Sun™ Cluster and Sun StorageTek™ Availability Suite 4.0 Software Integration Guide

Please
Recycle

Adobe PostScript

Contents

Preface

The *Sun Cluster and Sun StorageTek Availability Suite 4.0 Software Integration Guide* describes how to integrate the Sun StorageTek™ Availability Suite Remote Mirror and Point-in-Time Copy software in a Sun™ Cluster environment.

This guide is intended for system administrators who have experience with the Solaris™ operating system (Solaris OS), Sun Cluster software, and related disk storage systems.

Note – Before you install the Sun StorageTek Availability Suite software as described in the installation and release documentation in "Related Documentation" on page x, see Chapter 2.

To fully use the information in this document, you must have thorough knowledge of the topics discussed in the books in "Related Documentation" on page x.

How This Book Is Organized

Chapter 1 is an overview of the Sun Cluster and Sun StorageTek Availability Suite software integration.

Chapter 2 provides information on configuring the Sun StorageTek Availability Suite software for use in a Sun Cluster environment.

Chapter 3 describes using the Sun StorageTek Availability Suite software commands in a Sun Cluster environment.

Chapter 4 provides information on exporting, importing, and joining shadows in a Sun Cluster OE.

Using UNIX Commands

Use this section to alert readers that not all UNIX commands are provided.
For example:

This document might not contain information on basic UNIX® commands and procedures such as shutting down the system, booting the system, and configuring devices. Refer to the following for this information:

■ Software documentation that you received with your system

■ Solaris™ Operating System documentation, which is at

http://docs.sun.com

Shell Prompts

Shell	Prompt
C shell	*machine-name*%
C shell superuser	*machine-name*#
Bourne shell and Korn shell	$
Bourne shell and Korn shell superuser	#

Typographic Conventions

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

* The settings on your browser might differ from these settings.

Related Documentation

Application	Title	Part Number
Sun StorageTek Remote Mirror Software	*Sun StorageTek Availability Suite 4.0 Remote Mirror Software Administration Guide*	819-6148
Sun StorageTek Point-in-Time Copy Software	*Sun StorageTek Availability Suite 4.0 Point-in-Time Copy Software Administration Guide*	819-6149
Sun StorageTek Availability Suite Installation and Configuration	*Sun StorageTek Availability Suite 4.0 Software Installation and Configuration Guide*	819-6147
Sun StorageTek Availability Suite Troubleshooting	*Sun StorageTek Availability Suite 4.0 Software Troubleshooting Guide*	819-6151
Sun StorageTek Availability Suite Release Notes	*Sun StorageTek Availability Suite 4.0 Software Release Notes*	819-6152
Sun Cluster Hardware	*Sun Cluster 3.0 U1 Hardware Guide*	806-7070
	Sun Cluster 3.0 12/01 Hardware Guide	816-2023
Sun Cluster Software Installation	*Sun Cluster 3.0 U1 Installation Guide*	806-7069
	Sun Cluster 3.0 12/01 Software Installation Guide	816-2022
Sun Cluster Data Services	*Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide*	806-7071
	Sun Cluster 3.0 12/01 Data Services Installation and Configuration Guide	816-2024
Sun Cluster Concepts	*Sun Cluster 3.0 U1 Concepts*	806-7074
	Sun Cluster 3.0 12/01 Concepts	816-2027
Sun Cluster Error Messages	*Sun Cluster 3.0 U1 Error Messages Manual*	806-7076
	Sun Cluster 3.0 12/01 Error Messages Manual	816-2028
Sun Cluster Release Notes	*Sun Cluster 3.0 U1 Release Notes*	806-7078
	Sun Cluster 3.0 12/01 Release Notes	816-2029

Application	Title	Part Number
	Sun Cluster 3.0 Release Notes Supplement	806-7079
	Sun Cluster 3.0 12/01 Release Notes Supplement	816-3753
Sun Cluster System Administration	*Sun Cluster 3.0 U1 System Administration Guide*	806-7073
	Sun Cluster 3.0 12/01 System Administration Guide	816-2026

Accessing Sun Documentation

You can view, print, or purchase a broad selection of Sun documentation, including localized versions, at:

http://www.sun.com/documentation

Third-Party Web Sites

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

Contacting Sun Technical Support

If you have technical questions about this product that are not answered in this document, go to:

http://www.sun.com/service/contacting

Sun Welcomes Your Comments

Sun is interested in improving its documentation and welcomes your comments and suggestions. You can submit your comments by going to:

http://www.sun.com/hwdocs/feedback

Please include the title and part number of your document with your feedback:

Sun Cluster and Sun StorageTek Availability Suite 4.0 Software Integration Guide, part number 819-6150-10

Overview of Sun Cluster and Sun StorageTek Availability Suite Software

The Sun Cluster and Sun StorageTek Availability Suite software combine to provide a highly available environment for cluster storage. The Remote Mirror software is a data replication application that provides access to data as part of business continuance and disaster recovery plans. The Point-in-Time Copy software is a point-in-time snapshot copy application that enables you to create copies of application or test data.

For a description of the terminology used in this guide, refer to the "Glossary" on page 55.

The topics in this chapter include:

- "Supported Software and Hardware" on page 2
- "Using the Sun StorageTek Availability Suite Software in a Sun Cluster Environment" on page 3
- "Exporting, Importing, and Joining Shadows in a Sun Cluster Operating Environment" on page 6
- "VTOC Information" on page 7

Supported Software and Hardware

TABLE 1-1 Supported Software and Hardware

Operating Environment Software	Solaris 10 Update 1 and higher.
Sun Cluster Software	Sun Cluster 3.1 Update 4 and Sun Cluster 3.2
Devices and Volume Manager Software	Solaris Raw Device `/dev/rdsk/c?t?d?s2` Solaris Volume Manager Sun Cluster Global Device Veritas Volume Manager (VxVM)
Sun StorageTek Availability Suite Software	Sun StorageTek Availability Suite Remote Mirror and Point-in-Time Copy software.
Supported Cluster Configuration	The Sun Cluster 3.0 Update 3 release, the Sun Cluster 3.1 initial release, and the Sun StorageTek Availability Suite software are supported in a two-node cluster environment only.
Hardware	If you plan to install the software from the product CD, a CD-ROM drive connected to the host server where the software is to be installed.
Disk space requirements	Disk space requirements for either SPARC or x86 platforms: 15 Mbytes • Sun StorageTek Availability Suite core packages: 2.2 Mbytes • Remote Mirror packages: 1.1 Mbytes • Point-in-Time Copy packages: 0.6 Mbytes • Availability Suite local configuration file: 5.4 Mbytes • Availability Suite cluster configuration file: 5.4 Mbytes

Using the Sun StorageTek Availability Suite Software in a Sun Cluster Environment

To use Sun Cluster failover features with the Sun StorageTek Availability Suite software, your software environment requires the installation and configuration of Sun Cluster software on all nodes that are sharing storage in use by Availability Suite software.

As a failover data service, Availability Suite can be configured on some, or all nodes in a Sun Cluster Operating Environment. In this environment, the Sun StorageTek Availability Suite software is cluster aware. A software product is Sun Cluster aware if it can coexist within the Sun Cluster environment, failing over and failing back as the device group, or resource group containing a device group and logical host, fails over and fails back between configured nodes. As a Sun Cluster aware product, the Availability Suite data services and associated resource types can be configured as highly available, utilizing the High Availability framework that Sun Cluster provides.

Global and Local Use of the Sun StorageTek Availability Suite Software

---

Note – See "Rules for the Remote Mirror Software" on page 17.

The Sun StorageTek Availability Suite software can use volumes that are local or global devices. Global devices are those Sun StorageTek Availability Suite software or other volumes accessible from any cluster node and which will fail over under the control of the Sun Cluster framework. Local devices are volumes that are local to the individual node (host machine), not defined in a disk device or resource group, and not managed within a cluster file system. Local devices do not fail over and switch back.

To access local devices, use the `C local` or `-C local` options as part of the `sndradm` commands, or the `-C local` option with `iiadm` commands. To access global devices, use the command options `C` *tag* and `-C` *tag*. (Typically, you do not need to specify the `-C` *tag* option as `iiadm` and `sndradm` automatically detect the disk device group.)

See Chapter 3 in this guide and the Sun StorageTek Availability Suite Administration Guides listed in "Related Documentation" on page x.

Switching Over Global Devices Only

The `scswitch`(1M) command enables you to change all resource groups and device groups manually from the primary mastering node to the next preferred node. The Sun Cluster documentation describes how to perform these tasks.

Local devices do not fail over and switch back so do not configure them as part of your cluster. A file system mounted on a volume and designated as a local device must not be configured as a device to fail over and switch back in the Sun Cluster environment.

Volumes Eligible for Use

Note – When creating shadow volume sets, do not create shadow or bitmap volumes using partitions that include cylinder 0 because data loss might occur. See "VTOC Information" on page 7.

You can replicate the following critical volumes using the Remote Mirror software:

- Database and database management system (DBMS) logs (the total database or online DBMS log)
- Access control files

You can exclude volumes from replication if they can be reconstructed at the recovery site or if they seldom change:

- Temporary volumes (such as those used in sort operations)
- Spool files
- Paging volumes

When selecting a volume to be used in the volume set (including the configuration location), ensure that volume does not contain disk label private areas (for example, slice 2 on a Solaris OS-formatted volume). The disk label region is contained in the first sectors of cylinder 0 of a disk.

The Point-in-Time Copy software supports all Sun-supported storage. It works independently of the underlying data reliability software (for example, RAID-1, RAID-5, or volume manager). Additionally, you can use it as a tool when migrating data to and from differing storage types.

Typical uses for the Point-in-Time Copy software include:

- Backup of live application data
- Load data warehouses and fast resynchronization of data warehouses at predefined intervals
- Application development and test on a point-in-time snapshot of live data
- Migrate data across different types of storage platforms and volumes
- Hot back up of application data from frequent point-in-time snapshots

Exporting, Importing, and Joining Shadows in a Sun Cluster Operating Environment

Overview

The Sun StorageTek Availability Suite Point-in-Time Copy software is a volume-based snapshot facility that runs in the Solaris Operating System and allows creation and management of instant snapshot (point-in-time) copies, allowing repurposing of data for applications such as hot backup, decision support, or application development and testing.

This section provides an overview of exporting, importing, and joining shadows in a Sun Cluster OE. For detailed information, refer to "Exporting, Importing, and Joining Shadows in a Sun Cluster OE" on page 41

Requirements

Sun Cluster High Availability support for the Sun StorageTek Availability Suite Point-in-Time Copy software is achieved by using the Sun Cluster `HAStoragePlus` resource type. This resource type can be used to make the global devices or volume manager controlled volumes, and any file systems mounted on those volumes, highly available within a Sun Cluster Operating Environment. When this configured resource type includes volumes that are also configured with the Point-in-Time Copy software, the volumes are also highly available.

The Sun Cluster `HAStoragePlus` resource type is a failover resource type and as such is only active on one Sun Cluster node at a time. When one or more global devices or the file systems mounted on them are configured with the `HAStoragePlus` resource type, they become part of a Sun Cluster resource group and under Sun Cluster Resource Group Manager (RGM) control. The RGM causes a configured resource group to voluntarily or involuntarily failover or switchover to another configured Sun Cluster system. The global devices or volume manager controlled volumes and the file systems mounted on them are unmounted on the current Sun Cluster node and remounted on the failover Sun Cluster node under RGM control.

On a per Point-in-Time Copy set basis, each volume set requires a minimum of three raw global device partitions or volume manager controlled volumes (for example, Solaris Volume Manager or VERITAS Volume Manager).

Note – Sun Cluster DID devices are not supported as master, shadow, or bitmap volumes, due to disk data fencing functionality when Sun Cluster failure events are active. Sun Cluster Global devices that are symmetrical in name to the DID devices are supported.

Redundancy

To provide high availability to the data contained on a Point-in-Time Copy set when using global devices, it is assumed that the master, shadow, and bitmap volumes are on redundant storage, since there is no means today to provide for host-based data service redundancy on raw global devices. If controller-based redundancy is not available, then a Sun Cluster supported volume manager must be used. For performance reasons, it is recommended under all scenarios that bitmap volumes are NOT placed on RAID-5 volumes (either host-based or controller-based), since the nature of bitmap I/O processing within a Point-in-Time Copy volume set can be I/O intensive.

VTOC Information

The Solaris system administrator must be knowledgeable about the virtual table of contents (VTOC) that is created on raw devices by Solaris.

The creation and updating of a physical disk's VTOC is a standard function of Solaris. Software applications like Availability Suite, the growth of storage virtualization, and the appearance of SAN-based controllers have made it easy for an uninformed Solaris system administrator to inadvertently allow a VTOC to become altered. Altering the VTOC increases the possibility of data loss.

Remember these points about the VTOC:

- A VTOC is a software generated virtual table of contents based on the geometry of a device and written to the first cylinder of that device by the Solaris `format`(1M) utility.
- Various software components such as `dd`(1M), backup utilities, Point-in-Time Copy software, and Remote Mirror software can copy the VTOC of one volume to another volume if that volume includes cylinder 0 in its mapping.

- If the VTOC of the source and destination volumes are not 100% identical, then there is a possibility of some type of data loss occurring.

  This data loss may not be initially detectable, but can be detected later when other utilities are used, like fsck(1M).

  When first configuring and validating volume replication, save copies of all affected device's VTOCs using the prtvtoc(1M) utility. The fmthard(1M) utility can be used to restore them later, if necessary.

- When using volume managers like VxVM and SVM, copying between individual volumes created under these volume mangers is safe. VTOC issues are avoided because the VTOC is excluded from volumes created by these volume managers.

- When formatting individual partitions on a raw device, for all partitions except the backup partition, make sure they do not map cylinder 0, which contains the VTOC. When using raw partitions as volumes, you are the volume manager and you need to exclude the VTOC from partitions that you configure.

- When formatting the backup partition of a raw device, make sure that the physical geometries of the source and destination devices are identical. (Partition 2, by default, maps all cylinders under the backup partition.) If identical device sizing is not possible, make sure that the source backup partition is smaller then the destination partition, and that the destination partition does not map cylinder 0.

Configuring the Sun StorageTek Availability Suite Software

Note – This guide assumes that you have already installed the volume manager software and the Sun Cluster software on each node in your cluster.

Caution – Do not install the Sun StorageTek Availability Suite software on a system running the initial release of the Sun Cluster 3.0 software.

The *Sun StorageTek Availability Suite 4.0 Software Installation and Configuration Guide,* listed in "Related Documentation" on page x, describes how to install the Sun StorageTek Availability Suite software in a non-clustered environment. The installation steps to install this software in a Sun Cluster environment are generally the same as described in the installation guides. This chapter describes the differences when you install the software in a Sun Cluster environment.

The topics in this chapter include:

- "Setting the Location of the AVS Cluster-Specific Configuration Database" on page 10
- "Backing Up Configuration Information" on page 12
- "Editing the Bitmap Parameter Files" on page 14
- "Supported Configurations for the Remote Mirror Software" on page 16
- "Supported Configurations for the Point-in-Time Copy Software" on page 19
- "Disk Device Groups and the Sun StorageTek Availability Suite Software" on page 20
- "Handling Raw Devices in a Sun Cluster OE" on page 20
- "Configuring the Sun Cluster Environment" on page 25

Setting the Location of the AVS Cluster-Specific Configuration Database

When installed in a Sun Cluster operating environment, AVS software requires a configuration database for information specific to Sun Cluster. This Sun Cluster configuration database is in addition to the local configuration database, which is still required (see the *Sun StorageTek Availability Suite 4.0 Software Installation and Configuration Guide* for more details).

A single Sun Cluster configuration location must be shared by all cluster nodes, and must be explicitly set on each node. Running dscfgadm with no arguments will prompt you to set the cluster configuration location, if it has not yet been set on that node. The location may also be changed later by running dscfgadm -s on all nodes of the cluster.

The cluster configuration database must be located on a raw slice of a did device. For help in finding an unused did device, the scdidadm -L command shows the local and shared disks by device ID. Refer to TABLE 2-1 for the requirements for this configuration location.

Note – Ensure that the slice does not contain disk label private areas (for example, slice 2 on a Solaris OS-formatted volume). The disk label region is contained in the first sectors of cylinder 0 of a disk. See "VTOC Information" on page 7.

.

TABLE 2-1   Configuration Location Requirements and Considerations

Item	Requirement or Consideration
Location	A raw device on a `did` service. For example: `/dev/did/rdsk/d0s7`.
Availability	• The raw device must be accessible by all nodes of the cluster. • The location must be writable by the superuser user. • The location is available or persistent at system startup and reboots. • The slice used for the configuration database cannot be used by any other application (for example a file system or a database).
Disk space	The configuration location requires 5.5 Mbytes of disk space. If you specify a file for the configuration location during the installation, the file of the appropriate size is automatically created. **Note**: If you specify a volume or a slice for the configuration location, only 5.5 Mbytes of the space is used, the remainder is unused.
Mirroring	Consider configuring RAID (such as mirrored partitions) for the location and ensure that you mirror the location to another disk in the array. The location cannot be stored on the same disk as the replicated volumes.

Backing Up Configuration Information

In addition to the local configuration information (see the *Sun StorageTek Availability Suite 4.0 Software Installation and Configuration Guide*), the cluster-specific configuration information should also be periodically backed up. You may wish to set up a `cron(1M)` job to periodically back up the Sun Cluster configuration database automatically. In addition, back up the configuration information whenever you change your configuration, for example, adding and deleting volumes.

▼ To Back up Configuration Information

1. **On each node of the Sun Cluster, back up the local configuration database (see the *Sun StorageTek Availability Suite 4.0 Software Installation and Configuration Guide*).**

   ```
   # cp /etc/dscfg_local /var/backups/dscfg_local
   ```

2. **On one node of the Sun Cluster, perform the following two steps:**

   a. **Back up the `/etc/dscfg_cluster` reference file, containing the path to the AVS cluster database.**

   ```
   # cp /etc/dscfg_cluster /var/backups/dscfg_cluster
   ```

   b. **Back up the contents of the cluster-specific configuration database to a file, using the `dd` command. The size of the database is 5.5 MB.**

   ```
   # dd if=/dev/did/rdsk/d3s4 of=/var/backups/dscfg_cluster_data
   bs=512k count=11
   ```

3. **To determine the device you have specified for the cluster-specific configuration database, run a `cat` command on the `/etc/dscfg_cluster` file or run `dscfgadm -i`.**

   a.

   ```
   # cat /etc/dscfg_cluster
   /dev/did/rdsk/d3s4
   ```

b. or

```
# dscfgadm -i
SERVICE         STATE           ENABLED
nws_scm         online          true
nws_sv          online          true
nws_ii          online          true
nws_rdc         online          true
nws_rdcsyncd    online          true

Availability Suite Configuration:
Local configuration database: valid
cluster configuration database: valid
cluster configuration location: /dev/did/rdsk/d3s4
```

Editing the Bitmap Parameter Files

Bitmap volumes are used by the Remote Mirror and Point-in-Time Copy software to track differences between volumes and provide information for volume updates. The Sun StorageTek Availability Suite software documentation listed in "Related Documentation" on page x describes the bitmap size and other requirements.

In a Sun Cluster environment, bitmap volumes must be part of the same disk device group or cluster resource group as the corresponding primary host or secondary hosts data volume.

The Remote Mirror and Point-in-Time Copy software include two configuration files that determine how bitmap volumes are written to and saved:

- Remote Mirror software –/usr/kernel/drv/rdc.conf
- Point-in-Time Copy software –/usr/kernel/drv/ii.conf

Note – The Sun StorageTek Availability Suite Remote Mirror and Point-in-Time Copy software do not support bitmap files. The software uses regular raw devices to store bitmaps. These raw devices must be located on a disk separate from the disk that contains your data.

Setting the Bitmap Operation Mode

A bitmap maintained on disk can persist across a system crash, depending on the setting of rdc_bitmap_mode in /usr/kernel/drv/rdc.conf. The default setting is 0. Set the bitmap mode to 1, as in the following example:

● **Open the `rdc.conf` file and locate the following section. Edit the value for the bitmap mode, save the file, and close it.**

```
# rdc_bitmap_mode
# - Sets the mode of the RDC bitmap operation, acceptable values are:
#   0 - autodetect bitmap mode depending on the state of SDBC (default).
#   1 - force bitmap writes for every write operation, so an update resync
#       can be performed after a crash or reboot.
#   2 - only write the bitmap on shutdown, so a full resync is
#       required after a crash, but an update resync is required after
#       a reboot.
#
rdc_bitmap_mode=1;
```

The `/usr/kernel/drv/ii.conf` File

The `/usr/kernel/drv/ii.conf` file contains one setting that sets the Point-in-Time Copy bitmap save mode:

- `ii_bitmap` – modify to change how the bitmap volume is saved during a shut down or system crash. In a Sun Cluster environment, set this to `1`.

A bitmap maintained on disk can persist across a system crash when this field is set to `1`.

▼ To Edit the `ii.conf` File

1. **Open the `/usr/kernel/drv/ii.conf` file using a text editor such as `vi`(1).**

2. **In a Sun Cluster environment, set the bitmap mode to `1`.**

   For example:

```
# bitmap volume storage strategy:
# 0 indicates kernel memory loaded from bitmap volume when shadow is resumed
#   and saved to bitmap volume when shadow is suspended.
# 1 indicates permanent SDBC storage, bitmap volume is updated directly as
#   bits are changed.
# 2 indicates that if FWC is present strategy 1 is used, otherwise strategy 0.
ii_bitmap=1;
```

3. **Save and exit the file.**

4. **Disable and re-enable the data services as follows:**

```
# dscfgadm -d
# dscfgadm -e
```

Supported Configurations for the Remote Mirror Software

Adding Host Names

This step ensures that the host names in the `/etc/hosts` file are read and known by machines running the Availability Suite software.

▼ To Edit the `/etc/hosts` File

Place the names and IP addresses of all machines you plan to use with the Remote Mirror software in the `/etc/hosts` file. Make sure you include the logical host names and IP addresses of the logical hosts you plan to use with the Remote Mirror software in the `/etc/hosts` file.

● **Add the names and IP addresses of all machines you plan to use with the remote mirror software to the** `/etc/hosts` **file.**

Edit this file on each machine where you are installing and running the remote mirror software.

Using Autosynchronization

Consider the following when using autosynchronization with Sun Cluster:

■ If you want automatic resynchronization to occur in the event of a cluster failover, turn on the autosync feature. With this feature enabled, any cluster failover will automatically put the remote mirror volume sets back into replication mode after an update occurs.

■ If you want to manually force clusters to failover, you must be careful to allow all remote mirror components to fully recognize the condition, including the remote host that is not in the cluster. In practice, this means that you should not attempt an immediate update sync after the failover. You should wait at least thirty seconds after the completion of the scswitch command and before starting an update sync to allow time for Sun Cluster to complete its logical host interface reconfiguration.

Rules for the Remote Mirror Software

- **The primary volume and its bitmap volume and possible disk queue volume or the secondary volume and its bitmap volume in a remote mirror volume set must reside in the same disk device group per node.** (A remote mirror volume set also includes information about primary and secondary hosts and operating mode.)

  For example, you cannot have a primary volume with a disk device group name of `sndrdg` and a primary bitmap volume with a disk device group name of `sndrdg2` in the same remote mirror volume set.

- **With the Remote Mirror software, you can use more than one disk device group for cluster switchover and failover,** but each primary or secondary disk device component in the cluster node's volume set must reside in the same disk device group.

- **The Remote Mirror software requires the creation of a resource group containing a disk device group and logical failover host.** The resource group must contain a single instance of the `SUNW.HAStoragePlus` resource type, referencing the device group of the Remote Mirror primary or secondary data and bitmap volumes, plus a single instance of the `SUNW.LogicalHostname` resource type, referencing the failover IP address associated with the Remote Mirror's primary or secondary host name, plus and any other available resource types.

  To improve the failover affinity between a configured resource group name and the Remote Mirror device group configured within, it is suggested that the resource group name you specify consist of the disk device group name appended with `-stor-rg`. For example, if the group name is `sndrdg`, then the resource group name would be `sndrdg-stor-rg`.

- **Remote mirror replication within the cluster is not supported**. An example is when the primary host is cluster node 1 and the secondary host is cluster node 2 in the cluster and the primary, secondary, and bitmap volumes in a volume set reside in the same disk device group.

  Typically, the remote mirror primary host is part of one cluster configuration, while the replicating secondary host might or might not be part of a different cluster.

Three configurations for the Remote Mirror software are supported:

Remote Mirror Primary Host On a Cluster Node

In this configuration, the remote mirror primary host is the logical host you created in the remote mirror resource group for the remote mirror disk group using the `scrgadm` command; for example, see "To configure Sun Cluster for HAStorage or HAStoragePlus" on page 26.

If you have configured the remote mirror autosynchronization feature on the primary host, the Remote Mirror software starts an update resynchronization from the primary host for all affected remote mirror volume sets following a switchover or failover event, if the autosynchronization feature is enabled for those volume sets. This operation is performed after the resource group and network switchover operation is complete. See the `sndradm` man page and the *Sun StorageTek Availability Suite 4.0 Remote Mirror Software Administration Guide* for a description of the `sndradm -a` command to set the autosynchronization feature.

Remote Mirror Secondary Host On a Cluster Node

In this configuration, the remote mirror secondary host is the logical host you created in the remote mirror resource group for the remote mirror disk group using the `scrgadm` command; for example, see "To configure Sun Cluster for HAStorage or HAStoragePlus" on page 26.

Operations such as update resynchronizations occur and are issued from the primary host machine. Following a switchover or failover event, the Remote Mirror software attempts to start an update resynchronization for all affected remote mirror volume sets, if the autosynchronization feature is enabled for those volume sets. However, the remote mirror secondary host in a remote mirror volume set cannot initiate an update resynchronization.

This operation is performed after the resource group and network switchover operation is complete. In this case, the remote mirror secondary host switchover appears to be a short network outage to the remote mirror primary host.

If you have configured the remote mirror autosynchronization feature on the primary host, the `sndrsyncd` synchronization daemon attempts to resynchronize the volume sets if the system reboots or link failures occur. See the `sndradm` man page and the *Sun StorageTek Availability Suite 4.0 Remote Mirror Software Administration Guide* for a description of the `sndradm -a` command to set the autosynchronization feature.

If this feature is disabled (its default setting) and volume sets are logging but not replicating, perform the updates manually using the `sndradm` command.

Remote Mirror Primary and Secondary Hosts On a Cluster Node

Remote mirror replication within the cluster is not supported. That is, remote mirror replication when the primary and secondary hosts reside in the same cluster and the primary, secondary, and bitmap volumes in a volume set reside in the same disk device group.

If the remote mirror primary and secondary hosts are configured in different clusters, for operating considerations see "Remote Mirror Primary Host On a Cluster Node" on page 18 and "Remote Mirror Secondary Host On a Cluster Node" on page 18.

Supported Configurations for the Point-in-Time Copy Software

Following are some rules for Point-in-Time Copy software:

- **All Point-in-Time Copy volume set components must reside in the same disk device group, except when Export / Import / Join is also being used in a Sun Cluster.** A Point-in-Time Copy volume set includes the master, shadow, bitmap, and optional overflow volumes.

- **With the Point-in-Time Copy software, you can use more than one disk device group for cluster switchover and failover**. Each component in the volume set must reside in the same disk device group.

  For example, you cannot have a master volume with a disk device group name of `ii-group` and a shadow volume with a disk device group name of `ii-group2` in the same volume set.

- **If a Solaris OS failure or Sun Cluster failover occurs during a point-in-time copy or update operation to the master volume,** specifically where the shadow volume is copying (`iiadm -c m`) or updating (`iiadm -u m`) data to the master volume, the master volume might be in an inconsistent state (that is, the copy or update operation might be incomplete). "Preserving Point-in-Time Copy Volume Data" on page 40 describes how to avoid this situation.

Disk Device Groups and the Sun StorageTek Availability Suite Software

The Solstice DiskSuite™ and VERITAS Volume Manager (VxVM) can arrange disk devices into a group to be mastered by a cluster node. You can then configure these disk device groups to fail over to another cluster node, as described in "Configuring the Sun Cluster Environment" on page 25.

The Solstice DiskSuite and VxVM device paths contain the disk device group. When operating in a Sun Cluster environment, the Sun StorageTek Availability Suite commands `sndradm` and `iiadm` automatically detect and use the disk device group as configured in "Configuring the Sun Cluster Environment" on page 25.

You can also use the `sndradm` and `iiadm` commands to select specified disk device groups or to operate on a volume set as a local node-only configuration entry. See "Using the Sun StorageTek Availability Suite `iiadm` and `sndradm` Commands" on page 31

Handling Raw Devices in a Sun Cluster OE

The Sun StorageTek Availability Suite supports the use of raw devices in a Sun Cluster OE (Operating Environment), although the volume configuration procedures to handle raw device are very different from devices under the control of VERITAS Volume Manager (VxVM) or Solaris Volume Manager (SVM).

In a Sun Cluster OE, the Availability Suite supports raw devices only through the use of Sun Cluster Global Devices (as in `/dev/global/rdsk/d8s0`). Availability Suite can not use Sun Cluster DID devices (such as `/dev/did/rdsk/d8s0`), because Sun Cluster's disk fencing and data pathing software causes connectivity problems when directly using DID devices.

For example:

```
iiadm -e ind /dev/global/rdsk/d8s0 /dev/global/rdsk/d8s1 /dev/global/rdsk/d8s2
```

When the Point-in-Time Copy software is configured on a Global Device, any node in a Sun Cluster can concurrently access the data on the Point-in-Time Copy software master or shadow volume. One node in the Sun Cluster has direct I/O access; all other nodes have cluster-interconnect I/O access to the Point-in-Time Copy software sets that are configured.

Depending on Point-in-Time Copy software set usage patterns, when a specific node is dominating the I/O workload of either the master or shadow volume, the Global Device can be moved to the node of high application usage as follows:

```
scswitch -z -D dsk/d<n> -h Sun-Cluster-Nodename
```

Doing so will improve performance and decrease the amount of cluster interconnect traffic.

Global devices are very much like SVM and VxVM volumes, and can be switched between nodes. Although Availability Suite is supported on Global Devices, because of the default DID device behavior of Sun Cluster, each global device is in its own "named" Sun Cluster device group. For example, /dev/global/rdsk/d8 is in device group "dsk/d8". This default Sun Cluster behavior (all constituent volumes in a Point-in-Time Copy must be in the same Sun Cluster device group) would force you to put the Point-in-Time Copy software master, shadow and bitmap volume into the same Global Device, yielding very poor I/O performance.

By default in a Sun Cluster OE, all DID devices are mapped to global devices of the same name. For instance the following DID devices:

```
/dev/did/rdsk/d6, dsk/d7 & dsk/d8
```

are mapped to the following global devices:

```
/dev/global/rdsk/d6, dsk/d7 & dsk/d8
```

If you tried to put a Point-in-Time Copy software master, shadow and bitmap on separate global devices, this enable operation would fail as follows because of the constituent volume rule stated above.

▼ To Create a Named Global Device Group

1. **Bring the devices to the node.**

   ```
   # scswitch -z -D dsk/d6,dsk/d7,dsk/d8 -h Sun-Cluster-Nodename
   ```

2. **Try to enable a Point-in-Time Copy set spanning the device groups.**

   In this example, the groups are dsk/d6, dsk/d7, and dsk/d8.

   ```
   # iiadm -e dep /dev/global/rdsk/d6s0 /dev/global/rdsk/d7s0
   /dev/global/rdsk/d8s0
   iiadm: Volumes are not in same disk group:
   ```

   To resolve this issue, you need to reconfigure these different global devices into a "named" global device group.

3. **Sun Cluster "switch" the device groups into maintenance mode.**

   ```
   # scswitch -m -D dsk/d6
   # scswitch -m -D dsk/d7
   # scswitch -m -D dsk/d8
   ```

4. **Sun Cluster unconfigure the old names from the configuration.**

   ```
   # scconf -r -D name=dsk/d6
   # scconf -r -D name=dsk/d7
   # scconf -r -D name=dsk/d8
   ```

5. **Define a new named device group configuration name containing the DID devices.**

   In this example, the named group is "AVSuite."

   ```
   # scconf -a -D type=rawdisk, name=AVSuite,
   nodelist=Sun-Cluster-Node1, Sun-Cluster-Node2, .., Sun-Cluster-NodeN,
   preferenced=false,
   failback=disabled, numsecondaries=,
   globaldev=d6, globaldev=d7, globaldev=d8
   ```

6. **Bring the named global device group (AVSuite) to the current Sun Cluster node.**

   In this example, the named group is "AVSuite.**"**

   ```
   # scswitch -z -D AVSuite -h Sun-Cluster-Node1
   ```

7. **Try again to enable a Point-in-Time Copy set spanning the named groups.**

   In this example, the named groups are dsk/d6, dsk/d7, and dsk/d8.

   ```
   # iiadm -e dep /dev/global/rdsk/d6s0 /dev/global/rdsk/d7s0
   /dev/global/rdsk/d8s0
   ```

   This time, the command is successful.

8. **Verify that the cluster tag was set successfully.**

   Note that after running iiadm -i /dev/global/rdsk/d7s0, the cluster tag is
   AVSuite, and not dsk/d7.

   ```
   # iiadm -i /dev/global/rdsk/d7s0

   /dev/global/rdsk/d6s0: (master volume)
   /dev/global/rdsk/d7s0: (shadow volume)
   /dev/global/rdsk/d8s0: (bitmap volume)
   Cluster tag: AVSuite
   Dependent copy
   Volume size: 212127744
   Shadow chunks total: 3314496 Shadow chunks used: 0
   Percent of bitmap set: 0
           (bitmap clean)
   ```

9. **The newly named global device can now be switched to other Sun Cluster nodes,
   using its new device group name.**

   ```
   # scswitch -z -D AVSuite -h Sun-Cluster-Node2
   ```

▼ To Remove a Named Global Device Group

After you have created "named" global devices, there may be a need to remove them at some future time. Following are the steps to restore the Global Device configuration back to its initial state.

1. **Disable the use of the "named" Global Devices.**

   ```
   # iiadm -d /dev/global/rdsk/d7s0
   ```

2. **Sun Cluster "switch" the device groups into maintenance mode.**

   In this example, the named group is "AVSuite."

   ```
   # scswitch -m -D AVSuite
   ```

3. **Sun Cluster unconfigure the named global device configuration.**

   ```
   # scconf -r -D name=AVSuite
   ```

4. **Define a (default) dsk/dn device group configuration name for all previously configured DID devices.**

```
# scconf -a -D type=rawdisk, name="dsk/d6",
nodelist=Sun-Cluster-Node1,Sun-Cluster-Node2,..,Sun-Cluster-NodeN,
preferenced=false,failback=disabled, numsecondaries=, globaldev=d6
# scconf -a -D type=rawdisk, name="dsk/d7",
nodelist=Sun-Cluster-Node1,Sun-Cluster-Node2,..,Sun-Cluster-NodeN,
preferenced=false, failback=disabled, numsecondaries=, globaldev=d7
# scconf -a -D type=rawdisk, name="dsk/d8",
nodelist=Sun-Cluster-Node1,Sun-Cluster-Node2,..,Sun-Cluster-NodeN,
preferenced=false, failback=disabled, numsecondaries=, globaldev=d8
```

5. **Bring the named global device groups to the current Sun Cluster node.**

   ```
   # scswitch -z -D dsk/d6,dsk/d7,dsk/d8 -h Sun-Cluster-Node1
   ```

6. Verify that the original restriction is now back.

```
# iiadm -e dep /dev/global/rdsk/d6s0 /dev/global/rdsk/d7s0
/dev/global/rdsk/d8s0
iiadm: Volumes are not in same disk group:
```

Configuring the Sun Cluster Environment

The procedures in this section describe how to configure the Sun Cluster software for use with the Remote Mirror and Point-in-Time Copy software. The *Sun Cluster 3.2 Data Installation and Configuration Guide* contains more information about configuring and administering Sun Cluster data services. See the scrgadm(1M) and scswitch(1M) man pages for more information.

The general configuration steps are:

1. Log on to any node in the cluster.

2. Configure a disk device group using your volume manager.

3. Register the SUNW.HAStorage or SUNW.HAStoragePlus resource type.

4. Create a resource group.

5. Add SUNW.HAStorage or SUNW.HAStoragePlus to the disk device group.

6. (Remote Mirror step only) Add a logical failover host to the resource group.

7. Enable the resource group and bring it online.

▼ To configure Sun Cluster for HAStorage or HAStoragePlus

1. **Log on as the root user on any node in the cluster.**

2. **Configure a disk device group using your volume manager software.**

   See the documentation that came with your volume manager software. Also you might check the currently configured groups before configuring a new disk device group. For example, use the metaset(1M), vxdg, or vxprint commands, depending on your volume manager software.

3. **Register** SUNW.HAStorage **or** SUNW.HAStoragePlus **as a resource type.**

   ```
   # scrgadm -a -t SUNW.HAStorage
   ```

   ```
   # scrgadm -a -t SUNW.HAStoragePlus
   ```

   Note – For information on HAStorage and HAStoragePlus resource types, please refer to the SunCluster documentation. The *SunCluster 3.0 5/02 Supplement* (Part No. 816-3380-10) contains detailed information.

   ---

4. **Create a resource group for the** *devicegroup.*

   ```
   # scrgadm -a -g devicegroup-stor-rg -h node1,node2
   ```

devicegroup	is the required disk device group name.
–h *node1,node2*	specifies the cluster nodes that can master this resource group. If you do not specify these nodes, it defaults to all the nodes in the cluster.

5. **For a** SUNW.HAStorage **resource, use the following command to add the resource to the resource group.**

   ```
   # scrgadm -a -j devicegroup-stor -g devicegroup-stor-rg \
   -t SUNW.HAStorage \
   -x ServicePaths=devicegroup -x AffinityOn=True
   ```

devicegroup	Disk device group name.
`-x ServicePaths=`	Specifies the extension property that the Sun StorageTek Availability Suite software relies on. In this case, use the disk device *devicegroup*.
`-x AffinityOn=True`	Specifies that the `SUNW.HAStorage` resource needs to perform an affinity switchover for the global devices and cluster file systems defined in `-x ServicePaths`.
	It also enforces co-location of resource groups and disk device groups on the same node, thus enhancing the performance of disk-intensive data services.
	If the device group is switched to another node while the `SUNW.HAstorage` resource is online, `AffinityOn` has no effect and the resource group does not migrate along with the device group. On the other hand, if the resource group is switched to another node, `AffinityOn` being set to `True` causes the device group to follow the resource group to the new node.

6. **For a a** `SUNW.HAStoragePlus` **resource, use the following command to add the resource to the resource group.**

```
# scrgadm -a -j devicegroup-stor -g devicegroup-stor-rg \
-t SUNW.HAStoragePlus \
-x GlobalDevicePaths=devicegroup -x AffinityOn=True
```

`-x GlobalDevicePaths=`	specifies the extension property that the Sun StorageTek Availability Suite software relies on. In this case, use the disk device *devicegroup*.
`-x AffinityOn=True`	specifies that the `SUNW.HAStoragePlus` resource needs to perform an affinity switchover for the global devices and cluster file systems defined in `-x GlobalDevicePaths`.
	It also enforces co-location of resource groups and disk device groups on the same node, thus enhancing the performance of disk-intensive data services.
	If the device group is switched to another node while the `SUNW.HAstoragePlus` resource is online, `AffinityOn` has no effect and the resource group does not migrate along with the device group. On the other hand, if the resource group is switched to another node, `AffinityOn` being set to `True` causes the device group to follow the resource group to the new node.

7. **Add a logical hostname resource to the resource group.**

> **Note –** Perform this step for the remote mirror volumes only. This step is not needed for Point-in-Time Copy volumes.

```
# scrgadm -a -L [-j lhost-stor] -g devicegroup-stor-rg \
-l lhost1, lhost2,..., lhostN
-n nafo0@node,nafo0@node
```

-j *lhost*-stor	Optional resource *lhost*-**stor**. If you do not specify this option and resource, the name defaults to the first logical hostname specified in the -l option.
-l *lhost1,lhost2,...lhostN*	Specifies a comma-separated list of UNIX hostnames (logical hostnames) by which clients communicate with the Sun StorageTek Availability Suite software in the resource group.
-n nafo0@*node*,nafo0@*node*	Specifies the comma-separated list of Network Adapter Failover (NAFO) groups on each node.
	node can be a node name or ID. You can display the node ID using the scconf -p command.

8. **Enable the resources in the resource group, manage the resource group, and bring the resource group online.**

```
# scswitch -Z -g devicegroup-stor-rg
```

9. **Verify that the resource is online.**

   a. **Run the following command on any cluster node.**

```
# scstat -g
```

   b. **Look for the resource group state field to determine if the resource group is online on the nodes specified in the node list.**

10. **For the HAStoragePlus resource, verify that the resource group can be failed between nodes.**

```
# scswitch -z -g devicegroup-stor-rg -h fail-to node
```

The above command fails the resource group to the specified node.

```
# scswitch -S -h fail-from node
```

The above command fails ALL resources from the specified node.

Configuring the HAStoragePlus Resource Types with Volume Sets

This example shows how to configure a resource group on a locally mounted Sun Cluster global device partition.

You can configure the HAStoragePlus resource to fail over resource groups as well as individual volume sets to another node in the cluster. When configuring a resource type with volume sets, consider the following:

- When you add a new volume set to the Sun StorageTek Availability Suite software, you must disable the configured resource group and place it offline.

- You must specify each volume in the set. For example, the following command shows how to define a volume set to an existing resource group using the HAStoragePlus resource.

```
# scrgadm -a -j iidg-rs -g iidg -t SUNW.HAStoragePlus \
-x GlobalDevicePaths=/dev/vx/rdsk/iidg/ii01,/dev/vx/rdsk/ii02, \
/dev/vx/rdsk/iidg/ii11,/dev/vx/rdsk/iidg/ii12,/dev/vx/rdsk/iidg/iibitmap1, \
/dev/vx/rdsk/iidg/iibitmap2
```

where:

`-j iidg-rs`	is the resource name.
`-g iidg`	is the resource group name.
`-x GlobalDevicePaths=`	specifies the extension property `GlobalDevicePath` and raw device volume names for the point-in-time copy volume set.

Using the Sun StorageTek Availability Suite `iiadm` and `sndradm` Commands

This chapter describes using the Sun StorageTek Availability Suite commands `iiadm` and `sndradm` in a Sun Cluster environment. The Sun StorageTek Availability Suite Software Administration Guides, listed in "Related Documentation" on page x, describe the full command syntax and options for `iiadm` and `sndradm`.

The Sun StorageTek Availability Suite software can use volumes that are global or local devices.

■ *Global devices* are Sun StorageTek Availability Suite or other volumes accessible from any cluster node and which fail over and switch back under the control of the Sun Cluster framework.

■ *Local devices* are Sun StorageTek Availability Suite software volumes that are local to the individual node (host machine), not defined in a disk or resource group, and not managed within a cluster file system. Local devices do not fail over and switch back.

The topics in this chapter include:

Mounting and Replicating Global Volume File Systems

If a volume contains a file system and you wish to replicate the file system using the Sun StorageTek Availability Suite software, you must create and mount a related global file system on all cluster nodes. These steps ensure that the file system is available to all nodes and hosts when you copy or update the volume sets.

Note – See the Sun Cluster documentation for information about administering cluster file systems, including creating and mounting global file systems. See also the mount(1M) and mount_ufs(1M) commands.

Use the following steps to create and mount a related global file system on all cluster nodes:

1. **Create the file systems on the appropriate diskset metadevices or disk group volumes.**

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

   For example, using the VERITAS Volume Manager, you might specify *raw-disk-device* as /dev/vx/rdsk/sndrdg/vol01.

2. **On each node, create a mount point directory for the file system.**

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

   - *device-group* is the name of the directory that corresponds to the name of the device group that contains the device.
   - *mount-point* is the name of the directory on which to mount the file system.

3. **On each node, add an entry to the** /etc/vfstab **file for the mount point and use the** global **mount option.**

4. **On a cluster node, use** sccheck**(1M) to verify the mount points and other entries.**

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

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

6. **Verify that the file system is mounted using the `mount` command with no options.**

Global Device Command Syntax

> **Note –** During the initial enable of the remote mirror or point-in-time copy volume sets, you can optionally specify the global device disk group with the `-C` *tag* cluster option when you use the `iiadm` or `sndradm` commands. As this section shows, however, you do not have to use the `-C` *tag* cluster option. Also see "The C tag and -C tag Options" on page 33.

The Sun StorageTek Availability Suite software automatically derives the disk device group name from the volume path when you first enable volume sets. During this initial enable operation, the Remote Mirror and Point-in-Time Copy software creates a configuration entry for each volume set. Part of the entry is the disk device group name for use in a cluster.

The Remote Mirror software shows this name as C *tag*, where *tag* is the disk device group name. The Point-in-Time Copy software shows this name as `Cluster tag:` *tag*.

The C *tag* and -C *tag* Options

C *tag* is displayed as part of a volume set's configuration information as shown in "Global Device Command Syntax" on page 33.

Typically, the Sun StorageTek Availability Suite software derives the disk device group name from the volume path and does not require the `-C` *tag* option.

Use the `-C` *tag* option and C *tag* volume set option to execute the `iiadm` and `sndradm` commands on the enabled volume sets in the disk device group name *tag*, when the disk device group name is not indicated by the volume path. The commands are not executed on any other volume sets in your configuration; `-C` *tag* excludes those volume sets not contained in the *tag* disk device group from the specified operation.

For example, the following command makes a point-in-time copy volume set in the `iigrp2` disk device group wait for all copy or update operations to finish before you can issue other point-in-time copy commands.

```
# iiadm -w /dev/vx/rdsk/iigrp2/nfsvol-shadow -C iigrp2
```

Remote Mirror Example

When you enable this remote mirror volume set where host1 is a logical failover host name:

```
# sndradm -e host1 /dev/vx/rdsk/sndrdg/datavol /dev/vx/rdsk/sndrdg/datavolbm1 \
host2 /dev/rdsk/c1t3d0s0 /dev/rdsk/c1t2d0s4 ip sync
```

The corresponding configuration information, as shown by the sndradm -i command, is:

```
# sndradm -i

host1 /dev/vx/rdsk/sndrdg/datavol /dev/vx/rdsk/sndrdg/datavolbm1 \
host2 /dev/rdsk/c1t3d0s0 /dev/rdsk/c1t2d0s4 ip sync \
C sndrdg
```

The C portion of the entry shows a disk device group name sndrdg.

Point-in-Time Copy Example

When you enable a point-in-time copy volume set on a cluster node (logical failover host):

```
# iiadm -e ind /dev/vx/rdsk/iidg/c1t3d0s0 /dev/vx/rdsk/iidg/c1t3d0s4 \
/dev/vx/rdsk/iidg/c1t2d0s5
```

the corresponding configuration as shown by iiadm -i command is:

```
# iiadm -i

/dev/vx/rdsk/iidg/c1t3s0d0: (master volume)
/dev/vx/rdsk/iidg/c1t3d0s4: (shadow volume)
/dev/vx/rdsk/iidg/c1t2d0s5: (bitmap volume)
Cluster tag: iidg
Independent copy
Volume size:    208278
Percent of bitmap set: 0
```

The Cluster tag entry shows the derived disk device group name iidg.

Local Device Command Syntax

Note – Enabling a local disk device group named `local` prevents you from configuring a cluster disk device group named `local`.

- When you enable a Point-in-Time Copy volume set, use the `-C local` option to specify that the volume set's disk device group name is `local`.

```
iiadm -C local -e {dep |ind} master shadow bitmap
```

- When you enable a remote mirror volume set, use the `C local` option as part of the *vol-set* volume set definition.

```
sndradm -e vol-set
```

where *vol-set* is:

```
phost pdev pbitmap shost sdev sbitmap ip {sync | async} [g io-groupname][C local]
```

The `local` disk device group is local to the individual cluster node and is not defined in a cluster disk or resource group. Local devices do not fail over and switch back. This initial configuration is similar to using the Sun StorageTek Availability Suite software in a non-clustered environment.

When you enable a volume set with the `local` disk device group, its configuration entry includes the name of its host machine.

Caution – Volumes and bitmaps used in a local remote mirror volume set cannot reside in a shared disk device group or metaset.

Point-in-Time Copy Example

When you enable this Point-in-Time Copy volume set where `local` indicates a disk device group:

```
# iiadm -C local -e ind /dev/rdsk/c1t90d0s5 /dev/rdsk/c1t90d0s6 \
/dev/rdsk/c1t90d0s7
```

The corresponding configuration as shown by `iiadm -i` command is:

```
# iiadm -i
/dev/rdsk/iidg/c1t90d0s5: (master volume)
/dev/rdsk/iidg/c1t90d0s6: (shadow volume)
/dev/rdsk/iidg/c1t90d0s7: (bitmap volume)
Cluster tag: (local)
Independent copy
Volume size:    208278
Percent of bitmap set: 0
```

where *localhost* is the local host name as returned by the `hostname(1)` command.

The corresponding configuration information as shown by the `dscfg -l` command is:

```
# dscfg -l | grep /dev/rdsk/c1t3d0s0
ii: /dev/rdsk/c1t90d0s5 /dev/rdsk/c1t90d0s6 /dev/rdsk/c1t90d0s7 I - - - -
```

Which Host To Issue Remote Mirror Commands From?

The Sun StorageTek Availability Suite software requires that you issue the `iiadm` or `sndradm` commands from the node that is the current primary host for the disk device group that the command applies to.

In a clustered environment, you can issue the command from the node mastering the disk device group you specified in Step 2 in "To configure Sun Cluster for HAStorage or HAStoragePlus" on page 26.

When you enable the Remote Mirror software for the first time, issue the `sndradm` enable command from the primary and secondary hosts. See TABLE 3-1.

TABLE 3-1   Which Host to Issue Remote Mirror Commands From

Task	Where Command Is Issued	Comments
Assign a new bitmap to a volume set.	Primary and secondary host	Perform this command first on the host where the new bitmap resides and is being assigned, and then perform it on the other host.
Disable the Remote Mirror software.	Primary or secondary host	You can disable on one host, leave the other host enabled, and then re-enable the disabled host.
		Perform this operation on both hosts if you are deleting a volume set.
Enable the Remote Mirror software.	Primary and secondary host	When enabling the Remote Mirror software for the first time, issue the command from both hosts.
Full forward or reverse synchronization (copy).	Primary host	Ensure that both hosts are enabled.
Forward or reverse synchronization (update).	Primary host	Ensure that both hosts are enabled.

TABLE 3-1   Which Host to Issue Remote Mirror Commands From *(Continued)*

Task	Where Command Is Issued	Comments
Log.	Primary host	Perform on the primary host only if a synchronization is in progress.
		Perform on the secondary host if the primary host failed.
	Primary or secondary host	Perform on either host if no synchronization is in progress.
Toggle the autosynchronization state.	Primary host	
Update an I/O group.	Primary and secondary hosts	

Putting All Cluster Volume Sets in an I/O Group

Note – Placing volume sets in an I/O group does not affect the cluster operations of all volume sets configured in disk device and resource groups.

Caution – Do not reverse synchronize the primary volume from more than one secondary volume or host at a time. You can group one-to-many sets that share a common primary volume into a single I/O group to forward synchronize all sets simultaneously instead of issuing a separate command for each set.

You cannot use the same technique to reverse synchronize volume sets, however. In this case, you must issue a separate command for each set and reverse update the primary volume by using a specific secondary volume.

The Remote Mirror and Point-in-Time Copy software enables you to assign volume sets to I/O groups. Instead of issuing one command for each volume set, you can:

- Assign specific volume sets to an I/O group
- Issue one command specifying the I/O group

- Perform operations on those volume sets only

Like the -C *tag* and C *tag* options, the I/O group name excludes all other enabled volume sets from operations you specify.

In a clustered environment, you can assign some or all volume sets in a specific disk device group to an I/O group when you enable each volume set.

▼ To Place Volume Sets in an I/O Group

1. **Enable three Point-in-Time Copy volume sets and place them in an I/O group named** cluster1.

```
# iiadm -g cluster1 -e ind /dev/rdsk/iigrp2/c1t3d0s0 \
/dev/rdsk/iigrp2/c1t3d0s4 /dev/rdsk/iigrp2/c1t2d0s5

# iiadm -g cluster1 -e dep /dev/rdsk/iigrp2/c1t4d0s0 \
/dev/rdsk/iigrp2/c1t4d0s4 /dev/rdsk/iigrp2/c1t3d0s5

# iiadm -g cluster1 -e ind /dev/rdsk/iigrp2/c1t5d0s0 \
/dev/rdsk/iigrp2/c1t5d0s4 /dev/rdsk/iigrp2/c1t4d0s5
```

2. **Wait for any disk write operations to complete before issuing another command.**

```
# iiadm -g cluster1 -w
```

3. **Allow your applications to write to the master volumes.**

4. **Update the shadow volumes.**

```
# iiadm -g cluster1 -u s
```

Preserving Point-in-Time Copy Volume Data

If a Solaris operating environment system failure or Sun Cluster failover occurs during a point-in-time copy or update operation to the master volume, specifically where the shadow volume is copying (`iiadm -c m`) or updating (`iiadm -u m`) data to the master volume, the master volume might be in an inconsistent state (that is, the copy or update operation might be incomplete).

To avoid or reduce the risk of inconsistent data if a system failover occurs during such a copy or update operation, perform the following before performing the shadow volume-to-master volume copy or update operation:

1. Create a second independent shadow volume copy of the master volume by issuing an `iiadm -e ind` command.

   This operation results in a full shadow volume copy of the master volume data.

2. Ensure that all copy or update operations to this second shadow volume are finished by issuing a wait command (`iiadm -w` *shadowvol*) after issuing the `iiadm -e ind` command.

You can now perform the copy or update operation from the original shadow volume to the master volume. If a system failure or failover occurs during this operation, you at least have a known good copy of your original master volume data. When this operation is complete, you can keep the second shadow volume under point-in-time copy control or return it to your storage pool.

Exporting, Importing, and Joining Shadows in a Sun Cluster OE

Overview

The Point-in-Time Copy software provides the functionality to allow for an independent shadow volume on a multi-ported storage device to be exported to a secondary node within a Sun Cluster Operating Environment (OE), while still under Point-in-Time Copy software control. This capability of Exporting, Importing, and Joining of a Point-in-Time Copy shadow set enables shadow volume processing by associated applications to be off-loaded to a secondary node within a Sun Cluster OE, without affecting the primary node's master volume or its associated applications. Since the Point-in-Time Copy software retains control of the shadow volume while it is imported on a secondary node within a Sun Cluster OE, point-in-time consistency is retained to facilitate fast resynchronization via update processing at a later time.

Prior versions of the Point-in-Time Copy software did not support the use of E/I/J functionality in a Sun Cluster OE. With the current version of the Availability Suite product, Point-in-Time Copy supports the Export, Import, Join processing of Point-in-Time Copy shadow volumes for hosts running Sun Cluster 3.1 (10/3) OE.

This section describes the proper use, configuration and control for the Exporting, Importing, and Joining of shadow volumes feature of the Sun StorageTek Availability Suite Point-In-Time Copy software in a Sun Cluster 3.1 (10/3) OE.

The Master and Bitmap volumes of a Point-in-Time Copy set can reside on the primary Sun Cluster node, while the shadow volume and optional bitmap(2) volume are exported to the secondary Sun Cluster node. Once on the secondary node, the shadow volume and bitmap(2) volume can be imported and used for off-host read/write data processing, while not impacting the performance of the primary

node, the master volume, or the Point-in-time Copy set. Once secondary node processing is complete, the shadow volume and bitmap(2) volume can be moved back to the primary node and joined back with the master volume, restoring the Point-in-Time Copy set and Point-in-Time copy state, as though the shadow volume had never be exported.

Requirements

For the Export, Import, and Join functionality to work correctly, the shadow volume must be on a different global device or volume manager controlled device group than its associated master and bitmap volumes. This allows the shadow volume's device group to be switched between various nodes in a Sun Cluster and to be used as an exportable shadow volume.

Export, Import, and Join Functionality

The Export, Import, and Join functionality of the Point-in-Time Copy software allows for a previously configured shadow volume contained on a dual-ported or Sun StorageTek SAN Foundation (SAN) accessible storage device to be exported from a configured Point-in-Time Copy volume set. Within the Sun Cluster OE, this shadow volume can be contained on a Sun Cluster global device or either of the two supported Sun Cluster volume managers, Solaris Volume Manager (SVM) or VERITAS Volume Manager (VxVM).

Note – Sun Cluster DID devices are not supported as a master, shadow, or bitmap volume, due to disk data fencing functionality when Sun Cluster failure events are active. Sun Cluster Global devices that are symmetrical in name to the DID devices are supported.

When a Point-in-Time Copy volume set is initially configured such that the master and bitmap volumes are in one disk group and the shadow volume is in another disk group, the Point-in-Time Copy Export functionality allows for an independent shadow volume (once the Point-in-Time Copy becomes fully independent) to be exported from the Point-in-Time Copy volume set.

Once exported from the Point-in-Time Copy volume set, this shadow volume can be accessed in a read-only manner on any Sun Cluster node without impacting the master volume. If the shadow volume needs to be accessed in a read-write manner, then the Point-in-Time Copy Import feature can be utilized to provide both read and write access to the shadow volume while retaining point-in-time consistency.

Once secondary shadow volume processing is no longer required on the secondary Sun Cluster node, the shadow volume is disabled, if the volume was being used in the imported state. The shadow volume is then switched back to the Sun Cluster node now containing the original Point-in-Time Copy volume set's master and bitmap volumes. Utilizing the Point-in-Time Copy Join command, the shadow volume and the secondary bitmap volume can be used to reassociate the shadow volume back with its original master and bitmap volumes, thus restoring the Point-in-Time Copy volume set. Upon completion of these operations, the Point-in-Time Copy volume set exists as though the shadow was never exported in the first place.

The entire Export, Import, and Join functionality enables retention of Point-in-Time Copy information across the entire process of moving the shadow volume from one Sun Cluster node to another Sun Cluster node and back again. Through the incorporation of a secondary bitmap volume on the secondary Sun Cluster node, in combination with the Import feature, writes occurring on the secondary node are tracked. The secondary bitmap volume tracking information is then reflected back into the original Point-in-Time Copy set, when a Join operation is used to associate the shadow volume back into its original Point-in-Time Copy set. While the shadow volume was exported from the Point-in-time Copy set, write operations occurring to the master volume are still being tracked in the bitmap volume on the primary Sun Cluster node. The Join operation merges all write operations occurring to both the master and shadow volumes, retaining a consistent Point-in-Time Copy set.

In summary, changes to the Availability Suite's Point-in-time Copy software, along with new Sun Cluster configuration guidelines for configuring an exportable Point-in-Time Copy volume set, provide both high availability (HA) to the Point-in-Time Copy volume set and the ability to have the Point-in-Time Copy volume set retain these HA characteristics while Export, Import, and Join processing is being utilized.

Point-in-Time Copy Set in a Sun Cluster OE

The master volume of an independent Point-in-Time Copy set can be located on Sun Cluster controlled devices as either a raw global device (for example, `/dev/global/rdsk/d4s0`), SVM (for example, `/dev/md/AVsuite/rdsk/d5`), or VxVM (for example, `/dev/vx/rdsk/AVsuite/m1`) controlled volume. The shadow volume of this set can be on the same type or a different Sun Cluster controlled device type, as long as it is in its own device group. Furthermore, when the master and shadow volumes are on different Sun Cluster device groups, the Export, Import, and Join functionality and the Sun Cluster device group and resource group functionality allow the shadow volume of a Point-in-Time Copy set to be relocated on different nodes of a Sun Cluster OE.

While exported from the Point-in-Time Copy set, the shadow volume is disabled from the highly available resource group in which the Point-in-Time Copy set is configured. Once the shadow volume is no longer needed as an exported shadow volume, it can be joined with the Point-in-Time Copy set and enabled under the set's highly available resource group.

If an exportable shadow volume is currently in the imported state on a Sun Cluster node, a new feature of the Point-in-Time Copy software is an automatic implicit Join operation. If, during a Sun Cluster voluntary or involuntary failover event, the node where the imported shadow volume is currently enabled is picked, the software will detect this fact and automatically rejoin the imported shadow volume back into the Point-in-Time Copy set. The behavior retains the High Availability of the Sun Cluster configured resource group, while retaining control of the shadow volume's data.

Point-in-Time Copy Sets

The restriction that all constituent volumes of a Point-in-Time Copy set must be in the same device group is not enforced for the shadow volume.

Since an exportable shadow volume must also be an independent shadow volume, the exportable shadow volume must be the same size (or larger) than the master volume it is associated with. If the exportable shadow volume is to be used in a read/write mode on another node in the Sun Cluster, it is advisable that the master and bitmap volumes be configured in one Sun Cluster device group and that the shadow volume and secondary bitmap volume be configured in a different Sun Cluster device group. Also ensure that the secondary bitmap volume is the same size (or larger) than the original bitmap volume.

When configuring Availability Suite volumes on Sun Cluster global devices (`/dev/global/rdsk/d`*n*`s`*n*), the device group associated with each global device is the `dsk/d`*n* part. Therefore when configuring a Point-in-Time Copy Volume, the master and associated bitmap volume must be on one global device and the shadow volume and secondary bitmap should be on another.

Due to the "global nature" of Sun Cluster global devices, it is advisable that the master and shadow volumes be on different global devices, so that off-host processing of the shadow volume avoids initiating I/Os over the Sun Cluster private interconnect. Once the shadow volume is exported from the Point-in-Time Copy set, it can be switched over to the Sun Cluster node where off-host processing will occur. In doing so, those I/Os to the shadow volume will not be impacting the Sun Cluster system as it pertains to the private interconnect.

Prior to configuring a Point-in-Time Copy set with an exportable shadow volume in a Sun Cluster OE, one must make sure that the device groups of both the master and bitmap volumes and the shadow volumes are both highly available. Failure to do so, will prevent the single Point-in-Time Copy set from being highly available.

Configuring a Point-in-Time Copy Set in a Sun Cluster OE

The steps to create a highly available Point-in-Time Copy volume set are listed in the following section.

There is no convention required for the RGM resource groups or resource types, although a planned and well-thought-out naming scheme, spanning the volume manager if one is used and the Sun Cluster resource groups and types, will be beneficial later if troubleshooting is required.

The setup creates a Point-in-Time Volume set on Sun Cluster nodes, node-A and node-B, with the exportable shadow volume available on node-C.

▼ To Configure a Point-in-Time Copy Set in a Sun Cluster OE

Note – This example uses two SVM device groups: "oracle" and "backup", where "oracle" is the master volume's device group and "backup" is the exportable shadow volume's device group.

This example is based on configuring the following Point-in-Time Copy set:

```
# iiadm –ne ind /dev/md/oracle/rdsk/d1 /dev/md/backup/rdsk/d1 \
/dev/md/oracle/rdsk/d2
```

Do not invoke the above, or similar, iiadm command at this time in the following sequence of steps. The device attributes of the constituent volumes of an enabled Point-in-Time Copy set are such that without Sun Cluster Resource Group Manager (RGM) control, the association of a master and bitmap volume, allowing an exportable shadow volume, makes all of the associated device groups no longer highly available. This fact could impact the high availability of the Sun Cluster as it pertains to these associated volumes.

5. **Create a resource group that will contain the HAStoragePlus resource type associated with the Point-in-Time Copy set.**

   This resource group should specify two or more nodes within the Sun Cluster, or be left blank if all nodes in the Sun Cluster are capable of supporting the Point-in-Time Copy set as a highly available resource.

   ```
   # scrgadm -a -g Availability_Suite_RG -h node-A,node-B[,node-C,...]
   ```

   Or, for all nodes in the Sun Cluster, as long as the Availability Suite software has been installed and configured on each:

   ```
   # scrgadm -a -g Availability_Suite_RG
   ```

   Note – Additional resource types for other Sun Cluster HA data services or applications may be added to this same resource group at your discretion.

   ---

6. **Ensure that the SUNW.HAStoragePlus resource type is registered. If not register it.**

   ```
   # scrgadm -p | grep "Res Type name:" | grep HAStoragePlus
   # scrgadm -a -t SUNW.HAStoragePlus
   ```

7. **Add an HAStoragePlus resource type to the previously created resource group.**

   The HAStoragePlus resource type will be used to specify two Sun Cluster device groups, one representing the master-bitmap volume pair, and one for the exportable shadow volume. The ordering of these device group is important; the exportable shadow volume must be the last one specified.

> **Note –** The `HAStoragePlus` resource type allows for its `GlobalDevicePaths` parameter to be either the full device path specification of a Sun Cluster device or the name of a Sun Cluster device group. The first format is used in this example, so that no doubt exists as to which devices are being associated. In doing so, the example will be setting a GlobalDevicePath with both a master and bitmap volume, each of which is in the same device group. As such, one of the device path specifications is redundant, and will be ignored.

```
# scrgadm -a -g Availability_Suite_RG -j \
Availability_Suite_RES -t SUNW.HAStoragePlus -x \
GlobalDevicePaths=/dev/md/oracle/rdsk/d1,\
/dev/md/oracle/rdsk/d2,/dev/md/backup/rdsk/d1 -x \
AffinityOn=False
```

The Sun Cluster resource type `SUNW.HAStoragePlus`, supports a configuration option `AffinityOn`, with the default value being `True`. This setting, along with the fact that the `GlobalDevicePaths` qualifier contains two device groups, with one of the device groups being the exportable shadow volume, implies that if the exportable shadow volume is currently in use on a secondary Sun Cluster node, this resource group will have a strong affinity to move to the node, regardless of the node list specified in step 2 above.

For example, in a three (or more) node Sun Cluster configuration, where two nodes have the system resources to support an HA enterprise class application like ORACLE® and the third Sun Cluster node is a low-end backup system, if the exportable shadow volume is in use on this third system, the HA application will move to the third Sun Cluster node even if there are not system resources capable of supporting its execution. This is the justification for setting `AffinityOn=False` in the example above.

8. **Bring the resource group online. Then, verify that the resource group is located on the Sun Cluster node where the Point-in-Time Copy enable command will be invoked.**

```
# scswitch -Z -g Availability_Suite_RG
# scswitch -z -g Availability_Suite_RG -h node-A
```

9. **Enable the Point-in-Time Copy set using the new option** `-n` **to enable exportable shadows.**

```
# iiadm -ne ind /dev/md/oracle/rdsk/d1 /dev/md/backup/rdsk/d1 \
/dev/md/oracle/rdsk/d2
```

10. **Validate that the Point-in-Time Copy set is available on this node.**

```
# iiadm -i /dev/md/backup/rdsk/d1
# scstat -g
# scstat -D
```

11. **Switch the resource group from this node to each of the other configured nodes, and validate the set.**

```
# scswitch -z -g Availability_Suite_RG -h node-B
# telnet node-B
<login to root account>
# iiadm -i /dev/md/backup/rdsk/d1
# scstat -g
# scstat -D
# ^D {logout}
```

12. **This Point-in-Time Copy volume set is now highly available and usable as a resource group to which other highly available applications (HA-NFS, HA-ORACLE, and so forth) can now be added.**

    For example:

```
# scrgadm -a -g Availability_Suite_RG -j nfs_res -t SUNW.nfs
# scswitch -e -j nfs
```

    To use the Point-in-Time shadow volume on another node within the Sun Cluster, it must be exported from its associated set and disabled as a device path within its HAStoragePlus resource type.

13. **Confirm that the II set is fully independent.**

    Prior to being exported from a Point-in-Time Copy set, the II set must be fully independent. This is confirmed by wait (iiadm -w), returning.

```
# iiadm -w /dev/md/backup/rdsk/d1
```

14. **Export the II shadow volume from its associated Point-in-Time Copy set.**

```
# iiadm -E /dev/md/backup/rdsk/d1
# iiadm -i /dev/md/backup/rdsk/d1
```

15. **The Point-in-Time exportable shadow volume can be switched to another node in the Sun Cluster.**

```
# scswitch -z -D backup -h node-C
```

Or the Point-in-Time Copy set can be switched to another node in the Sun Cluster.

```
# scswitch -z -g Availability_Suite_RG -h node-C
```

16. **Validate the correct behavior.**

```
# telnet node-C
<login to root account>
# iiadm -i /dev/md/backup/rdsk/d1
# scstat -g
# scstat -D
```

The Point-in-Time Shadow volume is now accessible independently from the Point-in-Time Copy set, off-host, yet the original Point-in-Time set is still active on the other Sun Cluster node.

If the shadow volume will be accessed in a read/write mode, a secondary bitmap should be used to import the shadow locally on this node, so that subsequent fast-resynchronization operations (`iiadm -u`) can be done, versus the full-synchronization (`iiadm -c`).

Note – The exportable shadow MUST be enabled with the `-C local` tag, so that the system can differentiate between the highly available Point-in-Time Copy set and the locally accessible exportable shadow, each of which have the exact same name.

```
# iiadm -C local -I /dev/md/backup/rdsk/d1 /dev/md/backup/rdsk/d2
# iiadm -i /dev/md/backup/rdsk/d1
```

Note – From this node, you will see the imported shadow volume and the shadow volume's Point-in-Time Volume set as suspended on this node and active on node-A (or B).

---

17. **While this imported shadow volume is active on this node, a test of steps #5 and #6 should be carried out to validate that the original Point-in-Time Copy set is still highly available.**

    Remember that the original Point-in-Time Copy set is not configured to be highly available on node-C, since an attempt to do so will fail as long as the shadow volume is imported on this node.

18. **When you are done using the imported shadow volume on this node (if it was decided to import it), disable the locally accessible imported shadow volume, switch it back to the node where the Point-in-Time Volume set is active, and enable the resource in the resource group.**

    ```
    # iiadm -C local -d /dev/md/backup/rdsk/d1
    ```

19. **Take the resource offline and back online, forcing the exportable shadow volume back to the Sun Cluster node where the rest of the Point-in-Time Copy set is enabled.**

    ```
    # scswitch -n -j Availability_Suite_RES
    # scswitch -e -j Availability_Suite_RES
    ```

    Note – From this node, we will still see the shadow volume's Point-in-Time volume set as suspended on this node, active on node-A (or B). The imported shadow volume is no longer listed.

    ```
    # iiadm -i /dev/md/backup/rdsk/d1
    # ^D {logout, back to node-A }
    ```

20. **Join the shadow volume (with possible changes) back with the original Point-in-Time Copy set.**

    ```
    # iiadm -J /dev/md/backup/rdsk/d1 /dev/md/backup/rdsk/d2
    ```

    The Point-in-Time Copy set is back in its original state, as though the shadow volume has never been exported.

Point-in-Time Copy Set Considerations in a Sun Cluster OE

Redundancy

To provide high availability to the data contained on a Point-in-Time Copy set when using global devices, it is assumed that the master, shadow, and bitmap volumes are on redundant storage since there is no means today to provide for host-based data service redundancy on raw global devices. If controller-based redundancy is not available, then a Sun Cluster supported volume manager must be used. For performance reasons, it is recommended under all scenarios that bitmap volumes are NOT placed on RAID-5 volumes (either host-based or controller-based) since the nature of bitmap I/O processing within a Point-in-Time Copy volume set can be I/O intensive.

Implicit Join Operation

The implicit join operation is a new feature of the Availability Suite software when using the Export, Import, and Join functionality in a Sun Cluster OE. By design, a Point-in-Time Copy set with its shadow volume exported cannot coexist on the same node as the Point-in-Time Copy set containing the imported shadow volume.

From an operational point-of-view, having the master volume and shadow volume on two different Sun Cluster nodes is the sole reason for using Export, Import, and Join functionality in a Sun Cluster OE. If you wanted both master and shadow volumes on the same node, you would have avoided using Export, Import, and Join in the first place.

If a Sun Cluster voluntary or involuntary failover event moves the master and associated bitmap volume to the Sun Cluster node containing the imported shadow volume, the design issue mentioned earlier would prevent the failover from completing successfully. To address this concern, the Availability Suite software detects this failover condition and performs an implicit join operation, merging the two Point-in-Time Copy sets back into one. This operation should have no impact on the master or shadow volume other then the fact that both volumes are now in the same Point-in-Time Copy set on the same node in the Sun Cluster.

Incomplete Export, Import, and Join Sequence

The operational procedures for Export, Import, and Join are based on the assumption that all three steps will be performed. Due to circumstances in system availability outside of Availability Suite, after exporting the shadow volume from the set it may not be possible to perform the import step immediately, yet it is desirable to join the exported shadow volume back into the Point-in-Time Copy set. In other words, it may be desirable at times to perform an Export, and Join sequence with no import step.

To perform a join operation there is still the requirement for a secondary bitmap volume, but since the secondary bitmap volume was NOT used during a recent import operation, it contains stale data or uninitialized data. Prior to performing a join operation, it is required to copy (using the Solaris `dd` utility) the current bitmap volume over the contents of the secondary bitmap volume so that the secondary bitmap volume's data is in a known state. The failure to perform this manual initialization step may result in the join operation failing, or, when state data is used, it may cause an inconsistency between what is actually on the shadow volume and the current state as recorded in the bitmap.

Glossary

Cluster Interconnect The hardware networking infrastructure that includes cables, cluster transport junctions, and cluster transport adapters. The Sun Cluster and data service software use this infrastructure for intra-cluster communication.

Cluster Member An active member of the current cluster incarnation. This member is capable of sharing resources with other cluster members and providing services both to other cluster members and to clients of the cluster.

Co-Location The property of being on the same node. Specific to Remote Mirror software, for a configured replica, the logical hostname and the associated device group are collocated on the same node, along with any other failover data services or highly available applications. For Point-in-Time Copy software, co-location refers to the device group and other software failover data services or highly available applications.

Data service Highly Available (HA) applications within the Sun Cluster environment are also known as data services. The term *data service* is used to describe a third-party application that has been configured to run on a cluster rather than on a single server. A data service includes the application software and Sun Cluster software that starts, stops, and monitors the application.

Device group A user-defined group of device resources, such as disks, that can be mastered from different nodes in a cluster HA configuration. This group can include device resources of disks, Solaris Volume Manager disksets, VERITAS Volume Manager disk groups, Remote Mirror volumes or Point-in-Time Copy sets, or any combination thereof.

DID name Used to identify global devices in a Sun Cluster system. It is a clustering identifier with a one-to-one or a one-to-many relationship with Solaris logical names. It takes the form dXsY, where X is an integer and Y is the slice name. See also .

Exportable Shadow A Point-in-Time Copy Set's shadow volume in a Sun Cluster OE, that is in a different device group than the set's master and bitmap volumes. The ability to create a set with this configuration is available only when using the `iiadm -n` option, when initially creating the Point-in-Time Volume set.

Failover	The automatic relocation of a resource group or a device group from a current primary node to a new primary node after a failure has occurred.
Failover resource	A resource, each of whose resources can correctly be mastered by only one node at a time.
Global device	A namespace that contains the logical, cluster-wide names for global devices. Local devices in the Solaris environment are defined in the `/dev/dsk/c?t?d?s?` and `/dev/rdsk/c?t?d?s?` device directories. For each local device accessible on any cluster node, there is a derived disk device name based on the unique DID name of the underlying physical device. For each unique DID named deviced `/dev/did/dsk/d<n>` and `/dev/did/rdsk/d<n>`, there is a corresponding `/dev/global/dsk/d<n>` and `/dev/global/rdsk/d<n>` name.
Multihost disk	A disk that is physically connected to multiple nodes.
Network resource	A resource that contains one or more logical hostnames or shared addresses.
Point-in-Time Copy	Using the Availability Suite command `iiadm`, one configures a master, shadow and bitmap volume as a Point-in-Time Copy set. A Point-in-Time Copy can be further configured, such that it is a highly available resource in a Sun Cluster OE.
Potential primary	A cluster member that is able to master a failover resource type if the primary node fails.
Primary	A node on which a resource group or device group is currently online. That is, a primary is a node that is currently hosting or implementing the service associated with the resource. See also Secondary,
Primary and secondary hosts and nodes	In this guide and the Remote Mirror software documentation, the terms *primary host* and *secondary host* are used as follows.
	The primary and secondary hosts are physically-separate servers running the Remote Mirror software. The primary host contains the primary volume and bitmap volume to be initially replicated to a remote server called a secondary host. The secondary hosts contains the secondary volume and bitmap volume
	The terms *primary node* and *secondary node* refer to the Sun Cluster logical hostname that represents the current node that is also master the device group associated with the Remote Mirror set.
Resource	An instance of a resource type. Many resources of the same type might exist, each resource having its own name and set of property values, so that many instances of the underlying application might run on the cluster.
Resource group	A collection of resources that are managed by the RGM as a unit. Each resource that is to be managed by the RGM must be configured in a resource group. Typically, related and interdependent resources are grouped.

Resource type	The unique name given to a data service or LogicalHostname cluster object. Data service resource types can either be failover types or scalable types, although the Availability Suite software is only a failover data service.
Rolling upgrade	In a Sun Cluster configuration, an upgrade that is performed sequentially on one cluster node at a time. During a rolling upgrade, the cluster remains in production and services continue to run on the other nodes.
Secondary	A cluster member that is available to master disk device groups and resource groups in the event that the primary fails. See also Primary.
Solaris logical name	The names typically used to manage Solaris devices. For disks, these usually look something like /dev/rdsk/c0t2d0s2. For each one of these Solaris logical device names, there is an underlying Solaris physical device name. See also "DID name" on page 55.
Solaris Volume Manager	A software product that provides data reliability through disk striping, concatenation, mirroring, and dynamic growth of metadevices or volumes.
Switchover	The orderly transfer of a resource group or device group from one master (node) in a cluster to another master (or multiple masters, if resource groups are configured for multiple primaries). A switchover is initiated by an administrator by using the scswitch(1M) command.
VERITAS Volume Manager (VxVM)	A software product that provides data reliability through disk striping, concatenation, mirroring, and dynamic growth of metadevices or volumes.

Index