

TruCluster Server

Cluster Release Notes

Part Number: AA-RHH1D-TE

July 2001

Product Version: TruCluster Server Version 5.1A

Operating System and Version: Tru64 UNIX Version 5.1A

This manual provides important information about TruCluster Server Version 5.1A.

For information that may become available after the publication of these release notes, go to the following URL:

http://www.tru64unix.compaq.com/docs/pub_page/tcr_update.html

© 2001 Compaq Computer Corporation

Compaq, the Compaq logo, AlphaServer, and TruCluster Registered in U.S. Patent and Trademark Office. Tru64 is a trademark of Compaq Information Technologies Group, L.P. in the United States and other countries.

UNIX and The Open Group are trademarks of The Open Group in the United States and other countries. All other product names mentioned herein may be 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 document is provided "as is" without warranty of any kind and is subject to change without notice. The warranties for Compaq products are set forth in the express limited warranty statements accompanying such products. Nothing herein should be construed as constituting an additional warranty.

Contents

About This Manual

1 Features

1.1	Changes from TruCluster Server Version 5.1	1-2
1.1.1	Name of Cluster Interconnect Virtual Device Changed	1-2

2 Restrictions

2.1	Hardware Restriction	2-1
2.1.1	Restrictions on Memory Channel 2 on a GS80, GS160, or GS320 System	2-1
2.2	CFS Restrictions	2-1
2.3	AdvFS Restriction	2-2
2.4	Restrictions on Use of Internet Protocol Version 6 (IPv6)	2-3
2.5	Restrictions on Multiple Network Interfaces	2-3

3 Issues and Known Problems

3.1	Installation	3-1
3.1.1	Update to Latest Firmware Before Installing Tru64 UNIX Version 5.1A	3-1
3.1.2	Do Not Use the Installation Branch of the Software Option of the SysMan Menu	3-2
3.2	Cluster Creation and Member Addition	3-2
3.2.1	Conflict in Use of the Default Physical Cluster Interconnect IP Name	3-2
3.2.2	Misleading LAN Interconnect Information Provided by <code>clu_create</code>	3-3
3.2.3	The <code>clu_create</code> Command Does Not Add the First Member's Fully Qualified Hostname to the <code>/etc/cfgmgr.auth</code> File	3-3
3.2.4	Re-adding a Member with <code>clu_add_member -c</code> Does Not Correctly Configure NetRAIN	3-3
3.2.5	Run <code>/sbin/kreg</code> After Building Kernel with No Kernel Layered Products	3-4
3.3	Bootting and Shutdown	3-4

3.3.1	Do Not Reboot All Members Simultaneously	3-4
3.3.2	AlphaServer GS80/160/320 May Hang on Boot	3-5
3.3.3	A Member May Hang on Boot in a Cluster Using a KZPBA-CB SCSI Bus Adapter	3-6
3.3.4	Booting Member Hangs While Setting Time and Date with ntpdate	3-6
3.3.5	CNX Panic During Boot	3-7
3.3.6	Member Hangs During Clusterwide Shutdown	3-7
3.3.7	Booting a New Member Without a Cluster License Displays ATTENTION Message	3-7
3.3.8	Boot All Cluster Members Before Starting Applications That Use the Memory Channel API Library	3-8
3.3.9	"Cannot make specified change" Message During Boot	3-8
3.3.10	Booting a Member During Disaster Recovery	3-8
3.4	Cluster Configuration	3-9
3.4.1	Configurations That Support Ethernet Hubs	3-9
3.4.2	Set Ethernet Switch Address Aging to 15 Seconds	3-10
3.4.3	On LAN-based Clusters, Add Physical Network Addresses to ifaccess.conf	3-11
3.4.4	Whenever Changing Network Interfaces, Update ifaccess.conf	3-12
3.4.5	Configuring a Cluster Member as a DHCP Server	3-13
3.4.6	After Entering TruCluster Server License Information, Run lmf reset	3-14
3.4.7	Deleting a Member with an Inaccessible or Bad Member Boot Disk	3-14
3.5	File System	3-14
3.5.1	Do Not Create UFS File Systems on a Quorum Disk	3-15
3.5.2	CFS Relocation Failures Involving Applications That Wire Memory	3-15
3.6	LSM	3-15
3.6.1	LSM Commands Fail When Disk Left in Failing State	3-15
3.6.2	Problem Encapsulating Swap in Clusters with Long Host Names	3-16
3.6.3	Problem Encapsulating Swap When Remote Member Swap Areas Are Already Queued for Encapsulation	3-16
3.6.4	The voldisk list Command Does Not Display All Non-LSM Disks in a Cluster	3-16
3.7	CAA	3-16
3.7.1	Failure Threshold Value Should Not Be Greater Than 10 .	3-16
3.7.2	Problems with caa_relocate When Multiple Interdependent Applications Are Specified	3-17
3.7.3	Action Scripts PATH Considerations	3-17

3.7.4	The <code>caa_register -u</code> Command Does Not Correctly Update a Nonapplication Resource's State	3-18
3.7.5	Changes to <code>HOSTING_MEMBERS</code> List Can Leave Applications Running on Disallowed Members	3-18
3.7.6	CAA Events Are Malformed When Viewed from the Event Manager (EVM) Viewer	3-18
3.7.7	CAA Lets You Update a Running Application's Required Resources	3-18
3.7.8	SysMan Station Shows CAA Application Resources in UNKNOWN State as Having an Error	3-19
3.8	Rolling Upgrade	3-19
3.8.1	Rolling Upgrade: Do Not Shut Down a Cluster Member To Single-User Mode	3-19
3.8.2	Stop <code>prpasswd</code> Prior to Rolling Upgrade on Clusters Running Enhanced Security Environment	3-20
3.8.3	Updating Worldwide Language Support from CD-ROM During a Rolling Upgrade	3-21
3.8.4	Run Updates to <code>dop</code> Database Only on Upgraded Members	3-21
3.8.5	Recovering from a Failed Install Stage During a Rolling Upgrade from V5.1 to V5.1A	3-21
3.9	Miscellaneous Administration	3-22
3.9.1	RIS Boot Failures When Cluster is RIS Server	3-22
3.9.2	The <code>cfsstat</code> Command Can Return Incorrect Statistics	3-22
3.9.3	All Tape Writes at Same Density and Compression	3-23
3.9.4	Tape Opens Take Long Time to Fail When No Tape Is Present	3-23
3.9.5	Avoid Using Block Sizes Greater Than 512K Bytes for Tape Transfers	3-23
3.9.6	Disabling Internet Services on a Per-Member Basis	3-23
3.9.7	The <code>hwmgr -show comp</code> Command May Report an Inconsistency Error When Creating a Clusterwide Name for a SCSI Device	3-24
3.9.8	The <code>hwmgr -scan scsi</code> Command Does Not Work Clusterwide	3-24
3.9.9	Adding a Disk to a Running Cluster	3-24
3.9.10	Running Process Accounting on Large Clusters Can Exhaust Member Process Quotas	3-25
3.10	SysMan Menu	3-25
3.10.1	Nonroot Users with No Home Directory Cannot Run System Management Applications	3-25
3.11	SysMan Station	3-26

3.11.1	SysMan Station Does Not Display Applications with Long Resource Names	3-26
3.11.2	Unable to Expand Host Object in a Cluster	3-26
3.11.3	SysMan Station Might Display Cluster Status Incorrectly	3-26
3.11.4	SysMan Station Might Display New Hardware Objects Incorrectly	3-28
3.11.5	Properties Might Not Be Displayed for Selected Objects ..	3-28
3.11.6	Some SysMan Station Applications Display Wrong Target Member Name	3-28
3.12	Documentation	3-28
3.12.1	Recovering Cluster Root File System	3-29
3.12.2	Correct Size of Member Boot Disk a (Boot) Partition	3-29
3.12.3	Error in Example of Hexadecimal Format of IP Address ..	3-29
3.12.4	Media Changer Utility, mcutil, Works on Remote Members	3-29
3.12.5	Using LSM to Mirror the Cluster Root File System	3-30
3.12.6	Migrating from Automount to AutoFS	3-30
3.12.6.1	Migrating Without a Reboot	3-31
3.12.6.2	Migrating When Rebooting a Cluster Member	3-34
3.12.6.3	Migrating When Rebooting the Cluster	3-37
3.12.7	Clusterwide IPC-Supported Mechanisms	3-39
3.12.8	Software Product Description (SPD) Replaced by QuickSpec	3-39

Index

Tables

1-1	TruCluster Server Version 5.1A Features	1-1
-----	---	-----

About This Manual

This manual provides important information about TruCluster™ Server Version 5.1A, which runs on the Compaq Tru64™ UNIX Version 5.1A operating system.

Audience

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

Organization

This manual is organized as follows:

- | | |
|------------------|--|
| <i>Chapter 1</i> | Lists new features in TruCluster Server Version 5.1A. |
| <i>Chapter 2</i> | Lists restrictions on the use of certain components in TruCluster Server Version 5.1A. |
| <i>Chapter 3</i> | Discusses issues and known problems with TruCluster Server Version 5.1A, including updates to manuals since their publication. |

Related Documents

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

- *TruCluster Server QuickSpec* — Describes TruCluster Server Version 5.1A, including information about its capabilities and the hardware it supports. The *QuickSpec* replaces the TruCluster Server *Software Product Description (SPD)*. You can find the latest version of the *QuickSpec* at the following URL:
http://tru64unix.compaq.com/docs/pub_page/spds.html
- *Cluster Technical Overview* — Provides an overview of the TruCluster Server technology.
- *Cluster Hardware Configuration* — Describes how to set up the systems that are to become cluster members, and how to configure cluster shared storage.

- *Cluster LAN Interconnect* — Describes how to configure and maintain local area network (LAN) hardware for use as a cluster interconnect in the TruCluster Server environment.
- *Cluster Installation* — Describes how to install the TruCluster Server software.
- *Cluster Highly Available Applications* — Describes how to deploy existing applications on a TruCluster Server cluster, and how to write cluster-aware applications.
- *Cluster Administration* — Describes cluster-specific administration tasks.

You can find the latest version of the TruCluster Server documentation at:

http://www.tru64unix.compaq.com/docs/pub_page/doc_list.html

For information that may become available after the publication of these release notes, go to the following URL:

http://www.tru64unix.compaq.com/docs/pub_page/tcr_update.html

Note

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 you install, configure, and use your TruCluster Server cluster.

Icons on Tru64 UNIX Printed Manuals

The printed version of the Tru64 UNIX documentation uses letter icons on the spines of the manuals to help specific audiences quickly find the manuals that meet their needs. (You can order the printed documentation from Compaq.) The following list describes this convention:

- G Manuals for general users
- S Manuals for system and network administrators
- P Manuals for programmers
- R Manuals for reference page users

Some manuals in the documentation help meet the needs of several audiences. For example, the information in some system manuals is also

used by programmers. Keep this in mind when searching for information on specific topics.

The *Documentation Overview* provides information on all of the manuals in the Tru64 UNIX documentation set.

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
```

Please include the following information along with your comments:

- The full title of the manual and the order number. (The order number appears on the title page of printed and PDF versions of a manual.)
- 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:

%

\$

A percent sign represents the C shell system prompt.
A dollar sign represents the system prompt for the Bourne, Korn, and POSIX shells.

#

A number sign represents the superuser prompt.

% cat	Boldface type in interactive examples indicates typed user input.
<i>file</i>	Italic (slanted) type indicates variable values, placeholders, and function argument names.
[] { }	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.
...	In syntax definitions, a horizontal ellipsis indicates that the preceding item can be repeated one or more times.
cat(1)	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.
Return	In an example, a key name enclosed in a box indicates that you press that key.
Ctrl/ <i>x</i>	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 5.1A supports all of the features of TruCluster Server Version 5.1 plus additional enhancements that are listed in Table 1–1.

Table 1–1: TruCluster Server Version 5.1A Features

Feature	Description	For more information, see:
Logical Storage Manager (LSM) mirroring of swap and cluster root	You can use LSM to mirror swap and the cluster root file system.	<i>Cluster Administration</i> and <i>volmigrate(8)</i>
Cluster File System (CFS) performance improvements	Improved file system read performance with fully connected storage.	<i>Cluster Technical Overview</i>
Local area network (LAN) cluster interconnect	The cluster interconnect, which serves as a private communications channel between cluster members, can now use either a private LAN or Memory Channel.	<i>Cluster LAN Interconnect</i>
CFS support for quotas	You can set Berkeley Standard Distribution (BSD)-type user and group quotas on AdvFS file systems.	<i>Cluster Administration</i>
Support for Network File System (NFS) server access via a non-default cluster alias	You can define an alternate cluster alias that allows you to limit NFS server activity to those members that are actually connected to the storage that contains exported file systems. NFS clients use this alternate alias when they mount file systems served by the cluster.	<i>Cluster Technical Overview, Cluster Administration</i>
Server-side UNIX File System (UFS) read/write and Server-side Memory File System (MFS) read/write	Read/write access is supported for UFS and MFS file systems that are mounted as partitioned file systems (the default for UFS and MFS). A partitioned file system is accessible only on the member where the file system is mounted. UFS and MFS file systems that are mounted as read-only are accessible via CFS anywhere in the cluster.	<i>Cluster Technical Overview, Cluster Administration</i>

1.1 Changes from TruCluster Server Version 5.1

This section describes differences between TruCluster Server Version 5.1 and Version 5.1A.

1.1.1 Name of Cluster Interconnect Virtual Device Changed

For Version 5.1A, the name of the virtual device for the cluster interconnect has changed from `mc0` to `ics0`. If you perform a full installation of Version 5.1A, or perform a rolling upgrade from Version 5.1, the `NETDEV_x` configuration variable in each member's `/etc/rc.config` file that corresponds to this device will be defined as `ics0`.

Restrictions

This chapter lists restrictions on the use of TruCluster Server Version 5.1A features.

2.1 Hardware Restriction

This section describes hardware restrictions in a cluster.

2.1.1 Restrictions on Memory Channel 2 on a GS80, GS160, or GS320 System

If you have a Memory Channel 2 (MC2) module installed on a peripheral component interconnect (PCI) bus of a GS80, GS160, or GS320 system, that bus can contain only another MC2 module or the CCMFB fiber-optic module. No other module can be installed on that PCI bus, not even the standard I/O module.

The section on Memory Channel restrictions in the *Cluster Hardware Configuration* manual incorrectly states that this restriction applies only to redundant MC2 modules jumpered for 512 MB. In fact, the restriction applies to all configurations of MC2 on a GS80, GS160, or GS320.

2.2 CFS Restrictions

The Cluster File System (CFS) has the following restrictions in TruCluster Server Version 5.1A:

- 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.

If the member that has mounted the NFS file system shuts down or fails, the file system is automatically unmounted and CFS begins to clean up the mount points. During the cleanup process, members that access 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 to 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 that were 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 that are listed in the `/etc/fstab` file that are not currently mounted and served in the cluster. (If you are using AutoFS or automount, the remount will happen automatically.)

- 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.

2.3 AdvFS Restriction

The following restrictions apply to the Advanced File System (AdvFS) in a cluster:

- The `cluster_root` domain should contain only one fileset, `root`.
The software does not prevent you from adding filesets to `cluster_root`, but additional filesets can cause a panic if `cluster_root` has to fail over.
- If you create a clone of `/`, the clusterwide `root`, and mount it, the cloned fileset is added to the `cluster_root` domain. If `cluster_root` has to fail over while the cloned fileset is mounted, the cluster will panic.
If you make backups of the clusterwide root from a cloned fileset, keep to a minimum the time during which the clone is mounted. Mount the cloned fileset, perform the backup, and unmount the clone as quickly as possible.
- You cannot use the `addvol` command to add volumes to a member's root domain (the `a` partition on the member's boot disk). Instead, you must delete the member from the cluster, use `diskconfig` or SysMan to configure the disk appropriately, and then add the member back into the

cluster. For the configuration requirements for a member boot disk, see the TruCluster Server *Cluster Installation* manual.

2.4 Restrictions on Use of Internet Protocol Version 6 (IPv6)

The following restrictions apply to using IPv6 addresses in a cluster:

- An IPv6 address cannot be associated with a cluster alias.
- An IPv6 address cannot be used for a cluster interconnect address; `clu_create` and `clu_add_member` do not accept IPv6 addresses for the interconnect address.

Cluster members can use and advertise IPv6 addresses.

2.5 Restrictions on Multiple Network Interfaces

Tru64 UNIX supports multiple network interfaces in the following configurations:

- Multiple network interfaces in the same subnet
- Redundant Array of Independent Network Adapters (NetRAIN)

Cluster alias supports NetRAIN; it does not support multiple network interfaces in the same subnet.

The LAN interconnect supports only NetRAIN; it does not support NIC-based link aggregation or multiple network interfaces in the same subnet.

Switch-to-switch link aggregation is supported, as discussed in the *Cluster LAN Interconnect* manual.)

For more information about multiple network interfaces, see the section on network interfaces in the Tru64 UNIX *Network Administration: Connections* manual.

Issues and Known Problems

The following sections describe issues and known problems with TruCluster Server Version 5.1A.

3.1 Installation

The information in this section applies to installation.

3.1.1 Update to Latest Firmware Before Installing Tru64 UNIX Version 5.1A

Before installing Tru64 UNIX and TruCluster Server Version 5.1A, update all systems that will become cluster members with the latest firmware. A cluster member running old firmware may not be able to use some hardware connected to the cluster. For example, with old firmware, a member with a boot disk behind an HSZ80 or HSG80 controller may fail to boot, indicating "Reservation Conflict" errors.

To update a system's firmware, do the following:

1. Insert the firmware CD-ROM in the drive and boot from it:

```
>>> boot cdrom_console_device_name
```

The firmware update utility automatically identifies your system type and model and determines the correct firmware revision required for your system.

2. Follow the instructions on the screen. There is an automatic display of the READ-ME-FIRST file, which describes the firmware changes included in the update.
3. Power off the processor for at least 10 seconds when the firmware update is complete to initialize the new firmware.

If you do not have access to a firmware CD-ROM, you can find the latest firmware at the following URL:

```
ftp.digital.com/pub/Digital/Alpha/firmware/readme.html
```

You can download the firmware and associated documentation with the anonymous File Transfer Protocol (FTP).

3.1.2 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 that are discussed in the TruCluster Server *Cluster Installation* and *Cluster Administration* manuals.

3.2 Cluster Creation and Member Addition

Information in this section applies to creating a cluster and adding cluster members.

3.2.1 Conflict in Use of the Default Physical Cluster Interconnect IP Name

The default physical cluster interconnect IP name has the form `membermemberID-icstcp0`.

The `clu_create` and `clu_add_member` commands use `ping` to determine whether the default name is already in use on the net. If this check finds a host already using the default IP name, then the command will fail. You are instructed:

```
Enter the physical cluster interconnect interface device name []
```

After this, the command fails. Depending on which command was executing at the time of failure, you get one of the following messages:

```
Error: clu_create: Bad configuration
```

```
Error: clu_add_member: Bad configuration
```

If you see either of these messages, look in `/cluster/admin/clu_create.log` or `/cluster/admin/clu_add_member.log`, as appropriate, for the following error message:

```
Error: A system with the name 'membermemberID-icstcp0'  
is currently running on your network.
```

If you find this message, contact your network administrator about changing the hostname of the system already using the default IP name, because the `clu_create` and `clu_add_member` commands do not allow you to change the default physical cluster interconnect IP name.

3.2.2 Misleading LAN Interconnect Information Provided by `clu_create`

When you select `Help` from the cluster interconnect selection, the information displayed by `clu_create` implies that Ethernet hubs used in a LAN interconnect must operate in full-duplex mode. In fact, 100 Mb/sec Ethernet hubs in half-duplex mode are supported, subject to certain restrictions specified in the *Cluster LAN Interconnect* manual.

More generally, a cluster must have a dedicated cluster interconnect to which all members are connected. The cluster interconnect serves as the primary communications channel between cluster members. For hardware, the cluster interconnect can use either Memory Channel or a private LAN. See the *Cluster Hardware Configuration* manual and the *Cluster LAN Interconnect* manual for configuration details.

3.2.3 The `clu_create` Command Does Not Add the 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 Re-adding a Member with `clu_add_member -c` Does Not Correctly Configure NetRAIN

If you have configured NetRAIN in a cluster with a LAN interconnect, and you re-add a member via the `clu_add_member -c` command using the configuration file created when the member was last added to the cluster, NetRAIN is not configured correctly for the re-added member.

When you boot the re-added member, you may see a message like the following:

```
CNX MGR: cannot form: quorum disk is in use. Unable to establish
contact with members using disk.
```

To resolve this problem, edit `/etc/sysconfigtab` on the re-added member, and change the lines in the `ics_ll_tcp` stanza. The value for

`ics_tcp_adapter0` incorrectly lists the device names of the Ethernet interfaces.

1. Set `ics_tcp_adapter0` to `nr0`:

```
ics_tcp_adapter0=nr0
```
2. For each network adapter that is in the NetRAIN set, assign the device name of the adapter to an `ics_tcp_nr0` array member.

For example, if the line in `/etc/sysconfigtab` looks like the following:

```
ics_tcp_adapter0=ee0,ee1,ee2,ee3
```

Then you would change it as follows:

```
ics_tcp_adapter0=nr0
ics_tcp_nr0[0]=ee0
ics_tcp_nr0[1]=ee1
ics_tcp_nr0[2]=ee2
ics_tcp_nr0[3]=ee3
```

3.2.5 Run `/sbin/kreg` After Building Kernel with No Kernel Layered Products

During cluster creation, if you build a clusterized kernel with no kernel layered products, you must then rebuild `/usr/sys/conf/.product.list`, which is the registration file for kernel layered products. To do this, run the following command:

```
# /sbin/kreg -l DEC TruCluster /usr/opt/TruCluster/sys
```

The situation where you build a clusterized kernel with no kernel layered products can arise if the initial build of the clusterized kernel fails. If that happens, you are offered as a default the option of building the kernel with no kernel layered products.

3.3 Booting and Shutdown

This section discusses requirements and restrictions for booting members into a cluster and for shutting down cluster members.

3.3.1 Do Not Reboot All Members Simultaneously

If you attempt to reboot all cluster members simultaneously, one or more members will hang on the way down due to quorum loss, and the other rebooting nodes may fail to boot because they are unable to rejoin the cluster (the hung nodes).

The method you use to reboot the entire cluster depends on your intent:

- To reboot all cluster nodes without halting the cluster, reboot one member at a time, allowing the rebooting member to rejoin the cluster before you reboot the next member.
- To reboot the entire cluster (cluster state will be lost), shut down the entire cluster with the `shutdown -c` command and then boot the members.

3.3.2 AlphaServer GS80/160/320 May Hang on Boot

The following applies only to an AlphaServer™ GS80, GS160, or GS320 that is a member in a cluster that uses Memory Channel for the cluster interconnect.

When you boot a cluster member that is an AlphaServer GS80, GS160, or GS320 and the member hangs, you must `halt` and `reset` the member, and then boot it again, if a message similar to the following is displayed:

```
>>> Registering CMS Services
```

or

```
>>> Registering CMS Services
```

```
>>> rm_meet_rail_requirements: No access to physical rail 0
```

For example, if you are managing the member from the partition's console head (connected to the SCM/SRM port of the hung partition), do the following:

1. Use the `[Esc]` sequence to get to the prompt for the system control manager (SCM):

```
[Esc] [Esc] scm
```

2. At the prompt for SCM, `halt` the partition and `quit`:

```
SLV_E1> halt
Slave request to master
SLV_E1> quit
```

3. At the system reference manual (SRM) console prompt, `reset` the system:

```
P00>>> reset
```

A `reset` from SRM is recommended because, if the system is partitioned, an SRM `reset` acts only on the partition where the command executes.

4. Boot the system.

3.3.3 A Member May Hang on Boot in a Cluster Using a KZPBA-CB SCSI Bus Adapter

When you boot a member of a cluster with a KZPBA-CB SCSI bus adapter, the member may hang during the boot. The console log for that member will display messages similar to the following:

```
cam_logger: SCSI event packet
cam_logger: bus 10 target 15 lun 0
ss_perform_timeout
timeout on disconnected request
Active CCB at time of error
cam_logger: SCSI event packet
cam_logger: bus 10 target 15 lun 0
isp_process_abort_queue
IO abort failure (mailbox status 0x0), chip reinit scheduled
Active CCB at time of error
cam_logger: SCSI event packet
cam_logger: bus 10
isp_reinit
Beginning Adapter/Chip reinitialization (0x3)
cam_logger: SCSI event packet
cam_logger: bus 10
isp_reinit
Fatal reinit error 1: Unable to bring Qlogic chip back online
```

If you see messages like this, you must reset the system and then boot it again.

On an AlphaServer GS80, GS160, or GS320, you can perform a system control manager (SCM) `halt` and `reset`. (See Section 3.3.2 for specifics.) On other systems, you will have to do a hardware reset before booting again.

If the reset fails to clear up the problem, you must cycle the power. On an AlphaServer GS80, GS160, or GS320, you can do this with the SCM `power off` and `power on` commands.

3.3.4 Booting Member Hangs While Setting Time and Date with `ntpdate`

A booting member may hang while attempting to use `ntpdate` to set the time and date. The last message before the hang is as follows:

```
Setting the current time and date with ntpdate
Press Ctrl/C, and the boot process will continue normally.
```

If the cluster uses the suggested configuration, it is running `xntpd`, the Network Time Protocol (NTP) daemon. In this situation, the member will get the time despite the hang.

3.3.5 CNX Panic During Boot

When you boot a member in a cluster with a large storage configuration, the member may panic and display the following message:

```
CNX MGR: Invalid configuration for cluster seq disk
```

If this occurs, reboot the member.

3.3.6 Member Hangs During Clusterwide Shutdown

This note applies only to clusters with a LAN interconnect.

During clusterwide shutdown, a member might hang for ten minutes or more before finally shutting down. If this delay is inconvenient, use the halt button to shut down the member.

3.3.7 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 the following messages:

```
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/cdslinvchk 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, the following message is displayed:

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

When the TruCluster Server license is not configured on a member, the cluster application availability (CAA) 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 *Cluster 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.

3.3.8 Boot 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.

3.3.9 "Cannot make specified change" Message During Boot

If your cluster is not configured to use Network Information Service (NIS), the following error messages are displayed when you first boot the cluster:

```
/cluster/admin/run/C30niscluster: Cannot make specified change.
```

```
/cluster/admin/run/C30niscluster : Bad exit code : 1
```

You can safely ignore these messages.

To prevent these messages from being displayed, remove the C30niscluster symbolic link and then recreate it, as follows:

```
# rm /cluster/admin/run/C30niscluster
# ln -sf /usr/ucb/true /cluster/admin/run/C30niscluster
```

3.3.10 Booting a Member During Disaster Recovery

When, as part of a disaster recovery process, you boot the first cluster member, the cluster may fail to form due to lack of quorum. This can happen because the cluster expected votes value is greater than the number of voting members.

If this situation occurs, you must boot one member interactively and set the value of the `cluster_expected_votes` variable of the `clubase` kernel subsystem.

For example:

1. Reboot one cluster member as follows:

```
>>> boot -fl "ia"
      (boot dkb200.2.0.7.0 -flags ia)
```



```
block 0 of dkb200.2.0.7.0 is a valid boot block
reading 18 blocks from dkb200.2.0.7.0
bootstrap code read in
base = 200000, image_start = 0, image_bytes = 2400
initializing HWRPB at 2000
initializing page table at fff0000
initializing machine state
setting affinity to the primary CPU
jumping to bootstrap code
```

```
Tru64 UNIX boot - Mon Jan 4 14:08:41 EDT 2000
```

```
Enter kernel_name [option_1 ... option_n]
Press Return to boot default kernel
'vmunix':vmunix clubase:cluster_expected_votes=1 [Return]
```

2. Reboot the other surviving cluster members in a similar fashion, incrementing the interactive boot value of `cluster_expected_votes` with each member you boot into the cluster. For example, when booting the second cluster member, specify the following when the interactive boot procedure prompts for the kernel name and options:

```
Enter kernel_name [option_1 ... option_n]
Press Return to boot default kernel
'vmunix':vmunix clubase:cluster_expected_votes=2 [Return]
```

3. Adjust cluster votes or membership based on your expectations of when the lost members can be restored. If they will be restored soon, use the `clu_quorum` command on a current member to temporarily remove their votes. If they will remain down for a long time, use the `clu_delete_member` command to remove them from the cluster. You may need to configure a quorum disk in the surviving cluster to improve its availability.

For more information, see the section on troubleshooting unfortunate expected vote settings in the chapter on managing cluster membership in the *Cluster Administration* manual.

3.4 Cluster Configuration

This section discusses problems with adding and deleting members, and configuring a quorum disk.

3.4.1 Configurations That Support Ethernet Hubs

The *Cluster LAN Interconnect* manual describes the Ethernet hardware that can be configured as a LAN interconnect. It is, however, ambiguous as to those configurations in which Ethernet hubs are supported.

All Ethernet hubs (also known as shared hubs to distinguish them from Ethernet switches) run in half-duplex mode. As a result, when a hub is used in a LAN interconnect, the Ethernet adapters connected to it must be set to (or must autonegotiate) 100 Mb/sec, half-duplex mode. (See Section 4.7.1 of the *Cluster LAN Interconnect* manual for additional information on how to accomplish this for the DE50x and DE60x families of adapters.)

Use of an Ethernet hub in a LAN interconnect is supported as follows:

- A single Ethernet adapter (or multiple adapters configured as a NetRAIN virtual interface) on each member, connected to a single Ethernet hub. Note that the use of NetRAIN in this configuration guards against the failure of a single adapter in a member's NetRAIN set. The hub remains a single point of failure.
- Multiple Ethernet adapters configured as a NetRAIN virtual interface on each member connected as depicted in Figure 2-5 of the *Cluster LAN Interconnect* manual to a pair of Ethernet hubs connected by a single crossover cable. This configuration guards against the failure of a single member adapter or a single hub failure. However, because the failure of the crossover cable link between the hubs can cause a cluster network partition (as described in Section 2.2.3 of the *Cluster LAN Interconnect* manual), this configuration is not recommended.

Unlike Ethernet switches, Ethernet hubs cannot be configured with multiple parallel crossover cables to guard against potential network partitions. Hubs do not provide features to detect and respond to routing loops.

Because of the performance characteristics of Ethernet hubs, use them only in small clusters (two or three members).

3.4.2 Set Ethernet Switch Address Aging to 15 Seconds

Ethernet switches maintain tables that associate MAC addresses (and virtual LAN (VLAN) identifiers) with ports, thus allowing the switches to efficiently forward packets. These forwarding data bases (also known as unicast address tables) provide a mechanism for setting the time interval when dynamically learned forwarding information grows stale and is invalidated. This mechanism is sometimes referred to as the aging time.

For any Ethernet switch participating in a LAN interconnect, set its aging time to 15 seconds. Failure to do so may cause the switch to erroneously continue to route packets for a given MAC address to a port listed in the forwarding table after the MAC address has moved to another port (for example, due to NetRAIN failover).

Ultimately, this can cause cluster nodes to lose communication and result in one or more nodes being removed from the cluster. The consequence may

be that one or more nodes hang due to loss of quorum, but may also result in one of several panic messages. For example:

```
CNX MGR: this node removed from cluster
```

```
CNX QDISK: Yielding to foreign owner
```

3.4.3 On LAN-based Clusters, Add Physical Network Addresses to `ifaccess.conf`

On each member of a cluster, an interface access filter configuration file (`ifaccess.conf`) is used to deny access from untrusted subnets to the cluster interconnect. Without this filtering, a system outside of the cluster could masquerade as a cluster member.

For clusters that use a LAN interconnect, each member's `/etc/ifaccess.conf` file must contain one entry for the virtual network address (for example, 10.0.0.0) and one entry for the physical network address (for example, 10.0.1.0) for each network interface.

For clusters that use a Memory Channel interconnect, each `/etc/ifaccess.conf` file must contain only the virtual network interface address.

In either case, the `/etc/ifaccess.conf` file is correctly configured on the initial cluster member, `memberid = 1`.

For additional cluster members, the configuration process automatically adds virtual network address entries in the `/etc/ifaccess.conf` file for each network interface. This is sufficient for a cluster that uses a Memory Channel interconnect.

However, in a LAN interconnect cluster, in addition to an entry for the virtual network address, an entry is needed for the physical network address for each network interface. The member addition process does not add these required entries to `/etc/ifaccess.conf`.

On each cluster member other than member 1, you must manually edit `/etc/ifaccess.conf` and add a line for the physical network addresses for each network interface.

The line should have the following format:

```
interface-name interconnect_net_address 255.255.255.0 deny
```

For example, if the physical cluster interconnect network address of a member is 10.0.1.0 (the default) and the virtual address is 10.0.0.0 (the default) and the interface name is `tu0`, the following lines are required in `/etc/ifaccess.conf`:

```
tu0 10.0.1.0 255.255.255.0 deny
tu0 10.0.0.0 255.255.255.0 deny
```

3.4.4 Whenever Changing Network Interfaces, Update ifaccess.conf

Whenever you add a new network interface, or change or replace an existing one, you must update the `/etc/ifaccess.conf` file on the cluster member where the change occurred. Do this to deny access from untrusted subnets to the cluster interconnect.

In a cluster with a Memory Channel interconnect, you need to add only a line for the virtual network address of the new or changed network interface.

In a cluster with a LAN interconnect, you must add a line for the virtual network address and another line for the physical network address for the new or changed network interface.

Make the change as follows:

1. Log in to the member where the network interface changed.
2. Use the `ifconfig` command to learn the names of the network interfaces. For example:

```
# ifconfig -l
ics0 lo0 sl0 ee0 ee1 ee2 tu0 tun0
```

3. Add a line to `/etc/ifaccess.conf` for each network interface. Edit the file appropriately for each changed interface.

Note

On a cluster that uses the LAN interconnect, you must add a line for the physical network address and another line for the virtual network address.

Each line should have the following format:

```
interface-name interconnect_net_address 255.255.255.0 deny
```

For example, if the virtual cluster interconnect network address of a member is 10.0.0.0 (the default) and a new interface card has been added to a cluster member and the interface name is `tu0`, add the following line to `/etc/ifaccess.conf`:

```
tu0 10.0.0.0 255.255.255.0 deny
```

If the cluster has a LAN interconnect, you must also add a line for the physical network address. Suppose the physical cluster interconnect network address of the member is 10.0.1.0 (the default). In addition to the line for the virtual network address, you must add the following line:

```
tu0 10.0.1.0 255.255.255.0 deny
```

The cluster interconnect network address is common to all members. To see the address used by your cluster, look in `/etc/ifaccess.conf`.

On clusters using the LAN interconnect, the network interface used by a member for the interconnect must not appear in the member's `/etc/ifaccess.conf` file.

To learn the name of a member's network interface for the LAN interconnect, log in to that member and use the `sysconfig` command to query the `ics_ll_tcp` subsystem and the `ics_tcp_adapter0` attribute.

In the following example, the member where the `sysconfig` command executes is using `nr0` as the network interface for the LAN interconnect:

```
# sysconfig -q ics_ll_tcp ics_tcp_adapter0
ics_ll_tcp:
ics_tcp_adapter0 = nr0
```

For more information, see `ifaccess.conf(4)`.

3.4.5 Configuring a Cluster Member as a DHCP Server

The description of configuring the Dynamic Host Configuration Protocol (DHCP) in the *Cluster Administration* manual is incorrect. Do not perform step 4, which reads "Under Server/Security Parameters, set the Canonical Name entry to the default cluster alias."

Perform all other steps as documented:

- Familiarize yourself with the DHCP server configuration process that is described in the chapter on DHCP in the *Tru64 UNIX Network Administration: Connections* manual.
- On the cluster member that you want to act as the initial DHCP server, run `/usr/bin/X11/xjoin` and configure DHCP.
- Select Server/Security.
- From the pulldown menu that currently shows Server/Security Parameters, select IP Ranges.
- Set the DHCP Server entry to the IP address of the default cluster alias.

There can be multiple entries for the DHCP Server IP address in the DHCP database. You might find it more convenient to use the `jdbdump` command to generate a text file representation of the DHCP database. Then use a text editor to change all the occurrences of the original DHCP server IP address to the cluster alias IP address. Finally, use `jdbmod` to repopulate the DHCP database from the file you edited. For example:

```
# jdbdump > dhcp_db.txt
# vi dhcp_db.txt
```

Edit `dhcp_db.txt` and change the owner IP address to the IP address of the default cluster alias.

Update the database with your changes by entering the following command:

```
# jdbmod -e dhcp_db.txt
```

- When you finish with `xjoin`, make DHCP a highly available application. DHCP already has an action script and a resource profile, and it is already registered with the CAA daemon. To start DHCP with CAA, enter the following command:

```
# caa_start dhcp
```

3.4.6 After Entering TruCluster Server License Information, Run `lmf reset`

When you add a cluster member, the `clu_add_member` command prompts you to indicate whether you want to register the TruCluster Server license for the new member at this time. If you answer "yes," you are prompted to enter either the license information or the name of the file with the license information. If you enter the license information, after `clu_add_member` completes, you must run the following command on the member where `clu_add_member` executed:

```
# lmf reset
```

3.4.7 Deleting a Member with an Inaccessible or Bad Member Boot Disk

The `clu_delete_member -f` command can delete a member with an inaccessible or bad boot disk. However, be aware that, when you delete a member with an inaccessible boot disk, `clu_delete_member -f` does not adjust expected votes in the running cluster (as a normal `clu_delete_member` command does). 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.

3.5 File System

This section discusses issues with the CFS, AdvFS, and NFS file systems in a cluster.

3.5.1 Do Not Create UFS File Systems on a Quorum Disk

Do not use the quorum disk for user data. The `mkfdmn` command prevents you from creating an AdvFS domain on a quorum disk. However, the `newfs` command incorrectly allows you to create a file system on a quorum disk. Do not use `newfs` to create a file system on a quorum disk.

3.5.2 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 that contains 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
```

To allow the relocation to complete for the domain or file system on which the executables reside, kill the processes that are running the executables using the `plock()` and `mlock()` system calls. Find out whether `collect` is running. If it is, kill `collect` and restart it with the `-l` (do not lock pages into memory) option.

3.6 LSM

This section discusses known problems with using the Logical Storage Manager (LSM) in a TruCluster Server cluster.

3.6.1 LSM Commands Fail When Disk Left in Failing State

When a node boots into an existing cluster and has connectivity to a failed device, it automatically brings the device online and reestablishes the associations with appropriate disk media records. After this process, the disk is occasionally left in the failing state, which prevents the disk from being used when space is requested by commands such as `volassist`.

If this situation occurs, you must manually turn off the disk's failing state, as follows:

```
# voledit set failing=off device_name
```

3.6.2 Problem Encapsulating Swap in Clusters with Long Host Names

LSM has a problem encapsulating swap in a cluster on members with base host names greater than 24 characters, for example, `reallyreallyreallyverylonghostname.foo.bar.com`. To work around this problem, reduce the base hostname to fewer than 25 characters.

3.6.3 Problem Encapsulating Swap When Remote Member Swap Areas Are Already Queued for Encapsulation

LSM has a problem encapsulating swap in a cluster when you run the `volreconfig` command on one member and another member already has swap devices queued for encapsulation.

To encapsulate the swap devices in a cluster, run `volencap` followed by `volreconfig` on each cluster member in turn. Only after `volreconfig` completes and the member has rebooted can you begin the swap encapsulation process for the next member. Do not run `volencap` or `volreconfig` on the next member until after the current member has rebooted.

For more information, see the chapter on using LSM in a cluster in the *Cluster Administration* manual.

3.6.4 The `voldisk list` Command Does Not Display All Non-LSM Disks in a Cluster

The `voldisk list` command is supposed to show disks that are configured both under LSM and not under LSM, which gives disks that are 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.

3.7 CAA

This section discusses known problems in the cluster application availability (CAA) subsystem.

3.7.1 Failure Threshold Value Should Not Be Greater Than 10

The failure threshold value for any resource should not be greater than 10. Setting the failure threshold value greater than 10 may cause the `caad` daemon to crash if there are a large number of failures for this resource and the failure threshold is never exceeded within the failure interval.

An application's failure threshold is set in its CAA application resource profile. For more information, see `caa_profile(8)`.

3.7.2 Problems with `caa_relocate` When Multiple Interdependent Applications Are Specified

A forced relocate that is not directed to a specific cluster member and has multiple interdependent applications specified will relocate all of these applications multiple times. For example, if `app1` depends on `app2` and `app3`, then `caa_relocate -f app1 app2 app3` causes all of these applications to relocate three times.

A forced relocate directed to a specific cluster member with interdependent applications specified relocates the applications to the specified cluster member correctly, but responds with an error. For example, `caa_relocate -c member2 -f app1 app2 app3` successfully relocates all applications to `member2`, but error messages are displayed and the return code is non-zero.

If you are relocating interdependent applications, specify only *one* application, for example, `caa_relocate -f app1` or `caa_relocate -c member2 -f app1`. All dependent applications relocate with the application specified.

Also, `caa_relocate -s` fails when it is used with resources with dependencies. To avoid this, do the following:

1. Identify applications with dependencies by using the `caa_stat -p` command. Those applications with entries for `REQUIRED_RESOURCES` have dependencies.
2. Use the `caa_relocate -f` command to relocate the applications with dependencies.
3. Use `caa_relocate -s member1 -c member2` to relocate the other applications.

3.7.3 Action Scripts PATH Considerations

Newly generated CAA scripts do not set the `PATH` environment variable. When they are executed, the `PATH` is set to a default value `/sbin:/usr/sbin:/usr/bin`. Therefore, you must explicitly specify most path names that are used in scripts, or you must modify the resulting scripts to explicitly set the `PATH`. Action scripts generated in previous releases may have a `PATH` that includes the current directory (`.`). Because this situation may be a potential security issue, modify these scripts to remove the current directory from the path.

3.7.4 The `caa_register -u` Command Does Not Correctly Update a Nonapplication Resource's State

When you change the `SUBNET` value in a network resource's profile or the `DEVICE_NAME` value in a tape or changer resource's profile and run the `caa_register -u` command to update the CAA registry, CAA does not update the `STATE` value for the resource. To correctly update a nonapplication resource, follow these steps:

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

3.7.5 Changes to `HOSTING_MEMBERS` List Can Leave Applications Running on Disallowed Members

A change to the `HOSTING_MEMBERS` list only affects future relocations and starts. If you update the `HOSTING_MEMBERS` list in the profile of an `ONLINE` application resource with a restricted placement policy, make sure that the application is running on one of the cluster members in that list. If the application is not running on one of the allowed members, run the `caa_relocate` on the application after running the `caa_register -u` command.

3.7.6 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 that the EVM viewer displays.

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

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

resource is ONLINE or unless 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 its states.

3.7.8 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 Rolling Upgrade

This section discusses issues with rolling upgrade.

3.8.1 Rolling Upgrade: Do Not Shut Down a Cluster Member To Single-User Mode

This note applies only when you perform a rolling upgrade from TruCluster Server Version 5.1 to Version 5.1A.

Note

Use the procedure given here to get to single-user mode. Do not follow the procedure given in the rolling upgrade section of the Version 5.1 TruCluster Server *Software Installation* manual.

Before halting a cluster member, make sure that the cluster can maintain quorum without the member's vote. For more information, see the section on shutting down and starting a cluster member in the *Cluster Administration* manual.

To take a cluster member to single-user mode, use the `shutdown -h` command to halt the member, and then boot the member to single-user mode. When the system reaches single-user mode, run the `init s`, `bcheckrc`, `kloadsrv`, and `lmf reset` commands. For example:

```
# /sbin/shutdown -h now
>>> boot -fl s
# /sbin/init s
# /sbin/bcheckrc
# /sbin/kloadsrv
# /usr/sbin/lmf reset
```

We recommend this halt and boot method because it ensures that the member will provide a minimal set of services to the cluster, and conversely, that the running cluster will have minimal reliance on the member in single-user mode.

Shutting a member down to single-user mode can take a long time, and in some cases, services such as automount, autofs, or NIS might appear to be working from the point of view of other cluster members, when in fact the services are either down or partially up.

3.8.2 Stop `prpasswd` Prior to Rolling Upgrade on Clusters Running Enhanced Security Environment

Before starting the rolling upgrade process in a TruCluster Enhanced Security environment, set a checkpoint in your authentication databases. Do this as follows:

```
# /usr/tcb/bin/db_checkpoint -h /var/tcb/files -1
```

After the checkpoint has been set, you must then shut down the `prpasswd` daemon on each cluster member. Do this by issuing the following command on each member:

```
# /sbin/init.d/prpasswd stop
```

Users can still log in to the cluster; however, login performance may suffer.

During the period that the cluster is rolling, disable the starting of the `prpasswd` daemon. Do this by setting false startup arguments in your `rc.config.common` database by using the following command on one cluster member:

```
# rcmgr -c set PRPASSWDD_ARGS "WAIT FOR COMPLETED ROLL"
```

As nodes roll to the new revision, a message indicating the `prpasswd` daemon did not start will be issued. This is expected.

After all the cluster members have rolled to the new release, restart the `prpasswd` daemon. Do this by issuing the following command on one member:

```
# rcmgr -c delete PRPASSWDD_ARGS
```

Then restart the `prpasswd` daemon on each node as follows:

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

3.8.3 Updating Worldwide Language Support from CD-ROM During a Rolling Upgrade

Do not remove the Worldwide Language Support (WLS) product from a cluster while a rolling upgrade is in progress; otherwise the cluster is put into a state that prevents the rolling upgrade from completing successfully.

If you remove the WLS product prior to starting the rolling upgrade, use the `/usr/sbin/clu_upgrade` command to determine whether a rolling upgrade is in progress. See `clu_upgrade(8)` and the *TruCluster Server Cluster Installation* manual for additional details on how to determine whether a rolling upgrade is in progress.

3.8.4 Run Updates to dop Database Only on Upgraded Members

When performing a cluster rolling upgrade, do not issue any `dop` commands (for example, `dop -a`, `dop -d`, or `dop -w`) on cluster members running the older version of the operating system. Any `dop` additions or deletions made on the members running the older operating system might be lost after `dop` commands are issued from the upgraded members.

3.8.5 Recovering from a Failed Install Stage During a Rolling Upgrade from V5.1 to V5.1A

If the install stage of a `clu_upgrade` fails, do the following:

1. Halt the lead member.
2. Execute the following command from any member still UP:

```
# zcat < /var/adm/update/TruClusterKit/TCRBASE520 | tar -xf - \  
./usr/sbin/cluster/clu_common \  
./usr/sbin/clu_upgrade \  
./usr/lib/nls/msg/en_US.ISO8859-1/clu_upgrade.cat  
# clu_upgrade -undo install
```

The following is displayed:

```
This is the cluster upgrade program.  
You have indicated that you want to undo the 'install' stage of the  
upgrade.  
  
Do you want to continue to undo this stage of the upgrade? [yes]: Return  
Restoring tagged files.  
.....  
clu_rollprop: "/usr/lib/nls/msg/en_US.ISO59-1/.Old..clu_upgrade.cat" \  
does not exist  
.....  
/usr/lib/nls/msg/en_US.ISO8859-1/.Old..clu_upgrade.cat does not exist  
.....  
clu_rollprop: "/usr/sbin/.Old..clu_upgrade" does not exist  
clu_rollprop: "/usr/sbin/cluster/.Old..clu_common" does not exist  
.....  
/usr/sbin/.Old..clu_upgrade does not exist
```

```
/usr/sbin/cluster/.Old..clu_common does not exist
.....
The undo of the 'install' stage completed successfully.
```

3. Boot the lead member as follows:

```
>>> boot -fl s
```

4. When the system reaches single-user mode, run the following commands:

```
# init s
# bcheckrc
# kloadsrv
# lmf reset
```

5. Re-run `installupdate` as documented in *Cluster Installation*.

3.9 Miscellaneous Administration

This section discusses issues with various administration tools that are used in a cluster.

3.9.1 RIS Boot Failures When Cluster is RIS Server

If the system that became the initial cluster member was configured as a RIS server before the `clu_create` command was run, then the cluster creation process does not update the `sa` entry in `/etc/bootptab`. The `sa` remains the IP address of the standalone system. Because of this, attempts at RIS boots after clusterization fail to mount the root file system.

You must manually edit `/etc/bootptab` and update the `sa` entry to be the IP address for the default cluster alias.

3.9.2 The `cfsstat` Command Can Return Incorrect Statistics

The `cfsstat` command can return incorrect statistics for the following:

- `tokens` (statistics on tokens)

The values for the total number of token requests and the total number of token cache hits are always 0 (zero).

- `tokstats` (statistics on tokens traffic)

Some statistics might be returned as negative numbers.

- `cfsmem` (statistics on memory usage by CFS)

Incorrect values are returned for `tokens_t` structures (`cli_tokens.tokens_t`) and `token` structures (`cli_tokens.token`).

In addition, when displaying very large numbers, the `cfsstat` command might truncate the values or display them as negative numbers. To avoid this, reset CFS statistics with the command `cfsstat -z` before gathering statistics.

3.9.3 All Tape Writes at Same Density and Compression

All writes to a tape, regardless of the device file that is used, will have the same density and compression. This density and compression may not correspond to the density and compression documented for the device file in `tape(7)`.

3.9.4 Tape Opens Take Long Time to Fail When No Tape Is Present

When an attempt is made to open a tape when no tape is present, more than 5 minutes can elapse before an error is returned.

3.9.5 Avoid Using Block Sizes Greater Than 512K Bytes for Tape Transfers

If the block size on a tape operation is greater than 512K bytes, the actual read or write operation is divided into 512K-byte transfers. If an error occurs during a transfer greater than 512K, no EIO or other error is returned. Instead, the number of bytes transferred is reported.

3.9.6 Disabling Internet Services on a Per-Member Basis

The *Cluster Administration* manual describes a way to disable the Internet services daemon (`inetd`) on a particular member by using the `disable` keyword in the `/etc/inetd.conf.local` file. For example:

```
finger stream tcp      nowait root    disable      fingerd
```

If you add this entry, the following error appears in `/var/adm/syslog.dated/current/daemon.log`:

```
inetd: disable: file does not exist
```

The Internet services daemon continues to run normally, and the service is not disabled. The workaround is to remove the entry for the service from the global `/etc/inetd.conf` file and add it to the member-specific `/etc/inetd.conf.local` file on the members that you want to offer the service. For more information on configuring `inetd`, see `inetd.conf(4)`.

3.9.7 The hwmgr -show comp Command May Report an Inconsistency Error When Creating a Clusterwide Name for a SCSI Device

After you have used the `hwmgr -edit scsi` command to create a clusterwide unique name for a SCSI device, a subsequent `hwmgr -show comp` command may report an inconsistency on the SCSI device. The inconsistency appears when the `hwmgr -edit scsi` command is invoked on the second and subsequent members for the same device. You can ignore the inconsistency error in this situation.

For example:

```
root> hwmgr -show comp -id 373 -full

HWID:  HOSTNAME  FLAGS  SERVICE  COMPONENT NAME
-----
373:   rovel-qa1  rcd-i  iomap    SCSI-WWID:ff10000b:"media_chngr"

DSF GROUP
INSTANCE GRPFLAGS  GROUPID  SUBSYSTEM  BASENAME  L1          L2
-----
0         40         81       cam_changer mc2        media_changer generic

DEVICE NODE
ID  LBdevT  LCdevT  CBdevT  CCdevT  BFlags  CFlags  Class  Suffix  L3B  L3
-----
0   0       56008c0  0       13003b3 0x0    0x861  0x0    (null) (null) (null)

COMPONENT INCONSISTENCY
-----
Component should not have an entry in the cluster database but it does.
```

3.9.8 The hwmgr -scan scsi Command Does Not Work Clusterwide

The `hwmgr -scan scsi` command `-cluster` option does not work clusterwide. When entered on a cluster member, the `hwmgr -scan scsi -cluster` command updates the device databases for only that member, not for all cluster members.

To perform a clusterwide scan, enter the following command on each cluster member when you need to update the member device databases clusterwide:

```
# hwmgr -scan comp -cat scsi_bus
```

You usually use this command when you add a new disk to a cluster (see Section 3.9.9).

3.9.9 Adding a Disk to a Running Cluster

When you add a new disk to a running cluster (for example, when you replace 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. For all disk models, enter the following command on each member to scan SCSI buses clusterwide and configure any new devices.

```
# hwmgr -scan comp -cat scsi_bus
```

Allow a minute or two for the scans to complete.

2. If the disk that you are adding is an RZ26, RZ28, RZ29, or RZ1CB-CA model, enter the following command on each cluster member after you install 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.

3.9.10 Running Process Accounting on Large Clusters Can Exhaust 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 with no parameters to disable accounting.

3.10 SysMan Menu

This section discusses known problems that you may encounter when you use SysMan Menu in a cluster.

3.10.1 Nonroot Users with No Home Directory 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, who runs 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.

3.11 SysMan Station

This section discusses known problems that you may encounter when using SysMan Station in a cluster.

3.11.1 SysMan Station Does Not Display Applications with Long Resource Names

CAA application resources with names of 64 or more characters are not displayed by SysMan Station. If you plan on managing your system with SysMan Station, limit the length of resource names to 63 characters. Resources with names beyond this limit can still be successfully managed with either SysMan Menu or the command line CAA commands.

3.11.2 Unable to Expand Host Object in a Cluster

The SysMan Station client may occasionally encounter a Java class exception error when you attempt to expand a Host object.

If you encounter this error, click on a different view and then reselect the Hardware view and retry the expand operation. If the problem persists, restart the SysMan Station client and retry the expand operation.

3.11.3 SysMan Station Might Display Cluster Status Incorrectly

The SysMan Station relies on events generated by the Event Manager (EVM) subsystem in order 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 situation 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 situation can happen when the TruCluster Server 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 that are 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
```

3.11.4 SysMan Station Might 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
```

3.11.5 Properties Might Not Be Displayed for Selected Objects

Properties may not be displayed for selected objects. The Properties dialog box may appear briefly on the screen or may not be displayed at all.

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

3.11.6 Some SysMan Station Applications Display Wrong Target Member Name

When the following applications are launched from the SysMan Station, their title bars 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.

3.12 Documentation

This section discusses TruCluster Server Version 5.1A documentation issues.

3.12.1 Recovering Cluster Root File System

Take the following into consideration when using the procedures for recovering the cluster root file system described in the chapter on troubleshooting clusters in the *Cluster Administration* manual:

- When booting the initial cluster member, you may need to adjust expected quorum votes.

In the *Cluster Administration* manual, see the section on forming a cluster when members do not have enough votes.

- The recovery process is not complete until the `h` partition of each member's boot disk is updated with the correct information about the devices used for the cluster root file system.

You can do this by booting each member.

You can also update the `h` partition of a member's boot disk with the `clu_bdmgr` command. For more information, see `clu_bdmgr(8)`.

3.12.2 Correct Size of Member Boot Disk a (Boot) Partition

In the *Cluster Administration* manual, the section on backing up and repairing a member's boot disk gives the wrong size for the `a` (boot) partition.

The size of the `a` partition is 256 MB.

3.12.3 Error in Example of Hexadecimal Format of IP Address

In the *Cluster Administration* manual, section 3.11, Enabling Cluster Alias vMAC Support, has a typographical error in the example of the hexadecimal format of an IP address. The following line is incorrect:

IP address in hex format: 10.8C.70:D1

The colon (:) in the IP address should be a period (.):

IP address in hex format: 10.8C.70.D1

3.12.4 Media Changer Utility, `mcutil`, Works on Remote Members

The *Cluster Administration* manual incorrectly states that the `mcutil` command works only on a device that is directly connected to the member where the command is executed.

The `mcutil` command works on devices clusterwide, regardless of the member on which the command is executed.

3.12.5 Using LSM to Mirror the Cluster Root File System

The section of the *Cluster Administration* manual on using Logical Storage Manager in a cluster is missing an example of how to use LSM to mirror the cluster root file system.

Use the `volmigrate` command to mirror `cluster_root`, the cluster root file system. Doing so moves `cluster_root` from its current volume to the LSM volume you specify in the `volmigrate` command.

To mirror `cluster_root` requires that the LSM `rootdg` disk group contains at least two disks of the appropriate size and type to create a volume for the cluster root domain. These disks must be accessible to all cluster members.

You can display the size of the cluster root domain as follows:

```
# showfdmn cluster_root
```

To learn the connectivity of disks, use the `hwmgr` command:

```
# hwmgr -view devices -cluster
```

If a disk is suitable for mirroring `cluster_root`, then the output of the `hwmgr` command will list the device special file name of the disk multiple times, once for each member of the cluster.

If necessary, you can add disks to `rootdg` with the `voldg` command:

```
# voldg adddisk disk_name
```

To create the mirrored volume, use the `volmigrate` command. For example, if you want to use `dsk5` and `dsk7` as LSM mirrors for the cluster root file system, use the following command:

```
# volmigrate -m 2 cluster_root dsk5 dsk7
```

For more information about mirroring the cluster root file system, see `volmigrate(8)`.

3.12.6 Migrating from Automount to AutoFS

This section describes how to migrate from Automount to AutoFS. The information in this section was not included in the *Cluster Administration* manual.

The `autofs` daemon automatically and transparently mounts and unmounts NFS file systems on an as-needed basis. Like the `automount` daemon, it provides another alternative to using the `/etc/fstab` file for mounting NFS file systems on client machines. However, AutoFS is more efficient than the `automount` daemon because it requires less communication between the kernel and the user space daemon. The

autofs daemon also provides higher availability than the automount daemon.

Three possible migration scenarios are presented: migrating without rebooting any cluster member, migrating when you reboot the active Automount server node, and migrating when rebooting the entire cluster.

Regardless of which approach you select, first use the CAA `caa_stat` command to verify that the `autofs` CAA resource is registered. For example,

```
# /usr/bin/caa_stat autofs
Could not find resource autofs.
```

If `autofs` is not registered as a CAA resource, then register it as follows:

```
# /usr/sbin/caa_register autofs
```

3.12.6.1 Migrating Without a Reboot

Migrating without rebooting any cluster member requires the largest number of procedural steps, but provides the highest availability. The additional steps are required to ensure cleanup of the Automount intercept points and to automatically start AutoFS.

Do the following:

1. Change the `rc.config.common` file.
 - a. Determine any arguments to pass to the `autofsmount` command. These arguments are typically a subset of those already specified by the `AUTOMOUNT_ARGS` environment variable. To view the value of that variable, use the `rcmgr -get` command, as shown in the following example:

```
# /usr/sbin/rcmgr -c get AUTOMOUNT_ARGS
-m /net -hosts -D MACH=alpha -D NET=f
```

The `-m` option ignores directory-mapname pairs listed in the `auto.master` file. This option might be useful for debugging purposes if you suspect there is a syntax error in the `auto.master` file.

Environment variables set by using the `-D` option resolve placeholders in the definition of auto-mount map file entries. For example, the associated `NET` entry might appear in the map file as follows:

```
vsx ${NET}system:/usr/projects2/vsx
```

and would resolve to

```
vsx fsystem:/usr/projects2/vsx
```

- b. Set the arguments to pass to the `autofs` command, as determined in the previous step. To do this, use the `rcmgr -set` command. For example:

```
# /usr/sbin/rcmgr -c set AUTOFSMOUNT_ARGS -m /net -hosts
-D MACH=alpha -D NET=f
```

- c. Set the arguments to pass to the `autofs` daemon. For example:

```
# /usr/sbin/rcmgr -c set AUTOFS_ARGS -D
MACH=alpha -D NET=f
```

These arguments must match the environment variables, specified with the `-D` option, as set for `AUTOMOUNT_ARGS`.

- d. Use the `mount -e` command to identify an automounted file system.

```
# mount -e | grep nfs
deli.zk3.dec.com:(pid524825) on /net type nfs (v2, ro,
nogrpid, udp, hard, intr, noac,timeo=350, retrans=5)
```

The automounted file system is indicated by `hostname:(PID)`.

- e. Determine which cluster member is the Automount server node for the NFS file system you identified in the previous step, as shown in the following example:

```
# cfsmgr -p /net
Domain or filesystem name = /net
Server Name = swiss
Server Status : OK
```

- f. Stop the Automount service on all cluster members other than the Automount server you identified in the previous step. To do this, use the `ps -ef` command to display process identifiers, search the output for instances of `automount`, and then use the `kill` command to kill each process. This causes the automount daemon to unmount all file systems that it has mounted, and to exit.

```
# ps -ef | grep automount
root      1049132 1048577  0.0   May 10 ??          0:00.00
/usr/sbin/automount -m /net -hosts
```

```
# kill 1049132
```

Note that, as of Tru64 UNIX Version 5.1A, the `kill` command is clusterwide and you can kill a process from any cluster member.

- g. Disable Automount and enable AutoFS in the `rc.config.common` file, as follows:

```
# /usr/sbin/rcmgr -c set AUTOMOUNT 0
# /usr/sbin/rcmgr -c set AUTOFS 1
```


2. Allow all auto-mounted file systems to become quiescent.
3. Stop the Automount service on the cluster member operating as the server. To do this, use the `ps -ef` command to display process identifiers, search the output for instances of `automount`, and then use the `kill` command to kill each process.

```
# ps -ef | grep automount
root      524825 524289  0.0   May 10 ??   0:00.01
/usr/sbin/automount -m /net -hosts
```

```
# kill 524825
```

4. Use the `mount -e` command and search the output for `tmp_mnt`, or the directory specified with the `automount -M` command, to verify that auto-mounted file systems are no longer mounted. No file systems should be mounted on `tmp_mnt`.

```
# mount -e | grep tmp_mnt
```

If some mount points still exist, they will no longer be usable via the expected pathnames. However, they are still usable under the full `/tmp_mnt/...` pathnames. Because AutoFS does not use the `/tmp_mnt` mount point, there is no conflict and the full auto-mount name space is available for AutoFS. If these `tmp_mnt` mount points later become idle, you can unmount them by using the `-f` option of the `umount` command, which unmounts remote file systems without notifying the server.

5. Start AutoFS. AutoFS provides automatic failover of the automounting service by means of CAA: one cluster member acts as the CFS server for auto-mounted file systems, and runs the one active copy of the AutoFS daemon. If this cluster member fails, CAA starts the `autofs` resource on another member.

If you do not care which node serves AutoFS, use the `/usr/sbin/caa_start autofs` command without specifying a cluster member; otherwise, use the `/usr/sbin/caa_start autofs -c member-name` command to specify the cluster member that you want to serve AutoFS.

```
# /usr/sbin/caa_start autofs
```

The `-c` option starts the `autofs` resource on the specified member if the cluster member is allowed by the placement policy and resource dependencies. If the cluster member specified is not allowed by the placement policy and resource dependencies, the `caa_start` command fails. If the specified member is not available, the command fails.

See the discussion of the resource file options in `caa_profile(8)`.

6. Use the `caa_stat autofs` command to make sure that the `autofs` resource started as expected.

```
# /usr/bin/caa_stat autofs
NAME=autofs
TYPE=application
TARGET=ONLINE
STATE=ONLINE on swiss
```

3.12.6.2 Migrating When Rebooting a Cluster Member

Migrating when rebooting a cluster member requires fewer procedural steps than migrating without a reboot, at the expense of availability.

Note

Before you shut down a cluster member, you need to determine whether the cluster member you are shutting down is a critical voting member, and whether it is the only hosting member for one or more applications with a restricted placement policy. Both of these issues are described in Chapter 5 of the *Cluster Administration* manual.

Follow these steps to migrate from Automount to AutoFS when rebooting a cluster member:

1. Change the `rc.config.common` file.
 - a. Determine any arguments to pass to the `autofsmount` command. These arguments are typically a subset of those already specified by the `AUTOMOUNT_ARGS` environment variable. To view the value of that variable, use the `rcmgr -get` command, as shown in the following example:

```
# /usr/sbin/rcmgr -c get AUTOMOUNT_ARGS
-m /net -hosts -D MACH=alpha -D NET=f
```

Environment variables set by using the `-D` option resolve placeholders in the definition of auto-mount map file entries. For example, the associated `NET` entry might appear in the map file as follows:

```
vsx ${NET}system:/usr/projects2/vsx
```

and would resolve to

```
vsx fsystem:/usr/projects2/vsx
```

- b. Set the arguments to pass to the `autofsmount` command, as determined in the previous step. To do this, use the `rcmgr -set` command. For example:

```
# /usr/sbin/rcmgr -c set AUTOFSMOUNT_ARGS -m /net -hosts
-D MACH=alpha -D NET=f
```

- c. Set the arguments to pass to the `autofs` daemon. For example:

```
# /usr/sbin/rcmgr -c set AUTOFS_ARGS -D
MACH=alpha -D NET=f
```

These arguments must match the environment variables, specified with the `-D` option, as set for `AUTOMOUNT_ARGS`.

- d. Use the `mount -e` command to identify a file system served by Automount:

```
# mount -e | grep nfs
deli.zk3.dec.com: (pid524825) on /net type nfs (v2, ro, nogrpid,
udp, hard, intr, noac,timeo=350, retrans=5)
```

The automounted file system is indicated by `hostname: (PID)`.

- e. Determine which cluster member is the Automount server node for the NFS file system you identified in the previous step.

```
# cfsmgr -p /net
Domain or filesystem name = /net
Server Name = swiss
Server Status : OK
```

- f. Stop the Automount service on all cluster members other than the Automount server you identified in the previous step. To do this, use the `ps -ef` command to display process identifiers, search the output for instances of `automount`, and then use the `kill` command to kill each process. This causes the automount daemon to unmount all file systems that it has mounted, and to exit.

```
# ps -ef | grep automount
root      1049132 1048577  0.0   May 10 ??  0:00.00
/usr/sbin/automount -m /net -hosts
```

```
# kill 1049132
```

Note that, as of Tru64 UNIX Version 5.1A, the `kill` command is clusterwide and you can kill a process from any cluster member.

- g. Disable Automount and enable AutoFS in the `rc.config.common` file, as follows:

```
# /usr/sbin/rcmgr -c set AUTOMOUNT 0
# /usr/sbin/rcmgr -c set AUTOFS 1
```

2. (Optional) Specify the AutoFS server. AutoFS provides automatic failover of the automounting service by means of CAA: one cluster member acts as the CFS server for auto-mounted file systems, and runs

the one active copy of the AutoFS daemon. If this cluster member fails, CAA starts the `autofs` resource on another member.

You can use the `caa_profile autofs -print` command to view the CAA hosting and placement policy, if any. The hosting policy specifies an ordered list of members, separated by white space, that can host the application resource. The placement policy specifies the policy according to which CAA selects the member on which to start or restart the application resource.

```
# /usr/sbin/caa_profile autofs -print
NAME=autofs
TYPE=application
ACTION_SCRIPT=autofs.scr
ACTIVE_PLACEMENT=0
AUTO_START=0
CHECK_INTERVAL=0
DESCRIPTION=Autofs Services
FAILOVER_DELAY=0
FAILURE_INTERVAL=0
FAILURE_THRESHOLD=0
HOSTING_MEMBERS=
OPTIONAL_RESOURCES=
PLACEMENT=balanced
REQUIRED_RESOURCES=
RESTART_ATTEMPTS=3
SCRIPT_TIMEOUT=3600
```

The default, and recommended, behavior is to run on any cluster member, with a placement policy of `balanced`. If this is not suitable for your environment, use the `/usr/sbin/caa_profile -update` command to change the `autofs` resource profile.

See the discussion of the resource file options in `caa_profile(8)`.

If you make a change, use the following command to have the update take effect:

```
# /usr/sbin/caa_register -u autofs
```

3. Reboot the cluster member. Before you shut down the cluster member, make sure that it is not a critical voting member or the only hosting member for one or more applications with a restricted placement policy.

When it reboots, Automount will no longer be running in the cluster, and AutoFS will start.

```
# /sbin/shutdown -r now
```

3.12.6.3 Migrating When Rebooting the Cluster

Migrating when rebooting the entire cluster requires fewer procedural steps than migrating without a reboot or migrating when rebooting a single member. The trade-off is a loss of cluster availability.

Rebooting the cluster is a drastic measure and is not the preferred migration method.

Follow these steps to migrate from Automount to AutoFS when rebooting the cluster:

1. Change the `rc.config.common` file.
 - a. Determine any arguments to pass to the `autofs` command. These arguments are typically a subset of those already specified by the `AUTOMOUNT_ARGS` environment variable. To view the value of that variable, use the `rcmgr -get` command, as shown in the following example:

```
# /usr/sbin/rcmgr -c get AUTOMOUNT_ARGS
-m /net -hosts -D MACH=alpha -D NET=f
```

Environment variables set by using the `-D` option resolve placeholders in the definition of auto-mount map file entries. For example, the associated `NET` entry might appear in the map file as follows:

```
vsx ${NET}system:/usr/projects2/vsx
```

and would resolve to

```
vsx fsystem:/usr/projects2/vsx
```

- b. Set the arguments to pass to the `autofs` command, as determined in the previous step. To do this, use the `rcmgr -set` command. For example:

```
# /usr/sbin/rcmgr -c set AUTOFSMOUNT_ARGS -m /net
-hosts -D MACH=alpha -D NET=f
```

- c. Set the arguments to pass to the `autofs` daemon. For example:

```
# /usr/sbin/rcmgr -c set AUTOFSD_ARGS -D MACH=alpha -D NET=f
```

These arguments must match the environment variables, specified with the `-D` option, as set for `AUTOMOUNT_ARGS`.

- d. Use the `mount -e` command to identify a file system served by Automount:

```
# mount -e | grep nfs
deli.zk3.dec.com:(pid524825) on /net type nfs (v2, ro,
noprpid, udp, hard, intr, noac,timeo=350, retrans=5)
```

The automounted file system is indicated by `hostname: (PID)`.

- e. Determine which cluster member is the Automount server node for the NFS file system you identified in the previous step.

```
# cfsmgr -p /net
Domain or filesystem name = /net
Server Name = swiss
Server Status : OK
```

- f. Disable Automount and enable AutoFS in the `rc.config.common` file, as follows:

```
# /usr/sbin/rcmgr -c set AUTOMOUNT 0
# /usr/sbin/rcmgr -c set AUTOFS 1
```

2. (Optional) Specify the AutoFS server. AutoFS provides automatic failover of the automounting service by means of CAA: one cluster member acts as the CFS server for auto-mounted file systems, and runs the one active copy of the AutoFS daemon. If this cluster member fails, CAA starts the `autofs` resource on another member.

You can use the `/usr/bin/caa_profile autofs -print` command to view the CAA hosting and placement policy, if any. The hosting policy specifies an ordered list of members, separated by white space, that can host the application resource. The placement policy specifies the policy according to which CAA selects the member on which to start or restart the application resource.

```
# /usr/bin/caa_profile autofs -print
NAME=autofs
TYPE=application
ACTION_SCRIPT=autofs.scr
ACTIVE_PLACEMENT=0
AUTO_START=0
CHECK_INTERVAL=0
DESCRIPTION=Autofs Services
FAILOVER_DELAY=0
FAILURE_INTERVAL=0
FAILURE_THRESHOLD=0
HOSTING_MEMBERS=
OPTIONAL_RESOURCES=
PLACEMENT=balanced
REQUIRED_RESOURCES=
RESTART_ATTEMPTS=3
SCRIPT_TIMEOUT=3600
```

The default, and recommended, behavior is to run on any cluster member, with a placement policy of `balanced`. If this is not suitable for your environment, use the `/usr/bin/caa_profile -update` command to change the `autofs` resource profile.

See the discussion of the resource file options in `caa_profile(8)`.

If you make a change, use the following command to have the update take effect:

```
# /usr/sbin/caa_register -u autofs
```

3. Reboot the cluster. When it reboots, Automount will no longer be running in the cluster, and AutoFS will start.

```
# /sbin/shutdown -c now
```

3.12.7 Clusterwide IPC--Supported Mechanisms

This section describes how to migrate from Automount to AutoFS. The information in this section was not included in the *Cluster Administration* manual.

The following mechanisms for clusterwide IPC are supported:

- TCP/IP with sockets
- Memory Channel API:
 - memory windows
 - low level locks
 - signals
- Files:
 - Buffered I/O or memory mapped
 - UNIX API file locks
- Distributed Lock Manager Locks
- Clusterwide kill signals

The following mechanisms are not supported for clusterwide IPC:

- UNIX domain sockets
- Named pipes (FIFO special files)
- Signals
- System V IPC (messages, shared memory, semaphores)

3.12.8 Software Product Description (SPD) Replaced by QuickSpec

For TruCluster Server Version 5.1A, the SPD has been replaced by the TruCluster Server *QuickSpec*. For a description of TruCluster Server Version 5.1A and information about its capabilities and the hardware it supports, see the *QuickSpec*. You can find it online at http://www.tru64unix.compaq.com/docs/pub_page/spds.html,

and in the TruCluster/DOCS directory on the Associated Products Volume 2 CD-ROM.

Index

A

accounting

exhausting process quotas in large clusters, 3–25

adding a disk, 3–24

addvol

member root domain, 2–2

AdvFS

restrictions, 2–2

AlphaServer GS80/160/320

booting, 3–5

application resources

no Sysman Station display of resources with long names, 3–26

ATTENTION message while

booting, 3–7

B

backup

supported tape densities, 3–23

bcheckrc command, 3–19

boot partition

size of on member boot disk, 3–29

booting

adjusting quorum votes during

disaster recovery, 3–8

ATTENTION message during, 3–7

CNX message during, 3–7

hang during, 3–6

ict_mct messages during, 3–8

member on cluster with KZPBA

SCSI bus adapter, 3–6

RIS boot failures, 3–22

C

CAA

action scripts PATH considerations, 3–17

caa_relocate -f problems with multiple interdependent applications, 3–17

consequences of changing HOSTING_MEMBERS list, 3–18

failure threshold should not exceed 10, 3–16

incorrect updates of a nonapplication resource's state, 3–18

malformed event messages by EVM viewer, 3–18

no Sysman Station display of applications with long resource names, 3–26

updating a running application's required resources, 3–18

caa_register command

incorrect update of nonapplication resource's state, 3–18

updating a running application's required resources, 3–18

CD-ROM file system

(See CDFS)

CDFS

read-only access supported, 2–1

CFS

performance improvements, 1–1t

relocation failure, 3–15

restrictions on, 2–1

cfsmgr command
failure to relocate file systems
with memory-wired applications,
3-15

cfsstat command
incorrect statistics, 3-22

changes
cluster interconnect virtual device
name, 1-2

cloning filesets
root, 2-2

clu_add_member
NetRAIN configuration, 3-3

clu_check_config command
ATTENTION message from, 3-7

clu_create and clu_add_member
default physical cluster interconnect
IP name, 3-2

clu_create command, 3-2
incorrect Help message, 3-3

clu_delete_member command
deleting member with a boot disk
with a bad disk label, 3-14

cluster alias
IPv6 not supported for, 2-3
support for alternate cluster alias
with NFS, 1-1t
support for NetRAIN and link
aggregation, 2-3

cluster application availability
(See CAA)

cluster configuration check
ATTENTION message from, 3-7

Cluster File System
(See CFS)

cluster interconnect address
IPv6 not supported for, 2-3

**cluster interconnect virtual
device**
name change, 1-2

cluster recovery
additional information, 3-29

cluster root

mirroring cluster root with LSM,
3-30

cluster_root
restrictions on filesets in, 2-2

CNX message while booting, 3-7

configuring DHCP, 3-13

D

**deleting a member with
inaccessible boot disk**
expected votes not adjusted by
clu_delete_member command,
3-14

density
supported tape densities, 3-23

DHCP
configuring, 3-13

disabling Internet services, 3-23

disaster recovery
booting the cluster, 3-8

disk
left in failing state, 3-15
replacing a failed disk, 3-24

documentation
errors in, 3-28

dop command restriction, 3-21

**Dynamic Host Configuration
Protocol**
(See DHCP)

E

encapsulating swap
restrictions on, 3-16

Enhanced Security environment
stop prpasswd prior to rolling
upgrade, 3-20

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

/etc/sysconfigtab
NetRAIN configuration, 3-3

Ethernet hubs

configurations that support hubs,
3-9
support for LAN interconnect, 3-3

Ethernet switch

set address aging to 15 seconds,
3-10

Event Manager

(See EVM)

EVM

malformed CAA event messages
displayed by EVM viewer, 3-18

F

failure threshold

CAA failure threshold should not
exceed 10, 3-16

features, 1-1

FFM file system

not supported, 2-2

File-on-File Mount

(See FFM file system)

fileset cloning

root, 2-2

firmware

updating, 3-1

G

GS80/160/320

booting, 3-5

H

hangs

booting member hangs while setting
time and date, 3-6
booting member on cluster with
KZPBA SCSI bus adapter, 3-6
do not reboot all members
simultaneously, 3-4

insufficient Memory Channel
address space, 3-8
member hangs during clusterwide
shutdown, 3-7
while booting AlphaServer
GS80/160/320, 3-5

hardware

restrictions, 2-1

HOSTING_MEMBERS

consequences of changing CAA
HOSTING_MEMBERS list, 3-18

hostname

duplicate hostnames, 3-2
failure to add to /etc/cfgmgr.auth
file, 3-3

hwmgr command

-cluster option does not work
clusterwide, 3-24
-show comp option may report an
inconsistency error when creating
clusterwide names for SCSI
devices, 3-24

I

icssvr_daemon_from_pool

processes, 3-25

ifaccess.conf

updating, 3-12

inetd

disabling Internet services, 3-23

init -s command, 3-19

installation

do not use SysMan Menu to
perform, 3-2
update firmware, 3-1

Internet Protocol version 6

(See IPv6)

Internet services

disabling, 3-23

interprocess communication

(See IPC)

IP address
documentation error in hexadecimal
format of, 3–29

IPC, 3–39

IPv6
not supported for cluster alias, 2–3

K

**kernel build with no kernel
layered products**
rebuild /usr/sys/conf/product.list,
3–4

**kernel configuration management
database**
adding first member to, 3–3

kloadsrv command, 3–19

kreg command
run after build kernel with no
kernel layered products, 3–4

KZPBA-CB SCSI bus adapter
cluster member may hang on boot,
3–6

L

LAN
updating ifaccess.conf, 3–12

LAN interconnect
configurations that support
Ethernet hubs, 3–9
restrictions on multiple network
interfaces, 2–3
set switch address aging to 15
seconds, 3–10
supported hubs, 3–3

layered products registration file
rebuild after building kernel with
no kernel layered products, 3–4

license PAK
run lmf reset after entering license
information, 3–14

lmf reset command, 3–19

Logical Storage Manager

(*See* LSM)

long host names
restrictions on, 3–16

LSM
disk left in failing state, 3–15
mirroring cluster root, 3–30
voldisk list command does not
display all non-LSM disks within
a cluster, 3–16

M

mcutil
works on remote members, 3–29

media changer utility
(*See* mcutil)

member boot disk
adding volumes to a member root
domain, 2–2
size of a (boot) partition, 3–29

member root domain
addvol command, 2–2

Memory Channel
errors when booting a member into
a cluster, 3–8

Memory Channel 2
restrictions, 2–1

Memory File System
(*See* MFS)

messages
ATTENTION from boottime cluster
check, 3–7
C30niscluster: Cannot make
specified change, 3–8
CNX invalid configuration for seq
disk, 3–7
Error: Bad configuration, 3–2
ict_mct, 3–8
Server Relocation Failed from
cfsmgr command, 3–15
System V IPC not supported for
clusterwide IPC, 3–39

MFS, 2–1

mirroring cluster root, 3–30

mlock system call
CFS relocation problems, 3–15

N

named pipes
not supported, 2–2
not supported for clusterwide IPC,
3–39

NetRAIN
clu_add_member -c command, 3–3
cluster alias support for, 2–3

Network File System
(See NFS)

network interfaces
update ifaccess.conf, 3–11
update ifaccess.conf when changing,
3–12

new features, 1–1

newfs command, 3–15

NFS
failover of client file system not
supported, 2–1
support for alternate cluster alias,
1–1t

ntpdate
booting member hangs while setting
time and date, 3–6

P

PAK for TruCluster Server license
run lmf reset after entering license
information, 3–14

plock system call
CFS relocation problems, 3–15

/proc file system
not supported, 2–2

process accounting
exhausting process quotas in large
clusters, 3–25

product.product.list file

rebuild after building kernel with
no kernel layered products, 3–4

prpasswd
stop prpasswd prior to rolling
upgrade, 3–20

Q

QuickSpec
replaces Software Product
Description, 3–39

quorum disk
do not use for data, 3–15

quorum votes
adjusting when booting a member,
3–8

R

rebooting
AlphaServer GS80/160/320, 3–5
do not reboot all members
simultaneously, 3–4
member on cluster with KZPBA
SCSI bus adapter, 3–6

recovery
additional information, 3–29

**Redundant Array of Independent
Network Adapters**
(See NetRAIN)

remote installation services
(See RIS)

replacing failed disks, 3–24

restrictions
hardware restrictions, 2–1
IPv6 not supported for cluster alias
nor cluster interconnect address,
2–3
Memory Channel 2, 2–1

RIS
boot failures, 3–22

rolling upgrade

prpasswdd on Clusters and
Enhanced Security environment,
3-20
recovering from a failed install
stage, 3-21
run dop only on rolled members,
3-21
updating of WLS from CD-ROM,
3-21

root file system
recovering, 3-29

S

semaphores
System V IPC not supported for
clusterwide IPC, 3-39

Server Relocation Failed message
from cfsmgr command, 3-15

shared memory
System V IPC not supported for
clusterwide IPC, 3-39

shutdown
member hangs during, 3-7

signals
not supported for clusterwide IPC,
3-39

sockets
UNIX domain sockets not supported
for clusterwide IPC, 3-39

Software Product Description
(See SPD)

SPD
replaced by QuickSpec, 3-39

statistics
cfsstat can return incorrect
statistics, 3-22

storage backup
supported tape densities, 3-23

SysMan Menu
do not use Installation branch of
Software menu of, 3-2
nonroot users cannot perform
management tasks, 3-25

SysMan Station

applications display wrong target
member name, 3-28
Filesystem Attention group may not
be updated for nonoptimal file
system states, 3-27
may display cluster status
incorrectly, 3-26
may display new hardware objects
incorrectly, 3-28
no display of application resources
with long names, 3-26
properties may not display for
selected objects, 3-28
shows CAA application resources in
UNKNOWN state as having an
error, 3-19
unable to expand Host object in a
cluster, 3-26

System V IPC
not supported for clusterwide IPC,
3-39

T

tape
long failure time when no tape is
present, 3-23
recommendation on transfer block
size, 3-23
supported densities, 3-23

U

UFS, 2-1

UNIX domain sockets
not supported for clusterwide IPC,
3-39

UNIX File System
(See UFS)

update
ifaccess.conf and changes to
network interfaces, 3-12

ifaccess.conf on LAN clusters, 3–11
update firmware, 3–1
/usr/sys/conf/product.list
rebuild after building kernel with
no kernel layered products, 3–4

V

**virtual device name for cluster
interconnect**
name change, 1–2
voldisk command
list command does not display all
non-LSM disks within a cluster,
3–16

votes

deleting member with a boot disk
with a bad disk label, 3–14

W

wired memory

CFS relocation problems, 3–15

X

xjoin

configuring DHCP, 3–13