for `sccheck` errors?	Yes | No

For the Typical configuration of Sun Cluster software, `scinstall` automatically specifies the following 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`

TABLE 2–3 Interactive `scinstall` Configuration Worksheet (Custom)

Component	Description/Example	Enter Answers Here
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	Do you need to use DES authentication?	No | Yes
Network Address for the Cluster Transport	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	If this is a two-node cluster, does this cluster use transport junctions?	Yes | No

TABLE 2–3 Interactive `scinstall` Configuration Worksheet (Custom) *(Continued)*

Component	Description/Example	Enter Answers Here	
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	Node name (*the node from which you run scinstall*):		
	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?		
	Do you want to use autodiscovery to list the available adapters for the other nodes? If no, supply the following information for each additional node:	Yes | No	
	Node name:		
	Transport adapters:	*First*	*Second*
	Where does each transport adapter connect to (*a transport junction or another adapter*)? 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	
Global-Devices File System (*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?		
Validation Checks	Do you want to run the `sccheck` utility to validate the cluster?	Yes | No	

Note – 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 98. Then perform the procedures in "How to Install Sun Cluster Software Packages" on page 51 and in this procedure to reinstall the software and configure the node with the correct information.

---

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

6. **Start the `scinstall` utility.**

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

7. **Follow these guidelines to use the interactive `scinstall` utility:**

   - 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.

8. **From the `Main Menu`, choose `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) Print release information for this cluster node

        * ?) Help with menu options
        * q) Quit

      Option:  1
   ```

9. **From the `Install Menu`, choose `Install all nodes of a new cluster`.**

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

11. **Follow the menu prompts to supply your answers to Table 2–2 or Table 2–3, which you completed in Step 4.**

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.

12. SPARC: Do you intend to install VERITAS File System?

- If yes, go to "SPARC: How to Install VERITAS File System Software" on page 88.
- If no, set up the name-service look-up order. Go to "How to Configure the Name-Service Switch" on page 88.

Example – Configuring Sun Cluster Software on All Nodes

The following example shows the `scinstall` progress messages that are logged as `scinstall` completes configuration tasks on a two-node cluster. The cluster node names are `phys-schost-1` and `phys-schost-2`. The specified adapter names are `qfe2` and `hme2`.

```
Installation and Configuration

  Log file - /var/cluster/logs/install/scinstall.log.834

  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.

  Probing ..

  The following connections were discovered:

      phys-schost-1:qfe2  switch1  phys-schost-2:qfe2
      phys-schost-1:hme2  switch2  phys-schost-2:hme2

  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".

  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.834

Rebooting ...
```

Using SunPlex Manager to Install Sun Cluster Software

Note – To add a new node to an existing cluster, do not use SunPlex Manager. Instead, follow the procedures "How to Configure Sun Cluster Software on Additional Cluster Nodes (`scinstall`)" on page 83.

This section describes how to install SunPlex Manager. This section also describes how to use SunPlex Manager to install Sun Cluster software and to establish new cluster nodes. You can also use SunPlex Manager 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 Manager configures up to three metasets and associated metadevices. SunPlex Manager also creates and mounts cluster file systems for each metaset.
- (On Solaris 9 only) Solaris Volume Manager software – SunPlex Manager configures up to three Solaris Volume Manager volumes. SunPlex Manager 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 Manager installation requirements for these additional software products.

TABLE 2–4 Requirements to Use SunPlex Manager to Install Software

Software Package	Installation Requirements
Solstice DiskSuite/Solaris Volume Manager	A partition that uses `/sds` as the mount–point name. The partition must be at least 20 Mbytes in size.

TABLE 2–4 Requirements to Use SunPlex Manager to Install Software *(Continued)*

Software Package	Installation Requirements
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. ■ Solstice DiskSuite software installed, or Solaris Volume Manager software configured, by SunPlex Manager. ■ 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 the base hostnames of the cluster nodes. ■ A test IP address for each node of the cluster. SunPlex Manager 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. ■ Solstice DiskSuite software installed, or Solaris Volume Manager software configured, by SunPlex Manager. ■ 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 the base hostnames of the cluster nodes. ■ A test IP address for each node of the cluster. SunPlex Manager 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 Manager. The number of metasets and mount points that SunPlex Manager 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 Manager creates the `mirror-1` and `mirror-2` metasets. However, SunPlex Manager does not create the `mirror-3` metaset, because the node does not have enough shared disks to create a third metaset.

TABLE 2–5 Metasets Installed by SunPlex Manager

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 Manager still installs the Solstice DiskSuite packages. However, without sufficient shared disks, SunPlex Manager cannot configure the metasets, metadevices, or volumes. SunPlex Manager then cannot configure the cluster file systems that are needed to create instances of the data service.

Character-Set Limitations

SunPlex Manager 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 Manager server. The following characters are accepted by SunPlex Manager:

```
()+,-./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 Install SunPlex Manager Software

This procedure describes how to install SunPlex Manager software on your cluster.

Perform this procedure on each node of the cluster.

1. **Ensure that Solaris software and patches are installed on each node of the cluster.**

You must install Solaris software as described in "How to Install Solaris Software" on page 47. Or, if Solaris software is already installed on the node, you must ensure that the Solaris installation meets the requirements for Sun Cluster software. You must also ensure that the installation meets the requirements for any other software that you intend to install on the cluster.

2. **Review the requirements and guidelines in "Using SunPlex Manager to Install Sun Cluster Software" on page 57.**

3. **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, go to Step 4.
   - If you are using Internet Explorer, go to Step 5.

4. **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 6.
      - 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 6.**

5. **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 version 1.4.1 or a later version is shown, skip to Step 6.
- If an earlier version is shown, proceed to Step c.

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.

6. Become superuser on a cluster node.

7. Ensure that Apache software packages are installed on the node.

```
# pkginfo SUNWapchr SUNWapchu SUNWapchd
```

If necessary, install any missing Apache software packages by performing the following steps.

a. Insert the Solaris 8 or Solaris 9 Software 2 of 2 CD-ROM into the CD-ROM drive of 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.

b. Change to the `Product` directory.

- For Solaris 8, change to the `/cdrom/sol_8_sparc/Solaris_8/Product` directory.

```
# cd /cdrom/sol_8_sparc/Solaris_8/Product
```

- For Solaris 9, change to the `/cdrom/cdrom0/Solaris_9/Product` directory.

```
# cd /cdrom/cdrom0/Solaris_9/Product
```

c. Install the Apache software packages in the order that is shown in this step.

```
# pkgadd -d . SUNWapchr SUNWapchu SUNWapchd
```

d. Install any Apache software patches.

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

8. Install the SunPlex Manager software packages.

a. Insert the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM into the CD-ROM drive of the node.

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

b. From the `/cdrom/cdrom0` directory, 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) or `9` (for Solaris 9) .

The following example shows the path to the SPARC Solaris 8 version of Sun Cluster software.

```
# cd /cdrom/cdrom0/Solaris_sparc/Product/sun_cluster/Solaris_8/Packages
```

c. Install the SunPlex Manager software packages.

```
# pkgadd -d . SUNWscva SUNWscvr SUNWscvw
```

9. Repeat Step 6 through Step 8 on each node of the cluster.

10. Determine whether the `root` password is the same on every node of the cluster.

- If yes, proceed to Step 11.
- If no, set the `root` password to the same value on each node of the cluster. If necessary, also use the `chkey` command to update the RPC key pair. See the `chkey`(1) man page.

```
# passwd
    Enter new password
# chkey -p
```

To use the `root` password to access SunPlex Manager, the `root` password must be the same on all nodes in the cluster.

11. Use SunPlex Manager to configure Sun Cluster software.

Go to "How to Install and Configure Sun Cluster Software (SunPlex Manager)" on page 62.

▼ How to Install and Configure Sun Cluster Software (SunPlex Manager)

Note – To add a new node to an existing cluster, do not use SunPlex Manager. Instead, go to "How to Configure Sun Cluster Software on Additional Cluster Nodes (`scinstall`)" on page 83.

Perform this procedure to use SunPlex Manager to install and configure Sun Cluster software and 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) or to configure Solaris Volume Manager mirrored disksets (Solaris 9).

If you use SunPlex Manager to install Solstice DiskSuite software or to configure Solaris Volume Manager disksets, you can also install one or both of the following data services:

- Sun Cluster HA for NFS data service
- Scalable Sun Cluster HA for Apache data service

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.

1. **Ensure that the cluster configuration meets the requirements to use SunPlex Manager to install software.**

   See "Using SunPlex Manager to Install Sun Cluster Software" on page 57 for installation requirements and restrictions.

2. **Do you intend to install Sun Cluster HA for NFS or Sun Cluster HA for Apache?**

   - If no, proceed to Step 3.
   - If yes, ensure that the cluster configuration meets all applicable requirements. See "Using SunPlex Manager to Install Sun Cluster Software" on page 57.

3. **Ensure that SunPlex Manager software is installed on each node of the cluster.**

   See the installation procedures in "How to Install SunPlex Manager Software" on page 59.

4. **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.

5. **Do you intend to use the Remote Shared Memory Application Programming Interface (RSMAPI) or use SCI-PCI adapters for the interconnect transport?**

   - If no, proceed to Step 6.
   - If yes, install the additional packages from the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM that are required to support the RSMAPI or SCI-PCI adapters. SunPlex Manager does not automatically install these packages.

a. Determine which packages you must install.

The following table lists the Sun Cluster 3.1 4/04 packages that each feature requires and the order in which you must install each group of packages.

Feature	Additional Sun Cluster 3.1 4/04 Packages to Install
RSMAPI	`SUNWscrif`
SCI-PCI adapters	`SUNWsci SUNWscid SUNWscidx`

b. Use the following command to install the additional packages.

Replace *arch* with `sparc` or `x86` and replace *ver* with `8` (for Solaris 8) or `9` (for Solaris 9).

```
# cd Solaris_arch/Product/sun_cluster/Solaris_ver/Packages
# pkgadd -d . packages
```

6. Are there any patches that are required to support Sun Cluster or Solstice DiskSuite software?

- If yes, proceed to Step 7.
- If no, skip to Step 10.

7. Do you intend to use SunPlex Manager to install patches?

- If yes, proceed to Step 8.
- If no, manually install all patches that are required to support Sun Cluster or Solstice DiskSuite software *before* you use SunPlex Manager, then skip to Step 10.

8. Copy patches that are required for Sun Cluster or Solstice DiskSuite software into a single directory. This directory must reside on a file system that is available to each node.

a. Ensure that only one version of each patch is present in this patch directory.

If the patch directory contains multiple versions of the same patch, SunPlex Manager cannot determine the correct patch dependency order.

b. Ensure that the patches are uncompressed.

9. Complete the following installation worksheet.

TABLE 2–6 SunPlex Manager Installation and Configuration Worksheet

Component	Description/Example	Enter Answers Here
Cluster Name	What is the name of the cluster that you want to establish?	
	How many nodes are you installing in the cluster?	

TABLE 2–6 SunPlex Manager Installation and Configuration Worksheet *(Continued)*

Component	Description/Example	Enter Answers Here
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? ■ Solaris 9: Do you want to configure Solaris Volume Manager?	Yes | No
Sun Cluster HA for NFS *Requires Solstice DiskSuite or Solaris Volume Manager*	Do you want to install Sun Cluster HA for NFS? 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? *Supply one test IP address for each node in the cluster.*	
Sun Cluster HA for Apache (scalable) *Requires Solstice DiskSuite or Solaris Volume Manager*	Do you want to install scalable Sun Cluster HA for Apache? 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? *Supply one test IP address for each node in the cluster.*	
CD-ROM Paths	What is the path for each of the following components that you want to install? *The CD-ROM path must end with the directory that contains the* `.cdtoc` *file.*	
	Solstice DiskSuite:	
	Sun Cluster (base):	
	Sun Cluster data services (agents):	
	Patches:	
Validation Checks	Do you want to run the `sccheck` utility to validate the cluster?	Yes | No

---

Note – SunPlex Manager installation automatically specifies the default private-network address (`172.16.0.0`) and netmask (`255.255.0.0`). If you need to use a different address, do not use SunPlex Manager to install Sun Cluster software. Instead, follow procedures in "How to Install Sun Cluster Software Packages" on page 51 and in "How to Configure Sun Cluster Software on All Nodes (`scinstall`)" on page 52 to install and configure the cluster.

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 98. Then repeat this procedure to reinstall and configure the node with the correct information.

See "Planning the Solaris Operating Environment" on page 16 and "Planning the Sun Cluster Environment" on page 21 for planning guidelines. See the *Sun Cluster Data Service Planning and Administration Guide for Solaris OS* for data-service planning guidelines.

10. **Start SunPlex Manager.**

    a. **From the administrative console or any other machine outside the cluster, launch a browser.**

    b. **Disable the browser's Web proxy.**

       SunPlex Manager installation functionality is incompatible with Web proxies.

    c. **Ensure that disk caching and memory caching is enabled.**

       The disk cache and memory cache size must be greater than 0.

    d. **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 Manager displays the administration interface instead of the Sun Cluster Installation screen, Sun Cluster 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.

       ---

    e. **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 Manager.**

    If you meet all listed requirements, click Next to continue to the next screen.

13. **Follow the menu prompts to supply your answers to Table 2–6, which you completed in Step 9.**

14. **Click Begin Installation to start the installation process.**

    Follow these guidelines:

    - Do *not* close the browser window nor 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.

    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.

    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 Manager 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. **Use SunPlex Manager to verify the quorum assignments and to modify those assignments, if necessary.**

    For clusters with three or more nodes, the use of shared quorum devices is optional. SunPlex Manager 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.

16. **SPARC: Do you intend to install VERITAS File System?**

    - If yes, go to "SPARC: How to Install VERITAS File System Software" on page 88.
    - If no, set up the name-service look-up order. Go to "How to Configure the Name-Service Switch" on page 88.

▼ 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 and Sun Cluster software on all cluster nodes in a single operation and establishes the cluster. You can also use this procedure to add new nodes to an existing cluster.

1. **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 on how to set up the hardware.

2. **Ensure that your cluster configuration planning is complete.**

   See "How to Prepare for Cluster Software Installation" on page 42 for requirements and guidelines.

3. **Have available the following information:**

   - The Ethernet address of each cluster node
   - The following completed installation worksheet

TABLE 2–7 JumpStart Installation and Configuration Worksheet

Component	Description/Example	Enter Answers Here	
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	Do you need to use DES authentication?	No | Yes	
Network Address for the Cluster Transport	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	Does this cluster use transport junctions (*two-node clusters only*)?	Yes | No	
Cluster-Transport Junctions	If used, what are the names of the two transport junctions? Defaults: `switch1` and `switch2`	*First*	*Second*

TABLE 2–7 JumpStart Installation and Configuration Worksheet *(Continued)*

Component	Description/Example	Enter Answers Here	
Cluster-Transport Adapters and Cables	First node name:		
	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?		
	Do you want to use autodiscovery to list the available adapters for the other nodes? If no, supply the following information for each additional node:	Yes | No	
	Node name:		
	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	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?		
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	

See "Planning the Solaris Operating Environment" on page 16 and "Planning the Sun Cluster Environment" on page 21 for planning guidelines.

Note – 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 98. Then repeat this procedure to reinstall and configure the node with the correct information.

---

4. **Do you use a naming service?**

   - If no, proceed to Step 5. You set up the necessary hostname information in Step 15.
   - If yes, add the following information to any naming services that clients use to access cluster services:
     - Address-to-name mappings for all public hostnames and logical addresses
     - The IP address and hostname of the JumpStart server

   See "IP Addresses" on page 22 for planning guidelines. See your Solaris system-administrator documentation for information about using Solaris naming services.

5. **Are you installing a new node to an existing cluster?**

   - If no, proceed to Step 6.
   - If yes, run `scsetup`(1M) from another cluster node that is active. 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 Cluster Node to the Authorized Node List" in "Adding and Removing a Cluster Node" in *Sun Cluster System Administration Guide for Solaris OS*.

6. **Set up your JumpStart installation server for Solaris operating-environment installation.**

   - SPARC: As superuser, set up your JumpStart installation server.
   - x86: Set up and configure your JumpStart Dynamic Host Configuration Protocol (DHCP) server and your Solaris network for Preboot Execution Environment (PXE) installations.

     The *Sun Fire V60x and Sun Fire V65x Server Solaris Operating Environment Installation Guide* contains instructions about how to set up a DHCP server and Solaris network for PXE installations.

     ---

     x86 only – You must use a DHCP server and PXE to install Sun Cluster software with JumpStart over a network.

See "Preparing Custom JumpStart Installations (Tasks)" in *Solaris 9 Installation Guide* for instructions on how to set up a JumpStart installation server. See also the `setup_install_server`(1M) and `add_install_client`(1M) man pages.

When you set up the installation server, ensure that the following requirements are met:

- The installation server is on the same subnet as the cluster nodes but the server is not itself a cluster node.
- The installation server installs a release of the Solaris operating environment 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 install client that uses the custom JumpStart directory that you set up for Sun Cluster installation.

7. **Create a directory on the JumpStart installation server to hold your copy of the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM.**

   Skip this step if a directory already exists.

   In the following example, the `/export/suncluster` directory is created for this purpose.

   ```
   # mkdir -m 755 /export/suncluster
   ```

8. **Copy the Sun Cluster CD-ROM to the JumpStart installation server.**

   a. **Insert the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM into the CD-ROM drive on the JumpStart installation server.**

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

   b. **From the `/cdrom/cdrom0` directory, 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) or `9` (for Solaris 9) .**

      The following example uses the path to the SPARC Solaris 8 version of Sun Cluster software.

      ```
      # cd /cdrom/cdrom0/Solaris_sparc/Product/sun_cluster/Solaris_8/Tools
      ```

   c. **Copy the CD-ROM to a new directory on the JumpStart installation server.**

      The `scinstall` command creates the new installation directory when the command copies the CD-ROM files. The following example uses the installation directory name `/export/suncluster/sc31`.

      ```
      # ./scinstall -a /export/suncluster/sc31
      ```

d. Unload the CD-ROM from the CD-ROM drive.

i. To ensure that the CD-ROM is not being used, change to a directory that does *not* reside on the CD-ROM.

ii. Eject the CD-ROM.

```
# eject cdrom
```

e. Ensure that the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM image on the JumpStart installation server is NFS exported for reading by the JumpStart installation server.

See "Solaris NFS Environment" in *System Administration Guide, Volume 3* or "Managing Network File Systems (Overview)" in *System Administration Guide: Resource Management and Network Services* for more information about automatic file sharing. See also the `share`(1M) and `dfstab`(4) man pages.

9. 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) or `9` (for Solaris 9).

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

10. Follow these guidelines to use the interactive `scinstall` utility:

- 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.

11. From the `Main Menu`, choose `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) Print release information for this cluster node

      * ?) Help with menu options
```

```
  * q) Quit

Option:  2
```

Note – If the JumpStart option 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, do the following:

a. Exit the `scinstall` utility.
b. Repeat Step 6 through Step 8 to correct JumpStart setup.
c. Restart the `scinstall` utility.

---

12. Follow the menu prompts to supply your answers to Table 2–7, which you completed in Step 3.

13. If necessary, make adjustments to the default `class` file, or profile, that `scinstall` creates.

The `scinstall` command creates the following `autoscinstall.class` default `class` file in the *jumpstart-dir*`/autoscinstall.d/3.1` directory.

```
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
```

The default `class` file installs the End User Solaris Software Group (`SUNWCuser`). If your configuration has additional Solaris software requirements, change the `class` file accordingly. See "Solaris Software Group Considerations" on page 17 for more information.

If you installed the End User Solaris Software Group, add to the default `class` file any additional Solaris software packages that you might need. The following table lists Solaris packages that are required to support some Sun Cluster functionality.

TABLE 2–8 Solaris Software Package Dependencies

Feature	Required Solaris Software Packages (shown in installation order)
RSMAPI, SCI-PCI adapters (SPARC based clusters only)	`SUNWrsm SUNWrsmx SUNWrsmo SUNWrsmox`
SunPlex Manager	`SUNWapchr SUNWapchu`

You can change the default `class` file in one of the following ways:

- Edit the `autoscinstall.class` file directly. These changes are applied to all nodes in all clusters that use this custom JumpStart directory.
- Update the `rules` file to point to other profiles, then run the `check` utility to validate the `rules` file.

As long as the Solaris operating-environment installation profile meets minimum Sun Cluster file-system allocation requirements, Sun Cluster software places no restrictions on other changes to the installation profile. See "System Disk Partitions" on page 18 for partitioning guidelines and requirements to support Sun Cluster software.

For more information about JumpStart profiles, see "Preparing Custom JumpStart Installations (Tasks)" in *Solaris 9 Installation Guide* or "Preparing Custom JumpStart Installations (Tasks)" in *Solaris 9 Installation Guide*.

14. Set up Solaris patch directories.

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 on 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.

15. Set up files to contain the necessary hostname information locally on each node.

a. On the JumpStart installation server, create files that are named *jumpstart-dir*`/autoscinstall.d/nodes/`*node*`/archive/etc/inet/hosts`.

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

b. Add the following entries into each file.

- IP address and hostname of the NFS server that holds a copy of the Sun Cluster CD-ROM image. The NFS server could be the JumpStart installation server or another machine.
- IP address and hostname of each node in the cluster.

16. **Do you intend to use the Remote Shared Memory Application Programming Interface (RSMAPI) or use SCI-PCI adapters for the interconnect transport?**

    - If no, proceed to Step 17 if you intend to add your own postinstallation finish script. Otherwise, skip to Step 18.
    - If yes, follow instructions in Step 17 to set up a postinstallation finish script to install the following additional packages. Install the appropriate packages from the `Solaris_`*arch*`/Product/sun_cluster/Solaris_`*ver*`/Packages` directory of the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM in the order that is given in the following table.

    Note – In the CD-ROM path, replace *arch* with `sparc` or `x86` and replace *ver* with `8` (for Solaris 8) or `9` (for Solaris 9).

Feature	Additional Sun Cluster 3.1 4/04 Packages to Install
RSMAPI	`SUNWscrif`
SCI-PCI adapters	`SUNWsci SUNWscid SUNWscidx`

17. **(Optional) Add your own postinstallation finish script.**

    Note – If you intend to use the Remote Shared Memory Application Programming Interface (RSMAPI) or use SCI-PCI adapters for the interconnect transport, you must modify the finish script to install the Sun Cluster `SUNWscrif` software package. This package is not automatically installed by `scinstall`.

    You can add your own finish script, which is run after the standard finish script installed by the `scinstall` command. See "Preparing Custom JumpStart Installations" in *Solaris 8 Advanced Installation Guide* or "Preparing Custom JumpStart Installations (Tasks)" in *Solaris 9 Installation Guide* for information about creating a JumpStart finish script.

    a. **Name your finish script `finish`.**

    b. **Copy your finish script to the** *jumpstart-dir*`/autoscinstall.d/nodes/`*node* **directory.**

       Create one directory for each node in the cluster. Alternately, use this naming convention to create symbolic links to a shared finish script.

18. **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, you can use the `cconsole`(1M) utility to display the individual console screens. The `cconsole` utility also opens a master window from which you can send your input to all individual console windows at the same time. Use the following command to start `cconsole`:

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

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

19. Shut down each node.

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

20. 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 second 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 DHCP 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 software and Sun Cluster software on each node.

Note – 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 107 for information on how to suppress these messages under otherwise normal cluster conditions.

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.

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 98. Then repeat this procedure to reinstall and configure the node with the correct information.

21. Are you installing a new node to an existing cluster?

- If no, proceed to Step 22.
- If yes, 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 Step 24.

c. Is VERITAS Volume Manager (VxVM) installed on any nodes that are already in the cluster?

- If no, proceed to Step 22.
- If yes, ensure that the same `vxio` number is used on each of the VxVM-installed nodes. Also ensure that the `vxio` number is available for use on each of the nodes that do not have VxVM installed.

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

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.

22. (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.

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. Did you add a new node to an existing cluster or install Sun Cluster software patches that require you to reboot the entire cluster, or both?

- If no, reboot the individual node if any patches that you installed require a node reboot. Also reboot if any other changes that you made require a reboot to become active, then proceed to Step 25.
- If yes, perform a reconfiguration reboot of the cluster as instructed in the following steps.

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.

---

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
```

Note – 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 you run the `scsetup`(1M) command, during the procedure "How to Perform Postinstallation Setup and Configure Quorum Devices" on page 95.

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. SPARC: Do you intend to install VERITAS File System?

- If yes, go to "SPARC: How to Install VERITAS File System Software" on page 88.
- If no, set up the name-service look-up order. Go to "How to Configure the Name-Service Switch" on page 88.

How to Install Sun Cluster Software on a Single-Node Cluster

Perform this task to install Sun Cluster software and establish the cluster on a single node by using the `scinstall` command. See the `scinstall`(1M) man page for details.

Note – You cannot use SunPlex Manager or the interactive form of the `scinstall` utility to install Sun Cluster software on a single-node cluster.

The `scinstall -iFo` command establishes the following defaults during installation:

- Creates the required device IDs
- Uses the default `/globaldevices` file system, unless the global-devices file system is specified with the `-G` option

- Establishes a default cluster name, which is the same as the name of the node being installed, unless the *clustername* is specified with the `-C` option

Some steps that are required for multinode cluster installations are not necessary for single-node cluster installations. When you install a single-node cluster, you do not need to perform the following steps:

- You do not need to configure quorum.
- You do not need to configure interconnect adapters or junctions.

Tip – If you anticipate eventually adding a second node to your cluster, you can configure the transport interconnect during initial cluster installation. The transport interconnect is then available for later use. See the `scinstall`(1M) man page for details.

You can later expand a single-node cluster into a multinode cluster by following the appropriate procedures provided in "How to Configure Sun Cluster Software on Additional Cluster Nodes (`scinstall`)" on page 83.

---

1. **Ensure that the Solaris operating environment 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 47 for more information about installing Solaris software to meet Sun Cluster software requirements.

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

3. **Insert the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM into the CD-ROM drive of the node to install and configure.**

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

4. **From the `/cdrom/cdrom0` directory, 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) or `9` (for Solaris 9) .**

   The following example uses the path to the SPARC Solaris 8 version of Sun Cluster software.

   ```
   # cd /cdrom/cdrom0/Solaris_sparc/Product/sun_cluster/Solaris_8/Tools
   ```

5. **Install the Sun Cluster software and necessary patches by using the `scinstall` command.**

   ```
   # ./scinstall -iFo [-M patchdir=dirname]
   ```

`-i`
Specifies the install form of the `scinstall` command. The `scinstall` command installs Sun Cluster software and initializes the node as a new cluster.

`-F`
Establishes the node as the first node in a new cluster. All `-F` options can be used when installing a single-node cluster.

`-o`
Specifies that only one node is being installed for a single-node cluster. The `-o` option is only legal when used with both the `-i` and the `-F` forms of the command. When the `-o` option is used, cluster installation mode is preset to the disabled state.

`-M patchdir=`*dirname*`[,patchlistfile=`*filename*`]`
Specifies the path to patch information so that the specified patches can be installed by using the `scinstall` command. If you do not specify a patch-list file, the `scinstall` command installs all the patches in the directory *dirname*. This includes tarred, jarred, and zipped patches.

The `-M` option is not required with the `scinstall -iFo` command. The `-M` option is shown in this procedure because the use of this option is the most efficient method of installing patches during a single-node cluster installation. However, you can use any method that you prefer to install patches.

6. **Reboot the node.**

   This reboot after Sun Cluster software installation establishes the node as the cluster.

7. **(Optional) Change the cluster name.**

   A single-node cluster is created with the same name as the cluster node. If you prefer, you can change the cluster name. Use either the `scsetup` utility or use the following `scconf` command:

   ```
   # /usr/cluster/bin/scconf -c -C cluster=newclustername
   ```

8. **Verify the installation by using the `scstat` command.**

   ```
   # /usr/cluster/bin/scstat -n
   ```

   See the `scstat`(1M) man page for details.

9. **Ensure that cluster installation mode is disabled.**

   ```
   # /usr/cluster/bin/scconf -pv | grep "install mode"
   ```

10. **SPARC: Do you intend to install VERITAS File System?**

    - If yes, go to "SPARC: How to Install VERITAS File System Software" on page 88.
    - If no, set up the name-service look-up order. Go to "How to Configure the Name-Service Switch" on page 88.

Example—Installing Sun Cluster Software on a Single-Node Cluster

The following example shows how to use the `scinstall` and `scstat` commands to install and verify a single-node cluster. The example includes installation of all patches. See the `scinstall`(1M) and `scstat`(1M) man pages for details.

```
# scinstall -iFo -M patchdir=/var/cluster/patches

Checking device to use for global devices file system ... done
** Installing SunCluster 3.1 framework **
...
Installing patches ... done

Initializing cluster name to "phys-schost-1" ... done
Initializing authentication options ... done

Setting the node ID for "phys-schost-1" ... done (id=1)

Checking for global devices global file system ... done
Updating vfstab ... done

Verifying that "cluster" is set for "hosts" in nsswitch.conf ... done
Adding the "cluster" switch to "hosts" in nsswitch.conf ... done

Verifying that "cluster" is set for "netmasks" in nsswitch.conf ... done
Adding the "cluster" switch to "netmasks" in nsswitch.conf ... done

Verifying that power management is NOT configured ... done

Ensure that the EEPROM parameter "local-mac-address?" is set to "true" ... done

Ensure network routing is disabled ... done

Please reboot this machine.

# reboot
# scstat -n
-- Cluster Nodes --

                    Node name           Status
                    ---------           ------
  Cluster node:     phys-schost-1       Online
# scconf -pv | grep "install mode"
Cluster install mode:                   disabled
```

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

Perform this procedure to add new nodes to an existing cluster.

1. **Install the hardware on the new cluster node.**

   a. **Install the host adapter on the new node and verify that any existing cluster interconnects can support the new node.**

      See the *Sun Cluster Hardware Administration Manual for Solaris OS*.

   b. **Install any additional storage.**

      See the appropriate manual from the Sun Cluster 3.x Hardware Administration Collection.

2. **Ensure that the Solaris operating environment 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 47 for more information about installing Solaris software to meet Sun Cluster software requirements.

3. **Ensure that Sun Cluster software packages are installed on the node.**

   See "How to Install Sun Cluster Software Packages" on page 51.

4. **Complete the following configuration worksheet.**

TABLE 2–9 Added Node Configuration Worksheet

Component	Description/Example	Enter Answers Here
Software Patch Installation	Do you want `scinstall` to install patches for you?	Yes | No
	If yes, what is the patch directory?	
	Do you want to use a patch list?	Yes | No
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

TABLE 2–9 Added Node Configuration Worksheet *(Continued)*

Component	Description/Example	Enter Answers Here	
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?		
	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	

See "Planning the Solaris Operating Environment" on page 16 and "Planning the Sun Cluster Environment" on page 21 for planning guidelines.

5. **Are you adding this node to a single-node cluster?**
   - If yes, go to Step 6.
   - If no, skip to Step 8.

6. **From the existing cluster node, determine whether two cluster interconnects already exist.**

   You must have at least two cables or two adapters configured.

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

   - If the output shows configuration information for two cables or for two adapters, skip to Step 8.
   - If the output shows no configuration information for either cables or adapters, or shows configuration information for only one cable or adapter, proceed to Step 7.

7. **Configure new cluster interconnects.**

   a. **On the existing cluster node, start the `scsetup`(1M) utility.**

   ```
   # scsetup
   ```

   The Main Menu displays.

b. Select `Cluster interconnect`.

c. Select `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.

8. 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. Select `New nodes`.

c. Select `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.

e. Verify that the task has succeeded.

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

f. Quit the `scsetup` utility.

9. Become superuser on the cluster node to configure.

10. Start the `scinstall` utility.

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

11. Follow these guidelines to use the interactive `scinstall` utility:

- 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.

12. From the `Main Menu`, choose `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) Print release information for this cluster node

    * ?) Help with menu options
    * q) Quit

  Option:  1
```

13. From the `Install Menu`, choose `Add this machine as a node in an existing cluster`.

14. Follow the menu prompts to supply your answers to Table 2–9, which you completed in Step 4.

The `scinstall` utility configures the node and boots the node into the cluster.

15. Repeat this procedure on any other node to add to the cluster until all additional nodes are fully configured.

16. From an active cluster member, prevent any other nodes from joining the cluster.

```
# /usr/cluster/bin/scconf -a -T node=.
```

`-a`	Add
`-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 Cluster Node to the Authorized Node List" in "Adding and Removing a Cluster Node" in *Sun Cluster System Administration Guide for Solaris OS* for procedures.

17. SPARC: Do you intend to install VERITAS File System?

- If yes, go to "SPARC: How to Install VERITAS File System Software" on page 88.
- If no, set up the name-service look-up order. Go to "How to Configure the Name-Service Switch" on page 88.

Example – Configuring Sun Cluster Software on an Additional Node

The following example shows the `scinstall` command executed and the messages that the utility logs as `scinstall` completes configuration tasks on the node `phys-schost-3`. The sponsoring node is `phys-schost-1`.

```
 >>> Confirmation <<<

    Your responses indicate the following options to scinstall:

      scinstall -ik \
           -C sc-cluster \
           -N phys-schost-1 \
           -A trtype=dlpi,name=hme1 -A trtype=dlpi,name=hme3 \
           -m endpoint=:hme1,endpoint=switch1 \
           -m endpoint=:hme3,endpoint=switch2

    Are these the options you want to use (yes/no) [yes]?

    Do you want to continue with the install (yes/no) [yes]?

Checking device to use for global devices file system ... done

Adding node "phys-schost-3" to the cluster configuration ... done
Adding adapter "hme1" to the cluster configuration ... done
Adding adapter "hme3" 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
Setting the node ID for "phys-schost-3" ... done (id=3)

Verifying the major number for the "did" driver with "phys-schost-1" ...done

Checking for global devices global file system ... done
Updating vfstab ... done

Verifying that NTP is configured ... done
Installing a default NTP configuration ... done
Please complete the NTP configuration after scinstall has finished.

Verifying that "cluster" is set for "hosts" in nsswitch.conf ... done
Adding the "cluster" switch to "hosts" in nsswitch.conf ... done

Verifying that "cluster" is set for "netmasks" in nsswitch.conf ... done
Adding the "cluster" switch to "netmasks" in nsswitch.conf ... done

Verifying that power management is NOT configured ... done
Unconfiguring power management ... done
/etc/power.conf has been renamed to /etc/power.conf.61501001054
Power management is incompatible with the HA goals of the cluster.
Please do not attempt to re-configure power management.
```

```
Ensure that the EEPROM parameter "local-mac-address?" is set to "true" ...done
Ensure network routing is disabled ... done
Network routing has been disabled on this node by creating /etc/notrouter.
Having a cluster node act as a router is not supported by Sun Cluster.
Please do not re-enable network routing.

Log file - /var/cluster/logs/install/scinstall.log.9853


Rebooting ...
```

▼ SPARC: How to Install VERITAS File System Software

Perform this procedure on each node of the cluster.

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 Release Notes* for the location of patches and installation instructions.

3. **In the `/etc/system` file on each node, set the value for the `rpcmod:svc_default_stksize` variable to `0x8000` and set the value of the `lwp_default_stksize` variable to `0x6000`.**

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

   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.

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

4. **Set up the name-service look-up order.**

   Go to "How to Configure the Name-Service Switch" on page 88.

▼ How to Configure the Name-Service Switch

Perform this task on each node in the cluster.

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

2. **Edit the `/etc/nsswitch.conf` file.**

   a. **Verify that `cluster` is the first source look-up for the `hosts` and `netmasks` database entries.**

      This order is necessary for Sun Cluster software to function properly. The `scinstall`(1M) command adds `cluster` to these entries during installation.

   b. **(Optional) To increase availability to data services if the naming service becomes unavailable, change the look-up order of the following entries:**

      - For the `hosts` and `netmasks` database entries, follow `cluster` with `files`.
      - For Sun Cluster HA for NFS, also insert `[SUCCESS=return]` after `cluster files` and before name services.

        ```
        hosts:      cluster files [SUCCESS=return] nis
        ```

        This look-up order ensures that, if the node resolves a name locally, the node does not contact the listed name service(s). Instead, the node returns success immediately.
      - For all other database entries, place `files` first in the look-up order.
      - If the `[NOTFOUND=return]` criterion becomes the last item of an entry after you modify the lookup order, the criterion is no longer necessary. You can either delete the `[NOTFOUND=return]` criterion from the entry or leave the criterion in the entry. A `[NOTFOUND=return]` criterion at the end of an entry is ignored.

   c. **Make any other changes that are required by specific data services.**

      See each manual for the data services that you installed.

   The following example shows partial contents of an `/etc/nsswitch.conf` file. The look-up order for the `hosts` and `netmasks` database entries is first `cluster`, then `files`. The look-up order for other entries begins with `files`. The `[NOTFOUND=return]` criterion is removed from the entries.

   ```
   # vi /etc/nsswitch.conf
   ...
   passwd:     files nis
   group:      files nis
   ...
   hosts:      cluster files nis
   ...
   netmasks:   cluster files nis
   ...
   ```

   See the `nsswitch.conf`(4) man page for more information about `nsswitch.conf` file entries.

3. **Set up your root user's environment.**

   Go to "How to Set Up the Root Environment" on page 90.

▼ How to Set Up the Root Environment

Perform this procedure on each node in the cluster.

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 "Customizing a User's Work Environment" in *System Administration Guide: Basic Administration* (Solaris 9) for more information.

---

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`.**

   SPARC: If you are using VERITAS Volume Manager and VERITAS File System, also include the following paths that apply to your configuration:

Software Product	`PATH`
VERITAS Volume Manager (VxVM)	`/etc/vx/bin`
VxVM 3.2 GUI	`/opt/VRTSvmsa/bin`
VxVM 3.5 GUI	`/opt/VRTSob/bin`
VERITAS File System (VxFS)	`/opt/VRTSvxfs/sbin`, `/usr/lib/fs/vxfs/bin`, and `/etc/fs/vxfs`

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

   If you are using Solstice DiskSuite/Solaris Volume Manager, VERITAS Volume Manager, or VERITAS File System, also include the paths that apply to your configuration:

Software Product	`MANPATH`
Solstice DiskSuite/Solaris Volume Manager	`/usr/share/man`
SPARC: VxVM	`/opt/VRTS/man`
SPARC: VxVM GUI	`/opt/VRTSvmsa/man`

Software Product	MANPATH
SPARC: VxFS	/opt/VRTS/man

3. **(Optional) For ease of administration, set the same root password on each node, if you have not already done so.**
4. **Repeat Step 1 through Step 3 on each remaining cluster node.**
5. **Install data-service software packages.**
   - To use the `scinstall` utility, go to "How to Install Data-Service Software Packages (`scinstall`)" on page 93.
   - To use the `installer` program to install data services from the Sun Cluster 3.1 4/04 Data Services release, go to "How to Install Data-Service Software Packages (`installer`)" on page 91.

▼ How to Install Data-Service Software Packages (`installer`)

If you install data services from the Sun Cluster 3.1 4/04 release, you can use the `installer` program to install the packages. To install data services from an earlier release, follow the procedures in "How to Install Data-Service Software Packages (`scinstall`)" on page 93.

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.

1. **Become superuser on a cluster node.**
2. **(Optional) If you intend to use the `installer` program with a GUI, ensure that the `DISPLAY` environment variable is set.**
3. **Load the Java Enterprise System Accessory CD 3 CD-ROM into the CD-ROM drive.**

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

   ```
   # cd /cdrom/cdrom0
   ```
5. **Start the `installer` program.**

   ```
   # ./installer
   ```

6. **When you are prompted, select the type of installation.**
   - 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.
7. **When you are prompted, select the locale to install.**
   - To install only the C locale, select Typical.
   - To install other locales, select Custom.
8. **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.
9. **Exit the `installer` program.**
10. **Unload the CD–ROM from the CD-ROM drive.**
    a. **To ensure that the CD-ROM is not being used, change to a directory that does *not* reside on the CD-ROM.**
    b. **Eject the CD-ROM.**

       ```
       # eject cdrom
       ```
11. **Repeat Step 1 through Step 10 on each remaining cluster node.**
12. **Install any Sun Cluster data-service patches.**

    See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 Release Notes* 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. **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 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. Cluster nodes remain in installation mode until the first time that you run the `scsetup`(1M) command, during the procedure "How to Perform Postinstallation Setup and Configure Quorum Devices" on page 95.

---

13. Is this a single-node cluster?

- If no, perform postinstallation setup and assign quorum votes.

  Go to "How to Perform Postinstallation Setup and Configure Quorum Devices" on page 95.
- If yes, cluster installation is complete. You are now ready to install volume management software and to configure the cluster.
  - To install Solstice DiskSuite software or configure Solaris Volume Manager software, go to "Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software" on page 116.
  - SPARC: To install VERITAS Volume Manager software, go to "SPARC: Installing and Configuring VxVM Software" on page 149.

▼ How to Install Data-Service Software Packages (`scinstall`)

Perform this task on each cluster node to install data services. If you install data services from the Sun Cluster 3.1 4/04 release, you can instead use the `installer` program to install the packages. See "How to Install Data-Service Software Packages (`installer`)" on page 91.

Note – You do not need to perform this procedure if you used SunPlex Manager to install Sun Cluster HA for NFS or Sun Cluster HA for Apache or both and if you do not intend to install any other data services. Instead, go to "How to Perform Postinstallation Setup and Configure Quorum Devices" on page 95.

---

1. **Become superuser on a cluster node.**
2. **Load the Java Enterprise System Accessory CD 3 CD-ROM into the CD-ROM drive on the node.**

   If the volume management daemon `vold`(1M) is running and 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. **Follow these guidelines to use the interactive `scinstall` utility:**

- 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.

6. **From the `Main Menu`, choose `Add support for new data services to this cluster node`.**

7. **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.

8. **After the data services are installed, quit the `scinstall` utility.**

9. **Unload the Java Enterprise System Accessory CD 3 CD-ROM from the CD-ROM drive.**

   a. **To ensure that the CD-ROM is not being used, change to a directory that does *not* reside on the CD-ROM.**

   b. **Eject the CD-ROM.**

      ```
      # eject cdrom
      ```

10. **Repeat Step 1 through Step 9 on each cluster node where you are installing data services.**

11. **Install any Sun Cluster data-service patches.**

    See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 Release Notes* 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. **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 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.

Cluster nodes remain in installation mode until the first time that you run the `scsetup`(1M) command, during the procedure "How to Perform Postinstallation Setup and Configure Quorum Devices" on page 95.

12. Is this a single-node cluster?

- If no, perform postinstallation setup and assign quorum votes.

  Go to "How to Perform Postinstallation Setup and Configure Quorum Devices" on page 95.
- If yes, cluster installation is complete. You are now ready to install volume management software and to configure the cluster.
  - To install Solstice DiskSuite software or configure Solaris Volume Manager software, go to "Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software" on page 116.
  - SPARC: To install VERITAS Volume Manager software, go to "SPARC: Installing and Configuring VxVM Software" on page 149.

▼ How to Perform Postinstallation Setup and Configure Quorum Devices

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.

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 Manager to install the cluster. SunPlex Manager 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 98.

1. **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
   ```

2. **On each node, verify device connectivity to the cluster nodes.**

   Run the `scdidadm`(1M) command to 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
   ```

   The list on each node should be the same. 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
   ...
   ```

3. **Are you adding a new node to an existing cluster?**

   - If no, proceed to Step 4.
   - If yes, you might need to update the quorum configuration to accommodate your cluster's new configuration.

     See "Quorum Devices" in *Sun Cluster Overview for Solaris OS* and *Sun Cluster Concepts Guide for Solaris OS* for information about quorum. To change the quorum configuration, follow procedures in "Administering Quorum" in *Sun*

Cluster System Administration Guide for Solaris OS. When the quorum configuration is satisfactory, go to "How to Verify the Quorum Configuration and Installation Mode" on page 98.

4. **Determine the global device-ID name of each shared disk that you are configuring as a quorum device.**

   Use the `scdidadm` output from Step 2 to identify the device–ID name of each shared disk that you are configuring as a quorum device. For example, the output in Step 2 shows that global device `d2` is shared by `phys-schost-1` and `phys-schost-2`. You use this information in Step 7.

   See "Quorum Devices" on page 28 for further information about choosing quorum devices.

5. **Become superuser on one node of the cluster.**

6. **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 9.

   Tip – If the quorum setup process is interrupted or fails to be completed successfully, rerun `scsetup`.

   ---

7. **At the prompt `Do you want to add any quorum disks?`, type `Yes` if your cluster is a two-node cluster. Configure at least one shared quorum device.**

   If your cluster has three or more nodes, quorum device configuration is optional.

8. **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`.

9. **Quit the `scsetup` utility.**

10. **Verify the quorum configuration and that installation mode is disabled.**

    Go to "How to Verify the Quorum Configuration and Installation Mode" on page 98.

▼ How to Verify the Quorum Configuration and Installation Mode

Perform this procedure to verify that quorum configuration completed successfully and that cluster installation mode is disabled.

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. You are now ready to install volume management software and to configure the cluster.

   - To install Solstice DiskSuite software or configure Solaris Volume Manager software, go to "Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software" on page 116.
   - SPARC: To install VERITAS Volume Manager software, go to "SPARC: Installing and Configuring VxVM Software" on page 149.

▼ 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 (see Step 2 of "How to Verify the Quorum Configuration and Installation Mode" on page 98), 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*.

---

1. **Attempt to reinstall the node.**

   You can correct certain failed installations by repeating Sun Cluster software installation on the node. If you have already tried to reinstall the node without success, proceed to Step 2 to uninstall Sun Cluster software from the node.

2. **Become superuser on an active cluster member other than the node that you are uninstalling.**

3. **From an active cluster member, add the node that you intend to uninstall to the cluster node-authentication list.**

   Skip this step if you are uninstalling a single-node cluster.

   ```
   # /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

   Alternately, you can use the `scsetup`(1M) utility. See "How to Add a Cluster Node to the Authorized Node List" in "Adding and Removing a Cluster Node" in *Sun Cluster System Administration Guide for Solaris OS* for procedures.

4. **Become superuser on the node that you intend to uninstall.**

5. **Shut down the node that you intend to uninstall.**

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

6. **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
     ```

7. **Change to a directory, such as the root (/) directory, that does not contain any files that are delivered by the Sun Cluster packages.**

   ```
   # cd /
   ```

8. **Uninstall Sun Cluster software from the node.**

   ```
   # /usr/cluster/bin/scinstall -r
   ```

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

9. **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.

Configuring the Cluster

This section provides information and procedures to configure the software that you installed on the cluster.

Task Map: Configuring the Cluster

The following table lists the tasks to perform to configure your cluster. Before you start to perform these tasks, ensure that you completed the following tasks:

- Cluster framework installation as described in "Installing the Software" on page 40
- Volume manager installation and configuration as described in "Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software" on page 116 or "SPARC: Installing and Configuring VxVM Software" on page 149

TABLE 2–10 Task Map: Configuring the Cluster

Task	Instructions
Create and mount cluster file systems.	"How to Add Cluster File Systems" on page 100
Configure IP Network Multipathing groups.	"How to Configure Internet Protocol (IP) Network Multipathing Groups" on page 105
(Optional) Change a node's private hostname.	"How to Change Private Hostnames" on page 106
Create or modify the NTP configuration file.	"How to Configure Network Time Protocol (NTP)" on page 107
(Optional) SPARC: Install the Sun Cluster module to Sun Management Center software.	"SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 109 Sun Management Center documentation
Install third-party applications and configure the applications, data services, and resource groups.	*Sun Cluster Data Service Planning and Administration Guide for Solaris OS* Third-party application documentation

▼ How to Add Cluster File Systems

Perform this procedure for each cluster file system that you add. A cluster file system is accessible from any node in the cluster.

Caution – Any data on the disks is destroyed when you create a file system. Be sure you specify the correct disk device name. If you specify the wrong device name, you erase data that you might not intend to delete.

If you used SunPlex Manager to install data services, SunPlex Manager might have already created one or more cluster file systems.

1. **Ensure that volume-manager software is installed and configured.**

   For volume-manager installation procedures, see "Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software" on page 116 or "SPARC: Installing and Configuring VxVM Software" on page 149.

2. **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.

3. **Create a file system.**

   - For a UFS file system, use the `newfs`(1M) command.

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

   - SPARC: For a VxFS file system, follow procedures that are provided in your VxFS documentation.

   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` diskset
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`

4. **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

5. **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. See Table 2–11 or Table 2–12 for the list of required mount options.**

      Note – Logging is required for all cluster file systems. However, 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 operating environment in an upcoming Solaris release. Solaris UFS logging provides the same capabilities but superior performance, as well as lower system administration requirements and overhead.

TABLE 2–11 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.
`logging`	**Required**. This option enables logging.
`forcedirectio`	**Required** for cluster file systems that will host Oracle Parallel Server/Real Application Clusters RDBMS data files, log files, and control files. **Note –** Oracle Parallel Server/Real Application Clusters is supported for use only in SPARC based clusters.

TABLE 2–11 Mount Options for UFS Cluster File Systems *(Continued)*

Mount Option	Description
`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.

Note – 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:

- 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.
- 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.

A node might require rebooting to recover from these states.

Only the `onerror=panic` mount option is supported by Sun Cluster software. You do not have to 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.

See the `mount_ufs`(1M) man page for more information about UFS mount options.

TABLE 2–12 SPARC: 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.

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`.

6. 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.

7. Mount the cluster file system.

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

- For UFS, mount the cluster file system from any node in the cluster.
- SPARC: For VERITAS File System (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.

---

8. 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.

9. Configure IP Network Multipathing groups.

Go to "How to Configure Internet Protocol (IP) Network Multipathing Groups" on page 105.

Example – 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
```

▼ How to Configure Internet Protocol (IP) Network Multipathing Groups

Perform this task on each node of the cluster. If you used SunPlex Manager to install Sun Cluster HA for Apache or Sun Cluster HA for NFS, SunPlex Manager 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.

---

1. **Have available your completed "Public Networks Worksheet" on page 240.**

2. **Configure IP Network Multipathing groups.**

   Perform procedures for IPv4 addresses in "Deploying Network Multipathing" in *IP Network Multipathing Administration Guide* (Solaris 8) or "Administering Network Multipathing (Task)" in *System Administration Guide: IP Services* (Solaris 9).

   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.

- For multipathing groups that contain two or more adapters, you must configure a test IP address for each adapter in the group. 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, do no change the value of `TRACK_INTERFACES_ONLY_WITH_GROUPS` from `yes` to `no`.
- The name of a multipathing group has no requirements or restrictions.

3. **Do you intend to change any private hostnames?**
   - If no, proceed to Step 4.
   - If yes, go to "How to Change Private Hostnames" on page 106.
4. **Did you install your own `/etc/inet/ntp.conf` file before you installed Sun Cluster software?**
   - If yes, proceed to Step 5.
   - If no, go to "How to Configure Network Time Protocol (NTP)" on page 107 to install or create the NTP configuration file.
5. **Are you using Sun Cluster on a SPARC based system and do you intend to use Sun Management Center to monitor the cluster?**
   - If yes, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 109.
   - If no, install third-party applications, register resource types, set up resource groups, and configure data services. Follow procedures in the *Sun Cluster Data Service Planning and Administration Guide for Solaris OS* as well as 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.

---

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

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

   ```
   # scsetup
   ```

3. **From the `Main Menu`, choose `Private hostnames`.**

4. **From the `Private Hostname Menu`, choose `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
   ```

7. **Did you install your own `/etc/inet/ntp.conf` file before you installed Sun Cluster software?**

   - If no, go to "How to Configure Network Time Protocol (NTP)" on page 107 to install or create the NTP configuration file.
   - If yes, proceed to Step 8.

8. **Are you using Sun Cluster on a SPARC based system and you intend to use Sun Management Center to monitor the cluster?**

   - If yes, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 109.
   - If no, 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 Service Planning and Administration Guide for Solaris OS*.

▼ How to Configure Network Time Protocol (NTP)

Perform this task to create or modify the NTP configuration file after you install Sun Cluster software. You must also modify the NTP configuration file when you add a node to an existing cluster or when you change the private hostname of a node in the cluster.

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.

1. **Did you install your own `/etc/inet/ntp.conf` file before you installed Sun Cluster software?**
   - If yes, you do not need to modify your `ntp.conf` file. Skip to Step 8.
   - If no, proceed to Step 2.
2. **Become superuser on a cluster node.**
3. **Do you have your own `/etc/inet/ntp.conf` file to install on the cluster nodes?**
   - If no, proceed to Step 4.
   - If yes, copy your `/etc/inet/ntp.conf` file to **each** node of the cluster, then skip to Step 6.
4. **On one node of the cluster, edit the private hostnames in the `/etc/inet/ntp.conf.cluster` 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. **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.

   b. **Remove any unused private hostnames.**

      The `ntp.conf.cluster` file might contain nonexistent private hostnames. When a node is rebooted, the system generates error messages as the node attempts to contact those nonexistent private hostnames.

   c. **If necessary, make other modifications to meet your NTP requirements.**
5. **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.
6. **Stop the NTP daemon on each node.**

Wait for the stop command to complete successfully on each node before you proceed to Step 7.

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

7. **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 that 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 the following command:

     ```
     # /etc/init.d/xntpd start
     ```

8. **Are you using Sun Cluster on a SPARC based system and you intend to use Sun Management Center to monitor the cluster?**
   - If yes, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 109.
   - If no, 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 Service Planning and Administration Guide for Solaris OS for Solaris OS*.

SPARC: Installing the Sun Cluster Module for Sun Management Center

This section provides information and procedures to install the Sun Cluster module to Sun Management Center software.

SPARC: Task Map: Installing the Sun Cluster Module for Sun Management Center

The following table lists the tasks to perform to install the Sun Cluster–module software for Sun Management Center.

TABLE 2–13 Task Map: Installing the Sun Cluster Module for Sun Management Center

Task	Instructions
Install Sun Management Center server, help-server, agent, and console packages.	Sun Management Center documentation "SPARC: Installation Requirements for Sun Cluster Monitoring" on page 110
Install Sun Cluster–module packages.	"SPARC: How to Install the Sun Cluster Module for Sun Management Center" on page 111
Start Sun Management Center server, console, and agent processes.	"SPARC: How to Start Sun Management Center" on page 112
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 112
Load the Sun Cluster module to begin to monitor the cluster.	"SPARC: How to Load the Sun Cluster Module" on page 113

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 packages –** You must install the Sun Management Center server, help–server, and console packages on noncluster nodes, and you must install the Sun Management Center agent package on each cluster node.

  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.

  Follow procedures in your Sun Management Center documentation to install the Sun Management Center packages.
- **Simple Network Management Protocol (SNMP) port –** 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.

▼ 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.

---

1. **Ensure that all Sun Management Center core packages are installed on the appropriate machines.**

   This step includes installing Sun Management Center agent packages on each cluster node. See your Sun Management Center documentation for installation instructions.

2. **On the server machine, install the Sun Cluster–module server package `SUNWscssv`.**

   a. **Become superuser.**

   b. **Insert the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM into the CD-ROM drive. If the volume management daemon `vold`(1M) is running and configured to manage CD-ROM devices, the daemon automatically mounts the CD-ROM on the `/cdrom/cdrom0` directory.**

   c. **From the `/cdrom/cdrom0` directory, 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) or `9` (for Solaris 9) .**

      The following example uses the path to the SPARC Solaris 8 version of Sun Cluster software.

      ```
      # cd /cdrom/cdrom0/Solaris_sparc/Product/sun_cluster/Solaris_8/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, then eject the CD-ROM.**

3. **On the help-server machine, install the Sun Cluster–module help–server package `SUNWscshl`.**

   Use the same procedure as in the previous step.

4. **Install any Sun Cluster–module patches.**

See “Patches and Required Firmware Levels” in *Sun Cluster 3.1 Release Notes* for the location of patches and installation instructions.

5. **Start Sun Management Center.**
   Go to “SPARC: How to Start Sun Management Center” on page 112.

▼ SPARC: How to Start Sun Management Center

Perform this procedure to start the Sun Management Center server, agent, and console processes.

1. **As superuser, on the Sun Management Center server machine, start the Sun Management Center server process.**

   ```
   # /opt/SUNWsymon/sbin/es-start -S
   ```

2. **As superuser, on each Sun Management Center agent machine (cluster node), start the Sun Management Center agent process.**

   ```
   # /opt/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.

   ```
   % /opt/SUNWsymon/sbin/es-start -c
   ```

5. **Type your login name, password, and server hostname, then click Login.**

6. **Add cluster nodes as monitored host objects.**
   Go to “SPARC: How to Add a Cluster Node as a Sun Management Center Agent Host Object” on page 112.

▼ 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.

Note – 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.

1. **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.

2. **Choose Edit⇒Create an Object from the pull-down menu.**

3. **Click the Node tab.**

4. **From the Monitor Via pull-down list, select Sun Management Center Agent - Host.**

5. **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.

6. **In the Port text field, type the port number that you chose when you installed the Sun Management Center agent machine.**

7. **Click OK.**

   A Sun Management Center agent host object is created in the domain.

8. **Load the Sun Cluster module.**

   Go to "SPARC: How to Load the Sun Cluster Module" on page 113.

▼ SPARC: How to Load the Sun Cluster Module

Perform this procedure to start cluster monitoring.

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. **In the Details window 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.

6. **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:

     file:/opt/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.

7. **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 Service Planning and Administration Guide for Solaris OS*.

CHAPTER 3

Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software

Install and configure your local and multihost disks for Solstice DiskSuite/Solaris Volume Manager software by using the procedures in this chapter, along with the planning information in "Planning Volume Management" on page 31. See your Solstice DiskSuite/Solaris Volume Manager documentation for additional details.

The following information and procedures are in this chapter:

Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software

This section provides information and procedures to install and configure Solstice DiskSuite/Solaris Volume Manager software.

Task Map: Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software

The following table lists the tasks that you perform to install and configure Solstice DiskSuite/Solaris Volume Manager software for Sun Cluster configurations. You can skip certain procedures in the following conditions:

- If you used SunPlex Manager to install Solstice DiskSuite software (Solaris 8), the procedures "How to Install Solstice DiskSuite Software" on page 118 through "How to Create State Database Replicas" on page 122 are already completed. Go to "Mirroring the Root Disk" on page 123 or "How to Create a Diskset" on page 138 to continue to configure Solstice DiskSuite software.
- If you installed Solaris 9 software, Solaris Volume Manager is already installed. You can start at "How to Set the Number of Metadevice or Volume Names and Disksets" on page 120.

TABLE 3–1 Task Map: Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software

Task	Instructions
1. Plan the layout of your Solstice DiskSuite/Solaris Volume Manager configuration.	■ "Planning Volume Management" on page 31 ■ "Solstice DiskSuite/Solaris Volume Manager Configuration Example" on page 117
2. (*Solaris 8 only*) Install Solstice DiskSuite software.	"How to Install Solstice DiskSuite Software" on page 118
3. Calculate the number of metadevice names and disksets needed for your configuration, and modify the `/kernel/drv/md.conf` file.	"How to Set the Number of Metadevice or Volume Names and Disksets" on page 120

TABLE 3–1 Task Map: Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software *(Continued)*

Task	Instructions
4. Create state database replicas on the local disks.	"How to Create State Database Replicas" on page 122
5. (*Optional*) Mirror file systems on the root disk.	"Mirroring the Root Disk" on page 123
6. Create disksets by using the `metaset` command.	"How to Create a Diskset" on page 138
7. Add disk drives to the disksets.	"How to Add Disk Drives to a Diskset" on page 141
8. (*Optional*) Repartition disk drives in a diskset to allocate space to slices 1 through 6.	"How to Repartition Disk Drives in a Diskset" on page 142
9. 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 143
10. Initialize the `md.tab` files.	"How to Activate Metadevices or Volumes" on page 144
11. (*Dual-string configurations only*) Configure dual-string mediator hosts, check the status of mediator data, and, if necessary, fix bad mediator data.	1. "How to Add Mediator Hosts" on page 147 2. "How to Check the Status of Mediator Data" on page 147
12. Configure the cluster.	"Configuring the Cluster" on page 100

Solstice DiskSuite/Solaris Volume Manager Configuration Example

The following example helps to explain the process for determining the number of disk drives to place in each diskset. In this example, three storage devices are used. Existing applications are run over NFS (two file systems of 5 Gbytes each) and two ORACLE databases (one 5 Gbytes and one 10 Gbytes).

The following table shows the calculations that are used to determine the number of disk drives needed in the sample configuration. In a configuration with three storage devices, you would need 28 disk drives, which would be divided as evenly as possible among each of the three storage devices. Note that the 5-Gbyte file systems were given an additional 1 Gbyte of disk space because the number of disk drives needed was rounded up.

TABLE 3–2 Determining the Number of Disk Drives Needed for a Configuration

Use	Data	Disk Storage Needed	Disk Drives Needed
`nfs1`	5 Gbytes	3x2.1 Gbyte disks * 2 (Mirror)	6
`nfs2`	5 Gbytes	3x2.1 Gbyte disks * 2 (Mirror)	6
SPARC: `oracle1`	5 Gbytes	3x2.1 Gbyte disks * 2 (Mirror)	6
SPARC: `oracle2`	10 Gbytes	5x2.1 Gbyte disks * 2 (Mirror)	10

The following table shows the allocation of disk drives among the two disksets and four data services.

TABLE 3–3 Division of Disksets

Diskset	Data Services	Disk Drives	Storage Device 1	Storage Device 2	Storage Device 3
`dg-schost-1`	`nfs1`, `oracle1`	12	4	4	4
`dg-schost-2`	`nfs2`, `oracle2`	16	5	6	5

Initially, four disk drives on each storage device (a total of 12 disks) are assigned to `dg-schost-1`, and five or six disk drives on each (a total of 16 disks) are assigned to `dg-schost-2`.

No hot spare disks are assigned to either diskset. A minimum of one hot spare disk per storage device per diskset enables one drive to be hot spared, which restores full two-way mirroring.

▼ How to Install Solstice DiskSuite Software

Note – If you used SunPlex Manager to install Solstice DiskSuite software, do not perform this procedure. Instead, go to "Mirroring the Root Disk" on page 123.

If you installed Solaris 9 software, do not perform this procedure. Solaris Volume Manager software is installed with Solaris 9 software. Instead, go to "How to Set the Number of Metadevice or Volume Names and Disksets" on page 120.

Perform this task on each node in the cluster.

1. **Have available the following information.**
   - Mappings of your storage disk drives.
   - The following completed configuration planning worksheets. See "Planning Volume Management" on page 31 for planning guidelines.
     - "Local File System Layout Worksheet" on page 238
     - "Disk Device Group Configurations Worksheet" on page 244
     - "Volume-Manager Configurations Worksheet" on page 246
     - "Metadevices Worksheet (Solstice DiskSuite/Solaris Volume Manager)" on page 248
2. **Become superuser on the cluster node.**
3. **If you install from the CD-ROM, insert the Solaris 8 Software 2 of 2 CD-ROM into 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.
4. **Install the Solstice DiskSuite software 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
```

Note – If you have Solstice DiskSuite software patches to install, do not reboot after you install the Solstice DiskSuite software.

   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.
5. **If you installed from a CD-ROM, eject the CD-ROM.**
6. **Install any Solstice DiskSuite patches.**

   See "Patches and Required Firmware Levels" in *Sun Cluster 3.1 Release Notes* for the location of patches and installation instructions.
7. **Repeat Step 1 through Step 6 on the other nodes of the cluster.**
8. **From one node of the cluster, manually populate the global-device namespace for Solstice DiskSuite.**

```
# scgdevs
```

Note – The `scgdevs` command might return a message similar to the following:

```
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.

9. **Set the number of metadevice names and disksets that are expected in the cluster.**

   Go to "How to Set the Number of Metadevice or Volume Names and Disksets" on page 120.

▼ How to Set the Number of Metadevice or Volume Names and Disksets

Note – If you used SunPlex Manager to install Solstice DiskSuite software, do not perform this procedure. Instead, go to "Mirroring the Root Disk" on page 123.

This procedure describes how to determine the number of Solstice DiskSuite metadevice or Solaris Volume Manager volume names and disksets 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 diskset 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 `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.

1. **Have available the "Disk Device Group Configurations Worksheet" on page 244.**

2. **Determine the total number of disksets you expect to need in the cluster, then add one more diskset for private disk management.**

   The cluster can have a maximum of 32 disksets, 31 disksets for general use plus one diskset for private disk management. The default number of disksets is 4. You supply this value for the `md_nsets` field in Step 4.

3. **Determine the largest metadevice or volume name you expect to need for any diskset in the cluster.**

   Each diskset can have a maximum of 8192 metadevice or volume names. You supply this value for the `nmd` field in Step 4.

   a. **Determine the quantity of metadevice or volume names you expect to need for each diskset.**

      If you use local metadevices or volumes, ensure that each local metadevice or volume name 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. **Determine the highest of the metadevice or volume names you expect to use in any diskset.**

      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/Solaris Volume Manager software requires that you set the value at 1000 names, not 50.

4. **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 disksets served by each node. Failure to follow this guideline can result in serious Solstice DiskSuite/Solaris Volume Manager errors and possible loss of data.

   ---

   a. **Set the `md_nsets` field to the value that was determined in Step 2.**

   b. **Set the `nmd` field to the value that was determined in Step 3.**

5. **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.

6. **Create local state database replicas.**
   Go to "How to Create State Database Replicas" on page 122.

▼ How to Create State Database Replicas

Note – If you used SunPlex Manager to install Solstice DiskSuite software, do not perform this procedure. Instead, go to "Mirroring the Root Disk" on page 123.

Perform this procedure on each node in the cluster.

1. **Become superuser on the cluster node.**
2. **Create state database replicas on one or more local disks for each cluster node by using the `metadb` command.**
   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/Solaris Volume Manager software, create at least three replicas for each node. Also, you can place replicas on more than one disk to provide protection if one of the disks fails.

   See the `metadb`(1M) man page and your Solstice DiskSuite/Solaris Volume Manager documentation for details.
3. **Verify the replicas.**

   ```
   # metadb
   ```

   The `metadb` command displays the list of replicas.
4. **Do you intend to mirror file systems on the root disk?**
   - If yes, go to "Mirroring the Root Disk" on page 123.
   - If no, go to "How to Create a Diskset" on page 138 to create Solstice DiskSuite/Solaris Volume Manager disksets.

Example—Creating State Database Replicas

The following example shows three Solstice DiskSuite state database replicas. Each replica is created on a different disk. 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
```

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 123
- "How to Mirror the Global Namespace" on page 127
- "How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted" on page 130
- "How to Mirror File Systems That Can Be Unmounted" on page 134

Note – Some of the steps in these mirroring procedures can cause an error message similar to the following, which is harmless and can be ignored.

```
metainit: dg-schost-1: d1s0: not a metadevice
```

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.

1. **Become superuser on the node.**

2. **Use the `metainit`(1M) command to put 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 – 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. **Is the disk that is used to mirror the root disk physically connected to more than one node (multiported)?**

    - If no, proceed to Step 11.
    - If yes, perform the following steps to enable the `localonly` property of the raw-disk device group for the disk 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. Does the node list contain more than one node name?

- If yes, proceed to Step d.
- If no, skip to Step e.

d. Remove all nodes from the node list for the raw-disk device group except the node whose root disk you mirrored.

Only the node whose root disk you mirrored should remain in the node list.

```
# scconf -r -D name=dsk/dN,nodelist=node
```

`-D name=dsk/d`*N*	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

e. 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 "Troubleshooting the System" in *Solstice DiskSuite 4.2.1 User's Guide* or "Mirroring root (/) Special Considerations" 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 is unique throughout the cluster.

13. **Do you intend to mirror the global namespace, `/global/.devices/node@`*nodeid*?**

    - If yes, go to "How to Mirror the Global Namespace" on page 127.
    - If no, proceed to Step 14.

14. **Do you intend to mirror file systems than cannot be unmounted?**

    - If yes, go to "How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted" on page 130.
    - If no, proceed to Step 15.

15. **Do you intend to mirror user-defined file systems?**

    - If yes, go to "How to Mirror File Systems That Can Be Unmounted" on page 134.
    - If no, go to "How to Create a Diskset" on page 138 to create a diskset.

Example—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`. Disk `c2t2d0` is a multiported 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
```

▼ How to Mirror the Global Namespace

Use this procedure to mirror the global namespace, `/global/.devices/node@`*nodeid*.

1. **Become superuser on a node of the cluster.**
2. **Put 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 the mirror *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 complete.**

   Use the `metastat`(1M) command to view mirror status and to verify that mirror synchronization if complete.

   ```
   # metastat mirror
   ```

9. **Is the disk that is used to mirror the global namespace physically connected to more than one node (multiported)?**

   - If no, proceed to Step 10.
   - If yes, perform the following steps to enable the `localonly` property of the raw-disk device group for the disk 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. **Does the node list contain more than one node name?**

      - If yes, proceed to Step d.
      - If no, skip to Step e.

   d. **Remove all nodes from the node list for the raw-disk device group, except the node whose disk is mirrored.**

Only the node whose disk is mirrored should remain in the node list.

```
# scconf -r -D name=dsk/dN,nodelist=node
```

`-D name=dsk/d`*N*	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

e. 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.

10. Do you intend to mirror file systems other than root (/) that cannot be unmounted?

- If yes, go to "How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted" on page 130.
- If no, proceed to Step 11.

11. Do you intend to mirror user-defined file systems?

- If yes, go to "How to Mirror File Systems That Can Be Unmounted" on page 134.
- If no, go to "How to Create a Diskset" on page 138 to create a diskset.

Example—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`. Disk `c2t2d0` is a multiported 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
```

▼ 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`.

1. **Become superuser on a node of the cluster.**
2. **Put 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 (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 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. Is the disk that is used to mirror the unmountable file system physically connected to more than one node (multiported)?

- If no, proceed to Step 12.
- If yes, perform the following steps to enable the `localonly` property of the raw-disk device group for the disk 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. Does the node list contain more than one node name?

- If yes, proceed to Step d.
- If no, skip to Step e.

d. Remove all nodes from the node list for the raw-disk device group except the node whose root disk is mirrored.

Only the node whose root disk is mirrored should remain in the node list.

```
# scconf -r -D name=dsk/dN,nodelist=node
```

`-D name=dsk/d`*N*	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

e. 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.

12. Do you intend to mirror user-defined file systems?

- If yes, go to "How to Mirror File Systems That Can Be Unmounted" on page 134.
- If no, go to "How to Create a Diskset" on page 138 to create a diskset.

Example—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`. Disk `c2t2d0` is a multiported 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
```

▼ 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.

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 "Mounting and Unmounting File Systems" in *System Administration Guide: Basic Administration* for more information.
3. **Put 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. **Is the disk that is used to mirror the user-defined file system physically connected to more than one node (multiported)?**

    - If no, proceed to Step 12.
    - If yes, perform the following steps to enable the `localonly` property of the raw-disk device group for the disk 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. Does the node list contain more than one node name?

- If yes, proceed to Step d.
- If no, skip to Step e.

d. Remove all nodes from the node list for the raw-disk device group except the node whose root disk you mirrored.

Only the node whose root disk you mirrored should remain in the node list.

```
# scconf -r -D name=dsk/dN,nodelist=node
```

`-D name=dsk/d`*N*	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

e. 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. Mount the mirrored file system.

```
# mount /mount-point
```

See the `mount`(1M) man page and "Mounting and Unmounting File Systems" in *System Administration Guide: Basic Administration* for more information.

12. Create a diskset.

Go to "How to Create a Diskset" on page 138.

Example—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`. Disk `c2t2d0` is a multiported 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
```

▼ How to Create a Diskset

Perform this procedure for each diskset you create.

Note – If you used SunPlex Manager to install Solstice DiskSuite, one to three disksets might already exist. See "Using SunPlex Manager to Install Sun Cluster Software" on page 57 for information about the metasets that were created by SunPlex Manager.

---

1. **Do you intend to create more than three disksets in the cluster?**
   - If no, skip to Step 6.
   - If yes, proceed to Step 2 to prepare the cluster for more than three disksets. You must perform that task whether you are installing disksets for the first time or whether you are adding more disksets to a fully configured cluster.
2. **Ensure that the value of the `md_nsets` variable is set high enough to accommodate the total number of disksets you intend to create in the cluster.**
   a. **On any node of the cluster, check the value of the `md_nsets` variable in the `/kernel/drv/md.conf` file.**
   b. **If the total number of disksets in the cluster will be greater than the existing value of `md_nsets` minus one, on each node increase the value of `md_nsets` to the desired value.**

      The maximum permissible number of disksets is one less than the configured value of `md_nsets`. The maximum possible value of `md_nsets` is 32.
   c. **Ensure that the `/kernel/drv/md.conf` file is identical on each node of the cluster.**

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

---

d. From one node, shut down the cluster.

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

e. Reboot each node of the cluster.

```
ok> boot
```

3. **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.

4. **From one node of the cluster, run the `scgdevs`(1M) command to update the global-devices namespace.**

5. **On each node, verify that the `scgdevs` command has completed processing before you attempt to create any disksets.**

   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
   ```

6. **Ensure that the diskset you intend to create meets one of the following requirements.**

   - If the diskset is configured with exactly two disk strings, the diskset must connect to exactly two nodes and use exactly two mediator hosts, which must be the same two hosts used for the diskset. See "Configuring Dual-String Mediators" on page 146 for details on how to configure dual-string mediators.
   - If the diskset is configured with more than two disk strings, ensure that for any two disk strings S1 and S2, the sum of the number of disk drives on those strings exceeds the number of disk drives on the third string S3. Stated as a formula, the requirement is that count(S1) + count(S2) > count(S3).

7. **Ensure that the local state database replicas exist.**

   For instructions, see "How to Create State Database Replicas" on page 122.

8. **Become superuser on the cluster node that will master the diskset.**

9. **Create the diskset.**

   The following command creates the diskset and registers the diskset as a Sun Cluster disk device group.

   ```
   # metaset -s setname -a -h node1 node2
   ```

`-s` *setname*	Specifies the diskset name
`-a`	Adds (creates) the diskset
`-h` *node1*	Specifies the name of the primary node to master the diskset
node2	Specifies the name of the secondary node to master the diskset

Note – When you run the `metaset` command to configure a Solstice DiskSuite/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.

10. Verify the status of the new diskset.

```
# metaset -s setname
```

11. Add disk drives to the diskset.

Go to "Adding Disk Drives to a Diskset" on page 140.

Example—Creating a Diskset

The following command creates two disksets, `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
```

Adding Disk Drives to a Diskset

When you add a disk drive to a diskset, Solstice DiskSuite/Solaris Volume Manager repartitions the disk drive as follows so that the state database for the diskset can be placed on the disk drive.

- A small portion of each disk drive is reserved in slice 7 for use by Solstice DiskSuite/Solaris Volume Manager software. The remainder of the space on each disk drive is placed into slice 0.
- Disk drives are repartitioned when they are added to the diskset only if slice 7 is not configured correctly.
- Any existing data on the disk drives is lost by the repartitioning.
- If slice 7 starts at cylinder 0, and the disk drive partition is large enough to contain a state database replica, the disk drive is not repartitioned.

▼ How to Add Disk Drives to a Diskset

1. **Become superuser on the node.**

2. **Ensure that the diskset has been created.**

   For instructions, see "How to Create a Diskset" on page 138.

3. **List the DID mappings.**

   ```
   # scdidadm -L
   ```

   - Choose disk drives that are shared by the cluster nodes that will master or potentially master the diskset.
   - Use the full device-ID path names when you add disk drives to a diskset.

   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 disk 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 disk 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
   ...
   ```

4. **Take ownership of the diskset.**

   ```
   # metaset -s setname -t
   ```

`-s` *setname*	Specifies the diskset name
`-t`	Takes ownership of the diskset

5. **Add the disk drives to the diskset.**

   Use the full device-ID path name.

   ```
   # metaset -s setname -a drivename
   ```

`-a`	Adds the disk drive to the diskset
drivename	Full device-ID path name of the shared disk drive

Note – Do *not* use the lower-level device name (c*N*t*X*d*Y*) when you add a disk drive to a diskset. 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.

---

6. **Verify the status of the diskset and disk drives.**

   # **metaset -s** *setname*

7. **Do you intend to repartition disk drives for use in metadevices or volumes?**

   - If yes, go to "How to Repartition Disk Drives in a Diskset" on page 142.
   - If no, go to "How to Create an `md.tab` File" on page 143 to define metadevices or volumes by using an `md.tab` file.

Example—Adding Disk Drives to a Diskset

The `metaset` command adds the disk drives `/dev/did/rdsk/d1` and `/dev/did/rdsk/d2` to the diskset `dg-schost-1`.

```
# metaset -s dg-schost-1 -a /dev/did/rdsk/d1 /dev/did/rdsk/d2
```

▼ How to Repartition Disk Drives in a Diskset

The `metaset`(1M) command repartitions disk drives in a diskset so that a small portion of each disk drive is reserved in slice 7 for use by Solstice DiskSuite/Solaris Volume Manager software. The remainder of the space on each disk drive is placed into slice 0. To make more effective use of the disk 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.

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

2. **Use the `format` command to change the disk partitioning for each disk drive in the diskset.**

   When you repartition a disk drive, you must meet the following conditions to prevent the `metaset`(1M) command from repartitioning the disk drive.

   - Create slice 7 starting at cylinder 0, large enough to hold a state database replica (approximately 2 Mbytes).
   - 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 disk drive.

   See the `format`(1M) man page for details.

3. **Define metadevices or volumes by using an `md.tab` file.**

   Go to "How to Create an `md.tab` File" on page 143.

▼ 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 disksets 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 disksets. For example, if the device-ID name `/dev/did/dsk/d3` is used in a diskset, 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#`.

Tip – 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.

---

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.**

   See your Solstice DiskSuite/Solaris Volume Manager documentation and the `md.tab`(4) man page for details on how to create an `md.tab` file.

Note – If you have existing data on the disk 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.

4. **Activate the metadevices or volumes that are defined in the `md.tab` files.**
   Go to "How to Activate Metadevices or Volumes" on page 144.

Example—Sample `md.tab` File

The following sample `md.tab` file defines the diskset 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
   ```

▼ 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.

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 diskset on the node where the command will be executed.**

4. **Take ownership of the diskset.**

   # **metaset -s** *setname* **-t**

   -s *setname* Specifies the diskset name

   -t Takes ownership of the diskset

5. **Activate the diskset's metadevices or volumes, which are defined in the `md.tab` file.**

   # **metainit -s** *setname* **-a**

   -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 by hand.

   # **metattach** *mirror submirror2*

7. **Repeat Step 3 through Step 6 for each diskset in the cluster.**

   If necessary, run the `metainit`(1M) command from another node that has connectivity to the disk drives. This step is required for cluster-pair topologies, where the disk 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.

9. **Does your cluster contain disksets that are configured with exactly two disk enclosures and two nodes?**

   - If yes, those disksets require dual-string mediators. Go to "Configuring Dual-String Mediators" on page 146 to add mediator hosts.
   - If no, go to "How to Add Cluster File Systems" on page 100 to create a cluster file system.

Example—Activating Metadevices or Volumes in the `md.tab` File

In the following example, all metadevices that are defined in the `md.tab` file for diskset `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
```

Configuring Dual-String Mediators

This section contains the following information and procedures:

Requirements for Dual-String Mediators

A *dual-string mediator*, or mediator host, is a cluster node that stores mediator data. Mediator data provides information on the location of other mediators and contains a commit count that is identical to the commit count 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.

Dual-string mediators are required for all Solstice DiskSuite/Solaris Volume Manager disksets that are configured with exactly two disk strings and two cluster nodes. A *disk string* consists of a disk enclosure, its physical disk drives, cables from the enclosure to the node(s), and the interface adapter cards. 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. The following rules apply to dual-string configurations that use mediators.

- Disksets must be configured with exactly two mediator hosts. Those two mediator hosts must be the same two cluster nodes that are used for the diskset.
- A diskset cannot have more than two mediator hosts.
- Mediators cannot be configured for disksets 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 disksets 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.

1. **Become superuser on the node that currently masters the diskset to which you intend to add mediator hosts.**

2. **Run the `metaset`(1M) command to add each node with connectivity to the diskset as a mediator host for that diskset.**

   # **`metaset -s`** *setname* **`-a -m`** *mediator-host-list*

`-s` *setname*	Specifies the diskset name
`-a`	Adds to the diskset
`-m` *mediator-host-list*	Specifies the name of the node to add as a mediator host for the diskset

   See the `mediator`(7D) man page for details about mediator-specific options to the `metaset` command.

3. **Check the status of mediator data.**

   Go to "How to Check the Status of Mediator Data" on page 147.

Example—Adding Mediator Hosts

The following example adds the nodes `phys-schost-1` and `phys-schost-2` as mediator hosts for the diskset `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
```

▼ How to Check the Status of Mediator Data

1. **Add mediator hosts as described in "How to Add Mediator Hosts" on page 147.**

2. **Run the `medstat` command.**

   # **`medstat -s`** *setname*

`-s` *setname*	Specifies the diskset name

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

3. **Is `Bad` the value in the Status field?**

   - If yes, go to "How to Fix Bad Mediator Data" on page 148 to repair the affected mediator host.

- If no, go to "How to Add Cluster File Systems" on page 100 to create a cluster file system.

▼ How to Fix Bad Mediator Data

Perform this procedure to repair bad mediator data.

1. **Identify all mediator hosts with bad mediator data as described in the procedure "How to Check the Status of Mediator Data" on page 147.**
2. **Become superuser on the node that owns the affected diskset.**
3. **Remove all mediator hosts with bad mediator data from all affected disksets.**

   `# metaset -s` *setname* `-d -m` *mediator-host-list*

`-s` *setname*	Specifies the diskset name
`-d`	Deletes from the diskset
`-m` *mediator-host-list*	Specifies the name of the node to remove as a mediator host for the diskset

4. **Restore each mediator host that you removed in Step 3.**

   `# metaset -s` *setname* `-a -m` *mediator-host-list*

`-a`	Adds to the diskset
`-m` *mediator-host-list*	Specifies the name of the node to add as a mediator host for the diskset

   See the `mediator`(7D) man page for details about mediator-specific options to the `metaset` command.

5. **Create cluster file systems.**

   Go to "How to Add Cluster File Systems" on page 100.

CHAPTER 4

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 31. See your VxVM documentation for additional details.

The following information and procedures are in this chapter:

SPARC: Installing and Configuring VxVM Software

This section provides information and procedures to install and configure VxVM software on a Sun Cluster configuration.

SPARC: Task Map: Installing and Configuring VxVM Software

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 31
2. Determine how you will create the root disk group on each node.	"SPARC: Setting Up a Root Disk Group Overview" on page 150
3. Install VxVM software and create the root disk group:	
■ **Method 1** – Install VxVM software and encapsulate the root disk by using the `scvxinstall` command, and optionally mirror the encapsulated root disk.	1. "SPARC: How to Install VERITAS Volume Manager Software and Encapsulate the Root Disk" on page 151 2. "SPARC: How to Mirror the Encapsulated Root Disk" on page 154
■ **Method 2** – Install VxVM software and create the root disk group on local, nonroot disks.	1. "SPARC: How to Install VERITAS Volume Manager Software Only" on page 156 2. "SPARC: How to Create a Root Disk Group on a Nonroot Disk" on page 158
4. Create disk groups and volumes.	"SPARC: How to Create and Register a Disk Group" on page 159
5. 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 161
6. Verify the disk groups and volumes.	"SPARC: How to Verify the Disk Group Configuration" on page 162
7. Configure the cluster.	"Configuring the Cluster" on page 100

SPARC: Setting Up a Root Disk Group Overview

Each cluster node requires the creation of a root disk group after VxVM is installed. This 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 operating environment 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: Where to Go From Here

Install VxVM by using one of the following installation methods, depending on how you intend to create the root disk group.

- If you intend to encapsulate the root disk, go to "SPARC: How to Install VERITAS Volume Manager Software and Encapsulate the Root Disk" on page 151.
- If you intend to create the root disk group on local nonroot disks, go to "SPARC: How to Install VERITAS Volume Manager Software Only" on page 156.

▼ SPARC: How to Install VERITAS Volume Manager Software and Encapsulate the Root Disk

This procedure uses the `scvxinstall`(1M) command to install VxVM software and encapsulate the root disk in one operation.

Note – If you intend to create the root disk group on local, nonroot disks, go instead to "SPARC: How to Install VERITAS Volume Manager Software Only" on page 156.

Perform this procedure on each node that you intend to install with VxVM. You can install VERITAS Volume Manager (VxVM) on all nodes of the cluster, or on only those nodes that are physically connected to the storage devices that VxVM will manage.

If you later need to unencapsulate the root disk, follow the procedures in "SPARC: How to Unencapsulate the Root Disk" on page 163.

1. **Ensure that the cluster meets the following prerequisites.**
   - All nodes in the cluster are running in cluster mode.
   - The root disk of the node you install has two free (unassigned) partitions.
2. **Have available the following information.**
   - Mappings of your storage disk drives.
   - The following completed configuration planning worksheets. See "Planning Volume Management" on page 31 for planning guidelines.
     - "Local File System Layout Worksheet" on page 238
     - "Disk Device Group Configurations Worksheet" on page 244
     - "Volume-Manager Configurations Worksheet" on page 246
3. **Become superuser on a node that you intend to install with VxVM.**
4. **Insert the VxVM CD-ROM into the CD-ROM drive on the node.**
5. **Start `scvxinstall` in interactive mode.**

   Press Ctrl-C at any time to abort the `scvxinstall` command.

   ```
   # scvxinstall
   ```

   See the `scvxinstall`(1M) man page for more information.
6. **When prompted whether to encapsulate root, type `yes`.**

   ```
   Do you want Volume Manager to encapsulate root [no]? y
   ```
7. **When prompted, provide the location of the VxVM CD-ROM.**
   - If the appropriate VxVM CD-ROM is found, the location is displayed as part of the prompt within brackets. Press Enter to accept this default location.

     ```
     Where is the volume manager cdrom [default]?
     ```
   - If the VxVM CD-ROM is not found, the prompt is displayed without a default location. Type the location of the CD-ROM or CD-ROM image.

     ```
     Where is the volume manager cdrom?
     ```
8. **When prompted, type your VxVM license key.**

   ```
   Please enter license key: license
   ```

   The `scvxinstall` command automatically performs the following tasks:
   - Installs the required VxVM software, licensing, and man-page packages, but does not install the GUI packages

- Selects a cluster-wide `vxio` driver major number
- Creates a root disk group by encapsulating the root disk
- Updates the `/global/.devices` entry in the `/etc/vfstab` file

See the `scvxinstall`(1M) man page for further details.

Note – Two automatic reboots occur during installation. After all installation tasks are completed, `scvxinstall` automatically reboots the node the second time unless you press Ctrl-C when prompted. If you press Ctrl-C to abort the second reboot, you must reboot the node later to complete VxVM installation.

---

9. If you intend to enable the VxVM cluster feature, supply the cluster feature license key.

See your VxVM documentation for information about how to add a license.

10. (Optional) Install the VxVM GUI.

See your VxVM documentation for information about installing the VxVM GUI.

11. Eject the CD-ROM.

12. Install any VxVM patches.

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

13. (Optional) If you prefer not to have VxVM man pages reside on the cluster node, remove the man-page package.

```
# pkgrm VRTSvmman
```

14. Do you intend to install VxVM on another node?

- If yes, repeat Step 3 through Step 13.
- If no, proceed to Step 15.

15. Do you *not* intend to install one or more nodes with VxVM?

Note – If you intend to enable the VxVM cluster feature, you *must* install VxVM on all nodes of the cluster.

- If yes, proceed to Step 16.
- If no, skip to Step 17.

16. Modify the `/etc/name_to_major` file on each non-VxVM node.

a. On a node 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 b 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.

17. Do you intend to mirror the encapsulated root disk?

- If yes, go to "SPARC: How to Mirror the Encapsulated Root Disk" on page 154.
- If no, go to "SPARC: How to Create and Register a Disk Group" on page 159.

▼ 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.

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 37 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. **Does the node list contain more than one node name?**

   - If yes, proceed to Step 7.
   - If no, skip to Step 9.

7. **Remove from the node list for the raw-disk device group all nodes except the node whose root disk you mirrored.**

   Only the node whose root disk you mirrored should remain in the node list.

```
# scconf -r -D name=dsk/dN,nodelist=node
```

-D name=dsk/d*N*	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

8. **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.

9. **Repeat this procedure for each node in the cluster whose encapsulated root disk you intend to mirror.**

10. **Create disk groups.**

    Go to "SPARC: How to Create and Register a Disk Group" on page 159.

SPARC: Example—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 multiported 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
```

▼ SPARC: How to Install VERITAS Volume Manager Software Only

This procedure uses the `scvxinstall` command to install VERITAS Volume Manager (VxVM) software only.

Note – To create the root disk group by encapsulating the root disk, do not use this procedure. Instead, go to "SPARC: How to Install VERITAS Volume Manager Software and Encapsulate the Root Disk" on page 151 to install VxVM software and encapsulate the root disk in one operation.

Perform this procedure on each node that you want to install with VxVM. You can install VxVM on all nodes of the cluster, or on only those nodes that are physically connected to the storage devices that VxVM will manage.

1. **Ensure that all nodes in the cluster are running in cluster mode.**
2. **Become superuser on a cluster node that you intend to install with VxVM.**
3. **Insert the VxVM CD-ROM into the CD-ROM drive on the node.**
4. **Start `scvxinstall` in noninteractive installation mode.**

   ```
   # scvxinstall -i
   ```

   The `scvxinstall` command automatically performs the following tasks.

- Installs the required VxVM software, licensing, and man-page packages, but does not install the GUI packages
- Selects a cluster-wide `vxio` driver major number

Note – You add VxVM licenses during the next procedure, "SPARC: How to Create a Root Disk Group on a Nonroot Disk" on page 158.

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

5. **(Optional) Install the VxVM GUI.**

   See your VxVM documentation for information about installing the VxVM GUI.

6. **Eject the CD-ROM.**

7. **Install any VxVM patches.**

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

8. **(Optional) If you prefer not to have VxVM man pages reside on the cluster node, remove the man-page package.**

   ```
   # pkgrm VRTSvmman
   ```

9. **Do you intend to install VxVM on another node?**

   - If yes, repeat Step 2 through Step 8.
   - If no, proceed to Step 10.

10. **Do you *not* intend to install one or more nodes with VxVM?**

Note – If you intend to enable the VxVM cluster feature, you *must* install VxVM on all nodes of the cluster.

   - If yes, proceed to Step 11.
   - If no, skip to Step 12.

11. **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 c 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. **Create a root disk group.**

Go to "SPARC: How to Create a Root Disk Group on a Nonroot Disk" on page 158.

▼ 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.

1. **Have available the VERITAS Volume Manager (VxVM) license keys.**
2. **Become superuser on the node.**
3. **(Optional) 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.

4. **Start the `vxinstall` utility.**

```
# vxinstall
```

When prompted, make the following choices or entries.

- Supply the VxVM license key.
- 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.

5. **Does the root disk group that you created contain one or more disks that connect to more than one node?**

- If yes, enable the `localonly` property of the raw-disk device group for each of these shared disks in the root disk 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 the device that is used by the root disk group if that 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.
- If no, proceed to Step 6.

6. **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

7. **Reboot the node.**

   ```
   # shutdown -g0 -y -i6
   ```

8. **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.

9. **Create disk groups.**

   Go to "SPARC: How to Create and Register a Disk Group" on page 159.

▼ 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 on 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 you add.

1. **Have available the following information.**

   - 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.
   - The following completed configuration planning worksheets.
     - "Local File System Layout Worksheet" on page 238
     - "Disk Device Group Configurations Worksheet" on page 244
     - "Volume-Manager Configurations Worksheet" on page 246

     See "Planning Volume Management" on page 31 for planning guidelines.

2. **Become superuser on the node that will have ownership of the disk group.**

3. **Create a VxVM disk group and volume.**

   If you are installing Oracle Parallel Server/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.

   ---

4. **Is the VxVM cluster feature enabled?**

   - If no, proceed to Step 5.
   - If yes, skip to Step 7. If the VxVM cluster feature is enabled, do *not* register a shared disk group as a Sun Cluster disk device group.

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

   a. **Start the `scsetup`(1M) utility.**

      ```
      # scsetup
      ```

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

   c. **To register a disk device group, type `1` (Register a VxVM disk group).**

      Follow the instructions and type the VxVM disk device group to be registered as a Sun Cluster disk device group.

   d. **If you encounter the following error message when you attempt to register the disk device group, reminor the disk device group.**

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

To reminor the disk device group, use the procedure "SPARC: How to Assign a New Minor Number to a Disk Device Group" on page 161. This procedure enables you to assign a new minor number that does not conflict with a minor number used by existing disk device groups.

e. When finished, type `q` (Quit) to leave the `scsetup` utility.

6. 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
```

Tip – If you experience a stack overflow when the disk device group is brought online, the default value of the thread stack size might be insufficient. Add the following entry to the `/etc/system` file on each node, where *size* is a number greater than 8000, the default setting:

```
set cl_comm:rm_thread_stacksize=0xsize
```

Note – 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.

---

7. Verify the configuration of your VxVM disk groups and volumes.

Go to "SPARC: How to Verify the Disk Group Configuration" on page 162.

▼ 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.

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*

5. **Go to Step 5 of "SPARC: How to Create and Register a Disk Group" on page 159 to register the disk group as a Sun Cluster disk device group.**

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

This example uses the minor numbers 16000-16002 and 4000-4001. The `vxdg reminor` command 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
```

▼ SPARC: How to Verify the Disk Group Configuration

Perform this procedure on each node of the cluster.

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
   ```

4. **Configure the cluster.**

Go to “Configuring the Cluster” on page 100.

▼ SPARC: How to Unencapsulate the Root Disk

Perform this procedure to unencapsulate the root disk.

1. **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`. If any other file systems reside on the root disk, back them up and remove them from the root disk.

2. **Become superuser on the node that you intend to unencapsulate.**

3. **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

4. **Determine the node-ID number of the node.**

   ```
   # clinfo -nN
   ```

5. **Unmount the global-devices file system for this node, where *N* is the node ID number that is returned in Step 4.**

   ```
   # umount /global/.devices/node@N
   ```

6. **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
   ```

7. **Remove the VxVM volume that corresponds to the global-devices file system from the root disk group.**

   ```
   # vxedit -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.

8. **Unencapsulate the root disk.**

   Note – Do **not** accept the shutdown request from the command.

   ```
   # /etc/vx/bin/vxunroot
   ```

   See your VxVM documentation for details.

9. **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.

10. **Set up a file system on the partition that you created in Step 9.**

    ```
    # newfs /dev/rdsk/cNtXdYsZ
    ```

11. **Determine the DID name of the root disk.**

    ```
    # scdidadm -l cNtXdY
    1        phys-schost-1:/dev/rdsk/cNtXdY   /dev/did/rdsk/dN
    ```

12. **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 11.**

    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
```

13. **Mount the global-devices file system.**

    ```
    # mount /global/.devices/node@N
    ```

14. **From one node of the cluster, repopulate the global-devices file system with device nodes for any raw disk and Solstice DiskSuite/Solaris Volume Manager devices.**

    ```
    # scgdevs
    ```

    VxVM devices are recreated during the next reboot.

15. **Reboot the node.**

    ```
    # reboot
    ```

16. **Repeat this procedure on each node of the cluster to unencapsulate the root disk on those nodes.**

CHAPTER 5

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 4/04 software:

Overview of Upgrading a Sun Cluster Configuration

This section provides the following guidelines to upgrade a Sun Cluster configuration:

Upgrade Requirements and Restrictions

Observe the following requirements and restrictions when you upgrade to Sun Cluster 3.1 4/04 software:

- The cluster must run on or be upgraded to at least Solaris 8 2/02 software, including the most current required patches.
- The cluster hardware must be a supported configuration for Sun Cluster 3.1 4/04 software. Contact your Sun representative for information about current supported Sun Cluster configurations.
- You must upgrade all software to a version that is supported by Sun Cluster 3.1 4/04 software. For example, if a data service is supported on Sun Cluster 3.0 software but is not supported on Sun Cluster 3.1 4/04 software, you must upgrade that data service to the version of that data service that is supported on Sun Cluster 3.1 4/04 software. If the related application of that data service is not supported on Sun Cluster 3.1 4/04 software, you must also upgrade that application to a supported release.
- The `scinstall` upgrade utility only upgrades those data services that are provided with Sun Cluster 3.1 4/04 software. You must manually upgrade any custom or third-party data services.
- 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 *System Administration Guide: IP Services* (Solaris 9) for information about test IP addresses for IP Network Multipathing groups.
- Sun Cluster 3.1 4/04 software supports only nonrolling upgrade from Solaris 8 software to Solaris 9 software.
- Sun Cluster 3.1 4/04 software supports direct upgrade only from Sun Cluster 3.x software.
- Sun Cluster 3.1 4/04 software does not support any downgrade of Sun Cluster software.
- Sun Cluster 3.1 4/04 software does not support upgrade between architectures.
- Sun Cluster 3.1 4/04 software does not support the Live Upgrade method to upgrade Solaris software in a Sun Cluster configuration.

Choosing a Sun Cluster Upgrade Method

Choose one of the following methods to upgrade your cluster to Sun Cluster 3.1 4/04 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.
  - 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 also upgrading VxVM.
- **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 Solaris software only to a Solaris Update release, 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.

For overview information about planning your Sun Cluster 3.1 4/04 configuration, see Chapter 1.

Upgrading to Sun Cluster 3.1 4/04 Software (Nonrolling)

Follow the tasks in this section to perform a nonrolling upgrade from Sun Cluster 3.x software to Sun Cluster 3.1 4/04 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.

Note – To perform a rolling upgrade to Sun Cluster 3.1 4/04 software, instead follow the procedures in "Upgrading to Sun Cluster 3.1 4/04 Software (Rolling)" on page 187.

Task Map: Upgrading to Sun Cluster 3.1 4/04 Software (Nonrolling)

TABLE 5–1 Task Map: Upgrading to Sun Cluster 3.1 4/04 Software (Nonrolling)

Task	Instructions
1. Read the upgrade requirements and restrictions.	"Upgrade Requirements and Restrictions" on page 168
2. Take the cluster out of production, disable resources, and back up shared data and system disks. If the cluster uses dual-string mediators for Solstice DiskSuite/Solaris Volume Manager, unconfigure the mediators.	"How to Prepare the Cluster for Upgrade (Nonrolling)" on page 170
3. Upgrade the Solaris software, if necessary, to a supported Solaris update release. Optionally, upgrade VERITAS Volume Manager (VxVM).	"How to Upgrade the Solaris Operating Environment (Nonrolling)" on page 174
4. Upgrade to Sun Cluster 3.1 4/04 framework and data-service software. If necessary, upgrade applications. If the cluster uses dual-string mediators, reconfigure the mediators. If you upgraded VxVM, upgrade disk groups.	"How to Upgrade to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 177
5. (*Optional*) Upgrade the Sun Cluster module to Sun Management Center, if needed.	"How to Upgrade Sun Cluster-Module Software for Sun Management Center (Nonrolling)" on page 183
6. Register new resource types, migrate existing resources to new resource types, modify resource type extension properties as needed, enable resources, and bring resource groups online.	"How to Finish Upgrading to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 185

▼ How to Prepare the Cluster for Upgrade (Nonrolling)

Before you upgrade the software, perform the following steps to take the cluster out of production:

1. **Ensure that the configuration meets requirements for upgrade.**
   See "Upgrade Requirements and Restrictions" on page 168.

2. **Have available the CD-ROMs, documentation, and patches for all software products you are upgrading.**
   - Solaris 8 or Solaris 9 operating environment
   - Sun Cluster 3.1 4/04 framework
   - Sun Cluster 3.1 4/04 data services (agents)
   - Applications that are managed by Sun Cluster 3.1 4/04 data-service agents
   - VERITAS Volume Manager

See “Patches and Required Firmware Levels” in *Sun Cluster 3.1 Release Notes* for the location of patches and installation instructions.

3. **(Optional) Install Sun Cluster 3.1 4/04 documentation.**

   Install the documentation packages on your preferred location, such as an administrative console or a documentation server. See the `index.html` file at the top level of the Java Enterprise System Accessory CD 3 CD-ROM to access installation instructions.

4. **Are you upgrading from Sun Cluster 3.0 software?**

   - If no, proceed to Step 5.
   - If yes, have available your list of test IP addresses, one for each public-network adapter in the cluster.

     A test IP address is required for each public-network adapter in the cluster, regardless of whether the adapter is the active adapter or the backup adapter in the group. The test IP addresses will be 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 the *IP Network Multipathing Administration Guide* (Solaris 8) or *System Administration Guide: IP Services* (Solaris 9) for more information about test IP addresses for IP Network Multipathing.

5. **Notify users that cluster services will be unavailable during upgrade.**

6. **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 volume-manager status.

7. **Become superuser on a node of the cluster.**

8. **Switch each resource group offline.**

   ```
   # scswitch -F -g resource-group
   ```

-F	Switches a resource group offline
-g *resource-group*	Specifies the name of the resource group to take offline

9. 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.

Note – If you are upgrading from a Sun Cluster 3.1 release, you can use the `scsetup`(1M) utility instead of the command line. From the Main Menu, choose Resource Groups, then choose Enable/Disable Resources.

---

a. From any node, list all enabled resources in the cluster.

```
# scrgadm -pv | grep "Res enabled"
(resource-group:resource) Res enabled: True
```

b. Identify those resources that depend on other resources.

You must disable dependent resources first before you disable the resources that they depend on.

c. Disable each enabled resource in the cluster.

```
scswitch -n -j resource
```

-n	Disables
-j *resource*	Specifies the resource

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

d. Verify that all resources are disabled.

```
# scrgadm -pv | grep "Res enabled"
(resource-group:resource) Res enabled: False
```

10. Move each resource group to the unmanaged state.

```
# scswitch -u -g resource-group
```

-u	Moves the specified resource group to the unmanaged state
-g *resource-group*	Specifies the name of the resource group to move into the unmanaged state

11. Verify that all resources on all nodes are `Offline` and that all resource groups are in the `Unmanaged` state.

```
# scstat -g
```

12. Does your cluster use dual-string mediators for Solstice DiskSuite/Solaris Volume Manager?

- If no, proceed to Step 13.
- If yes, unconfigure your mediators.

  See "Configuring Dual-String Mediators" on page 146 for more information.

a. Run the following command to verify that no mediator data problems exist.

```
# medstat -s setname
```

`-s` *setname* Specifies the diskset 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 148.

b. List all mediators.

Use this information for when you restore the mediators during the procedure "How to Upgrade to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 177.

c. For a diskset that uses mediators, take ownership of the diskset if no node already has ownership.

```
# metaset -s setname -t
```

`-t` Takes ownership of the diskset

d. Unconfigure all mediators for the diskset.

```
# metaset -s setname -d -m mediator-host-list
```

`-s` *setname*	Specifies the diskset name
`-d`	Deletes from the diskset
`-m` *mediator-host-list*	Specifies the name of the node to remove as a mediator host for the diskset

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 diskset that uses mediators.

13. Stop all applications that are running on each node of the cluster.

14. Ensure that all shared data is backed up.

15. From one node, shut down the cluster.

```
# scshutdown -g -y
```

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

16. Boot each node into noncluster mode.

```
ok boot -x
```

17. **Ensure that each system disk is backed up.**

18. **Determine whether to upgrade the Solaris operating environment.**

    - If Sun Cluster 3.1 4/04 software does not support the release of the Solaris environment that you currently run on your cluster, you must upgrade the Solaris software to a supported release. Go to "How to Upgrade the Solaris Operating Environment (Nonrolling)" on page 174.
    - If your cluster configuration already runs on a release of the Solaris environment that supports Sun Cluster 3.1 4/04 software, further Solaris software upgrade is optional.
      - To upgrade Sun Cluster software, go to "How to Upgrade to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 177.
      - To upgrade Solaris software, go to "How to Upgrade the Solaris Operating Environment (Nonrolling)" on page 174.

    See "Supported Products" in *Sun Cluster Release Notes for Solaris OS* for more information.

▼ How to Upgrade the Solaris Operating Environment (Nonrolling)

Perform this procedure on each node in the cluster to upgrade the Solaris operating environment. If the cluster already runs on a version of the Solaris environment that supports Sun Cluster 3.1 4/04 software, this procedure is optional.

Note – The cluster must already run on, or be upgraded to, at least the minimum required level of the Solaris 8 or Solaris 9 environment to support Sun Cluster 3.1 4/04 software. See "Supported Products" in *Sun Cluster Release Notes for Solaris OS* for more information.

---

1. **Ensure that all steps in "How to Prepare the Cluster for Upgrade (Nonrolling)" on page 170 are completed.**

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

3. **Determine whether the following Apache links already exist, and if so, whether the file names contain an uppercase K or S:**

```
/etc/rc0.d/K16apache
/etc/rc1.d/K16apache
/etc/rc2.d/K16apache
/etc/rc3.d/S50apache
/etc/rcS.d/K16apache
```

- If these links already exist and do contain an uppercase K or S in the file name, no further action is necessary for these links.
- If these links do not exist, or if these links exist but instead contain a lowercase k or s in the file name, you move aside these links in Step 8.

4. **Comment out all entries for globally mounted file systems in the `/etc/vfstab` file.**

   a. **Make a record of all entries that are already commented out for later reference.**

   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 operating environment.**

Volume Manager	Procedure to Use	Location of Instructions
Solstice DiskSuite/Solaris Volume Manager	Any Solaris upgrade method *except* the Live Upgrade method.	Solaris 8 or Solaris 9 installation documentation
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.**

   Note the following special instructions:

   - Do not perform the final reboot instruction in the Solaris software upgrade. Instead, return to this procedure to perform Step 7 and Step 8, then reboot into noncluster mode in Step 9 to complete Solaris software upgrade.
   - When 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:

```
# reboot -- -xs
ok boot -xs
```

7. **In the `/a/etc/vfstab` file, uncomment those entries for globally mounted file systems that you commented out in Step 4.**

8. **Move aside restored Apache links if either of the following conditions was true *before* you upgraded the Solaris software:**

   - The Apache links listed in Step 3 did not exist.
   - The Apache links listed in Step 3 existed *and* contained a lowercase k or s in the file names.

   To move aside restored Apache links, which contain an uppercase K or S in the name, 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
```

9. **Reboot the node into noncluster mode.**

   Include the double dashes (--) in the following command:

```
# reboot -- -x
```

10. **If your cluster runs VxVM, perform the remaining steps in the procedure "Upgrading VxVM and Solaris" to reinstall or upgrade VxVM.**

    Note the following special instructions:

    - 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
```

    - When the VxVM procedures instruct you to perform a final reconfiguration reboot by using the `-r` option, reboot into noncluster mode by using the `-x` option instead.

```
# reboot -- -x
```

    - After VxVM upgrade is complete, 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.

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 Release Notes* for the location of patches and installation instructions.

12. **Upgrade to Sun Cluster 3.1 4/04 software.**

    Go to "How to Upgrade to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 177.

Note – To complete upgrade from Solaris 8 to Solaris 9 software, you must also upgrade to the Solaris 9 version of Sun Cluster 3.1 4/04 software, even if the cluster already runs on Sun Cluster 3.1 4/04 software.

▼ How to Upgrade to Sun Cluster 3.1 4/04 Software (Nonrolling)

This procedure describes how to upgrade the cluster to Sun Cluster 3.1 4/04 software. You must also perform this procedure to complete cluster upgrade from Solaris 8 to Solaris 9 software.

Tip – You can perform this procedure on more than one node at the same time.

---

1. **Ensure that all steps in "How to Prepare the Cluster for Upgrade (Nonrolling)" on page 170 are completed.**

   If you upgraded from Solaris 8 to Solaris 9 software, also ensure that all steps in "How to Upgrade the Solaris Operating Environment (Nonrolling)" on page 174 are completed.

2. **Become superuser on a node of the cluster.**

3. **Ensure that you have installed all required Solaris software patches and hardware-related patches.**

   For Solstice DiskSuite software (Solaris 8), also ensure that you have installed all required Solstice DiskSuite software patches.

4. **Insert the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM into the CD-ROM drive on the node.**

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

5. **Upgrade the node to Sun Cluster 3.1 4/04 software.**

   a. **From the `/cdrom/cdrom0` directory, change to the `Solaris_sparc/Product/sun_cluster/Solaris_`*ver*`/Tools` directory, where *ver* is `8` (for Solaris 8) or `9` (for Solaris 9) .**

   The following example uses the path to the Solaris 8 version of Sun Cluster software.

   ```
   # cd /cdrom/cdrom0/Solaris_sparc/Product/sun_cluster/Solaris_8/Tools
   ```

   b. **Upgrade the cluster framework software.**

   - To upgrade from Sun Cluster 3.0 software, run the following command:

     ```
     # ./scinstall -u update -S interact -M patchdir=dirname
     ```

     `-S`
     Specifies the test IP addresses to use to convert NAFO groups to IP Network Multipathing groups

     `interact`
     Specifies that `scinstall` prompts the user for each test IP address needed

     `-M patchdir=`*dirname*`[,patchlistfile=`*filename*`]`
     Specifies the path to patch information so that the specified patches can be installed using the `scinstall` command. If you do not specify a patch-list file, the `scinstall` command installs all the patches in the directory *dirname*, including tarred, jarred, and zipped patches.

     The `-M` option is not required. You can use any method you prefer for installing patches.

   - To upgrade from Sun Cluster 3.1 software, run the following command:

     ```
     # ./scinstall -u update -M patchdir=dirname
     ```

     `-M patchdir=`*dirname*`[,patchlistfile=`*filename*`]`
     Specifies the path to patch information so that the specified patches can be installed using the `scinstall` command. If you do not specify a patch-list file, the `scinstall` command installs all the patches in the directory *dirname*, including tarred, jarred, and zipped patches.

   The `-M` option is not required. You can use any method you prefer for installing patches.

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

Upgrade processing is finished when the system displays the message `Completed Sun Cluster framework upgrade` and the path to the upgrade log.

See the `scinstall`(1M) man page for more information. See the *IP Network Multipathing Administration Guide* (Solaris 8) or *System Administration Guide: IP Services* (Solaris 9) for information about test addresses for IP Network Multipathing.

Note – Sun Cluster 3.1 4/04 software requires at least version 3.5.1 of Sun Explorer software. Upgrade to Sun Cluster software includes installing Sun Explorer data collector software, to be used in conjunction with the `sccheck` utility. If another version of Sun Explorer software was already installed before Sun Cluster upgrade, it is replaced by the version that is provided with Sun Cluster software. Options such as user identity and data delivery are preserved, but `crontab` entries must be manually recreated.

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.
- 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.

c. Change to the CD-ROM root directory and eject the CD-ROM.

6. Upgrade software applications that are installed on the cluster.

Ensure that application levels are compatible with the current version 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 4/04 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:

```
# reboot -- -xs
ok boot -xs
```

7. **(Optional) Upgrade Sun Cluster data services to the Sun Cluster 3.1 4/04 software versions.**

Note – You must upgrade the Sun Cluster HA for Oracle 3.0 64–bit for Solaris 9 data service to the Sun Cluster 3.1 4/04 version. Otherwise, you can continue to use Sun Cluster 3.0 data services after upgrade to Sun Cluster 3.1 4/04 software.

Only those data services that are delivered on the Java Enterprise System Accessory CD 3 CD-ROM are automatically upgraded by the `scinstall`(1M) utility. You must manually upgrade any custom or third-party data services. Follow the procedures provided with those data services.

a. **Insert the Java Enterprise System Accessory CD 3 CD-ROM into the CD-ROM drive on the node to upgrade.**

b. **Upgrade the data-service software.**

```
# scinstall -u update -s all -d /cdrom/cdrom0
```

`-u update`	Specifies upgrade
`-s all`	Updates all Sun Cluster data services that are installed on the node

Upgrade processing is finished when the system displays the message `Completed upgrade of Sun Cluster data services agents` and the path to the upgrade log.

c. **Change to the CD-ROM root directory and eject the CD-ROM.**

d. **As needed, manually upgrade any custom data services that are not supplied on the Java Enterprise System Accessory CD 3 CD-ROM.**

e. **Install any Sun Cluster 3.1 4/04 data-service patches.**

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

8. **After all nodes are upgraded, reboot each node into the cluster.**

```
# reboot
```

9. **Verify that all upgraded software is at the same version on all upgraded nodes.**

a. **On each upgraded node, view the installed levels of Sun Cluster software.**

```
# scinstall -pv
```

b. **From one 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.

10. Did you upgrade from Solaris 8 to Solaris 9 software?

- If no, skip to Step 14.
- If yes, proceed to Step 11.

11. On each node, run the following command to verify the consistency of the storage configuration:

```
# scdidadm -c
```

`-c` Perform a consistency check

Caution – Do not proceed to Step 12 until your configuration passes this consistency check. Failure to do so 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 to Take
`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 199 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 12.
No output message	None

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

12. On each node, migrate the Sun Cluster storage database to Solaris 9 device IDs.

```
# scdidadm -R all
```

`-R` Perform repair procedures

`all` Specify all devices

13. 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 11 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. If device-ID migration is verified on all cluster nodes, proceed to Step 14.

14. Does your configuration use dual-string mediators for Solstice DiskSuite/Solaris Volume Manager?

- If no, proceed to Step 15.
- If yes, restore the mediator configurations.

a. Determine which node has ownership of a diskset to which you will add the mediator hosts.

```
# metaset -s setname
```

`-s` *setname* Specifies the diskset name

b. If no node has ownership, take ownership of the diskset.

```
# metaset -s setname -t
```

`-t` Takes ownership of the diskset

c. Recreate the mediators.

```
# metaset -s setname -a -m mediator-host-list
```

`-a`	Adds to the diskset
`-m` *mediator-host-list*	Specifies the names of the nodes to add as mediator hosts for the diskset

d. Repeat Step a through Step c for each diskset in the cluster that uses mediators.

15. Did you upgrade VxVM?

- If no, proceed to Step 16.
- If yes, upgrade all disk groups.

  To upgrade a disk group to the highest version supported by the VxVM release you installed, run the following command from the primary node of the disk group:

  ```
  # vxdg upgrade dgname
  ```

  See your VxVM administration documentation for more information about upgrading disk groups.

16. Do you use Sun Management Center to monitor the cluster?

- If yes, go to "How to Upgrade Sun Cluster-Module Software for Sun Management Center (Nonrolling)" on page 183.

- If no, go to "How to Finish Upgrading to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 185.

Example—Upgrade From Sun Cluster 3.0 to Sun Cluster 3.1 4/04 Software

The following example shows the process of a nonrolling upgrade of a two-node cluster from Sun Cluster 3.0 to Sun Cluster 3.1 4/04 software on the Solaris 8 operating environment. The cluster node names are `phys-schost-1` and `phys-schost-2`.

```
    (On the first node, upgrade framework software from the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM)
phys-schost-1# cd /cdrom/cdrom0/Solaris_sparc/Product/sun_cluster/ \
Solaris_8/Tools
phys-schost-1# ./scinstall -u update -S interact

    (On the first node, upgrade data services from the Java Enterprise System Accessory CD 3 CD-ROM)
phys-schost-1# ./scinstall -u update -s all -d /cdrom/cdrom0

    (On the second node, upgrade framework software from the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM)
phys-schost-2# cd /cdrom/cdrom0/Solaris_sparc/Product/sun_cluster/ \
Solaris_8/Tools
phys-schost-2# ./scinstall -u update -S interact

    (On the second node, upgrade data services from the Java Enterprise System Accessory CD 3 CD-ROM)
phys-schost-2# ./scinstall -u update -s all -d /cdrom/cdrom0

    (Reboot each node into the cluster)
phys-schost-1# reboot
phys-schost-2# reboot

    (Verify cluster membership)
# scstat
-- Cluster Nodes --
                    Node name      Status
                    ---------      ------
  Cluster node:     phys-schost-1  Online
  Cluster node:     phys-schost-2  Online
```

▼ How to Upgrade Sun Cluster-Module Software for Sun Management Center (Nonrolling)

Perform the following steps to upgrade Sun Cluster-module software on the Sun Management Center server machine, help-server machine, and console machine.

If you intend to upgrade the Sun Management Center software itself, do not perform this procedure. Instead, proceed to "How to Finish Upgrading to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 185 to finish Sun Cluster software upgrade. Then go to "How to Upgrade Sun Management Center Software" on page 201 to upgrade Sun Management Center software and the Sun Cluster module.

1. **As superuser, remove the existing Sun Cluster–module packages.**

   Use the `pkgrm`(1M) command to remove all Sun Cluster–module packages from all locations listed in the following table.

   ```
   # pkgrm module-package
   ```

Location	Module Package to Remove
Sun Management Center console machine	`SUNWscscn`
Sun Management Center server machine	`SUNWscssv`
Sun Management Center help-server machine	`SUNWscshl`

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 from the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM to the locations listed in the following table.**

   In the CD-ROM path, the value of *ver* is `8` (for Solaris 8) or `9` (for Solaris 9).

   ```
   # cd /cdrom/cdrom0/Solaris_sparc/Product/sun_cluster/Solaris_ver/Packages/
   # pkgadd module-package
   ```

Location	Module Package to Install
Sun Management Center console machine	`SUNWscshl`
Sun Management Center server machine	`SUNWscssv`
Sun Management Center help-server machine	`SUNWscshl`

   You install the help-server package `SUNWscshl` on both the console machine and the help-server machine. You do not upgrade to a new `SUNWscscn` package on the console machine.

3. **Finish the upgrade.**

Go to "How to Finish Upgrading to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 185.

▼ How to Finish Upgrading to Sun Cluster 3.1 4/04 Software (Nonrolling)

Perform this procedure to reregister all resource types that received a new version from the upgrade, to modify eligible resources to use the new version of its resource type, then to re-enable resources and bring resource groups back online.

Note – To upgrade future versions of resource types, see "Upgrading a Resource Type" in *Sun Cluster Data Service Planning and Administration Guide for Solaris OS*.

---

1. **Ensure that all steps in "How to Upgrade to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 177 are completed.**
2. **If you upgraded any data services that are not supplied on the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM or the Java Enterprise System Accessory CD 3 CD-ROM, register the new resource types for those data services.**

   Follow the documentation that comes with the data services.
3. **From any node, start the `scsetup`(1M) utility.**

   ```
   # scsetup
   ```
4. **Register the new resource types.**

   a. **From the Main Menu, choose `Resource groups`.**

   b. **Choose `Resource type registration`.**

   c. **Choose `Register all resource types which are not yet registered`.**

      The `scsetup` utility displays all resource types that are not registered.

      Follow the prompts to register the new resource types.
5. **Migrate all eligible resources to the new version of its resource type.**

   a. **From the Resource Group menu, choose `Change properties of a resource`.**

   b. **Choose `Manage resource versioning`.**

   c. **Choose `Show versioning status`.**

The `scsetup` utility displays all resources for which a new version of its resource type was installed during the upgrade. Make note of which new resource types you will upgrade the resources to.

d. Choose `Re-version all eligible resources`.

Follow the prompts to upgrade eligible resources to the new version of its resource type.

e. Return to the Change properties of a resource menu.

6. Modify extension properties for new resource type versions.

a. For each new resource type that you migrated existing resources to, determine whether the new resource type requires additional modifications to its extension properties.

Refer to each related data service manual for the requirements of each new resource type.

Note – You do not need to change the `Type_version` property of a new resource type. That property was modified when you migrated resources to their new resource types in Step 5.

- If no resource type requires additional modifications other than the `Type_version` property, go to Step 7.
- If one or more resource types require additional modifications to extension properties, proceed to Step b.

b. From the Change properties of a resource menu, choose `Change extension resource properties`.

c. Follow the prompts to modify the necessary extension properties.

Refer to your data service documentation for the names of extension properties and values to modify.

d. Repeat for each resource type that requires modifications.

e. Return to the Resource Groups menu.

7. Re-enable all disabled resources.

a. From the Resource Group Menu, choose `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 `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.

   The cluster upgrade is complete. You can now return the cluster to production.

Upgrading to Sun Cluster 3.1 4/04 Software (Rolling)

This section provides procedures to perform a rolling upgrade from Sun Cluster 3.1 software to Sun Cluster 3.1 4/04 software. In a rolling upgrade, you upgrade one cluster node at a time, while the other cluster nodes remain in production.

To upgrade from Sun Cluster 3.0 software, follow the procedures in "Upgrading to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 169.

Note – Sun Cluster 3.1 4/04 software does not support rolling upgrade from Solaris 8 software to Solaris 9 software. You can 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, perform the procedures in "Upgrading to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 169.

Task Map: Upgrading to Sun Cluster 3.1 4/04 Software (Rolling)

TABLE 5–2 Task Map: Upgrading to Sun Cluster 3.1 4/04 Software (Rolling)

Task	Instructions
1. Read the upgrade requirements and restrictions.	"Upgrade Requirements and Restrictions" on page 168

TABLE 5–2 Task Map: Upgrading to Sun Cluster 3.1 4/04 Software (Rolling) *(Continued)*

Task	Instructions
2. Take the cluster out of production, disable resources, and ensure that shared data and system disks are backed up. If the cluster uses dual-string mediators for Solstice DiskSuite/Solaris Volume Manager, unconfigure the mediators.	"How to Prepare the Cluster for Upgrade (Rolling)" on page 188
3. Upgrade the Solaris software, if necessary, to a supported Solaris update release. Optionally, upgrade VERITAS Volume Manager (VxVM).	"How to Upgrade to a Solaris Maintenance Update Release (Rolling)" on page 191
4. Upgrade to Sun Cluster 3.1 4/04 framework and data-service software. If necessary, upgrade applications. If you upgraded VxVM, upgrade disk groups.	"How to Upgrade to Sun Cluster 3.1 4/04 Software (Rolling)" on page 192
5. If the cluster uses dual-string mediators, reconfigure the mediators. Register new resource types, migrate existing resources to new resource types, modify resource type extension properties as needed, enable resources, and bring resource groups online. Upgrade the Sun Cluster module to Sun Management Center, if needed.	"How to Finish Upgrading to Sun Cluster 3.1 4/04 Software (Rolling)" on page 195

▼ How to Prepare the Cluster for Upgrade (Rolling)

Perform this procedure on one node at a time. The upgraded node is taken out of the cluster, while the remaining nodes continue to function as active cluster members.

Note – Do not use any new features of the update release, install new data services, or issue any administrative configuration commands until all nodes of the cluster are successfully upgraded.

---

1. **Ensure that the configuration meets requirements for upgrade.**

   See "Upgrade Requirements and Restrictions" on page 168.

2. **Have available the CD-ROMs, documentation, and patches for all the software products you are upgrading before you begin to upgrade the cluster.**

   - Solaris 8 or Solaris 9 operating environment
   - Sun Cluster 3.1 4/04 framework
   - Sun Cluster 3.1 4/04 data services (agents)
   - Applications that are managed by Sun Cluster 3.1 4/04 data-service agents

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

3. **(Optional) Install Sun Cluster 3.1 4/04 documentation.**

   Install the documentation packages on your preferred location, such as an administrative console or a documentation server. See the `index.html` file at the top level of the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM to access installation instructions.

4. **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.

5. **Become superuser on one node of the cluster to upgrade.**

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 completed successfully.**

   ```
   # scstat -g -D
   ```

`-g`	Show status for all resource groups
`-D`	Show status for all disk device groups

8. **Ensure that the system disk and data is backed up.**

9. **Does your cluster use dual-string mediators for Solstice DiskSuite/Solaris Volume Manager?**

   - If no, proceed to Step 10.
   - If yes, unconfigure your mediators.

     See "Configuring Dual-String Mediators" on page 146 for more information.

   a. **Run the following command to verify that no mediator data problems exist.**

      `# medstat -s` *setname*

      `-s` *setname* Specifies the diskset 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 148.

   b. **List all mediators.**

Use this information for when you restore the mediators during the procedure "How to Finish Upgrading to Sun Cluster 3.1 4/04 Software (Rolling)" on page 195.

c. For a diskset that uses mediators, take ownership of the diskset if no node already has ownership.

```
# metaset -s setname -t
```

`-t` Takes ownership of the diskset

d. Unconfigure all mediators for the diskset.

```
# metaset -s setname -d -m mediator-host-list
```

`-s` *setname*	Specifies the diskset name
`-d`	Deletes from the diskset
`-m` *mediator-host-list*	Specifies the name of the node ot remove as a mediator host for the diskset

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 diskset that uses mediators.

10. Shut down the node that you want to upgrade and boot it into noncluster mode.

```
# shutdown -y -g0
ok boot -x
```

The other nodes of the cluster continue to function as active cluster members.

11. Do you intend to upgrade the Solaris software to a Maintenance Update release?

Note – The cluster must already run on, or be upgraded to, at least the minimum required level of the Solaris operating environment to support Sun Cluster 3.1 4/04 software. See the *Sun Cluster Release Notes for Solaris OS* for information about supported releases of the Solaris operating environment.

- If yes, go to "How to Upgrade to a Solaris Maintenance Update Release (Rolling)" on page 191.
- If no, go to "How to Upgrade to Sun Cluster 3.1 4/04 Software (Rolling)" on page 192.

How to Upgrade to a Solaris Maintenance Update Release (Rolling)

Perform this procedure to upgrade the Solaris 8 or Solaris 9 operating environment to a supported Maintenance Update release.

Note – To upgrade a cluster from Solaris 8 to Solaris 9 software, with or without upgrading Sun Cluster software as well, you must perform a nonrolling upgrade. Go to "Upgrading to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 169.

---

1. **Ensure that all steps in "How to Prepare the Cluster for Upgrade (Rolling)" on page 188 are completed.**

2. **Temporarily comment out all entries for globally mounted file systems in the `/etc/vfstab` file.**

   Perform this step to prevent the Solaris upgrade from attempting to mount the global devices.

3. **Follow 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.

   ---

4. **Uncomment all entries in the `/a/etc/vfstab` file for globally mounted file systems that you commented out in Step 2.**

5. **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 6.

   ---

6. **Reboot the node into noncluster mode.**

   Include the double dashes (--) in the following command:

   ```
   # reboot -- -x
   ```

7. **Upgrade the Sun Cluster software.**

   Go to "How to Upgrade to Sun Cluster 3.1 4/04 Software (Rolling)" on page 192.

How to Upgrade to Sun Cluster 3.1 4/04 Software (Rolling)

Perform this procedure to upgrade a node to Sun Cluster 3.1 4/04 software while the remaining cluster nodes are in cluster mode.

Note – Do not use any new features provided in the Sun Cluster 3.1 4/04 software until all nodes of the cluster are upgraded.

---

1. **Ensure that all steps in "How to Prepare the Cluster for Upgrade (Rolling)" on page 188 are completed.**

   If you upgraded the Solaris operating environment to a Maintenance Update release, also ensure that all steps in "How to Upgrade to a Solaris Maintenance Update Release (Rolling)" on page 191 are completed.

2. **Upgrade to Sun Cluster 3.1 4/04 software.**

   a. **Insert the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM into the CD-ROM drive on the node.**

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

   b. **From the `/cdrom/cdrom0` directory, change to the `Solaris_sparc/Product/sun_cluster/Solaris_`*ver*`/Tools` directory, where *ver* is `8` (for Solaris 8) or `9` (for Solaris 9) .**

      The following example uses the path to the Solaris 8 version of Sun Cluster software.

      ```
      # cd /cdrom/cdrom0/Solaris_sparc/Product/sun_cluster/Solaris_8/Tools
      ```

   c. **Install the Sun Cluster 3.1 4/04 software.**

      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 Java Enterprise System 2004Q2 2 of 2 CD-ROM.

      ```
      # ./scinstall -u update -M patchdir=dirname
      ```

      `-M patchdir=`*dirname*`[,patchlistfile=`*filename*`]`
      Specifies the path to patch information so that the specified patches can be installed using the `scinstall` command. If you do not specify a patch-list file, the `scinstall` command installs all the patches in the directory *dirname*, including tarred, jarred, and zipped patches.

The `-M` option is not required. You can use any method you prefer for installing patches.

Upgrade processing is finished when the system displays the message `Completed Sun Cluster framework upgrade` and the path to the upgrade log.

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

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

Note – Sun Cluster 3.1 4/04 software requires at least version 3.5.1 of Sun Explorer software. Upgrade to Sun Cluster software includes installing Sun Explorer data collector software, to be used in conjunction with the `sccheck` utility. If another version of Sun Explorer software was already installed before Sun Cluster upgrade, it is replaced by the version that is provided with Sun Cluster software. Options such as user identity and data delivery are preserved, but `crontab` entries must be manually recreated.

---

d. Change to the CD-ROM root directory and eject the CD-ROM.

e. Install any Sun Cluster 3.1 4/04 software patches.

3. Do you intend to upgrade any data services?

- If yes, proceed to Step 4.
- If no, skip to Step 6.

4. Upgrade applications as needed.

Note – Do not upgrade an application if the newer version cannot coexist in the cluster with the older version.

Follow the instructions that are provided in your third-party documentation.

5. (Optional) For each node on which data services are installed, upgrade to the Sun Cluster 3.1 4/04 data-service update software.

Note – You must upgrade the Sun Cluster HA for Oracle 3.0 64–bit for Solaris 9 data service to the Sun Cluster 3.1 4/04 version. Otherwise, you can continue to use Sun Cluster 3.0 data services after upgrade to Sun Cluster 3.1 4/04 software.

---

a. Insert the Java Enterprise System Accessory CD 3 CD-ROM into the CD-ROM drive on the node.

b. Install the Sun Cluster 3.1 4/04 data-service–update patches.

Use one of the following methods:

- To upgrade one or more specified data services, type the following command.

```
# scinstall -u update -s srvc[,srvc,...] -d cdrom-image
```

`-u update`	Upgrades a cluster node to a later Sun Cluster software release
`-s` *srvc*	Upgrades the specified data service
`-d` *cdrom-image*	Specifies an alternate directory location for the CD-ROM image

- To upgrade all data services present on the node, type the following command.

```
# scinstall -u update -s all -d cdrom-image
```

`-s all`	Upgrades all data services

This command assumes that updates for all installed data services exist on the update release. If an update for a particular data service does not exist in the update release, that data service is not upgraded.

Upgrade processing is finished when the system displays the message `Completed upgrade of Sun Cluster data services agents` and the path to the upgrade log.

c. Change to the CD-ROM root directory and eject the CD-ROM.

d. Install any Sun Cluster 3.1 4/04 data-service software patches.

e. Verify that each data-service update patch is installed successfully.

View the upgrade log file that is referenced at the end of the upgrade output messages.

6. Reboot the node into the cluster.

```
# reboot
```

7. Run the following command on the upgraded node to verify that Sun Cluster 3.1 4/04 software is installed successfully.

```
# scinstall -pv
```

8. From any node, verify the status of the cluster configuration.

```
% scstat
% scrgadm -pv[v]
```

Output should be the same as for Step 4 in "How to Prepare the Cluster for Upgrade (Rolling)" on page 188.

9. **Do you have another node to upgrade?**
   - If yes, return to "How to Prepare the Cluster for Upgrade (Rolling)" on page 188 and repeat all upgrade procedures on the next node to upgrade.
   - If no, go to "How to Finish Upgrading to Sun Cluster 3.1 4/04 Software (Rolling)" on page 195.

How to Finish Upgrading to Sun Cluster 3.1 4/04 Software (Rolling)

1. **Ensure that all upgrade procedures are completed for all cluster nodes that you are upgrading.**
2. **Does your configuration use dual-string mediators for Solstice DiskSuite/Solaris Volume Manager?**
   - If no, proceed to Step 3.
   - If yes, restore the mediator configurations.

   a. **Determine which node has ownership of a diskset to which you will add the mediator hosts.**

      `# metaset -s` *setname*

      `-s` *setname* Specifies the diskset name

   b. **If no node has ownership, take ownership of the diskset.**

      `# metaset -s` *setname* `-t`

      `-t` Takes ownership of the diskset

   c. **Recreate the mediators.**

      `# metaset -s` *setname* `-a -m` *mediator-host-list*

      `-a` Adds to the diskset

      `-m` *mediator-host-list* Specifies the names of the nodes to add as mediator hosts for the diskset

   d. **Repeat Step a through Step c for each diskset in the cluster that uses mediators.**

3. **Are you using Sun Management Center to monitor your Sun Cluster configuration?**
   - If no, proceed to Step 4.

- If yes, perform the following steps to upgrade the Sun Cluster-module software on the Sun Management Center server machine, help-server machine, and console machine.

Note – Sun Cluster-module software on the cluster nodes was already upgraded during the cluster-framework upgrade.

If you intend to upgrade the Sun Management Center software itself, do not perform this step. Instead, proceed to Step 4 to finish Sun Cluster software upgrade. Then go to "How to Upgrade Sun Management Center Software" on page 201 to upgrade Sun Management Center software and the Sun Cluster module.

a. As superuser, remove the existing Sun Cluster–module packages.

Use the `pkgrm`(1M) command to remove all Sun Cluster–module packages from all locations listed in the following table.

```
# pkgrm module-package
```

Location	Module Package to Remove
Sun Management Center console machine	`SUNWscscn`
Sun Management Center server machine	`SUNWscssv`
Sun Management Center help-server machine	`SUNWscshl`

Note – Sun Cluster-module software on the cluster nodes was already upgraded during the cluster-framework upgrade.

---

b. As superuser, reinstall Sun Cluster–module packages from the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM to the locations listed in the following table.

In the CD-ROM path, the value for *ver* is `8` (for Solaris 8) or `9` (for Solaris 9).

```
# cd Solaris_sparc/Product/sun_cluster/Solaris_ver/Packages/
# pkgadd -d . module-package
```

Location	Module Package to Install
Sun Management Center console machine	`SUNWscshl`

Location	Module Package to Install
Sun Management Center server machine	SUNWscssv
Sun Management Center help-server machine	SUNWscshl

You install the help-server package SUNWscshl on both the console machine and the help-server machine. You do not upgrade to a new SUNWscscn package on the console machine.

4. **If you upgraded any data services that are not supplied on the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM or the Java Enterprise System Accessory CD 3 CD-ROM, register the new resource types for those data services.**

   Follow the documentation that comes with the data services.

5. **From any node, start the `scsetup`(1M) utility.**

   ```
   # scsetup
   ```

6. **Register the new resource types.**

   a. **From the Main Menu, choose `Resource groups`.**

   b. **Choose `Resource type registration`.**

   c. **Choose `Register all resource types which are not yet registered`.**

      The scsetup utility displays all resource types that are not registered.

      Follow the prompts to register the new resource types.

7. **Migrate all eligible resources to the new version of its resource type.**

   a. **From the Resource Group menu, choose `Change properties of a resource`.**

   b. **Choose `Manage resource versioning`.**

   c. **Choose `Show versioning status`.**

      The scsetup utility displays all resources for which a new version of its resource type was installed during the upgrade. Make note of which new resource types you will upgrade the resources to.

   d. **Choose `Re-version all eligible resources`.**

      Follow the prompts to upgrade eligible resources to the new version of its resource type.

   e. **Return to the Change properties of a resource menu.**

8. **Modify extension properties for new resource type versions.**

   a. **For each new resource type that you migrated existing resources to, determine whether the new resource type requires additional modifications to its extension properties.**

      Refer to each related data service manual for the requirements of each new resource type.

      Note – You do not need to change the `Type_version` property of a new resource type. That property was modified when you migrated resources to their new resource types in Step 7.

      - If no resource type requires additional modifications other than the `Type_version` property, go to Step 9.
      - If one or more resource types require additional modifications to extension properties, proceed to Step b.

   b. **From the Change properties of a resource menu, choose `Change extension resource properties`.**

   c. **Follow the prompts to modify the necessary extension properties.**

      Refer to your data service documentation for the names of extension properties and values to modify.

   d. **Repeat for each resource type that requires modifications.**

   e. **Return to the Resource Groups menu.**

9. **Re-enable all disabled resources.**

   a. **From the Resource Group Menu, choose `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.**

10. **Bring each resource group back online.**

    a. **From the Resource Group Menu, choose `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.**

11. **When all resource groups are back online, exit the `scsetup` utility.**

Type **q** to back out of each submenu, or press Ctrl-C.

The cluster upgrade is complete. You can now return the cluster to production.

12. **Restart any applications.**

    Follow the instructions that are provided in your third-party documentation.

    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 199
- "How to Resolve Mistaken Storage Changes During an Upgrade" on page 200

▼ 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.

1. **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.

2. **Become superuser on a node that is attached to the unverified device.**

3. **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.

4. **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

5. **Repeat Step 2 through Step 4 on all other nodes that are attached to the unverified device.**

6. **Return to the remaining upgrade tasks.**

   - For nonrolling upgrade, go to Step 11 in "How to Upgrade to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 177.
   - For rolling upgrade, go to Step 12 in "How to Prepare the Cluster for Upgrade (Rolling)" on page 188.

▼ 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 change the storage configuration back 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 procedures in "How to Handle Storage Reconfiguration During an Upgrade" on page 199.

---

1. **Change the storage topology back to its original configuration.**

   Check the configuration of the devices that were flagged as possibly being replaced, including the cabling.

2. **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.

3. **Did the `scdidadm` command return any error messages in Step 2?**

   - If no, proceed to Step 4.

- If yes, return to Step 1 to make further modifications to correct the storage configuration, then repeat Step 2.

4. **Return to the remaining upgrade tasks.**
   - For nonrolling upgrade, go to Step 11 in "How to Upgrade to Sun Cluster 3.1 4/04 Software (Nonrolling)" on page 177.
   - For rolling upgrade, go to Step 12 in "How to Prepare the Cluster for Upgrade (Rolling)" on page 188.

Sun Management Center Software Upgrade

This section describes how to upgrade from Sun Management Center 2.1.1 to either Sun Management Center 3.0 software or Sun Management Center 3.5 software on a Sun Cluster 3.1 4/04 configuration.

▼ How to Upgrade Sun Management Center Software

1. **Have available the following items:**
   - **Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM or the path to the CD-ROM image.** You use the CD-ROM to reinstall the Sun Cluster 3.1 4/04 version of the Sun Cluster–module packages after you upgrade Sun Management Center software.
   - **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 Release Notes* for the location of patches and installation instructions.

2. **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 from the menu bar.

   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
```

3. As superuser, remove Sun Cluster–module packages.

Use the `pkgrm`(1M) command to remove all Sun Cluster–module packages from all locations listed in the following table.

```
# pkgrm module-package
```

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 help-server machine	`SUNWscshl`

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 5, after you upgrade Sun Management Center software.

4. Upgrade the Sun Management Center software.

Follow the upgrade procedures in your Sun Management Center documentation.

5. As superuser, reinstall Sun Cluster–module packages from the Sun Java Enterprise System 2004Q2 2 of 2 CD-ROM to the locations listed in the table below.

In the CD-ROM path, the value of *ver* is **8** (for Solaris 8) or **9** (for Solaris 9)

```
# cd Solaris_sparc/Product/sun_cluster/Solaris_ver/Packages/
# pkgadd -d . module-package
```

Location	Module Package to Install
Each cluster node	`SUNWscsam`, `SUNWscsal`
Sun Management Center server machine	`SUNWscssv`
Sun Management Center console machine	`SUNWscshl`
Sun Management Center help-server machine	`SUNWscshl`

You install the help-server package `SUNWscshl` on both the console machine and the help-server machine.

6. **Apply any Sun Management Center patches and any Sun Cluster–module patches to each node of the cluster.**

7. **Restart Sun Management Center agent, server, and console processes.**

   Follow procedures in "SPARC: How to Start Sun Management Center" on page 112.

8. **Load the Sun Cluster module.**

   Follow procedures in "SPARC: How to Load the Sun Cluster Module" on page 113.

   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 Module⇒Unload Module from the console's Details window.

CHAPTER 6

Configuring Data Replication With Sun StorEdge Availability Suite 3.1 Software

This chapter provides guidelines for configuring data replication between clusters by using Sun StorEdge Availability Suite 3.1 software.

This chapter also contains an example of how data replication was configured for an NFS application by using Sun StorEdge Availability Suite 3.1 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:

The following procedures are in this chapter:

Introduction to Data Replication

This section introduces disaster tolerance and describes the data replication methods used by Sun StorEdge Availability Suite 3.1 software.

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 3.1 Software

This section describes the remote mirror replication method and the point-in-time snapshot method used by Sun StorEdge Availability Suite 3.1 software. This software uses the `sndradm`(1RPC) and `iiadm`(1II) commands to replicate data. For more information about these commands, see the *Sun Cluster 3.0 and Sun StorEdge 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.

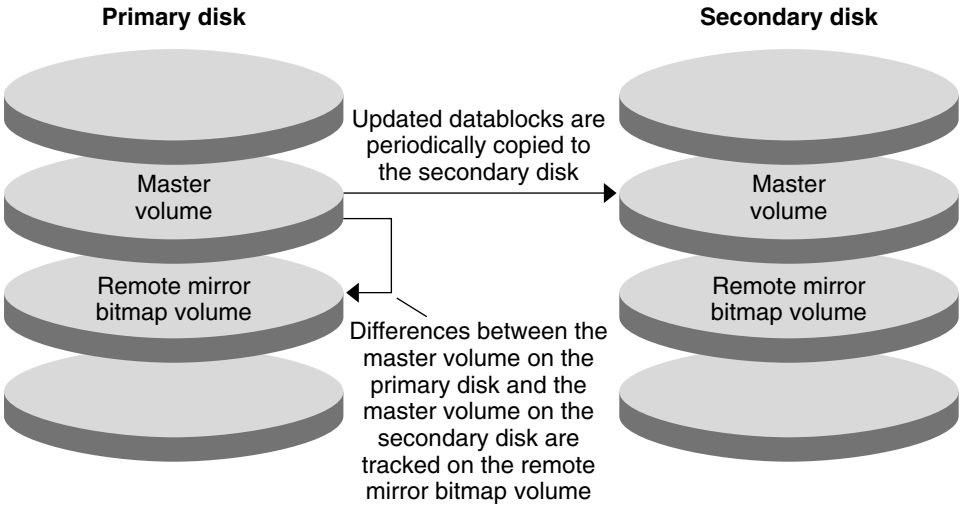

FIGURE 6–1 Remote Mirror Replication

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.

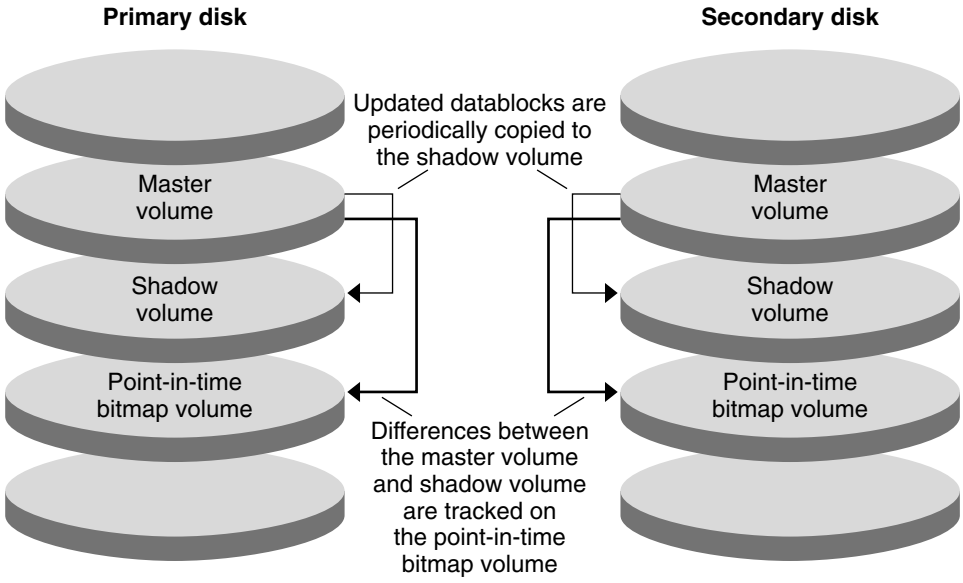

FIGURE 6–2 Point-in-Time Snapshot

Replication in the Example Configuration

The following figure illustrates how remote mirror replication and point-in-time snapshot are used in "Example Configuration" on page 214.

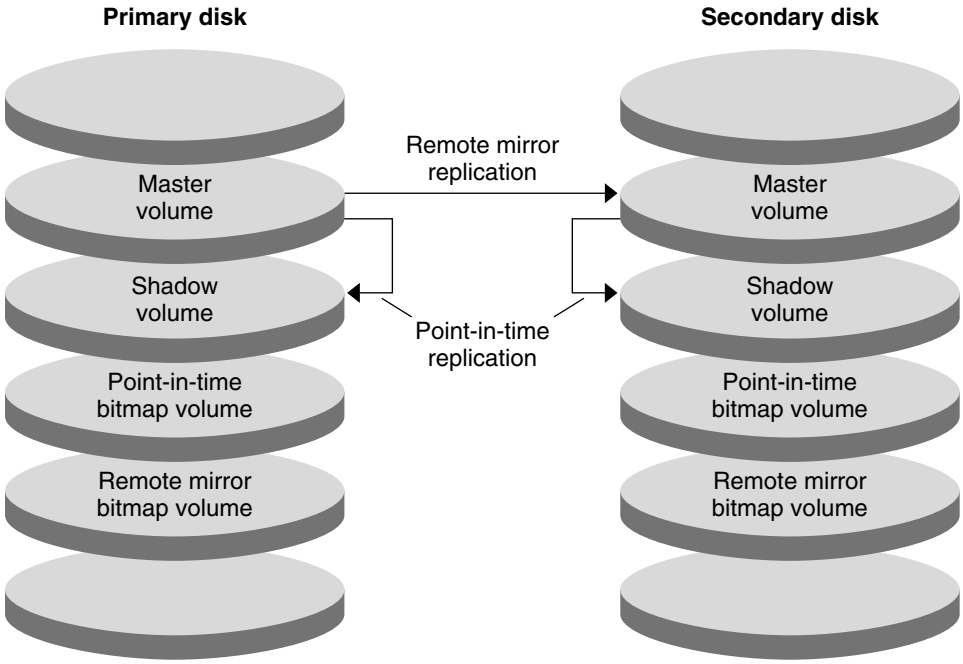

FIGURE 6–3 Replication in the Example Configuration

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

Replication resource groups collocate the device group under Sun StorEdge Availability Suite 3.1 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 collocated, 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 collocated, 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" on page 211
- "Configuring Resource Groups for a Scalable Application" on page 212

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 collocated 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 compulsory but is advised.

  – If the application data is mounted locally, the presence of an HAStoragePlus resource in the application resource group is compulsory.

    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.

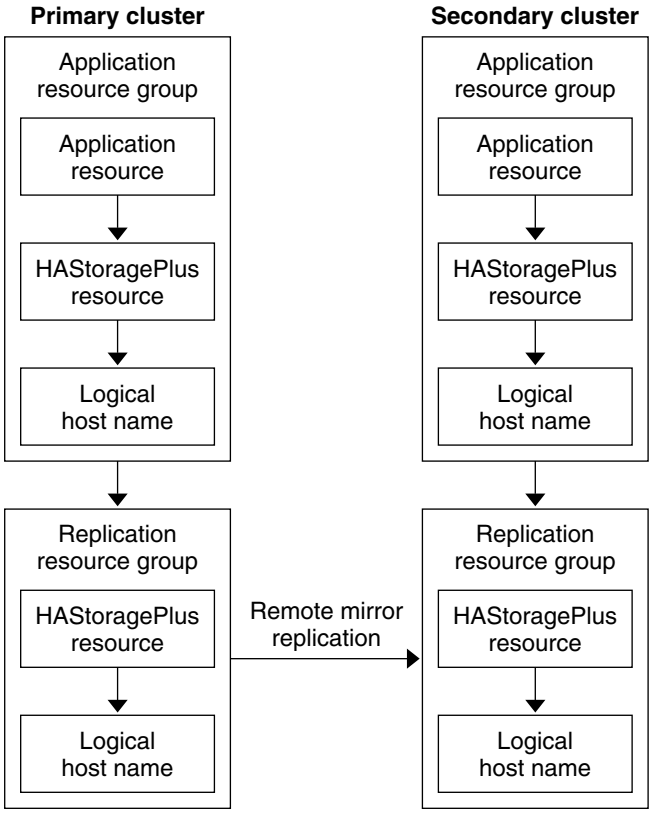

FIGURE 6–4 Configuration of Resource Groups in a Failover Application

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 collocate 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.

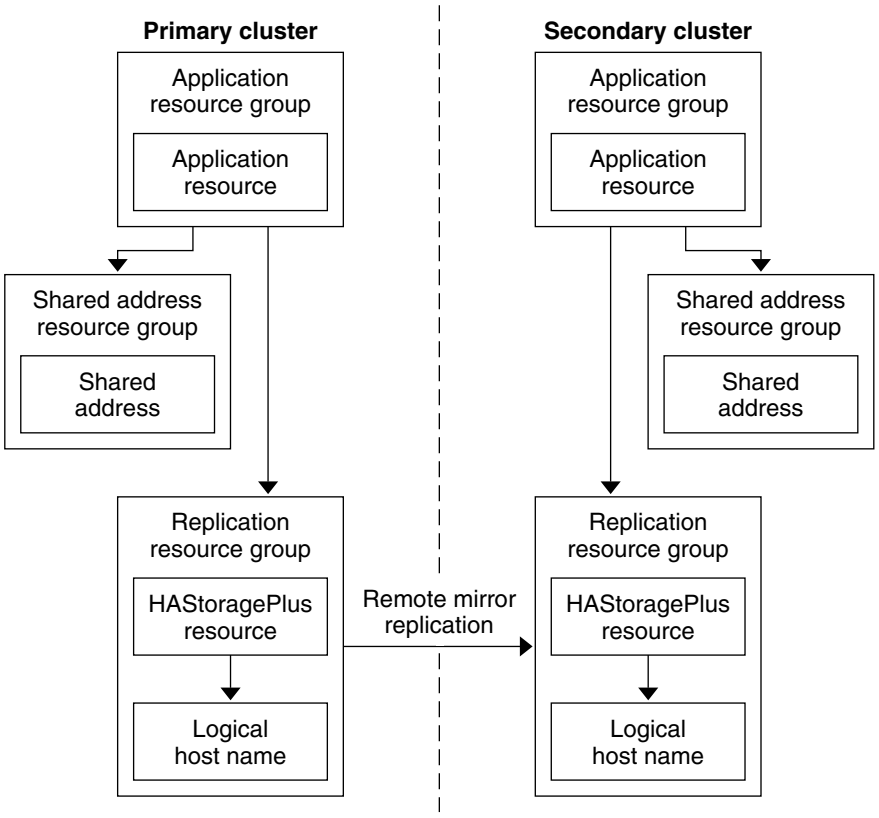

FIGURE 6–5 Configuration of Resource Groups in a Scalable Application

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. In addition, the secondary volume must be mounted on the mount point directory for the application file system.

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.

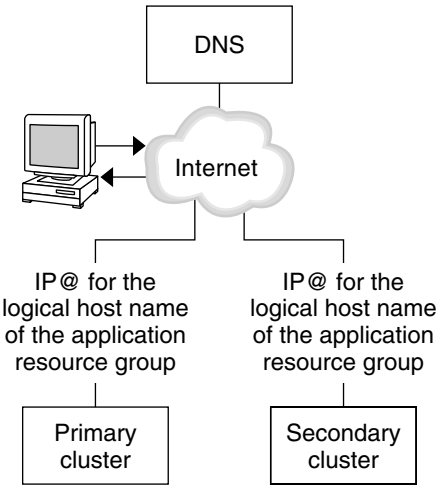

FIGURE 6–6 DNS Mapping of a Client to a Cluster

To update the DNS, use the `nsupdate` command. For information, see the `nsupdate`(1M) man page. For an example of how to cope with a failover or switchover, see "Example of How to Cope With a Failover or Switchover" on page 232.

After repair, the primary cluster can be brought back online. To switch back to the original primary cluster, perform the following steps:

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.
3. Mount the primary volume on to the mount point directory for the application file system.

Example Configuration

This section provides a step-by-step example of how data replication was configured for an NFS application by using Sun StorEdge Availability Suite 3.1 software.

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.

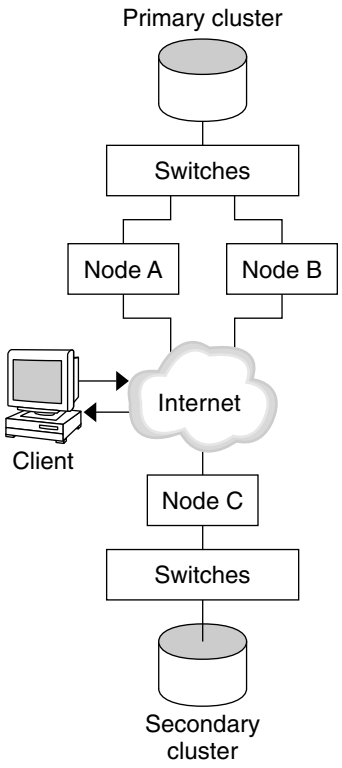

FIGURE 6–7 Example Cluster Configuration

Table 6–1 summarizes the hardware and software required by the example configuration. The operating environment, Sun Cluster software, and volume manager software must be installed on the cluster nodes *before* you install Sun StorEdge Availability Suite 3.1 software and patches.

TABLE 6–1 Required Hardware and Software

Hardware or Software	Requirement
Node hardware	Sun StorEdge Availability Suite 3.1 software is supported on all servers using the Solaris operating environment. For information about which hardware to use, see the *Sun Cluster 3.x Hardware Administration Manual*.
Disk space	Approximately 11 Mbytes.
Operating environment	Solaris 8 or Solaris 9 releases that are supported by Sun Cluster software. All nodes must use the same version of operating environment. For information about installation, see"Installing the Software" on page 40.
Sun Cluster software	Sun Cluster 3.1 4/04 software. For information about installation, see Chapter 2 and "How to Install Sun Cluster Software on a Single-Node Cluster" on page 79.
Volume manager software	Solstice DiskSuite/Solaris Volume Manager or VERITAS Volume Manager (VxVM). All nodes must use the same version of volume manager software. Information about installation is in"Installing and Configuring Solstice DiskSuite/Solaris Volume Manager Software" on page 116 and "SPARC: Installing and Configuring VxVM Software" on page 149.
Sun StorEdge Availability Suite 3.1 software	For information about how to install the software, see the *Sun StorEdge Availability Suite 3.1 Point-in-Time Copy Software Installation Guide* and the *Sun StorEdge Availability Suite 3.1 Remote Mirror Software Installation Guide*.
Sun StorEdge Availability Suite 3.1 software patches	For information about the latest patches, see `http://sunsolve.sun.com`.

Example of How to Configure Device Groups and Resource Groups

This chapter describes how disk device groups and resource groups were configured for an NFS application. The following table lists the names of the groups and resources that were created for the example configuration.

TABLE 6–2 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`.

Configuring Disk Device Groups

This section describes how to configure a disk device group on the primary cluster and the secondary cluster. This example configuration uses VxVM software. For information about Solstice DiskSuite/Solaris Volume Manager software, see Chapter 3.

The following figure illustrates the volumes that were created in the disk device group.

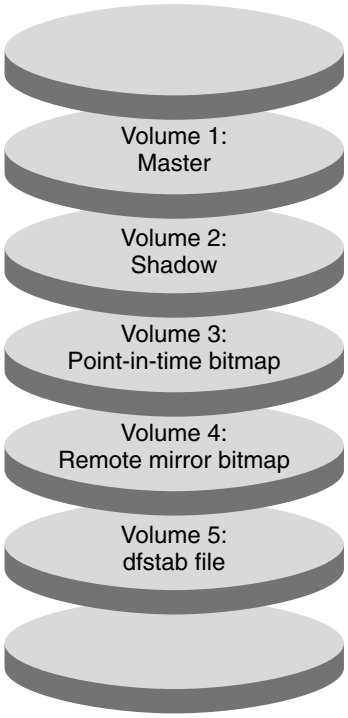

FIGURE 6–8 Volumes for the Disk Device Group

Note – The volumes defined in this section 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*

1. **Create a disk group that contains four volumes, volume 1 through volume 4.**

   For information about configuring a disk group by using the VxVM software, see Chapter 4.

2. **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.

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. **Start the disk device group.**

   nodeA# **/usr/cluster/bin/scswitch -z -D** *devicegroup* **-h** *nodeA*

5. **Synchronize the disk device group with the Sun Cluster software.**

   nodeA# **/usr/cluster/bin/scconf -c -D name=***devicegroup***,sync**

6. **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**
   nodeA# **/usr/sbin/newfs /dev/vx/rdsk/***devicegroup***/vol03 < /dev/null**
   nodeA# **/usr/sbin/newfs /dev/vx/rdsk/***devicegroup***/vol04 < /dev/null**

7. **Enable remote access between the nodes in the primary cluster and secondary cluster by adding the following entities to the `/.rhosts` file on `nodeA` and `nodeB`.**

   ```
   nodeC +
   + root
   ```

▼ *How to Configure a Disk Device Group on the Secondary Cluster*

- **Follow the procedure in "How to Configure a Disk Device Group on the Primary Cluster" on page 218, with these exceptions:**
  - Replace `nodeA` with `nodeC`.
  - Do not use `nodeB`.
  - In Step 3, include `nodeC` only in the node list. For example:

    nodeC# **/usr/cluster/bin/scconf -a -D type=vxvm,name=***devicegroup* \
    ,nodelist=*nodeC*

  - In Step 7, add the following entities to the `/.rhosts` file on `nodeC` only:

    ```
    nodeA +
    nodeB +
    + root
    ```

Configuring File Systems for the NFS Application

This section describes how the file systems were configured for the NFS application.

▼ *How to Configure the File System on the Primary Cluster for the NFS Application*

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 Sun StorEdge Availability Suite 3.1 software.**

   ```
   nodeA# /usr/sbin/vxassist -g devicegroup make vol05 120m disk1
   ```

   Volume 5 contains the file system information that is used by Sun StorEdge Availability Suite 3.1 software.

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 volume 5.**

   ```
   nodeA# /usr/sbin/newfs /dev/vx/rdsk/devicegroup/vol05
   ```

6. **On `nodeA` and `nodeB`, create a mount point for volume 5.**

   For example:

   ```
   nodeA# mkdir /global/etc
   ```

7. **On `nodeA` and `nodeB`, configure volume 5 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 volume 5 on `nodeA`.**

   ```
   nodeA# mount /global/etc
   ```

9. **Make volume 5 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
```

▼ *How to Configure the File System on the Secondary Cluster for the NFS Application*

- **Repeat the procedure in "How to Configure the File System on the Primary Cluster for the NFS Application" on page 219, with these exceptions:**
  - Replace `nodeA` with `nodeC`.
  - Do not use `nodeB`.

Creating Replication Resource Groups

This section describes how a replication resource group was created on the primary cluster and on the secondary cluster.

▼ *How to Create a Replication Resource Group on the Primary Cluster*

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 3.1 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
```

Where *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 for `nodeA` and `nodeB`.

▼ *How to Create a Replication Resource Group on the Secondary Cluster*

- **Repeat the procedure in "How to Create a Replication Resource Group on the Primary Cluster" on page 221, with these exceptions:**
  - Replace `nodeA` with `nodeC`.
  - Do not use `nodeB`.
  - Replace references to `lhost-reprg-prim` with `lhost-reprg-sec`.

Creating Application Resource Groups

This section describes how application resource groups were created for an NFS application. The procedures in this section are specific to the application. The procedures cannot be used for another type of application.

▼ *How to Create an Application Resource Group on the Primary Cluster*

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
   Is 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
```

Where *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 Application Resource Group on the Secondary Cluster*

1. **Create the application group resource as described in Step 1 through Step 6 of "How to Create an Application Resource Group on the Primary Cluster" on page 223, with the following exceptions:**

   - Replace `nodeA` with `nodeC`.
   - Ignore references to `nodeB`.
   - Replace references to *lhost-nfsrg-prim* with *lhost-nfsrg-sec*.

2. **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*.

3. **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.

Example of How to Enable Data Replication

This section describes how data replication was enabled for the example configuration. This section uses the Sun StorEdge Availability Suite 3.1 software commands `sndradm` and `iiadm`. For more information about these commands, see the *Sun Cluster 3.0 and Sun StorEdge Software Integration Guide*.

▼ How to Enable Replication on the Primary Cluster

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**

   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 of the primary cluster to the master volume of the secondary cluster. In addition, this step enables replication to the remote mirror bitmap on volume 4.

   - 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 look like this:

```
/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 \
/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 of the primary disk to be copied to the shadow volume on the same disk. In this example, the master volume is volume 1, the shadow volume is volume 2, and the point-in-time bitmap volume is volume 3.

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 3.1 software ensures that a point-in-time snapshot is taken before remote mirror replication can occur.

▼ How to Enable Replication on the Secondary Cluster

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`**

Example of How to Perform Data Replication

This section describes how data replication was performed for the example configuration. This section uses the Sun StorEdge Availability Suite 3.1 software commands `sndradm` and `iiadm`. For more information about these commands, see the *Sun Cluster 3.0 and Sun StorEdge Software Integration Guide*.

▼ 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 volume 1 and the remote mirror bitmap volume is volume 4.

1. **Access `nodeA` as superuser.**

2. **Verify that the cluster is in logging mode.**

   nodeA# **/usr/opt/SUNWesm/sbin/sndradm -P**

   The output should look like this:

   /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 look like this:

   /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 3.1 software.

▼ How to Perform a Point-in-Time Snapshot

In this procedure, point-in-time snapshot was used to synchronize the shadow volume of the primary cluster to the master volume of the primary cluster. The master volume is volume 1 and the shadow volume is volume 2.

1. **Access `nodeA` as superuser.**

2. **Quiesce the application that is running on `nodeA`.**

   nodeA# **`/usr/cluster/bin/scswitch -n -j`** *nfs-rs*

3. **Put the primary cluster in 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`**

Example of How to Verify That Replication Is Configured Correctly

This section describes how the replication configuration was confirmed in the example configuration.

▼ How to Verify That Replication Is Configured Correctly

1. **Verify that the primary cluster is in replicating mode, with autosynchronization on.**

   nodeA# **/usr/opt/SUNWesm/sbin/sndradm -P**

   The output should look like this:

   /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 3.1 software.

   If the primary cluster is not in replicating mode, put it in to replicating mode, as follows:

   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**

2. **Make a directory on a client machine.**

   a. **Log in to a client machine as superuser.**

      You see a prompt like this:

      client-machine#

   b. **Make a directory on the client machine.**

      client-machine# **mkdir /***dir*

3. **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*

4. **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. Put the primary cluster in 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. Bring the application resource group online on the secondary cluster.

```
nodeC# /usr/cluster/bin/scswitch -Z -g nfs-rg
```

e. Access the client machine as superuser.

You see a prompt like this:

```
client-machine#
```

f. Mount the directory that was created in Step 2 to the application on the secondary cluster.

```
client-machine# mount -o rw lhost-nfsrg-sec:/global/mountpoint /dir
```

g. Display the mounted directory.

```
client-machine# ls /dir
```

5. Ensure that the directory displayed in Step 3 is the same as that displayed in Step 4.

6. 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. Put the primary cluster in to 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
```

When the primary volume is written to, the secondary volume is updated by Sun StorEdge Availability Suite 3.1 software.

Example of How to Cope With a Failover or Switchover

This section describes how a switchover was provoked and how the application was transferred to the secondary cluster. After a switchover or failover, you must update the DNS entry and configure the application to read and write to the secondary volume.

▼ How to Provoke a Switchover

1. Put the primary cluster into 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.

2. Confirm that the primary cluster and the secondary cluster are in logging mode, with autosynchronization off.

a. On `nodeA`, run this command:

```
nodeA# /usr/opt/SUNWesm/sbin/sndradm -P
```

The output should look like this:

```
/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`, run this command:

```
nodeC# /usr/opt/SUNWesm/sbin/sndradm -P
```

The output should look like this:

```
/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*
   nodeC# **scswitch -Z -g** *nfs-rg* **-h** *nodeC*

▼ How to Update the DNS Entry

For an illustration of how the DNS maps a client to a cluster, see Figure 6–6.

1. **Start the nsupdate command.**

   For information, see the nsupdate(1M) man page.

2. **Remove the current DNS mapping between the client machine and the logical hostname of the application resource group on the primary cluster.**

   > **update delete** *client-machine* **A**
   > **update delete** *IPaddress1***.in-addr.arpa** *TTL* **PTR** *client machine*

   client-machine
   Is the full name of the client. For example, mymachine.mycompany.com.

   IPaddress1
   Is the IP address is of the logical hostname lhost-nfsrg-prim, in reverse order.

   TTL
   Is the time to live, in seconds. A typical value is 3600.

3. **Create the new DNS mapping between the client machine and the logical hostname of the application resource group on the secondary cluster.**

   > **update add** *client-machine TTL* **A** *IPaddress2*
   > **update add** *IPaddress3***.in-addr.arpa** *TTL* **PTR** *client-machine*

   IPaddress2
   Is the IP address is of the logical hostname lhost-nfsrg-sec, in forward order.

   IPaddress3
   Is the IP address is of the logical hostname lhost-nfsrg-sec, in reverse order.

▼ How to Configure the Application to Read and Write to the Secondary Volume

1. **Configure the secondary volume to be mounted to the mount point directory for the NFS file system.**

   client-machine# **mount -o rw** *lhost-nfsrg-sec*:/global/*mountpoint* **/xxx**

   The mount point was created in Step 1 of "How to Configure the File System on the Primary Cluster for the NFS Application" on page 219.

2. **Confirm that the secondary cluster has write access to the mount point.**

   client-machine# **touch /xxx/data.1**
   client-machine# **umount /xxx**

APPENDIX **A**

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 Service 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 238	"Example: Local File System Layout Worksheets, With and Without Mirrored Root" on page 239	"System Disk Partitions" on page 18 "Guidelines for Mirroring the Root Disk" on page 37
"Public Networks Worksheet" on page 240	"Example: Public Networks Worksheet" on page 241	"Public Networks" on page 23 "IP Network Multipathing Groups" on page 27
"Local Devices Worksheets" on page 242	"Example: Local Devices Worksheets" on page 243	---
"Disk Device Group Configurations Worksheet" on page 244	"Example: Disk Device Group Configurations Worksheet" on page 245	"Disk Device Groups" on page 29 "Planning Volume Management" on page 31
"Volume-Manager Configurations Worksheet" on page 246	"Example: Volume-Manager Configurations Worksheet" on page 247	"Planning Volume Management" on page 31 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/Solaris Volume Manager)" on page 248	"Example: Metadevices Worksheet (Solstice DiskSuite/Solaris Volume Manager)" on page 249	"Planning Volume Management" on page 31 *Solstice DiskSuite 4.2.1 Installation and Product Notes* or *Solaris Volume Manager Administration Guide*

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/ Diskset Name	Node Names (indicate priority if ordered list)	Ordered priority? (circle one)	Failback? 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/ Diskset Name	Node Names (indicate priority if ordered list)	Ordered priority? (circle one)	Failback? circle one)
`dg-schost-1`	`1) phys-schost-1,` `2) phys-schost-2`	**`Yes`**	**`Yes`**
		Yes ǀ No	Yes ǀ No
		Yes ǀ No	Yes ǀ No
		Yes ǀ No	Yes ǀ No
		Yes ǀ No	Yes ǀ No
		Yes ǀ No	Yes ǀ No
		Yes ǀ No	Yes ǀ No
		Yes ǀ No	Yes ǀ No
		Yes ǀ No	Yes ǀ No
		Yes ǀ No	Yes ǀ No
		Yes ǀ No	Yes ǀ No
		Yes ǀ No	Yes ǀ 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/Solaris Volume Manager)

TABLE A–16 Metadevices Worksheet (Solstice DiskSuite/Solaris Volume Manager)

File System	Metatrans	Metamirrors		Submirrors		Hot Spare Pool	Physical Device	
		(Data)	(Log)	(Data)	(Log)		(Data)	(Log)

Example: Metadevices Worksheet (Solstice DiskSuite/Solaris Volume Manager)

TABLE A–17 Example: Metadevices Worksheet (Solstice DiskSuite/Solaris Volume Manager)

		Metamirrors		Submirrors			Physical Device	
File System	**Metatrans**	**(Data)**	**(Log)**	**(Data)**	**(Log)**	**Hot Spare Pool**	**(Data)**	**(Log)**
`/A`	`d10`	`d11`		`d12, d13`		`hsp000`	`c1t0d0s0, c2t0d1s0`	
			`d14`		`d15`	`hsp006`		`c1t0d1s6, c2t1d1s6`

Index

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

X