Sun Cluster 3.0 U1 Installation Guide

Sun Microsystems, Inc.
901 San Antonio Road
Palo Alto, CA 94303-4900 U.S.A.
650-960-1300

Part No. 806-7069-10
July 2001, Revision A

Contents

Preface

The *Sun Cluster 3.0 U1 Installation Guide* contains guidelines for planning a Sun™ Cluster 3.0 configuration, and provides procedures for installing, upgrading, and configuring the Sun Cluster software.

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 used with Sun Cluster software.

Using UNIX Commands

This document contains information on commands 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.

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

Typographic Conventions

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

Shell Prompts

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

Related Documentation

Application	Title	Part Number
Hardware	*Sun Cluster 3.0 U1 Hardware Guide*	806-7070
Data Services	*Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide*	805-7071
API Development	*Sun Cluster 3.0 U1 Data Services Developers' Guide*	805-7072
Administration	*Sun Cluster 3.0 U1 System Administration Guide*	806-7073
Concepts	*Sun Cluster 3.0 U1 Concepts*	806-7074
Error Messages	*Sun Cluster 3.0 U1 Error Messages Manual*	806-7076
Release Notes	*Sun Cluster 3.0 U1 Release Notes*	805-7078

Accessing Sun Documentation Online

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

`http://docs.sun.com`

Ordering Sun Documentation

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

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

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

Getting Help

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

- Your name and email address (if available)
- Your company name, address, and phone number
- The model 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 3.0)

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

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

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

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 Sun Cluster software installation tasks and the order in which you should perform them.

TABLE 1-1 Location of Sun Cluster Software Installation Task Information (Sheet 1 of 2)

Task	For Instructions, Go To ...
Set up cluster hardware.	*Sun Cluster 3.0 U1 Hardware Guide* Documentation shipped with your server and storage devices
Plan cluster software installation.	This chapter "Sun Cluster Installation Configuration Worksheets and Examples" in the *Sun Cluster 3.0 U1 Release Notes*
Install a new cluster or add nodes to an existing cluster.	
Install Solaris operating environment, Cluster Control Panel (optional), SunPlex Manager (optional), cluster framework, and data service software packages.	"Installing the Software" on page 24
Install and configure volume manager software.	
Solstice DiskSuite	"Installing and Configuring Solstice DiskSuite Software" on page 122 Solstice DiskSuite documentation
VERITAS Volume Manager (VxVM)	"Installing and Configuring VxVM Software" on page 162 VxVM documentation
Configure cluster framework software and optionally install and configure Sun Management Center.	"Configuring the Cluster" on page 78
Plan, install, and configure resource groups and data services.	*Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide* "Data Service Configuration Worksheets and Examples" in the *Sun Cluster 3.0 U1 Release Notes*

TABLE 1-1 Location of Sun Cluster Software Installation Task Information (Sheet 2 of 2)

Task	For Instructions, Go To ...
Upgrade Solaris operating environment, cluster framework, data services, and volume manager software from Sun Cluster 2.2 to Sun Cluster 3.0.	"Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 Update 1 Software" on page 96 "Installing and Configuring Solstice DiskSuite Software" on page 122 or "Installing and Configuring VxVM Software" on page 162 Volume manager documentation
Develop custom data services.	*Sun Cluster 3.0 U1 Data Services Developers' Guide*

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 the 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 JumpStart. If you are installing several cluster nodes, consider a network installation.

See "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 59 for details about the `scinstall` JumpStart installation method. See the Solaris installation documentation for details about standard Solaris installation methods.

System Disk Partitions

Add this information to the "Local File System Layout Worksheet" in the *Sun Cluster 3.0 U1 Release Notes*.

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 –** Allocate at least 750 Mbytes or twice the physical memory, whichever is greater.
- `/globaldevices` **–** Create a 100-Mbyte file system that will be used by the `scinstall`(1M) utility for global devices.
- **Volume manager –** Create a 10-Mbyte partition for volume manager use on a slice at the end of the disk (slice 7). If your cluster uses VERITAS Volume Manager (VxVM) and you intend to encapsulate the root disk, you need 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™ software requires less than 5 Mbytes, and VxVM software requires less than 15 Mbytes. For best results, you need to configure ample additional space and inode capacity for the creation of both block special devices and character special devices used by either Solstice DiskSuite or VxVM software, especially if a large number of shared disks are in the cluster. Therefore, add at least 100 Mbytes to the amount of space you would normally allocate for your root (/) file system.
- /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 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 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. If you use Sun™ Management Center software (formerly named Sun Enterprise SyMON™) 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 swap Partition

The minimum size of the swap partition must be either 750 Mbytes or twice the amount of physical memory on the machine, whichever is greater. In addition, any third-party applications you install might also have swap requirements. See your third-party application documentation for any swap requirements.

Guidelines for the `/globaldevices` File System

Sun Cluster software requires that you set aside a special file system on one of the local disks for use in managing global devices. This file system must be separate, as it will later be mounted as a cluster file system. Name this file system `/globaldevices`, which is the default name recognized by the `scinstall`(1M) command. The `scinstall`(1M) command later renames the file system `/global/.devices/node@`*nodeid*, where *nodeid* represents the number assigned to a node when it becomes a cluster member, and the original `/globaldevices` mount point is removed.

The `/globaldevices` file system must have ample space and inode capacity for creating both block special devices and character special devices, especially if a large number of disks are in the cluster. A file system size of 100 Mbytes should be more than enough for most cluster configurations.

Volume Manager Requirements

If you use Solstice DiskSuite software, you must set aside a slice on the root disk for use in creating the replica database. 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 replica databases in the same slice for Solstice DiskSuite software to function properly. See the Solstice DiskSuite documentation for more information.

If you use VxVM and you intend to encapsulate the root disk, you need two unused slices available for use by VxVM, as well as some additional unassigned free space at either the beginning or the end of the disk. See the 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 will be installed with the Solaris operating environment End User System Support 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 software or VxVM. If you use Solstice DiskSuite software, you use slice 7 for the replica database. If you use VxVM, you later free slice 7 by assigning it a zero length. This layout provides the necessary two free slices, 4 and 7, and it provides for unused space at the end of the disk.

TABLE 1-2 Sample File System Allocation

Slice	Contents	Allocation (in Mbytes)	Description
0	`/`	1168	441 Mbytes for Solaris operating environment software. 100 Mbytes extra for root (`/`). 100 Mbytes extra for `/var`. 25 Mbytes for Sun Cluster software. 55 Mbytes for volume manager software. 1 Mbyte for Sun Cluster HA for NFS software. 25 Mbytes for the Sun Management Center agent and Sun Cluster module agent packages. 421 Mbytes (the remaining free space on the disk) for possible future use by database and application software.
1	swap	750	Minimum size when physical memory is less than 750 Mbytes.
2	overlap	2028	The entire disk.
3	`/globaldevices`	100	The Sun Cluster software later assigns this slice a different mount point and mounts it 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	10	Used by Solstice DiskSuite software for the replica database, or used by VxVM for installation after you free the slice.

Planning the Sun Cluster Environment

This section provides guidelines for planning and preparing for Sun Cluster software installation. For detailed information about Sun Cluster components, see *Sun Cluster 3.0 U1 Concepts*.

Licensing

Ensure that you have any necessary license certificates available 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 the *Sun Cluster 3.0 U1 Release Notes* or consult your Enterprise Services representative or service provider. See the *Sun Cluster 3.0 U1 System Administration Guide* for general guidelines and procedures for applying patches.

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 to them. Add these IP addresses to any naming services used. Also add these IP addresses to the local `/etc/inet/hosts` file on each cluster node after Solaris software is installed.

TABLE 1-3 Sun Cluster Components That Use IP Addresses

Component	IP Addresses Needed
Administrative console	1 per subnet
Cluster nodes	1 per node, per subnet
Terminal concentrator or System Service Processor	1
Logical addresses	1 per logical host resource, per subnet

Terminal Concentrator or System Service Processor

A terminal concentrator communicates between the administrative console and the cluster node consoles. Sun Enterprise™ E10000 servers use a System Service Processor (SSP) instead of a terminal concentrator. For more information about console access, see *Sun Cluster 3.0 U1 Concepts*.

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. See the *Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide* for information and worksheets for planning resource groups. For more information about data services and resources, also see *Sun Cluster 3.0 U1 Concepts*.

Sun Cluster Configurable Components

This section provides guidelines for the Sun Cluster components that you configure during installation.

Cluster Name

Add this planning information to the "Cluster and Node Names Worksheet" in the *Sun Cluster 3.0 U1 Release Notes*.

You specify a name for the cluster during Sun Cluster installation. The cluster name should be unique throughout the enterprise.

Node Names

Add this planning information to the "Cluster and Node Names Worksheet" in the *Sun Cluster 3.0 U1 Release Notes*. Information for most other worksheets is grouped by node name.

The node name is the name you assign to a machine when you install the Solaris operating environment. During Sun Cluster installation, you specify the names of all nodes that you are installing as a cluster.

Private Network

Add this planning information to the "Cluster and Node Names Worksheet" in the *Sun Cluster 3.0 U1 Release Notes*.

Sun Cluster software uses the private network for internal communication between nodes. Sun Cluster requires at least two connections to the cluster interconnect on the private network. You specify the private network address and netmask when you install 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 you have successfully installed the node as a cluster member, you cannot change the private network address and netmask.

If you specify a private network address other than the default, it 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

  See the *TCP/IP and Data Communications Administration Guide* for instructions on obtaining copies of RFCs.

If you specify a netmask other than the default, it must meet the following requirements.

- Minimally mask all bits given in the private network address
- Have no "holes"

Private Hostnames

Add this planning information to the "Cluster and Node Names Worksheet" in the *Sun Cluster 3.0 U1 Release Notes*.

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

Cluster Interconnect

Add this planning information to the "Cluster Interconnect Worksheet" in the *Sun Cluster 3.0 U1 Release Notes*.

The cluster interconnect provides the hardware pathway for private network communication between cluster nodes. Each interconnect consists of a cable connected between two transport adapters, a transport adapter and a transport junction, or two transport junctions. During Sun Cluster installation, you specify the following configuration 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. If you specify a transport junction, it will be easier to add another node to the cluster in the future.
- **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 automatically assigned during installation, or create other names.

  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.

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 installation by using the `scsetup`(1M) utility.

For more information about the cluster interconnect, see *Sun Cluster 3.0 U1 Concepts*.

Public Networks

Add this planning information to the "Public Networks Worksheet" in the *Sun Cluster 3.0 U1 Release Notes*.

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 `false`. Sun Cluster software does not support a `local-mac-address` value of `true`.

See also "NAFO Groups" on page 13 for guidelines on planning public network adapter backup groups. For more information about public network interfaces, see *Sun Cluster 3.0 U1 Concepts*.

Disk Device Groups

Add this planning information to the "Disk Device Group Configurations Worksheet" in the *Sun Cluster 3.0 U1 Release Notes*.

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 so 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 20 for additional guidelines. See "Installing and Configuring Solstice DiskSuite Software" on page 122 or "Installing and Configuring VxVM Software" on page 162 and your volume manager documentation for instructions on mirroring.

For more information about disk device groups, see *Sun Cluster 3.0 U1 Concepts*.

NAFO Groups

Add this planning information to the "Public Networks Worksheet" in the *Sun Cluster 3.0 U1 Release Notes*.

A Network Adapter Failover (NAFO) group provides public network adapter monitoring and failover, and is the foundation for a network address resource. If a NAFO group is configured with two or more adapters and the active adapter fails, all of the NAFO group's addresses fail over to another adapter in the NAFO group. In this way, the active NAFO group adapter maintains public network connectivity to the subnet to which the adapters in the NAFO group connect.

Consider the following points when you plan your NAFO groups.

- Each public network adapter must belong to a NAFO group.
- Each node can have only one NAFO group per subnet.
- No more than one adapter in a given NAFO group can have a hostname association, in the form of an `/etc/hostname.`*adapter* file.
- The NAFO group naming convention is `nafo`*N*, where *N* is the number you supply when you create the NAFO group.

For more information about Network Adapter Failover, see *Sun Cluster 3.0 U1 Concepts*.

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.

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 so that the quorum devices have completely independent failure pathways.
- **Connection –** Do not connect a quorum device to more than two nodes.

For more information about quorum, see *Sun Cluster 3.0 U1 Concepts*.

Planning the Global Devices and Cluster File Systems

This section provides guidelines for planning global devices and cluster file systems. For more information about global devices and cluster files systems, see *Sun Cluster 3.0 U1 Concepts*.

Guidelines for Highly Available Global Devices and Cluster File Systems

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

- **Mirroring –** You must mirror all global devices for the global device to be considered highly available.
- **Disks –** When you mirror, lay out disks so that they 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.

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 in the `/global` directory, unless prohibited by other software products. Using a `/global` directory enables you to easily distinguish cluster file systems, which are globally available, from local file systems.
- **Nesting mount points –** Normally, you should not nest the mount points for cluster file systems. For example, do not set up one file system mounted on `/global/a` and another file system mounted on `/global/a/b`. Ignoring this rule can cause availability and node boot order problems 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 (for example, different slices on the same disk).

Planning Volume Management

Add this planning information to the "Disk Device Group Configurations Worksheet" and the "Volume Manager Configurations Worksheet" in the *Sun Cluster 3.0 U1 Release Notes*. For Solstice DiskSuite, also add this planning information to the "Metadevices Worksheet (Solstice DiskSuite)."

This section provides guidelines for planning volume management of your cluster configuration.

Sun Cluster uses volume manager software to group disks into disk device groups that can then be administered as one unit. Sun Cluster supports Solstice DiskSuite software and VERITAS Volume Manager (VxVM).

- If you use Solstice DiskSuite software, you must install it on all nodes of the cluster, regardless of whether you use VxVM on some nodes to manage disks.
- If you use VxVM and enable the VxVM cluster feature, you must install and license VxVM on all nodes of the cluster.
- If you use VxVM and do *not* enable the VxVM cluster feature, you only need to install and license VxVM on the nodes which are attached to storage devices which VxVM will manage.
- If you install both Solstice DiskSuite software and VxVM on a node, you must use Solstice DiskSuite software to manage disks local to each node (such as the root disk) and you must use VxVM to manage all shared disks.

See your volume manager documentation and "Installing and Configuring Solstice DiskSuite Software" on page 122 or "Installing and Configuring VxVM Software" on page 162 for instructions on how to install and configure the volume manager software. For more information about volume management in a cluster configuration, see *Sun Cluster 3.0 U1 Concepts*.

Guidelines for Volume Manager Software

Consider the following general guidelines when configuring your disks.

- **Mirrored multihost disks –** You must mirror all multihost disks across disk expansion units. See "Mirroring Multihost Disks" on page 20 for guidelines on mirroring multihost disks.
- **Mirrored root –** Mirroring the root disk ensures high availability, but such mirroring is not required. See "Mirroring Guidelines" on page 20 for guidelines on deciding whether to mirror the root disk.

- **Unique naming –** On any cluster node, if a local Solstice DiskSuite metadevice or VxVM volume is used as the device on which the `/global/.devices/node@`*nodeid* file system is mounted, the name of that metadevice or 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 3.0 U1 Data Services Installation and Configuration Guide* 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 configured in the node list for that device group. Solstice DiskSuite software is able to automatically check for this 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. In addition, when you use the clustering software to register Solstice DiskSuite disksets, VxVM disk groups, or individual sets of global devices as global device groups, you can perform only limited connectivity checking.
- **Hot spare disks –** You can use hot spare disks to increase availability, but they are not required.

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

Guidelines for Solstice DiskSuite Software

Consider the following points when you plan Solstice DiskSuite configurations.

- **Local metadevice names –** Each local metadevice name must be unique throughout the cluster and cannot be the same as any device ID (DID) name.
- **Mediators –** Each diskset configured with exactly two disk strings and mastered by exactly two nodes must have Solstice DiskSuite 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. You must configure each diskset with exactly two nodes acting as mediator hosts. You must use the same two nodes for all disksets requiring mediators and 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`(7) man page for details.
- `/kernel/drv/md.conf` **settings –** All metadevices used by each diskset are created in advance, at reconfiguration boot time, based on configuration parameters found in the `/kernel/drv/md.conf` file. The fields in the `md.conf` file are described in the Solstice DiskSuite documentation. You must modify the `nmd` and `md_nsets` fields as follows to support a Sun Cluster configuration.

– `nmd` – The `nmd` field defines the number of metadevices created for each diskset. You must set the value of `nmd` to the predicted largest number of metadevices used by any one of the disksets in the cluster. For example, if a cluster uses 10 metadevices in its first 15 disksets, but 1000 metadevices in the 16th diskset, you must set the value of `nmd` to at least 1000. Also, the value of `nmd` must be large enough to ensure that there are enough numbers for each DID name and each local metadevice name to be unique throughout the cluster. The maximum number of metadevices allowed per diskset is 8192. The default number of metadevices per diskset is 128.

– `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. You must set the value of `md_nsets` to the expected number of disksets in the cluster, plus one to allow Solstice DiskSuite software to manage the private disks on the local host (that is, those metadevices that are not in the local diskset). The maximum number of disksets allowed per cluster is 32. The default number of disksets is 4.

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

Caution – All cluster nodes 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 errors and possible loss of data.

Guidelines for VERITAS Volume Manager Software

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

- **Root disk group –** You must create a default root disk group (`rootdg`) on each node. The `rootdg` disk group can be created on the following disks.
  - The root disk, which must be encapsulated
  - One or more local non-root disks, which can be encapsulated or initialized
  - A combination of root and local non-root disks

  The `rootdg` 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 will 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) is highly recommended but not required. Using DRL decreases volume recovery time after a node failure. Using DRL might decrease I/O throughput.

File System Logging

Logging is required for cluster file systems. Sun Cluster supports the following logging file systems.

- Solaris UFS logging
- Solstice DiskSuite trans-metadevice UNIX file system (UFS) logging

For information about Solstice DiskSuite trans metadevice UFS logging, see your Solstice DiskSuite documentation. For information about Solaris UFS logging, see the `mount_ufs`(1M) man page. provided

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

TABLE 1-4 Supported File System Logging Matrix

Volume Manager	Supported File System Logging
Solstice DiskSuite	Solaris UFS loggingSolstice DiskSuite trans metadevice UFS logging,
VERITAS Volume Manager	Solaris UFS logging

Consider the following points when choosing between Solaris UFS logging and Solstice DiskSuite trans metadevice UFS logging for your Solstice DiskSuite volume manager.

- **Solaris UFS log size –** Solaris UFS logging always allocates the log 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, up to a maximum of 64 Mbytes.

- **Log metadevice –** A Solstice DiskSuite trans metadevice manages UFS logging. The logging device component of a trans metadevice is a metadevice 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. See your Solstice DiskSuite documentation for information about logging with trans metadevices.

Mirroring Guidelines

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

Mirroring Multihost Disks

Mirroring 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.

Consider the following points when mirroring 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 software and VERITAS Volume Manager (VxVM) support three-way mirroring. However, Sun Cluster requires only two-way mirroring.
- **Number of metadevices –** Under Solstice DiskSuite software, mirrors consist of other metadevices such as concatenations or stripes. Large configurations might contain a large number of metadevices. For example, seven metadevices are created for each logging UFS file system.
- **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 *Sun Cluster 3.0 U1 Concepts*.

Mirroring the Root Disk

Add this planning information to the "Local File System Layout Worksheet" in the *Sun Cluster 3.0 U1 Release Notes*.

For maximum availability, you should mirror root (`/`), `/usr`, `/var`, `/opt`, and `swap` on the local disks. Under VxVM, you encapsulate the root disk and mirror the generated subdisks. However, mirroring the root disk is not a requirement of Sun Cluster.

Before deciding whether to mirror the root disk, consider the risks, complexity, cost, and service time for the various alternatives concerning the root disk. There is no single mirroring strategy that works for all configurations. You might want to consider your local Enterprise Services representative's preferred solution when deciding whether to mirror root.

See your volume manager documentation and "Installing and Configuring Solstice DiskSuite Software" on page 122 or "Installing and Configuring VxVM Software" on page 162 for instructions on mirroring the root disk.

Consider the following issues and guidelines when deciding whether to mirror the root disk.

- **Complexity –** Mirroring the root disk adds complexity to system administration and complicates booting in single-user mode.
- **Backups –** Regardless of whether or not 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 configured as a quorum device to mirror a root disk.
- **Quorum –** Under Solstice DiskSuite software, in failure scenarios in which metadevice state database quorum is lost, you cannot reboot the system until maintenance is performed. See the Solstice DiskSuite documentation for information about the metadevice state database and state database replicas.
- **Separate controllers –** Highest availability includes mirroring the root disk on a separate controller.
- **Boot disk –** You can set up the mirror to be a bootable root disk so that you can boot from the mirror if the primary boot disk fails.
- **Secondary root disk –** With a mirrored root disk, the primary root disk can fail but work can continue on the secondary (mirror) root disk. At a later point, the primary root disk might return to service (perhaps after a power cycle or transient I/O errors) and subsequent boots are performed by using the primary root disk specified in the OpenBoot™ PROM `boot-device` field. In this situation no manual repair task occurs, but the drive starts working well enough to boot. Note that a Solstice DiskSuite 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 (causing a stale submirror). For example, changes to the `/etc/system` file would be lost. Some Solstice DiskSuite 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 it is booting from a mirror or an underlying physical device, and the mirroring becomes active partway through the boot process (after the metadevices are loaded). Before this point, the system is 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 procedures are in this chapter.

Installing the Software

The following table lists the tasks you perform to install the software.

TABLE 2-1 Task Map: Installing the Software

Task	For Instructions, Go To ...
Plan the layout of your cluster configuration and prepare to install software.	"How to Prepare for Cluster Software Installation" on page 25
(*Optional*) Install the Cluster Control Panel (CCP) software on the administrative console.	"How to Install Cluster Control Panel Software on the Administrative Console" on page 27
Install the Solaris operating environment and Sun Cluster software to establish new cluster nodes. Choose one of the following three methods.	
• **Method 1** – (*New clusters or added nodes*) Install Solaris software, then install the Sun Cluster software by using the `scinstall` utility.	"How to Install Solaris Software" on page 30 "How to Install Sun Cluster Software (scinstall)" on page 35
• **Method 2** – (*New clusters only*) Install Solaris software, then install SunPlex™ Manager and use it to install the Sun Cluster software.	"How to Install Solaris Software" on page 30 "Using SunPlex Manager to Install Sun Cluster Software" on page 47
• **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 59
Configure the name service look-up order.	"How to Configure the Name Service Switch" on page 71
Set up directory paths.	"How to Set Up the Root Environment" on page 72
Install data service software packages.	"How to Install Data Service Software Packages" on page 73
Perform post-installation setup and assign quorum votes.	"How to Perform Post-Installation Setup" on page 75
Install and configure volume manager software.	
• Install and configure Solstice DiskSuite software.	"Installing and Configuring Solstice DiskSuite Software" on page 122 Solstice DiskSuite documentation
• Install and configure VERITAS Volume Manager software.	"Installing and Configuring VxVM Software" on page 162 VERITAS Volume Manager documentation
Configure the cluster.	"Configuring the Cluster" on page 78

▼ 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 will help you plan your cluster configuration and prepare your installation strategy.**
   - *Sun Cluster 3.0 U1 Release Notes*—Restrictions, bug workarounds, and other late-breaking information.
   - *Sun Cluster 3.0 U1 Release Notes Supplement*—Post-release documentation about additional restrictions, bug workarounds, new features, and other late-breaking information. This document is regularly updated and is published online at the following Web site.

     `http://docs.sun.com`
   - *Sun Cluster 3.0 U1 Concepts*—Overview of the Sun Cluster product.
   - *Sun Cluster 3.0 U1 Installation Guide* (this manual)—Planning guidelines and procedures for installing and configuring Solaris, Sun Cluster, and volume manager software.
   - *Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide*—Planning guidelines and procedures for installing and configuring data services.
2. **Plan your cluster configuration.**
   - Use the planning guidelines in Chapter 1 and in the *Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide* to determine how you will install and configure your cluster.
   - Fill out the cluster framework and data services configuration worksheets in the *Sun Cluster 3.0 U1 Release Notes*. Use your completed worksheets for reference during the installation and configuration tasks.
3. **Have available all related documentation, including third-party documents.**

   The following is a partial list of product documentation you might need for reference during cluster installation.
   - Solaris software
   - Solstice DiskSuite software
   - VERITAS Volume Manager
   - Sun Management Center
   - Third-party applications such as Oracle
4. **Get all necessary patches for your cluster configuration.**

   See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.

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 the Administrative Console" on page 27.
   - If no, proceed to one of the following procedures.
     - If you intend to install Sun Cluster software by using either SunPlex Manager (GUI-based method) or the `scinstall` utility (text-based method), go to "How to Install Solaris Software" on page 30 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 59.

▼ How to Install Cluster Control Panel Software on the Administrative Console

This procedure describes how to install the Cluster Control Panel (CCP) software on the 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, plus a common window that sends input to all nodes at one time.

You can use any desktop machine that runs the Solaris 8 operating environment as an administrative console. In addition, you can also use the administrative console as a Sun Management Center console and/or server, and as an AnswerBook server. See Sun Management Center documentation for information on how to install Sun Management Center software. See the *Sun Cluster 3.0 U1 Release Notes* for information on how to install an AnswerBook server.

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.

---

1. **Ensure that the Solaris 8 operating environment and any Solaris patches are installed on the administrative console.**

   All platforms require Solaris 8 with at least the End User System Support software group.

2. **If you install from the CD-ROM, insert the Sun Cluster 3.0 7/01 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, it automatically mounts the CD-ROM on the `/cdrom/suncluster_3_0u1` directory.

3. **Change to the `/cdrom/suncluster_3_0u1/SunCluster_3.0/Packages` directory.**

   ```
   # cd /cdrom/suncluster_3_0u1/SunCluster_3.0/Packages
   ```

4. **Install the `SUNWccon` package.**

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

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

```
# pkgadd -d . SUNWscman
```

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

6. **If you installed from a CD-ROM, eject the CD-ROM.**

7. **Create an `/etc/clusters` file.**

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.

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

Add the physical node name of each cluster node, the terminal concentrator (TC) or System Service Processor (SSP) name, and the serial port numbers to the file.

```
# vi /etc/serialports
node1  TC-hostname  500N
node2  TC-hostname  500N
```

node1	Physical name of the cluster node
TC-hostname	Name of the TC or SSP
500*N*	Serial (`telnet`) port number

Note – Use the `telnet`(1) port numbers, not the physical port numbers, for the serial port numbers in the `/etc/serialports` file. To determine the serial port number, add 5000 to the physical port number. For example, if a physical port number is 6, the serial port number should be 5006.

For Sun Enterprise E10000 server servers, see the `/opt/SUNWcluster/bin/serialports`(4) man page for details and special considerations.

9. **For convenience, add the `/opt/SUNWcluster/bin` directory to the PATH and the `/opt/SUNWcluster/man` directory to the MANPATH on the administrative console.**

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

10. **Start the CCP utility.**

```
# /opt/SUNWcluster/bin/ccp clustername
```

    See the procedure "How to Remotely Log In to Sun Cluster" in the *Sun Cluster 3.0 U1 System Administration Guide* and the `/opt/SUNWcluster/bin/ccp`(1M) man page for information about how to use the CCP.

11. **Install the Solaris operating environment.**
    - To install Solaris software, go to "How to Install Solaris Software" on page 30.
    - To install Solaris and Sun Cluster software by using the `scinstall` JumpStart option, go to "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 59.

▼ How to Install Solaris Software

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

Note – If your nodes are already installed with the Solaris operating environment, you must still reinstall the Solaris software as described in this procedure to ensure successful installation of Sun Cluster software.

---

1. **Ensure that the hardware setup is complete and connections are verified before you install Solaris software.**

   See the *Sun Cluster 3.0 U1 Hardware Guide* and your server and storage device documentation for details.

2. **Have available your completed "Local File System Layout Worksheet" from the *Sun Cluster 3.0 Release Notes*.**

3. **Are you using a naming service?**

   - If no, proceed to Step 4. You will set up local hostname information in Step 12.
   - If yes, add address-to-name mappings for all public hostnames and logical addresses to any naming services (such as NIS, NIS+, or DNS) used by clients for access to cluster services. See "IP Addresses" on page 8 for planning guidelines. See your Solaris system administrator documentation for information about using Solaris naming services.

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

   If Cluster Control Panel (CCP) is installed and configured on your administrative console, you can use the `cconsole`(1M) utility to display the individual console screens. CCP also opens a master window from which you can send your input to all individual console windows at the same time.

   If you do not use CCP, connect to the consoles of each node individually.

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

---

5. **On each node of the cluster, determine whether the `local-mac-address` variable is correctly set to `false`.**

   Sun Cluster software does not support the `local-mac-address` variable set to `true`.

   a. Display the value of the `local-mac-address` variable.

   - If the node is preinstalled with Solaris software, as superuser run the following command.

```
# /usr/sbin/eeprom local-mac-address?
```

   - If the node is not yet installed with Solaris software, run the following command from the `ok` prompt.

```
ok printenv local-mac-address?
```

   b. Does the command return `local-mac-address?=false` on each node?

   - If yes, the variable settings are correct. Proceed to Step 6.
   - If no, change the variable setting on any node that is not set to `false`.
     - If the node is preinstalled with Solaris software, as superuser run the following command.

```
# /usr/sbin/eeprom local-mac-address?=false
```

     - If the node is not yet installed with Solaris software, run the following command from the `ok` prompt.

```
ok setenv local-mac-address? false
```

   c. Repeat Step a to verify any changes you made in Step b.

   The new setting becomes effective at the next system reboot.

6. **Install the Solaris operating environment as instructed in the 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 normally used to install the Solaris operating environment to install the software on new nodes to be installed into a clustered environment. These methods include the Solaris interactive installation program, Solaris JumpStart, and Solaris Web Start.

During Solaris software installation, do the following.

a. Install at least the End User System Support software group.

Note – Sun Enterprise E10000 server servers require the Entire Distribution + OEM software group.

You might need to install other Solaris software packages which are not part of the End User System Support software group, for example, the Apache HTTP server packages. Third-party software, such as Oracle, might also require additional Solaris packages. See your third-party documentation for any Solaris software requirements.

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

- Create a file system of at least 100 MBytes for use by the global-devices subsystem. If you intend to use SunPlex Manager to install Sun Cluster software, you must create the file system with a mount point of `/globaldevices`. This mount point is the default used by `scinstall`.

Note – A global-devices file system is required for Sun Cluster software installation to succeed.

- If you plan to use SunPlex Manager to install Solstice DiskSuite while installing Sun Cluster software, create a file system on slice 7 of at least 10 Mbytes with a mount point of `/sds`. Otherwise, create any file system partitions needed to support your volume manager software as described in "System Disk Partitions" on page 4.

c. Choose auto reboot.

Note – Solaris software is installed and the node reboots before the next prompts display.

---

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

e. Answer `no` when asked whether to enable automatic power-saving shutdown.

You must disable automatic shutdown in Sun Cluster configurations. See the `pmconfig`(1M) and `power.conf`(4) man pages for more information.

Note – The Solaris interface groups feature is disabled by default during Solaris software installation. Interface groups are not supported in a Sun Cluster configuration and should not be enabled. See the `ifconfig`(1M) man page for more information about Solaris interface groups.

---

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

- If yes, go to Step 8.
- If no, skip to Step 10.

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

- If yes, proceed to Step 9.
- If no, run `scsetup`(1M) from another, active cluster node to add the new node's name to the list of authorized cluster nodes. See "How to Add a Cluster Node to the Authorized Node List" in the *Sun Cluster 3.0 U1 System Administration Guide* for procedures.

9. Create a mount point on the new node for each cluster file system in the cluster.

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

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

b. 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.

10. Install any Solaris software patches.

See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions. If necessary, view the `/etc/release` file to see the exact version of Solaris software that is installed on a node.

11. **Install any hardware-related patches and download any needed firmware contained in the hardware patches.**

    See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.

12. **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.

13. **Install Sun Cluster software on your cluster nodes.**

    - To use SunPlex Manager, go to "Using SunPlex Manager to Install Sun Cluster Software" on page 47.
    - To use `scinstall`, go to "How to Install Sun Cluster Software (scinstall)" on page 35.

▼ How to Install Sun Cluster Software (`scinstall`)

After you install the Solaris operating environment, perform this task on each node of the cluster to install Sun Cluster software and establish new cluster nodes. You can also use this procedure to add new nodes to an existing cluster.

Note – If you used the `scinstall`(1M) custom JumpStart or SunPlex Manager installation method, the Sun Cluster software is already installed. Proceed to "How to Configure the Name Service Switch" on page 71.

---

1. **Have available the following completed configuration planning worksheets from the** *Sun Cluster 3.0 Release Notes***.**
   - "Cluster and Node Names Worksheet"
   - "Cluster Interconnect Worksheet"

   See "Planning the Sun Cluster Environment" on page 8 for planning guidelines.
2. **Become superuser on the cluster node.**
3. **If you install from the CD-ROM, insert the Sun Cluster 3.0 7/01 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, it automatically mounts the CD-ROM on the `/cdrom/suncluster_3_0u1` directory.
4. **Change to the `/cdrom/suncluster_3_0u1/SunCluster_3.0/Tools` directory.**

   ```
   # cd /cdrom/suncluster_3_0u1/SunCluster_3.0/Tools
   ```
5. **Are you installing a new node to an existing cluster?**
   - If yes, skip to Step 8.
   - If no, go to Step 6.

6. **Install the first node and establish the new cluster.**

   Follow the prompts to install Sun Cluster software, using the information from your configuration planning worksheets.

   a. Start the `scinstall`(1M) utility.

```
# ./scinstall
```

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

   - Interactive `scinstall` enables you to type ahead. Therefore, do not press Return 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.
   - Your session answers are stored as defaults for the next time you run this menu option. Default answers display between brackets ([ ]) at the end of the prompt.

Tip – Until the node is successfully booted in cluster mode, you can rerun `scinstall` and change the configuration information as needed. However, if bad configuration data for the node was pushed over to the established portion of the cluster, you might first need to remove the bad information. To do this, log in to one of the active cluster nodes, then use the `scconf`(1M) command to remove the bad adapter, junction, or cable information.

---

   b. From the Main Menu, type `1` (Establish a new cluster).

```
 *** Main Menu ***

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

      * 1) Establish a new cluster using this machine as the first node
      * 2) Add this machine as a node in an established cluster
        3) Configure a cluster to be JumpStarted from this install server
        4) Add support for new data services to this cluster node
        5) Print release information for this cluster node

      * ?) Help with menu options
      * q) Quit

    Option:  1

 *** Establishing a New Cluster ***
...
 Do you want to continue (yes/no) [yes]?  y
```

c. Specify the cluster name.

```
 >>> Cluster Name <<<
...
    What is the name of the cluster you want to establish?  clustername
```

d. Specify the names of the *other* nodes that will become part of this cluster.

```
 >>> Cluster Nodes <<<
...
    Node name:  node2
    Node name (Ctrl-D to finish):  <Control-D>

    This is the complete list of nodes:
...
    Is it correct (yes/no) [yes]?
```

e. Specify whether to use data encryption standard (DES) authentication.

By default, Sun Cluster software permits a node to connect to the cluster only if the node is physically connected to the private interconnect and if the node name was specified in Step d. However, the node actually communicates with the sponsoring node over the public network, since the private interconnect is not yet fully configured. DES authentication provides an additional level of security at installation time by enabling the sponsoring node to more reliably authenticate nodes that attempt to contact it to update the cluster configuration.

If you choose to use DES authentication for additional security, you must configure all necessary encryption keys before any node can join the cluster. See the `keyserv`(1M) and `publickey`(4) man pages for details.

```
 >>> Authenticating Requests to Add Nodes <<<
...
    Do you need to use DES authentication (yes/no) [no]?
```

f. Specify the private network address and netmask.

Note – You cannot change the private network address after the cluster is successfully formed.

```
 >>> Network Address for the Cluster Transport <<<
...
    Is it okay to accept the default network address (yes/no) [yes]?
    Is it okay to accept the default netmask (yes/no) [yes]?
```

g. If this is a two-node cluster, specify whether the cluster uses transport junctions.

Tip – You can specify that the cluster uses transport junctions, regardless of whether the nodes are directly connected to each other. If you specify that the cluster uses transport junctions, you can more easily add new nodes to the cluster in the future.

```
 >>> Point-to-Point Cables <<<
 ...
    Does this two-node cluster use transport junctions (yes/no) [yes]?
```

h. If this cluster uses transport junctions, specify names for the transport junctions.

You must use transport junctions if a cluster contains three or more nodes. You can use the default names `switch`*N* or create your own names.

```
 >>> Cluster Transport Junctions <<<
 ...
    What is the name of the first junction in the cluster [switch1]?
    What is the name of the second junction in the cluster [switch2]?
```

i. Specify the cluster interconnect transport adapters and, if used, the names of the transport junctions they connect to.

You can configure up to two adapters by using the `scinstall` command. You can configure additional adapters after Sun Cluster software is installed by using the `scsetup` utility.

```
 >>> Cluster Transport Adapters and Cables <<<
...
    What is the name of the first cluster transport adapter?  adapter
...
    Name of the junction to which "adapter" is connected [switch1]?
...
    What is the name of the second cluster transport adapter?  adapter
...
    Okay to use the default for the "qfe0" connection (yes/no) [yes]?

    What is the name of the second cluster transport adapter?  adapter

    Name of the junction to which "adapter" is connected [switch2]?

    Use the default port for the "adapter" connection [yes]?
```

j. Specify the global devices file system name.

```
 >>> Global Devices File System <<<
...
    The default is to use /globaldevices.

    Is it okay to use this default (yes/no) [yes]?
```

k. Do you have any Sun Cluster software patches to install?

- If yes, type **no** in the Automatic Reboot screen to decline automatic reboot.
- If no, type **yes** to accept automatic reboot.

```
 >>> Automatic Reboot <<<
...
    Do you want scinstall to reboot for you (yes/no) [yes]?
```

l. **Accept or decline the generated `scinstall` command.**

The `scinstall` command generated from your input is displayed for confirmation.

```
 >>> Confirmation <<<

    Your responses indicate the following options to scinstall:

      scinstall -i  \
...
    Are these the options you want to use (yes/no) [yes]?
    Do you want to continue with the install (yes/no) [yes]?
```

- If you accept the command and continue the installation, `scinstall` processing continues. "Example—Installing Sun Cluster Software" on page 45 shows an example of the output you might see during `scinstall` processing.
- If you decline the command, the `scinstall` utility returns you to the Main Menu. From there you can rerun menu option `1` and provide different answers. Your previous answers display as the defaults.

Sun Cluster installation output is logged in the `/var/cluster/logs/install/scinstall.log.`*pid* file, where *pid* is the process ID number of the `scinstall` instance.

Note – Unless you have installed your own `/etc/inet/ntp.conf` file, the `scinstall` command installs a default `ntp.conf` file for you. Because the default file is shipped with references to the maximum possible number of nodes, 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 Update Network Time Protocol (NTP)" on page 87 for information on how to suppress these messages under otherwise normal cluster conditions.

---

7. **Do you have any Sun Cluster software patches to install?**

- If yes, install any Sun Cluster software patches on the node and reboot the node. See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.
- If no, and `scinstall` rebooted the node during installation, go to Step 8. If `scinstall` did not reboot the node, manually reboot the node to establish the cluster before you proceed to Step 8.

The first node reboot after Sun Cluster software installation forms the cluster and establishes this node as the *first-installed* node of the cluster.

8. **Install the second node of the cluster.**

   Follow the prompts to install Sun Cluster software. Refer to the information from your configuration planning worksheets.

Note – *Do not* reboot or shut down the first-installed node while any other nodes are being installed, even if you use another node as the sponsoring node. Until quorum votes are assigned to the cluster nodes and cluster install mode is disabled, the first-installed node, which established the cluster, is the only node that has a quorum vote. Rebooting or shutting down the first-installed node will therefore cause a system panic because of lost quorum.

---

   a. **Start the `scinstall` utility.**

      You can start this step while software is still being installed on the first-installed node. If necessary, the second node waits for the first node to complete installation.

```
# ./scinstall
```

   b. **From the Main Menu, type `2` (Add this machine as a node).**

```
 *** Main Menu ***

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

      * 1) Establish a new cluster using this machine as the first node
      * 2) Add this machine as a node in an established cluster
        3) Configure a cluster to be JumpStarted from this install server
        4) Add support for new data services to this cluster node
        5) Print release information for this cluster node

      * ?) Help with menu options
      * q) Quit

    Option:  2

  *** Adding a Node to an Established Cluster ***
...
    Do you want to continue (yes/no) [yes]? y
```

c. Specify the name of any existing cluster node, referred to as the *sponsoring node*.

```
 >>> Sponsoring Node <<<
...
    What is the name of the sponsoring node?  node1
```

d. Specify the cluster name.

```
 >>> Cluster Name <<<
...
    What is the name of the cluster you want to join?  clustername
```

e. Specify whether this is a two-node cluster and whether the cluster uses transport junctions.

You must use transport junctions if a cluster contains three or more nodes.

```
 >>> Point-to-Point Cables <<<
...
    Is this a two-node cluster (yes/no) [yes]?

    Does this two-node cluster use transport junctions (yes/no) [yes]?
```

f. Specify the cluster interconnect transport adapters and transport junctions, if any.

```
 >>> Cluster Transport Adapters and Cables <<<
...
    What is the name of the first cluster transport adapter?  adapter
...
    Name of adapter on "node1" to which "adapter" is connected?  adapter

    What is the name of the second cluster transport adapter?  adapter
    Name of adapter on "node1" to which "adapter" is connected?  adapter
```

g. Specify the global devices file system name.

```
 >>> Global Devices File System <<<
...
    The default is to use /globaldevices.

    Is it okay to use this default (yes/no) [yes]?
```

h. Do you have any Sun Cluster software patches to install?

- If yes, type **no** in the Automatic Reboot screen to decline automatic reboot.
- If no, type **yes** to accept automatic reboot.

```
 >>> Automatic Reboot <<<
...
    Do you want scinstall to reboot for you (yes/no) [yes]?
```

i. Accept or decline the generated `scinstall` command.

The `scinstall` command generated from your input is displayed for confirmation.

```
 >>> Confirmation <<<

    Your responses indicate the following options to scinstall:

      scinstall -i  \
...
    Are these the options you want to use (yes/no) [yes]?
    Do you want to continue with the install (yes/no) [yes]?
```

- If you accept the command and continue the installation, `scinstall` processing continues. "Example—Installing Sun Cluster Software" on page 45 shows an example of the output you might see during `scinstall` processing. If the sponsoring node is not yet established in the cluster, `scinstall` waits for the sponsoring node to become available.
- If you decline the command, the `scinstall` utility returns you to the Main Menu. From there you can rerun menu option 2 and provide different answers. Your previous answers display as the defaults.

Sun Cluster installation output is logged in the `/var/cluster/logs/install/scinstall.log.`*pid* file, where *pid* is the process ID number of the `scinstall` instance.

Note – Unless you have installed your own `/etc/inet/ntp.conf` file, the `scinstall` command installs a default `ntp.conf` file for you. Because the default file is shipped with references to eight nodes, 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 Update Network Time Protocol (NTP)" on page 87 for information on how to suppress these messages under otherwise normal cluster conditions.

---

9. **Do you have any Sun Cluster software patches to install?**

   - If yes, install the Sun Cluster software patches on the node and reboot the node. See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.

Note – *Do not* reboot or shut down the first-installed node while any other nodes are being installed, even if you use another node as the sponsoring node. Until quorum votes are assigned to the cluster nodes and cluster install mode is disabled, the first-installed node, which established the cluster, is the only node that has a quorum vote. Rebooting or shutting down the first-installed node will therefore cause a system panic because of lost quorum. Cluster nodes remain in install mode until the first time you run the `scsetup`(1M) command, during the procedure "How to Perform Post-Installation Setup" on page 75.

   - If no, and `scinstall` rebooted the node during installation, go to Step 10. If `scinstall` did not reboot the node, manually reboot the node to establish the cluster before you proceed to Step 10.

10. **Repeat Step 8 and Step 9 on each additional node until all nodes are fully configured.**

    You do not need to wait for the second node to complete installation and reboot into the node before you begin installation on additional nodes.

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

    Go to "How to Configure the Name Service Switch" on page 71.

Example—Installing Sun Cluster Software

The following example shows the progress messages displayed as `scinstall` installation tasks are completed on the node `phys-schost-1`, which is the first node to be installed in the cluster.

```
** Installing SunCluster 3.0 **
        SUNWscr.....done.
        SUNWscdev...done.
        SUNWscu.....done.
        SUNWscman...done.
        SUNWscsal...done.
        SUNWscsam...done.
        SUNWscrsmop.done.
        SUNWsci.....done.
        SUNWscid....done.
        SUNWscidx...done.
        SUNWscvm....done.
        SUNWmdm.....done.


Initializing cluster name to "sccluster" ... done
Initializing authentication options ... done
Initializing configuration for adapter "hme2" ... done
Initializing configuration for adapter "hme4" ... done
Initializing configuration for junction "switch1" ... done
Initializing configuration for junction "switch2" ... done
Initializing configuration for cable ... done
Initializing configuration for cable ... done
Setting the node ID for "phys-schost-1" ... done (id=1)


Checking for global devices global file system ... done
Checking device to use for global devices 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.060199105132
Power management is incompatible with the HA goals of the cluster.
Please do not attempt to re-configure power management.


Ensure 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.276

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, go to "How to Install Sun Cluster Software (scinstall)" on page 35.

This section describes how to install SunPlex Manager and use it to install Sun Cluster software and establish new cluster nodes. You can also use SunPlex Manager to install one or more of the following additional software products.

- Solstice DiskSuite software—Also configures up to three metasets and associated metadevices, and creates and mounts cluster file systems for each.
- Sun Cluster HA for NFS data service
- Sun Cluster HA for Apache scalable data service

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

TABLE 2-2 Requirements to Use SunPlex Manager to Install Software

Software Package	Installation Requirements
Solstice DiskSuite	A 10-Mbyte partition that uses `/sds` as the file system name.
Sun Cluster HA for NFS data service	At least two shared disks of the same size which are connected to the same set of nodes.
	Solstice DiskSuite software installed 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 and is on the same subnet as the base hostnames of the cluster nodes.
Sun Cluster HA for Apache scalable data service	At least two shared disks of the same size which are connected to the same set of nodes.
	Solstice DiskSuite software installed 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 and is on the same subnet as the base hostnames of the cluster nodes.

The following table lists each metaset name and cluster file system mount point created by SunPlex Manager, depending on the number of shared disks connected to the node. For example, if a node has four shared disks connected to it, SunPlex

Manager creates the `mirror-1` and `stripe-1` metasets, but does not create the `concat-1` metaset because the node does not have enough shared disks to create a third metaset.

TABLE 2-3 Metasets Installed by SunPlex Manager

Shared Disks[1]	Metaset Name	Cluster File System Mount Point	Purpose
First pair of shared disks	`mirror-1`	`/global/mirror-1`	Sun Cluster HA for NFS or Sun Cluster HA for Apache scalable data service, or both
Second pair of shared disks	`stripe-1`	`/global/stripe-1`	unused
Third pair of shared disks	`concat-1`	`/global/concat-1`	unused

1. If the cluster does not meet the minimum shared disk requirement, SunPlex Manager still installs the Solstice DiskSuite packages. But without sufficient shared disks, SunPlex Manager cannot configure the metasets, metadevices, or cluster file systems needed to create instances of the data service.

▼ How to Install SunPlex Manager Software

The SunPlex Manager graphical user interface (GUI) provides an easy way to install and administer Sun Cluster software. Follow this procedure to install SunPlex Manager software on your cluster.

Note – If you intend to install Sun Cluster software by using another method, you do not need to perform this procedure. The `scinstall` command installs SunPlex Manager for you as part of the installation process.

Perform this procedure on each node of the cluster.

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

   See the installation procedures in "How to Install Solaris Software" on page 30.

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

3. **Install Apache software packages.**

   The Apache software packages are included in the Solaris Entire Distribution software group and all higher-level software groups. If you installed a lower-level software group, use the `pkginfo`(1) command to determine whether the software packages in Step c are already installed. If they are already installed, proceed to Step 4.

   a. **If you install from the CD-ROM, insert the Solaris 8 Software 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, it automatically mounts the CD-ROM.

   b. **Change to the `/cdrom/sol_8_sparc/Solaris_8/Product` directory.**

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

   c. **Install the Apache software packages in the following order.**

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

   d. **Eject the Solaris CD-ROM.**

   e. **Install any Apache software patches.**

      See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.

4. **Install the SunPlex Manager software packages.**

   a. **If you install from the CD-ROM, insert the Sun Cluster 3.0 7/01 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, it automatically mounts the CD-ROM on the `/cdrom/suncluster_3_0u1` directory.

   b. **Change to the `/cdrom/suncluster_3_0u1/SunCluster_3.0/Packages` directory.**

```
# cd /cdrom/suncluster_3_0u1/SunCluster_3.0/Packages
```

   c. **Install the SunPlex Manager software packages and answer `yes` for all prompts.**

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

   d. **Eject the Sun Cluster CD-ROM.**

5. **Repeat Step 2 through Step 4 on each node of the cluster.**

6. **Is the `root` password the same on every node of the cluster?**

   - If yes, go to Step 7.
   - If no, set the root password to the same value on each node of the cluster. If necessary, also use chkey(1) to update the RPC key pair.

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

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

7. **Do you intend to set up additional user accounts to access SunPlex Manager?**

   - If yes, go to Step 8.
   - If no, go to "How to Install Sun Cluster Software (SunPlex Manager)" on page 54 to install Sun Cluster software on your cluster nodes.

   Users who do not use the `root` system account nor have a user account set up on a particular node cannot access the cluster through SunPlex Manager from that node. Also, users cannot manage that node through another cluster node to which the users do have access.

8. **Determine how to set up user accounts to access SunPlex Manager.**

   In addition to `root`-user access, users can log in to SunPlex Manager with a user account that has role-based access control (RBAC). Go to one of the procedures listed in the following table to set up user accounts.

TABLE 2-4 Methods to Set Up SunPlex Manager User Accounts

Method	Go to This Procedure
Add RBAC authorization to an existing user account.	"How to Add RBAC Authorization to an Existing User Account" on page 52
Create a new user account that has RBAC authorization.	"How to Create a New User Account" on page 53

Note – If you assign RBAC authorization to a non-`root` user account, that user account can perform administrative actions usually performed only by `root`.

See "Role-Based Access Control" in the Solaris *System Administration Guide, Volume 2* for more information.

▼ How to Add RBAC Authorization to an Existing User Account

Add RBAC authorization to an existing user account. This enables the user to log in to SunPlex Manager by using the user's regular system password and have access to full SunPlex Manager functionality.

Note – If you assign RBAC authorization to a non-`root` user account, that user account can perform a set of administrative actions usually performed only by `root`.

---

1. **Become superuser on a node of the cluster.**
2. **Add the following entry to the `/etc/user_attr` file.**

```
# vi /etc/user_attr
username::::type=normal;auths=solaris.cluster.admin
```

3. **Repeat on each remaining node of the cluster.**
4. **Use SunPlex Manager to install Sun Cluster software.**

   Go to "How to Install Sun Cluster Software (SunPlex Manager)" on page 54.

▼ How to Create a New User Account

Create a new user account on all nodes the cluster.

Note – If you assign RBAC authorization to a non-`root` user account, that user account can perform a set of administrative actions usually performed only by `root`.

---

1. **Become superuser on a node of the cluster.**
2. **Create the new user account.**

```
# useradd -d dir -A solaris.cluster.admin login
```

`-d` *dir*	Specifies the home directory of the new user
`-A solaris.cluster.admin`	Assigns `solaris.cluster.admin` authorization to the new user account
login	Name of the new user account

Note – The user name must be unique and must not already exist either on the local machine or in the network name service.

See the `useradd`(1M) man page for more information about creating user accounts.

3. **Set the password.**

```
# passwd login
```

4. **Repeat on each remaining node of the cluster.**

   Ensure that the password for the user account is the same on all nodes of the cluster.

5. **Use SunPlex Manager to install Sun Cluster software.**

   Go to "How to Install Sun Cluster Software (SunPlex Manager)" on page 54.

▼ How to Install 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 Install Sun Cluster Software (scinstall)" on page 35.

Perform this procedure to use SunPlex Manager to install 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, and to install the Sun Cluster HA for NFS data service or scalable Sun Cluster HA for Apache data service or both.

The installation process might take from 30 minutes to two or more hours, depending on the number of cluster nodes, choice of data services, and number of disks in your cluster configuration.

1. **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 49. See "Using SunPlex Manager to Install Sun Cluster Software" on page 47 for installation requirements.

2. **Do you intend to install Sun Cluster HA for NFS or Sun Cluster HA for Apache?**
   - If no, go 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 47.

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

   a. **Provide each CD-ROM image in a location that is available to each node.**

      The CD-ROM images must be 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 exported to the network from machines outside the cluster.
      - Exported file systems on machines outside the cluster.
      - CD-ROM images copied to local file systems on each node of the cluster. The local file system must use the same name on each node.

   b. **Record the path to each CD-ROM image.**

      You will provide this information to SunPlex Manager in Step 19.

4. **Are there any patches required to support Sun Cluster or Solstice DiskSuite software?**
   - If yes, go to Step 5.
   - If no, go to Step 7.

5. **Do you intend to use SunPlex Manager to install patches?**
   - If yes, go to Step 6.
   - If no, manually install all patches required to support Sun Cluster or Solstice DiskSuite software *before* you use SunPlex Manager, then proceed to Step 7.

6. **Copy patches required for Sun Cluster or Solstice DiskSuite software into a single directory 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.**

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

      You will provide this information to SunPlex Manager in Step 19.

7. **Have available the following completed configuration planning worksheets from the** *Sun Cluster 3.0 Release Notes***.**
   - "Cluster and Node Names Worksheet"
   - "Cluster Interconnect Worksheet"
   - "Network Resources" worksheet

   See "Planning the Sun Cluster Configuration" on page 1 and the *Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide* for planning guidelines.

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

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

   SunPlex Manager installation functionality is incompatible with Web proxies.

10. **Ensure the disk caching and memory caching is enabled.**

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

11. **From the browser, connect to port 3000 on one node of the cluster.**

```
https://node:3000/
```

The Sun Cluster Installation screen displays in the browser window.

Note – If the SunPlex Manager displays the administration interface instead of the Sun Cluster Installation screen, Sun Cluster software is already installed on that node. Check that the name of the node in the URL is the correct name of the cluster node to install.

---

12. **If the browser displays a New Site Certification window, follow the onscreen instructions to accept the certificate.**

13. **In the Sun Cluster Installation screen, verify that the cluster meets the listed requirements for using SunPlex Manager.**

    - The Solaris End User Software Group or higher is installed.
    - Root disk partitions include a 100-MByte slice with the mount point `/globaldevices`.
    - Root disk partitions include a 10-MByte slice with the mount point `/sds`, if you will install Solstice DiskSuite.
    - File system paths to all needed CD-ROM images and patches are set up, as described in Step 3 through Step 6.

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

14. **Type a name for the cluster and select the number of nodes in your cluster.**

    Click Next to continue.

Tip – You can use the Back button to return to a previous screen and change your information. However, SunPlex Manager does not save the information you supplied in the later screens. When you click Next, you must again type or select your configuration information in those screens.

---

15. **Type the name of each cluster node.**

    Click Next to continue.

16. **From the pull-down lists for each node, select the names of the two adapters used for the private interconnects.**

    Refer to your completed "Cluster Interconnect Worksheet" for the appropriate adapter names for each node.

    Click Next to continue.

17. **Choose whether to install Solstice DiskSuite software.**

    You must install Solstice DiskSuite software if you intend to install the Sun Cluster HA for NFS or Sun Cluster HA for Apache data service.

Caution – When Solstice DiskSuite is installed, any data on all shared disks will be lost.

Click Next to continue.

18. **Choose whether to install Sun Cluster HA for NFS, Sun Cluster HA for Apache, or both.**

    Refer to your completed "Network Resources" worksheet for the appropriate logical hostname or shared address.

    - For Sun Cluster HA for NFS, also specify the logical hostname the data service will use.
    - For Sun Cluster HA for Apache, also specify the shared address the data service will use.

    Click Next to continue.

19. **Type the path for each CD-ROM image needed to install the packages you specified, and optionally the path for the patch directory.**

    - Type each path in the appropriate path field for each software package, as shown in TABLE 2-5.
    - Each specified path for a CD-ROM image must be the directory which contains the `.cdtoc` file for the CD-ROM.
    - For any software package you do not install, leave the relevant path field blank.
    - If you have already installed the required patches, leave the Patch Directory Path field blank.

    TABLE 2-5 CD-ROM Image Path Fields for Software Packages

Software Package to Install	Name of CD-ROM Image Path Field
Solstice DiskSuite	Solaris CD-ROM Path
Sun Cluster	Sun Cluster 3.0 7/01 CD-ROM Path
Sun Cluster HA for NFS, Sun Cluster HA for Apache	Sun Cluster 3.0 Agents 7/01 CD-ROM Path
Sun Cluster patches, Solstice DiskSuite patches	Patch Directory Path

    Click Next to continue.

20. **Is the information you supplied correct as displayed in the Confirm Information screen?**

    - If yes, proceed to Step 21.
    - If no, perform the following steps to correct the configuration information.

    a. Click Back until you return to the screen with the information to change.

    Note – When you click Back to back up to a previous screen, any information you typed in the subsequent screens is lost.

b. Type the correct information and click Next.

c. Retype or reselect the information in each screen until you return to the Confirm Information screen.

d. Verify that the information in the Confirm Information screen is now correct.

21. Click Begin Installation to start the installation process.

Note – Do *not* close the browser window or change the URL during the installation process.

---

a. If the browser displays a New Site Certification window, follow the onscreen instructions to accept the certificate.

b. If the browser prompts for login information, type the appropriate user ID and password for the node you connect to.

During installation, the screen displays brief messages about the status of the cluster installation process. When installation is complete, the browser displays the cluster monitoring and administration GUI.

SunPlex Manager installation output is logged in the `/var/cluster/spm` directory. Sun Cluster installation output is logged in the `/var/cluster/logs/install/scinstall.log.`*pid* file, where *pid* is the process ID number of the `scinstall` instance.

22. Use SunPlex Manager to verify quorum assignments and modify them, if necessary.

For clusters with three or more nodes, using 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 reassign quorum votes in the cluster.

23. Set up the name service look-up order.

Go to "How to Configure the Name Service Switch" on page 71.

▼ 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 establish 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 connections are verified before you install Solaris software.**

   See the *Sun Cluster 3.0 U1 Hardware Guide* and your server and storage device documentation for details on how to set up the hardware.

2. **Have available the following information.**

   - The Ethernet address of each cluster node
   - The following completed configuration planning worksheets from the *Sun Cluster 3.0 Release Notes*.
     - "Local File System Layout Worksheet"
     - "Cluster and Node Names Worksheet"
     - "Cluster Interconnect Worksheet"

   See "Planning the Solaris Operating Environment" on page 4 and "Planning the Sun Cluster Environment" on page 8 for planning guidelines.

3. **Are you using a naming service?**

   - If no, proceed to Step 4. You will set up the necessary hostname information in Step 13.
   - If yes, add address-to-name mappings for all public hostnames and logical addresses, as well as the IP address and hostname of the JumpStart server, to any naming services (such as NIS, NIS+, or DNS) used by clients for access to cluster services. See "IP Addresses" on page 8 for planning guidelines. See your Solaris system administrator documentation for information about using Solaris naming services.

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

   - If yes, run `scsetup`(1M) from another, active cluster node to add the new node's name to the list of authorized cluster nodes. See "How to Add a Cluster Node to the Authorized Node List" in the *Sun Cluster 3.0 U1 System Administration Guide* for procedures.
   - If no, go to Step 5.

5. **As superuser, set up the JumpStart install server for Solaris operating environment installation.**

   See the `setup_install_server`(1M) and `add_install_client`(1M) man pages and the *Solaris Advanced Installation Guide* for instructions on how to set up a JumpStart install server.

   When you set up the install server, ensure that the following requirements are met.

   - The install server is on the same subnet as the cluster nodes, but is not itself a cluster node.
   - The install server installs the release of the Solaris operating environment required by the Sun Cluster software.
   - A custom JumpStart directory exists for JumpStart installation of Sun Cluster. This *jumpstart-dir* directory must contain a copy of the `check`(1M) utility and be NFS exported for reading by the JumpStart install server.
   - Each new cluster node is configured as a custom JumpStart install client that uses the custom JumpStart directory set up for Sun Cluster installation.

6. **Create a directory on the JumpStart install server to hold your copy of the Sun Cluster 3.0 7/01 CD-ROM, if one does not already exist.**

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

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

7. **Copy the Sun Cluster CD-ROM to the JumpStart install server.**

   a. **Insert the Sun Cluster 3.0 7/01 CD-ROM into the CD-ROM drive on the JumpStart install server.**

      If the Volume Management daemon `vold`(1M) is running and configured to manage CD-ROM devices, it automatically mounts the CD-ROM on the `/cdrom/suncluster_3_0u1` directory.

   b. **Change to the `/cdrom/suncluster_3_0u1/SunCluster_3.0/Tools` directory.**

```
# cd /cdrom/suncluster_3_0u1/SunCluster_3.0/Tools
```

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

The `scinstall` command creates the new installation directory as it copies the CD-ROM files. The installation directory name `/export/suncluster/sc30` is used here as an example.

```
# ./scinstall -a /export/suncluster/sc30
```

d. Eject the CD-ROM.

```
# cd /
# eject cdrom
```

e. Ensure that the Sun Cluster 3.0 7/01 CD-ROM image on the JumpStart install server is NFS exported for reading by the JumpStart install server.

See the *NFS Administration Guide* and the `share`(1M) and `dfstab`(4) man pages for more information about automatic file sharing.

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

- If yes, go to Step 9.
- If no, skip to Step 10.

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

- If yes, proceed to Step 10.
- If no, run `scsetup`(1M) from any existing cluster node to add the new node's name to the list of authorized cluster nodes. See "How to Add a Cluster Node to the Authorized Node List" in the *Sun Cluster 3.0 U1 System Administration Guide* for procedures.

10. Use `scinstall` to configure custom JumpStart finish scripts.

JumpStart uses these finish scripts to install the Sun Cluster software.

a. From the JumpStart install server, start the `scinstall`(1M) utility.

The path `/export/suncluster/sc30` is used here as an example of the installation directory you created.

```
# cd /export/suncluster/sc30/SunCluster_3.0/Tools
# ./scinstall
```

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

- Interactive `scinstall` enables you to type ahead. Therefore, do not press Return 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.
- Your session answers are stored as defaults for the next time you run this menu option. Default answers display between brackets ([ ]) at the end of the prompt.

b. From the Main Menu, type 3 (Configure a cluster to be JumpStarted from this install server).

If option 3 does not have an asterisk in front, the option is disabled because JumpStart setup is not complete or has an error. Exit the `scinstall` utility, repeat Step 5 through Step 7 to correct JumpStart setup, then restart the `scinstall` utility.

```
*** Main Menu ***

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

       1) Establish a new cluster using this machine as the first node
       2) Add this machine as a node in an established cluster
     * 3) Configure a cluster to be JumpStarted from this install server
       4) Add support for new data services to this cluster node
       5) Print release information for this cluster node

     * ?) Help with menu options
     * q) Quit

   Option:  3

*** Custom JumpStart ***
...
   Do you want to continue (yes/no) [yes]?
```

c. Specify the JumpStart directory name.

```
>>> Custom JumpStart Directory <<<
....
   What is your JumpStart directory name?  jumpstart-dir
```

d. Specify the name of the cluster.

```
>>> Cluster Name <<<
...
   What is the name of the cluster you want to establish?  clustername
```

e. Specify the names of all cluster nodes.

```
 >>> Cluster Nodes <<<
...
    Please list the names of all cluster nodes planned for the initial
    cluster configuration. You must enter at least two nodes. List one
    node name per line. When finished, type Control-D:

    Node name:  node1
    Node name:  node2
    Node name (Ctrl-D to finish): <Control-D>

    This is the complete list of nodes:
...
    Is it correct (yes/no) [yes]?
```

f. Specify whether to use data encryption standard (DES) authentication.

By default, Sun Cluster software permits a node to connect to the cluster only if the node is physically connected to the private interconnect and if the node name was specified in Step e. However, the node actually communicates with the sponsoring node over the public network, since the private interconnect is not yet fully configured. DES authentication provides an additional level of security at installation time by enabling the sponsoring node to more reliably authenticate nodes that attempt to contact it to update the cluster configuration.

If you choose to use DES authentication for additional security, you must configure all necessary encryption keys before any node can join the cluster. See the `keyserv`(1M) and `publickey`(4) man pages for details.

```
 >>> Authenticating Requests to Add Nodes <<<
...
    Do you need to use DES authentication (yes/no) [no]?
```

g. Specify the private network address and netmask.

Note – You cannot change the private network address after the cluster is successfully formed.

```
 >>> Network Address for the Cluster Transport <<<
...
    Is it okay to accept the default network address (yes/no) [yes]?
    Is it okay to accept the default netmask (yes/no) [yes]?
```

h. If this is a two-node cluster, specify whether the cluster uses transport junctions.

Tip – You can specify that the cluster uses transport junctions, regardless of whether the nodes are directly connected to each other. If you specify that the cluster uses transport junctions, you can more easily add new nodes to the cluster in the future.

```
 >>> Point-to-Point Cables <<<
...
    Does this two-node cluster use transport junctions (yes/no) [yes]?
```

i. If this cluster uses transport junctions, specify the transport junction names.

You must use transport junctions if a cluster contains three or more nodes. You can use the default names `switch`*N* or create your own names.

```
 >>> Cluster Transport Junctions <<<
...
    What is the name of the first junction in the cluster [switch1]?
    What is the name of the second junction in the cluster [switch2]?
```

j. Specify the cluster interconnect transport adapters and, if used, the names of the transport junctions they connect to.

You can configure up to two adapters by using the `scinstall` command. You can configure additional adapters after Sun Cluster software is installed by using the `scsetup` utility.

```
 >>> Cluster Transport Adapters and Cables <<<
...
 For node "node1",
    What is the name of the first cluster transport adapter?  adapter
...
 For node "node1",
    Name of the junction to which "adapter" is connected [switch1]?
...
 For node "node1",
    Okay to use the default for the "adapter" connection (yes/no) [yes]?
```

```
 For node "node1",
    What is the name of the second cluster transport adapter?  adapter
 For node "node1",
    Name of the junction to which "adapter" is connected [switch2]?
 For node "node1",
    Use the default port for the "adapter" connection (yes/no) [yes]?


 For node "node2",
    What is the name of the first cluster transport adapter?  adapter
 For node "node2",
    Name of the junction to which "adapter" is connected [switch1]?
 For node "node2",
    Okay to use the default for the "adapter" connection (yes/no) [yes]?


 For node "node2",
    What is the name of the second cluster transport adapter?  adapter
 For node "node2",
    Name of the junction to which "adapter" is connected [switch2]?
 For node "node2",
    Use the default port for the "adapter" connection (yes/no) [yes]?
```

k. Specify the global devices file system name.

```
 >>> Global Devices File System <<<
...
    The default is to use /globaldevices.

 For node "node1",
    Is it okay to use this default (yes/no) [yes]?

 For node "node2",
    Is it okay to use this default (yes/no) [yes]?
```

l. Accept or decline the generated `scinstall` commands.

The `scinstall` command generated from your input is displayed for confirmation.

```
 >>> Confirmation <<<

    Your responses indicate the following options to scinstall:
------------------------------------------
 For node "node1",
      scinstall -c jumpstart-dir -h node1  \
...
    Are these the options you want to use (yes/no) [yes]?
------------------------------------------
 For node "node2",
      scinstall -c jumpstart-dir -h node2  \
...
    Are these the options you want to use (yes/no) [yes]?
------------------------------------------
    Do you want to continue with JumpStart set up (yes/no) [yes]?
```

If you do not accept the generated commands, the `scinstall` utility returns you to the Main Menu. From there you can rerun menu option `3` and provide different answers. Your previous answers display as the defaults.

11. If necessary, make adjustments to the default `class` file, or profile, created by `scinstall`.

The `scinstall` command creates the following `autoscinstall.class` default `class` file in the *jumpstart-dir*`/autoscinstall.d/3.0` directory.

```
install_type    initial_install
system_type     standalone
partitioning    explicit
filesys         rootdisk.s0 free /
filesys         rootdisk.s1 750 swap
filesys         rootdisk.s3 100  /globaldevices
filesys         rootdisk.s7 10
cluster         SUNWCuser       add
package         SUNWman         add
```

Note – The default `class` file installs the End User System Support software group (`SUNWCuser`) of Solaris software. For Sun Enterprise E10000 server servers, you must install the Entire Distribution + OEM software group. Also, some third-party software, such as Oracle, might require additional Solaris packages. See your third-party documentation for any Solaris software requirements.

You can change the profile 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 install profile meets minimum Sun Cluster file system allocation requirements, there are no restrictions on other changes to the install profile. See "System Disk Partitions" on page 4 for partitioning guidelines and requirements to support Sun Cluster 3.0 software.

12. Set up Solaris patch directories.

a. Create *jumpstart-dir*`/autoscinstall.d/nodes/`*node*`/patches` directories on the JumpStart install 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.

Also place copies of any hardware-related patches that must be installed after Solaris software is installed into each of these directories.

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

a. On the JumpStart install server, create files 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. This could be the JumpStart install server or another machine.
- IP address and hostname of each node in the cluster.

14. **(Optional) Add your own post-installation finish script.**

    You can add your own finish script, which is run after the standard finish script installed by the `scinstall` command.

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

    b. **Copy your finish script to the *jumpstart-dir*`/autoscinstall.d/nodes/`*node* directory, one directory for each node in the cluster.**

       Alternately, use this naming convention to create symbolic links to a shared finish script.

15. **If you use an administrative console, display a console screen for each node in the cluster.**

    If `cconsole`(1M) is installed and configured on your administrative console, you can use it to display the individual console screens. Otherwise, you must connect to the consoles of each node individually.

16. **From the `ok` PROM prompt on the console of each node, type the `boot net - install` command to begin the network JumpStart installation of each node.**

Note – The dash (-) in the command must be surrounded by a space on each side.

```
ok boot net - install
```

Sun Cluster installation output is logged in the `/var/cluster/logs/install/scinstall.log.`*pid* file, where *pid* is the process ID number of the `scinstall` instance.

Note – Unless you have installed your own `ntp.conf` file in the `/etc/inet` directory, the `scinstall` command installs a default `ntp.conf` file for you. Because the default file is shipped with references to the maximum possible number of nodes, 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 Update Network Time Protocol (NTP)" on page 87 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.

Note – The Solaris interface groups feature is disabled by default during Solaris software installation. Interface groups are not supported in a Sun Cluster configuration and should not be reenabled. See the `ifconfig`(1M) man page for more information about Solaris interface groups.

---

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

- If no, proceed to Step 18.
- If yes, create mount points on the new node for all existing cluster file systems.

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

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

b. On the node 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 returned by the mount command is `/global/dg-schost-1`, run `mkdir -p /global/dg-schost-1` on the node being added to the cluster.

Note – The mount points become active after you reboot the cluster in Step 19.

---

18. Install any Sun Cluster software patches.

See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.

19. **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 you installed require a node reboot.
    - If yes, perform a reconfiguration reboot.

    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.

    ```
    ok boot
    ```

    Until cluster install mode is disabled, only the first-installed node, which established the cluster, has a quorum vote. In an established cluster that is still in install mode, if the cluster is not shut down before the first-installed node is rebooted, the remaining cluster nodes cannot obtain quorum and the entire cluster shuts down. Cluster nodes remain in install mode until the first time you run the `scsetup`(1M) command, during the procedure "How to Perform Post-Installation Setup" on page 75.

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

    Go to "How to Configure the Name Service Switch" on page 71.

▼ 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 lookup order of the following entries.**

   - For the `hosts` and `netmasks` database entries, place `files` after `cluster`.
   - For all other database entries, place `files` first in 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 it in, in which case it is ignored.

   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 `nsswitch.conf`(4) for more information about `nsswitch.conf` entries.

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

   Go to "How to Set Up the Root Environment" on page 72.

▼ How to Set Up the Root Environment

Perform these tasks 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 before they attempt to output to the terminal. Otherwise, unexpected behavior or interference with data services might occur. See the Solaris *System Administration Guide, Volume 1* for more information about how to customize a user's work environment.

---

1. **Become superuser on a cluster node.**
2. **Modify the `.cshrc` file PATH and MANPATH entries.**

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

   For VERITAS Volume Manager, also set your PATH to include `/etc/vx/bin`. If you will install the `VRTSvmsa` package, also add `/opt/VRTSvmsa/bin` to your PATH.

   b. **Set the MANPATH to include `/usr/cluster/man`. Also include the volume manager-specific paths.**

   - For Solstice DiskSuite software, set your MANPATH to include `/usr/share/man`.
   - For VERITAS Volume Manager, set your MANPATH to include `/opt/VRTSvxvm/man`. If you will install the `VRTSvmsa` package, also add `/opt/VRTSvmsa/man` to your MANPATH.

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.**

   Go to "How to Install Data Service Software Packages" on page 73.

▼ How to Install Data Service Software Packages

Perform this task on each cluster node.

Note – If you used SunPlex Manager to install Sun Cluster HA for NFS or Sun Cluster HA for Apache, or both, and you do not intend to install any other data services, you do not need to perform this procedure. Instead, go to "How to Perform Post-Installation Setup" on page 75.

---

1. **Become superuser on a cluster node.**
2. **If you install from the CD-ROM, insert the Sun Cluster 3.0 Agents 7/01 CD-ROM into the CD-ROM drive on the node.**
3. **Start the `scinstall`(1M) utility.**

```
# scinstall
```

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

- Interactive `scinstall` enables you to type ahead. Therefore, do not press Return 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.

4. **To add data services, type `4` (Add support for a new data service to this cluster node).**

Follow the prompts to select all data services to install.July 2001

Note – You must install the same set of data service packages on each node, even if a node is not expected to host resources for an installed data service.

---

5. **If you installed from a CD-ROM, eject the CD-ROM.**
6. **Install any Sun Cluster data service patches.**

See the *Sun Cluster 3.0 U1 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, first shut down the cluster by using the `scshutdown`(1M) command, then reboot the each node in the cluster.

Note – Until cluster install mode is disabled, only the first-installed node, which established the cluster, has a quorum vote. In an established cluster which is still in install mode, if the cluster is not shut down before the first-installed node is rebooted, the remaining cluster nodes cannot obtain quorum and the entire cluster shuts down. Cluster nodes remain in install mode until the first time you run the `scsetup`(1M) command, during the procedure "How to Perform Post-Installation Setup" on page 75.

7. **Repeat Step 1 through Step 6 on each remaining cluster node.**
8. **Perform post-installation setup and assign quorum votes.**

   Go to "How to Perform Post-Installation Setup" on page 75.

▼ How to Perform Post-Installation Setup

Perform this procedure one time only, after the cluster is fully formed.

Verify that all nodes have joined the cluster.

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. **Determine the global device ID (DID) of each shared disk you will configure as a quorum device.**

   Use the `scdidadm` output from Step 2 to identify the DID name of each shared disk you will configure as a quorum device. For example, the output in the previous substep shows that global device `d2` is shared by `phys-schost-1` and `phys-schost-2`. You will use this information in Step 8. See "Quorum Devices" on page 13 for further information about planning quorum devices.

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

   - If yes, you might need to update the quorum configuration to accommodate your cluster's new configuration. See *Sun Cluster 3.0 U1 Concepts* for information about quorum. To change the quorum configuration, follow procedures in the *Sun Cluster 3.0 U1 System Administration Guide*. When the quorum configuration is satisfactory, go to Step 12.
   - If no, go to Step 6.

5. **Did you use** SunPlex Manager **to install** Sun Cluster **software?**

   - If yes, skip to Step 11. During Sun Cluster installation, SunPlex Manager assigns quorum votes and removes the cluster from install mode for you.
   - If no, go to Step 6.

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

7. **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 11.

   If the quorum setup process is interrupted or fails to complete successfully, rerun `scsetup`.

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

   A two-node cluster remains in install mode until a shared quorum device is configured. After the `scsetup` utility configures the quorum device, the message `Command completed successfully` is displayed. If your cluster has three or more nodes, quorum device configuration is optional.

9. **At the prompt** `Is it okay to reset "installmode"?`**, answer** `Yes`**.**

   After the `scsetup` utility sets quorum configurations and vote counts for the cluster, the message `Cluster initialization is complete` is displayed and the utility returns you to the Main Menu.

10. **From any node, verify the device and node quorum configurations.**

```
% scstat -q
```

11. **From any node, verify that cluster install mode is disabled.**

    You do not need to be superuser to run this command.

```
% scconf -p | grep 'Cluster install mode:'
Cluster install mode:                                   disabled
```

12. **Install volume management software.**

    - To install Solstice DiskSuite software, go to "Installing and Configuring Solstice DiskSuite Software" on page 122.
    - To install VERITAS Volume Manager software, go to "Installing and Configuring VxVM Software" on page 162.

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 24
- Volume manager installation and configuration as described in "Installing and Configuring Solstice DiskSuite Software" on page 122 or "Installing and Configuring VxVM Software" on page 162

TABLE 2-6 Task Map: Configuring the Cluster

Task	For Instructions, Go To ...
Create and mount cluster file systems.	"How to Add Cluster File Systems" on page 79
(Optional) Configure additional public network adapters.	"How to Configure Additional Public Network Adapters" on page 83
Configure Public Network Management (PNM) and set up NAFO groups.	"How to Configure Public Network Management (PNM)" on page 84
(Optional) Change a node's private hostname.	"How to Change Private Hostnames" on page 86
Edit the `/etc/inet/ntp.conf` file to update node name entries.	"How to Update Network Time Protocol (NTP)" on page 87
(Optional) Install the Sun Cluster module to Sun Management Center software.	"Installing the Sun Cluster Module for Sun Management Center" on page 88 Sun Management Center documentation
Install third-party applications and configure the applications, data services, and resource groups.	*Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide* "Data Service Configuration Worksheets and Examples" in the *Sun Cluster 3.0 Release Notes* Third-party application documentation

▼ How to Add Cluster File Systems

Perform this procedure for each cluster file system you add.

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 will erase data that you might not intend to delete.

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

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

   For volume manager installation procedures, see "Installing and Configuring Solstice DiskSuite Software" on page 122 or "Installing and Configuring VxVM Software" on page 162.

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 you create a file system for.

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

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

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

TABLE 2-7 Sample Raw Disk Device Names

Volume Manager	Sample Disk Device Name	Description
Solstice DiskSuite	`/dev/md/oracle/rdsk/d1`	Raw disk device `d1` within the `oracle` diskset
VERITAS Volume Manager	`/dev/vx/rdsk/oradg/vol01`	Raw disk device `vol01` within the `oradg` disk group
None	`/dev/global/rdsk/d1s3`	Raw disk device `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 will not be 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.**

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

      Logging is required for all cluster file systems.

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

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

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

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

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

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

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

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

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

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

```
# sccheck
```

If no errors occur, nothing is returned.

7. From any node in the cluster, mount the cluster file system.

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

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. Are your cluster nodes are connected to more than one public subnet?

- If yes, go to "How to Configure Additional Public Network Adapters" on page 83 to configure additional public network adapters.
- If no, go to "How to Configure Public Network Management (PNM)" on page 84 to configure PNM and set up NAFO groups.

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 Additional Public Network Adapters

If the nodes in the cluster are connected to more than one public subnet, you can configure additional public network adapters for the secondary subnets. This task is optional.

Note – Configure only public network adapters, not private network adapters.

---

1. **Have available your completed "Public Networks Worksheet" from the** *Sun Cluster 3.0 Release Notes***.**
2. **Become superuser on the node to configure for additional public network adapters.**
3. **Create a file named** `/etc/hostname.`*adapter*, **where** *adapter* **is the adapter name.**

Note – In each NAFO group, an `/etc/hostname.`*adapter* file should exist for only one adapter in the group.

---

4. **Type the hostname of the public network adapter IP address in the** `/etc/hostname.`*adapter* **file.**

   The following example shows the file `/etc/hostname.hme3`, created for the adapter `hme3`, which contains the hostname `phys-schost-1`.

```
# vi /etc/hostname.hme3
phys-schost-1
```

5. **On each cluster node, ensure that the** `/etc/inet/hosts` **file contains the IP address and corresponding hostname assigned to the public network adapter.**

   The following example shows the entry for `phys-schost-1`.

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

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

---

6. **On each cluster node, turn on the adapter.**

```
# ifconfig adapter plumb
# ifconfig adapter hostname netmask + broadcast + -trailers up
```

7. **Verify that the adapter is configured correctly.**

```
# ifconfig adapter
```

The output should contain the correct IP address for the adapter.

8. **Configure PNM and set up NAFO groups.**

Go to "How to Configure Public Network Management (PNM)" on page 84.

Each public network adapter to be managed by the Resource Group Manager (RGM) must belong to a NAFO group.

▼ How to Configure Public Network Management (PNM)

Perform this task on each node of the cluster.

Note – All public network adapters *must* belong to a Network Adapter Failover (NAFO) group. Also, each node can have only one NAFO group per subnet.

---

1. **Have available your completed "Public Networks Worksheet" from the** *Sun Cluster 3.0 Release Notes***.**
2. **Become superuser on the node to configure for a NAFO group.**
3. **Create the NAFO group.**

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

`-c` *nafo-group*	Configures the NAFO group *nafo-group*
`-o create` *adapter*	Creates a new NAFO group that contains one or more public network adapters

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

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

```
# pnmstat -l
```

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

5. **Do you intend to** change any private hostnames?
   - If yes, go to "How to Change Private Hostnames" on page 86.
   - If no, go to "How to Update Network Time Protocol (NTP)" on page 87 to update the `/etc/inet/ntp.conf` file.

Example—Configuring PNM

The following example creates NAFO group `nafo0`, which uses public network adapters `qfe1` and `qfe5`.

```
# pnmset -c nafo0 -o create qfe1 qfe5
# pnmstat -l
group  adapters       status  fo_time    act_adp
nafo0  qfe1:qfe5      OK      NEVER      qfe5
nafo1  qfe6           OK      NEVER      qfe6
```

▼ How to Change Private Hostnames

Perform this task if you do not want to use the default private hostnames (`clusternode`*nodeid*`-priv`) assigned during Sun Cluster software installation.

Note – Do *not* perform after applications and data services have been configured and started. Otherwise, an application or data service might continue to use the old private hostname after it 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. **To work with private hostnames, type `5` (Private hostnames).**
4. **To change a private hostname, type `1` (Change a private hostname).**

   Follow the prompts to change the private hostname. Repeat for each private hostname to change.

5. **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
```

6. **Update the `/etc/inet/ntp.conf` file.**

   Go to "How to Update Network Time Protocol (NTP)" on page 87.

▼ How to Update Network Time Protocol (NTP)

Perform this task on each node.

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

2. **Edit the `/etc/inet/ntp.conf` file.**

   The `scinstall`(1M) command copies a template file, `ntp.cluster`, to `/etc/inet/ntp.conf` as part of standard cluster installation. But if an `ntp.conf` file already exists before Sun Cluster software is installed, that existing file remains unchanged. If cluster packages are installed by using other means, such as direct use of `pkgadd`(1M), you need to configure NTP.

   a. **Remove all entries for private hostnames that are not used by the cluster.**

      If the `ntp.conf` file contains non-existent private hostnames, when a node is rebooted, error messages are generated on the node's attempts to contact those private hostnames.

   b. **If you changed any private hostnames after Sun Cluster software installation, update each file entry with the new private hostname.**

   c. **If necessary, make other modifications to meet your NTP requirements.**

      The primary requirement when you configure NTP, or any time synchronization facility, within the cluster is that all cluster nodes be synchronized to the same time. Consider accuracy of time on individual nodes secondary to the synchronization of time among nodes. You are free to configure NTP as best meets your individual needs, as long as this basic requirement for synchronization is met.

      See *Sun Cluster 3.0 U1 Concepts* for further information about cluster time. See the `ntp.cluster` template for guidelines on how to configure NTP for a Sun Cluster configuration.

3. **Restart the NTP daemon.**

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

4. **Do you intend to use Sun Management Center to configure resource groups or monitor the cluster?**

   - If yes, go to "Installing the Sun Cluster Module for Sun Management Center" on page 88.
   - If no, install third-party applications, register resource types, set up resource groups, and configure data services. See the documentation supplied with the application software and the *Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide*.

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 2-8 Task Map: Installing the Sun Cluster Module for Sun Management Center

Task	For Instructions, Go To...
Install Sun Management Center server, help server, agent, and console packages.	Sun Management Center documentation "Installation Requirements for Sun Cluster Monitoring" on page 89
Install Sun Cluster module packages.	"How to Install the Sun Cluster Module for Sun Management Center" on page 90
Start Sun Management Center server, console, and agent processes.	"How to Start Sun Management Center" on page 91
Add each cluster node as a Sun Management Center agent host object.	"How to Add a Cluster Node as a Sun Management Center Agent Host Object" on page 92
Load the Sun Cluster module to begin to monitor of the cluster.	"How to Load the Sun Cluster Module" on page 93

Installation Requirements for Sun Cluster Monitoring

The Sun Cluster module for Sun Management Center (formerly Sun Enterprise SyMON) is used to configure resource groups and monitor clusters. 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 non-cluster nodes. If you have an administrative console or other dedicated machine, you can realize improved performance if you run the console on the administrative console and the server on a separate machine. You must install the Sun Management Center agent package on each cluster node.

  Follow procedures in the Sun Management Center documentation to install the Sun Management Center packages.
- **Simple Network Management Protocol (SNMP) port –** When you install Sun Management Center on the agent, 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 you choose for reference later when you configure the cluster for monitoring.

▼ How to Install the Sun Cluster Module for Sun Management Center

Perform this procedure to install the Sun Cluster module console, server, and help server packages.

Note – The Sun Cluster module agent packages (`SUNWscsal` and `SUNWscsam`) were added to cluster nodes during Sun Cluster software installation.

---

1. **Ensure that the Sun Management Center core packages are installed.**

   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 administrative console, install the Sun Cluster module console package.**

   a. **Become superuser.**

   b. **If you install from the CD-ROM, insert the Sun Cluster 3.0 7/01 CD-ROM into the CD-ROM drive.**

   c. **Change to the `/cdrom/suncluster_3_0u1/SunCluster_3.0/Packages` directory.**

   d. **Install the Sun Cluster module console package.**

   ```
   # pkgadd -d . SUNWscscn
   ```

   e. **If you installed from a CD-ROM, eject the CD-ROM.**

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

   Use the same procedure as in Step 2.

4. **On the help server machine, install the Sun Cluster module help server package `SUNWscshl`.**

   Use the same procedure as in Step 2.

5. **Install any Sun Cluster module patches.**

   See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.

6. **Start Sun Management Center.**

   Go to "How to Start Sun Management Center" on page 91.

▼ 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 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
```

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

5. **Add cluster nodes as monitored host objects.**

   Go to "How to Add a Cluster Node as a Sun Management Center Agent Host Object" on page 92.

▼ 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 will contain the Sun Management Center agent host object 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 one.

   See your Sun Management Center documentation for information about how to create Sun Management Center domains.

2. **Select Edit>Create an Object from the pull-down menu.**

3. **Select 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 you chose when you installed the Sun Management Center agent.**

7. **Click OK.**

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

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

   Go to "How to Load the Sun Cluster Module" on page 93.

▼ How to Load the Sun Cluster Module

Perform this procedure to start cluster monitoring.

1. **From the Sun Management Center main window, point to an icon of a cluster node, right-click to display the pull-down menu, and select Load Module.**

   The Load Module window lists each available Sun Management Center module and whether it is currently loaded.

2. **Select Sun Cluster: Not loaded, which is usually at the bottom of the list, and click OK.**

   The Module Loader window shows the current parameter information for the selected module.

3. **Click OK.**

   After a few moments the module is loaded and a Sun Cluster icon is displayed in the Details window.

4. **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.

5. **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, 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, 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 URL `file:/opt/SUNWsymon/lib/locale/C/help/main.top.html`.

Note – The Help button in the Sun Management Center browser accesses Sun Management Center online help, not 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.

Where to Go From Here

Install third-party applications, register resource types, set up resource groups, and configure data services. See the documentation supplied with the application software and the *Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide*.

CHAPTER 3

Upgrading Sun Cluster Software

This chapter provides step-by-step procedures for upgrading a two-node Sun Cluster 2.2 configuration to Sun Cluster 3.0 Update 1 (U1).

Note – Do not perform this procedure if your cluster is already installed with Sun Cluster 3.0 software. Instead, to update your Sun Cluster 3.0 cluster to the U1 release, follow instructions in the `SunCluster_3.0/Tools/Upgrade/README` file on the Sun Cluster 3.0 7/01 CD-ROM.

The following step-by-step instructions are in this chapter.

For overview information about planning your Sun Cluster 3.0 configuration, see Chapter 1. For a high-level description of the related procedures in this chapter, see "Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 Update 1 Software" on page 96.

Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 Update 1 Software

Perform the following tasks to upgrade your two-node cluster from Sun Cluster 2.2 to Sun Cluster 3.0 Update 1 (U1) software.

TABLE 3-1 Task Map: Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 U1 Software

Task	For Instructions, Go To ...
Read upgrade conditions and restrictions, and plan a root disk partitioning scheme to support Sun Cluster 3.0 U1 software.	"Overview of Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 U1 Software" on page 97
Take the cluster out of production. For VERITAS Volume Manager (VxVM), also disable shared CCD.	"How to Shut Down the Cluster" on page 99
If your cluster uses VxVM, deport disk groups and remove VxVM software packages.	"How to Uninstall VERITAS Volume Manager Software From a Sun Cluster 2.2 Configuration" on page 103
Upgrade to the Solaris 8 operating environment if necessary, add a new `/globaldevices` file system, and change file system allocations to support Sun Cluster 3.0 U1 software. If your cluster uses Solstice DiskSuite software, also remove mediators and upgrade Solstice DiskSuite software.	"How to Upgrade the Solaris Operating Environment" on page 104
Upgrade to Sun Cluster 3.0 U1 framework software. If your cluster uses Solstice DiskSuite software, also recreate mediators.	"How to Upgrade Cluster Software Packages" on page 108
Update the PATH and MANPATH.	"How to Update the Root Environment" on page 112
Upgrade to Sun Cluster 3.0 U1 data services software. If necessary, upgrade third-party applications.	"How to Upgrade Data Service Software Packages" on page 113
Assign a quorum device, finish the cluster software upgrade, and start device groups and data services. If your cluster uses VERITAS Volume Manager (VxVM), reinstall VxVM software packages and import and register disk groups. If your cluster uses Solstice DiskSuite software, restore mediators.	"How to Finish Upgrading Cluster Software" on page 115
Verify that all nodes have joined the cluster.	"How to Verify Cluster Membership" on page 119

Overview of Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 U1 Software

This section provides conditions, restrictions, and planning guidelines for upgrading from Sun Cluster 2.2 to Sun Cluster 3.0 U1 software.

Conditions and Restrictions

The following conditions must be met to upgrade from Sun Cluster 2.2 to Sun Cluster 3.0 U1 software.

- The cluster must have exactly two nodes and be a supported configuration for Sun Cluster 3.0 U1 software. The upgrade does not support clusters of three or more nodes.
- Only Ethernet adapters are supported. Transport adapters must have a transmission rate of 100 Mbit/sec or greater.
- All cluster hardware must be stable and working properly.
- All third-party applications must be functioning properly.
- The cluster must be running on, or upgraded to, the Solaris 8 operating environment.
- You must upgrade all Sun Cluster software, framework, and data services at the same time.
- Sun Cluster 3.0 U1 software does not support upgrading directly to Sun Cluster 3.0 U1 software from Solstice HA 1.3, Sun Cluster 2.0, or Sun Cluster 2.1 software.
- Sun Cluster 3.0 U1 software does not support converting from one volume manager product to another during upgrade.
- The upgrade from Sun Cluster 2.2 to Sun Cluster 3.0 U1 software cannot be reversed after the `scinstall`(1M) command has been started on a node, even if the command does not complete successfully. To restart a failed upgrade, you must first reinstall Sun Cluster 2.2 software on the node.

Planning the Upgrade

To support Sun Cluster 3.0 U1 software, you probably need to change your current system disk layout. Consider the following when planning your new partitioning scheme.

- **Global devices namespace –** On each node you must create a file system of at least 100 MBytes and set its mount point as `/globaldevices`. This file system will be converted during upgrade to the appropriate global device namespace. If necessary, you can remove some of the swap space for this purpose, or use an external disk that is not shared with any other node.

- **Mirrored root –** If your root disks are mirrored, you must unmirror them before you modify partitions. The mirror can be used to recover the original configuration if the upgrade procedure fails. See your volume manager documentation for information.
- **Root (/) file system allocation –** If you intend to upgrade your configuration to the Solaris 8 operating environment, you probably need to increase the size of your root (/) partition on the root disks of all Sun Cluster nodes.

See "System Disk Partitions" on page 4 for more information about disk space requirements to support Sun Cluster 3.0 U1 software.

▼ How to Shut Down the Cluster

Before you upgrade the software, take the cluster out of production.

1. **Have available the CD-ROMs, documentation, and patches for all the software products you are upgrading.**
   - Solaris 8 operating environment
   - Solstice DiskSuite software or VERITAS Volume Manager
   - Sun Cluster 3.0 U1 framework
   - Sun Cluster 3.0 U1 data services (agents)
   - Third-party applications

   Solstice DiskSuite software and documentation are now part of the Solaris 8 product.

   Note – These procedures assume you are installing from CD-ROMs. If you are installing from a network, ensure that the CD-ROM image for each software product is loaded on the network.

   See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.

2. **Notify users that the cluster will be down.**
3. **Become superuser on each node of the cluster.**
4. **Search the `/var/adm/messages` log for unresolved error or warning messages.**

   Correct any problems.

5. **Verify that no logical hosts are in the maintenance state.**

   a. **Become superuser on a node of the cluster.**

   b. **Use the `hastat`(1M) command to display the status of the cluster.**

```
# hastat
HIGH AVAILABILITY CONFIGURATION AND STATUS
-----------------------------------------
...
LOGICAL HOSTS IN MAINTENANCE STATE
```

   If the screen output displays `NONE`, no logical hosts are in the maintenance state. Proceed to Step 6.

c. If a logical host is in the maintenance state, use the `haswitch`(1M) command to perform switchover.

```
# haswitch hostname logical-hostname
```

hostname Specifies the name of the node that is to own the logical host

logical-hostname Specifies the name of the logical host

d. Run the `hastat` command to verify the switchover completed successfully.

6. Ensure that the size of each logical host administrative file system is at least 10 Mbytes.

```
# df -k /logical-hostname
```

A logical host administrative file system that is not the required minimum size of 10 Mbytes will not be mountable after upgrade to Sun Cluster 3.0 U1. If a logical host administrative file system is smaller than 10 Mbytes, follow your volume manager documentation procedure for growing this file system.

7. Back up your system.

Ensure that all users are logged off the system before you back it up.

8. **(VxVM only) Disable the shared Cluster Configuration Database (CCD).**

   a. **From either node, create a backup copy of the shared CCD.**

```
# ccdadm -c backup-filename
```

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

   b. **On each node of the cluster, remove the shared CCD.**

```
# scconf clustername -S none
```

   c. **On each node, run the `mount`(1M) command to determine on which node the `ccdvol` is mounted.**

   The `ccdvol` entry looks similar to the following.

```
# mount
...
/dev/vx/dsk/sc_dg/ccdvol    /etc/opt/SUNWcluster/conf/ccdssa
ufs     suid,rw,largefiles,dev=27105b8  982479320
```

   d. **Run the `cksum`(1) command on each node to ensure that the `ccd.database` file is identical on both nodes.**

```
# cksum ccd.database
```

   e. **If the `ccd.database` files are different, from either node restore the shared CCD backup that you created in Step a.**

```
# ccdadm -r backup-filename
```

   f. **Stop the Sun Cluster 2.2 software on the node on which the `ccdvol` is mounted.**

```
# scadmin stopnode
```

   g. **From the same node, unmount the `ccdvol`.**

```
# umount /etc/opt/SUNWcluster/conf/ccdssa
```

9. **Stop the** Sun Cluster **2.2 software on each node of the cluster.**

```
# scadmin stopnode
```

10. **Run the** `hastat` **command to verify that no nodes are in the cluster.**

11. **Does the cluster use VERITAS Volume Manager**?

    - If yes, go to "How to Uninstall VERITAS Volume Manager Software From a Sun Cluster 2.2 Configuration" on page 103.
    - If no, go to "How to Upgrade the Solaris Operating Environment" on page 104.

▼ How to Uninstall VERITAS Volume Manager Software From a Sun Cluster 2.2 Configuration

If your cluster uses VERITAS Volume Manager (VxVM), perform this procedure on each node of the cluster to uninstall the VxVM software. Existing disk groups are retained and automatically reimported after you have upgraded all software.

Note – To upgrade to Sun Cluster 3.0 U1 software, you must remove VxVM software and later reinstall it, regardless of whether you have the latest version of VxVM installed.

---

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

2. **Uninstall VxVM.**

   Follow procedures in your VxVM documentation. This process involves the following tasks.

   - Deport all VxVM disk groups. Ensure that disks containing data to be preserved are not used for other purposes during the upgrade.
   - Unencapsulate the root disk, if it is encapsulated.
   - Shut down VxVM.
   - Remove all installed VxVM software packages.

3. **Remove the VxVM device namespace.**

   ```
   # rm -rf /dev/vx
   ```

4. **Repeat Step 1 through Step 3 on the other cluster node.**

5. **Upgrade or prepare the Solaris operating environment to support Sun Cluster 3.0 U1 software.**

   Go to "How to Upgrade the Solaris Operating Environment" on page 104.

▼ How to Upgrade the Solaris Operating Environment

Perform this procedure on each node in the cluster to upgrade or prepare the Solaris operating environment to support Sun Cluster 3.0 U1 software.

1. **Become superuser on the cluster node.**
2. **If your volume manager is Solstice DiskSuite and you are using mediators, unconfigure mediators.**

   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 by following the procedure "How to Fix Bad Mediator Data" on page 158.

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

   b. **List all mediators.**

   Use this information to determine which node, if any, has ownership of the diskset from which you will remove mediators.

```
# metaset -s setname
```

   Save this information for when you restore the mediators during the procedure "How to Upgrade Cluster Software Packages" on page 108.

   c. **If no node has ownership, take ownership of the diskset.**

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

   `-t` Takes ownership of the diskset

d. Unconfigure all mediators.

```
# 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`(7) man page for further information about mediator-specific options to the `metaset` command.

e. Remove the mediator software.

```
# pkgrm SUNWmdm
```

3. Does your configuration currently run Solaris 8 software?

- If no, go to Step 4.
- If yes,

a. Create a file system of at least 100 Mbytes and set its mount point as `/globaldevices`.

Note – The `/globaldevices` file system is necessary for Sun Cluster 3.0 U1 software installation to succeed.

---

b. Reallocate space in other partitions as needed to support Sun Cluster 3.0 U1 software.

See "System Disk Partitions" on page 4 for guidelines.

c. Go to Step 6.

4. Determine which procedure to use to upgrade to Solaris 8 software.

Volume Manager	Procedure to Use	For Instructions, Go To ...
Solstice DiskSuite	Upgrading both Solaris and Solstice DiskSuite software	Solstice DiskSuite installation documentation
VxVM	Performing a standard Solaris software installation	Solaris 8 installation documentation

5. **Upgrade to Solaris 8 software, following the procedure you selected in Step 4.**

   During installation, make the following changes to the root disk partitioning scheme.

   - Create a file system of at least 100 Mbytes and set its mount point as `/globaldevices`. The `/globaldevices` file system is necessary for Sun Cluster 3.0 U1 software installation to succeed.
   - Reallocate space in other partitions as needed to support Sun Cluster 3.0 U1 software.

   See "System Disk Partitions" on page 4 for partitioning guidelines.

Note – The Solaris interface groups feature is disabled by default during Solaris software installation. Interface groups are not supported in a Sun Cluster configuration and should not be enabled. See the `ifconfig`(1M) man page for more information about Solaris interface groups.

---

6. **Install any Solaris software patches.**

   See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.

7. **Install any hardware-related patches.**

   See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.

8. **For Solstice DiskSuite software, install any Solstice DiskSuite software patches.**

   See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.

9. **Upgrade to Sun Cluster 3.0 U1 software.**

   Go to "How to Upgrade Cluster Software Packages" on page 108.

Example—Unconfiguring Mediators

The following example shows the mediator host `phys-schost-1` unconfigured from the Solstice DiskSuite diskset `schost-1` before the upgrade to Solaris 8 software.

```
(Check mediator status)
# medstat -s schost-1

(List all mediators)
# metaset -s schost-1

(Unconfigure the mediator)
# metaset -s schost-1 -d -m phys-schost-1

(Remove mediator software)
# pkgrm SUNWmdm

(Begin software upgrade)
```

▼ How to Upgrade Cluster Software Packages

Perform this procedure on each node. You can perform this procedure on both nodes simultaneously if you have two copies of the Sun Cluster 3.0 7/01 CD-ROM.

Note – The `scinstall`(1M) upgrade command is divided into a two-step process—the `-u begin` option and the `-u finish` option. This procedure runs the `begin` option. The `finish` option is run in "How to Finish Upgrading Cluster Software" on page 115.

---

1. **Become superuser on a cluster node.**
2. **If your volume manager is Solstice DiskSuite, install the latest Solstice DiskSuite mediator package (`SUNWmdm`) on each node.**

   a. **If you are installing from the CD-ROM, insert the Sun Cluster 3.0 7/01 CD-ROM into the CD-ROM drive on a node.**

   If the volume daemon `vold`(1M) is running and configured to manage CD-ROM devices, it automatically mounts the CD-ROM on the `/cdrom/suncluster_3_0u1` directory.

   b. **Change to the `/cdrom/suncluster_3_0u1/SunCluster_3.0/Packages` directory.**

```
# cd /cdrom/suncluster_3_0u1/SunCluster_3.0/Packages
```

   c. **Add the `SUNWmdm` package.**

```
# pkgadd -d . SUNWmdm
```

   d. **Reboot the node.**

```
# shutdown -g0 -y -i6
```

   e. **Repeat on the other node.**

3. **Reconfigure mediators.**

   a. **Determine which node has ownership of the 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 for each diskset.**

4. **Begin upgrade to Sun Cluster 3.0 U1 software.**

   a. **On one node, change to the** `/cdrom/suncluster_3_0u1/SunCluster_3.0/Tools` **directory.**

```
# cd /cdrom/suncluster_3_0u1/SunCluster_3.0/Tools
```

b. Upgrade the cluster software framework.

Node To Upgrade	Command to Use
First node	**./scinstall -u begin -F**
Second node	**./scinstall -u begin -N** *node1*

`-F`	Specifies that this is the first-installed node in the cluster
`-N` *node1*	Specifies the name of the first-installed node in the cluster, *not* the name of the second node to be installed

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

c. Reboot the node.

```
# shutdown -g0 -y -i6
```

When the first node reboots into cluster mode, it establishes the cluster. The second node waits if necessary for the cluster to be established before completing its own processes and joining the cluster.

d. Repeat on the other cluster node.

5. On each node, install any Sun Cluster patches.

See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.

6. Update the directory paths.

Go to "How to Update the Root Environment" on page 112.

Example—Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 U1 Software – Begin Process

The following example shows the beginning process of upgrading a two-node cluster from Sun Cluster 2.2 to Sun Cluster 3.0 U1 software. The cluster node names are `phys-schost-1`, the first-installed node, and `phys-schost-2`, which joins the cluster that `phys-schost-1` established. The volume manager is Solstice DiskSuite and both nodes are used as mediator hosts for the diskset `schost-1`.

```
(Install the latest Solstice DiskSuite mediator package on each node)
# cd /cdrom/suncluster_3_0u1/SunCluster_3.0/Packages
# pkgadd -d . SUNWmdm

(Restore the mediator)
# metaset -s schost-1 -t
# metaset -s schost-1 -a -m phys-schost-1 phys-schost-2

(Begin upgrade on the first node)
phys-schost-1# cd /cdrom/suncluster_3_0u1/SunCluster_3.0/Tools
phys-schost-1# ./scinstall -u begin -F

(Begin upgrade on the second node)
phys-schost-2# cd /cdrom/suncluster_3_0u1/SunCluster_3.0/Tools
phys-schost-2# ./scinstall -u begin -N phys-schost-1

(Reboot each node)
# shutdown -g0 -y -i6
```

▼ How to Update the Root Environment

Perform the following tasks on each node of 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 before attempting to output to the terminal. Otherwise, unexpected behavior or interference with data services might occur. See the Solaris system administration documentation for more information on customizing a user's work environment.

---

1. **Become superuser on a cluster node.**
2. **Modify the `.cshrc` file PATH and MANPATH entries.**
   a. **Set the PATH to include `/usr/sbin` and `/usr/cluster/bin`.**

      For VERITAS Volume Manager, also set your PATH to include `/etc/vx/bin`. If you installed the `VRTSvmsa` package, also add `/opt/VRTSvmsa/bin` to your PATH.

   b. **Set the MANPATH to include `/usr/cluster/man`. Also include the volume manager-specific paths.**

   - For Solstice DiskSuite software, also set your MANPATH to include `/usr/share/man`.
   - For VERITAS Volume Manager, also set your MANPATH to include `/opt/VRTSvxvm/man`. If you installed the `VRTSvmsa` package, also add `/opt/VRTSvmsa/man` to your MANPATH.

3. **(Optional) For ease of administration, set the same root password on each node, if you have not already done so.**
4. **Start a new shell to activate the environment changes.**
5. **Repeat Step 1 through Step 4 on the other node.**
6. **Upgrade to Sun Cluster 3.0 U1 data service software.**

   Go to "How to Upgrade Data Service Software Packages" on page 113.

▼ How to Upgrade Data Service Software Packages

Perform this procedure on each cluster node.

1. **Become superuser on a node of the cluster.**
2. **Upgrade applications and apply application patches as needed.**

   See your application documentation for installation instructions.

   Note – 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.

   ---

3. **Add data services.**

   a. **Insert the Sun Cluster 3.0 Agents 7/01 CD-ROM into the CD-ROM drive on the node.**

   b. **Enter the `scinstall`(1M) utility.**

   ```
   # scinstall
   ```

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

   - Interactive `scinstall` enables you to type ahead. Therefore, do not press Return 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.

   c. **To add data services, type `4` (Add support for a new data service to this cluster node).**

   Follow the prompts to add data services.

   d. **Eject the CD-ROM.**

4. **Install any Sun Cluster data service patches.**

   See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.

5. **Repeat Step 1 through Step 4 on the other node of the cluster.**

6. **Shut down the second node to be upgraded to Sun Cluster 3.0 U1 software.**

```
phys-schost-2# shutdown -g0 -y -i0
```

Leave the second node shut down until after the first-installed node is rebooted.

7. **Reboot the first-installed node of the cluster.**

Ensure that the second node is shut down before rebooting the first-installed node. Otherwise, the second node will panic because quorum votes are not yet assigned.

```
phys-schost-1# shutdown -g0 -y -i6
```

8. **After the first-installed node has completed booting, boot the second node.**

```
ok boot
```

9. **After both nodes are rebooted, verify from either node that both nodes are cluster members.**

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

See the `scstat`(1M) man page for more information about displaying cluster status.

10. **Assign a quorum device and finish the upgrade.**

Go to "How to Finish Upgrading Cluster Software" on page 115.

▼ How to Finish Upgrading Cluster Software

This procedure finishes the `scinstall`(1M) upgrade process begun in "How to Upgrade Cluster Software Packages" on page 108. Perform these steps on each node of the cluster.

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

2. **Choose a shared disk to be the quorum device.**

   You can use any disk shared by both nodes as a quorum device. From either node, use the `scdidadm`(1M) command to determine the shared disk's device ID (DID) name. You specify this device name in Step 5, in the `-q globaldev=`*DIDname* option to `scinstall`.

   ```
   # scdidadm -L
   ```

3. **If your volume manager is VxVM, reinstall and configure the VxVM software on each node of the cluster, including any patches.**

   Follow the procedures in "Installing and Configuring VxVM Software" on page 162.

   Note – If you must reboot the first-installed node, first shut down the cluster by using the `scshutdown`(1M) command, then reboot. Do not reboot the first-installed node of the cluster until *after* the cluster is shut down.

   Until cluster install mode is disabled, only the *first-installed* node, which established the cluster, has a quorum vote. In an established cluster that is still in install mode, if the cluster is not shut down before the first-installed node is rebooted, the remaining cluster nodes cannot obtain quorum and the entire cluster shuts down.

   To determine which is the first-installed node, view quorum vote assignments by using the `scconf -p` command. The only node that has a quorum vote is the first-installed node.

   After you perform Step 7, quorum votes are assigned and this reboot restriction is no longer necessary.

4. **Insert the Sun Cluster 3.0 Agents 7/01 CD-ROM into the CD-ROM drive on a node.**

   This step assumes that the volume daemon `vold`(1M) is running and configured to manage CD-ROM devices.

5. **Finish the cluster software upgrade on that node.**

```
# scinstall -u finish -q globaldev=DIDname \
-d /cdrom/scdataservices_3_0 -s srvc[,srvc]
```

`-q globaldev=`*DIDname*	Specifies the device ID (DID) name of the quorum device
`-d /cdrom/scdataservices_3_0`	Specifies the directory location of the CD-ROM image
`-s` *srvc*	Specifies the name of the data service to configure

---

Note – An error message similar to the following might be generated. You can safely ignore it.

```
** Installing Sun Cluster - Highly Available NFS Server **
Skipping "SUNWscnfs" - already installed
```

6. **Eject the CD-ROM.**

7. **Repeat Step 4 through Step 6 on the other node.**

   When completed on both nodes, cluster install mode is disabled and all quorum votes are assigned.

8. **If your volume manager is Solstice DiskSuite, from either node bring pre-existing disk device groups online.**

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

`-z`	Performs the switch
`-D` *disk-device-group*	Specifies the name of the disk device group, which for Solstice DiskSuite software is the same as the diskset name
`-h` *node*	Specifies the name of the cluster node that serves as the primary of the disk device group

9. **From either node, bring pre-existing data service resource groups online.**

   At this point, Sun Cluster 2.2 logical hosts are converted to Sun Cluster 3.0 U1 resource groups, and the names of logical hosts are appended with the suffix `-lh`. For example, a logical host named `lhost-1` is upgraded to a resource group named `lhost-1-lh`. Use these converted resource group names in the following command.

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

   `-g` *resource-group* Specifies the name of the resource group to bring online

   You can use the `scrgadm -p` command to display a list of all resource types and resource groups in the cluster. The `scrgadm -pv` command displays this list with more detail.

10. **If you are using Sun Management Center to monitor your Sun Cluster configuration, install the Sun Cluster module for Sun Management Center.**

    a. **Ensure that you are using the most recent version of Sun Management Center.**

       See your Sun Management Center documentation for installation or upgrade procedures.

    b. **Follow guidelines and procedures in "Installation Requirements for Sun Cluster Monitoring" on page 89 to install the Sun Cluster module packages.**

11. **Verify that all nodes have joined the cluster.**

    Go to "How to Verify Cluster Membership" on page 119.

Example—Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 U1 Software – Finish Process

The following example shows the finish process of upgrading a two-node cluster upgraded from Sun Cluster 2.2 to Sun Cluster 3.0 U1 software. The cluster node names are `phys-schost-1` and `phys-schost-2`, the device group names are dg-schost-1 and dg-schost-2, and the data service resource group names are `lh-schost-1` and `lh-schost-2`.

```
(Determine the DID of the shared quorum device)
phys-schost-1# scdidadm -L

(Finish upgrade on each node)
phys-schost-1# scinstall -u finish -q globaldev=d1 \
-d /cdrom/suncluster_3_0u1 -s nfs
phys-schost-2# scinstall -u finish -q globaldev=d1 \
-d /cdrom/suncluster_3_0u1 -s nfs

(Bring device groups and data service resource groups on each node online)
phys-schost-1# scswitch -z -D dg-schost-1 -h phys-schost-1
phys-schost-1# scswitch -z -g lh-schost-1 -h phys-schost-1
phys-schost-1# scswitch -z -D dg-schost-2 -h phys-schost-2
phys-schost-1# scswitch -z -g lh-schost-2 -h phys-schost-2
```

▼ How to Verify Cluster Membership

Perform this procedure to verify that all nodes have joined the cluster.

1. **Become superuser on any node in the cluster.**
2. **Display cluster status.**

   Verify that cluster nodes are online and that the quorum device, device groups, and data services resource groups are configured and online.

   ```
   # scstat
   ```

   See the `scstat`(1M) man page for more information about displaying cluster status.

3. **On each node, display a list of all devices the system checks to verify their connectivity to the cluster nodes.**

   The output on each node should be the same.

   ```
   # scdidadm -L
   ```

   The cluster upgrade is complete. You can now return the cluster to production.

APPENDIX **A**

Installing and Configuring Solstice DiskSuite Software for Sun Cluster Configurations

Install and configure your local and multihost disks for Solstice DiskSuite software by using the procedures in this appendix, along with the planning information in "Planning Volume Management" on page 16. See your Solstice DiskSuite documentation for additional details.

The following procedures are in this appendix.

Installing and Configuring Solstice DiskSuite Software

Before you begin, have available the following information.

- Mappings of your storage disk drives.
- The following completed configuration planning worksheets from the *Sun Cluster 3.0 Release Notes*. See "Planning Volume Management" on page 16 for planning guidelines.
  - "Local File System Layout Worksheet"
  - "Disk Device Group Configurations Worksheet"
  - "Volume Manager Configurations Worksheet"
  - "Metadevices Worksheet (Solstice DiskSuite)"

The following table lists the tasks to perform to install and configure Solstice DiskSuite software for Sun Cluster configurations.

Note – If you used SunPlex Manager to install Solstice DiskSuite software, the procedures "How to Install Solstice DiskSuite Software" on page 124 through "How to Create Metadevice State Database Replicas" on page 128 are already completed.

TABLE A-1 Task Map: Installing and Configuring Solstice DiskSuite Software (Sheet 1 of 2)

Task	For Instructions, Go To ...
Plan the layout of your Solstice DiskSuite configuration.	"Planning Volume Management" on page 16 "Solstice DiskSuite Configuration Example" on page 159
Install Solstice DiskSuite software.	"How to Install Solstice DiskSuite Software" on page 124
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 Names and Disksets" on page 126
Create metadevice state database replicas on the local disks.	"How to Create Metadevice State Database Replicas" on page 128
(*Optional*) Mirror file systems on the root disk.	"Mirroring the Root Disk" on page 129
Create disksets by using the `metaset` command.	"How to Create a Diskset" on page 145
Add disk drives to the disksets.	"How to Add Drives to a Diskset" on page 147
Repartition drives in a diskset to allocate space to slices 1 through 6.	"How to Repartition Drives in a Diskset" on page 149

TABLE A-1 Task Map: Installing and Configuring Solstice DiskSuite Software (Sheet 2 of 2)

Task	For Instructions, Go To ...
List device ID pseudo-driver mappings and define metadevices in the `/etc/lvm/md.tab` files.	"How to Create an md.tab File" on page 150
Initialize the `md.tab` files.	"How to Activate Metadevices" on page 153
For dual-string configurations, configure mediator hosts, check the status of mediator data, and, if necessary, fix bad mediator data.	"How to Add Mediator Hosts" on page 156 "How to Check the Status of Mediator Data" on page 157 "How to Fix Bad Mediator Data" on page 158
Configure the cluster.	"Configuring the Cluster" on page 78

▼ How to Install Solstice DiskSuite Software

Note – If you used SunPlex Manager to install Solstice DiskSuite software, do not perform this procedure. Instead, proceed to "Mirroring the Root Disk" on page 129.

Perform this task on each node in the cluster.

1. **Become superuser on the cluster node.**
2. **If you install from the CD-ROM, insert the Solaris 8 Software 2 of 2 CD-ROM 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.
3. **Install the Solstice DiskSuite software packages.**

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

Install software packages in the order shown in the following example.

```
# cd /cdrom/sol_8_sparc_2/Solaris_8/EA/products/DiskSuite_4.2.1/sparc/Packages
# pkgadd -d . SUNWmdr SUNWmdu [SUNWmdx] optional-pkgs
```

The `SUNWmdr` and `SUNWmdu` packages are required for all Solstice DiskSuite installations. The `SUNWmdx` package is also required for the 64-bit Solstice DiskSuite installation.

See your Solstice DiskSuite installation documentation for information about optional software packages.

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

   See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.
6. **Repeat Step 1 through Step 5 on the other nodes of the cluster.**
7. **From one node of the cluster, manually populate the global device namespace for** Solstice DiskSuite**.**

```
# scgdevs
```

8. **Set the number of metadevice names and disksets expected in the cluster.**

   Go to "How to Set the Number of Metadevice Names and Disksets" on page 126.

▼ How to Set the Number of Metadevice Names and Disksets

Note – If you used SunPlex Manager to install Solstice DiskSuite software, do not perform this procedure. Instead, proceed to "Mirroring the Root Disk" on page 129.

This procedure describes how to calculate the quantity of metadevice names needed for your configuration, and how to modify the `/kernel/drv/md.conf` file.

Tip – The default number of metadevice 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.

---

1. **Calculate the largest metadevice name you need for any diskset in the cluster.**

   Each diskset can have a maximum of 8192 metadevice names. You will supply this calculated value for the `nmd` field.

   a. **Calculate the quantity of metadevice names you need for each diskset.**

      If you use local metadevices, ensure that each local metadevice name is unique throughout the cluster and does not use the same name as any device ID (DID) in the cluster.

Tip – Choose a range of numbers to use exclusively for DID names and a range for each node to use exclusively for its local metadevice names. For example, DIDs would use names in the range from `d1` to `d1000`, local metadevices on node 1 would use names in the range from `d1100` to `d1199`, local metadevices on node 2 would use `d1200` to `d1299`, and so on.

---

   b. **Determine the largest of the metadevice names to be used in any diskset.**

      The quantity of metadevice names to set is based on the metadevice name *value* rather than on the *actual quantity*. For example, if your metadevice names range from `d950` to `d1000`, Solstice DiskSuite software requires 1000 names, not 50.

2. **Calculate the total expected number of disksets in the cluster, then add one for private disk management.**

   The cluster can have a maximum of 32 disksets. The default number of disksets is 4. You will supply this calculated value for the `md_nsets` field.

3. **On each node, 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 errors and possible loss of data.

---

   a. **If the calculated quantity of metadevice names exceeds 128, set the `nmd` field to the quantity calculated in Step 1.**

   b. **Set the `md_nsets` field to the quantity calculated in Step 2.**

4. **On each node, perform a reconfiguration reboot.**

```
# touch /reconfigure
# shutdown -g0 -y -i6
```

   Changes to the `/kernel/drv/md.conf` file become operative after you perform a reconfiguration reboot.

5. **Create local replicas.**

   Go to "How to Create Metadevice State Database Replicas" on page 128.

▼ How to Create Metadevice State Database Replicas

Note – If you used SunPlex Manager to install Solstice DiskSuite software, do not perform this procedure. Instead, proceed to "Mirroring the Root Disk" on page 129.

Perform this procedure on each node in the cluster.

1. **Become superuser on the cluster node.**
2. **Create replicas on one or more local disks for each cluster node by using the `metadb` command.**

   See the `metadb`(1M) man page and your Solstice DiskSuite documentation for details.

Tip – To provide protection of metadevice state data, which is necessary to run Solstice DiskSuite 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.

---

3. **Verify the replicas.**

```
# metadb
```

4. **Do you intend to mirror file systems on the root disk?**
   - If yes, go to "Mirroring the Root Disk" on page 129.
   - If no, go to "How to Create a Diskset" on page 145 to create Solstice DiskSuite disksets.

Example—Creating Metadevice State Database Replicas

The following example shows three metadevice state database replicas, each created on a different disk.

```
# metadb -af c0t0d0s7 c0t1d0s7 c1t0d0s7
# metadb
flags           first blk      block count
     a       u       16            1034         /dev/dsk/c0t0d0s7
     a       u       1050          1034         /dev/dsk/c0t1d0s7
     a       u       2084          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 130
- "How to Mirror the Global Namespace" on page 134
- "How to Mirror File Systems That Cannot Be Unmounted" on page 137
- "How to Mirror User-Defined File Systems" on page 141

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 a node of the cluster.**
2. **Use the `metainit`(1M) command to put the root slice in a single-slice (one-way) concatenation.**

```
# metainit -f submirror1 1 1 root-disk-slice
```

3. **Create a second concatenation.**

```
# metainit -f submirror2 1 1 submirror-disk-slice
```

4. **Create a one-way mirror with one submirror.**

```
# metainit mirror -m submirror1
```

Note – The metadevice 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.

```
# 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. **Evacuate any resource groups or device groups from the node.**

```
# scswitch -S -h node
```

`-S`	Evacuates all resource groups and device groups
`-h` *node*	Specifies the name of the node from which to evacuate resource or device groups

8. **Reboot the node.**

```
# shutdown -g0 -y -i6
```

9. **Use the `metattach`(1M) command to attach the second submirror to the mirror.**

```
# metattach mirror submirror2
```

10. **If the disk used to mirror the root disk is physically connected to more than one node (multiported), 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 -L` command to display the full device ID (DID) pseudo-driver 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 DID pseudo-driver name.

```
# scdidadm -L
...
1         phys-schost-3:/dev/rdsk/c1t1d0     /dev/did/rdsk/d2
# scconf -c -D name=dsk/d2,localonly=true
```

       For more information about the `localonly` property, see the `scconf_dg_rawdisk`(1M) man page.

b. Use the `scconf`(1M) command to enable the `localonly` property.

```
# scconf -c -D name=rawdisk-groupname,localonly=true
```

`-D name=`*rawdisk-groupname* Specifies the name of the raw disk device group

11. Record the alternate boot path for possible future use.

```
# 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 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 134.
- If no, go to Step 14.

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

- If yes, go to "How to Mirror File Systems That Cannot Be Unmounted" on page 137.
- If no, go to Step 15.

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

- If yes, go to "How to Mirror User-Defined File Systems" on page 141.
- If no, go to "How to Create a Diskset" on page 145 to create a diskset.

Example—Mirroring the Root (/) File System

The following example shows 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 -f d20 1 1 c2t2d0s0
d12: Concat/Stripe is setup
# metainit d0 -m d10
d10: Mirror is setup
# metaroot d0
# lockfs -fa

(Reboot the node)
# scswitch -S -h phys-schost-1
# shutdown -g0 -y -i6

(Attach the second submirror)
# metattach d0 d20
d0: Submirror d20 is attached

(Enable the localonly property of the mirrored disk's raw disk device group)
# 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.**

```
# metainit -f submirror1 1 1 diskslice
```

3. **Create a second concatenation.**

```
# metainit -f submirror2 1 1 submirror-diskslice
```

4. **Create a one-way mirror with one submirror.**

```
# metainit mirror -m submirror1
```

Note – The metadevice name for the mirror *must* be unique throughout the cluster.

5. **Attach the second submirror to the mirror.**

   This attachment starts a sync 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.**

   Ensure that each metadevice name for a mirror is unique throughout the cluster.

8. **Wait for the sync of the mirrors, started in Step 5, to complete.**

   Use the `metastat`(1M) command to view mirror status.

```
# metastat mirror
```

9. **If the disk used to mirror the global namespace is physically connected to more than one node (multiported), 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 -L` command to display the full device ID (DID) pseudo-driver 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 DID pseudo-driver name.

```
# scdidadm -L
...
1         phys-schost-3:/dev/rdsk/c1t1d0     /dev/did/rdsk/d2
# scconf -c -D name=dsk/d2,localonly=true
```

      For more information about the `localonly` property, see the `scconf_dg_rawdisk`(1M) man page.

   b. **Use the `scconf`(1M) command to enable the `localonly` property.**

```
# scconf -c -D name=rawdisk-groupname,localonly=true
```

      `-D name=`*rawdisk-groupname* Specifies the name of the raw disk device group

10. **Do you intend to mirror file systems than cannot be unmounted?**
    - If yes, go to "How to Mirror File Systems That Cannot Be Unmounted" on page 137.
    - If no, go to Step 11.

11. **Do you intend to mirror user-defined file systems?**
    - If yes, go to "How to Mirror User-Defined File Systems" on page 141.
    - If no, go to "How to Create a Diskset" on page 145 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 -f 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 DID name of the mirrored disk's raw disk device group)
# scdidadm -L
...
1         phys-schost-3:/dev/rdsk/c2t2d0     /dev/did/rdsk/d2

(Enable the localonly property of the mirrored disk's raw disk device group)
# scconf -c -D name=dsk/d2,localonly=true
```

▼ How to Mirror File Systems That Cannot Be Unmounted

Use this procedure to mirror file systems 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.**

```
# metainit -f submirror1 1 1 diskslice
```

3. **Create a second concatenation.**

```
# metainit -f submirror2 1 1 submirror-diskslice
```

4. **Create a one-way mirror with one submirror.**

```
# metainit mirror -m submirror1
```

Note – The metadevice name for the mirror does *not* need to be unique throughout the cluster.

---

5. **Repeat Step 1 through Step 4 for each unmountable file system to be mirrored.**

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. **Evacuate any resource groups or device groups from the node.**

```
# scswitch -S -h node
```

`-S`	Evacuates all resource groups and device groups
`-h` *node*	Specifies the name of the node from which to evacuate resource or device groups

8. **Reboot the node.**

```
# shutdown -g0 -y -i6
```

9. **Attach the second submirror to each mirror.**

   This attachment starts a sync of the submirrors.

```
# metattach mirror submirror2
```

10. **Wait for the sync of the mirrors, started in Step 9, to complete.**

    Use the `metastat`(1M) command to view mirror status.

```
# metastat mirror
```

11. **If the disk used to mirror the unmountable file system is physically connected to more than one node (multiported), 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 (DID) pseudo-driver 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 DID pseudo-driver name.

```
# scdidadm -L
...
1            phys-schost-3:/dev/rdsk/c1t1d0    /dev/did/rdsk/d2
# scconf -c -D name=dsk/d2,localonly=true
```

For more information about the `localonly` property, see the `scconf_dg_rawdisk`(1M) man page.

b. Use the `scconf`(1M) command to enable the `localonly` property.

```
# scconf -c -D name=rawdisk-groupname,localonly=true
```

`-D name=`*rawdisk-groupname* Specifies the name of the raw disk device group

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

- If yes, go to "How to Mirror User-Defined File Systems" on page 141.
- If no, go to "How to Create a Diskset" on page 145 to create a diskset.

Example—Mirroring Unmountable File Systems

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 -f 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


(Reboot the node)
# scswitch -S -h phys-schost-1
# 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 DID name of the mirrored disk's raw disk device group)
# scdidadm -L
...
1         phys-schost-3:/dev/rdsk/c2t2d0     /dev/did/rdsk/d2


(Enable the localonly property of the mirrored disk's raw disk device group)
# scconf -c -D name=dsk/d2,localonly=true
```

▼ How to Mirror User-Defined File Systems

Use this procedure to mirror user-defined file systems. In this procedure, nodes do not need to be rebooted.

1. **Become superuser on a node of the cluster.**
2. **Put the slice on which a user-defined file system resides in a single slice (one-way) concatenation.**

```
# metainit -f submirror1 1 1 diskslice
```

3. **Create a second concatenation.**

```
# metainit -f submirror2 1 1 submirror-diskslice
```

4. **Create a one-way mirror with one submirror.**

```
# metainit mirror -m submirror1
```

Note – The metadevice name for the mirror does *not* need to be unique throughout the cluster.

---

5. **Repeat Step 1 through Step 4 for each user-defined file system to be mirrored.**
6. **On each node, edit the `/etc/vfstab` file entry for each user-defined 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. **Attach the second submirror to the mirror.**

   This attachment starts a sync of the submirrors.

```
# metattach mirror submirror2
```

8. **Wait for the sync of the mirrors, started in Step 7, to complete.**

   Use the `metastat`(1M) command to view mirror status.

```
# metastat mirror
```

9. **If the disk used to mirror the user-defined file system is physically connected to more than one node (multiported), 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 (DID) pseudo-driver 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 DID pseudo-driver name.

```
# scdidadm -L
...
1         phys-schost-3:/dev/rdsk/c1t1d0     /dev/did/rdsk/d2
# scconf -c -D name=dsk/d2,localonly=true
```

      For more information about the `localonly` property, see the `scconf_dg_rawdisk`(1M) man page.

   b. **Use the `scconf`(1M) command to enable the `localonly` property.**

```
# scconf -c -D name=rawdisk-groupname,localonly=true
```

      `-D name=`*rawdisk-groupname* Specifies the name of the raw disk device group

10. **Create a diskset.**

    Go to "How to Create a Diskset" on page 145.

Example—Mirroring User-Defined File Systems

The following example shows creation of mirror `d4` to mirror `/home`, 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 `/home` is updated to use the mirror name `d4`. Disk `c2t2d0` is a multiported disk, so the `localonly` property is enabled.

```
(Create the mirror)
# metainit -f d14 1 1 c0t0d0s4
d14: Concat/Stripe is setup
# metainit -f 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 /home           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 DID name of the mirrored disk's raw disk device group*)

```
# scdidadm -L
...
1         phys-schost-3:/dev/rdsk/c2t2d0     /dev/did/rdsk/d2
```

(*Enable the localonly property of the mirrored disk's raw disk device group*)

```
# scconf -c -D name=dsk/d2,localonly=true
```

▼ 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 47 for information about the metasets created by SunPlex Manager.

---

1. **Ensure that the diskset you intend to create meets one of the following requirements.**
   - If 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 "Mediators Overview" on page 155 for details on how to set up mediators.
   - If configured with more than two disk strings, ensure that for any two disk strings S1 and S2, the sum of the number of disks on those strings exceeds the number of disks on the third string S3. Stated as a formula, the requirement is that count(S1) + count(S2) > count(S3).
2. **Ensure that `root` is a member of group 14.**

```
# vi /etc/group
...
sysadmin::14:root
...
```

3. **Ensure that the local metadevice state database replicas exist.**

   For instructions, see "How to Create Metadevice State Database Replicas" on page 128.
4. **Become superuser on the cluster node that will master the diskset.**

5. **Create the diskset.**

   This command also 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

6. **Verify the status of the new diskset.**

```
# metaset -s setname
```

7. **Add drives to the diskset.**

   Go to "Adding Drives to a Diskset" on page 147.

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` assigned 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 Drives to a Diskset

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

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

▼ How to Add 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 145.
3. **List the device ID (DID) mappings.**

```
# scdidadm -L
```

   - Choose drives that are shared by the cluster nodes that will master or potentially master the diskset.
   - Use the full DID pseudo-driver names when you add drives to a diskset.

   The first column of output is the DID instance number, the second column is the full path (physical path), and the third column is the full DID pseudo-driver name (pseudo path). A shared drive has more than one entry for the same DID instance number.

   In the following example, the entries for DID instance number 2 indicate a drive that is shared by `phys-schost-1` and `phys-schost-2`, and the full DID 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 drives to the diskset.**

   Use the full DID pseudo-driver name.

```
# metaset -s setname -a DIDname
```

`-a`	Adds the disk drive to the diskset
DIDname	Device ID (DID) name of the shared disk

Note – Do *not* use the lower-level device name (`cNtXdY`) when you add a 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 drives.**

```
# metaset -s setname
```

7. **Do you intend to repartition drives for use in metadevices?**
   - If yes, go to "How to Repartition Drives in a Diskset" on page 149.
   - If no, go to "How to Create an md.tab File" on page 150 to define metadevices by using an `md.tab` file.

Example—Adding Drives to a Diskset

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

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

▼ How to Repartition Drives in a Diskset

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

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

2. **Use the `format(1M)` command to change the disk partitioning for each drive in the diskset.**

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

   - Create a partition 7 starting at cylinder 0 that is large enough to hold a state database replica (approximately 2 Mbytes).
   - Set the `Flag` field in slice 7 to `V_UNMT` (unmountable) and do not set it to read-only.
   - Do not allow slice 7 to overlap any other slice on the disk.

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

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

   Go to "How to Create an md.tab File" on page 150.

▼ 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 metadevices for the disksets you created.

Note – If you are using local metadevices, ensure that local metadevices names are distinct from the device ID (DID) names used to form disksets. For example, if the DID name `/dev/did/dsk/d3` is used in a diskset, do not use the name `/dev/md/dsk/d3` for a local metadevice. This requirement does not apply to shared metadevices, which use the naming convention `/dev/md/`*setname*`/{r}dsk/d#`.

Tip – To avoid possible confusion between local metadevices in a cluster environment, use a naming scheme that makes each local metadevice name unique throughout the cluster. For example, for node 1 choose names from d100-d199, for node 2 use d200-d299, and so on.

---

1. **Become superuser on the cluster node.**
2. **List the DID mappings for reference when you create your `md.tab` file.**

   Use the full DID pseudo-driver 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 path (physical path), and the third column is the full DID pseudo-driver 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 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 disks that will be used for the submirrors, you must back up the data before metadevice setup and restore it onto the mirror.

---

4. **Activate the metadevices defined in the `md.tab` files.**

   Go to "How to Activate Metadevices" on page 153.

Example—Sample `md.tab` File

The following sample `md.tab` file defines the metadevices for the diskset named `dg-schost-1`. The ordering of lines in the `md.tab` file is not important.

```
dg-schost-1/d0 -t dg-schost-1/d1 dg-schost-1/d4
    dg-schost-1/d1 -m dg-schost-1/d2
        dg-schost-1/d2 1 1 /dev/did/rdsk/d1s4
        dg-schost-1/d3 1 1 /dev/did/rdsk/d55s4
    dg-schost-1/d4 -m dg-schost-1/d5
        dg-schost-1/d5 1 1 /dev/did/rdsk/d3s5
        dg-schost-1/d6 1 1 /dev/did/rdsk/d57s5
```

The sample `md.tab` file is constructed as follows.

- The first line defines the trans metadevice `d0` to consist of a master (UFS) metadevice `d1` and a log device `d4`. The `-t` signifies this is a trans metadevice. The master and log devices are specified by their position after the `-t` flag.

```
dg-schost-1/d0 -t dg-schost-1/d1 dg-schost-1/d4
```

- The second line defines the master device as a mirror of the metadevices. The `-m` in this definition signifies a mirror device, and one of the submirrors, `d2`, is associated with the mirror device, `d1`.

```
dg-schost-1/d1 -m dg-schost-1/d2
```

- The fifth line similarly defines the log device, `d4`, as a mirror of metadevices.

```
dg-schost-1/d4 -m dg-schost-1/d5
```

- The third line defines the first submirror of the master device, `d2`, as a one-way stripe.

```
dg-schost-1/d2 1 1 /dev/did/rdsk/d1s4
```

- The fourth line defines the second submirror of the master device, `d3`.

```
dg-schost-1/d3 1 1 /dev/did/rdsk/d55s4
```

- Finally, the log device submirrors, `d5` and `d6`, are defined. In this example, simple metadevices for each submirror are created.

```
dg-schost-1/d5 1 1 /dev/did/rdsk/d3s5
dg-schost-1/d6 1 1 /dev/did/rdsk/d57s5
```

▼ How to Activate Metadevices

Perform this procedure to activate metadevices defined in the `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, 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 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 disks. This step is required for cluster-pair topologies, where the disks are not accessible by all nodes.

8. **Check the status of the metadevices.**

```
# metastat -s setname
```

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

9. **Does your cluster contain disksets configured with exactly two disk enclosures and two nodes?**
   - If yes, those disksets require mediators. Go to "Mediators Overview" on page 155 to add mediator hosts.
   - If no, go to "How to Add Cluster File Systems" on page 79 to create a cluster file system.

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

In the following example, all metadevices 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
```

Mediators Overview

A 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.

Mediators are required for all Solstice DiskSuite disksets configured with exactly two disk strings and two cluster nodes. A *disk string* consists of a disk enclosure, its physical disks, 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, and those two mediator hosts must be the same two cluster nodes 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 have exactly two nodes. Rather, they only require that 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 mediators.

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

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`(7) 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 157.

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 156.**
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 158 to repair the affected mediator host.
   - If no, go to "How to Add Cluster File Systems" on page 79 to create a cluster file system.

▼ How to Fix Bad Mediator Data

Perform this procedure to repair bad mediator data.

1. **Identify the mediator host(s) with bad mediator data as described in the procedure "How to Check the Status of Mediator Data" on page 157.**
2. **Become superuser on the node that owns the affected diskset.**
3. **Remove the mediator host(s) 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 the mediator host.**

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

5. **Create a cluster file system.**

Go to "How to Add Cluster File Systems" on page 79.

Solstice DiskSuite Configuration Example

The following example helps to explain the process for determining the number of disks to place in each diskset when you use Solstice DiskSuite software. In this example three storage devices are used, and existing applications 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 used to determine the number of drives needed in the sample configuration. If you have three storage devices, you would need 28 drives that 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 Gbyte of disk space because the number of disks needed was rounded up.

TABLE A-2 Determining Drives Needed for a Configuration

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

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

TABLE A-3 Division of Disksets

Diskset	Data Services	Disks	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 disks on each storage device (a total of 12 disks) are assigned to `dg-schost-1`, and five or six disks on each (a total of 16) 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).

APPENDIX **B**

Installing and Configuring VERITAS Volume Manager for Sun Cluster Configurations

Install and configure your local and multihost disks for VERITAS Volume Manager (VxVM) by using the procedures in this appendix, along with the planning information in "Planning Volume Management" on page 16. See your VxVM documentation for additional details.

The following procedures are in this appendix.

Installing and Configuring VxVM Software

Before you begin, have available the following information.

- Mappings of your storage disk drives.
- The following completed configuration planning worksheets from the *Sun Cluster 3.0 Release Notes*. See "Planning Volume Management" on page 16 for planning guidelines.
  - "Local File System Layout Worksheet"
  - "Disk Device Group Configurations Worksheet"
  - "Volume Manager Configurations Worksheet"

The following table lists the tasks to perform to install and configure VxVM software for Sun Cluster configurations.

TABLE B-1 Task Map: Installing and Configuring VxVM Software

Task	For Instructions, Go To ...
Plan the layout of your VxVM configuration.	"Planning Volume Management" on page 16
Determine how you will create the `rootdg` disk group on each node.	"Setting Up a rootdg Disk Group Overview" on page 163
Install VxVM software and create the `rootdg` disk group.	
Method 1 – Install VxVM software and encapsulate the root disk by using the `scvxinstall` command, and optionally mirror the encapsulated root disk.	"How to Install VERITAS Volume Manager Software and Encapsulate the Root Disk" on page 164 "How to Mirror the Encapsulated Root Disk" on page 168
Method 2 – Install VxVM software and create `rootdg` on local, non-root disks.	"How to Install VERITAS Volume Manager Software Only" on page 171 "How to Create a rootdg Disk Group on a Non-Root Disk" on page 174
Create shared disk groups and volumes.	"How to Create and Register a Shared Disk Group" on page 176
Resolve any minor number conflicts between disk device groups by assigning a new minor number.	"How to Assign a New Minor Number to a Disk Device Group" on page 178
Verify the shared disk groups and volumes.	"How to Verify the Disk Group Configuration" on page 179
Configure the cluster.	"Configuring the Cluster" on page 78

Setting Up a `rootdg` Disk Group Overview

Each cluster node requires the creation of a `rootdg` 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 `rootdg` disk group must be restricted to only that node.
- Remote nodes must never access data stored in another node's `rootdg`.
- Do not use the `scconf`(1M) command to register the `rootdg` disk group as a shared disk group.
- Whenever possible, configure the `rootdg` for each node on a non-shared disk.

Sun Cluster software supports the following methods to configure the `rootdg` 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 non-root 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 non-root disks for use as `rootdg`. A `rootdg` disk group that is created on local non-root disks is local to that node, neither globally accessible nor highly available. As with the root disk, to encapsulate a non-root 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.

Where to Go From Here

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

- If you intend to encapsulate the root disk, go to "How to Install VERITAS Volume Manager Software and Encapsulate the Root Disk" on page 164.
- If you intend to create the `rootdg` disk group on local non-root disks, go to "How to Install VERITAS Volume Manager Software Only" on page 171.

▼ 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 `rootdg` disk group on local, non-root disks, go instead to "How to Install VERITAS Volume Manager Software Only" on page 171.

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 device(s) VxVM will manage.

Note – Although the `scvxinstall` utility disables Dynamic Multipathing (DMP) at the start of installation processing, DMP is automatically re-enabled by VxVM 3.1.1 when the VRTSvxvm package is installed. Earlier versions of VxVM must still run with DMP disabled.

---

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. **Become superuser on a node you intend to install with VxVM.**
3. **Ensure that the node's `/etc/name_to_major` file does not already contain a `vxio` entry.**

   If a `vxio` entry exists, remove it from the file. This entry might exist if VxVM was previously installed on other nodes of the cluster. The correct `vxio` entry is added automatically during installation of the VxVM software packages.
4. **Insert the VxVM CD-ROM into the CD-ROM drive on the node.**
5. **Start `scvxinstall` in interactive mode.**

   Press Control-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.

   - Disables Dynamic Multipathing (DMP)
   - Installs the `VRTSvxvm`, `VRTSvmdev`, and `VRTSvmman` packages
   - Sets the `vxio` driver major number to `210`
   - Creates a `rootdg` 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 – After the installation tasks are completed, the `scvxinstall` command automatically reboots the node unless you press Control-C when prompted. If you press Control-C, you must reboot the node later to complete VxVM installation.

---

9. **If you intend to enable the VxVM cluster feature, run the `vxlicense` command to supply the cluster feature license key.**

   See your VxVM documentation for information about the `vxlicense` command.

10. **(Optional) Install the** VxVM **GUI.**

```
# pkgadd VRTSvmsa
```

See your VxVM documentation for information about the VxVM GUI.

11. **Eject the CD-ROM.**

12. **Install any VxVM patches.**

See the *Sun Cluster 3.0 U1 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 2 through Step 13.
- If no, go to Step 15.

15. **Are there one or more nodes that you do *not* intend to install with** VxVM**?**

Note – If you intend to enable the VxVM cluster feature, you *must* install VxVM on all nodes of the cluster. Proceed to Step 20.

- If yes, go to Step 16.
- If no, go to Step 20.

16. **Become superuser on a node that you do *not* intend to install with** VxVM**.**

17. **Edit the `/etc/name_to_major` file and add an entry to set the `vxio` major number to `210`.**

```
# vi /etc/name_to_major
vxio 210
```

Note – If you later install this node with VxVM, you must remove the `vxio` entry before you install VxVM.

---

18. **Initialize the `vxio` entry.**

```
# drvconfig -b -i vxio -m 210
```

Note – The next time you reboot this node, you might see messages similar to the following. These messages are harmless and can be ignored.

```
/sbin/rcS: /usr/sbin/vxrecover: not found
/etc/rc2.d/S75MOUNTGFSYS: /usr/sbin/vxdctl: not found
```

19. **Repeat Step 16 through Step 18 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.

20. **Do you intend to mirror the encapsulated root disk?**

    - If yes, go to "How to Mirror the Encapsulated Root Disk" on page 168.
    - If no, go to "How to Create and Register a Shared Disk Group" on page 176.

Note – If you later need to unencapsulate the root disk, follow the procedures in "How to Unencapsulate the Root Disk" on page 180.

▼ 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 following the procedures in your VxVM documentation.**

   For maximum availability and simplified administration, use a local disk for the mirror. See "Mirroring the Root Disk" on page 21 for additional guidelines.

   Caution – Do not use a quorum device to mirror a root disk. Doing so might prevent the node from booting from the root disk mirror under certain circumstances.

   ---

2. **Display the device ID (DID) mappings.**

   ```
   # scdidadm -L
   ```

3. **From the DID mappings, locate the disk used to mirror the root disk.**

4. **Extract the raw disk device group name from the DID 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/cNtXdY      /dev/did/rdsk/dN
   ```

5. **View the node list of the raw disk device group.**

   Output will look 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, go to Step 7.
   - If no, go 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 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 shared disk groups.**

    Go to "How to Create and Register a Shared Disk Group" on page 176.

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 for the raw disk device group)
# scconf -r -D name=dsk/d2,nodelist=phys-schost-3

(Enable the localonly property for the mirror disk's raw disk device group)
# scconf -c -D name=dsk/d2,localonly=true
```

▼ 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 `rootdg` disk group by encapsulating the root disk, do not use this procedure. Instead, go to "How to Install VERITAS Volume Manager Software and Encapsulate the Root Disk" on page 164 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 device(s) VxVM will manage.

Note – Although the `scvxinstall` utility disables Dynamic Multipathing (DMP) at the start of installation processing, DMP is automatically re-enabled by VxVM 3.1.1 when the VRTSvxvm package is installed. Earlier versions of VxVM must still run with DMP disabled.

---

1. **Ensure that all nodes in the cluster are running in cluster mode.**
2. **Become superuser on a cluster node you intend to install with VxVM.**
3. **Ensure that the node's `/etc/name_to_major` file does not already contain a `vxio` entry.**

   If a `vxio` entry exists, remove it from the file. This entry might exist if VxVM was previously installed on other nodes of the cluster. The correct `vxio` entry is added automatically during installation of the VxVM software packages.
4. **Insert the VxVM CD-ROM into the CD-ROM drive on the node.**
5. **Start `scvxinstall` in interactive installation mode.**

```
# scvxinstall -i
```

   The `scvxinstall` command automatically performs the following tasks.

   - Disables Dynamic Multipathing (DMP)
   - Installs the `VRTSvxvm`, `VRTSvmdev`, and `VRTSvmman` packages
   - Sets the `vxio` driver major number to `210`

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

6. **(Optional) Install the** VxVM **GUI.**

```
# pkgadd VRTSvmsa
```

See your VxVM documentation for information about the VxVM GUI.

7. **Eject the CD-ROM.**

8. **Install any VxVM patches.**

See the *Sun Cluster 3.0 U1 Release Notes* for the location of patches and installation instructions.

9. **(Optional) If you prefer not to have** VxVM **man pages reside on the cluster node, remove the man page package.**

```
# pkgrm VRTSvmman
```

10. **Do you intend to install VxVM on another node?**

- If yes, repeat Step 2 through Step 9.
- If no, go to Step 11.

11. **Are there one or more nodes that you do *not* intend to install with** VxVM**?**

Note – If you intend to enable the VxVM cluster feature, you *must* install VxVM on all nodes of the cluster. Proceed to Step 16.

- If yes, go to Step 12.
- If no, go to Step 16.

12. **Become superuser on a node that you do *not* intend to install with** VxVM**.**

13. **Edit the** `/etc/name_to_major` **file and add an entry to set the** `vxio` **major number to** `210`**.**

```
# vi /etc/name_to_major
vxio 210
```

Note – If you later install this node with VxVM, you must remove the `vxio` entry before you install VxVM.

---

14. **Initialize the `vxio` entry.**

```
# drvconfig -b -i vxio -m 210
```

Note – The next time you reboot this node, you might see messages similar to the following. These messages are harmless and can be ignored.

```
/sbin/rcS: /usr/sbin/vxrecover: not found
/etc/rc2.d/S75MOUNTGFSYS: /usr/sbin/vxdctl: not found
```

15. **Repeat Step 12 through Step 14 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.

16. **Create a `rootdg` disk group.**

    Go to "How to Create a rootdg Disk Group on a Non-Root Disk" on page 174.

▼ How to Create a `rootdg` Disk Group on a Non-Root Disk

Use this procedure to create a `rootdg` 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 will 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`(1M) 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 root disk.
   - Choose any disks to add to the `rootdg` disk group.
   - Do not accept automatic reboot.
5. **Evacuate any resource groups or device groups from the node.**

   ```
   # scswitch -S -h node
   ```

`-S`	Evacuates all resource groups and device groups
`-h` *node*	Specifies the name of the node from which to evacuate resource or device groups

6. **Reboot the node.**

   ```
   # shutdown -g0 -y -i6
   ```

7. **Use the `vxdiskadm`(1M) command to add multiple disks to the `rootdg` disk group.**

   The rootdg disk group becomes tolerant of a disk failure when it contains multiple disks. See VxVM documentation for procedures.

8. **Create shared disk groups.**

   Go to "How to Create and Register a Shared Disk Group" on page 176.

▼ How to Create and Register a Shared 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 the *Sun Cluster 3.0 U1 System Administration Guide* 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 up the disk group you add.

1. **Have available the following information.**
   - Mappings of your storage disk drives. See the appropriate *Sun Cluster 3.0 U1 Hardware Guide* chapter on how to perform an initial installation for your storage device.
   - The following completed configuration planning worksheets from the *Sun Cluster 3.0 U1 Release Notes*.
     - "Local File System Layout Worksheet"
     - "Disk Device Group Configurations Worksheet"
     - "Volume Manager Configurations Worksheet"

     See "Planning Volume Management" on page 16 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, create shared VxVM disk groups by using the cluster functionality of VxVM as described in the *VERITAS Volume Manager Administrator's Reference Guide*. Otherwise, create VxVM disk groups by using the standard procedures documented in the VxVM documentation.

   Note – You can use Dirty Region Logging (DRL) to decrease volume recovery time in the event of a node failure. However, DRL might decrease I/O throughput.

   ---

4. **Is the** VxVM **cluster feature enabled?**
   - If yes, go to Step 7. Do not register a disk group as a Sun Cluster disk device group if the VxVM cluster feature is enabled.
   - If no, go to Step 5.

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 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 "How to Assign a New Minor Number to a Disk Device Group" on page 178. 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 displayed by the following command.

```
# scstat -D
```

Note – If you change any configuration information for a VxVM disk group or volume, you must reregister the Sun Cluster disk device group by using `scsetup`. Such configuration changes include adding or removing volumes, as well as changing the group, owner, or permissions of existing volumes. Reregistration after configuration changes ensures that the global namespace is in the correct state. See the *Sun Cluster 3.0 U1 System Administration Guide* for procedures for how to reregister a disk device group.

---

7. **Verify the configuration of your VxVM disk groups and volumes.**

   Go to "How to Verify the Disk Group Configuration" on page 179.

▼ 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 "How to Create and Register a Shared Disk Group" on page 176 to register the disk group as a** Sun Cluster **disk device group.**

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
```

▼ 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 (`rootdg`), and shared 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 shared 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 78.

▼ How to Unencapsulate the Root Disk

Perform this procedure to unencapsulate the root disk.

Note – This procedure is valid for Sun Cluster 3.0 configurations. To unencapsulate the root disk on a Sun Cluster 2.2 configuration, use procedures provided in your VxVM documentation.

---

1. **Ensure that only Solaris root file systems—root (/), swap, the global devices namespace, `/usr`, `/var`, `/opt`, and `/home`—are present on the root disk.**

   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 you intend to unencapsulate.**

3. **Evacuate all resource groups and device groups from the node.**

```
# scswitch -S -h node
```

-S	Evacuates all resource groups and device groups
-h *node*	Specifies the name of the node from which to evacuate resource or device groups

4. **Determine the node ID number of the node.**

```
# clinfo -n
N
```

5. **Unmount the global-devices file system for this node, where *N* is the node ID number 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 `rootdg` disk group.**

```
# vxedit -rf rm rootdiskxNvol
```

Note – All data in the global-devices file system is destroyed when you remove the VxVM volume, but is restored after the root disk is unencapsulated.

8. **Unencapsulate the root disk.**

```
# /etc/vx/bin/vxunroot
```

See your VxVM documentation for details.

9. **Use the `format`(1M) command to add a 100-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 you created in Step 9.**

```
# newfs /dev/rdsk/cNtXdYsZ
```

11. **Determine the device ID (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 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.**

    You do not need to perform a global mount.

```
# 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 devices.**

    VxVM devices are re-created during the next reboot.

```
# scgdevs
```

15. **Reboot the node.**

```
# reboot
```

16. **Repeat this procedure on each node of the cluster to unencapsulate the root disk on those nodes.**