

TruCluster Server

Release Notes

Part Number: AA-RHH1B-TE

April 2000

Product Version: TruCluster Server Version 5.0A

Operating System and Version: Tru64 UNIX Version 5.0A

This manual provides important information about TruCluster Server Version 5.0A.

© 2000 Compaq Computer Corporation

COMPAQ and the Compaq logo Registered in U.S. Patent and Trademark Office. TruCluster and Tru64 are trademarks of Compaq Information Technologies Group, L.P.

Microsoft and Windows are trademarks of Microsoft Corporation. UNIX and The Open Group are trademarks of The Open Group. All other product names mentioned herein may be trademarks or registered trademarks of their respective companies.

Confidential computer software. Valid license from Compaq required for possession, use, or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Compaq shall not be liable for technical or editorial errors or omissions contained herein. The information in this publication is subject to change without notice and is provided "as is" without warranty of any kind. The entire risk arising out of the use of this information remains with recipient. In no event shall Compaq be liable for any direct, consequential, incidental, special, punitive, or other damages whatsoever (including without limitation, damages for loss of business profits, business interruption or loss of business information), even if Compaq has been advised of the possibility of such damages. The foregoing shall apply regardless of the negligence or other fault of either party and regardless of whether such liability sounds in contract, negligence, tort, or any other theory of legal liability, and notwithstanding any failure of essential purpose of any limited remedy.

The limited warranties for Compaq products are exclusively set forth in the documentation accompanying such products. Nothing herein should be construed as constituting a further or additional warranty.

Contents

About This Manual

1 Features

2 Restrictions

2.1	CFS Restrictions	2-1
2.2	LSM Restrictions	2-2
2.3	Restricted Support for Tape Devices on a Shared Bus	2-3

3 Known Problems

3.1	Do Not Use the Installation Branch of the Software Option of the SysMan Menu	3-1
3.2	Cluster Creation Notes	3-1
3.2.1	Using HSZ80 and HSG80 Devices as Member Boot Disks	3-1
3.2.2	Do Not Configure NIS Master or Slave Server Before Creating the Cluster	3-2
3.2.3	The clu_create Command Forgets to Add First Member's Fully Qualified Hostname to the /etc/cfgmgr.auth File	3-2
3.2.4	The clu_create.log File Contains "malformed CDSL" Messages	3-2
3.3	Member Addition Notes	3-2
3.3.1	What to Do when the Kernel Build of a Newly Added Member Fails	3-3
3.4	Bootting Notes	3-3
3.4.1	Bootting a New Member Without a Cluster License Displays ATTENTION Message	3-3
3.4.2	Bootting All Cluster Members Before Starting Applications that Use the Memory Channel API Library	3-4
3.5	Member Deletion Notes	3-4
3.5.1	Deleting a Member with a Bad Member Boot Disk	3-4
3.5.2	The clu_delete_member Command Displays Confusing Messages	3-5
3.6	File System Notes	3-6

3.6.1	Cluster Member Shutdown Unmounts File Systems that Are Not Busy	3-6
3.6.2	CFS Relocation Failures Involving Applications that Wire Memory	3-6
3.6.3	The cfsmgr Command Dumps Core when the -p or -m Option Is Used with the -F raw Option	3-6
3.6.4	Wrong errno Value Returned on Opens of mmapped Files for Direct I/O	3-7
3.7	LSM Notes	3-7
3.7.1	Use Only Shared Disks with LSM	3-7
3.7.2	Failed LSM Disk May Prevent Members from Performing LSM Commands	3-8
3.7.3	The voldisk list Command Does Not Display All Non-LSM Disks In a Cluster	3-8
3.7.4	LSM Commands Hang or Fail in a Cluster	3-9
3.8	CAA Notes	3-9
3.8.1	CAA Does Not Detect a Restored Path to a Tape Device ...	3-9
3.8.2	The caa_register -u Command Does Not Correctly Update a Network Resource's State	3-10
3.8.3	Default Media Changer Resource Profile Template Contains Incorrect Resource Type	3-10
3.8.4	CAA Events Are Malformed when Viewed from the Event Manager (EVM) Viewer	3-10
3.8.5	CAA Application Resources May Time Out During Quorum Loss	3-11
3.8.6	CAA Does Not Recognize Fully Qualified Host Names	3-11
3.8.7	SysMan Menu Does Not Handle Balanced Placement of Application Resources Correctly	3-11
3.8.8	CAA Stops Required Resources When a Resource Requiring Them Fails to Start	3-12
3.8.9	Certain CAA Resource Profile Options Cannot Be Updated	3-13
3.8.10	CAA Lets You Update a Running Application's Required Resources	3-13
3.8.11	CAA Lets You Register or Update Application Resources With Nonexistent or Null Action Script	3-13
3.8.12	SysMan Station Shows CAA Application Resources in UNKNOWN State As Having an Error	3-14
3.8.13	The caa_relocate -s Command Does Not Relocate All Application Resources from the Specified Member	3-14
3.8.14	CAA Displays "Cannot Communicate with the CAA daemon" Message During Member Boot	3-15

3.8.15	caa_stat -t -v Output May Be Corrupted by Long Resource Names	3-16
3.8.16	Default CAA Action Scripts Can Generate Too Many Log Files	3-16
3.9	Miscellaneous Administration Notes	3-16
3.9.1	The -cluster Option of hwmgr -scan scsi Command Does Not Work Correctly	3-16
3.9.2	Adding a Disk to a Running Cluster	3-16
3.9.3	Tape Device Becomes Inaccessible to the Cluster If Client System Goes Down Before Open Completes	3-17
3.9.4	Do Not Use the rcinet restart Option in a Cluster	3-17
3.9.5	The clu_get_info Command Returns Failure Messages	3-18
3.9.6	The sysman -clone Command Succeeds Inappropriately in a Cluster	3-18
3.9.7	Running Process Accounting on Large Clusters Can Exhaust a Member Process Quotas	3-18
3.9.8	Clusterwide Shutdown Causes gated to Dump Core and Generate Many Task Terminate Messages	3-18
3.9.9	Characters May Be Missing or Jumbled in Terminal Window Displays	3-18
3.9.10	The addvol and rmvol Commands Do Not Automatically Update Member-Specific etc/clu_bdmgr.conf Files	3-19
3.9.11	Spurious Message from clu_upgrade undo install Command	3-19
3.10	SysMan Menu Notes	3-19
3.10.1	Do Not Use SysMan Menu to Configure Remote Member Network Interface Cards	3-19
3.10.2	Clusterwide nfsconfig Deconfigure Operation May Stop Daemons on Members with Overrides	3-20
3.10.3	Nonroot Users Cannot Run System Management Applications	3-20
3.11	SysMan Station Notes	3-21
3.11.1	SysMan Station Daemon Might Dump Core when Multiple Members Are Booted Simultaneously	3-21
3.11.2	SysMan Station Might Display Cluster Status Incorrectly	3-22
3.11.3	SysMan Station May Display New Hardware Objects Incorrectly	3-23
3.11.4	Properties Might Not Display for Selected Objects	3-23
3.11.5	Some SysMan Station Applications Display Wrong Target Member Name	3-23
3.11.6	Documentation Notes	3-24

Index

Tables

1-1	TruCluster Server Version 5.0A Features	1-1
-----	---	-----

About This Manual

This manual provides important information about TruCluster™ Server Version 5.0A, which runs on the Tru64™ UNIX® Version 5.0A operating system.

Audience

TruCluster Server users, administrators, and programmers should read this manual.

Organization

This manual contains the following chapters:

- *Chapter 1* lists the special features of TruCluster Server Version 5.0A.
- *Chapter 2* lists restrictions on the use of certain components in TruCluster Server Version 5.0A.
- *Chapter 3* discusses known problems with TruCluster Server Version 5.0A.

Related Documents

Consult the following manuals for assistance in cluster hardware configuration, installation, administration, and programming tasks:

- TruCluster Server *Software Product Description* (SPD) — The comprehensive description of the TruCluster Server Version 5.0A product. You can find the latest version of the SPD and other TruCluster Server documentation at the following URL:

http://www.unix.digital.com/faqs/publications/pub_page/cluster_list.html

- TruCluster Server *Technical Overview* — Provides an overview of the TruCluster Server technology.
- TruCluster Server *Hardware Configuration* — Describes how to set up the processors that are to become cluster members, and how to configure cluster shared storage.
- TruCluster Server *Software Installation* — Describes how to install the TruCluster Server software.

- TruCluster Server *Highly Available Applications* — Describes how to deploy existing applications on a TruCluster Server cluster, and how to write cluster-aware applications.
- TruCluster Server *Cluster Administration* — Describes cluster-specific administration tasks.

It is well worth your time to read the Tru64 UNIX *Release Notes* and the Tru64 UNIX *System Administration* manual to become familiar with restrictions and new features in the base operating system before installing, configuring, and using your TruCluster Server cluster.

Reader's Comments

Compaq welcomes any comments and suggestions you have on this and other Tru64 UNIX manuals.

You can send your comments in the following ways:

- Fax: 603-884-0120 Attn: UBPG Publications, ZKO3-3/Y32
- Internet electronic mail: `readers_comment@zk3.dec.com`

A Reader's Comment form is located on your system in the following location:

```
/usr/doc/readers_comment.txt
```

- Mail:

Compaq Computer Corporation
 UBPG Publications Manager
 ZKO3-3/Y32
 110 Spit Brook Road
 Nashua, NH 03062-2698

A Reader's Comment form is located in the back of each printed manual. The form is postage paid if you mail it in the United States.

Please include the following information along with your comments:

- The full title of the book and the order number. (The order number is printed on the title page of this book and on its back cover.)
- The section numbers and page numbers of the information on which you are commenting.
- The version of Tru64 UNIX that you are using.
- If known, the type of processor that is running the Tru64 UNIX software.

The Tru64 UNIX Publications group cannot respond to system problems or technical support inquiries. Please address technical questions to your local system vendor or to the appropriate Compaq technical support office.

Information provided with the software media explains how to send problem reports to Compaq.

Conventions

The following typographical conventions are used in this manual:

<code>%</code>	
<code>\$</code>	A percent sign represents the C shell system prompt. A dollar sign represents the system prompt for the Bourne, Korn, and POSIX shells.
<code>#</code>	A number sign represents the superuser prompt.
<code>% cat</code>	Boldface type in interactive examples indicates typed user input.
<code><i>file</i></code>	Italic (slanted) type indicates variable values, placeholders, and function argument names.
<code>[]</code> <code>{ }</code>	In syntax definitions, brackets indicate items that are optional and braces indicate items that are required. Vertical bars separating items inside brackets or braces indicate that you choose one item from among those listed.
<code>...</code>	In syntax definitions, a horizontal ellipsis indicates that the preceding item can be repeated one or more times.
<code>cat(1)</code>	A cross-reference to a reference page includes the appropriate section number in parentheses. For example, <code>cat(1)</code> indicates that you can find information on the <code>cat</code> command in Section 1 of the reference pages.
<code>Return</code>	In an example, a key name enclosed in a box indicates that you press that key.
<code>Ctrl/x</code>	This symbol indicates that you hold down the first named key while pressing the key or mouse button that follows the slash. In examples, this

key combination is enclosed in a box (for example, Ctrl/C).

1

Features

TruCluster Server Version 5.0A provides the features listed in Table 1–1.

Table 1–1: TruCluster Server Version 5.0A Features

Feature	Description	For more information, see:
Clusterwide namespace	The cluster file system (CFS) supports a single clusterwide namespace and uniform, coherent access to all file systems in a cluster. Context-dependent symbolic links (CDSLs) are used to maintain per-system configuration and data files within the shared CFS root (/), /usr, and /var file systems.	<i>Technical Overview, Cluster Administration</i>
Clusterwide access to disk and tape storage	The device request dispatcher facility provides highly available clusterwide access to both character and block disk devices, as well as tape devices. All cluster disk and tape I/O passes through the device request dispatcher.	<i>Technical Overview, Cluster Administration</i>
Logical Storage Manager (LSM)	The semantics of LSM have been extended to a cluster environment.	<i>Technical Overview, Cluster Administration</i>
Connection manager	The connection manager ensures that all cluster members communicate with each other in order to control the formation and continued operation of a cluster. The connection manager calculates the votes required for quorum and decides when members are added to and removed from the cluster.	<i>Technical Overview, Cluster Administration</i>
Cluster application availability (CAA)	The CAA facility provides resource monitoring and application restart capabilities. It provides the same type of application availability provided by user-defined services in the TruCluster Available Server Software and TruCluster Production Server Software products.	<i>Technical Overview, Cluster Administration, Highly Available Applications</i>

Table 1–1: TruCluster Server Version 5.0A Features (cont.)

Feature	Description	For more information, see:
Cluster alias	The cluster alias subsystem lets TCP and UDP applications address the cluster as though it were a single system. When the cluster is created, a default alias is defined that addresses all cluster members. A site can define additional aliases that address some or all cluster members.	<i>Technical Overview, Cluster Administration, Highly Available Applications</i>
Highly available NFS server using cluster alias	As shipped, the cluster is a highly available NFS server. CFS ensures that file systems exported from a TruCluster Server cluster are highly available to clients. Clients use the default cluster alias as the name of the NFS server when mounting file systems exported by the cluster.	<i>Cluster Administration</i>
Single-system management	Because a cluster uses CFS, all systems' configuration files are available for management. The SysMan suite of graphical management utilities provides an integrated view of the cluster environment, letting you manage a single member or the entire cluster.	<i>Technical Overview, Cluster Administration</i>
Single security domain	Because a cluster uses CFS, there is a single copy of security administration files, such as <code>/etc/passwd</code> and <code>/etc/group</code> . A user authenticated on one member has access to all members. A user with access to a file on one member has access to that file from any member. Access control lists (ACLs) are uniformly available to all members.	<i>Cluster Administration</i>
Memory Channel interconnect	The Memory Channel interconnect is a high-speed interconnect designed specifically for the needs of clusters. The Memory Channel interconnect provides both broadcast and point-to-point connections between cluster members. TruCluster Server provides a Memory Channel application programming interface (API) library, which is the same as that provided in the TruCluster Production Server Software product.	<i>Hardware Configuration, Technical Overview, Highly Available Applications</i>

Table 1–1: TruCluster Server Version 5.0A Features (cont.)

Feature	Description	For more information, see:
Distributed lock manager (DLM)	TruCluster Server supports the DLM and its API, which is the same as that provided in the TruCluster Production Server Software product.	<i>Technical Overview, Highly Available Applications</i>
Rolling upgrade and rolling patch	TruCluster Server Version 5.0A contains the software infrastructure required to support rolling upgrades and patches. Customers who install TruCluster Server Version 5.0A will be able to perform a rolling upgrade to the next TruCluster Server release, and roll patches onto a Version 5.0A cluster.	<code>clu_upgrade(8)</code> reference page

Restrictions

This chapter lists restrictions on the use of TruCluster Server Version 5.0A features.

2.1 CFS Restrictions

The cluster file system (CFS) has the following restrictions in TruCluster Server Version 5.0A:

- CFS supports the Network File System (NFS) client for read/write access.

When a file system is NFS mounted in a cluster, CFS makes it available for read/write access from all cluster members. The member that has actually mounted it serves the file system to other cluster members.

Should the member that has mounted the NFS file system be shut down or fail, the file system is automatically unmounted and CFS begins to clean up the mount points. During the cleanup process, members accessing these mount points may see various types of behavior, depending upon how far the cleanup has progressed:

- If members still have files open on that file system, their writes will be sent to a local cache instead of the actual NFS mounted file system.
- After all of the files on that file system have been closed, attempts to open a file on that file system will fail with an `EIO` error until the file system is remounted. Applications may encounter "Stale NFS handle" messages. This is normal behavior on a standalone system, as well as in a cluster.

Until the CFS cleanup is complete, members may still be able to create new files at the NFS file system's local mount point (or in any directories created locally beneath that mount point).

An NFS file system does not automatically fail over to another cluster member. Rather, you must manually remount it — on the same mount point or another — from another cluster member to make it available again. Alternatively, booting a cluster member will remount those file systems listed in the `/etc/fstab` file not currently mounted and served in the cluster. (If you are using automount, the remount will happen automatically.)

- CFS supports the UNIX File System (UFS) for read-only access.

- CFS supports the CD-ROM File System (CDFS) for read-only access. Because TruCluster Server Version 5.0A does not support CD-ROM devices on a shared bus, a CD-ROM device becomes inaccessible to the cluster when the member to which it is locally connected fails, even if it is being served by another member.

The CDFS management utilities are not cluster-aware and must be run on the member that is acting as the CFS server for the CDFS filesystem to be managed.

- CFS does not support the Memory File System (MFS).
- CFS does not support a clusterwide `/proc` file system. A `/proc` file system can be mounted and accessed only on the local member.
- CFS does not support a clusterwide File-on-File Mount (FFM) file system. An FFM file system can be mounted and accessed only on the local member.
- CFS does not support clusterwide named pipes. The reader and writer of a named pipe must reside on the same member.
- User and group file system quotas are not supported in a cluster. The `quota` and `vquota` series of commands return a failure. AdvFS fileset quotas are supported.

2.2 LSM Restrictions

The following restrictions apply to use of the Logical Storage Manager (LSM) in a cluster:

- The cluster root domain cannot be an LSM volume. Because LSM mirroring is not available for cluster root, use hardware RAID to mirror the root file system.
- Because there is no LSM support for the clusterwide root domain, the `volrootmir` and `volunroot` commands, as well as the `volencap` command on the root domain, are not supported for clusters.
- You cannot add an LSM volume to the `cluster_root` domain. An attempt to use the `addvol` command in this manner fails.
- LSM cannot be used to mirror member boot partitions.
- A quorum disk cannot be used with LSM. Therefore, LSM volumes cannot reside on the same disk as a quorum disk, and a quorum disk cannot be mirrored using LSM.
- LSM volumes cannot be used for either primary or secondary swap devices in a cluster. Because LSM mirroring is not available for swap devices, use hardware RAID to protect swap space.
- LSM RAID 5 volumes cannot be used in clusters.

2.3 Restricted Support for Tape Devices on a Shared Bus

In TruCluster Server Version 5.0A, you can access a tape device in the cluster from any member, regardless of whether it is located on that member's private bus, a shared bus, or another member's private bus.

However, if the tape device is located on the shared bus, applications that access it must be written to react appropriately to certain events on the shared SCSI bus, such as bus and device resets. Bus and device resets (such as those that result from cluster membership transitions) cause any tape device on the shared SCSI bus to rewind. Therefore, a tape server application will inspect the `errno` value and extended error information returned from its I/O call and reposition the tape. The commonly used utilities `tar`, `cpio`, `dump`, and `vdump` are not designed in this way, so they may unexpectedly terminate when used on a tape device that resides on a shared bus in a cluster. Currently, the only advantage to situating a tape device on a shared bus in this release is that multiple systems are physically connected to it, and any one of those systems can access it.

Known Problems

The following sections describe known problems with TruCluster Server Version 5.0A.

3.1 Do Not Use the Installation Branch of the Software Option of the SysMan Menu

The Installation branch of the Software menu of the SysMan Menu application is not supported in a cluster. Use the mechanisms for installing and deinstalling the TruCluster Server product and layered product software discussed in the TruCluster Server *Software Installation* and *Cluster Administration* manuals.

3.2 Cluster Creation Notes

The following notes apply to cluster creation.

3.2.1 Using HSZ80 and HSG80 Devices as Member Boot Disks

Alpha System Reference Manual (SRM) console firmware Version 5.7 or later must be installed on any cluster member that boots from a disk behind an HSZ80 or HSG80 controller. If the cluster member is using Version 5.6 firmware (the version on the latest Alpha Systems Firmware Update CD-ROM) or earlier, the member may fail to boot, indicating "Reservation Conflict" errors.

Alpha SRM console firmware Version 5.7 can be obtained from the following Web location:

<http://ftp.digital.com/pub/DEC/Alpha/firmware>

Note

Boot support for Fibre Channel disk devices on nonclustered Tru64 UNIX Version 5.0A systems is fully supported with Alpha SRM console firmware Version 5.6.

3.2.2 Do Not Configure NIS Master or Slave Server Before Creating the Cluster

If the cluster will be a Network Information Service (NIS) master or slave, the TruCluster Server *Software Installation* and *Cluster Administration* manuals instruct you to configure NIS before creating the cluster. If you do so, however, the `clu_create` command will not correctly reconfigure NIS to respond to the cluster alias rather than to the host name of a single cluster member.

If you do mistakenly configure NIS before creating the cluster, the easiest workaround is to remove the NIS configuration and reconfigure it.

3.2.3 The `clu_create` Command Forgets to Add First Member's Fully Qualified Hostname to the `/etc/cfgmgr.auth` File

The `clu_create` command fails to add the first member's fully qualified hostname to the `/etc/cfgmgr.auth` file. The `clu_add_member` command, however, does add subsequent members' hostnames to the file.

To avoid problems with remote kernel configuration management in a cluster, manually add the first member's fully qualified hostname to the `/etc/cfgmgr.auth` file. For example:

```
member1.zk3.dec.com
member2.zk3.dec.com
member3.zk3.dec.com
```

3.2.4 The `clu_create.log` File Contains "malformed CDSL" Messages

The `/cluster/admin/clu_create.log` file may contain messages of the following type:

```
clu_create.log:Fixing malformed CDSL
clu_create.log:Keeping listed malformed CDSL
```

These messages should cause you no concern. They do not indicate a problem with the cluster and its context-dependent symbolic links (CDSLs).

3.3 Member Addition Notes

The following note applies to adding new members to the cluster.

3.3.1 What to Do when the Kernel Build of a Newly Added Member Fails

When you first boot a newly added member, the new kernel for the member is built. If the kernel build fails, you are asked if you would like to build a kernel without the optional kernel options.

It is important that you answer *no* to this question. A kernel built without the optional kernel options would be incompatible with the existing cluster and will not boot.

If the kernel build fails for a newly added member, take steps to determine why the build failed and, after diagnosing and fixing the problem, use the `doconfig` command to rebuild the kernel.

3.4 Booting Notes

The following notes discuss requirements and restrictions for booting members into a cluster.

3.4.1 Booting a New Member Without a Cluster License Displays ATTENTION Message

When you boot a newly added member, the `clu_check_config` utility performs a series of configuration checks. If you have not yet installed the TruCluster Server license, the TCS-UA product authorization key (PAK), on the member, the boot procedure will display:

```
Starting Cluster Configuration Check...
The boottime cluster check found a potential problem.
For details search for !!!!!ATTENTION!!!! in /cluster/admin/clu_check_log_<hostname>
check_cdsl_config : Boot Mode : Running /usr/sbin/cdslinrchk in the background
check_cdsl_config : Results can be found in : /var/adm/cdsl_check_list
clu_check_config : no configuration errors or warnings were detected
```

When you inspect the `/cluster/admin/clu_check_log_<hostname>` file, you will see the message:

```
/usr/sbin/caad is NOT_RUNNING !!!!!ATTENTION!!!!
```

When the TruCluster Server license is not configured on a member, the cluster application availability daemon (`caad`) is not automatically started on that member. This is normal and expected behavior.

If you did not configure the license from within `clu_add_member` when you added the new member (as discussed in the TruCluster Server *Software Installation* manual), you can configure it later using the `lmf register` command. After the license has been installed, you can start the CAA daemon on that member using the `usr/sbin/caad` command.

3.4.2 Booting All Cluster Members Before Starting Applications that Use the Memory Channel API Library

If there is insufficient Memory Channel address space in a cluster, a booting node may have problems joining the cluster. When this is the case, one or more members may panic with an assertion failure (ICS MCT Assertion Failed), or the booting member may hang early in its boot.

Memory Channel resources are dynamically allocated as new members join the cluster. Running applications that call the Memory Channel application programming interface (API) library functions can consume required Memory Channel resources, and prevent a member from getting the resources it needs to join the cluster. To avoid this problem, boot all cluster members before starting any applications that call the Memory Channel API library functions.

3.5 Member Deletion Notes

The following sections discuss problems with the `clu_delete_member` command.

3.5.1 Deleting a Member with a Bad Member Boot Disk

The `clu_delete_member -f` command can delete a member with a bad boot disk. However, there are two issues of which you should be aware:

1. When you delete a member with a bad boot disk, `clu_delete_member -f` does not adjust expected votes in the running cluster (as a normal `clu_delete_member` command would do). If the deleted member was a voting member, use the `clu_quorum -e` command after the member has been deleted to adjust expected votes appropriately.
2. The `clu_delete_member -f` command cannot delete a member if the disk label on its boot disk has been overwritten or corrupted.

To work around this problem, you must manually edit the `label:` field on the boot disk so that it contains a string of the form `clu_membern`, where `n` is the cluster member's member ID. You can then enter the `clu_delete_member` command to delete the member from the cluster.

For example:

```
% disklabel -e dsk6
# /dev/rdisk/dsk6c:
type: SCSI
disk: RZ28
label: clu_member5
flags:
bytes/sector: 512
sectors/track: 99
tracks/cylinder: 16
sectors/cylinder: 1584
```

```

cylinders: 2595
sectors/unit: 4110480
rpm: 5400
interleave: 1
trackskew: 6
cylinderskew: 30
headswitch: 0          # milliseconds
track-to-track seek: 0 # milliseconds
drivedata: 0

8 partitions:
#          size      offset   fstype   [fsize bsize  cpg]       # NOTE:
values n
ot exact
write new label? [y]:
re-edit the label? [n]:

```

3.5.2 The `clu_delete_member` Command Displays Confusing Messages

The `clu_delete_member` command requires that the member you intend to delete be halted at the time you enter the command. If you specify an active member to the `clu_delete_member -m <memberid>` command, the command should display the following message:

```

Error: Cluster member ID <memberid> is UP.
Members with a state of UP cannot be deleted

```

Instead, it displays the message:

```

Error: clu_delete_member: Bad configuration

```

This message is a cryptic way of telling you to halt the member you intend to delete.

Additionally, as `clu_delete_member` proceeds to delete a member, it displays rather self-reflexive NTP peer messages and a few extraneous mail messages, such as the following:

```

Initial cluster deletion successful, member '3' can no longer join the
cluster. Deletion continuing with cleanup.
expected_votes: reconfigured
  Removing cluster interconnect interface 'pepicelli-mc0' from /.rhosts
  Configuring Network Time Protocol for new member
  Deleting interface 'pepperoni.zk3.dec.com' as an NTP peer
    to member 'pepperoni.zk3.dec.com'
  Deleting interface 'polishham.zk3.dec.com' as an NTP peer
    to member 'polishham.zk3.dec.com'
Clusterizing mail...
Mail has not been configured for the cluster.
Please run /usr/sbin/mailconfig or /usr/sbin/mailsetup
to configure mail in the cluster.

```

You can safely ignore these messages.

3.6 File System Notes

The following sections discuss issues with the CFS, AdvFS, and NFS file systems in a cluster.

3.6.1 Cluster Member Shutdown Unmounts File Systems that Are Not Busy

When a cluster member shuts down (using the `shutdown` command), the `/sbin/rc0` run-level transition script will unmount all file systems in the cluster that are not in active use. To avoid this problem for all file systems other than NFS file systems, locate the following line in the `/sbin/rc0` script:

```
/sbin/umount -Af 2> /dev/null
```

Change that line to read:

```
/sbin/umount -Af -t nfs 2> /dev/null
```

Even after this change, NFS file systems that are not in active use will still be unmounted when a cluster member shuts down.

3.6.2 CFS Relocation Failures Involving Applications that Wire Memory

Applications that use the `plock()` or `mlock()` system call to lock pages of physical memory can cause the `cfsmgr` command to fail when performing a manual relocation.

If the application uses `plock()`, the domain or file system containing the application executable cannot relocate. In the case of `mlock()`, if the locked pages are associated with files, then the file systems where those files reside cannot relocate.

In the event of failure, the `cfsmgr` command returns the following message:

```
Server Relocation Failed  
Failure Reason: Invalid Relocation
```

Killing the processes (if you are able to identify them) that are running the executables using the `plock()` and `mlock()` system calls allows the relocation to complete for the domain or file system on which the executables reside.

3.6.3 The `cfsmgr` Command Dumps Core when the `-p` or `-m` Option Is Used with the `-F raw` Option

Explicitly specifying the `-p` or `-m` option on the command line with the `-F raw` option causes the `cfsmgr` command to dump core. The `-F raw` option

selects raw-formatted output when the `cfsmgr` command is invoked to output CFS data. Scripts can be written to process the raw CFS data.

The parameters used with the `-p` and `-m` options, and with the `-d` option, specify (in different ways) the file system for which information is to be output. When none of these options is specified, the `cfsmgr` command takes the name of the directory on which the file system is mounted. You can precede that directory name with a `-p` option or not. Specifying the `-p` option with the directory name when using the `-F raw` option results in a core dump. We recommend omitting the `-p` option.

The `-m` option to the `cfsmgr` command takes the mounted-from location of the file system. Using the `-m` option with `-F raw` results in a core dump. There is no workaround.

3.6.4 Wrong errno Value Returned on Opens of mmapped Files for Direct I/O

When an application opens a memory-mapped file in a cluster for direct-I/O (that is, using the `open()` or `creat()` system call with the `O_DIRECTIO` flag), the call should fail with an `errno` value of `EINVAL`. Instead, it can return a value of `EINVAL` (if the file is opened on its CFS server) or `ENOTSUP` (if it is opened on a client of its CFS server).

3.7 LSM Notes

The following notes discuss known problems with using the Logical Storage Manager (LSM) in a TruCluster Server cluster.

3.7.1 Use Only Shared Disks with LSM

Do not put disks under LSM control if they are not directly connected to all cluster members. A problem prevents the LSM configuration daemons (`vold` daemons) on booting members from synchronizing and agreeing on the current LSM configuration. The `vold` daemons exit, and a message indicating that `vold` is not enabled for transactions is displayed in the `Starting LSM` phase of the member boot procedure. When the member has finished booting, any LSM commands run from it will fail with the following message:

```
lsm:voldisk: ERROR: IPC failure: Configuration daemon is not accessible
```

In addition, a member failure will prevent LSM access to the disk and possibly cause a volume to fail.

You may be able to perform LSM configuration changes from other members.

3.7.2 Failed LSM Disk May Prevent Members from Performing LSM Commands

If the disks on a member's private bus are configured under LSM in a running cluster and that member fails or is shut down, subsequent activity before that member reboots may cause LSM to mark that member's private disks as failed. As a result, when the member with the private disks does reboot and rejoin the cluster, it cannot synchronize with and participate in LSM-related operations.

For example, consider a three-member cluster. Member C has a private disk, `dsk10`. All members are up and all disks are available. Members B and C go down (not necessarily together, but at some point they are both down). Member B reboots while member C is still down. When the LSM configuration daemon (`vold`) on member B starts, it determines that `dsk10` is unavailable and fails that disk. When member C reboots and tries to synchronize with the cluster, `dsk10` is already marked as failed. As a result, LSM does not add the disk back into the LSM configuration.

If the cluster did detect and process the failed LSM disk, you might need to recover or remove the failed disk before that member can participate in LSM operations. To replace and recover a failed LSM disk, see the Tru64 UNIX *Logical Storage Manager* manual. To remove a failed LSM disk, either move the LSM objects currently configured on the failed disk to another LSM disk or remove the failed objects as follows:

- To move LSM objects to another LSM disk, use the `volevac` or `volassist mv` command.
- To remove LSM objects, disassociate any failed plexes from the disk and use the `voledit rm`, `volplex -o`, `voldg rmdisk`, or `voldisk rm` command, as appropriate, to remove them.

For example, to remove the detached/failed plex XYZ on the failed disk, enter:

```
# volplex -o rm dis XYZ
```

To remove a failed disk (`dsk23`) from `rootdg`, enter:

```
# voldg rmdisk dsk23
```

To remove disk `dsk23` from LSM, enter:

```
# voldisk rm dsk23
```

3.7.3 The `voldisk list` Command Does Not Display All Non-LSM Disks In a Cluster

The `voldisk list` command is supposed to show disks configured both under LSM and not under LSM, giving disks not configured under LSM a

status of unknown. In a cluster, however, the `voldisk list` command displays only those non-LSM disks that are directly connected to the member on which the command was executed.

For disks configured with LSM, `voldisk list` will display consistent data from all members.

3.7.4 LSM Commands Hang or Fail in a Cluster

The LSM configuration daemon (`vold`) may fail and dump core on a cluster member under certain scenarios involving failed disks or simultaneous LSM disk operations (such as a `volsetup -s` command run while other LSM commands are active). Additionally, the `vold` may encounter a checksum error at boot time and fail to start on one or more cluster members. When the `vold` daemon fails, LSM commands may appear to hang on one or more members.

If you suspect you have encountered this problem, follow these steps:

1. Enter the following command on each member to verify that the `vold` daemon is still running:

```
# ps -ax | grep vold
```

2. Check for a `vold` core dump by entering the following command:

```
# file /core
```

If a core dump exists, save it for later diagnosis by Customer Support Services.

3. Enter the following command to restart the `vold` daemon on any member it is not running on:

```
# vold -k
```

Usually, when you restart the `vold` daemon, LSM commands will resume.

3.8 CAA Notes

The following notes discuss known problems in the cluster application availability (CAA) subsystem.

3.8.1 CAA Does Not Detect a Restored Path to a Tape Device

Because underlying hardware does not detect the restoration of a lost connection between a cluster member and a tape device, the CAA tape monitor does not receive an event that would cause it to put the tape resource `ONLINE`. To update the state of a tape device after it has been

reconnected to the cluster, execute the following commands on the member to which the connection has been restored:

```
% hwmgr -scan scsi
% drdmgr -a CHECK_PATH <tape-device-name>
```

The state displayed in `caa_stat` may not reflect the true state because it is not updated immediately after a device is connected.

Additionally, when an application resource with a tape device defined as a required resource attempts to start, CAA checks the SCSI bus to see if the device is truly `ONLINE` for that member. If this check fails, the state of all tape devices is scanned and updated.

3.8.2 The `caa_register -u` Command Does Not Correctly Update a Network Resource's State

When you change the `SUBNET` value in a network resource's profile and run the `caa_register -u` command to update the CAA registry, the command will not update the `STATE` value for the resource. To correctly update a network resource when you have changed its `SUBNET` attribute, follow these steps:

1. Unregister the resource using the `caa_unregister` command.
2. Change the `SUBNET` attribute in the resource profile.
3. Reregister the resource using the `caa_register` command.

3.8.3 Default Media Changer Resource Profile Template Contains Incorrect Resource Type

When you create a media changer resource profile using the `caa_profile -create` command or the SysMan Menu, CAA incorrectly sets the `TYPE` attribute of the resource to `tape` instead of `changer`. To work around this problem, edit the `/var/cluster/caa/template/TYPE_changer.cap` file before creating any changer resources. Locate the following line:

```
-t |TYPE=tape
```

Change it to:

```
-t |TYPE=changer
```

3.8.4 CAA Events Are Malformed when Viewed from the Event Manager (EVM) Viewer

The Event Manager (EVM) viewer may display malformed CAA event messages, or messages with missing information. For example, the message:

```
CAA named is transitioning from state ONLINE to state OFFLINE on skiing
```

is displayed as:

```
CAA named is transitioning from state to state skiing
```

To work around this problem, examine the messages in the `daemon.log` file for more complete information. The messages in the `daemon log` file are in a slightly different format from those displayed by the EVM viewer.

3.8.5 CAA Application Resources May Time Out During Quorum Loss

While a cluster has lost quorum, the action scripts of online CAA application resources may time out. When the cluster regains quorum, the resources may fail over to other members, depending upon the contents of their profiles and action scripts. This should not be a problem for applications that define placement policies and dependencies in their profiles. For applications without a defined placement policy that are sensitive to the members on which they run, use the `caa_stat` and `caa_relocate` commands to restore application placement to its state before the quorum loss.

3.8.6 CAA Does Not Recognize Fully Qualified Host Names

Do not specify a fully qualified host name in the `-c` option of any CAA command. The CAA command will fail in the following manner:

```
% caa_relocate test -c pepperoni.zk3.dec.com
Trying to relocate to a dead member ...
Resource test has placement error.
```

To work around this problem, specify the cluster member host name without the domain name (for example, `pepperoni`).

3.8.7 SysMan Menu Does Not Handle Balanced Placement of Application Resources Correctly

The balanced placement policy does not always work correctly for application resources that are started from the SysMan Menu. If you start an application resource with a balanced policy on a specific member, all subsequent applications you start during that SysMan Menu session are started on that same member unless you explicitly start those applications on other members.

The `caa_start` command always starts application resources with a balanced placement policy properly.

3.8.8 CAA Stops Required Resources When a Resource Requiring Them Fails to Start

When CAA starts or relocates an application resource, it automatically starts or relocates those application resources on which the application depends. However, if the application's start script fails, CAA not only does not start the application's required resources, it also stops any of its required resources that are running at the time of the failure. Moreover, CAA will stop any other application resource that shares a mutual dependency on any of those required resources.

For example, applications A and B depend on application C. Application A fails to start on member 1; CAA tries to relocate it to member 2, then member 3, and so on. At the same time, CAA tries to relocate applications B and C along with with application A. If application A's start script continues to fail on each member, application A will eventually fail completely and go OFFLINE. Applications B and C, and any applications that depend on them, will also be set to OFFLINE, even though those applications do not depend on application A.

To avoid this problem, make sure that application resource profiles contain a valid ACTION_SCRIPT entry for an existing script and that the script has been well tested before starting the application resource in a production environment. Upon success, the script should return a status of 0 (zero).

If you encounter this problem, follow these steps:

1. Identify the application resource whose failure to start or automatically fail over caused the other resources to be stopped. Refer to the CAA messages in the event manager (EVM) viewer or the `daemon.log` for assistance.
2. Enter a `caa_stop` command for the application resource to set its target state to OFFLINE. For example:

```
# caa_stop A
```

3. Restart the required resources and any others that may also have been stopped. For example:

```
# caa_start B  
# caa_start C
```

4. Fix the start script for the problematic application resource and retest the script to ensure that it works correctly.
5. Restart the application resource. For example:

```
# caa_start A
```

3.8.9 Certain CAA Resource Profile Options Cannot Be Updated

When you attempt to update the following profile attributes, using either the `caa_profile -update` command or the CAA GUI, CAA fails to update them:

- `FAILURE_THRESHOLD`
- `FAILURE_INTERVAL`
- `CHECK_INTERVAL`
- `RESTART_ATTEMPTS`
- `FAILOVER_DELAY`
- `AUTO_START`
- `SCRIPT_TIMEOUT`
- `ACTIVE_PLACEMENT`

To update these attributes, use a text editor on the application profile and then execute the `caa_register -u` command.

3.8.10 CAA Lets You Update a Running Application's Required Resources

The `caa_register -u` command and the CAA GUI allow you to successfully update the `REQUIRED_RESOURCES` field in the profile of an `ONLINE` resource with the name of resource that is `OFFLINE`. (It should not allow you to update this field unless the required resource is `ONLINE` or the resource whose profile you are modifying is `OFFLINE`.)

If you accidentally update the `REQUIRED_RESOURCES` field incorrectly, you must manually start the required resource or stop the updated resource to correct their states.

3.8.11 CAA Lets You Register or Update Application Resources With Nonexistent or Null Action Script

The `caa_profile` command and the CAA GUI erroneously validate an application resource profile that specifies a nonexistent or null action script. Moreover, when you register such an application resource, using either the `caa_register` command or the CAA GUI, registration will succeed.

When you attempt to start an application resource without a valid action script, CAA returns the following message:

```
# caa_start test
Attempting to start 'test' on member 'polishham'
'test' on member 'polishham' has experienced an unrecoverable failure.
Human intervention required to resume its availability.
Could not start resource test.
```

When you receive this message, you should inspect the application resource's profile and correct the `ACTION_SCRIPT` entry, and update its registration (using either the `caa_register -update` command or the CAA GUI).

In addition, CAA lets you update the `ACTION_SCRIPT` field in the profile of an `ONLINE` application resource with a nonexistent or null action script. In this case, CAA will become unable to monitor and manage the resource. You may need to enter the `caa_stop -f` command (or perform the equivalent action within the CAA GUI) to forcibly stop the application (from CAA's point of view) and explicitly kill the application. Only after you have ensured that the application is stopped, can you proceed to correct the profile, update its registration, and restart the application resource under CAA.

3.8.12 SysMan Station Shows CAA Application Resources in UNKNOWN State As Having an Error

The SysMan Station shows any application resources that are in the `UNKNOWN` state as having an error and does not show what member the application is `UNKNOWN` on. For example, an application resource named `foo` in the `UNKNOWN` state is shown underneath the cluster icon as `foo (error)`.

3.8.13 The `caa_relocate -s` Command Does Not Relocate All Application Resources from the Specified Member

The `caa_relocate -s` command should relocate all application resources from a member but does not. To relocate all application resources from a member, you must explicitly specify the name of each resource in the `caa_relocate` command line. For example, to relocate all resources from `member1`, enter the following commands:

```
# caa_stat -c member1
NAME=clock
TYPE=application
TARGET=ONLINE
STATE=ONLINE on member1

NAME=dhcp
TYPE=application
TARGET=ONLINE
STATE=ONLINE on member1

NAME=cluster_lockd
TYPE=application
```

```
TARGET=ONLINE
STATE=ONLINE on member2
# caa_relocate clock dhcp
```

3.8.14 CAA Displays "Cannot Communicate with the CAA daemon" Message During Member Boot

CAA may display the following message during a member boot:

```
Cannot communicate with the CAA daemon.
```

This message may or may not be preceded by the message:

```
Error: could not start up CAA Applications
Cannot communicate with the CAA daemon.
```

These messages indicate any of the following situations:

- You neglected to register the TruCluster Server license.
When the member finishes booting, enter the following command:

```
# lmf list for TCS-UA
```


If the TCS-UA license is not present, register it as instructed in the TruCluster Server *Cluster Administration* manual and start the CAA daemon (caad) as follows:

```
#/usr/sbin/caad
```
- The CAA daemon was started successfully and the message is spurious.
When the member finishes booting, enter the following command:

```
# caa_stat
```


If the command runs successfully, no further action is required. If CAA complains again that it cannot communicate with its daemon, restart the daemon by entering the following command:

```
# /usr/sbin/caad
```
- The CAA daemon was not be able to complete its initialization before the boot procedure began to start CAA application resources.
When the member finishes booting, follow these steps:
 1. Enter the following command to check whether the daemon is running:

```
% ps ax | grep caad
```
 2. If the daemon is running, no further action is required. If the daemon is not running, start it using the following command:

```
% /usr/sbin/caad -1
```
 3. If the command returns an error, repeat steps 1 and 2 until the daemon starts without error messages.

3.8.15 `caa_stat -t -v` Output May Be Corrupted by Long Resource Names

The `caa_stat -t -v` command may display corrupted information for resources that have very long names. To work around this problem, use the `caa_stat -v` command to display the output for these resources in a nontabular form.

3.8.16 Default CAA Action Scripts Can Generate Too Many Log Files

You can create a default action script for an application by using either the `caa_profile -B` command or the SysMan Menu. Each time the `start`, `stop`, or `check` entry point of a default action script entry point runs, it creates a new log file in the `/var/cluster/caa/log/` directory. This can result in an excessive number of log files in that directory and, over time, lead to disk space and performance problems.

To work around this problem, you can clean up unneeded log files by a `crontab` entry similar to the following:

```
# Remove log files older than three days.
0 1 * * * /usr/bin/find /var/cluster/caa/log -mtime +3 -exec /bin/rm {}
\;
```

This entry will run once a day at 1:00 AM and remove any log files that were modified more than three days ago.

3.9 Miscellaneous Administration Notes

The following notes discuss issues with various administration tools used in a cluster.

3.9.1 The `-cluster` Option of `hwmgr -scan scsi` Command Does Not Work Correctly

When entered from a single cluster member, the `hwmgr -scan scsi -cluster` command does not properly update the device databases of all cluster members. You would usually use this command when adding a new disk to a cluster (see Section 3.9.2).

To work around this problem, enter the following command on *each* cluster member when you need to update the member device databases clusterwide:

```
# hwmgr -scan scsi
```

3.9.2 Adding a Disk to a Running Cluster

When you add a new disk to a running cluster (for example, when replacing a failed disk), the cluster may not properly identify or configure the disk.

To ensure that all cluster members properly recognize a new disk, follow these steps:

1. If the disk you are adding is a RZ26, RZ28, RZ29, or RZ1CB-CA model, enter the following command on each cluster member after installing the disk:

```
# /usr/sbin/clu_disk_install
```

Note

This command may take several minutes to complete if the cluster has a large number of storage devices.

2. For all disk models, enter the following command on each member to scan SCSI buses clusterwide and configure any new devices.

```
# hwmgr -scan scsi
```

Allow a minute or two for the scans to complete.

3.9.3 Tape Device Becomes Inaccessible to the Cluster If Client System Goes Down Before Open Completes

If a cluster member goes down while it is in the process of opening a tape device served by another cluster member, the tape device may become inaccessible from the remaining cluster members. When run on the member serving the tape device, the `drdmgr` command shows the device to be present with a local connection, but any application that tries to open the device from a cluster member fails with a "device busy" error.

To work around this problem, relocate the server of the tape device to another member or, if that is not possible, reboot the server.

3.9.4 Do Not Use the `rcinet restart` Option in a Cluster

Do not use the `rcinet restart` command on a cluster member. Because the command is not cluster-aware, it will not restart the cluster aliases and will cause the `gated` daemon to be restarted with an incorrect configuration file. As a result, cluster members will lose some connectivity.

If you do use the `rcinet restart` command, you can undo these problems by entering the following command:

```
# /sbin/init.d/clu_alias restart
```

3.9.5 The `clu_get_info` Command Returns Failure Messages

The `clu_get_info` command fails to return certain information if the `/etc/hosts` file, or a member's `/etc/rc.config` or `/etc/sysconfigtab` file, is configured incorrectly; for example, if the `/etc/hosts` file is missing information, such as any cluster member's Memory Channel interface, or contains duplicate IP names and addresses. When it discovers a misconfigured configuration file, the `clu_get_info` command returns a message, such as `Bad Cluster interconnect IP name or address`.

If you see this message, examine the system configuration files for errors.

3.9.6 The `sysman -clone` Command Succeeds Inappropriately in a Cluster

The `sysman -clone` command is not supported in a cluster environment. However, if you do not have the `/usr/sbin` directory in your path, the `sysman -clone` command will not properly detect that it is running in a cluster and will mistakenly create a configuration description file (CDF). To avoid this problem, make sure `/usr/sbin` is in your path.

3.9.7 Running Process Accounting on Large Clusters Can Exhaust a Member Process Quotas

If process accounting is enabled on large clusters (six to eight members), cluster members may start swapping heavily and eventually exhaust their process quotas. A `ps` command on such a member will show tens of thousands of `icssvr_daemon_from_pool` processes.

If you see this situation developing in a cluster that is running process accounting, use the `accton` command to disable accounting.

3.9.8 Clusterwide Shutdown Causes `gated` to Dump Core and Generate Many Task Terminate Messages

When you perform a clusterwide shutdown using the `shutdown -c` command, the `gated` daemon may dump core to the `/usr/tmp` directory and generate many messages on the consoles of members that are shutting down. This behavior is benign and you can safely ignore it.

3.9.9 Characters May Be Missing or Jumbled in Terminal Window Displays

When you boot a cluster member and invoke a terminal window, you may notice that some characters are replaced by utterly unsuitable, but often entertaining, characters and that other characters (such as the colon) are

missing. This problem is annoying but benign. Although the font mapping to the display is incorrect, the correct character codes are being generated by the keyboard.

To resolve this situation, exit and reenter the desktop session and restart the terminal window.

3.9.10 The `addvol` and `rmvol` Commands Do Not Automatically Update Member-Specific `etc/clu_bdmgr.conf` Files

When you add a volume to or delete a volume from the cluster root domain (using the `addvol` and `rmvol` commands respectively), the member-specific files that record information from the `h` partition of the member boot disk (`etc/clu_bdmgr.conf`) files should automatically be updated. However, a problem exists where each member's file is not updated until the member is rebooted.

To update the member-specific `etc/clu_bdmgr.conf` files without rebooting each member, enter the following command on any cluster member after modifying the cluster root domain:

```
# /usr/sbin/clu_bdmgr -u
```

3.9.11 Spurious Message from `clu_upgrade undo install` Command

The `clu_upgrade undo install` command displays the following message:

```
> Cannot make temporary copy of /usr/share/man/man4/.Old...Old.4.gz
   in .RollTemp...Old.4.gz
```

You can safely ignore this message.

3.10 SysMan Menu Notes

The following notes discuss known problems that you may encounter when using SysMan Menu in a cluster.

3.10.1 Do Not Use SysMan Menu to Configure Remote Member Network Interface Cards

Although the `netconfig` utility and the Network Configuration option of the SysMan Menu give you the option to configure the network interface cards of remote members of a cluster, do not do so. Regardless of which remote member you chose to manage, the management utility displays only the local network devices from the node on which you have invoked it. If you proceed to configure these devices, the utility will erroneously copy the local member's configuration to the remote member.

To avoid this problem, always configure a member's network interface cards by running the `netconfig` utility and the Network Configuration option of the SysMan Menu locally on that member.

3.10.2 Clusterwide `nfscnfig` Deconfigure Operation May Stop Daemons on Members with Overrides

If NFS has been configured clusterwide and a cluster member has NFS configured with member-specific values, that member is considered to have set up local overrides. If a member has an override, it should not be subject to clusterwide configuration changes, whether they enable or disable NFS features. However, in the current implementation, when an administrator uses the Network File System (NFS) Configuration option of the SysMan Menu or the `nfscnfig` utility to deconfigure an NFS daemon (such as `proplisd`) clusterwide, the utility will stop the daemon on all members, including members that have local overrides defined.

If the Network File System (NFS) Configuration option of the SysMan Menu or the `nfscnfig` utility erroneously stops daemons on a member with local overrides, follow these steps to correct the situation:

1. Log on to the member where the daemons have stopped.
2. Run the Network File System (NFS) Configuration option of the SysMan Menu or the `nfscnfig` utility with focus to that member.
3. Restart the NFS daemons.

3.10.3 Nonroot Users Cannot Run System Management Applications

Most system management applications require `root` privileges to make configuration changes. Nonroot users are permitted to run system management applications only to view the current configuration. They are prevented from changing the configuration.

In a cluster, the system management applications use the remote shell command (`rsh`) to execute commands at a remote host. Part of the `rsh` command processing includes verifying access in the remote user's `$HOME/.rhosts` file in their home directory. For this reason, a nonroot user, without a home directory, running a system management application might encounter a core dump. Users can avoid these problems by ensuring that they have home directories set up before attempting to use the system management applications.

An additional problem exists where certain SysMan Menu tasks can be run only by the `root` user, regardless of whether a nonroot user's home directory has been established.

Each task in the SysMan Menu is associated with an action name. A task's action name is the same as its accelerator (displayed by the `sysman -list` command). These action names are associated with privileges by the Configure Division of Privilege (DOP) application in the SysMan Menu. Nonroot users can be granted the privilege to run specific actions.

Several SysMan Menu tasks do not have a required privilege associated with their actions. Such tasks can only be performed when a user is logged in as `root`. These tasks are:

- View hardware hierarchy
- View cluster
- View device information
- Manage Cluster File System
- Manage DRD Storage
- Cluster Alias Manager

3.11 SysMan Station Notes

The following notes discuss known problems that you may encounter when using SysMan Station in a cluster.

3.11.1 SysMan Station Daemon Might Dump Core when Multiple Members Are Booted Simultaneously

If multiple members of a cluster of three or more members are booted at the same time, the SysMan Station server (`smsd`) may fail on some members and dump core. This prevents you from using the the SysMan Station on those cluster members where the failure occurred.

When this occurs, you can safely delete the core dumps (named `core.smsd.node.n`) from the clusterwide root directory (`/`). To restart `smsd`, follow these steps:

1. Determine which members are still running the `smsd` daemon by entering the following command:

```
# ps -ax | grep smsd
```
2. On each of those members, stop `smsd` using the following command:

```
# /sbin/init.d/smsd stop
```
3. Enter the following command from one cluster member:

```
# rm /var/cluster/sms/*/objects/*
```

4. Restart `smsd` on all members by entering the following command on each member:

```
# /sbin/init.d/smsd start
```

3.11.2 SysMan Station Might Display Cluster Status Incorrectly

The SysMan Station relies on events generated by the Event Management (EVM) subsystem to monitor and display cluster status. In the following situations, the SysMan Station may reflect the state of the system incorrectly:

- The Filesystems light in the Monitor window may indicate a warning state (yellow) after all file system objects have returned to a normal state. This may occur after a new member has been added to the cluster. To clear this warning, restart the SysMan Station daemon (`smsd`) on the affected cluster members by following these steps:
 1. Close all open SysMan Station sessions.
 2. Enter the following command:

```
# /sbin/init.d/smsd restart
```
- After a cluster member has booted, the Network light in the Monitor window may indicate a warning state (yellow) when no network errors exist. This condition is caused by network events that are generated during the boot sequence. To clear this warning, follow these steps:
 1. Click on the Network light in the Monitor window to display the Network Event window.
 2. Click on the Clear Events button.
- If the cluster application availability daemon (`caad`) fails to start on a cluster member, the SysMan Station will not correctly display the state of CAA objects. For example, this can happen when the CAA license is not loaded on all the cluster members. To obtain accurate information on CAA applications from the SysMan Station, follow these steps:
 1. Start the `caad` daemon on the affected cluster members using the following command:

```
# /usr/sbin/caad
```
 2. Restart the SysMan Station daemon (`smsd`) using the following command:

```
# /sbin/init.d/smsd restart
```

Additionally, the Filesystem Attention group in the SysMan Station Monitor window does not properly update on all cluster members for suboptimal

file system states. If a file system becomes suboptimally configured, the Filesystem Attention group on only one cluster member will reflect the new state properly. SysMan Station clients connected to cluster members will not reflect this change in the Monitor window. However, the Physical Filesystem View on all cluster members will properly display this state information.

To correct this problem, stop and restart the SysMan Station daemon (`smsd`) on each cluster member where the Filesystems Attention Group is not reflecting the suboptimal state, as follows:

1. Close all open SysMan Station sessions.
2. Restart the SysMan Station daemon using the following command:

```
# /sbin/init.d/smsd restart
```

3.11.3 SysMan Station May Display New Hardware Objects Incorrectly

If a new disk device is added or an existing disk device is replaced in a running cluster, the SysMan Station's Hardware View may display the new or modified disk object incorrectly. The disk object may be positioned incorrectly in the hardware hierarchy; for example, the disk may be drawn as a child of the host object instead of as a child of a SCSI bus.

To correct the view, restart the SysMan Station daemon (`smsd`) on each cluster member by performing the following steps on all affected members:

1. Close all open SysMan Station sessions.
2. Enter the following command:

```
# /sbin/init.d/smsd restart
```

3.11.4 Properties Might Not Display for Selected Objects

Properties may not display for selected objects. The `Properties` dialog box may flash quickly on the screen or may not be displayed at all.

To work around this problem, continue to try and display properties in the current SysMan Station client or exit the SysMan Station client and start a new SysMan Station session.

3.11.5 Some SysMan Station Applications Display Wrong Target Member Name

When launched from the SysMan Station, the title bars of the following applications incorrectly display the name of the cluster member on which the SysMan Station client is running instead of the cluster member that is the target of the application's actions:

- Security Auditing Configuration
- Network Configuration Applications
- NFS Configuration Applications
- NTP Configuration Applications
- PPP Configuration Applications

The application is directed to the correct cluster member; only the name in the title bar is incorrect.

3.11.6 Documentation Notes

The following notes correct errors in the TruCluster Server Version 5.0A documentation:

- Figure 1-8 in the TruCluster Server *Hardware Configuration* manual is labeled incorrectly as being a no-single-point-of-failure (NSPOF) configuration. As depicted, the configuration is not NSPOF because the loss of a switch would result in the loss of an HSG80 RAID array controller pair. Refer to Figures 6-5 through 6-7 in the *Hardware Configuration* manual for illustrations of NSPOF multiple-bus failover configurations.
- Section 2.3 of the *Hardware Configuration* manual, in listing cluster Fibre Channel requirements and restrictions, incorrectly states that a maximum of four member systems is supported with Fibre Channel. As stated in Section 2.1 of the same manual, eight member systems may be connected to common storage over Fibre Channel in a fabric (switch) configuration.
- Table 2-1 of the TruCluster Server *Hardware Configuration* manual lists the number of shared KGPSA-BC PCI-to-Fibre Channel adapters supported in a cluster. This table also pertains to the KGPSA-CA PCI-to-Fibre Channel host bus adapter.
- Section 9.8.2 of the *Cluster Administration* manual has an incorrect form of the `vdump` command in the example of saving a dump image of a member boot partition. The correct command is as follows:

```
# vdump -ODf /var/cluster/members/member3/boot_part_vdump \
/cluster/members/member3/boot_partition
```

The extended example of restoring a member's boot partition should be replaced with the following:

1. Use `clu_get_info` to determine that `member3` is down:

```
# clu_get_info -m 3
Cluster memberid = 3
Hostname = member3.zk3.dec.com
```

```
Cluster interconnect IP name = member3-mc0
Member state = DOWN
```

2. **Select a new disk (in this example, dsk5) as the replacement boot disk for member3. Because the boot disk for member3 is dsk3, you are instructed to edit member3's /etc/sysconfigtab so that dsk5 is used as the new boot disk for member3.**

To configure dsk5 as the boot disk for member3, enter the following command:

```
# /usr/sbin/clu_bdmgr -c dsk5 3
The new member's disk, dsk5, is not the same name as the original disk
configured for domain root3_domain. If you continue the following
changes will be required in member3's/etc/sysconfigtab file:
vm:
  swapdevice=/dev/disk/dsk5b
clubase:
  cluster_seqdisk_major=19
  cluster_seqdisk_minor=175
```

3. **Mount member3's root domain (now on dsk5) so you can edit member3's /etc/sysconfigtab and restore the boot partitions:**

```
# mount root3_domain#root mnt
```

4. **Edit member3's /etc/sysconfigtab**

```
# cd /mnt/etc
# cp sysconfigtab sysconfigtab-bu
```

As indicated in the output from the clu_bdmgr command, change the values of the swapdevice attribute in the vm stanza and the cluster_seqdisk_major and cluster_seqdisk_minor attributes in the clubase stanza:

```
swapdevice=/dev/disk/dsk5b
clubase:
  cluster_seqdisk_major=19
  cluster_seqdisk_minor=175
```

5. **Restore the boot partition:**

```
# vrestore -xf /var/cluster/members/member3/boot_part_vdump -D /mnt
```

6. **Restore the h partition CNX information:**

```
# /usr/sbin/clu_bdmgr -h dsk5
```

The h partition information is copied from the cluster member where you run the clu_bdmgr command to the h partition on dsk5.

If the entire cluster is down, you need to boot one of the members from the clustered kernel. Once you have a single-member cluster running, you can restore the CNX h partition information to member3's new boot disk, dsk5, from /mnt/etc/clu_bdmgr.conf. Enter the following command:

```
# /usr/sbin/clu_bdmgr -h dsk5 /mnt/etc/clu_bdmgr.conf
```

7. Unmount the root domain for member3:

```
# umount root3_domain#root /mnt
```

8. You can now boot member3 into the cluster.

- Section 9.4.1.3 of the *Cluster Administration* manual presents an incorrect procedure for replacing RZ26, RZ28, RZ29, or RZ1CB-CA disks as direct-access I/O disks.

The following is the correct procedure:

If you replace an RZ26, RZ28, RZ29, or RZ1CB-CA direct-access I/O disk with a disk of the same type (for example, replace an RZ28-VA with another RZ28-VA), follow these steps to make the new disk a direct-access I/O disk:

1. Physically install the disk in the bus.
2. On every cluster member, enter the `hwmgrr` command to scan for the new disk as follows:

```
# hwmgrr -scan component -type disk
```

Allow a minute or two for the scans to complete.

3. If you want the new disk to have the same device name as the disk it replaced, you need to use the `hwmgrr -redirect scsi` command. For details, see `hwmgrr(8)` and the section on replacing a failed SCSI device in the *Tru64 UNIX System Administration* manual.
4. Enter the `clu_disk_install` command on each cluster member:

```
# clu_disk_install
```

Note

This command may take several minutes to complete if the cluster has a large number of storage devices.

Index

A

- accounting
 - can exhaust process quotas in large clusters, 3-18
- addvol command
 - does not automatically update clu_bdmgr.conf file, 3-19

B

- Bad configuration message
 - from clu_delete_member, 3-5
- boot disk
 - repairing a member's boot disk, 3-24
- booting
 - ATTENTION message during, 3-3
 - ict_mct messages during, 3-4
 - using HSG80 or HSZ80 disks, 3-1

C

- CAA
 - application resources may time out during quorum loss, 3-11
 - applications may not be fully initialized at boot time, 3-15
 - balanced placement does not work properly for applications started by SysMan Menu, 3-11
 - caa_register -u command does not correctly update a network resource's state, 3-10

- caa_relocate command does not relocate all application resources from the specified member, 3-14
- caa_stat tabular output corrupted by very long resource names, 3-16
- certain resource profile options cannot be updated, 3-13
- default action scripts can generate too many log files, 3-16
- default media changer resource template contains incorrect resource type, 3-10
- does not recognize fully qualified host names, 3-11
- erroneously lets you update a running application's required resources, 3-13
- event messages are malformed when displayed by the EVM viewer, 3-10
- human intervention message from caa_start command, 3-13
- lets you register application resources with a nonexistent or null action script, 3-13
- may stop required resources when a resource requiring them fails to start, 3-12
- purpose of, 1-1t
- tape resource problems, 3-9
- trying to relocate to down member message, 3-11
- caa_profile command

- validates application resource profiles with a nonexistent or null action script, 3-13
- caa_register command
 - erroneously lets you update a running application's required resources, 3-13
 - lets you register or update application resources with a nonexistent or null action script, 3-13
 - u option does not correctly update a network resource's state, 3-10
- caa_relocate command
 - does not relocate all application resources from the specified member, 3-14
- caa_start command
 - human intervention message, 3-13
- caa_stat
 - tabular output corrupted by very long resource names, 3-16
- Cannot communicate with CAA daemon message, 3-15
- CDFS
 - management utilities not cluster-aware, 2-2
 - supported for read-only access, 2-1
- CDSL
 - malformed, 3-2
 - purpose of, 1-1t
- CFS
 - purpose of, 1-1t
 - relocation failure, 3-6
 - restrictions on, 2-1
- cfsmgr command
 - dumps core when -p or -m option is used with -F raw option, 3-6
 - fails to relocate file systems containing applications that wire memory, 3-6
- characters
 - missing or strange in terminal window display, 3-18
- clu_add_member command, 3-2
 - if kernel build for newly added member fails on first boot, 3-3
- clu_check_config command
 - ATTENTION message from, 3-3
- clu_create command, 3-1
 - malformed CDSL message from, 3-2
- clu_delete_member command
 - Bad configuration message from, 3-5
 - cannot delete a member with a boot disk with a bad disk label, 3-4
 - does not adjust expected votes when deleting a member with bad boot disk, 3-4
 - NTP and mail messages from, 3-5
- clu_get_info command
 - failure messages, 3-18
- clu_upgrade command
 - spurious message from clu_upgrade undo install, 3-19
- cluster alias
 - purpose of, 1-2t
- cluster application availability
 - base operating system (*See* CAA)
- cluster configuration check
 - ATTENTION message from, 3-3
- cluster file system
 - (*See* CFS)
- cluster root
 - cannot add LSM volume to, 2-2
 - cannot use LSM on, 2-2
- clusterwide namespace, 1-1t
- connection manager
 - purpose of, 1-1
- context-dependent symbolic link
 - (*See* CDSL)

D

- device request dispatcher
 - (*See* DRD)

direct I/O
 wrong errno value returned on
 opens of memory-mapped files,
 3-7

direct-access I/O
 replacing disks, 3-26

disk
 replacing a failed, 3-16

distributed lock manager
 (*See* DLM)

DLM
 purpose of, 1-3t

documentation
 errors in, 3-24

DRD
 purpose of, 1-1

E

/etc/cfgmgr.auth file
 adding first member to, 3-2

etc/clu_bdmgr.conf file
 not automatically updated by addvol
 and rmvol commands, 3-19

F

FFM file system
 not supported, 2-2

file system
 CFS, 1-1t
 unmounted unexpectedly, 3-6

floppy disks
 and UFS read-only file system, 2-1

G

gated daemon
 dumps core and generates messages
 at shutdown, 3-18

H

hang
 in LSM commands, 3-9

HSG80 or HSZ80 controllers
 used for member boot disk, 3-1

hwmgr command
 -cluster option does not work
 properly with -scan scsi, 3-16

I

icssvr_daemon_from_pool processes,
 3-18

installation of cluster software
 do not use SysMan Menu to
 perform, 3-1

K

kernel build
 if fails for newly added member on
 first boot, 3-3

kernel configuration management
 database
 adding first member to, 3-2

L

Logical Storage Manager
 (*See* LSM)

LSM
 clusterwide, 1-1

 commands hang due to vold daemon
 failure, 3-9

 failed LSM disk can prevent
 members from performing LSM
 commands, 3-8

 restrictions on, 2-2

 use only on shared disks in a
 cluster, 3-7

vold daemon fails to start in cluster, 3-9
voldisk list command does not display all non-LSM disks within a cluster, 3-8

M

mail
 messages from `clu_delete_member` command, 3-5
malformed CDSL message, 3-2
media changer CAA resource
 default resource template contains incorrect resource type, 3-10
member boot disk
 repairing, 3-24
member boot partition
 cannot use LSM on, 2-2
 on HSZ80 or HSG80 storage, 3-1
Memory Channel
 errors when booting a member into a cluster, 3-4
 purpose of, 1-2t
messages
 ATTENTION from boottime cluster check, 3-3
 Bad Cluster interconnect IP name or address (from `clu_get_info` command), 3-18
 Bad configuration message from `clu_delete_member` command, 3-5
 Cannot communicate with CAA daemon, 3-15
 ERROR: Configuration daemon is not accessible, 3-7
 from `clu_upgrade undo install` command, 3-19
 human intervention messages from `caa_start` command, 3-13
 `ict_mct`, 3-4

mail messages from
 `clu_delete_member` command, 3-5
malformed CDSL, 3-2
NTP messages from
 `clu_delete_member` command, 3-5
Server Relocation Failed from `cfsmgr` command, 3-6
Trying to relocate to a dead member (from `caa_relocate` command), 3-11
MFS
 not supported, 2-2
mlock system call
 CFS relocation problems, 3-6

N

named pipes
 not supported, 2-2
namespace
 clusterwide, 1-1t
netconfig utility
 cannot configure remote member's network interface cards, 3-19
network
 loss of connectivity, 3-17
Network Information Services (*See* NIS)
network interface cards
 configuring, 3-19
NFS client file system
 failover not supported, 2-1
NFS server
 highly available, 1-2t
nfsconfig utility
 ignores local overrides when deconfiguring NFS daemons, 3-20
NIS
 do not configure before creating the cluster, 3-2
NTP

messages from `clu_delete_member` command, 3-5

P

`plock` system call
CFS relocation problems, 3-6
`/proc` file system
not supported, 2-2
process accounting
can exhaust process quotas in large clusters, 3-18

Q

quorum disk
cannot use LSM on, 2-2
quota command
not supported in a cluster, 2-2
quotas
not supported for file systems in a cluster, 2-2
supported for AdvFS filesets in a cluster, 2-2

R

RAID 5
not supported by LSM in a cluster, 2-2
`rcinet` restart command
do not use, 3-17
replacing direct-access I/O disks, 3-26
`rmvol` command
does not automatically update `clu_bdmgr.conf` file, 3-19
rolling upgrade, 1-3t
root file system
cannot use LSM on, 2-2
RZ26, RZ28, RZ29, or or RZ1CB-CA
replacing as direct-access I/O disks, 3-26

S

security
cluster as single security domain, 1-2t
Server Relocation Failed message
from `cfsmgr` command, 3-6
shutdown
clusterwide shutdown causes gated to dump core and generate many messages, 3-18
shutdown command
for single member unmounts
inactive file systems, 3-6
single-system management, 1-2t
swap
cannot use LSM on, 2-2
`sysman -clone` command
succeeds inappropriately in a cluster, 3-18
SysMan management tools, 1-2t
SysMan Menu
balanced placement does not work properly for applications started by, 3-11
cannot configure remote member's network interface cards, 3-19
do not use Installation branch of Software menu of, 3-1
ignores local overrides when deconfiguring NFS daemons, 3-20
nonroot users cannot perform management tasks, 3-20
SysMan Station
applications display wrong target member name, 3-23
Filesystem Attention group may not be updated for nonoptimal file system states, 3-22
may display cluster status incorrectly, 3-22

- may display new hardware objects incorrectly, 3-23
- may dump core when multiple members are booted simultaneously, 3-21
- properties may not display for selected objects, 3-23
- shows CAA application resources in UNKNOWN state as having an error, 3-14

T

- tape devices
 - become inaccessible to cluster if client system goes down before open completes, 3-17
 - on shared bus, 2-3
- tape resource
 - problems with, 3-9
- terminal window
 - is missing character or displays wrong ones, 3-18

U

- UFS

- and floppy disks, 2-1
- supported for read-only access, 2-1

V

- vold daemon
 - fails to start in a cluster, 3-9
- voldisk command
 - list command does not display all non-LSM disks within a cluster, 3-8
- volencap command
 - not supported for root domain, 2-2
- volrootmir command
 - not supported in a cluster, 2-2
- volunroot command
 - not supported in a cluster, 2-2
- vquota command
 - not supported in a cluster, 2-2

W

- wired memory
 - CFS relocation problems, 3-6

How to Order Tru64 UNIX Documentation

You can order documentation for the Tru64 UNIX operating system and related products at the following Web site:

<http://www.businesslink.digital.com/>

If you need help deciding which documentation best meets your needs, see the Tru64 UNIX *Documentation Overview* or call **800-344-4825** in the United States and Canada. In Puerto Rico, call **787-781-0505**. In other countries, contact your local Compaq subsidiary.

If you have access to Compaq's intranet, you can place an order at the following Web site:

<http://asmorder.nqo.dec.com/>

The following table provides the order numbers for the Tru64 UNIX operating system documentation kits. For additional information about ordering this and related documentation, see the *Documentation Overview* or contact Compaq.

Name	Order Number
Tru64 UNIX Documentation CD-ROM	QA-6ADAA-G8
Tru64 UNIX Documentation Kit	QA-6ADAA-GZ
End User Documentation Kit	QA-6ADAB-GZ
Startup Documentation Kit	QA-6ADAC-GZ
General User Documentation Kit	QA-6ADAD-GZ
System and Network Management Documentation Kit	QA-6ADAE-GZ
Developer's Documentation Kit	QA-6ADAG-GZ
Reference Pages Documentation Kit	QA-6ADAF-GZ

Reader's Comments

TruCluster Server

Release Notes

AA-RHH1B-TE

Compaq welcomes your comments and suggestions on this manual. Your input will help us to write documentation that meets your needs. Please send your suggestions using one of the following methods:

- This postage-paid form
- Internet electronic mail: readers_comment@zk3.dec.com
- Fax: (603) 884-0120, Attn: UBPG Publications, ZKO3-3/Y32

If you are not using this form, please be sure you include the name of the document, the page number, and the product name and version.

Please rate this manual:

	Excellent	Good	Fair	Poor
Accuracy (software works as manual says)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Clarity (easy to understand)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Organization (structure of subject matter)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Figures (useful)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Examples (useful)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Index (ability to find topic)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Usability (ability to access information quickly)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Please list errors you have found in this manual:

Page	Description
_____	_____
_____	_____
_____	_____
_____	_____

Additional comments or suggestions to improve this manual:

What version of the software described by this manual are you using? _____

Name, title, department _____

Mailing address _____

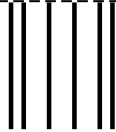
Electronic mail _____

Telephone _____

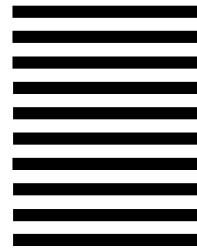
Date _____

----- Do Not Cut or Tear - Fold Here and Tape -----

COMPAQ



NO POSTAGE
NECESSARY IF
MAILED IN THE
UNITED STATES



BUSINESS REPLY MAIL

FIRST CLASS MAIL PERMIT NO. 33 MAYNARD MA

POSTAGE WILL BE PAID BY ADDRESSEE

COMPAQ COMPUTER CORPORATION
UBPG PUBLICATIONS MANAGER
ZKO3-3/Y32
110 SPIT BROOK RD
NASHUA NH 03062-2698



----- Do Not Cut or Tear - Fold Here -----

Cut on This Line