

PRIMECLUSTER™

Cluster Foundation (CF)

Configuration and Administration Guide (Solaris)

Comments... Suggestions... Corrections...

The User Documentation Department would like to know your opinion of this manual. Your feedback helps us optimize our documentation to suit your individual needs.

Fax forms for sending us your comments are included in the back of the manual.

There you will also find the addresses of the relevant User Documentation Department.

Certified documentation according DIN EN ISO 9001:2000

To ensure a consistently high quality standard and user-friendliness, this documentation was created to meet the regulations of a quality management system which complies with the requirements of the standard DIN EN ISO 9001:2000.

cognitas. Gesellschaft für Technik-Dokumentation mbH
www.cognitas.de

Copyright and Trademarks

Copyright © 2002, 2003 Fujitsu Siemens Computers inc. and Fujitsu LIMITED.

All rights reserved.

Delivery subject to availability; right of technical modifications reserved.

All hardware and software names used are trademarks of their respective manufacturers.

This manual is printed on
paper treated with
chlorine-free bleach.

Preface

Cluster Foundation

CF Registry and Integrity Monitor

Cluster resource management

GUI administration

LEFTCLUSTER state

CF topology table

Shutdown Facility

System console

CF over IP

Diagnostics and troubleshooting

CF messages and codes

Manual pages

Glossary

Abbreviations

Figures

Tables

Index

Contents

1	Preface	1
1.1	Contents of this manual	1
1.2	Documentation	2
1.2.1	Suggested documentation	3
1.3	Conventions	4
1.3.1	Notation	4
1.3.1.1	Prompts	4
1.3.1.2	The keyboard	4
1.3.1.3	Typefaces	4
1.3.1.4	Example 1	5
1.3.1.5	Example 2	5
1.3.2	Command syntax	5
1.4	Notation symbols	6
2	Cluster Foundation	7
2.1	CF, CIP, and CIM configuration	7
2.1.1	CIP versus CF over IP	11
2.1.2	cfset	13
2.1.3	CF security	15
2.1.4	An example of creating a cluster	16
2.2	CIP configuration file	38
2.3	Cluster Configuration Backup and Restore (CCBR)	40
3	CF Registry and Integrity Monitor	49
3.1	CF Registry	49
3.2	Cluster Integrity Monitor	50
3.2.1	Configuring CIM	50
3.2.2	Query of the quorum state	51
3.2.3	Reconfiguring quorum	51
4	Cluster resource management	55
4.1	Overview	55
4.2	Kernel parameters for Resource Database	56
4.3	Resource Database configuration	59
4.4	Registering hardware information	61
4.4.1	Prerequisite for EMC Symmetrix	62
4.4.2	Multi-path automatic generation	63
4.4.3	Automatic resource registration	64
4.5	Start up synchronization	65
4.5.1	Start up synchronization and the new node	67
4.6	Adding a new node	67

Contents

4.6.1	Backing up the Resource Database	68
4.6.2	Reconfiguring the Resource Database	69
4.6.3	Configuring the Resource Database on the new node	71
4.6.4	Adjusting StartingWaitTime	72
4.6.5	Restoring the Resource Database	72
5	GUI administration	75
5.1	Overview	75
5.2	Starting Cluster Admin GUI and logging in	76
5.3	Main CF table	80
5.4	Node details	81
5.5	Displaying the topology table	82
5.6	Starting and stopping CF	83
5.7	Marking nodes DOWN	86
5.8	Using CF log viewer	87
5.8.1	Search based on time filter	90
5.8.2	Search based on keyword	91
5.8.3	Search based on severity levels	92
5.9	Displaying statistics	93
5.10	Adding and removing a node from CIM	98
5.11	Unconfigure CF	100
5.12	CIM Override	101
6	LEFTCLUSTER state	103
6.1	Description of the LEFTCLUSTER state	104
6.2	Recovering from LEFTCLUSTER	106
6.2.1	Caused by a panic/hung node	106
6.2.2	Caused by staying in the kernel debugger too long	106
6.2.3	Caused by a cluster partition	107
6.2.4	Caused by reboot	109
7	CF topology table	111
7.1	Basic layout	113
7.2	Selecting devices	114
7.3	Examples	115
8	Shutdown Facility	119
8.1	Overview	119
8.2	Available Shutdown Agents	121
8.2.1	RCI	122
8.2.2	SCON	123
8.2.3	RCCU	123
8.2.4	RPS	124
8.2.5	NPS	124

8.3	SF split-brain handling	125
8.3.1	Administrative LAN	126
8.3.2	Overview of split-brain handling	126
8.3.3	Runtime processing	127
8.3.4	Split-brain resolution manager selection	128
8.3.5	Configuration notes	128
8.4	Configuring the Shutdown Facility	131
8.4.1	Invoking the Configuration Wizard	131
8.4.2	Configuration via CLI	151
8.4.2.1	Shutdown daemon	151
8.4.2.2	Shutdown Agents	153
8.5	SF facility administration	158
8.5.1	Starting and stopping SF	158
8.5.1.1	Starting and stopping SF manually	158
8.5.1.2	Starting and stopping SF automatically	158
8.6	Logging	159
9	System console	161
9.1	Overview	161
9.1.1	Role of the cluster console	161
9.1.2	Platforms	162
9.2	Topologies	163
9.2.1	Single cluster console	163
9.2.2	Distributed cluster console	164
9.3	Network considerations	165
9.4	Configuration on the cluster console	165
9.4.1	Updating the /etc/hosts file	165
9.4.2	Running the Configure script	166
9.4.3	Editing the rmshosts file	167
9.4.4	Additional steps for distributed cluster console	167
9.4.5	rmshosts method file	168
9.5	Updating a configuration on the cluster console	168
9.6	Configuration on the cluster nodes	169
9.6.1	Manually configuring the SCON SA	169
9.6.2	Configuration of the Shutdown Facility	169
9.6.3	Other configuration of the cluster nodes	169
9.6.3.1	Redirecting console input/output	169
9.6.3.2	Bootng with kadb	170
9.7	Using the cluster console	170
9.7.1	Without XSCON	171
9.7.2	With XSCON	171

Contents

10	CF over IP	173
10.1	Overview	173
10.2	Configuring CF over IP	175
11	Diagnostics and troubleshooting	177
11.1	Beginning the process	177
11.2	Symptoms and solutions	181
11.2.1	Join-related problems	182
11.3	Collecting troubleshooting information	191
11.3.1	Executing the fjsnap command	191
11.3.2	System dump	192
11.3.3	SCF dump	193
12	CF messages and codes	195
12.1	cfconfig messages	196
12.1.1	Usage message	196
12.1.2	Error messages	197
12.2	cipconfig messages	204
12.2.1	Usage message	204
12.2.2	Error messages	205
12.3	cftool messages	206
12.3.1	Usage message	207
12.3.2	Error messages	208
12.4	rcqconfig messages	211
12.4.1	Usage message	211
12.4.2	Error messages	211
12.5	rcquery messages	223
12.5.1	Usage message	223
12.5.2	Error messages	223
12.6	CF runtime messages	224
12.6.1	Alphabetical list of messages	225
12.7	CF Reason Code table	229
12.8	Error messages for different systems	240
12.9	Solaris/Linux ERRNO table	241
12.10	Resource Database messages	257
12.10.1	HALT messages	258
12.10.2	INFO messages	259
12.10.3	WARNING messages	260
12.10.4	ERROR messages	261
12.11	Shutdown Facility	288
12.12	Monitoring Agent messages	294
12.12.1	INFO message	295
12.12.2	WARNING message	296
12.12.3	ERROR message	296

13	Manual pages	301
13.1	CCBR	301
13.2	CF	301
13.3	CFS	302
13.4	CIP	302
13.5	Monitoring Agent	303
13.6	PAS	303
13.7	RCVM	303
13.8	Resource Database	304
13.9	RMS	305
13.10	RMS Wizards	306
13.11	SCON	306
13.12	SF	306
13.13	SIS	307
13.14	Web-Based Admin View	307
Glossary		309
Abbreviations		325
Figures		329
Tables		333
Index		335

Contents

1 Preface

The Cluster Foundation (CF) provides a comprehensive base of services that user applications and other PRIMECLUSTER services need to administrate and communicate in a cluster. These services include the following:

- Internode communications
- Node state management
- Cluster-wide configuration information
- Management and administration
- Distributed lock management

In addition, the foundation provides the following optional services:

- RCFS is a cluster-wide file share service
- RCVM is a cluster-wide volume management service

This document assumes that the reader is familiar with the contents of the *Concepts Guide* and that the PRIMECLUSTER software has been installed as described in the *Installation Guide*.

1.1 Contents of this manual

This manual contains the configuration and administration information for the PRIMECLUSTER components. This manual is organized as follows:

- The Chapter “Cluster Foundation” describes the administration and configuration of the Cluster Foundation.
- The Chapter “CF Registry and Integrity Monitor” discusses the purpose and physical characteristics of the CF synchronized registry, and it discusses the purpose and implementation of CIM.
- The Chapter “Cluster resource management” discusses the database which is a synchronized clusterwide database holding information specific to several PRIMECLUSTER products.
- The Chapter “GUI administration” describes the administration features in the CF portion of the Cluster Admin graphical user interface (GUI).

- The Chapter “LEFTCLUSTER state” discusses the LEFTCLUSTER state, describes this state in relation to the other states, and discusses the different ways a LEFTCLUSTER state is caused.
- The Chapter “CF topology table” discusses the CF topology table as it relates to the CF portion of the Cluster Admin GUI.
- The Chapter “Shutdown Facility” describes the components and advantages of PRIMECLUSTER SF and provides administration information.
- The Chapter “System console” discusses the SCON product functionality and configuration. The SCON product is installed on the cluster console.
- The Chapter “CF over IP” discusses CF communications based on the use of interconnects.
- The Chapter “Diagnostics and troubleshooting” provides help for troubleshooting and problem resolution for PRIMECLUSTER Cluster Foundation.
- The Chapter “CF messages and codes” provides a listing of messages and codes.
- The Chapter “Manual pages” lists the manual pages for PRIMECLUSTER.

1.2 Documentation

The documentation listed in this section contains information relevant to PRIMECLUSTER and can be ordered through your sales representative.

In addition to this manual, the following manuals are also available for PRIMECLUSTER:

- *Installation Guide (Solaris)*—Provides instructions for installing PRIMECLUSTER.
- *Concepts Guide (Solaris, Linux)*—Provides conceptual details on the PRIMECLUSTER family of products.
- *Reliant Monitor Services (RMS) Configuration and Administration Guide (Solaris)*—Provides instructions for configuring and administering RMS.
- *Scalable Internet Services (SIS) Configuration and Administration Guide (Solaris, Linux)*—Provides information on configuring and administering Scalable Internet Services.

- *Global Disk Services Configuration and Administration Guide (Solaris)*—Provides information on configuring and administering Global Disk Services (GDS).
- *Global File Services Configuration and Administration Guide (Solaris)*—Provides information on configuring and administering Global File Services (GFS).
- *Global Link Services Configuration and Administration Guide: Redundant Line Control Function (Solaris)*—Provides information on configuring and administering the redundant line control function for Global Link Services (GLS).
- *Global Link Services Configuration and Administration Guide: Multipath Function (Solaris)*—Provides information on configuring and administering the multipath function for Global Link Services (GLS).
- *Web-Based Admin View Operation Guide (Solaris)*—Provides information on using the Web-Based Admin View management GUI.
- *SNMP Reference Manual (Solaris)*—Provides reference information on the Simple Network Management Protocol (SNMP) product.
- Release notices for all products—These documentation files are included as html files on the PRIMECLUSTER Framework CD. Release notices provide late-breaking information about installation, configuration, and operations for PRIMECLUSTER. Read this information first.
- *RMS Wizards documentation package*—Available on the PRIMECLUSTER CD. These documents deal with topics like the configuration of file systems and IP addresses, or the different kinds of wizards.

1.2.1 Suggested documentation

The following manuals contain relevant information and can be ordered through your sales representative (not available in all areas):

- *ANSI C Programmer's Guide*
- *LAN Console Installation, Operation and Maintenance*
- *Terminal TM100/TM10 Operating Manual*
- *PRIMEPOWER User's Manual* (operating manual)



Your sales representative will need your operating system release and product version to place your order.

1.3 Conventions

In order to standardize the presentation of material, this manual uses a number of notational, typographical, and syntactical conventions.

1.3.1 Notation

This manual uses the following notational conventions.

1.3.1.1 Prompts

Command line examples that require system administrator (or root) privileges to execute are preceded by the system administrator prompt, the hash sign (#). In some examples, the notation `node#` indicates a root prompt on the specified node. For example, a command preceded by `fuji2#` would mean that the command was run as user `root` on the node named `fuji2`. Entries that do not require system administrator rights are preceded by a dollar sign (\$).

1.3.1.2 The keyboard

Keystrokes that represent nonprintable characters are displayed as key icons such as `[Enter]` or `[F1]`. For example, `[Enter]` means press the key labeled *Enter*; `[Ctrl-b]` means hold down the key labeled *Ctrl* or *Control* and then press the `[B]` key.

1.3.1.3 Typefaces

The following typefaces highlight specific elements in this manual.

Typeface	Usage
Constant Width	Computer output and program listings; commands, file names, manual page names and other literal programming elements in the main body of text.
<i>Italic</i>	Variables that you must replace with an actual value. Items or buttons in a GUI screen.
Bold	Items in a command line that you must type exactly as shown.

Typeface conventions are shown in the following examples.

1.3.1.4 Example 1

Several entries from an `/etc/passwd` file are shown below:

```
root:x:0:1:0000-Admin(0000):/:/sbin/sh
sysadm:x:0:0:System Admin.:/usr/admin:/usr/sbin/sysadm
setup:x:0:0:System Setup:/usr/admin:/usr/sbin/setup
daemon:x:1:1:0000-Admin(0000):/:
```

1.3.1.5 Example 2

To use the `cat` command to display the contents of a file, enter the following command line:

```
$ cat file
```

1.3.2 Command syntax

The command syntax observes the following conventions.

Symbol	Name	Meaning
[]	Brackets	Enclose an optional item.
{ }	Braces	Enclose two or more items of which only one is used. The items are separated from each other by a vertical bar ().
	Vertical bar	When enclosed in braces, it separates items of which only one is used. When not enclosed in braces, it is a literal element indicating that the output of one program is piped to the input of another.
()	Parentheses	Enclose items that must be grouped together when repeated.
...	Ellipsis	Signifies an item that may be repeated. If a group of items can be repeated, the group is enclosed in parentheses.

1.4 Notation symbols

Material of particular interest is preceded by the following symbols in this manual:



Contains important information about the subject at hand.



Caution

Indicates a situation that can cause harm to data.

2 Cluster Foundation

This chapter describes the administration and configuration of the Cluster Foundation (CF).

This chapter discusses the following:

- The Section “CF, CIP, and CIM configuration” describes CF, Cluster Internet Protocol (CIP) and Cluster Integrity Monitor (CIM) configuration that must be done prior to other cluster services.
- The Section “CIP configuration file” describes the format of the CIP configuration file.
- The Section “Cluster Configuration Backup and Restore (CCBR)” details a method to save and restore PRIMECLUSTER configuration information.

2.1 CF, CIP, and CIM configuration

CF configuration must be done before any other cluster services, such as Reliant Monitor Services (RMS) or Scalable Internet Services (SIS) configurations.

CF defines which nodes are in a given cluster. After CF configuration, SIS may be run on those nodes. After CF and CIP configuration is done, the nodes are ready for the Shutdown Facility (SF) and RMS to run on them.

Starting with PRIMECLUSTER, RMS is not responsible for node elimination. This is the responsibility of the Shutdown Facility (SF). This means that even if RMS is not installed or running in the cluster, missing CF heartbeats will cause node elimination by means of SF.

The CF Wizard in the Cluster Admin can be used to easily configure CF, CIP, and CIM for all nodes in the cluster. The SF Wizard in Cluster Admin may be used to configure SF.

A CF configuration consists of the following main attributes:

- Cluster name—This may be any name that you choose as long as it is 31 characters or less per name and each character comes from the set of printable ASCII characters, excluding white space, newline, and tab characters. Cluster names are always mapped to upper case.

- Set of interfaces on each node in the cluster used for CF networking—For example, the interface of an IP address on the local node may be an Ethernet device.
- CF node name—By default, in Cluster Admin, the CF node names are the same as the Web-Based Admin View names; however, you can use the CF Wizard to change them.

The dedicated network connections used by CF are known as interconnects. They typically consist of some form of high speed networking such as 100 MB or Gigabit Ethernet links. There are a number of special requirements that these interconnects must meet if they are to be used for CF:

1. The network links used for interconnects must have low latency and low error rates. This is required by the CF protocol. Private switches and hubs will meet this requirement. Public networks, bridges, and switches shared with other devices may not necessarily meet these requirements, and their use is not recommended.

It is recommended that each CF interface be connected to its own private network with each interconnect on its own switch or hub.

2. The interconnects should not be used on any network that might experience network outages of 5 seconds or more. A network outage of 10 seconds will, by default, cause a route to be marked as `DOWN`. `cfset(1M)` can be used to change the 10 second default. See Section “`cfset`”.

Since CF automatically attempts to bring up downed interconnects, the problem with split clusters only occurs if all interconnects experience a 10-second outage simultaneously. Nevertheless, CF expects highly reliable interconnects.

CF may also be run over IP. Any IP interface on the node may be chosen as an IP device, and CF will treat this device much as it does an Ethernet device. However, all the IP addresses for all the cluster nodes on that interconnect must be on the same IP subnetwork, and their IP broadcast addresses must be the same (refer to the Chapter “CF over IP” for more information).

The IP interfaces used by CF must be completely configured by the System Administrator before they are used by CF. You may run CF over both Ethernet devices and IP devices.

Higher level services, such as RMS, SF, GFS, and so forth, will not notice any difference when CF is run over IP.

You should carefully choose the number of interconnects you want in the cluster before you start the configuration process. If you decide to change the number of interconnects after you have configured CF across the cluster, you will need to bring down CF on each node to do the reconfiguration. Bringing down CF requires that higher level services, like RMS, SF, SIS and applications, be stopped on that node, so the reconfiguration process is neither trivial nor unobtrusive.



Your configuration should specify at least two interconnects to avoid a single point of failure in the cluster.

Before you begin the CF configuration process, you should make sure that all nodes are connected to the interconnects you have chosen and that all of the nodes can communicate with each other over those interconnects. The CF version 1.2 and beyond will allow a node to join the cluster if it has at least one working interconnect that can talk to all of the other nodes in the cluster. However, for proper CF configuration using Cluster Admin, all of the interconnects should be working during the configuration process.

CIP configuration involves defining virtual CIP interfaces and assigning IP addresses to them. Up to eight CIP interfaces may be defined per node. These virtual interfaces act like normal TCP/IP interfaces except that the IP traffic is carried over the CF interconnects. Because CF is typically configured with multiple interconnects, the CIP traffic will continue to flow even if an interconnect fails. This helps eliminate single points of failure as far as physical networking connections are concerned for intracluster TCP/IP traffic.

Except for their IP configuration, the eight possible CIP interfaces per node are all treated identically. There is no special priority for any interface, and each interface uses all of the CF interconnects equally. For this reason, many system administrators may choose to define only one CIP interface per node.

To ensure that you can communicate between nodes using CIP, the IP address on each node for a specific CIP interface should use the same subnet.

CIP traffic is really intended only to be routed within the cluster. The CIP addresses should not be used outside of the cluster. Because of this, you should use addresses from the non-routable reserved IP address range.

Address Allocation for Private Internets (RFC 1918) defines three address ranges that are set aside for private subnets:

Subnets(s)	Class	Subnetmask
10.0.0.0	A	255.0.0.0
172.16.0.0 ... 172.31.0.0	B	255.255.0.0
192.168.0.0 ... 192.168.255.0	C	255.255.255.0

For CIP nodenames, it is strongly recommended that you use the following convention for RMS:

*cfname*RMS

cfname is the CF name of the node and RMS is a literal suffix. This will be used for one of the CIP interfaces on a node. This naming convention is used in the Cluster Admin GUI to help map between normal nodenames and CIP names. In general, only one CIP interface per node is needed to be configured.



A proper CIP configuration uses */etc/hosts* to store CIP names. You should make sure that */etc/nsswitch.conf(4)* is properly set up to use *files* criteria first in looking up its nodes. Refer to the *PRIMECLUSTER Installation Guide (Solaris)* for more details.

The recommended way to configure CF, CIP and CIM is to use the Cluster Admin GUI. A CF/CIP Wizard in the GUI can be used to configure CF, CIP, and CIM on all nodes in the cluster in just a few screens. Before running the wizard, however, the following steps must have been completed:

1. CF/CIP, Web-Based Admin View, and Cluster Admin should be installed on all nodes in the cluster.
2. If you are running CF over Ethernet, then all of the interconnects in the cluster should be physically attached to their proper hubs or networking equipment and should be working.
3. If you are running CF over IP, then all interfaces used for CF over IP should be properly configured and be up and running. See Chapter “CF over IP” for details.
4. Web-Based Admin View configuration must be done. Refer to the *PRIMECLUSTER Installation Guide (Solaris)* for details.

In the *cf* tab in Cluster Admin, make sure that the CF driver is loaded on that node. Press the *Load Driver* button if necessary to load the driver. Then press the *Configure* button to start the CF Wizard.

The CF/CIP Wizard is invoked by starting the GUI on a node where CF has not yet been configured. When this is done, the GUI will automatically bring up the CF/CIP Wizard in the *cf* tab of the GUI. You can start the GUI by entering the following URL with a browser running a proper version of the Java plug-in:

http://management_server:8081/Plugin.cgi

management_server is the primary or secondary management server you configured for this cluster. Refer to the *PRIMECLUSTER Installation Guide (Solaris)* for details on configuring the primary and secondary management service and on which browsers and Java plug-ins are required for the Cluster Admin GUI.

2.1.1 CIP versus CF over IP

Although the two terms CF over IP and CIP (also known as IP over CF) sound similar, they are two very distinct technologies.

CIP defines a reliable IP interface for applications on top of the cluster foundation (CF). CIP itself distributes the traffic generated by the application over the configured cluster interconnects (see Figure 1).

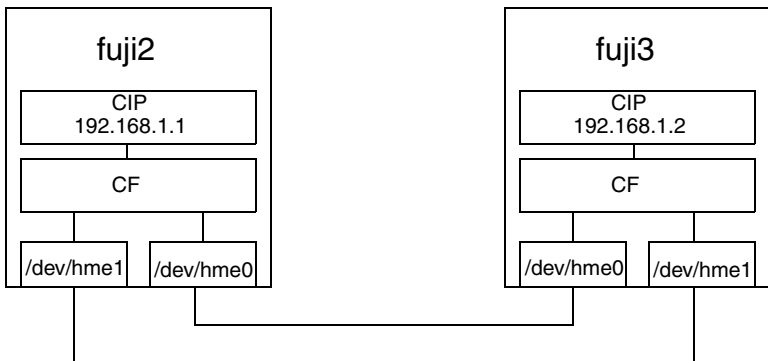


Figure 1: CIP diagram

CF over IP uses an IP interface, provided by the operating system, as a CF interconnect. The IP interface should not run over the public network. It should only be on a private network, which is also the local network. The IP interface over the private interconnect can be configured by using an IP address designed for the private network. The IP address normally uses the following address:

192.168.0.x

x is an integer between 1 and 254.

During the cluster joining process, CF sends broadcast messages to other nodes; therefore, all the nodes must be on the same local network. If one of the nodes is on a different network or subnet, the broadcast will not be received by that node. Therefore, the node will fail to join the cluster.

The following are possible scenarios for CF over IP:

- Where the cluster spans over two Ethernet segments of the same sub network. Each sub-level Ethernet protocol is not forwarded across the router but does pass IP traffic.
- When you need to reach beyond the physical cable length. Regular Ethernet is limited to the maximum physical length of the cable. Distances that are longer than the maximum cable length cannot be reached.
- If some of the network device cards that only support TCP/IP (for example, some Fiber channel) are not integrated into CF.

i Use CF with the Ethernet link-level connection whenever possible because CF over IP implies additional network/protocol information and usually will not perform as well (see Figure 2).

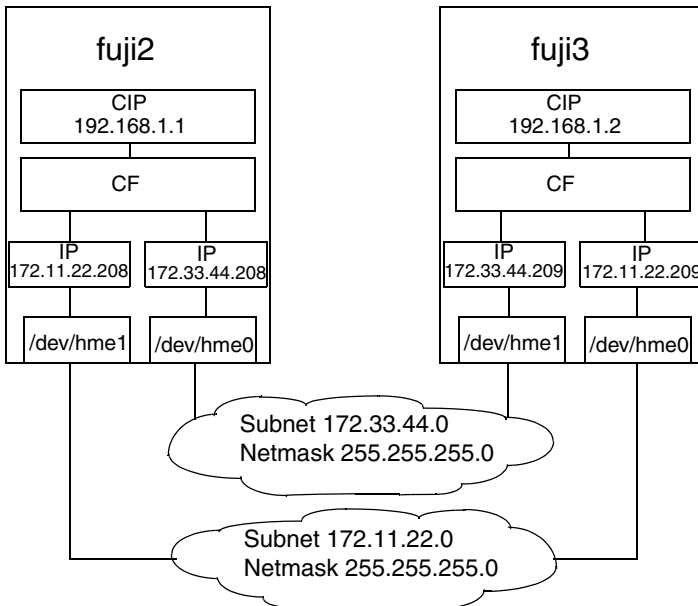


Figure 2: CF over IP diagram

2.1.2 cfset

The `cfset(1M)` utility is used to set certain tunable parameters in the CF driver. The values are stored in `/etc/default/cluster.config`. The `cfset(1M)` utility can be used to retrieve and display the values from the kernel or the file as follows:

- A new file under `/etc/default` called `cluster.config` is created.
- The values defined in `/etc/default/cluster.config` can be set or changed using the GUI (for `cfcp` and `cfsh`) or by using a text editor.
- The file consists of the following tuple entries, *Name* and *Value*:

Name:

- This is the name of a CF configuration parameter. It must be the first token in a line.
- Maximum length for *Name* is 31 bytes. The name must be unique.
- Duplicate names will be detected and reported as an error when the entries are applied by `cfconfig -l` and by the `cfset(1M)` utility (`cfset -r` and `-f` option). This will log invalid and duplicate entries to `/var/adm/messages`.
- `cfset(1M)` will change the *Value* for the *Name* in the kernel if the driver is already loaded and running.

Value:

- This represents the value to be assigned to the CF parameter. It is a string, enclosed in double quotes or single quotes. Maximum length for *Value* is 4K characters.
- New lines are not allowed inside the quotes.
- A new line or white space marks the close of a token.
- However, if double quotes or single quotes start the beginning of the line, treat the line as a continuation value from the previous value.

Example 1:

```
TEST      "abcde"
"1234"
```

The above becomes: TEST "abcde1234"

Example 2:

```
TEST
"abcde"
```

The above becomes:

```
TEST    "abcde"
```

"abcde" alone will be considered invalid format.

- The maximum number of *Name/Value* pair entries is 100.
- The hash sign (#) is used for the comment characters. It must be the first character in the line, and it causes the entries on that line to be ignored.
- Single quotes can be enclosed in double quotes or vice versa.

`cfset(1M)` has the following options:

- r Reloads all of the file entries into the CF driver.
- f Prints all *Name* and *Value* pairs from `/etc/default/cluster.config`. The file format will be verified, duplicate entries will be detected, and errors will be reported.
- o *Name* Prints only the specified *Name* and its corresponding *Value* entry from the file.
- a Prints from kernel all *Name* and *Value* pairs that CF is currently using.
- g *Name* Prints only the specified *Name* and its corresponding *Value* from the kernel.

The settable are as follows:

- `CLUSTER_TIMEOUT` (refer to the example that follows)
- `CFSH` (refer to the following Section “CF security”)
- `CFCP` (refer to the following Section “CF security”)

After any change to `cluster.config`, run the `cfset(1M)` command as follows:

```
# cfset -r
```

Example

Use `cfset(1M)` to tune timeout as follows:

```
CLUSTER_TIMEOUT "30"
```

This changes the default 10-second timeout to 30 seconds. The minimum value is 1 second. There is no maximum. It is strongly recommended that you use the same value on all cluster nodes.

`CLUSTER_TIMEOUT` represents the number of seconds that one cluster node waits while for a heartbeat response from another cluster node. Once `CLUSTER_TIMEOUT` seconds has passed, the non-responding node is declared to be in the `LEFTCLUSTER` state. The default value for `CLUSTER_TIMEOUT` is 10, which experience indicates is reasonable for most `PRIMECLUSTER` installations. We allow this value to be tuned for exceptional situations, such as networks which may experience long switching delays.

2.1.3 CF security

CF includes the ability to allow cluster nodes to execute commands on another node (`cfsh`) and to allow cluster nodes to copy files from one node to another (`cfcp`). However, this means that your cluster interconnects must be secure since any node that can join the cluster has access to these facilities. Because of this, these facilities are disabled by default.

`PRIMECLUSTER 4.1` offers a chance to configure these facilities. As one of the final steps of the CF Configuration Wizard in the Cluster Adm GUI, there are now two new checkboxes. Checking one will allow you to enable remote file copying and checking the other will enable remote command execution.

The `PRIMECLUSTER` family of products assume that the cluster interconnects are private networks; however, it is possible to use public networks as cluster interconnects because ICF does not interfere with other protocols running on the physical media. The security model for running `PRIMECLUSTER` depends on physical separation of the cluster interconnect networks from the public network.



For reasons of security, it is strongly recommended not to use public networks for the cluster interconnect.

The use of public networks for the cluster interconnects will allow any node on that public network to join the cluster (assuming that it is installed with the `PRIMECLUSTER` products). Once joined, an unauthorized user, through the node, would have full access to all cluster services.

In this release, we have included special functionality to be used in environments which do not support `.rhosts`.

If you do not wish to use `.rhosts` files, you should set the following parameters in `cluster.config` to enable remote access as follows:

```
CFCP "cfcp"  
CFSH "cfsh"
```

To deactivate, remove the settings from the `/etc/default/cluster.config` file and run `cfset -r`.

Refer to the Section “`cfset`” in this chapter for more information.

2.1.4 An example of creating a cluster

The following example shows what the Web-Based Admin View and Cluster Admin screens would look like when creating a two-node cluster. The nodes involved are named `fujii2` and `fujii3`, and the cluster name is `FUJII`.

This example assumes that Web-Based Admin View configuration has already been done. `fujii2` is assumed to be configured as the primary management server for Web-Based Admin View, and `fujii3` is the secondary management server.

The first step is to start Web-Based Admin View by entering the following URL in a java-enabled browser:

```
http://Management_Server:8081/Plugin.cgi
```

`fujii2` is a management server. Enter the following:

```
http://fujii2:8081/Plugin.cgi
```

After a few moments, a login password asking for a user name and password appears (see Figure 3).

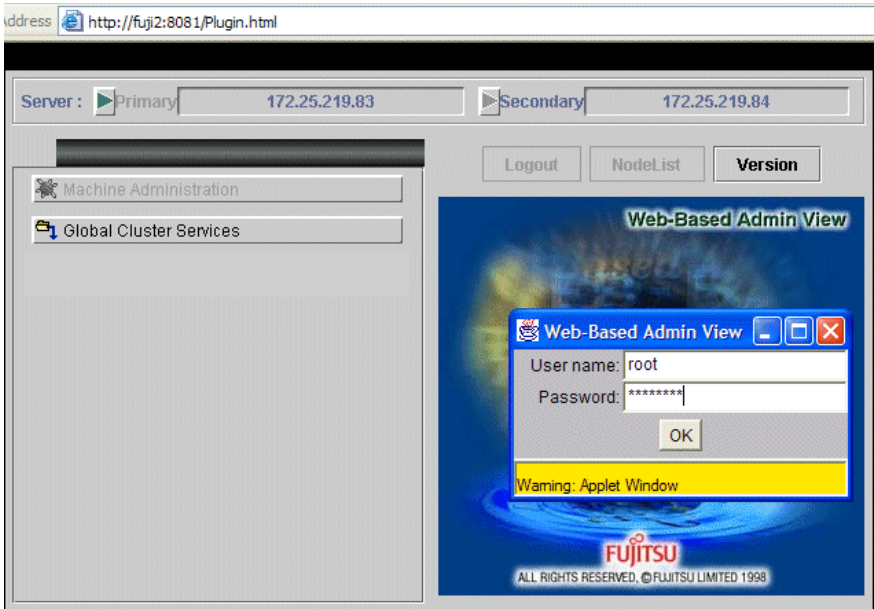


Figure 3: Login pop-up

Since you will be running the Cluster Admin CF Wizard, which does configuration work, you will need a privileged user ID such as `root`. There are three possible categories of users with sufficient privilege:

- The user `root`—You may enter `root` for the user name and `root`'s password on `fuji2`. The user `root` is always given the maximum privilege in Web-Based Admin View and Cluster Admin.
- A user in group `clroot`—You may enter the user name and password for a user on `fuji2` who is part of the UNIX group `clroot`. This user will have maximum privilege in Cluster Admin, but will be restricted in what Web-Based Admin View functions they can perform. This should be fine for CF configuration tasks.
- A user in group `wvroot`—You may enter the user name and password for a user on `fuji2` who is part of the UNIX group `wvroot`. Users in `wvroot` have maximum Web-Based Admin View privileges and are also granted maximum Cluster Admin privileges.

For further details on Web-Based Admin View and Cluster Admin privilege levels, refer to the PRIMECLUSTER *Installation Guide (Solaris)*.

After clicking on the *OK* button, the top menu appears (see Figure 4). Click on the button labeled *Global Cluster Services*.

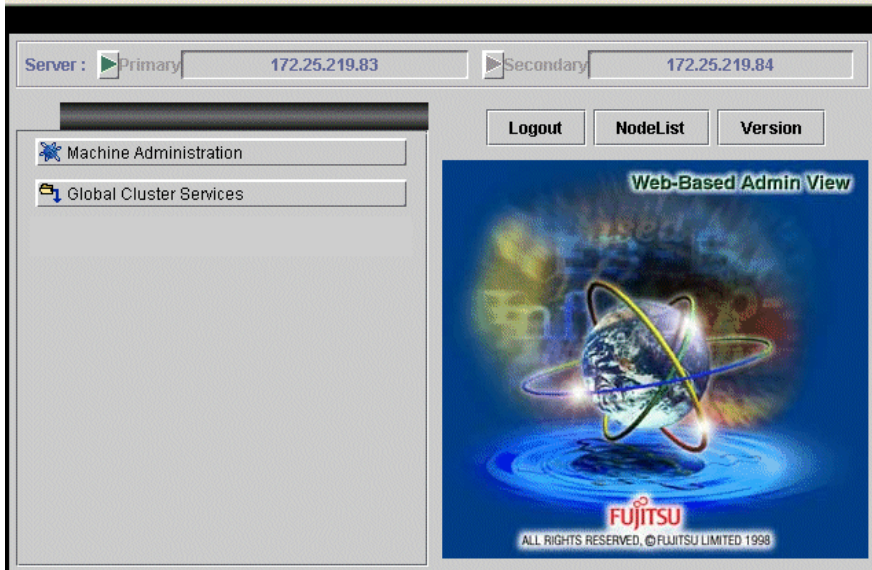


Figure 4: Main Web-Based Admin View screen after login

The Cluster Admin login screen appears (see Figure 5).

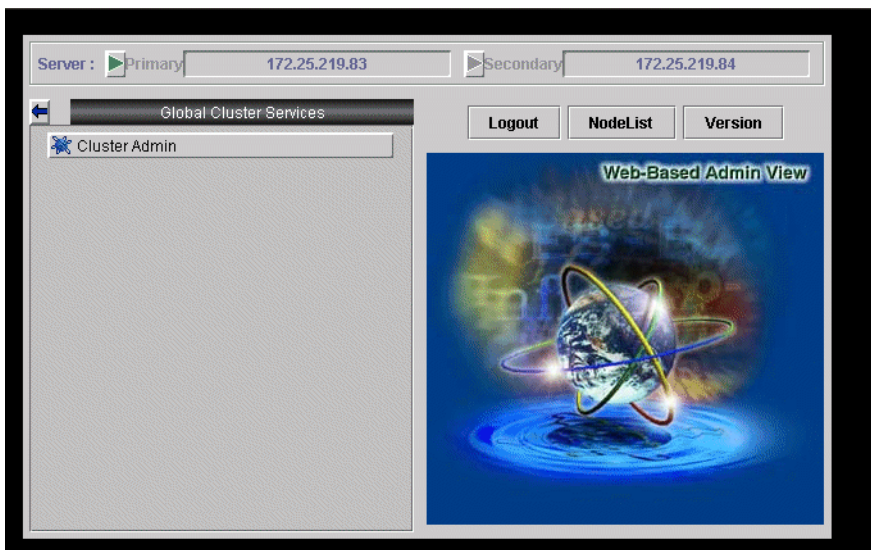


Figure 5: Global Cluster Services screen in Web-Based Admin View

Click on the button labeled *Cluster Admin* to launch the Cluster Admin GUI. The *Choose a node for initial connection* screen appears (see Figure 6).

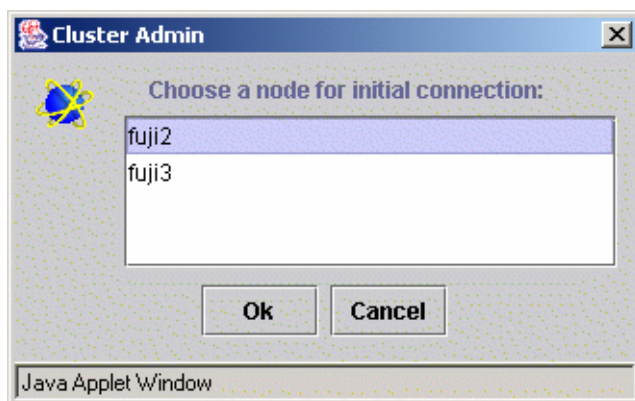


Figure 6: Initial connection pop-up

The *Choose a node for initial connection* screen (see Figure 6) lists the nodes that are known to the Web-Based Admin View management station. If you select a node where CF has not yet been configured, then Cluster Admin will let you run the CF Wizard on that node.

In this example, neither `fuji2` nor `fuji3` have had CF configured on them, so either would be acceptable as a choice. In Figure 6, `fuji2` is selected. Clicking on the *OK* button causes the main Cluster Admin GUI to appear. Since CF is not configured on `fuji2`, a screen similar to Figure 7 appears.

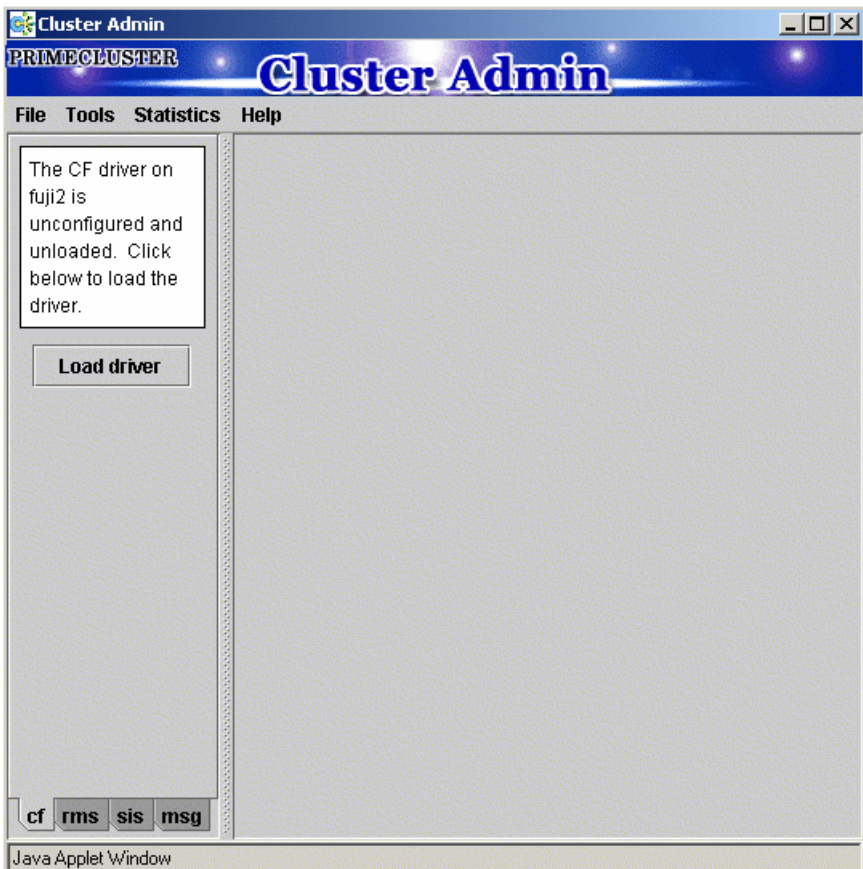


Figure 7: CF is unconfigured and unloaded

Click on the *Load driver* button to load the CF driver.

A screen indicating that CF is loaded but not configured appears (see Figure 8).



Figure 8: CF loaded but not configured

Click on the *Configure* button to bring up the CF Wizard.

The CF Wizard begins by looking for existing clusters (see Figure 9).

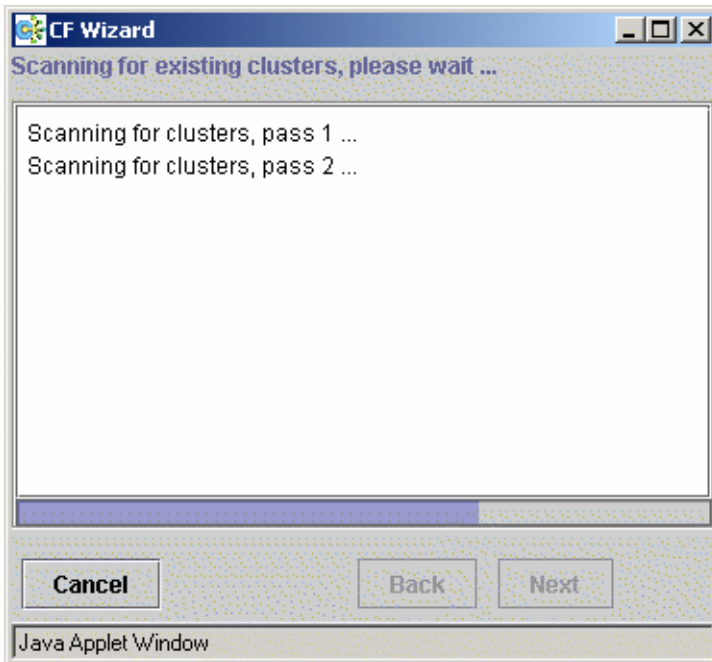


Figure 9: Scanning for clusters

After the CF Wizard finishes looking for clusters, a screen similar to Figure 10 appears.

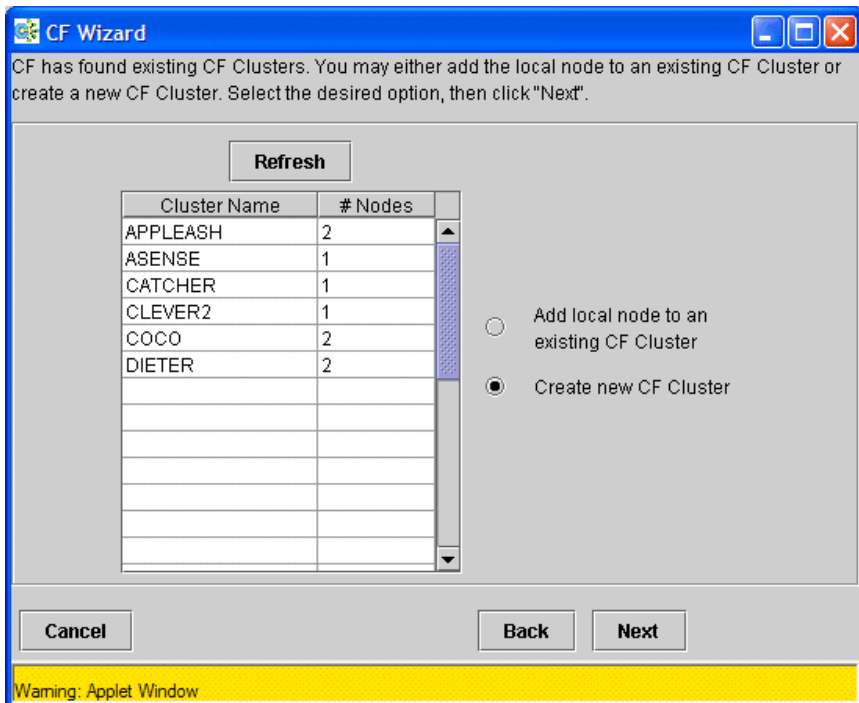


Figure 10: Creating or joining a cluster

This screen lets you decide if you want to join an existing cluster or create a new one. To create a new cluster, ensure that the *Create new CF Cluster* button is selected. Then, click on the *Next* button.

The screen for creating a new cluster appears (see Figure 11).

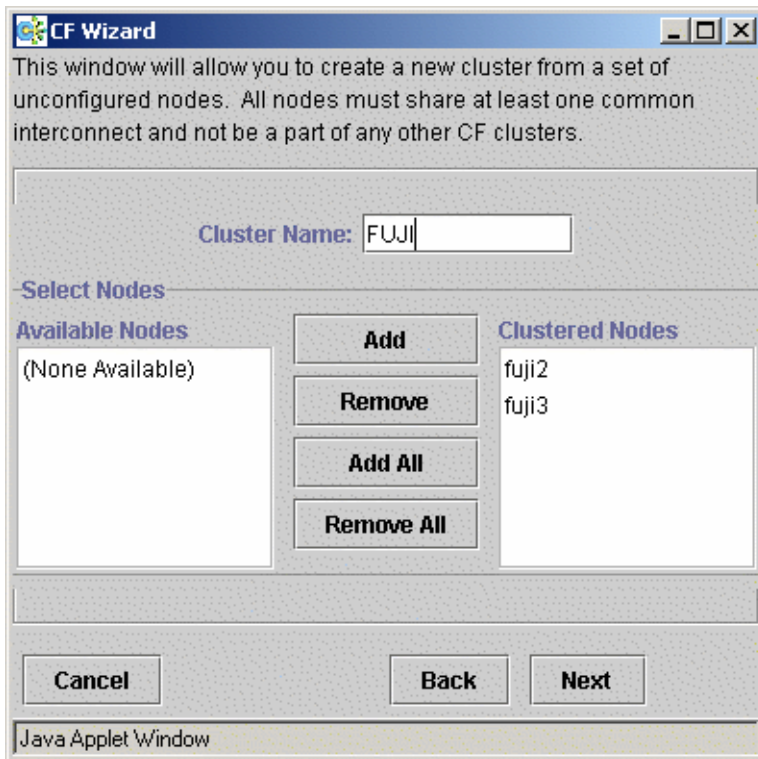


Figure 11: Selecting cluster nodes and the cluster name

This screen lets you choose the cluster name and also determine what nodes will be in the cluster. In the example above, we have chosen FUJI for the cluster name.

Below the cluster name are two boxes. The one on the right, under the label *Clustered Nodes*, contains all nodes that you want to become part of this CF cluster. The box on the left, under the label *Available Nodes*, contains all the other nodes known to the Web-Based Admin View management server. You should select nodes in the left box and move them to the right box using the *Add* or *Add All* button. If you want all of the nodes in the left box to be part of the CF cluster, then just click on the *Add All* button.

If you get to this screen and you do not see all of the nodes that you want to be part of this cluster, then there is a very good chance that you have not configured Web-Based Admin View properly. When Web-Based Admin View is initially installed on the nodes in a potential cluster, it configures each node as if it were a primary management server independent of every other node. If no additional Web-Based Admin View configuration were done, and you started up Cluster Admin on such a node, then Figure 11 would show only a single node in the right-hand box and no additional nodes on the left-hand side. If you see this, then it is a clear indication that proper Web-Based Admin View configuration has not been done.

Refer to the PRIMECLUSTER *Installation Guide (Solaris)* for more details on Web-Based Admin View configuration.

After you have chosen a cluster name and selected the nodes to be in the CF cluster, click on the *Next* button.

The screen that allows you to edit the CF node names for each node appears (see Figure 12). By default, the CF node names, which are shown in the right-hand column, are the same as the Web-Based Admin View names which are shown in the left-hand column.

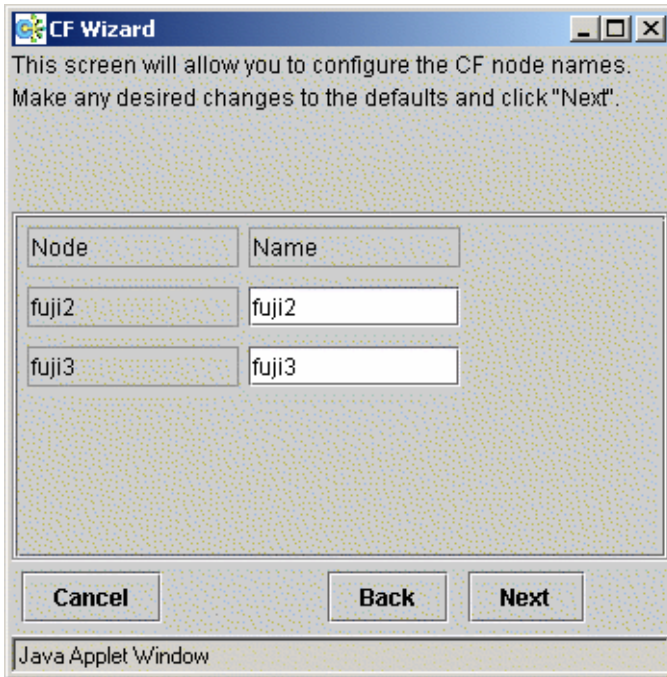


Figure 12: Edit CF node names

Make any changes to the CF node name and click *Next*.

The CF Wizard then loads CF on all the selected nodes and does CF pings to determine the network topology. While this activity is going on, a screen similar to Figure 13 appears.

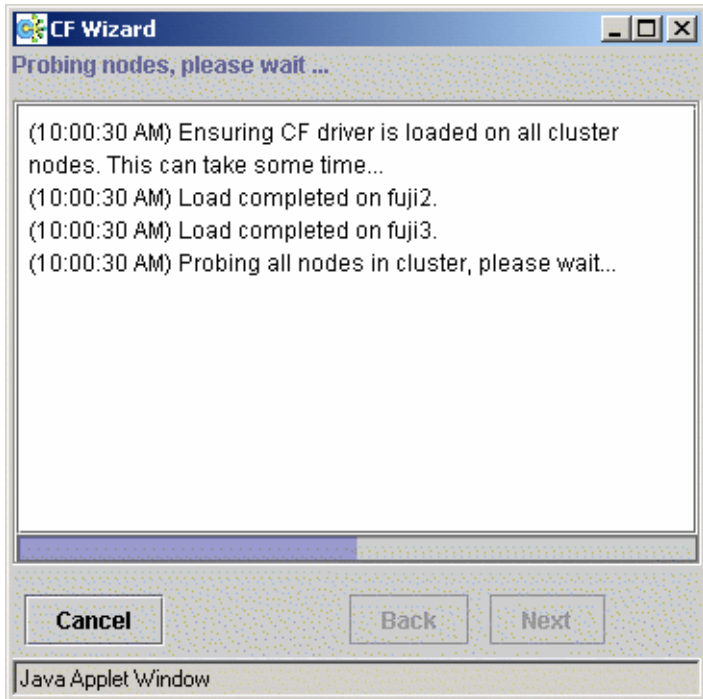


Figure 13: CF loads and pings

On most systems, loading the CF driver is a relatively quick process. However, on some systems that have certain types of large disk arrays, the first CF load can take up to 20 minutes or more.

After the CF Wizard has finished the loads and the pings, the CF topology and connection table appears (see Figure 14).

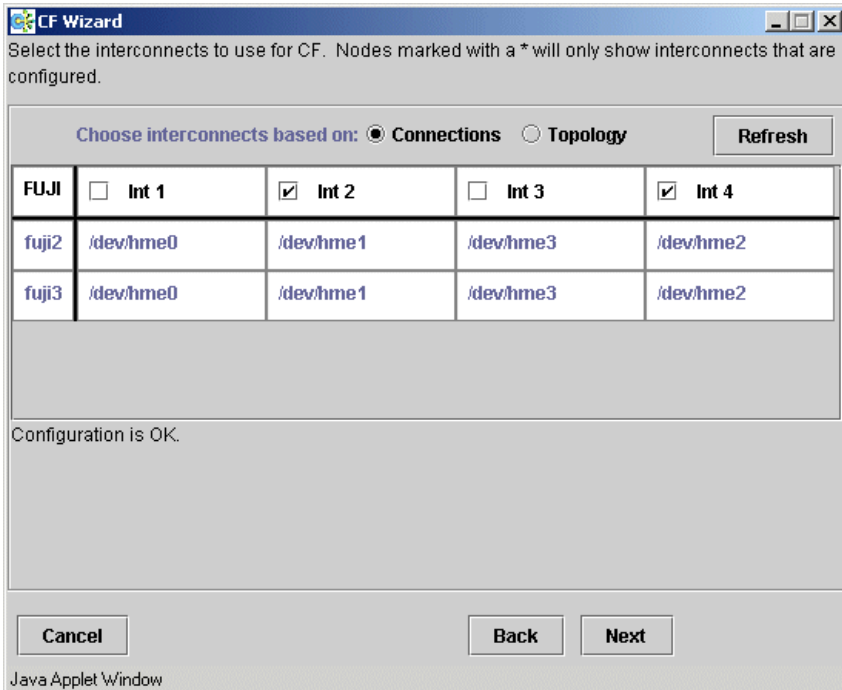


Figure 14: CF topology and connection table

Before using the CF topology and connection table in Figure 14, you should understand the following terms:

- Full interconnect—An interconnect where CF communication is possible to all nodes in the cluster.
- Partial interconnect—An interconnect where CF communication is possible between at least two nodes, but not to all nodes. If the devices on a partial interconnect are intended for CF communications, then there is a networking or cabling problem somewhere.
- Unconnected devices—These devices are potential candidates for CF configuration, but are not able to communicate with any other nodes in the cluster.

The CF Wizard determines all the full interconnects, partial interconnects, and unconnected devices in the cluster using CF pings. If there are one or more full interconnects, then it will display the connections table shown in Figure 14.

The connections table lists all full interconnects. Each column with an `Int` header represents a single interconnect. Each row represents the devices for the node whose name is given in the left-most column. The name of the CF cluster is given in the upper-left corner of the table.

In Figure 14, for example, Interconnect 1 (`Int 1`) has `/dev/hme0` on `fuji2` and `fuji3` attached to it. The cluster name is `FUJI`.

Although the CF Wizard may list `Int 1`, `Int 2`, and so on, it should be pointed out that this is simply a convention in the GUI. CF itself does not number interconnects. Instead, it keeps track of point-to-point routes to other nodes.

To configure CF using the connections table, click on the interconnects that have the devices that you wish to use. In Figure 14, Interconnects 2 and 4 have been selected. If you are satisfied with your choices, then you may click on *Next* to go to the CIP configuration screen.

Occasionally, there may be problems setting up the networking for the cluster. Cabling errors may mean that there are no full interconnects. If you click on the button next to *Topology*, the CF Wizard will display all the full interconnects, partial interconnects, and unconnected devices it has found. If a particular category is not found, it is omitted. For example, in Figure 14, only full interconnects are shown because no partial interconnects or unconnected devices were found on `fuji2` or `fuji3`.

The topology table gives more flexibility in configuration than the connection table. In the connection table, you could only select an interconnect, and all devices on that interconnect would be configured. In the topology table, you can individually select devices.

While you can configure CF using the topology table, you may wish to take a simpler approach. If no full interconnects are found, then display the topology table to see what your networking configuration looks like to CF. Using this information, correct any cabling or networking problems that prevented the full interconnects from being found. Then go back to the CF Wizard screen where the cluster name was entered and click on *Next* to cause the Wizard to reprobe the interfaces. If you are successful, then the connections table will show the full interconnects, and you can select them. Otherwise, you can repeat the process.

The text area at the bottom of the screen will list problems or warnings concerning the configuration.

When you are satisfied with your CF interconnect (and device) configuration, click on *Next*. This causes the CF over IP screen to appear (see Figure 15).

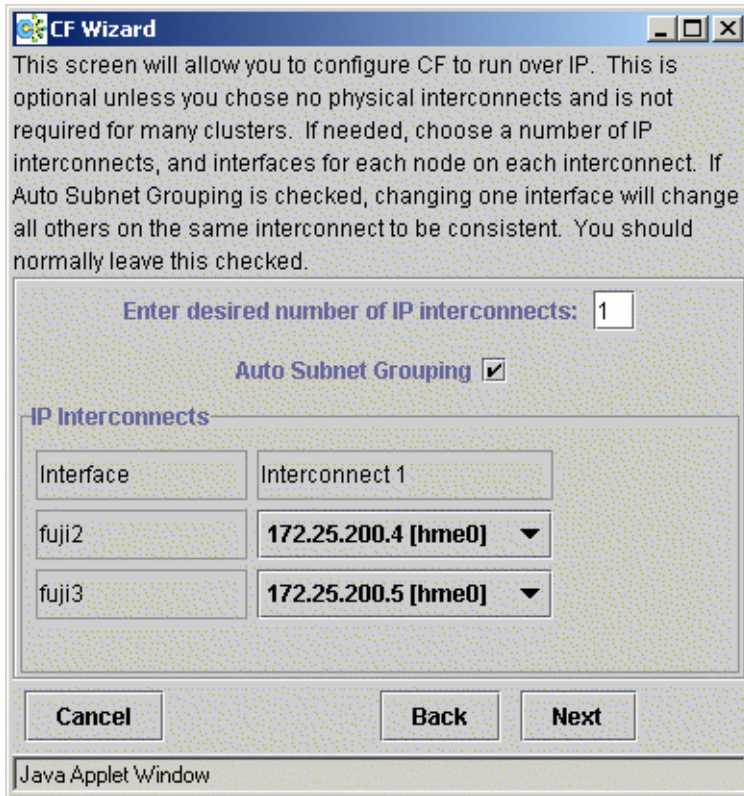


Figure 15: CF over IP screen

This is optional. If desired, enter the desired number of IP interconnects and press **Return**. The CF Wizard then displays interconnects sorted according to the valid subnetworks, netmasks, and broadcast addresses.

All the IP addresses for all the nodes on a given IP interconnect must be on the same IP subnetwork and should have the same netmask and broadcast address. CF over IP uses the IP broadcast address to find all the CF nodes during join process. So the dedicated network should be used for IP interconnects.

Auto Subnet Grouping should always be checked in this screen. If it is checked and you select one IP address for one node, then all of the other nodes in that column have their IP addresses changed to interfaces on the same subnetwork.

Choose the IP interconnects from the combo boxes on this screen, and click on *Next*. The CIP Wizard screen appears (see Figure 16).

CF Wizard

This screen will allow to configure IP over CF. Choose the number of subnets you would like, and for each subnet, choose a naming scheme, and an IP range. You may also mark one subnet for use by RMS.

Number of CIP subnets to configure:

Place all nodes on all CIP subnets:

CIP Subnets

Name	Host Suffix	Subnet Number	Subnet Mask	For RMS
net1	RMS	192.168.1.0	255.255.255.0	<input checked="" type="checkbox"/>

Cancel Back Next

Java Applet Window

Figure 16: CIP Wizard screen

This screen allows you to configure CIP. You can enter a number in the box after *Number of CIP subnets to configure* to set the number of CIP subnets to configure. The maximum number of CIP subnets is 8.

For each defined subnet, the CIP Wizard configures a CIP interface on each node defined in the CF cluster. The CIP interface will be assigned the following values:

- The IP address will be a unique IP number on the subnet specified in the *Subnet Number* field. The node portions of the address start at 1 and are incremented by 1 for each additional node.

The CIP Wizard will automatically fill in a default value for the subnet number for each CIP subnetwork requested. The default values are taken from the private IP address range specified by RFC 1918. Note that the values entered in the *Subnet Number* have 0 for their node portion even though the CIP Wizard starts the numbering at 1 when it assigns the actual node IP addresses.

- The IP name of the interface will be of the form *cfnameSuffix* where *cfname* is the name of a node from the CF Wizard, and the *Suffix* is specified in the field *Host Suffix*. If the checkbox *For RMS* is selected, then the host suffix will be set to RMS and will not be editable. If you are using RMS, one CIP network must be configured for RMS.
- The *Subnet Mask* will be the value specified.

In Figure 16, the system administrator has selected 1 CIP network. The *For RMS* checkbox is selected, so the RMS suffix will be used. Default values for the *Subnet Number* and *Subnet Mask* are also selected. The nodes defined in the CF cluster are *fuji2* and *fuji3*. This will result in the following configuration:

- On *fuji2*, a CIP interface will be configured with the following:
IP nodename: *fuji2RMS*
IP address: 192.168.1.1
Subnet Mask: 255.255.255.0
- On *fuji3*, a CIP interface will be configured with the following:
IP nodename: *fuji3RMS*
IP address: 192.168.1.2
Subnet Mask: 255.255.255.0

The CIP Wizard stores the configuration information in the file `/etc/cip.cf` on each node in the cluster. This is the default CIP configuration file. The Wizard will also update `/etc/hosts` on each node in the cluster to add the new IP nodenames. The cluster console will not be updated.

i The CIP Wizard always follows an orderly naming convention when configuring CIP names. If you have done some CIP configuration by hand before running the CIP Wizard, then you should consult the Wizard documentation to see how the Wizard handles irregular names.

When you click on the *Next* button, CIM configuration screen appears (see Figure 17).

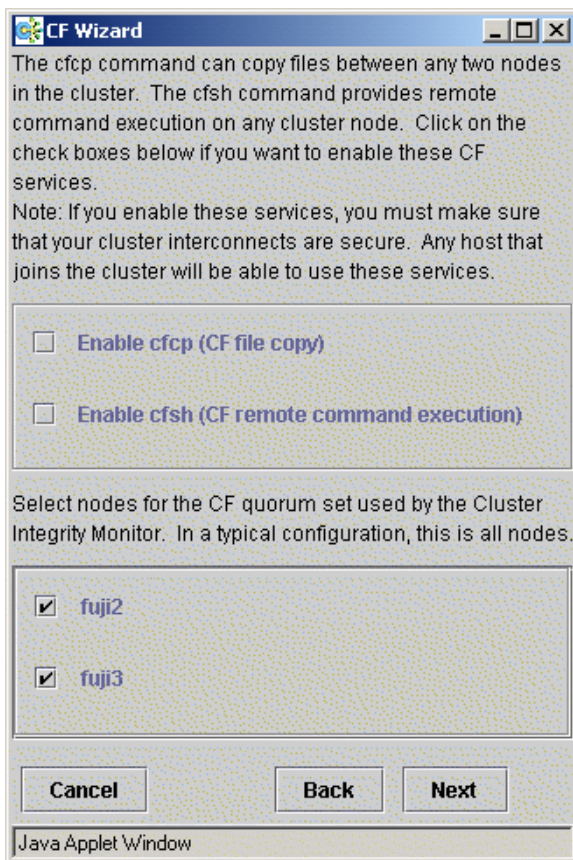


Figure 17: CIM configuration screen

The CIM configuration screen in Figure 17 has the following parts:

- The upper portion allows you to enable `cfcp` and `cfsh`.
`cfcp` is a CF-based file copy program. It allows files to be copied among the cluster hosts. `cfsh` is a remote command execution program that similarly works between nodes in the cluster. The use of these programs is optional. In this example these items are not selected. If you enable these services, however, any node that has access to the cluster interconnects can copy files or execute commands on any node with root privileges.
- The lower portion allows you to determine which nodes should be monitored by CIM.

This screen also lets you select which nodes should be part of the CF quorum set. The CF quorum set is used by the CIM to tell higher level services such as GDS when it is safe to access shared resources.

**Caution**

Do not change the default selection of the nodes that are members of the CIM set unless you fully understand the ramifications of this change.

A checkbox next to a node means that node will be monitored by CIM. By default, all nodes are checked. For almost all configurations, you will want to have all nodes monitored by CIM.

This screen will also allow you to configure CF Remote Services. You can enable either remote command execution, remove file copying, or both.

**Caution**

Enabling either of these means that you must trust all nodes on the CF interconnects and the CF interconnects must be secure. Otherwise any system able to connect to the CF interconnects will have access to these services.

Click on the *Next* button to go to the summary screen (see Figure 18).

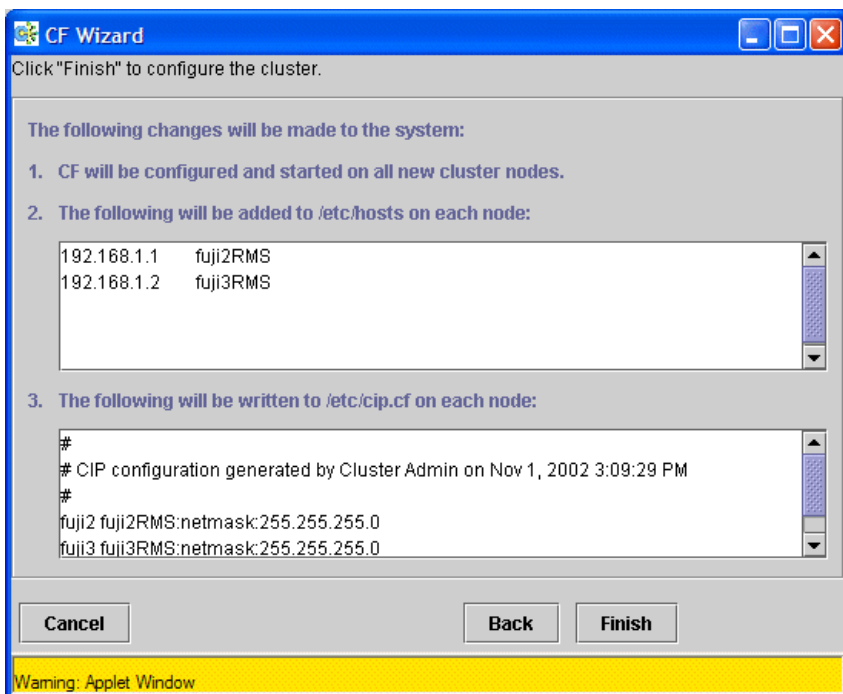


Figure 18: Summary screen

This screen summarizes the major changes that the CF, CIP, and CIM Wizards will perform. When you click on the *Finish* button, the CF Wizard performs the actual configuration on all nodes.

A screen similar to Figure 19 is displayed while the configuration is being done.

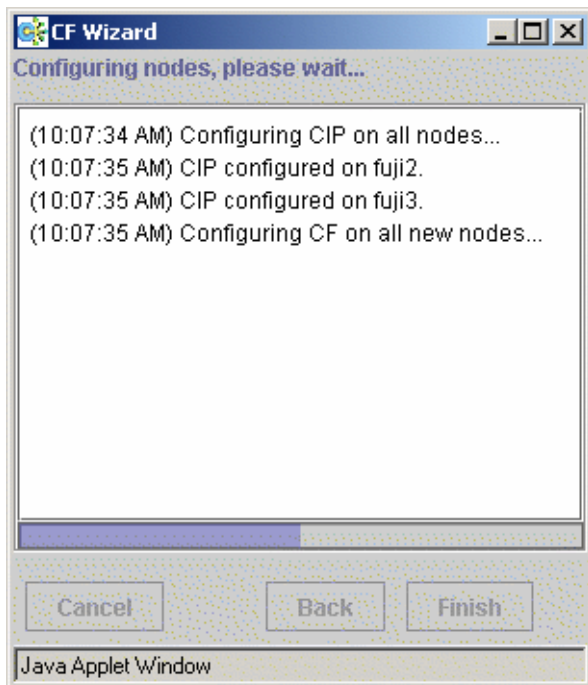


Figure 19: Configuration processing screen

This screen is updated after each configuration step. When configuration is complete, a pop-up appears announcing this fact (see Figure 20).

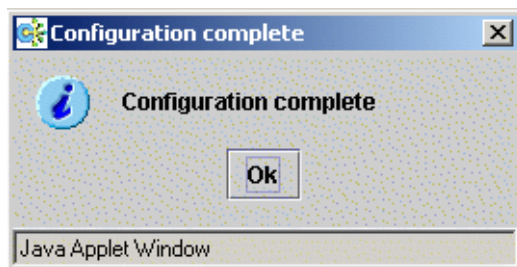


Figure 20: Configuration completion pop-up

Click on the *OK* button, and the pop-up is dismissed. The configuration processing screen now has a *Finish* button (see Figure 21).

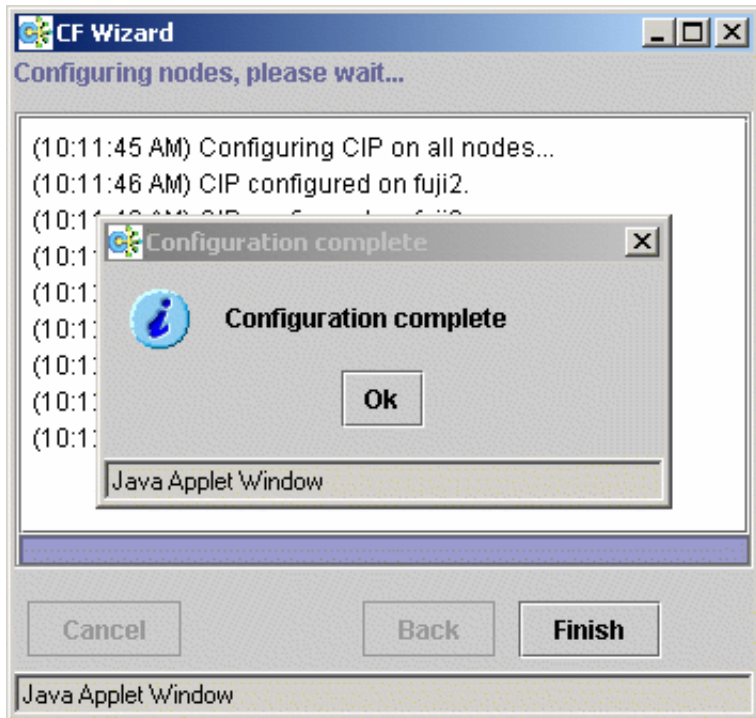


Figure 21: Configuration screen after completion

You might see the following error message in the screen shown in Figure 21:

```
cf:cfconfig      OSDU_stop: failed to unload cf_drv
```

Unless you are planning to use the dynamic hardware reconfiguration feature of PRIMEPOWER, then you can safely ignore this message.

When the CF Wizard is run on an unconfigured node, it will ask the CF driver to push its modules on every Ethernet device on the system. This allows CF to do CF pings on each interface so that the CF Wizard can discover the network topology.

Occasionally, this unload will fail. To correct this problem, you need to unload and reload the CF driver on the node in question. This can be done easily through the GUI (refer to the Section “Starting and stopping CF”).

Click on the *Finish* button to dismiss the screen in Figure 21. A small pop-up appears asking if you would like to run the SF Wizard. Click on *yes*, and run the SF Wizard (described in the Section “Invoking the Configuration Wizard”).

After the CF (and optionally the SF) Wizards are done, you will see the main Cluster Admin screen which will resemble Figure 22.

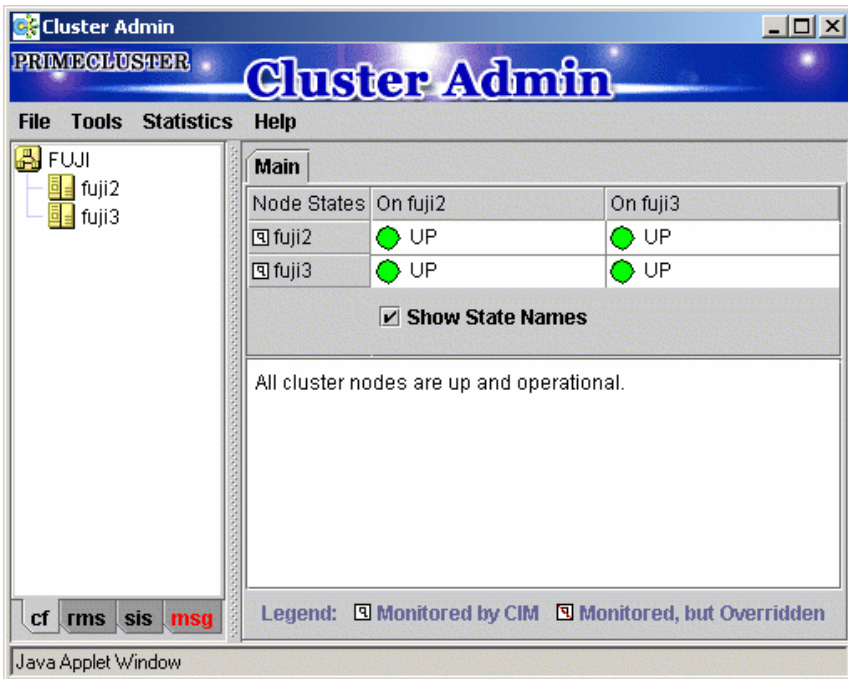


Figure 22: Main CF screen

2.2 CIP configuration file

The CIP configuration file is stored in `/etc/cip.cf` on each node in the cluster. Normally, you can use the GUI to create this file during cluster configuration time. However, there may be times when you wish to manually edit this file.

The format of a CIP configuration file entry is as follows:

```
cfname CIP_Interface_Info [ CIP_Interface_Info ... ]
```

The `cip.cf` configuration file typically contains configuration information for all CIP interfaces on all nodes in the cluster. The first field, *cfname*, tells what node the configuration information is for. When a node parses the `cip.cf` file, it can ignore all lines that do not start with its own CF node name.

The *CIP_Interface_Info* gives all of the IP information needed to configure a single CIP interface. At the minimum, it must consist of an IP address. The address may be specified as either a number in internet dotted-decimal notation or as a symbolic node name. If it is a symbolic node name, it must be specified in `/etc/hosts`.

The IP address may also have additional options following it. These options are passed to the configuration command `ifconfig`. They are separated from the IP address and each other by colons (:). No spaces may be used around the colons.

For example, the CIP configuration done in Section “An example of creating a cluster” would produce the following CIP configuration file:

```
fuji2    fuji2RMS:netmask:255.255.255.0
fuji3    fuji3RMS:netmask:255.255.255.0
```

Although not shown in this example, the CIP syntax does allow multiple CIP interfaces for a node to be defined on a single line. Alternately, additional CIP interfaces for a node could be defined on a subsequent line beginning with that node's CF node name. The `cip.cf` manual page has more details about the `cip.cf` file.

If you make changes to the `cip.cf` file by hand, you should be sure that the file exists on all nodes, and all nodes are specified in the file. Be sure to update all nodes in the cluster with the new file. Changes to the CIP configuration file will not take effect until CIP is stopped and restarted. If you stop CIP, be sure to stop all applications that use it. In particular, RMS needs to be shut down before CIP is stopped.

To stop CIP, use the following command:

```
# /opt/SMAW/SMAWcf/dep/stop.d/K98cip unload
```

To start or restart CIP, use the following command:

```
# /opt/SMAW/SMAWcf/dep/start.d/S01cip load
```

2.3 Cluster Configuration Backup and Restore (CCBR)



Caution

CCBR only saves PRIMECLUSTER configuration information. It does not replace an external, full backup facility.

CCBR provides a simple method to save the current PRIMECLUSTER configuration information of a cluster node. It also provides a method to restore the configuration information whenever a node update has caused severe trouble or failure, and the update (and any side-effects) must be removed. CCBR provides a node-focused backup and restore capability. Multiple cluster nodes must each be handled separately.

CCBR provides the following commands:

- `cfbackup(1M)`—Saves all information into a directory that is converted to a compressed tar archive file.
- `cfrestore(1M)`—Extracts and installs the saved configuration information from one of the `cfbackup(1M)` compressed tar archives.

After `cfrestore(1M)` is executed, you must reactivate the RMS configuration in order to start RMS.



To guarantee that the `cfrestore(1M)` command will restore a functional PRIMECLUSTER configuration, it is recommended that there be no hardware or operating system changes since the backup was taken, and that the same versions of the PRIMECLUSTER products are installed.

Because the installation or reinstallation of some PRIMECLUSTER products add kernel drivers, device reconfiguration may occur. This is usually not a problem. However, if Network Interface Cards (NICs) have been installed, removed, replaced, or moved, the device instance numbers (for example, the number 2 in `/dev/hme2`) can change. Any changes of this nature can, in turn, cause a restored PRIMECLUSTER configuration to be invalid.

`cfbackup(1M)` and `cfrestore(1M)` consist of a framework and plug-ins. The framework and plug-ins function as follows:

1. The framework calls the plug-in for the `SMAWcf` package.
2. This plug-in creates and updates the saved-files list, the log files, and error log files.

3. All the other plug-ins for installed PRIMECLUSTER products are called in name sequence.
4. Once all plug-ins have been successfully processed, the backup directory is archived (using `tar`) and compressed.
5. The backup is logged as complete and the file lock on the log file is released.

The `cfbackup(1M)` command runs on a PRIMECLUSTER node to save all the cluster configuration information. To avoid any problem, this command should be concurrently executed on every cluster node to save all relevant PRIMECLUSTER configuration information. This command must be executed as `root`.

If a backup operation is aborted, no `tar` archive is created. If the backup operation is not successful for one plug-in, the command processing will abort rather than continue with the next plug-in. `cfbackup(1M)` exits with a status of zero on success and non-zero on failure.

The `cfrestore(1M)` command runs on a PRIMECLUSTER node to restore all previously saved PRIMECLUSTER configuration information from a compressed `tar` archive. The node must be in single-user mode with CF not loaded. The node must not be an active member of a cluster. The command must be executed as `root`. `cfrestore(1M)` exits with a status of zero on success and non-zero on failure.

It is recommended to reboot once `cfrestore(1M)` returns successfully. If `cfrestore(1M)` aborts, the reason for this failure should be examined carefully since the configuration update may be incomplete.



There should only be one `cfbackup/cfrestore` command active at a time on one node.



Note that certain PRIMECLUSTER information is given to a node when it joins the cluster. The information restored is not used. In order to restore and to use this PRIMECLUSTER information, the entire cluster needs to be `DOWN` and the first node to create the cluster must be the node with the restored data.

The following files and directories that are fundamental to the operation of the `cfbackup(1M)` and `cfrestore(1M)` commands:

- The `/opt/SAW/ccbr/plugins` directory contains executable CCBP plug-in. The installed PRIMECLUSTER products supply them.

- The `/opt/SMAW/ccbr/ccbr.conf` file must exist and specifies the value for `CCBRHOME`, the pathname of the directory to be used for saving CCBR archive files. A default `ccbr.conf` file, with `CCBRHOME` set to `/var/spool/SMAW/SMAWccbr` is supplied as part of the `SMAWccbr` package.

The system administrator can change the `CCBRHOME` pathname at anytime. It is recommended that the system administrator verify that there is enough disk space available for the archive file before setting `CCBRHOME`. The system administrator might need to change the `CCBRHOME` pathname to a file system with sufficient disk space.



It is important to remember that re-installing the `SMAWccbr` package will reset the contents of the `/opt/SMAW/ccbr/ccbr.conf` file to the default package settings.

The following is an example of `ccbr.conf`:

```
#!/bin/ksh -
#ident "@(#)ccbr.conf Revision: 12.1 02/05/08 14:45:57"
#
# CCBR CONFIGURATION FILE
#
# set CCBR home directory
#
CCBRHOME=/var/spool/SMAW/SMAWccbr
export CCBRHOME
```

- The `/opt/SMAW/ccbr/ccbr.gen` (generation number) file is used to form the name of the CCBR archive to be saved into (or restored from) the `CCBRHOME` directory. This file contains the next backup sequence number. The generation number is appended to the archive name.

If this file is ever deleted, `cfbackup(1M)` and/or `cfrestore(1M)` will create a new file containing the value string of 1. Both commands will use either the generation number specified as a command argument, or the file value if no command argument is supplied. The `cfbackup(1M)` command additionally checks that the command argument is not less than the value of the `/opt/SMAW/ccbr/ccbr.gen` file. If the command argument is less than the value of the `/opt/SMAW/ccbr/ccbr.gen` file, the `cfbackup(1M)` command will use the file value instead.

Upon successful execution, the `cfbackup(1M)` command updates the value in this file to the next sequential generation number. The system administrator may update this file at any time.

- If `cfbackup(1M)` backs up successfully, a compressed tar archive file with the following name will be generated in the `CCBRHOME` directory:

`hostname_ccbrN.tar.Z`

`hostname` is the nodename and `N` is the number suffix for the generation number.

For example, in the cluster node `fuji2`, with the generation number 5, the archive file name will be:

`fuji2_ccbr5.tar.Z`

- Each backup request creates a backup tree directory. The directory is:

`CCBRHOME/nodename_ccbrN`

`nodename` is the node name and `N` is the number suffix for the generation number.

`CCBROOT` is set to this directory.

For example, on the node `fuji2`:

```
fuji2# cfbackup 5
```

Using the default setting for `CCBRHOME`, the following directory will be created:

```
/var/spool/SMAW/SMAWccbr/fuji2_ccbr5
```

This backup directory tree name is passed as an environment variable to each plug-in.

- The `CCBRHOME/ccbr.log` log file contains startup, completion messages, and error messages. All the messages are time stamped.
- The `CCBROOT/errlog` log file contains specific error information when a plug-in fails. All the messages are time stamped.
- The `CCBROOT/plugin.blog` or `CCBROOT/plugin.rlog` log files contain startup and completion messages from each backup /restore attempt for each plug-in. These messages are time stamped.

Refer to the Chapter “Manual pages” for more information on `cfbackup(1M)` and `cfrestore(1M)`.

Example 1: Backup**fuji2# cfbackup**

This command backs up and validates the configuration files for all CCBR plugins that exist on the system fuji2.

The output of the cfbackup(1M) command looks similar to the following:

```
# cfbackup
01/16/03 17:21:39 cfbackup 11 started
01/16/03 17:21:40 active cluster nodes:
Node Number State      Os      Cpu
fuji2 1      UP      Solaris Sparc
fuji3 2      UP      Solaris Sparc
01/16/03 17:21:40 installed ccb plugins:
FJSVclapm.pi
FJSVcldev.pi
FJSVwvbs.pi
SMAWcf.pi
SMAWdtcp.pi
_rmswizvalidate
_sample.pi
rmswizbackup
rscmgr.pi
sfbackup
sfvalidate
01/16/03 17:21:40 FJSVclapm: validate started
01/16/03 17:21:40 FJSVclapm: validate ended
01/16/03 17:21:40 FJSVcldev: validate started
01/16/03 17:21:40 FJSVcldev: validate ended
01/16/03 17:21:40 FJSVwvbs: validate started
01/16/03 17:21:40 FJSVwvbs: validate ended
01/16/03 17:21:40 SMAWcf: validate started for /var/spool/SMAW/
SMAWccbr/fuji2_ccbr11
01/16/03 17:21:40 SMAWcf: validate ended
01/16/03 17:21:41 SMAWdtcp: validate started
Checking for file /etc/dtcp.ap
Checking for file /etc/inittab
01/16/03 17:21:41 SMAWdtcp: validate ended
01/16/03 17:21:41 validation failed in /opt/SMAW/ccbr/plugins/
rmswizbackup
01/16/03 17:21:41 rscmgr.pi: validate started
01/16/03 17:21:41 rscmgr.pi: validate normal ended
SMAWsf validation begins .....
Validation done. No problems found .....
Please read the validation report /var/spool/SMAW/SMAWccbr/
fuji2_ccbr11/sf/backupvalidatelog
01/16/03 17:21:41 cfbackup 11 ended unsuccessfully.
```


The output shows that `cfbackup(1M)` ended unsuccessfully with the problem in the `rmswizbackup`. In this case, the subdirectory `/var/spool/SMAW/SMAWccbr/fuji2_ccbr11` will be created. Under this directory, `rmswizbackup.blog` and `errlog` will be found.

Output from the `rmswizbackup.blog` file:

```
01/16/03 17:21:40 rmswizbackup: validate started
01/16/03 17:21:40 rmswizbackup: validate ended
```

Output from `errlog`:

```
01/16/03 17:21:40 cfbackup 11 error log started
Environment variable $CCBR00T not set ...
/opt/SMAW/ccbr/plugins/rmswizbackup[66]:
/opt/SMAW/ccbr/plugins/rmswizvalidate: not found
```

Before doing `cfrestore(1M)`, CF needs to be unloaded. The system needs to be in single-user mode.

The following files are handled differently during `cfrestore(1M)`:

- **root files**—These are the files under the `CCBR00T/root` directory. They are copied from the `CCBR00T/root` file tree to their corresponding places in the system file tree.
- **OS files**—These files are the operating system files that are saved in the archive but not restored. The system administrator might need to merge the new OS files and the restored OS files to get the necessary changes.

Example 2: Restore

```
fuji2# cfrestore 11
```

The output of the `cfrestore(1M)` command looks similar to the following:

```
01/16/03 17:35:28 cfrestore 11 started
01/16/03 17:35:28 extract files from tar archive
x .. 0 bytes, 0 tape blocks
x ./root, 0 bytes, 0 tape blocks
x ./root/etc, 0 bytes, 0 tape blocks
x ./root/etc/opt, 0 bytes, 0 tape blocks
x ./root/etc/opt/FJSVwvbs, 0 bytes, 0 tape blocks
x ./root/etc/opt/FJSVwvbs/etc, 0 bytes, 0 tape blocks
x ./root/etc/opt/FJSVwvbs/etc/webview.cnf, 834 bytes, 2 tape
blocks
x ./root/etc/opt/FJSVwvbs/etc/wvlocal.cnf, 260 bytes, 1 tape
blocks
x ./root/etc/default, 0 bytes, 0 tape blocks
```

```
x ./root/etc/default/cluster, 136 bytes, 1 tape blocks
x ./root/etc/default/cluster.config, 144 bytes, 1 tape blocks
x ./root/etc/cip.cf, 279 bytes, 1 tape blocks
x ./root/var, 0 bytes, 0 tape blocks
x ./root/var/adm, 0 bytes, 0 tape blocks
x ./root/var/adm/cfreg.data, 216 bytes, 1 tape blocks
x ./OS, 0 bytes, 0 tape blocks
x ./OS/etc, 0 bytes, 0 tape blocks
x ./OS/etc/hosts, 195 bytes, 1 tape blocks
x ./errlog, 92 bytes, 1 tape blocks
x ./ccbr.cluster.list, 79 bytes, 1 tape blocks
x ./ccbr.plugin.list, 33 bytes, 1 tape blocks
x ./pirc, 2 bytes, 1 tape blocks
x ./FJSVwvbs.blog, 172 bytes, 1 tape blocks
x ./SMAWcf.blog, 242 bytes, 1 tape blocks
x ./FJSVwvbs.id, 36 bytes, 1 tape blocks
x ./saved.files, 160 bytes, 1 tape blocks
x ./SMAWcf.id, 20 bytes, 1 tape blocks
01/16/03 17:35:28 this backup (/var/spool/SMAW/SMAWccbr/
fuji2_ccbr11) created on
01/16/03 17:26:32
01/16/03 17:35:28 nodes in the cluster were:
Node Number State      Os      Cpu
fuji2 1      UP      Solaris Sparc
fuji3 2      UP      Solaris Sparc
Are you sure you want to continue (y/n) y
01/16/03 17:36:02 FJSVwvbs: validate started
01/16/03 17:36:02 FJSVwvbs: validate ended
01/16/03 17:36:02 SMAWcf: validate started for /var/spool/SMAW/
SMAWccbr/fuji2_ccbr11
01/16/03 17:36:02 SMAWcf: validate ended
01/16/03 17:36:02 cfrestore: The following files will be
automatically restored:
/etc/opt/FJSVwvbs/etc/webview.cnf
/etc/opt/FJSVwvbs/etc/wvlocal.cnf
/etc/opt/FJSVwvbs/etc
/etc/opt/FJSVwvbs
/etc/opt
/etc/default/cluster
/etc/default/cluster.config
/etc/default
/etc/cip.cf
/etc
/var/adm/cfreg.data
/var/adm
/var
/
112 blocks
```

```
01/16/03 17:36:02 FJSVwvbs: restore started
01/16/03 17:36:02 FJSVwvbs: restore ended
01/16/03 17:36:03 SMAWcf: restore started for /var/spool/SMAW/
SMAWccbr/fuji2_ccbr11
01/16/03 17:36:03 SMAWcf: restore ended
01/16/03 17:36:03 cfrestore: System Administrator please NOTE:
The following system (OS) files were saved but have not been
restored:
./etc/hosts
01/16/03 17:36:03 cfrestore 11 ended
```

3 CF Registry and Integrity Monitor

This chapter discusses the purpose and physical characteristics of the CF registry (CFREG), and it discusses the purpose and implementation of the Cluster Integrity Monitor (CIM).

This chapter discusses the following:

- The Section “CF Registry” discusses the purpose and physical characteristics of the CF synchronized registry.
- The Section “Cluster Integrity Monitor” discusses the purpose and implementation of CIM.

3.1 CF Registry

The CFREG provides a set of CF base product services that allows cluster applications to maintain cluster global data that must be consistent on all of the nodes in the cluster and must live through a clusterwide reboot.

Typical applications include cluster-aware configuration utilities that require the same configuration data to be present and consistent on all of the nodes in a cluster (for example, cluster volume management configuration data).

The data is maintained as named registry entries residing in a data file where each node in the cluster has a copy of the data file. The services will maintain the consistency of the data file throughout the cluster.

A user-level daemon (`cfregd`), runs on each node in the cluster, and is responsible for keeping the data file on the node where it is running synchronized with the rest of the cluster. The `cfregd` process will be the only process that ever modifies the data file. Only one synchronization daemon process will be allowed to run at a time on a node. If a daemon is started with an existing daemon running on the node, the started daemon will log messages that state that a daemon is already running and terminate itself. In such a case, all execution arguments for the second daemon will be ignored.

3.2 Cluster Integrity Monitor

The purpose of the CIM is to allow applications to determine when it is safe to perform operations on shared resources. It is safe to perform operations on shared resources when a node is a member of a cluster that is in a consistent state.

A consistent state means that all the nodes of a cluster that are members of the CIM set are in a known and safe state. The nodes that are members of the CIM set are specified in the CIM configuration. Only these nodes are considered when the CIM determines the state of the cluster. When a node first joins or forms a cluster, the CIM indicates that the cluster is consistent only if it can determine the status of the other nodes that make up the CIM set and that those nodes are in a safe state.

CIM currently supports Node State Management (NSM) method. The Remote Cabinet Interface (RCI) method is supported for PRIMEPOWER nodes.

The CIM reports on a node state that a cluster is consistent (`True`), or a cluster is not consistent (`False`) for the node. `True` and `False` are defined as follows:

`True`—All CIM nodes in the cluster are in a known state.

`False`—One or more CIM nodes in the cluster are in an unknown state.

3.2.1 Configuring CIM

You can perform CIM procedures through the following methods:

- Cluster Admin GUI—This is the preferred method of operation. Refer to the Section “Adding and removing a node from CIM” for the GUI procedures.
- CLI—Refer to the Chapter “Manual pages” for complete details on the CLI options and arguments, some of which are described in this section. For more complete details on CLI options and arguments, refer to the manual page. The commands can also be found in the following directory:

```
/opt/SMAW/SMAWcf/bin
```

CLI

The CIM is configured using the command `rcqconfig(1M)` after CF starts. `rcqconfig(1M)` is run to set up or change CIM configuration. This command is run manually if the cluster is not configured through Cluster Admin.

When `rcqconfig(1M)` is invoked, it checks that the node is part of the cluster. When the `rcqconfig(1M)` command is invoked without any option, after the node joins the cluster, it checks if any configuration is present in the `CFReg.database`. If there is none, it returns as error. This is done as part of the GUI configuration process.

`rcqconfig(1M)` configures a quorum set of nodes, among which CF decides the quorum state. `rcqconfig(1M)` is also used to show the current configuration. If `rcqconfig(1M)` is invoked without any configuration changes or with only the `-v` option, `rcqconfig(1M)` will apply any existing configuration to all the nodes in the cluster. It will then start or restart the quorum operation. `rcqconfig(1M)` can be invoked from the command line to configure or to start the quorum. It can also be invoked through `cfconfig -l`.

3.2.2 Query of the quorum state

CIM recalculates the quorum state when it is triggered by some node state change. However you can force the CIM to recalculate it by running `rcquery(1M)` at any time. Refer to the Chapter “Manual pages” for complete details on the CLI options and arguments.

`rcquery(1M)` queries the state of quorum and gives the result using the return code. It also gives you readable results if the verbose option is given. `rcquery(1M)` returns `True` if the states of all the nodes in the quorum set of nodes are known. If the state of any node is unknown, then it returns `false`.

`rcquery(1M)` exits with a status of zero when a quorum exists, and it exits with a status of 1 when a quorum does not exist. If an error occurs during the operation, then it exits with any other non-zero value other than 1.

3.2.3 Reconfiguring quorum

Refer to the Section “Adding and removing a node from CIM” for the GUI procedures.

CLI

The configuration can be changed at any time and is effective immediately. When a new node is added to the quorum set of nodes, the node being added must be part of the cluster so as to guarantee that the new node also has the same quorum configuration. Removing a node from the quorum set can be done without restriction.

When the configuration information is given to the command `rcqconfig(1M)` as arguments, it performs the transaction to `CFREG` to update the configuration information. The rest of the configuration procedure is the same. Until CIM is successfully configured and gets the initial state of the quorum, CIM has to respond with the quorum state of `False` to all queries.

Examples

Display the states of all the nodes in the cluster as follows:

```
fuji2# cftool -n
Node                Number      State      Os
Cpu
  fuji2              1          UP        Solaris
  fuji3              1          UP        Solaris
```

Display the current quorum configuration as follows:

```
fuji2# rcqconfig -g
```

Nothing is returned, since all nodes have been deleted from the quorum.

Add new nodes in a quorum set of nodes as follows:

```
fuji2# rcqconfig -a fuji2 fuji3
```

Display the current quorum configuration parameters as follows:

```
fuji2# rcqconfig -g
QUORUM_NODE_LIST= fuji2 fuji3
```

Delete nodes from a quorum set of nodes as follows:

```
fuji2# rcqconfig -d fuji2
```

Display the current quorum configuration parameters after one node is deleted as follows:

```
fuji2# rcqconfig -g
QUORUM_NODE_LIST= fuji3
```

The results of this query can only be `True` or `False`. `rcquery(1M)` will return `True` if the states of all the nodes in the quorum set of nodes are up. If any one of the nodes is down, then it will return `False`. In this case, since the cluster is still up and running, the results of `rcquery(1M)` will be set to `True`.

Add a new node, `fuji10` (which is not in the cluster), in a quorum set of nodes as follows:

```
fuji2# rcqconfig -a fuji2 fuji3 fuji10
Cannot add node fuji10 that is not up.
```


Since CF only configured the cluster to consist of `fuji2` and `fuji3`, `fuji10` does not exist. The quorum set remains empty.

```
fuji2# rcqconfig -g
```

Nothing will be returned, since no quorum configuration has been done.

4 Cluster resource management

This chapter discusses the Resource Database, which is a synchronized clusterwide database, holding information specific to several PRIMECLUSTER products.

This chapter discusses the following:

- The Section “Overview” introduces cluster resource management.
- The Section “Kernel parameters for Resource Database” discusses the default values of the Solaris OE kernel which have to be modified when the Resource Database is used.
- The Section “Resource Database configuration” details how to set up the Resource Database for the first time on a new cluster.
- The Section “Registering hardware information” explains how to register hardware information in the Resource Database.
- The Section “Start up synchronization” discusses how to implement a start up synchronization procedure for the Resource Database.
- The Section “Adding a new node” describes how to add a new node to the Resource Database.

4.1 Overview


The cluster Resource Database is a dedicated database used by some PRIMECLUSTER products. You must configure the Resource Database if you are using GDS or GFS. Fujitsu customers should always configure the Resource Database since it is used by many products from Fujitsu.

If you do not need to configure the Resource Database, then you can skip this chapter.

The Resource Database is intended to be used only by PRIMECLUSTER products. It is not a general purpose database which a customer could use for their own applications.


4.2 Kernel parameters for Resource Database

The default values of the Solaris operating environment (OE) kernel have to be modified when the Resource Database is used. This section lists the kernel parameters that have to be changed. In the case of kernel parameters that have already been set in the file `/etc/system`, the values recommended here should be added. In the case of kernel parameters that have not been defined in the file `/etc/system`, the values recommended here must be added to the default values.

 The values in the `/etc/system` file do not take effect until the system is rebooted.

If an additional node is added to the cluster, or if more disks are added after your cluster has been up and running, it is necessary to recalculate using the new number of nodes and/or disks after the expansion, change the values in `/etc/system`, and then reboot each node in the cluster.

Refer to the *PRIMECLUSTER Installation Guide (Solaris)* for details on meanings and methods of changing kernel parameters.

 The values used for product and user applications operated under the cluster system must also be reflected in kernel parameter values.

The recommended kernel parameters values are as follows:

- `seminfo_semmni`

Amount to add for Resource Database: 20

- `seminfo_semmns`

Amount to add for Resource Database: 30

- `seminfo_semmnu`

Amount to add for Resource Database: 30

- `shminfo_shmmni` value = 30

Amount to add for Resource Database: 30

- `shminfo_shmseg`

Amount to add to Resource Database: 30

- `shminfo_shmmax`

The value of `shminfo_shmmax` is calculated in the following way:

1. Remote resources:

$DISKS \times (NODES+1) \times 2$

DISKS is the number of shared disks. For disk array units, use the number of logical units (LUN). For devices other than disk array units, use the number of physical disks.

NODES is the number of nodes connected to the shared disks.

2. Local resources:

LOCAL_DISKS: Add up the number of local disks of all nodes in the cluster.

3. Total resources:

Total resources = (remote resources + local resources) x 2776 + 1048576.

4. Selecting the value:

If *shminfo_shmmax* has already been altered by another product (meaning, */etc/system* already has an entry for *shminfo_shmmax*), then set the value of *shminfo_shmmax* to the sum of the current value and the result from Step 3. Should this value be less than 4194394, set *shminfo_shmmax* to 4194394.

If *shminfo_shmmax* has not been altered from the default (meaning, there is no entry for *shminfo_shmmax* in */etc/system*) and the result from Step 3 is greater than 4194394, set *shminfo_shmmax* to the result of Step 3, otherwise set *shminfo_shmmax* to 4194394.

In summary, the formula to calculate the total resources is as follows:

Total resources = (*DISKS* x (*NODES*+1) x 2 + *LOCAL_DISKS*) x 2776 + 1048576 + *current value*.

The algorithm to set *shminfo_shmmax* is as follows:

```
if (Total Resources < 4194394)
then
    shminfo_shmmax =4194394
else
    shminfo_shmmax =Total Resources
endif
```

Example:

Referring to Figure 23, the following example shows how to calculate the total resources.

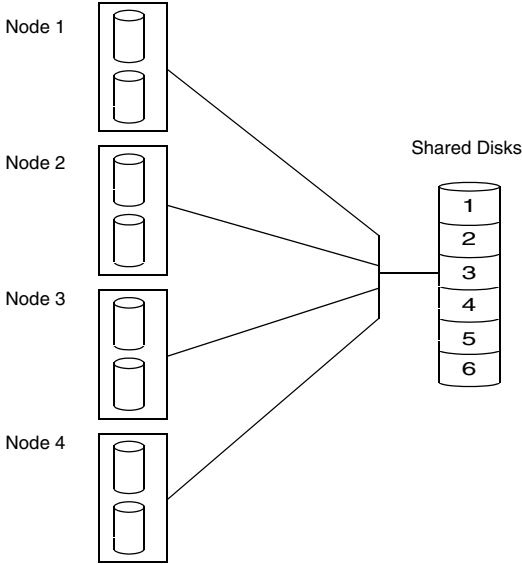


Figure 23: Cluster resource diagram

Referring to Figure 23, calculate the total resources as follows:

1. Remote resources:

$$\begin{aligned} \text{DISKS} &= 6, \text{ NODES} = 4 \\ \text{remote resources} &= 6 \times (4+1) \times 2 = 60 \end{aligned}$$

2. Local resources:

$$\text{local resources} = 2 \times 4 = 8$$

3. Total resources:

$$(60+8) \times 2776 + 1048576 = 1237344$$

4. Selecting the value:

$$\text{current value} = 0$$

The sum of 1237344 and the `current` value is less than 4194394, therefore `shminfo_shmmax` has to be set to 4194394. If the sum of 1237344 and the `Current` Value is more than 4194394, then set `shminfo_shmmax` to the new sum.

4.3 Resource Database configuration

This section discusses how to set up the Resource Database for the first time on a new cluster. The following procedure assumes that the Resource Database has not previously been configured on any of the nodes in the cluster.

If you need to add a new node to the cluster, and the existing nodes are already running the Resource Database, then a slightly different procedure needs to be followed. Refer to the Section “Adding a new node” for details.

Before you begin configuring the Resource Database, you must first make sure that CIP is properly configured on all nodes. The Resource Database uses CIP for communicating between nodes, so it is essential that CIP is working.

The Resource Database also uses the CIP configuration file `/etc/cip.cf` to establish the mapping between the CF node name and the CIP name for a node. If a particular node has multiple CIP interfaces, then only the first one is used. This will correspond to the first CIP entry for a node in `/etc/cip.cf`. It will also correspond to `cip0` on the node itself.

Because the Resource Database uses `/etc/cip.cf` to map between CF and CIP names, it is critical that this file be the same on all nodes. If you used the Cluster Admin CF Wizard to configure CIP, then this will already be the case. If you created some `/etc/cip.cf` files by hand, then you need to make sure that all nodes are specified and they are the same across the cluster.

In general, the CIP configuration is fairly simple. You can use the Cluster Admin CF Wizard to configure a CIP subnet after you have configured CF. If you use the Wizard, then you will not need to do any additional CIP configuration. See the Section “CF, CIP, and CIM configuration” for more details.

After CIP has been configured, you can configure the Resource Database on a new cluster by using the following procedure. This procedure must be done on all the nodes in the cluster.

1. Log in to the node with system administrator authority.

- Verify that the node can communicate with other nodes in the cluster over CIP. You can use the `ping(1M)` command to test CIP network connectivity. The file `/etc/cip.cf` contains the CIP names that you should use in the `ping` command.

If you are using RMS and you have only defined a single CIP subnetwork, then the CIP names will be of the following form:

```
cfnameRMS
```

For example, if you have two nodes in your cluster named `fuji2` and `fuji3`, then the CIP names for RMS would be `fuji2RMS` and `fuji3RMS`, respectively. You could then run the following commands:

```
fuji2# ping fuji3RMS
```

```
fuji3# ping fuji2RMS
```

This tests the CIP connectivity.

Be careful if you have configured multiple CIP interfaces for some nodes. In this case, only the first CIP interface on a node will be used by the Resource Database. This first interface may not necessarily be the one used by RMS.

- Execute the `clsetup(1M)` command. When used for the first time to set up the Resource Database on a node, it is called without any arguments as follows:

```
# /etc/opt/FJSVcluster/bin/clsetup
```

- Execute the `clgettree(1)` command to verify that the Resource Database was successfully configured on the node, as shown in the following:

```
##/etc/opt/FJSVcluster/bin/clgettree
```

The command should complete without producing any error messages, and you should see the Resource Database configuration displayed in a tree format.

For example, on a two-node cluster consisting of `fuji2` and `fuji3`, the `clgettree(1)` command might produce output similar to the following:

```
Cluster 1 cluster
  Domain 2 Domain0
    Shared 7 SHD_Domain0
      Node 3 fuji2 UNKNOWN
      Node 5 fuji3 UNKNOWN
```

If you need to change the CIP configuration to fix the problem, you will also need to run the `clinitreset(1M)` command and start the information process over.

The format of `clgettree(1)` is more fully described in its manual page. For the purpose of setting up the cluster, you need to check the following:

- Each node in the cluster should be referenced in a line that begins with the word `Node`.
- The `clgettree(1M)` output must be identical on all nodes.

If either of the above conditions are not met, then it is possible that you may have an error in the CIP configuration. Double-check the CIP configuration using the methods described earlier in this section. The actual steps are as follows:

1. Make sure that CIP is properly configured and running.
2. Run `clinitreset(1M)` on all nodes in the cluster.
3. Reboot each node.
4. Rerun the `clsetup(1M)` command on each node.
5. Use the `clgettree(1)` command to verify the configuration.

4.4 Registering hardware information



With RCVM, you do not need to register the shared disk unit in the Resource Database.

This section explains how to register hardware information in the Resource Database.

You can register the following hardware in the Resource Database by executing the `clautoconfig(1M)` command:

- Shared disk unit
- Network interface card
- Line switching unit

The command automatically detects the information. Refer to the Chapter “Manual pages” for additional details on this command.

4.4.1 Prerequisite for EMC Symmetrix

The following is an outline of prerequisite steps for EMC Symmetrix storage units:

1. PowerPath is required to use EMC Symmetrix.
2. When EMC Symmetrix is connected, devices such as native device configuring EMC power devices, BCV (Business Continuance Volume), R2 (SRDF target device), GK (GateKeeper), and CKD (Count Key Data) are ineligible for automatic resource registration.
3. Create a list of devices that are ineligible for automatic resource registration (an excluded device list) on all nodes after completing setups of BCV, GK and EMC PowerPath.

Set up the excluded device list in the `/etc/opt/FJSVcluster/etc/diskinfo` file as follows:

- a) List all the disks that should not be used for cluster services, except for BCV, R2, GK, CKD, and volume managed by Volume Logix in this exclusive device list.
- b) You can differentiate which disk is BCV, R2, GK, or CKD by executing the `syminq(1M)` command provided in SymCLI. Execute the `syminq(1M)` command and describe all devices (`cCtTdD`, `emcpower N`) indicated as BCV, R2, GK, or CKD in the excluded device list where the options are as follows:

C is the controller number

T is the target ID

D is the disk number

N is the `emcpower` device number

- c) Volume Configuration Management Data Base (VCMDB), used for Volume Logix, is not output by executing `syminq(1M)`. Check with an EMC customer engineer or a system administrator who set up the Volume Logix about which disk should not be used for cluster services, and add it to the list.

An example of the `/etc/opt/FJSVcluster/etc/diskinfo` file that has done its setup is as follows:

```
# cat /etc/opt/FJSVcluster/etc/diskinfo

c1t0d16
c1t0d17
c1t0d18
c1t0d19
.....
emcpower63
emcpower64
emcpower65
emcpower66
```

To create an example file, the `awk` script is provided for simplified setup. Take the following steps to edit the `diskinfo` file using this script:

```
# syminq | nawk -f /etc/opt/FJSVcluster/sys/clmkdiskinfo >
/etc/opt/FJSVcluster/etc/diskinfo
```

You need to use the `syminq` command path that is specified at the time of SymCLI installation. Normally, it should be `/usr/symcli/bin/syminq`.



Do not describe BCV and R2 devices used for GDS Snapshot in the excluded list. Also, do not describe the native device configuring BCV and R2 devices. For details of GDS Snapshot, refer to the *PRIMECLUSTER Global Disk Services Configuration and Administration Guide (Solaris)*.

If you do not include the R2 device of the SRDF pair in the excluded device list, you need to make the SRDF pair in the `split` state to run automatic resource registration.

4.4.2 Multi-path automatic generation

You can make a logic path generate automatically on all nodes in a PRIMECLUSTER system when you use a Multi-Path Disk Control Load Balance (MPLB) option.

With this feature, the logic path of the shared disk unit is managed, and the instance number of the logic path is identical on all nodes. This instance number is required on all the nodes, so we recommend that you use this feature for the logic path generation.

4.4.3 Automatic resource registration

This section explains how to register the detected hardware in the Resource Database

The registered network interface card should be displayed in the plumb-up state as a result of executing the `ifconfig(1M)` command.

Do not modify the volume name registered in VTOC using the `format(1M)` command after automatic resource registration. The volume name is required when the shared disk units are automatically detected.

The following prerequisites should be met:

- The Resource Database setup is done.
- Hardware is connected to each node.
- All nodes are started in the multi-user mode.

Take the following steps to register hardware in the Resource Database. This should be done on an arbitrary node in a cluster system.

1. Log in with system administrator access privileges.
2. Execute the `clautoconfig(1M)` command, using the following full path:

```
# /etc/opt/FJSCluster/bin/clautoconfig -r
```
3. Confirm registration.

Execute the `clgettree(1)` command for confirmation as follows:

```
# /etc/opt/FJSCluster/bin/clgettree
```

```
Cluster 1 cluster0
  Domain 2 domain0
    Shared 7 SHD_domain0
      SHD_DISK 9 shd001 UNKNOWN
      DISK 11 c1t1d0 UNKNOWN node0
      DISK 12 c2t2d0 UNKNOWN node1
      SHD_DISK 10 shd002 UNKNOWN
      DISK 13 c1t1d1 UNKNOWN node0
      DISK 14 c2t2d1 UNKNOWN node1
    Node 3 node0 ON
      Ethernet 20 hme0 UNKNOWN
      DISK 11 c1t1d0 UNKNOWN
      DISK 13 c1t1d1 UNKNOWN node0
    Node 5 node1 ON
      Ethernet 21 hme0 UNKNOWN
```

```
DISK 12 c2t2d0 UNKNOWN
DISK 14 c2t2d1 UNKNOWN
```

Reference

When deleting the resource of hardware registered by automatic registration, the following commands are used. Refer to the manual page for details of each command.

- `cldeIdevice(1M)`—Deletes the shared disk resource
- `cldeIrsc(1M)`—Deletes the network interface card resource
- `cldeIswursc(1M)`—Deletes the line switching unit resource

4.5 Start up synchronization

A copy of the Resource Database is stored locally on each node in the cluster. When the cluster is up and running, all of the local copies are kept in sync. However, if a node is taken down for maintenance, then its copy of the Resource Database may be out of date by the time it rejoins the cluster. Normally, this is not a problem. When a node joins a running cluster, then its copy of the Resource Database is automatically downloaded from the running cluster. Any stale data that it may have had is thus overwritten.

There is one potential problem. Suppose that the entire cluster is taken down before the node with the stale data had a chance to rejoin the cluster. Then suppose that all nodes are brought back up again. If the node with the stale data comes up long before any of the other nodes, then its copy of the Resource Database will become the master copy used by all nodes when they eventually join the cluster.

To avoid this situation, the Resource Database implements a start up synchronization procedure. If the Resource Database is not fully up and running anywhere in the cluster, then starting the Resource Database on a node will cause that node to enter into a synchronization phase. The node will wait up to `StartingWaitTime` seconds for other nodes to try to bring up their own copies of the Resource Database. During this period, the nodes will negotiate among themselves to see which one has the latest copy of the Resource Database. The synchronization phase ends when either all nodes have been accounted for or `StartingWaitTime` seconds have passed. After the synchronization period ends, the latest copy of the Resource Database that was found during the negotiations will be used as the master copy for the entire cluster.

The default value for `StartingWaitTime` is 60 seconds.

This synchronization method is intended to cover the case where all the nodes in a cluster are down, and then they are all rebooted together. For example, some businesses require high availability during normal business hours, but power their nodes down at night to reduce their electric bill. The nodes are then powered up shortly before the start of the working day. Since the boot time for each node may vary slightly, the synchronization period of up to `StartingWaitTime` ensures that the latest copy of the Resource Database among all of the booting nodes is used.

Another important scenario in which all nodes may be booted simultaneously involves the temporary loss and then restoration of power to the lab where the nodes are located.

However, for this scheme to work properly, you must verify that all nodes in the cluster have boot times that differ by less than `StartingWaitTime` seconds. Furthermore, you might need to modify the value of `StartingWaitTime` to a value that is appropriate for your cluster.

Modify the value of `StartingWaitTime` as follows:

1. Start up all of the nodes in your cluster simultaneously. You should probably start the nodes from a cold power on.
2. After the each node has come up, look in `/var/adm/messages` for message number 2200. This message is output by the Resource Database when it first starts. For example, enter the following command:

```
# grep 2200 /var/adm/messages
Feb 23 19:00:41 fuji2 dcmond[407]: [ID 888197 daemon.notice]
FJSVcluster: INFO: DCM: 2200: Cluster configuration
management facility initialization started.
```

Compare the timestamps for the messages on each node and calculate the difference between the fastest and the slowest nodes. This will tell you how long the fastest node has to wait for the slowest node.

3. Check the current value of `StartingWaitTime` by executing the `clsetparam(1M)` command on any of the nodes. For example, enter the following command:

```
# /etc/opt/FJSVcluster/bin/clsetparam -p StartingWaitTime 60
```

The output above shows that `StartingWaitTime` is set to 60 seconds.

4. If there is a difference in start up times found in Step 2, the `StartingWaitTime`, or if the two values are relatively close together, then you should increase the `StartingWaitTime` parameter. You can do this by running the `clsetparam(1M)` command on any node in the cluster. For example, enter the following command:

```
# /etc/opt/FJSVcluster/bin/clsetparam -p StartingWaitTime 300
```

This sets the `StartingWaitTime` to 300 seconds.

Refer to the Chapter “Manual pages” for more details on the possible values for `StartingWaitTime`.

4.5.1 Start up synchronization and the new node

After the Resource Database has successfully been brought up in the new node, then you need to check if the `StartingWaitTime` used in start up synchronization is still adequate. If the new node boots much faster or slower than the other nodes, then you may need to adjust the `StartingWaitTime` time.

4.6 Adding a new node

If you have a cluster where the Resource Database is already configured, and you would like to add a new node to the configuration, then you should follow the procedures in this section. You will need to make a configuration change to the currently running Resource Database and then configure the new node itself. The major steps involved are listed below:

1. Back up the currently running Resource Database. A copy of the backup is used in a later step to initialize the configuration on the new node. It also allows you to restore your configuration to its previous state if a serious error is encountered in the process.
2. Reconfigure the currently running Resource Database so it will recognize the new node.
3. Initialize the Resource Database on the new node.
4. Verify that the `StartingWaitTime` is sufficient for the new node, and modify this parameter if necessary.

Figure 24 shows these steps as a flow chart.

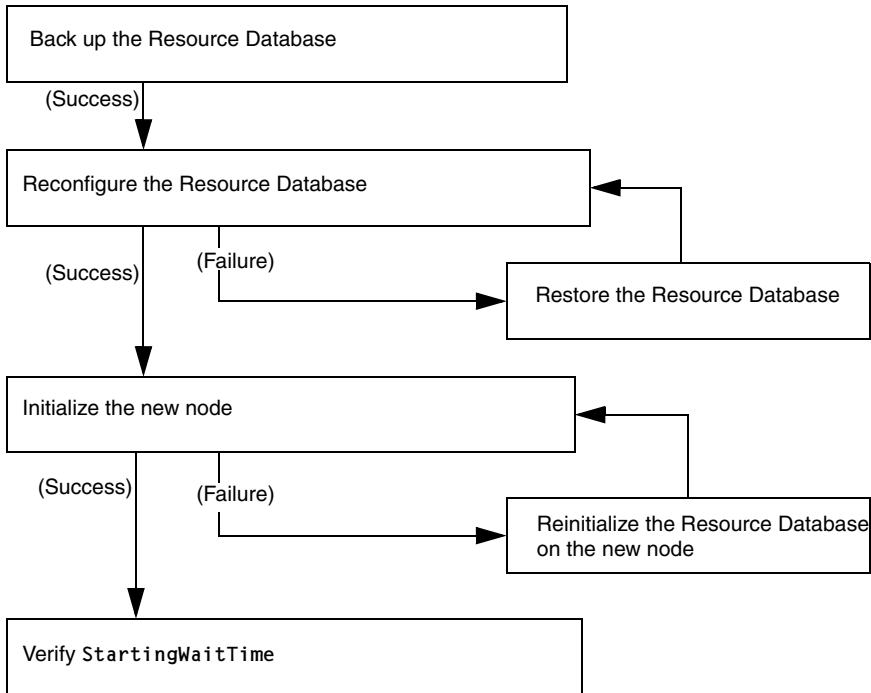


Figure 24: Adding a new node

The sections that follow describe each step in more detail.

4.6.1 Backing up the Resource Database

Before you add a new node to the Resource Database, you should first back up the current configuration. The backup will be used later to help initialize the new node. It is also a safeguard. If the configuration process is unexpectedly interrupted by a panic or some other serious error, then you may need to restore the Resource Database from the backup.

i The configuration process itself should not cause any panics. However, if some non-PRIMECLUSTER software panics or if the SF/SCON forces a panic because of a CF cluster partition, then the Resource Database configuration process could be so severely impacted that a restoration from the backup would be needed.

i The restoration process requires all nodes in the cluster to be in single user mode.

Since the Resource Database is synchronized across all of its nodes, the backup can be done on any node in the cluster where the Resource Database is running. The steps for performing the backup are as follows:

1. Log onto any node where the Resource Database is running with system administrator authority.
2. Run the command `clbackuprdb(1M)` to back the Resource Database up to a file. The syntax is as follows:

```
/etc/opt/FJsvcluster/bin/clbackuprdb -f file
```

For example:

```
# /etc/opt/FJsvcluster/bin/clbackuprdb -f /mydir/backup_rdb
```

`clbackuprdb(1M)` stores the Resource Database as a compressed tar file. Thus, in the above example, the Resource Database would be stored in `/mydir/backup_rdb.tar.Z`.

Make sure that you do not place the backup in a directory whose contents are automatically deleted upon reboot (for example, `/tmp`).

i The hardware configuration must not change between the time a backup is done and the time that the restore is done. If the hardware configuration changes, you will need to take another backup. Otherwise, the restored database would not match the actual hardware configuration, and new hardware resources would be ignored by the Resource Database.

4.6.2 Reconfiguring the Resource Database

After you have backed up the currently running Resource Database, you will need to reconfigure the database to recognize the new node. Before you do the reconfiguration, however, you need to perform some initial steps.

After these initial steps, you should reconfigure the Resource Database. This is done by running the `clsetup(1M)` command on any of the nodes which is currently running the Resource Database. Since the Resource Database is synchronized across all of its nodes, the reconfiguration takes effect on all nodes. The steps are as follows:

1. Log in to any node where the Resource Database is running. Log in with system administrator authority.
2. If this node is not the same one where you made the backup, then copy the backup to this node. Then run the `clsetup(1M)` command with the `-a` and `-g` options to reconfigure the database. The syntax in this case is as follows:

```
/etc/opt/FJSVcluster/bin/clsetup -a cfname -g file
```

cfname is the CF name of the new node to be added, and *file* is the name of the backup file without the `.tar.Z` suffix.

For example, suppose that you want to add a new node whose CF name is `fujii4` to a cluster. If the backup file on an existing node is named `/mydir/rdb.tar.Z`, then the following command would cause the Resource Database to be configured for the new node:

```
# cd /etc/opt/FJSVcluster/bin/  
# ./clsetup -a fujii4 -g /mydir/rdb.tar.Z
```

If `clsetup(1M)` is successful, then you should immediately make a new backup of the Resource Database. This backup will include the new node in it. Be sure to save the backup to a place where it will not be lost upon a system reboot.

If an unexpected failure such as a panic occurs, then you may need to restore the Resource Database from an earlier backup. See the Section “Restoring the Resource Database” for details.

3. To verify if the reconfiguration was successful, run the `clgettree(1)` command. Make sure that the new node is displayed in the output for that command. If it is not present, then recheck the CIP configuration to see if it omitted the new node. If the CIP configuration is in error, then you will need to do the following to recover:
 - a) Correct the CIP configuration on all nodes. Make sure that CIP is running with the new configuration on all nodes.

- b) Restore the Resource Database from backup.
- c) Rerun the `clsetup(1M)` command to reconfigure the Resource Database.

4.6.3 Configuring the Resource Database on the new node

After the Resource Database has been reconfigured on the existing nodes in the cluster, you are ready to set up the Resource Database on the new node itself.

The first step is to verify the CIP configuration on the new node. The file `/etc/cip.cf` should reference the new node. The file should be the same on the new node as it is on existing nodes in the cluster. If you used the Cluster Admin CF Wizard to configure CF and CIP for the new node, then CIP should already be properly configured.

You should also verify that the existing nodes in the cluster can ping the new node using the new node's CIP name. If the new node has multiple CIP subnetworks, then recall that the Resource Database only uses the first one that is defined in the CIP configuration file.

After verifying that CIP is correctly configured and working, then you should do the following:

1. Log in to the new node with system administrator authority.
2. Copy the latest Resource Database backup to the new node. This backup was made in Step 2 of the second list in the Section "Reconfiguring the Resource Database".
3. Run the command `clsetup(1M)` with the `-s` option. The syntax for this case is as follows:

```
/etc/opt/FJSVcluster/bin/clsetup -s file
```

file is the name of the backup file.

If we continue our example of adding `fujii4` to the cluster and we assume that the backup file `rdb.tar.Z` was copied to `/mydir`, then the command would be as follows:

```
# /etc/opt/FJSVcluster/bin/clsetup -s /mydir/rdb.tar.Z
```

If the new node unexpectedly fails before the `clsetup(1M)` command completes, then you should execute the `clinitreset(1M)` command. After `clinitreset(1M)` completes, you must reboot the node and then retry the `clsetup(1M)` command which was interrupted by the failure.

If the `clsetup(1M)` command completes successfully, then you should run the `clgettree(1)` command to verify that the configuration has been set-up properly. The output should include the new node. It should also be identical to output from `clgettree(1)` run on an existing node.

If the `clgettree(1)` output indicates an error, then recheck the CIP configuration. If you need to change the CIP configuration on the new node, then you will need to do the following on the new node after the CIP change:

- a) Run `clinitreset(1M)`.
- b) Reboot.
- c) Rerun the `clsetup(1M)` command described above.

4.6.4 Adjusting StartingWaitTime

After the Resource Database has successfully been brought up in the new node, then you need to check if the `StartingWaitTime` used in startup synchronization is still adequate. If the new node boots much faster or slower than the other nodes, then you may need to adjust the `StartingWaitTime` time. Refer to the Section “Start up synchronization” for further information.

4.6.5 Restoring the Resource Database

The procedure for restoring the Resource Database is as follows:

1. Copy the file containing the Resource Database to all nodes in the cluster.
2. Log in to each node in the cluster and shut it down with the following command:

```
# /usr/sbin/shutdown -y -i0
```

3. Reboot each node to single user mode with the following command:

```
{0} ok boot -s
```



The restore procedure requires that all nodes in the cluster must be in single user mode.

4. Mount the local file systems on each node with the following command:

```
# mountall -l
```

5. Restore the Resource Database on each node with the `clrestorerdb(1M)` command. The syntax is:

```
# clrestorerdb -f file
```

file is the backup file with the `.tar.Z` suffix omitted.

For example, suppose that a restoration was being done on a two-node cluster consisting of nodes `fuji2` and `fuji3`, and that the backup file was copied to `/mydir/backup_rdb.tar.Z` on both nodes. The command to restore the Resource Database on `fuji2` and `fuji3` would be as follows:

```
fuji2# cd /etc/opt/FJSVcluster/bin/
```

```
fuji2# ./clrestorerdb -f /mydir/backup_rdb.tar.Z
```

```
fuji3# cd /etc/opt/FJSVcluster/bin/
```

```
fuji3# ./clrestorerdb -f /mydir/backup_rdb.tar.Z
```

6. After Steps 1 through 5 have been completed on all nodes, then reboot all of the nodes with the following command:

```
# /usr/sbin/shutdown -y -i6
```

5 GUI administration

This chapter covers the administration of features in the Cluster Foundation (CF) portion of Cluster Admin.

This chapter discusses the following:

- The Section “Overview” introduces the Cluster Admin GUI.
- The Section “Starting Cluster Admin GUI and logging in” describes logging in and shows the first screens you will see.
- The Section “Main CF table” describes the features of the main table.
- The Section “Node details” explains how to get detailed information.
- The Section “Displaying the topology table” discusses the topology table, which allows you to display the physical connections in the cluster.
- The Section “Starting and stopping CF” describes how to start and stop CF.
- The Section “Marking nodes DOWN” details how to mark a node DOWN.
- The Section “Using CF log viewer” explains how to use the CF log viewer, including how to view and search `syslog` messages.
- The Section “Displaying statistics” discusses how to display statistics about CF operations.
- The Section “Adding and removing a node from CIM” describes how to add and remove a node from CIM.
- The Section “Unconfigure CF” explains how to use the GUI to unconfigure CF.
- The Section “CIM Override” discusses how to use the GUI to override CIM, which causes a node to be ignored when determining a quorum.

5.1 Overview

CF administration is done by means of the Cluster Admin GUI. The following sections describe the CF Cluster Admin GUI options.

5.2 Starting Cluster Admin GUI and logging in

The first step is to start Web-based Admin View by entering the following URL in a java-enabled browser:

```
http://Management_Server:8081/Plugin.cgi
```

In this example, if `fujj2` is a management server, enter the following:

```
http://fujj2:8081/Plugin.cgi
```

Figure 25 shows the opening screen.

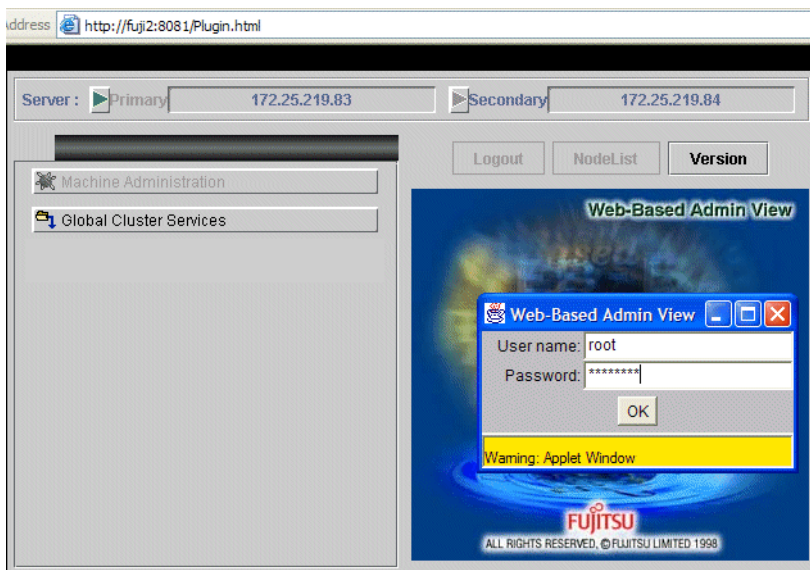


Figure 25: Invoking the Cluster Admin GUI

You can start the Cluster Admin GUI on the primary or secondary management station. Enter the user name and password and click the *OK* button

The top menu appears (see Figure 26).

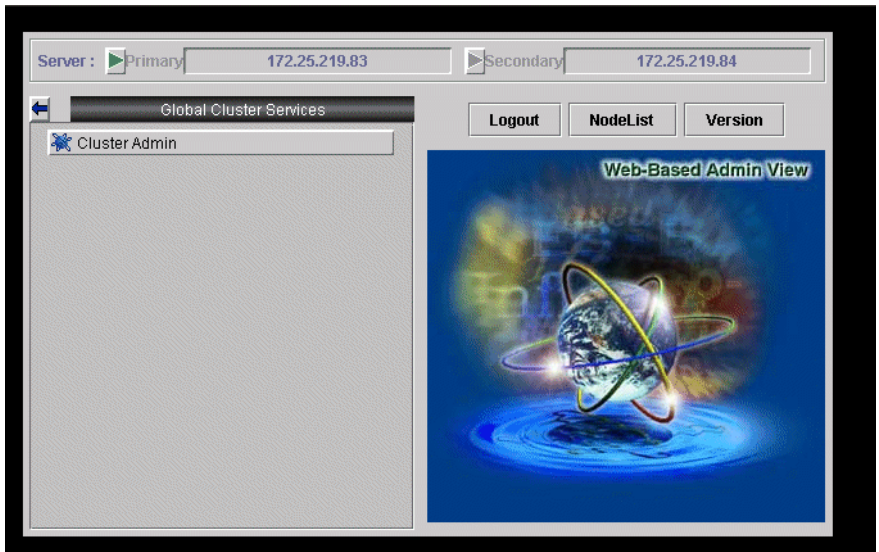


Figure 26: Top menu

Click on the *Cluster Admin* button to go to the next screen.

The *Choose a node for initial connection* menu appears (see Figure 27).

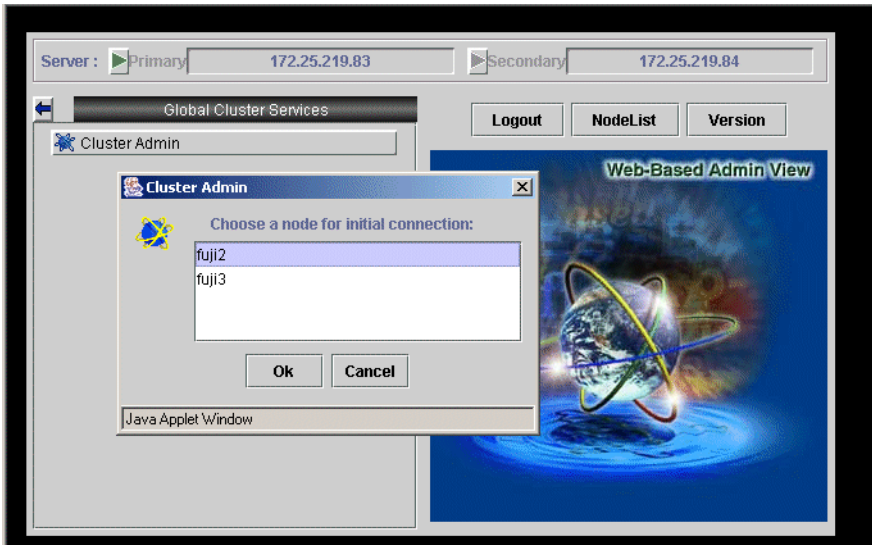


Figure 27: Cluster menu

Select a node and click on the *OK* button.

The Cluster Admin main screen appears (see Figure 28).

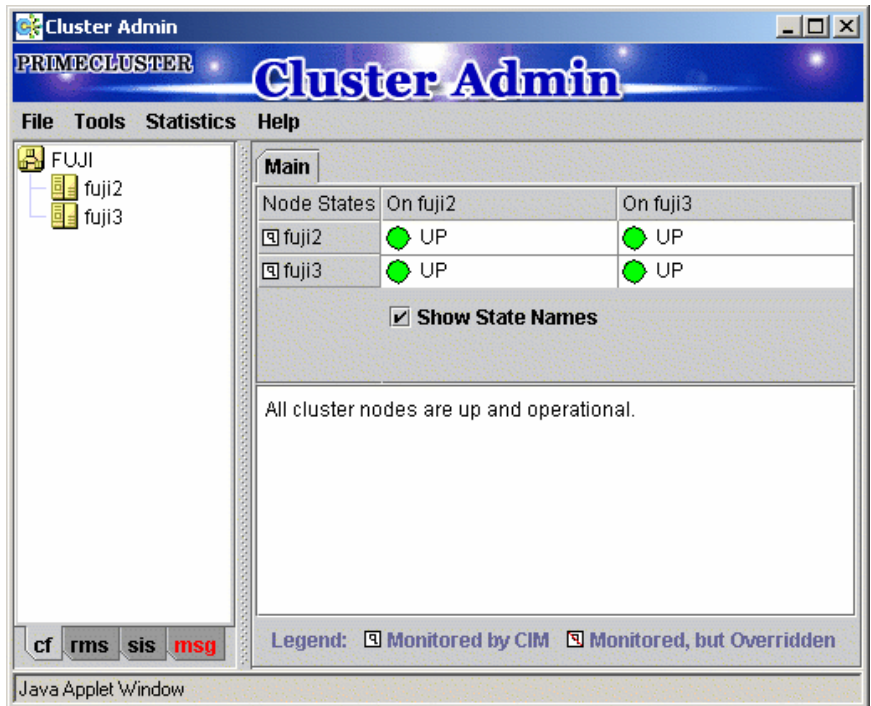


Figure 28: CF main screen

By default, the *cf* tab is selected and the CF main screen is presented. Use the appropriate privilege level while logging in. There are three privilege levels: *root* privileges, *administrative* privileges, and *operator* privileges.

With the *root* privileges, you can perform all actions including configuration, administration and viewing tasks. With *administrative* privileges, you can view as well as execute commands but cannot make configuration changes. With the *operator* privileges, you can only perform viewing tasks.

5.3 Main CF table

When the GUI is first started, or after the successful completion of the configuration wizard, the main CF table will be displayed in the main panel. A tree showing the cluster nodes will be displayed in the left panel. An example of this display is shown in Figure 28.

This table shows a list of the CF states of each node of the cluster as seen by the other nodes in the cluster. For instance, the cell in the second row and first column is the state of `fujii3` as seen by the node `fujii2`.

There are two types of CF states. Local states are the states a node can consider itself in. Remote states are the states a node can consider another node to be in. Table 1 and Table 2 list the different states.

CF state	Description
UNLOADED	The node does not have a CF driver loaded.
LOADED	The node has a CF driver loaded, but is not running.
COMINGUP	The node is in the process of starting and should be UP soon.
UP	The node is up and running normally.
INVALID	The node has an invalid configuration and must be reconfigured.
UNKNOWN	The GUI has no information from this node. This can be temporary, but if it persists, it probably means the GUI cannot contact that node.
UNCONFIGURED	The node is unconfigured.

Table 1: Local states

CF state	Description
UP	The node is up and part of this cluster.
DOWN	The node is down and not in the cluster.
UNKNOWN	The reporting node has no opinion on the reported node.
LEFTCLUSTER	The node has left the cluster unexpectedly, probably from a crash. To ensure cluster integrity, it will not be allowed to rejoin until marked DOWN.

Table 2: Remote states

5.4 Node details

To get detailed information on a cluster node, left-click on the node in the left tree. This replaces the main table with a display of detailed information. (To bring the main table back, left-click on the cluster name in the tree.) The panel displayed is similar to the display in Figure 29.

Main

Node Name:

Status:

Operating System:

CPU:

Cluster Integrity Monitored Host:

Cluster Integrity:

Interfaces used:

Routes

Remote Node	Remote Device	Local Device	State
fuji3	/dewhme1	/dewhme1	UP
fuji3	/dewhme2	/dewhme2	UP

Figure 29: CF node information

Shown are the node's name, its CF state(s), operating system, platform, and the interfaces configured for use by CF. The states listed will be all of the states the node is considered to be in. For instance, if the node considers itself UNLOADED and other nodes consider it DOWN, DOWN/UNLOADED will be displayed.

The bottom part of the display is a table of all of the routes being used by CF on this node. It is possible for a node to have routes go down if a network interface or interconnect fails, while the node itself is still accessible.

5.5 Displaying the topology table

To examine and diagnose physical connectivity in the cluster, select *Tools -> Topology*. This menu option will produce a display of the physical connections in the cluster. This produces a table with the nodes shown along the left side and the interconnects of the cluster shown along the top. Each cell of the table lists the interfaces on that node connected to the interconnect. There is also a checkbox next to each interface showing if it is being used by CF. This table makes it easy to locate cabling errors or configuration problems at a glance.

An example of the topology table is shown in Figure 30.

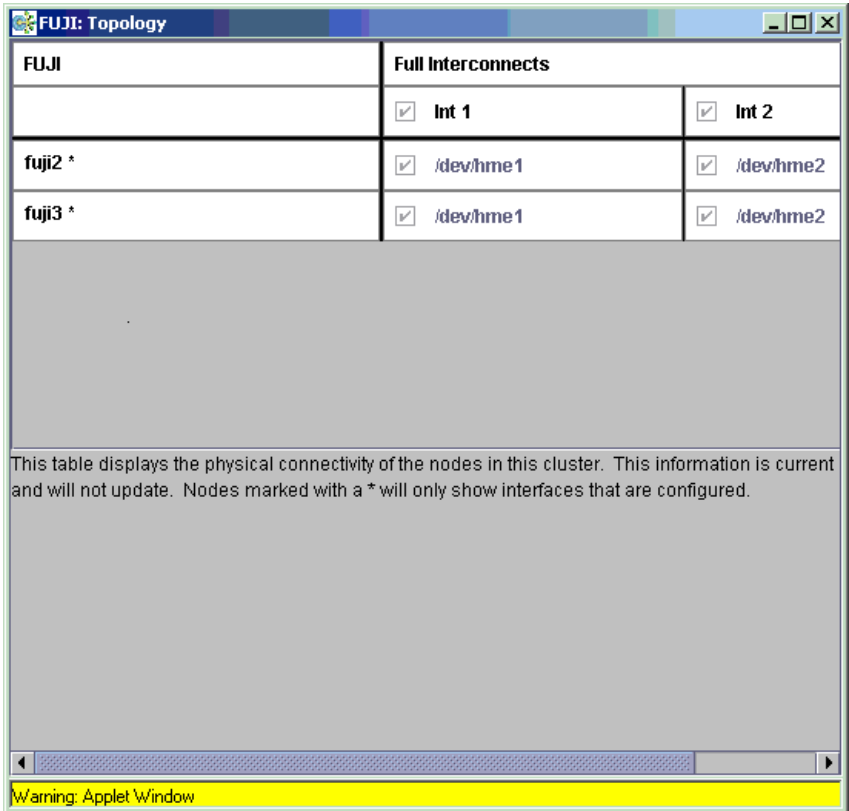


Figure 30: CF topology table

5.6 Starting and stopping CF

There are two ways that you can start or stop CF from the GUI. The first is to simply right-click on a particular node in the tree in the left-hand panel. A state sensitive command pop-up menu for that node will appear. If CF on the selected node is in a state where it can be started (or stopped), then the menu choice *Start CF* (or *Stop CF*) will be offered.

You can also go to the *Tools* pull-down menu and select either *Start CF* or *Stop CF*. A pop-up listing all the nodes where CF may be started or stopped will appear. You can then select the desired node to carry out the appropriate action. Figure 31 shows the pop-up when you select *Start CF*.

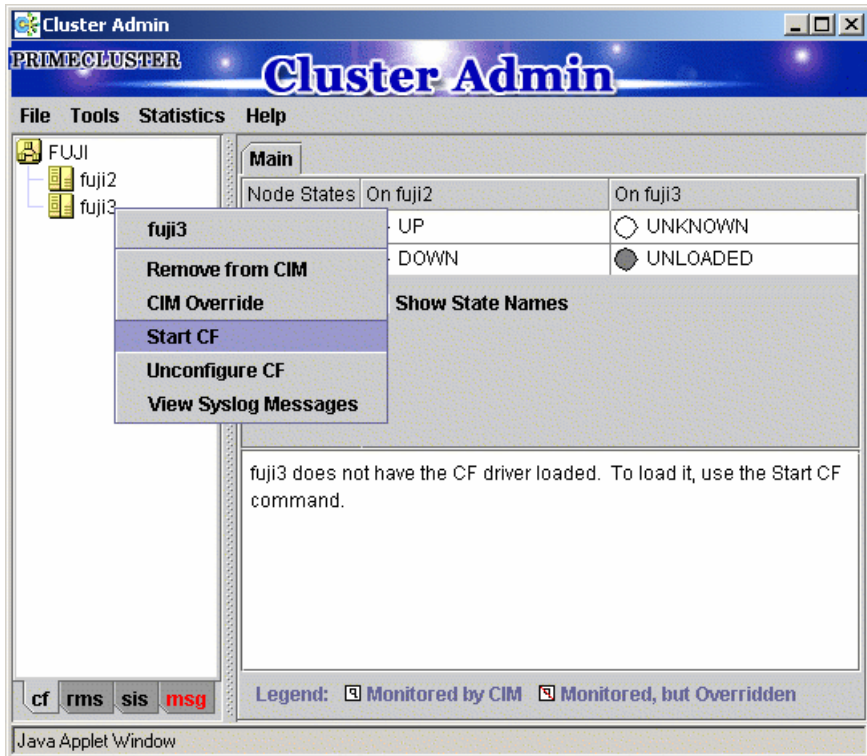


Figure 31: Starting CF

The CF GUI gets its list of CF nodes from the node used for the initial connection screen as shown in Figure 27. If CF is not up and running on the initial connection node, then the CF GUI will not display the list of nodes in the tree in the left panel.

Because of this, when you want to stop CF on multiple nodes (including the initial node) via the GUI, you should make sure that the initial connection node is the very last one on which you stop CF.

Right-click on a CF node name and select *Stop CF* (see Figure 32).

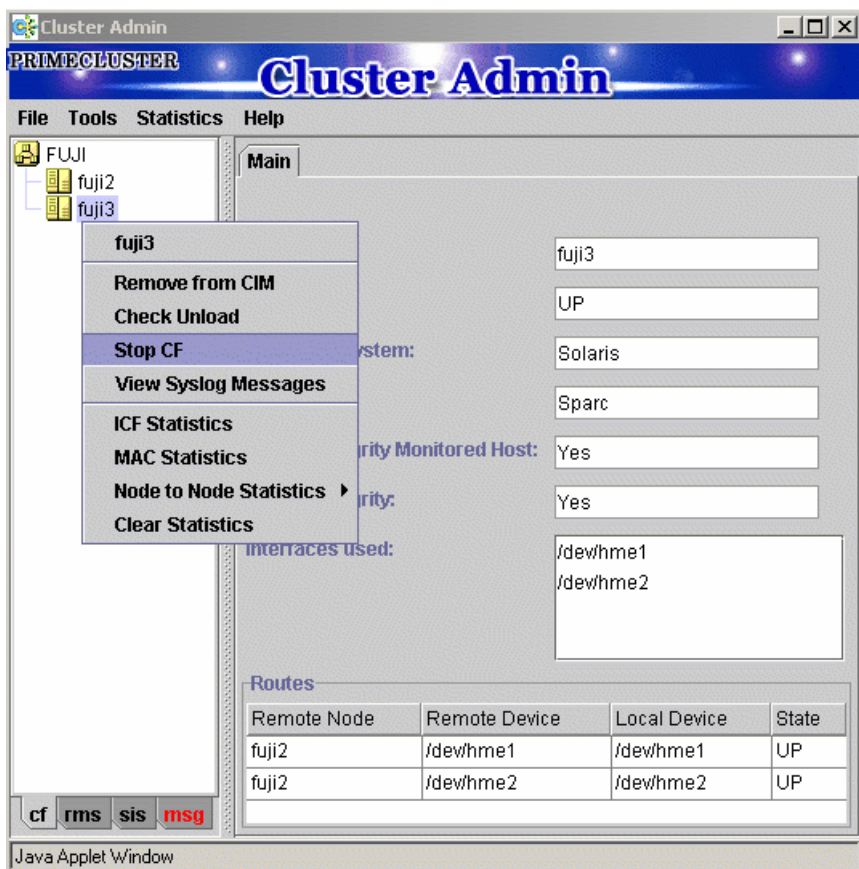


Figure 32: Stop CF

A confirmation pop-up appears (see Figure 33). Choose *Yes* to continue.

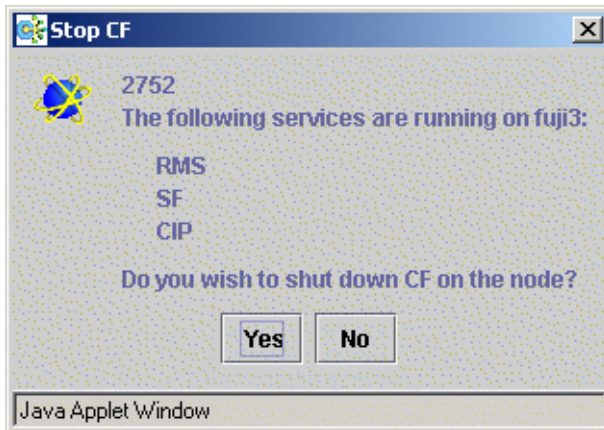


Figure 33: Stopping CF

Before stopping CF, all services that run over CF on that node should first be shut down. When you invoke *Stop CF* from the GUI, it will use the CF dependency scripts to see what services are still running. It will print out a list of these in a pop-up and ask you if you wish to continue. If you do continue, it will then run the dependency scripts to shut down these services. If any service does not shutdown, then the *Stop CF* operation will fail.

i The dependency scripts currently include only PRIMECLUSTER products. If third-party products, for example Oracle OPS, are using PAS or CF services, then the GUI will not know about them. In such cases, the third-party product should be shut down before you attempt to stop CF.

To stop CF on a node, the node's CF state must be UP, COMING UP, or INVALID.

To start CF on a node, its CF state must be UNLOADED or LOADED.

5.7 Marking nodes DOWN

If a node is shut down normally, it is considered DOWN by the remaining nodes. If it leaves the cluster unexpectedly, it will be considered LEFTCLUSTER. To ensure the integrity of the cluster, a node considered LEFTCLUSTER will not be allowed to rejoin the cluster until it has been marked DOWN. The menu option *Tools->Mark Node Down* allows nodes to be marked as DOWN.

To do this, select *Tools->Mark Node Down*. This displays a dialog of all of the nodes that consider another node to be LEFTCLUSTER. Clicking on one of them displays a list of all the nodes that node considered LEFTCLUSTER. Select one and then click *OK*. This clears the LEFTCLUSTER status on that node.

Refer to the Chapter “LEFTCLUSTER state” for more information on the LEFTCLUSTER state.

5.8 Using CF log viewer

The CF log messages for a given node may be displayed by right-clicking on the node in the tree and selecting *View Syslog Messages*.

Alternately, you may go to the *Tools* menu and select *View Syslog Messages*. This brings up a pop-up where you can select the node whose `syslog` messages you would like to view.



The CF log viewer shows only the CF messages that are found in the `syslog`. Non-CF messages in `syslog` are not shown.

Figure 34 shows an example of the CF log viewer.

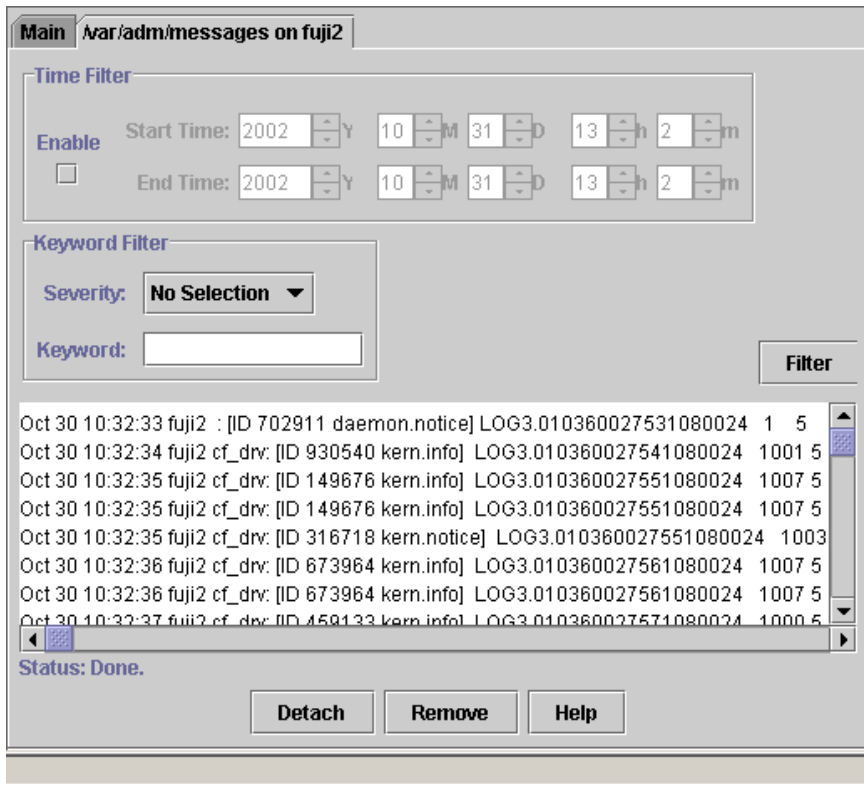


Figure 34: CF log viewer

The `syslog` messages appears in the right-hand panel. If you click on the *Detach* button on the bottom, then the `syslog` window appears as a separate window.

Figure 35 shows the detached CF log viewer window.

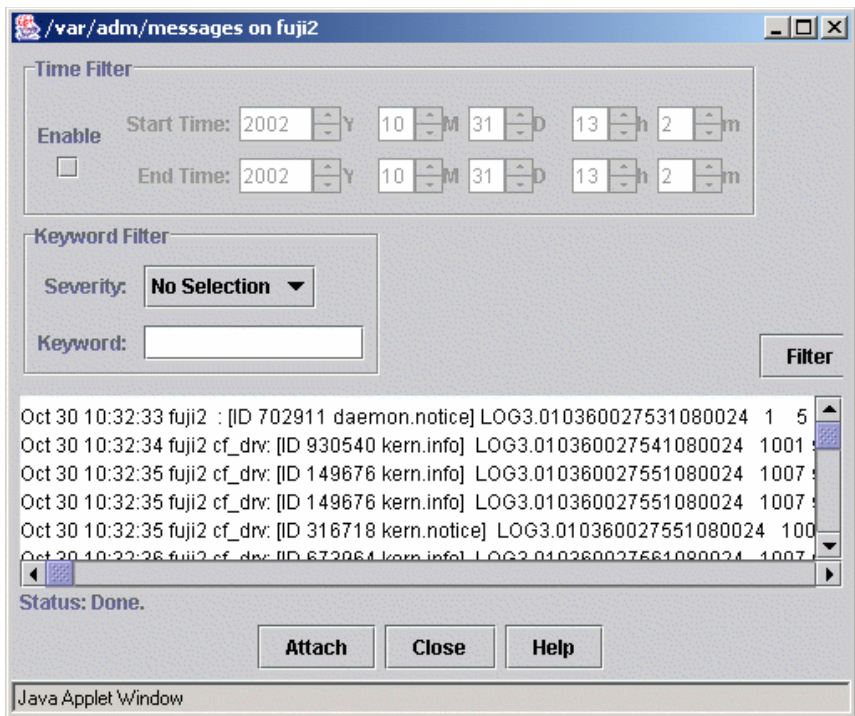


Figure 35: Detached CF log viewer

The CF log viewer has search filters based on date/time/keyword and severity levels.

5.8.1 Search based on time filter

To perform a search based on a start and end time, click the check box for *Enable*, specify the start and end times for the search range, and click on the *Filter* button (see Figure 36).

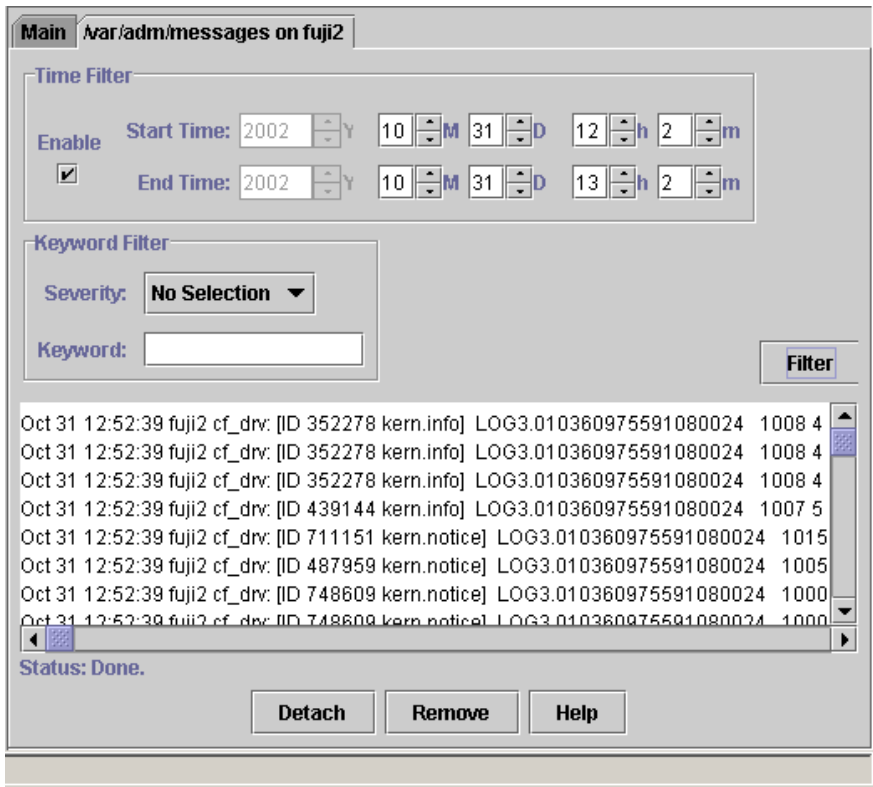


Figure 36: Search based on date/time

5.8.2 Search based on keyword

To perform a search based on a keyword, enter a keyword and click on the *Filter* button (see Figure 37).

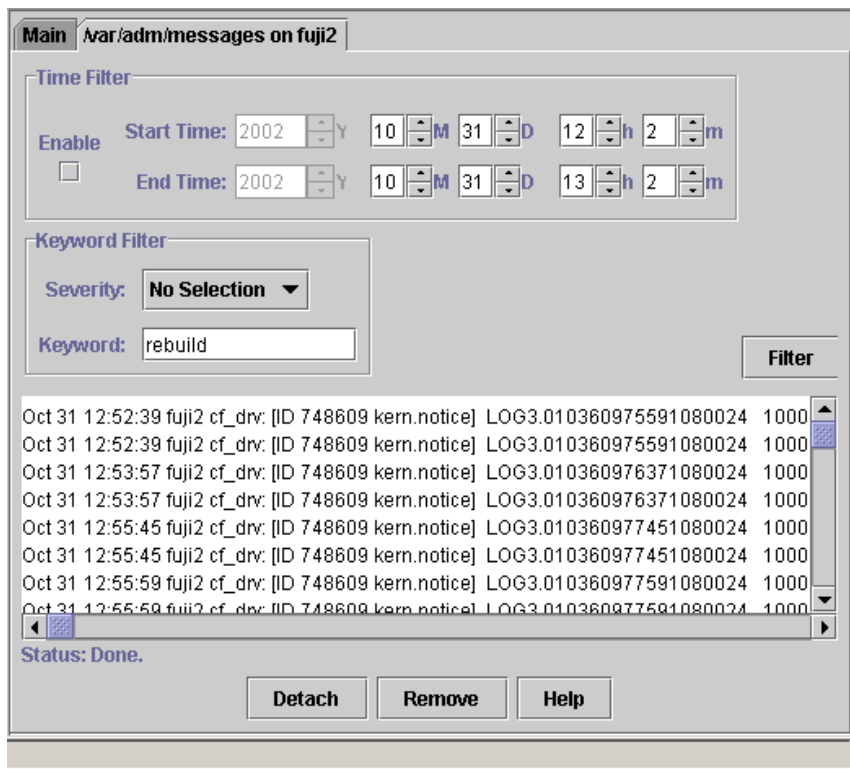


Figure 37: Search based on keyword

5.8.3 Search based on severity levels

To perform a search based severity levels, click on the *Severity* pull-down menu. You can choose from the severity levels shown in Table 3 and click on the *Filter* button. Figure 38 shows the log for a search based on severity level.

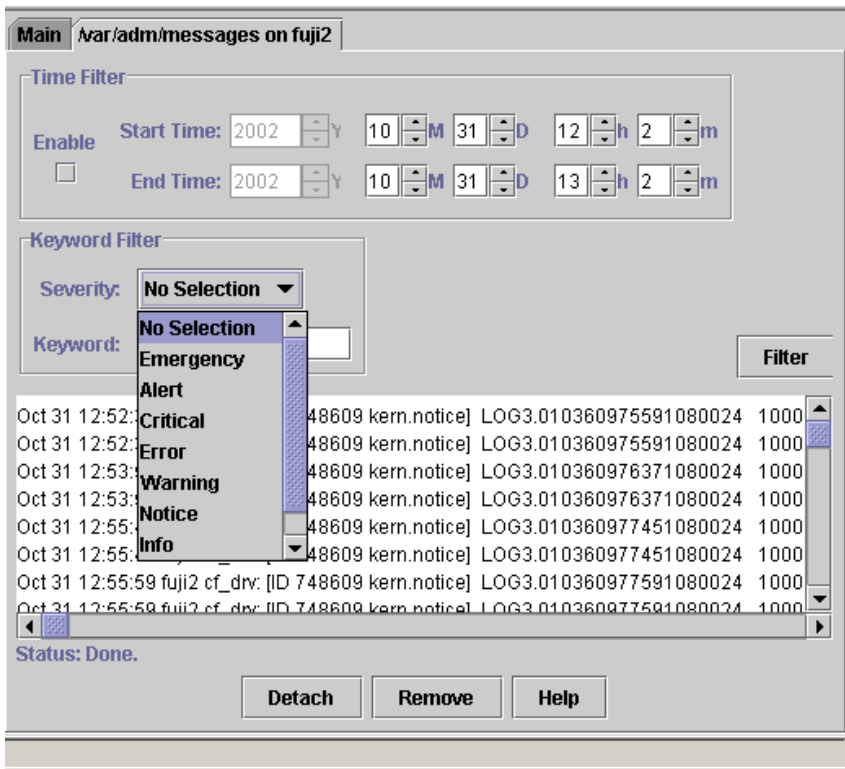


Figure 38: Search based on severity

Severity level	Severity description
<i>Emergency</i>	Systems cannot be used
<i>Alert</i>	Immediate action is necessary
<i>Critical</i>	Critical condition
<i>Error</i>	Error condition

Table 3: CF log viewer severity levels

Severity level	Severity description
<i>Warning</i>	Warning condition
<i>Notice</i>	Normal but important condition
<i>Info</i>	For information
<i>Debug</i>	Debug message

Table 3: CF log viewer severity levels

5.9 Displaying statistics

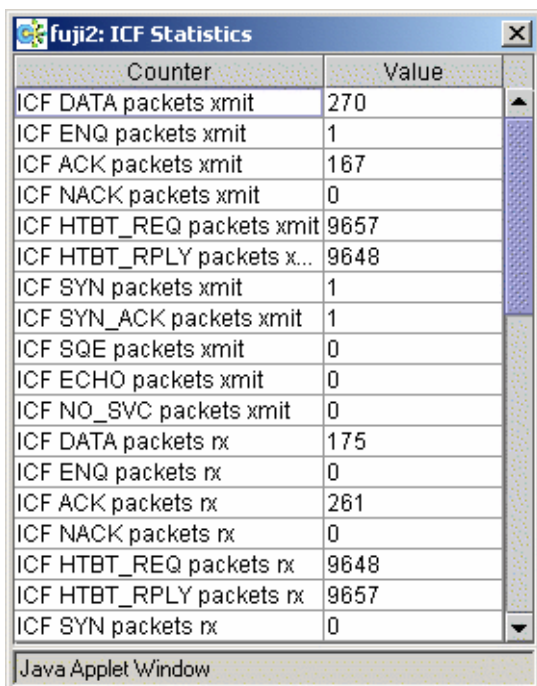
CF can display various statistics about its operation. There are three types of statistics available:

- ICF
- MAC
- Node to Node

To view the statistics for a particular node, right-click on that node in the tree and select the desired type of statistic.

Alternately, you can go to the *Statistics* menu and select the desired statistic. This will bring up a pop-up where you can select the node whose statistics you would like to view. The list of nodes presented in this pop-up will be all nodes whose states are UP as viewed from the login node.

Figure 39 shows the display of ICF Statistics.

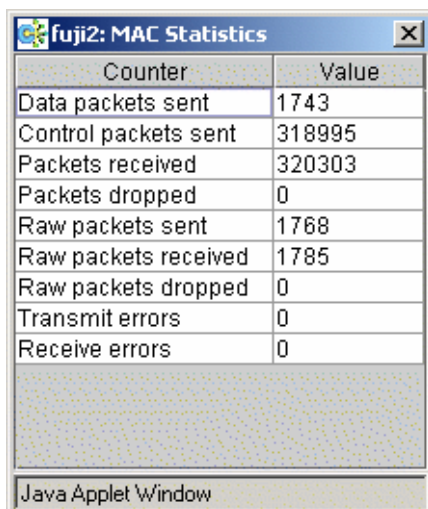


The screenshot shows a Java Applet Window titled "fuji2: ICF Statistics". The window contains a table with two columns: "Counter" and "Value". The table lists various ICF statistics, including packets transmitted (xmit) and received (rx) for different types of data and control packets. The values range from 0 to 9657. The window also has a scrollbar on the right side and a "Java Applet Window" label at the bottom.

Counter	Value
ICF DATA packets xmit	270
ICF ENQ packets xmit	1
ICF ACK packets xmit	167
ICF NACK packets xmit	0
ICF HTBT_REQ packets xmit	9657
ICF HTBT_RPLY packets x...	9648
ICF SYN packets xmit	1
ICF SYN_ACK packets xmit	1
ICF SQE packets xmit	0
ICF ECHO packets xmit	0
ICF NO_SVC packets xmit	0
ICF DATA packets rx	175
ICF ENQ packets rx	0
ICF ACK packets rx	261
ICF NACK packets rx	0
ICF HTBT_REQ packets rx	9648
ICF HTBT_RPLY packets rx	9657
ICF SYN packets rx	0

Figure 39: ICF statistics

Figure 40 shows the display of MAC Statistics,



The screenshot shows a Java Applet window titled "fuji2: MAC Statistics". The window contains a table with two columns: "Counter" and "Value". The table lists various network statistics and their corresponding values. Below the table is a large empty area with a dotted background, and at the bottom, it says "Java Applet Window".

Counter	Value
Data packets sent	1743
Control packets sent	318995
Packets received	320303
Packets dropped	0
Raw packets sent	1768
Raw packets received	1785
Raw packets dropped	0
Transmit errors	0
Receive errors	0

Figure 40: MAC statistics

To display node to node statistics, choose *Node to Node Statistics* and click on the desired node (see Figure 41).

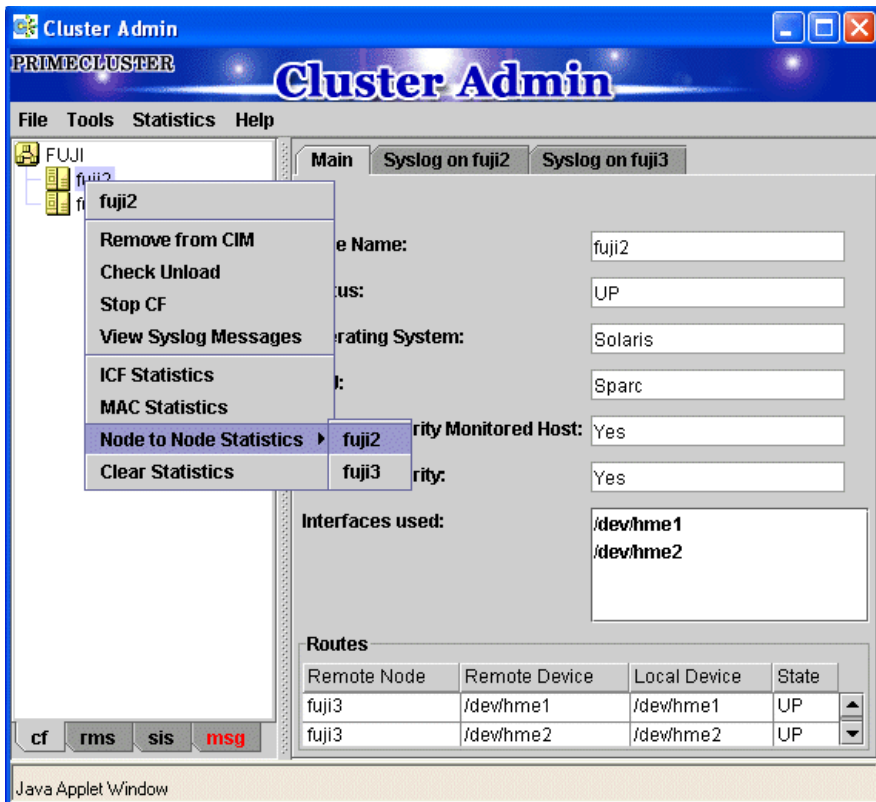
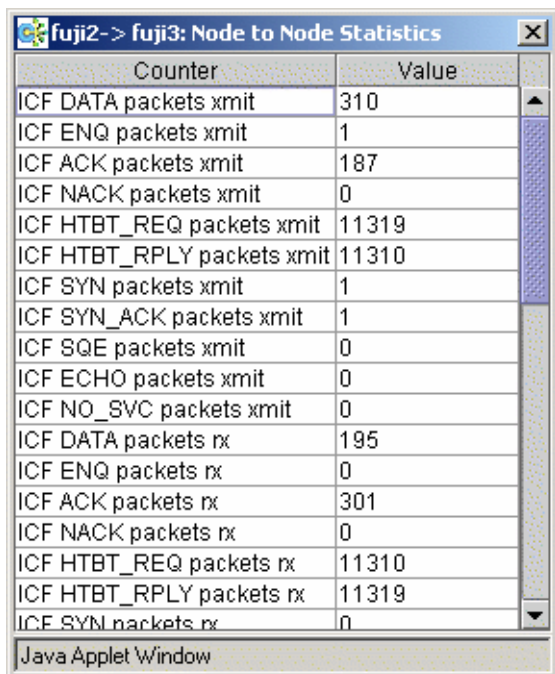


Figure 41: Selecting a node for node to node statistics

The screen for Node to Node Statistics appears (see Figure 42).



Counter	Value
ICF DATA packets xmit	310
ICF ENQ packets xmit	1
ICF ACK packets xmit	187
ICF NACK packets xmit	0
ICF HTBT_REQ packets xmit	11319
ICF HTBT_RPLY packets xmit	11310
ICF SYN packets xmit	1
ICF SYN_ACK packets xmit	1
ICF SQE packets xmit	0
ICF ECHO packets xmit	0
ICF NO_SVC packets xmit	0
ICF DATA packets rx	195
ICF ENQ packets rx	0
ICF ACK packets rx	301
ICF NACK packets rx	0
ICF HTBT_REQ packets rx	11310
ICF HTBT_RPLY packets rx	11319
ICF SYN packets rx	0

Figure 42: Node to Node statistics

The statistics counters for a node can be cleared by right-clicking on a node and selecting *Clear Statistics* from the command pop-up. The *Statistics* menu also offers the same option.

5.10 Adding and removing a node from CIM

To add a node to CIM, click on the *Tools* pull-down menu. Select *Cluster Integrity* and *Add to CIM* from the expandable pull-down menu (see Figure 43).

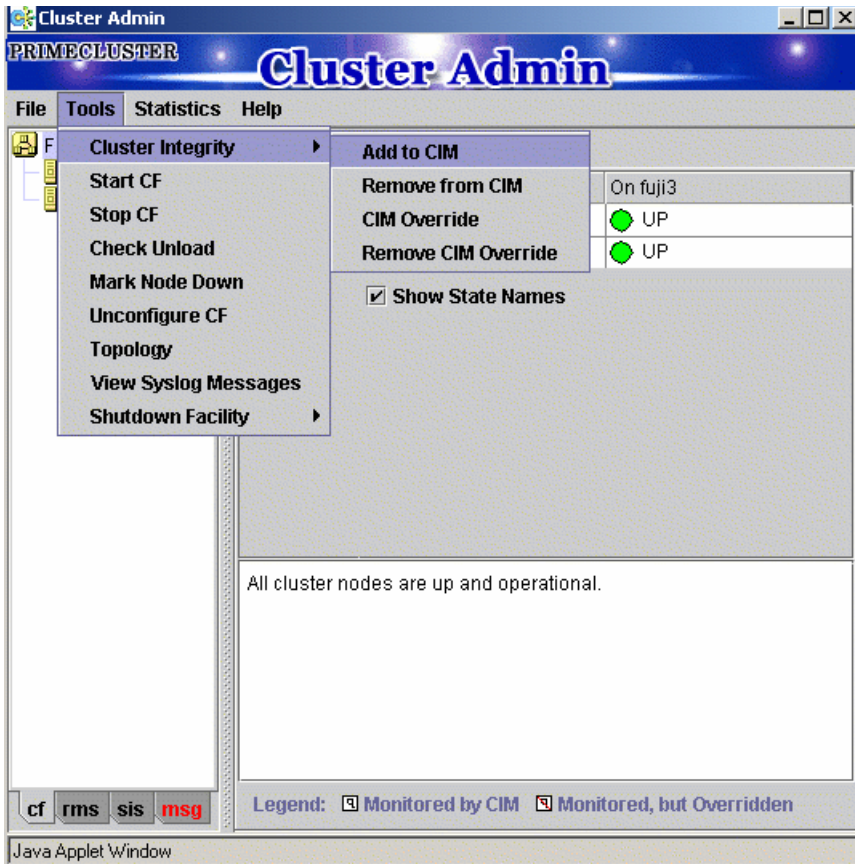


Figure 43: CIM options

The *Add to CIM* pop-up display appears. Choose the desired CF node and click on *Ok* (see Figure 44).

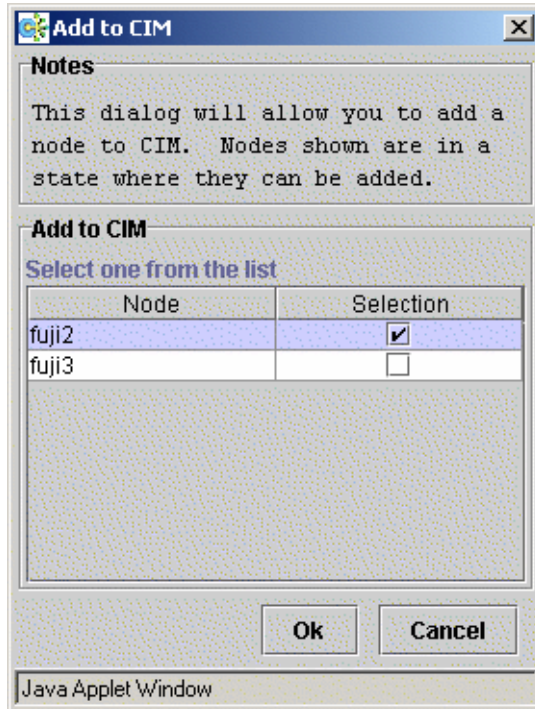


Figure 44: Add to CIM

To remove a node from CIM by means of the *Tools* pull-down menu, select *Cluster Integrity* and *Remove from CIM* from the expandable pull-down menu. Choose the CF node to be removed from the pop-up and click on *Ok*. A node can be removed at any time.

Refer to the Section “Cluster Integrity Monitor” for more details on CIM.

5.11 Unconfigure CF

To unconfigure a CF node, first stop CF on that node. Then, from the *Tools* pull-down menu, click on *Unconfigure CF*.

The *Unconfigure CF* pop-up display appears. Select the check box for the CF node to unconfigure, and click on *Ok* (see Figure 45).

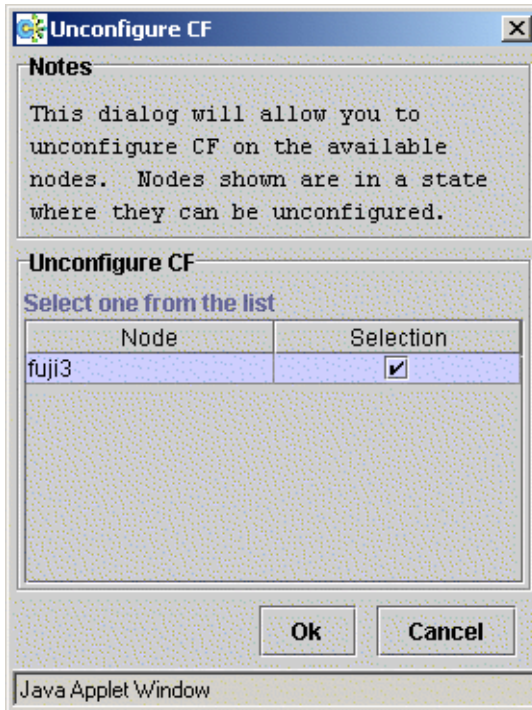


Figure 45: Unconfigure CF

The unconfigured node will no longer be part of the cluster. However, other cluster nodes will still show that node as *DOWN* until they are rebooted.

5.12 CIM Override

The CIM Override option causes a node to be ignored when determining a quorum. A node cannot be overridden if its CF state is UP. To select a node for CIM Override, right-click on a node and choose *CIM Override* (see Figure 46).

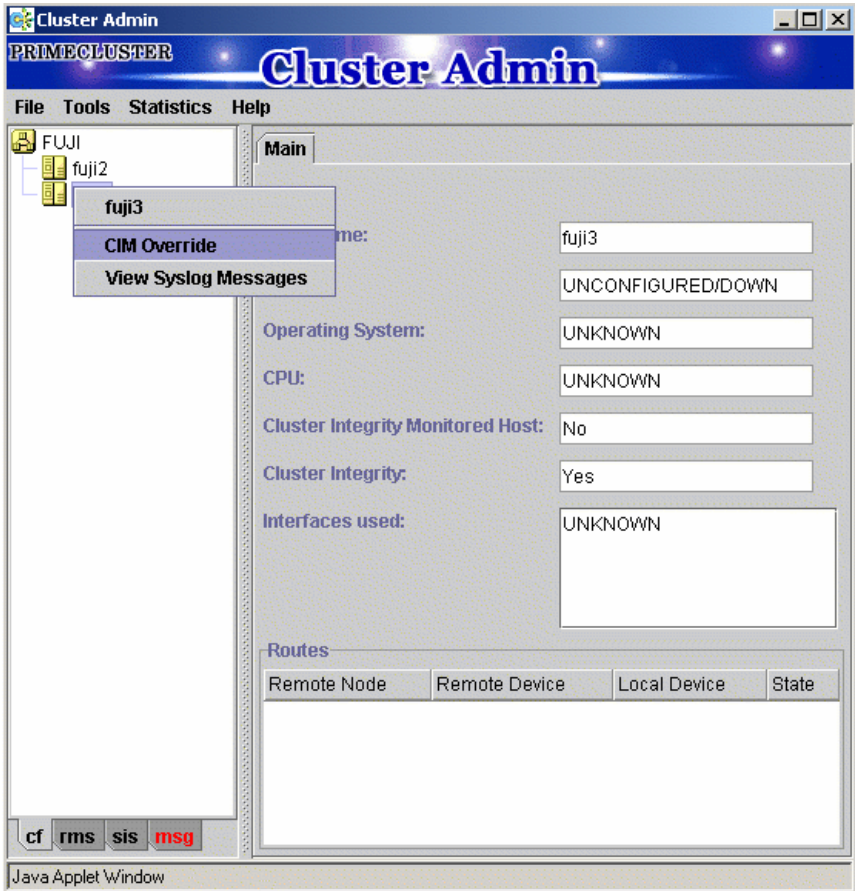


Figure 46: CIM Override

A confirmation pop-up appears (see Figure 47).

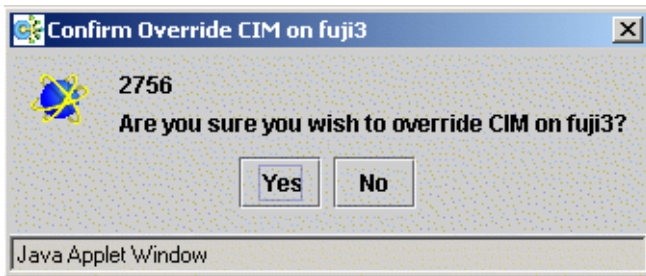


Figure 47: CIM Override confirmation

Click *Yes* to confirm.

6 LEFTCLUSTER state

This chapter defines and describes the `LEFTCLUSTER` state.

This chapter discusses the following:

- The Section “Description of the `LEFTCLUSTER` state” describes the `LEFTCLUSTER` state in relation to the other states.
- The Section “Recovering from `LEFTCLUSTER`” discusses the different ways a `LEFTCLUSTER` state is caused and how to clear it.

Occasionally, while `CF` is running, you may encounter the `LEFTCLUSTER` state, as shown by running the `cftool -n` command. A message will be printed to the console of the remaining nodes in the cluster. This can occur under the following circumstances:

- Broken interconnects—All cluster interconnects going to another node (or nodes) in the cluster are broken.
- Panicked nodes—A node panics.
- Node in kernel debugger—A node is left in the kernel debugger for too long and heartbeats are missed.
- Entering the firmware monitor `OBP`—Will cause missed heartbeats and will result in the `LEFTCLUSTER` state.
- Reboot—Shutting down a node with the `reboot` command.



Nodes running `CF` should normally be shut down with the `shutdown` command or with the `init` command. These commands will run the `rc` scripts that will allow `CF` to be cleanly shut down on that node. If you run the `reboot` command, the `rc` scripts are not run, and the node will go down while `CF` is running. This will cause the node to be declared to be in the `LEFTCLUSTER` state by the other nodes.

If `SF` is fully configured and running on all cluster nodes, it will try to resolve the `LEFTCLUSTER` state automatically. If `SF` is not configured and running, or the `SF` fails to clear the state, the state has to be cleared manually. This section explains the `LEFTCLUSTER` state and how to clear this state manually.

6.1 Description of the LEFTCLUSTER state

Each node in a CF cluster keeps track of the state of the other nodes in the cluster. For example, the other node's state may be UP, DOWN, or LEFTCLUSTER.

LEFTCLUSTER is an intermediate state between UP and DOWN, which means that the node cannot determine the state of another node in the cluster because of a break in communication.

For example, consider the three-node cluster shown in Figure 48.

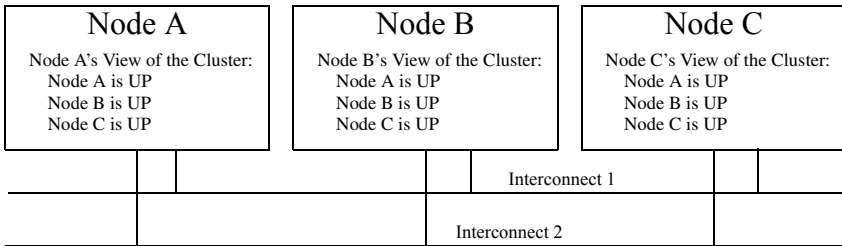


Figure 48: Three-node cluster with working connections

Each node maintains a table of what states it believes all the nodes in the cluster are in.

Now suppose that there is a cluster partition in which the connections to Node C are lost. The result is shown in Figure 49.

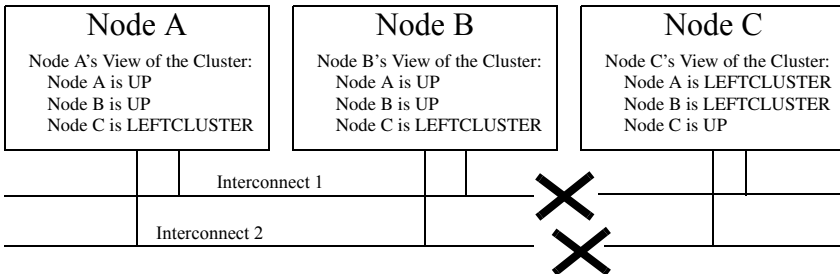


Figure 49: Three-node cluster where connection is lost

Because of the break in network communications, Nodes A and B cannot be sure of Node C's true state. They therefore update their state tables to say that Node C is in the LEFTCLUSTER state. Likewise, Node C cannot be sure of the true states of Nodes A and B, so it marks those nodes as being in the LEFTCLUSTER in its state table.



LEFTCLUSTER is a state that a particular node believes other nodes are in. It is never a state that a node believes that it is in. For example, in Figure 49, each node believes that it is UP.

The purpose of the LEFTCLUSTER state is to warn applications which use CF that contact with another node has been lost and that the state of such a node is uncertain. This is very important for RMS.

For example, suppose that an application on Node C was configured under RMS to fail over to Node B if Node C failed. Suppose further that Nodes C and B had a shared disk to which this application wrote.

RMS needs to make sure that the application is, at any given time, running on either Node C or B but not both, since running it on both would corrupt the data on the shared disk.

Now suppose for the sake of argument that there was no LEFTCLUSTER state, but as soon as network communication was lost, each node marked the node it could not communicate with as DOWN. RMS on Node B would notice that Node C was DOWN. It would then start an instance of the application on Node C as part of its cluster partition processing. Unfortunately, Node C isn't really DOWN. Only communication with it has been lost. The application is still running on Node C. The applications, which assume that they have exclusive access to the shared disk, would then corrupt data as their updates interfered with each other.

The LEFTCLUSTER state avoids the above scenario. It allows RMS and other application using CF to distinguish between lost communications (implying an unknown state of nodes beyond the communications break) and a node that is genuinely down.

When SF notices that a node is in the LEFTCLUSTER state, it uses a non-CF communications facility to contact the previously configured Shutdown Agent and requests that the node which is in the LEFTCLUSTER state be shut down. With PRIMECLUSTER, a weight calculation determines which node or nodes should survive and which ones should be shut down. SF has the capability to arbitrate among the shutdown requests and shut down a selected set of nodes in the cluster, such that the subcluster with the largest weight is left running and subclusters with lessor weights are shutdown.

In the example given, Node C would be shut down, leaving Nodes A and B running. After the SF software shuts down Node C, SF on Nodes A and B clear the LEFTCLUSTER state such that Nodes A and B see Node C as DOWN. Refer to the Chapter “Shutdown Facility” for details on configuring SF and shutdown agents.

i Note that a node cannot join an existing cluster when the nodes in that cluster believe that the node is in the LEFTCLUSTER state. The LEFTCLUSTER state must be cleared before the joining can be done.

6.2 Recovering from LEFTCLUSTER

If SF is not running on all nodes, or if SF is unable to shut down the node which left the cluster, and the LEFTCLUSTER condition occurs, then the system administrator must manually clear the LEFTCLUSTER state. The procedure for doing this depends on how the LEFTCLUSTER condition occurred.

6.2.1 Caused by a panic/hung node

The LEFTCLUSTER state may occur because a particular node panicked or hung. In this case, the procedure to clear LEFTCLUSTER is as follows:

1. Make sure the node is really down. If the node panicked and came back up, proceed to Step 2. If the node is in the debugger, exit the debugger. The node will reboot if it panicked, otherwise shut down the node, called the *offending node* in the following discussion.
2. While the offending node is down, use Cluster Admin to log on to one of the surviving nodes in the cluster. Invoke the CF GUI and select *Mark Node Down* from the *Tools* pull-down menu, then mark the offending node as DOWN. This may also be done from the command line by using the following command:

```
# cftool -k
```
3. Bring the offending node back up. It will rejoin the cluster as part of the reboot process.

6.2.2 Caused by staying in the kernel debugger too long

In Figure 50, Node C was placed in the kernel debugger too long so it appears as a hung node. Nodes A and B decided that Node C's state was LEFTCLUSTER.

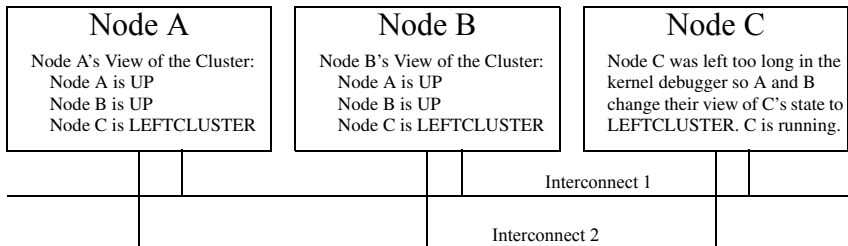


Figure 50: Node C placed in the kernel debugger too long

To recover from this situation, you would need to do the following:

1. Shut down the node.
2. While Node C is down, start up the Cluster Admin on Node A or B. Use *Mark Node Down* from the *Tools* pull-down menu in the CF portion of the GUI to mark Node C DOWN.
3. Bring Node C back up. It will rejoin the cluster as part of its reboot process.

6.2.3 Caused by a cluster partition

A cluster partition refers to a communications failure in which all CF communications between sets of nodes in the cluster are lost. In this case, the cluster itself is effectively partitioned into sub-clusters.

To manually recover from a cluster partition, you must do the following:

1. Decide which of the sub-clusters you want to survive. Typically, you will chose the sub-cluster that has the largest number of nodes in it or the one where the most important hardware is connected or the most important application is running.
2. Shut down all of the nodes in the sub-cluster which you don't want to survive.
3. While the nodes are down, use the Cluster Admin GUI to log on to one of the surviving nodes and run the CF portion of the GUI. Select *Mark Node Down* from the *Tools* menu to mark all of the shut down nodes as DOWN.
4. Fix the network break so that connectivity is restored between all nodes in the cluster.
5. Bring the nodes back up. They will rejoin the cluster as part of their reboot process.

For example, consider Figure 51.

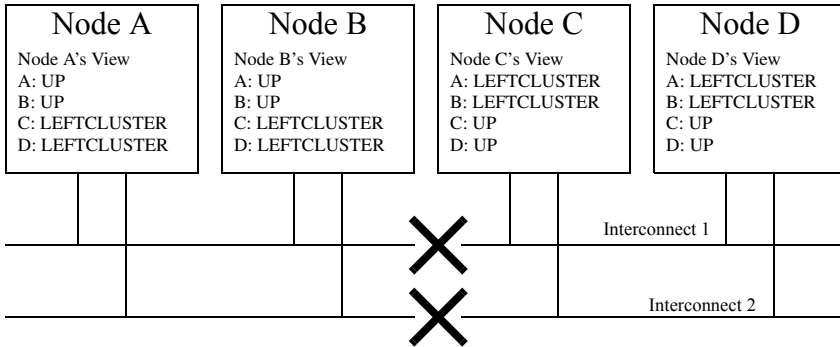


Figure 51: Four-node cluster with cluster partition

In Figure 51, a four-node cluster has suffered a cluster partition. Both of its CF interconnects (Interconnect 1 and Interconnect 2) have been severed. The cluster is now split into two sub-clusters. Nodes A and B are in one sub-cluster while Nodes C and D are in the other.

To recover from this situation, in instances where SF fails to resolve the problem, you would need to do the following:

1. Decide which sub-cluster you want to survive. In this example, let us arbitrarily decide that Nodes A and B will survive.
2. Shut down all of the nodes in the other sub-cluster, here Nodes C and D.
3. While Nodes C and D are down, run the Cluster Admin GUI on either Node A or Node B. Start the CF portion of the GUI and go to *Mark Node Down* from the *Tools* pull-down menu. Mark Nodes C and D as **DOWN**.
4. Fix the interconnect break on Interconnect 1 and Interconnect 2 so that both sub-clusters will be able to communicate with each other again.
5. Bring Nodes C and D back up.

6.2.4 Caused by reboot

The LEFTCLUSTER state may occur because a particular node (called the offending node) has been rebooted. In this case the procedure to clear LEFTCLUSTER is as follows:

1. Make sure the offending node is rebooted in multi-user mode.
2. Use Cluster Admin to log on to one of the surviving nodes in the cluster. Invoke the CF GUI by selecting *Mark Node Down* from the *Tools* pull-down menu. Mark the offending node as *DOWN*.
3. The offending node will rejoin the cluster automatically.

7 CF topology table

This chapter discusses the CF topology table as it relates to the CF portion of the Cluster Admin GUI.

This chapter discusses the following:

- The Section “Basic layout” discusses the physical layout of the topology table.
- The Section “Selecting devices” discusses how the GUI actually draws the topology table.
- The Section “Examples” shows various network configurations and what their topology tables would look like.

The CF topology table is part of the CF portion of the Cluster Admin GUI. The topology table may be invoked from the *Tools->Topology* menu item in the GUI (refer to the Section “Displaying the topology table” in the Chapter “GUI administration”). It is also available during CF configuration in the CF Wizard in the GUI.

The topology table is designed to show the network configuration from perspective of CF. It shows what devices are on the same interconnects and can communicate with each other.

The topology table only considers Ethernet devices. It does not include any IP interconnects that might be used for CF, even if CF over IP is configured.

Displayed devices

The topology table is generated by doing CF pings on all nodes in the cluster and then analyzing the results. On pre-4.0 systems, when the CF driver was loaded, it pushed its modules on all possible Ethernet devices on the system regardless of whether or not they were configured. This allowed CF pings to be done on all Ethernet devices on all nodes in the cluster. Thus, all Ethernet devices would show up in the topology table.

With PRIMECLUSTER 4.0, however, the behavior changed. Starting in 4.0, the CF product offered two different ways to load the driver. A `cfconfig -l` caused the driver to be loaded in the same way as it was on pre-4.0 systems. The CF modules would be pushed on all Ethernet devices on the system. However, the new option `cfconfig -L` caused CF to push CF modules only on the Ethernet devices which were configured for use with CF.

The `-L` option offers several advantages. On systems with large disk arrays, it meant that CF driver load time was often dramatically reduced. On PRIMEPOWER systems with dynamic hardware reconfiguration, Ethernet controllers that are not used by CF could be moved more easily between partitions. Because of these advantages, the `nc` scripts that load CF use the `-L` option.

However, the `-L` option restricts the devices which are capable of sending or receiving CF pings to only configured devices. CF has no knowledge of other Ethernet devices on the system. Thus, when the topology table displays devices for a node where CF has been loaded with the `-L` option, it only displays devices that have been configured for CF.

It is possible that a running cluster might have a mixture of nodes where some were loaded with `-l` and others were loaded with `-L`. In this case, the topology table would show all Ethernet devices for nodes loaded with `-l`, but only CF configured devices for nodes loaded with `-L`. The topology table indicates which nodes have been loaded with the `-L` option by adding an asterisk (*) after the node's name.

When a cluster is totally unconfigured, the CF Wizard will load the CF driver on each node using the `-l` option. This allows all devices on all nodes to be seen. After the configuration is complete, the CF Wizard will unload the CF driver on the newly configured nodes and reload it with `-L`. This means that if the topology table is subsequently invoked on a running cluster, only configured devices will typically be seen.

If you are using the CF Wizard to add a new CF node into an existing cluster where CF is already loaded, then the Wizard will load the CF driver on the new node with `-l` so all of its devices can be seen. However, it is likely that the already configured nodes will have had their CF drivers loaded with `-L`, so only configured devices will show up on these nodes.

The rest of this chapter discusses the format of the topology table. The examples implicitly assume that all devices can be seen on each node. Again, this would be the case when first configuring a CF cluster.

7.1 Basic layout

The basic layout of the topology table is shown in Table 4.

mycluster	Full interconnects		Partial interconnects		Unconnected devices
	Int 1	Int 2	Int 3	Int 4	
Node A	hme0 hme2	hme1	hme3	hme5	hme4 hme6
Node B	hme0	hme2	missing	hme1	
Node C	hme1	hme2	hme3	missing	hme4

Table 4: Basic layout for the CF topology table

The upper-left-hand corner of the topology table gives the CF cluster name. Below it, the names of all of the nodes in the cluster are listed.

The CF devices are organized into three major categories:

- Full interconnects—Have working CF communications to each of the nodes in the cluster.
- Partial interconnects—Have working CF communications to at least two nodes in the cluster, but not to all of the nodes.
- Unconnected devices—Have no working CF communications to any node in the cluster.

If a particular category is not present, it will be omitted from the topology table. For example, if the cluster in Table 4 had no partial interconnects, then the table headings would list only full interconnects and unconnected devices (as well as the left-most column giving the clustername and node names).

Within the full interconnects and partial interconnects category, the devices are further sorted into separate interconnects. Each column under an *Int* number heading represents all the devices on an interconnect. (The column header *Int* is an abbreviation for *Interconnect*.) For example, in Table 4, there are two full interconnects listed under the column headings of *Int 1* and *Int 2*.

Each row for a node represents possible CF devices for that node.

Thus, in Table 4, Interconnect 1 is a full interconnect. It is attached to hme0 and hme2 on Node A. On Node B, it is attached to hme0, and on Node C, it is attached to hme1.

Since CF runs over Ethernet devices, the *hmen* devices in Table 4 represent the Ethernet devices found on the various systems. The actual names of these devices will vary depending on the type of Ethernet controllers on the system. For nodes whose CF driver was loaded with *-L*, only configured devices will be shown.

It should be noted that the numbering used for the interconnects is purely a convention used only in the topology table to make the display easier to read. The underlying CF product does not number its interconnects. CF itself only knows about CF devices and point-to-point routes.

If a node does not have a device on a particular partial interconnect, then the word *missing* will be printed in that node's cell in the partial interconnects column. For example, in Table 4, Node B does not have a device for the partial interconnect labeled *Int 3*.

7.2 Selecting devices

The basic layout of the topology table is shown in Table 4. However, when the GUI actually draws the topology table, it puts check boxes next to all of the interconnects and CF devices as shown in Table 5.

mycluster	Full interconnects		Partial interconnects		Unconnected devices
	<input checked="" type="checkbox"/> Int 1	<input checked="" type="checkbox"/> Int 2	<input type="checkbox"/> Int 3	<input type="checkbox"/> Int 4	
Node A	<input checked="" type="checkbox"/> hme0 <input type="checkbox"/> hme2	<input checked="" type="checkbox"/> hme1	<input type="checkbox"/> hme3	<input type="checkbox"/> hme5	<input type="checkbox"/> hme4 <input type="checkbox"/> hme6
Node B	<input checked="" type="checkbox"/> hme0	<input checked="" type="checkbox"/> hme2	missing	<input type="checkbox"/> hme1	
Node C	<input checked="" type="checkbox"/> hme1	<input checked="" type="checkbox"/> hme2	<input type="checkbox"/> hme3	missing	<input type="checkbox"/> hme4

Table 5: Topology table with check boxes shown

The check boxes show which of the devices were selected for use in the CF configuration. (In the actual topology table, check marks appear instead of x's.)

When the topology table is used outside of the CF Wizard, these check boxes are read-only. They show what devices were previously selected for the configuration. In addition, the unchecked boxes (representing devices which were not configured for CF) will not be seen for nodes where *-L* was used to load CF.

When the topology table is used within the CF Wizard, then the check boxes may be used to select which devices will be included in the CF configuration. Clicking on the check box in an Int *number* heading will automatically select all devices attached to that interconnect. However, if a node has multiple devices connected to a single interconnect, then only one of the devices will be selected.

For example, in Table 5, Node A has both hme0 and hme2 attached to Interconnect 1. A valid CF configuration allows a given node to have only one CF device configured per interconnect. Thus, in the CF Wizard, the topology table will only allow hme0 or hme2 to be selected for Node A. In the above example, if hme2 were selected for Node A, then hme0 would automatically be unchecked.

If the CF Wizard is used to add a new node to an existing cluster, then the devices already configured in the running cluster will be displayed as read-only in the topology table. These existing devices may not be changed without unconfiguring CF on their respective nodes.

7.3 Examples

The following examples show various network configurations and what their topology tables would look like when the topology table is displayed in the CF Wizard on a totally unconfigured cluster. For simplicity, the check boxes are omitted.

Example 1

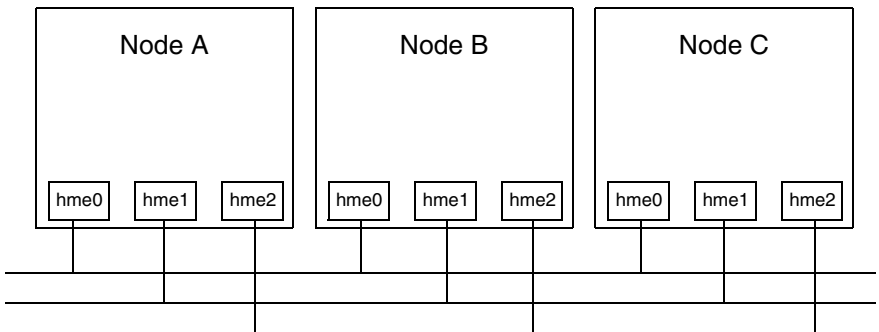


Figure 52: A three-node cluster with three full interconnects

The resulting topology table for Figure 52 is shown in Table 6.

mycluster	Full interconnects		
	Int 1	Int 2	Int 3
Node A	hme0	hme1	hme2
Node B	hme0	hme1	hme2
Node C	hme0	hme1	hme2

Table 6: Topology table for 3 full interconnects

Since there are no partial interconnects or unconnected devices, those columns are omitted from the topology table.

Example 2

In this example, Node A's Ethernet connection for hme1 has been broken. This is shown in Figure 53.

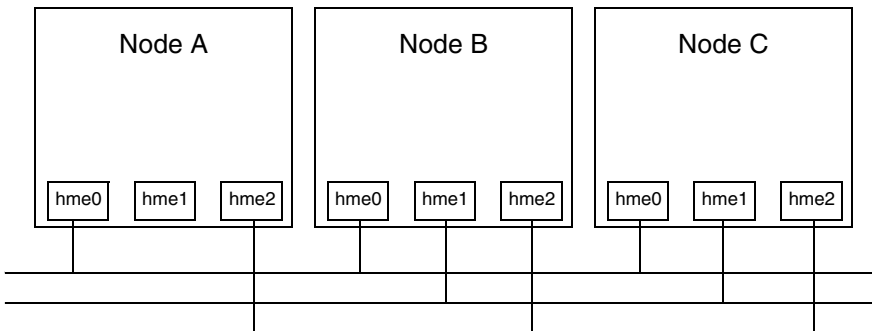


Figure 53: Broken Ethernet connection for hme1 on Node A

The resulting topology table for Figure 53 is shown in Table 7.

mycluster	Full interconnects		Partial interconnects	Unconnected devices
	Int 1	Int 2		
Node A	hme0	hme2	missing	hme1
Node B	hme0	hme2	hme1	
Node C	hme0	hme2	hme1	

Table 7: Topology table with broken Ethernet connection

In Table 7, hme1 for Node A now shows up as an unconnected device. Since one of the interconnects is missing a device for Node A, the Partial Interconnect column now shows up. Note that the relationship between interconnect numbering and the devices has changed between Table 6 and Table 7. In Table 6, for example, all hme1 devices were on Int 2. In Table 7, the hme1 devices for Nodes B and C are now on the partial interconnect Int 3. This change in numbering illustrates the fact that the numbers have no real significance beyond the topology table.

Example 3

This example shows a cluster with severe networking or cabling problems in which no full interconnects are found.

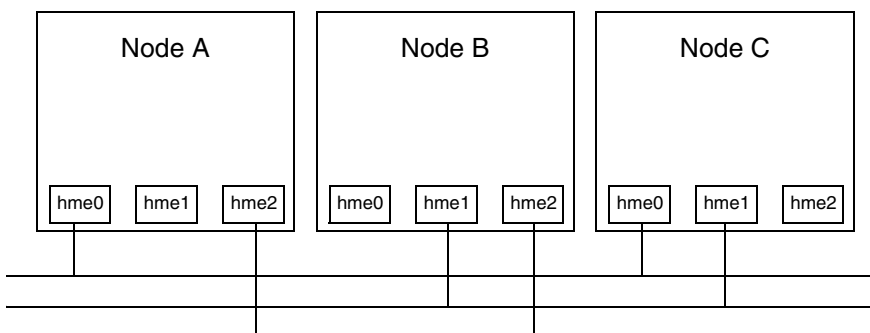


Figure 54: Cluster with no full interconnects

The resulting topology table for Figure 54 is shown in Table 8.

mycluster	Partial interconnects			Unconnected devices
	Int 1	Int 2	Int 3	
Node A	hme0	missing	hme2	hme1
Node B	missing	hme1	hme2	hme0
Node C	hme0	hme1	missing	hme2

Table 8: Topology table with no full interconnects

In Table 8, the full interconnects column is omitted since there are none. Note that if this configuration were present in the CF Wizard, the wizard would not allow you to do configuration. The wizard requires that at least one full interconnect must be present.

8 Shutdown Facility

This chapter describes the components and advantages of PRIMECLUSTER Shutdown Facility (SF) and provides administration information.

This chapter discusses the following:

- The Section “Overview” describes the components of SF.
- The Section “Available Shutdown Agents” describes the available agents for use by the SF.
- The Section “SF split-brain handling” describes the methods for resolving split cluster situations.
- The Section “Configuring the Shutdown Facility” describes the configuration of SF and its agents.
- The Section “SF facility administration” provides information on administering SF.
- The Section “Logging” describes the log files used by SF and its agents.

8.1 Overview

The SF provides the interface for managing the shutdown of cluster nodes when error conditions occur. The SF also cares for advising other PRIMECLUSTER products of the successful completion of node shutdown so that recovery operations can begin.

The SF is made up of the following major components:

- The Shutdown Daemon (SD)
- One or more Shutdown Agents (SA)
- Monitoring Agent (MA)
- `sdttool(1M)` command

Shutdown Daemon

The SD is started at system boot time and is responsible for the following:

- Monitoring the state of all cluster nodes
- Monitoring the state of all registered SAs

- Reacting to indications of cluster node failure and verify or manage node elimination
- Resolving split-brain conditions
- Advising other PRIMECLUSTER products of node elimination completion

The SD uses SAs to perform most of its work with regard to cluster node monitoring and elimination. In addition to SA's, the SD interfaces with the Cluster Foundation layer's ENS system to receive node failure indications and to advertise node elimination completion.

Shutdown Agents

The SA's role is to attempt to shut down a remote cluster node in a manner in which the shutdown can be guaranteed. Some of the SAs are shipped with the SF product, but may differ based on the architecture of the cluster node on which SF is installed. SF allows any PRIMECLUSTER service layer product to shut down a node whether RMS is running or not.

An SA is responsible for shutting down, and verifying the shutdown of a cluster node. Each SA uses a specific method for performing the node shutdown such as:

- SA_scon uses the cluster console running the SCON software.
- SA_pprcip and SA_pprcir use the RCI interface available on PRIMEPOWER nodes.
- SA_rccu uses the RCCU units available for PRIMEPOWER nodes.
- SA_wtinps uses an NPS unit.
- SA_rps uses an RPS unit.

The Section "Available Shutdown Agents" discuss SAs in more detail.

If more than one SA is used, the first SA in the configuration is used as the primary SA. SD always uses the primary SA. The other secondary SAs are used as fall back SAs only if the primary SA fails for some reason.

Monitoring Agent

The monitoring agent provides the following functions:

- Monitors the remote node—Monitors the state of the remote node that uses the hardware features. It also notifies the SD of a failure in the event of an unexpected system panic and shutoff.
- Eliminates the remote node—Provides a function to forcibly shut down the node as Shutdown Agent.

sdtool command

The `sdtool(1M)` utility is the command line interface for interacting with the SD. With it the administrator can:

- Start and stop the SD (although this is typically done with an RC script run at boot time)
- View the current state of the SA's
- Force the SD to reconfigure itself based on new contents of its configuration file
- Dump the contents of the current SF configuration
- Enable/disable SD debugging output
- Eliminate a cluster node



Although the `sdtool(1M)` utility provides a cluster node elimination capability, the preferred method for controlled shutdown of a cluster node is to use the `init(1M)` system utility. Refer to the `init(1M)` manual page for more details.

8.2 Available Shutdown Agents

This section describes the set of supported SAs:

- RCI—Remote Cabinet Interface
- NPS—Network Power Switch
- SCON—Single Console
- RCCU—Remote Console Control Unit
- RPS—Remote Power Switch

8.2.1 RCI

The RCI SA provides a shutdown method only for the PRIMEPOWER clusters on all PRIMEPOWER platforms.

There are two kinds of RCI SAs:

- `SA_pprcip`—Provides a shutdown mechanism by panicking the node through RCI.
- `SA_pprcir`—Provides a shutdown mechanism by resetting the node through RCI.

Setup and configuration

Hardware setup of the RCI is performed only by qualified support personnel. Contact them for more information. In addition, you can refer to the manual shipped with the unit and to any relevant PRIMECLUSTER Release Notices for more details on configuration.

Restrictions

- RCI node elimination does not work in heterogeneous clusters of PRIMEPOWER 200, 400, 600, 650, 850 and PRIMEPOWER 800, 1000, 2000.
- The RCI network is restricted to a maximum distance of 150 meters between all nodes.
- PRIMEPOWER nodes only reboot automatically after a panic if the setting of the `eeeprom` variable `boot-file` is not `kadb`. The `SCON` kill on PRIMEPOWER 200, 400, 600, 650, 850 nodes require the `kadb` setting. An automatic reboot after panic is not possible on those nodes if the elimination via panic is supposed to be a *fall back* elimination method after a failing `SCON` elimination.



Applies to `SCON` supported clusters only.

- PRIMEPOWER 200, 400, 600, 650, 850 nodes should have the `eeeprom` variable `auto-boot` set to `true`. If this variable is not `true` the nodes will not be booted up automatically after a power recycle.

Log file

/var/opt/SMAwsf/log/SA_pprcip.log

or

/var/opt/SMAwsf/log/SA_pprcir.log

8.2.2 SCON



Certain product options are region-specific. For information on the availability of SCON, contact your local customer-support service representative.

The Single Console (SCON) SA, `SA_scon`, provides an alternative SA for PRIMECLUSTER. SCON performs necessary node elimination tasks, coordinated with console usage.

Setup and configuration

To use the `SA_scon` SA, a system console (external to the cluster nodes) should be fully configured with the SCON product. Refer to the Chapter “System console” for details on the setup and configuration of SCON.

`SA_scon` is one of the SAs called by the Shutdown Facility when performing node elimination. The `SA_scon` process running on the cluster node communicates with the SCON running on the cluster console to request that a cluster node be eliminated. To communicate with the cluster console, the `SA_scon` SA must be properly configured.

Log file

/var/opt/SMAwsf/log/SA_scon.log

8.2.3 RCCU



Certain product options are region-specific. For information on the availability of RCCU, contact your local customer-support service representative.

The Remote Console Control Unit (RCCU) SA, `SA_rccu`, provides a SA using the RCCU.

Setup and configuration

The RCCU unit must be configured according to the directions in the manual shipped with the unit. The RCCU unit should be assigned an IP address and name, so that the cluster nodes can connect to it over the network. All the RCCU ports that will be connected to the cluster nodes console lines should be configured according to the instructions given in the manual.

i If the node is eliminated by the console monitoring agent, a break signal is sent to the node, and this node is stopped in the open-boot prompt (OBP) mode.

Log file

`/var/opt/SMAWsf/log/SA_rccu.log`

8.2.4 RPS

i Certain product options are region-specific. For information on the availability of RPS, contact your local customer-support service representative.

The Remote Power Switch (RPS) SA, `SA_rps`, provides a node shutdown function using the RPS unit.

Setup and configuration

The RPS must be configured according to the directions in the RPS manuals. The optional software `SMAWrsb` must be installed and working for power off and power on commands. The nodes must be connected to plugs with the plug-IDs given in the appropriate host entry.

Log file

`/var/opt/SMAWsf/log/SA_rps.log`

8.2.5 NPS

i Certain product options are region-specific. For information on the availability of SCON, contact your local customer-support service representative.


The Network Power Switch (NPS) SA is `SA_wtinps`. This SA provides a node shutdown function using the Western Telematic Inc. Network Power Switch (WTI NPS) unit to power-cycle selected nodes in the cluster.


Setup and configuration

The WTI NPS unit must be configured according to the directions in the manual shipped with the unit. At the very least, an IP address must be assigned to the unit and a password must be enabled. Make sure that the cluster node's power plugs are plugged into the NPS box and that the `command confirmation` setting on the NPS box is set to `on`.

It is advisable to have the NPS box on a robust LAN connected directly to the cluster nodes.

The boot delay of every configured plug in the NPS box should be set to 10 seconds.

 If you want to set the boot delay to any other value, make sure that the “timeout value” for the corresponding `SA_wtinps` agent should be set such that it is greater than this boot delay value by at least 10 seconds. To set this value, use the detailed configuration mode for SF.

 If more than a single plug is assigned to a name (which means that more than one plug will be operated per `/on`, `/off`, `/boot` command), the “boot delay” of these plugs must be assigned to a value larger than 10 seconds, otherwise timeouts may occur. The timeout value of the corresponding `SA_wtinps` should be set as follows:

```
timeout = boot_delay + (* 2 * no of plugs) + 10
```

Log file

```
/var/opt/SMAwsf/log/SA_wtinps.log
```

8.3 SF split-brain handling

The PRIMECLUSTER product provides the ability to gracefully resolve split-brain situations as described in this section.

8.3.1 Administrative LAN

Split-brain processing makes use of Administrative LAN. For details on setting up such a LAN, see the *PRIMECLUSTER Installation Guide (Solaris)*. The use of Admin LAN is optional, however the use of an Administrative LAN is recommended for faster and more accurate split-brain handling.

8.3.2 Overview of split-brain handling

A *split-brain* condition is one in which one or more cluster nodes have stopped receiving heartbeats from one or more other cluster nodes, yet those nodes have been determined to still be running. Each of these distinct sets of cluster nodes is called a sub-cluster, and when a split-brain condition occurs the Shutdown Facility has a choice to make as to which sub-cluster should remain running.

Only one of the sub-clusters created due to a split-brain condition can survive. The SF will attempt to determine which sub-cluster is most important and allow that sub-cluster to remain. The notion of importance is maintained within PRIMECLUSTER in two ways:

- The `ShutdownPriority` attribute of an RMS `userApplication` object
- The weight value assigned to each cluster node by the Shutdown Facility

RMS ShutdownPriority attribute

RMS supports the ability to set application importance in the form of a `ShutdownPriority` value for each `userApplication` object defined within the RMS configuration. These values are combined for all `userApplication` objects that are `Online` on a given cluster node to represent the total application weight of that node. When a `userApplication` object is switched from one node to another, the value of that `userApplication` object's `ShutdownPriority` is transferred to the new node.

The higher the value of the `ShutdownPriority` attribute, the more important the application.

Shutdown Facility weight assignment

The Shutdown Facility supports the ability to define node importance in the form of a weight setting in the configuration file. This value represents a node weight for the cluster node.

The higher the node weight value, the more important the node.

8.3.3 Runtime processing

Spit-brain handling may be performed by one of the following elements of the Shutdown Facility:

- The cluster console running the SCON software
- The Shutdown Facility internal algorithm

Both methods use the node weight calculation to determine which sub-cluster is of greater importance. The node weight is the added value of the node weight defined in the Shutdown Facility and the total application weight calculated within RMS.

SCON algorithm

When the SCON is selected as the split-brain resolution manager, SF passes the node weight to the `SA_scon` SA which in turn passes a shutdown request to the SCON.

All cluster nodes send shutdown requests to the SCON containing the name of the node requesting the shutdown, its node weight, and the name of the node to shutdown. These shutdown requests are passed to the SCON over an administrative network (which may or may not be the same network identified as `admIP` within the SF configuration file). The SCON collects these requests and determines which sub-cluster is the heaviest and proceeds to shut down all other nodes not in the heaviest sub-cluster.

SF internal algorithm

When the SF is selected as the split-brain resolution manager, the SF uses the node weight internally.

The SF on each cluster node identifies which cluster nodes are outside its sub-cluster and adds each one of them to an internal shutdown list. This shutdown list, along with the local nodes node weight, is advertised to the SF instances running on all other cluster nodes (both in the local sub-cluster and outside the local sub-cluster) via the `admIP` network defined in the SF configuration file. After the SFs on each cluster node receive the advertisements, they each calculate the heaviest sub-cluster. The heaviest sub-cluster shuts down all lower weight sub-clusters.

In addition to handling well-coordinated shutdown activities defined by the contents of the advertisements, the SF internal algorithm will also resolve split-brain if the advertisements fail to be received. If the advertisements are not received then the split-brain will still be resolved, but it may take a bit more time as some amount of delay will have to be incurred.

The split-brain resolution done by the SF in situations where advertisements have failed depends on a variable delay based on the inverse of the percentage of the available cluster weight the local sub-cluster contains. The more weight it contains the less it delays. After the delay expires (assuming the sub-cluster has not been shut down by a higher-weight sub-cluster) the SF in the sub-cluster begins shutting down all other nodes in all other sub-clusters.

If a sub-cluster contains greater than 50 percent of the available cluster weight, then the SF in that sub-cluster will immediately start shutting down all other nodes in all other sub-clusters.

8.3.4 Split-brain resolution manager selection

The selection of the method to use for split-brain resolution (SCON or SF) depends on site specific conditions. This is done automatically at startup.

SCON is selected as the split-brain resolution manager if SCON is the only SA for your cluster.

For all other situations, SF is selected as the split-brain resolution manager.

i If SF is selected as the split-brain resolution manager, SCON should be configured not to do split-brain processing. This can be done by changing the `rmshosts` method file. Refer to the Section “`rmshosts` method file” for more information.

This selection cannot be changed manually after startup.

8.3.5 Configuration notes

When configuring the Shutdown Facility, RMS, and defining the various weights, the administrator should consider what the eventual goal of a split-brain situation should be.

Typical scenarios that are implemented are as follows:

- Largest Sub-cluster Survival (LSS)

- Specific Hardware Survival (SHS)
- Specific Application Survival (SAS)

The weights applied to both cluster nodes and to defined applications allow considerable flexibility in defining what parts of a cluster configuration should survive a split-brain condition. Using the settings outlined below, administrators can advise the Shutdown Facility about what should be preserved during split-brain resolution.

Largest Sub-cluster Survival

In this scenario, the administrator does not care which physical nodes survive the split, just that the maximum number of nodes survive. If RMS is used to control applications, it will move the applications to the surviving cluster nodes after split-brain resolution has succeeded.

This scenario is achieved as follows:

- By means of Cluster Admin, set the SF node weight values to 1. 1 is the default value for this attribute, so new cluster installations may simply ignore it.
- By means of the RMS Wizard Tools, set the RMS attribute `ShutdownPriority` of all `userApplications` to 0. 0 is the default value for this attribute, so if you are creating new applications you may simply ignore this setting.

As can be seen from the default values of both the SF weight and the RMS `ShutdownPriority`, if no specific action is taken by the administrator to define a split-brain resolution outcome, LSS is selected by default.

Specific Hardware Survival

In this scenario, the administrator has determined that one or more nodes contain hardware that is critical to the successful functioning of the cluster as a whole.

This scenario is achieved as follows:

- By means of Cluster Admin, set the SF node weight of the cluster nodes containing the critical hardware to values more than double the combined value of cluster nodes not containing the critical hardware.
- By means of the RMS Wizard Tools, set the RMS attribute `ShutdownPriority` of all `userApplications` to 0. 0 is the default value for this attribute so if you are creating new applications you may simply ignore this setting.

As an example, in a four-node cluster in which two of the nodes contain critical hardware, set the SF weight of those critical nodes to 10 and set the SF weight of the non-critical nodes to 1. With these settings, the combined weights of both non-critical nodes will never exceed even a single critical node.

Specific Application Survival

In this scenario, the administrator has determined that application survival on the node where the application is currently `Online` is more important than node survival. This can only be implemented if RMS is used to control the application(s) under discussion. This can get complex if more than one application is deemed to be critical and those applications are running on different cluster nodes. In some split-brain situations, all applications will not survive and will need to be switched over by RMS after the split-brain has been resolved.

This scenario is achieved as follows:

- By means of Cluster Admin, set the SF node weight values to 1. 1 is the default value for this attribute, so new cluster installations may simply ignore it.
- By means of the RMS Wizard Tools, set the RMS attribute `ShutdownPriority` of the critical applications to more than double the combined values of all non-critical applications, plus any SF node weight.

As an example, in a four-node cluster there are three applications. Set the SF weight of all nodes to 1, and set the `ShutdownPriority` of the three applications to 50, 10, 10. This would define that the application with a `ShutdownPriority` of 50 would survive no matter what, and further that the sub-cluster containing the node on which this application was running would survive the split no matter what. To clarify this example, if the cluster nodes were A, B, C and D all with a weight of 1, and App1, App2 and App3 had `ShutdownPriority` of 50, 10 and 10 respectively, even in the worst-case split that node D with App1 was split from nodes A, B and C which had applications App2 and App3 the weights of the sub-clusters would be D with 51 and A,B,C with 23. The heaviest sub-cluster (D) would win.

8.4 Configuring the Shutdown Facility

This section describes how to use Cluster Admin and the CLI to configure the Shutdown Facility (SF).

8.4.1 Invoking the Configuration Wizard

Use the *Tools* pull-down menu to select *Shutdown Facility*, and then choose *Configuration Wizard* to invoke the SF Configuration Wizard (see Figure 55).

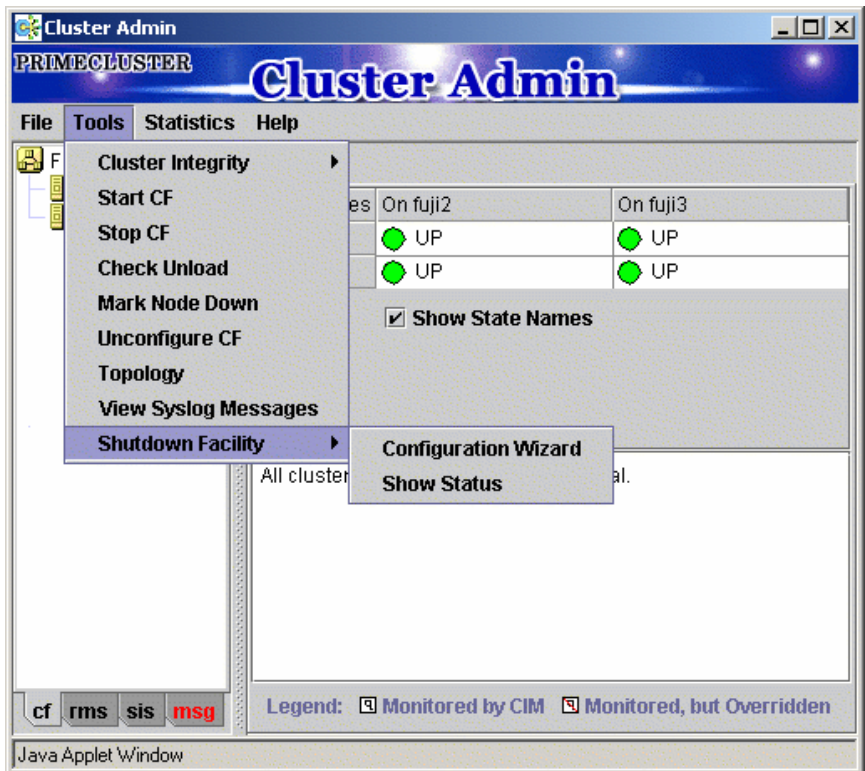


Figure 55: Opening the SF Configuration Wizard

Select the mode for configuration (see Figure 56). You can either choose the *Easy configuration* mode or the *Detailed configuration* mode. *Easy configuration* mode provides the most commonly used configurations. *Detailed configuration* provides complete flexibility in configuration. It is recommended that you use the *Easy configuration* mode.

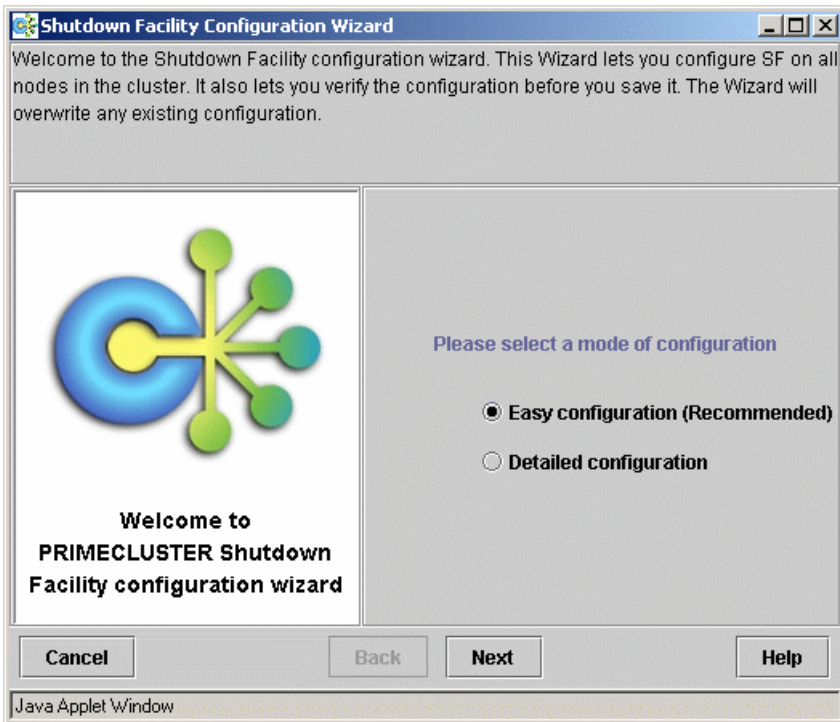


Figure 56: Selecting the mode of SF configuration

Choose the *Easy configuration* selection as shown in Figure 56 and click *Next*.

The screen for selecting the SA appears (see Figure 57). Now select the SAs to be configured. You can either select *SCON* as the primary SA and one or more backup agents from the list, or you can configure *RCI* as the only SA. If one or more backup agents are selected, you will have to enter their individual configuration information on subsequent screens.

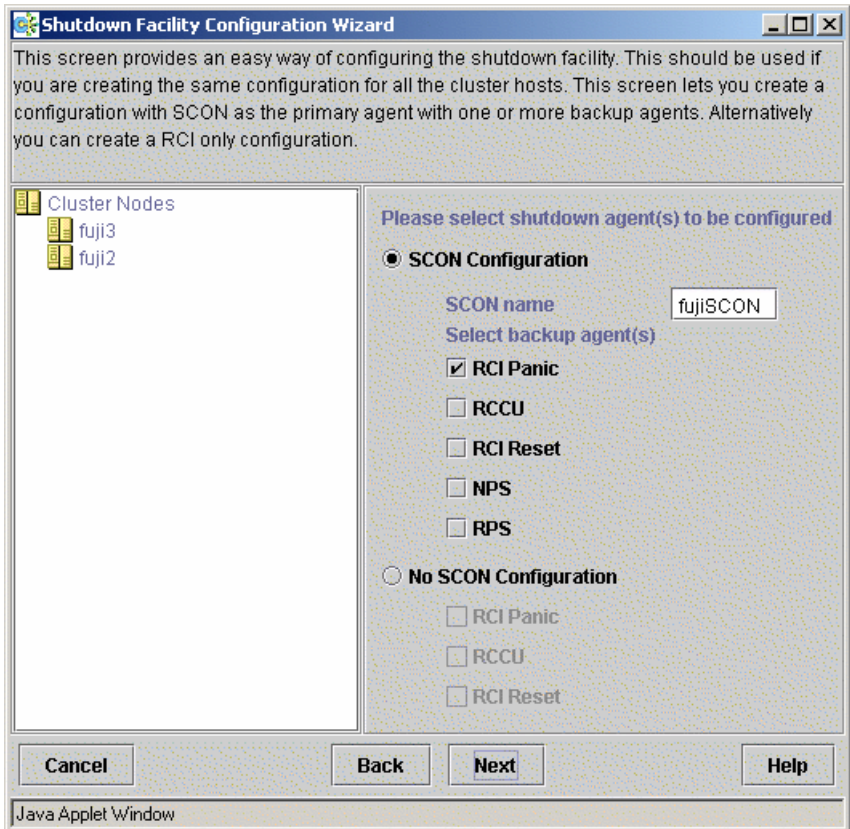


Figure 57: Easy mode of SF configuration

Choose the appropriate selection as shown in Figure 57 and click *Next*. If you choose *RCCU*, *NPS*, or *RPS* as backup agents, you will be taken to the individual SA's configuration screens, which are Figure 65, Figure 66, and Figure 67 respectively.



If you choose *SCON Configuration*, the *SCON name* field has to be filled with the name of the system console.

After you are done configuring individual SAs (if any), you are taken to the screen for finishing the configuration (see Figure 69).

You can also choose to create a new configuration file or edit an existing configuration. If you choose *Detailed configuration* in Figure 56 and click *Next*, a figure such as Figure 58 appears. Choose *Create* as shown in Figure 58 and click *Next*.

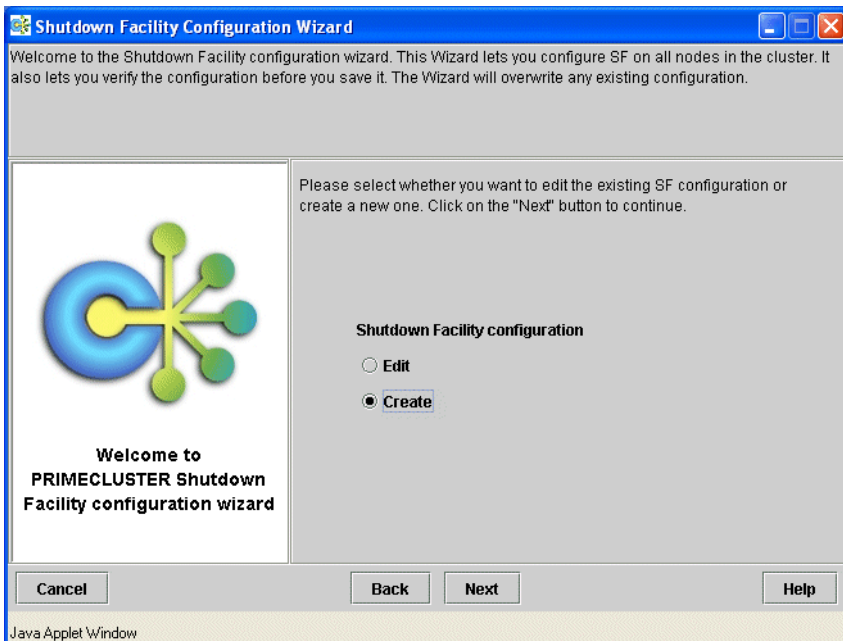


Figure 58: Detailed mode of SF configuration

Select a configuration with the same set of SAs for all the nodes or different SAs for the individual nodes as shown in Figure 59. Click *Next*.

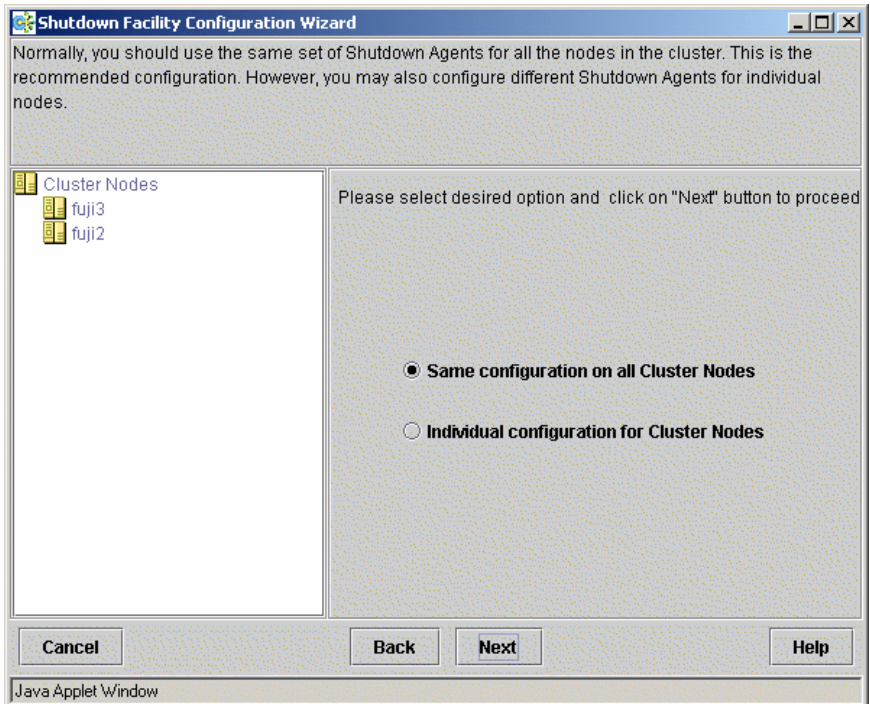


Figure 59: Choice of common configuration for all nodes

If you choose *Same configuration on all Cluster Nodes* and click *Next*, a screen such as Figure 61 appears. If you choose *Individual configuration for Cluster Nodes*, then a screen as shown in Figure 60 appears. In this case, you may configure SF individually at a later time for each of the nodes or groups of nodes.

i It is required that you have the same configuration on all cluster nodes.

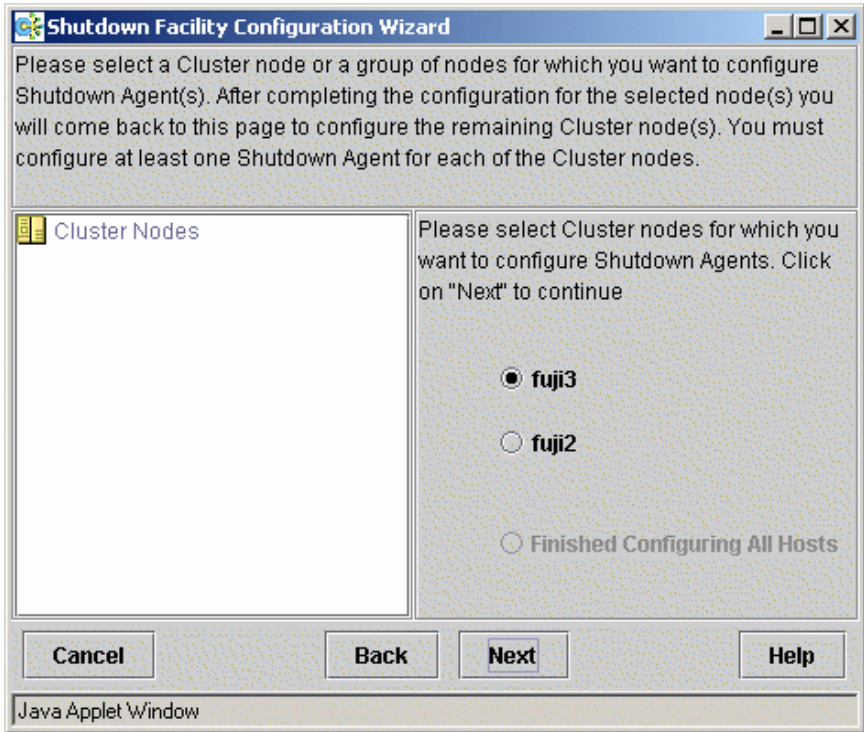


Figure 60: Individual configuration for Cluster Nodes

Choose the cluster node that you want to configure and click *Next*. Note that the left panel in the window displays the cluster nodes and will progressively show the SAs configured for each node.

If you chose *Same configuration on all Cluster Nodes* in Figure 59 and clicked *Next*, a screen such as Figure 61 appears.

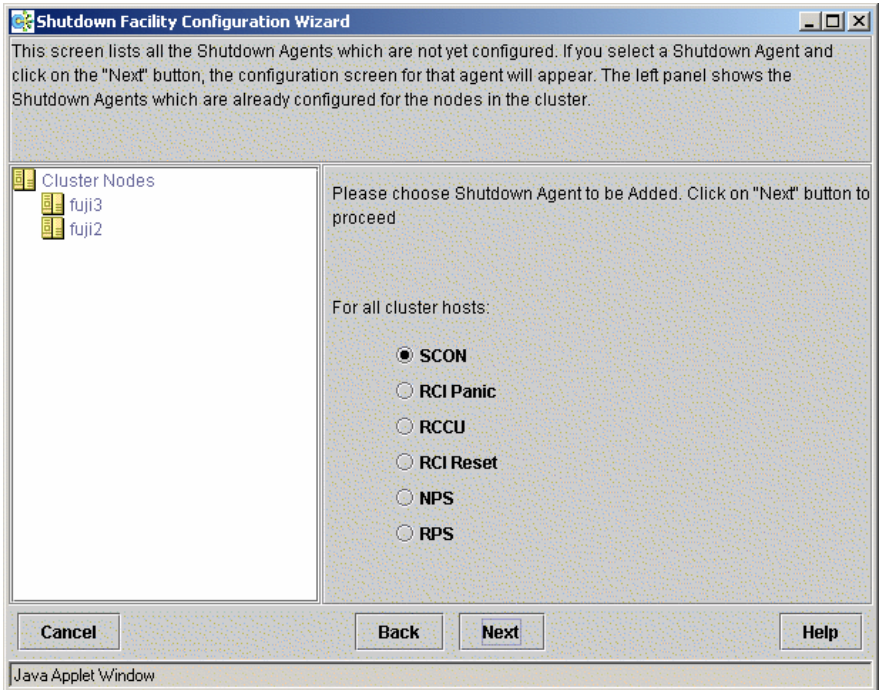


Figure 61: Choose Shutdown Agent to be added

Choose a SA from the given list and press the *Next* button. From here you will be taken to the individual SAs configuration screen, depending on the SAs selected here.

For *RCI Panic* and *RCI Reset* SAs, no further configuration is required.

If you select *SCON* from the list and click on the *Next* button, the screen to configure the SCON SA appears (see Figure 62).

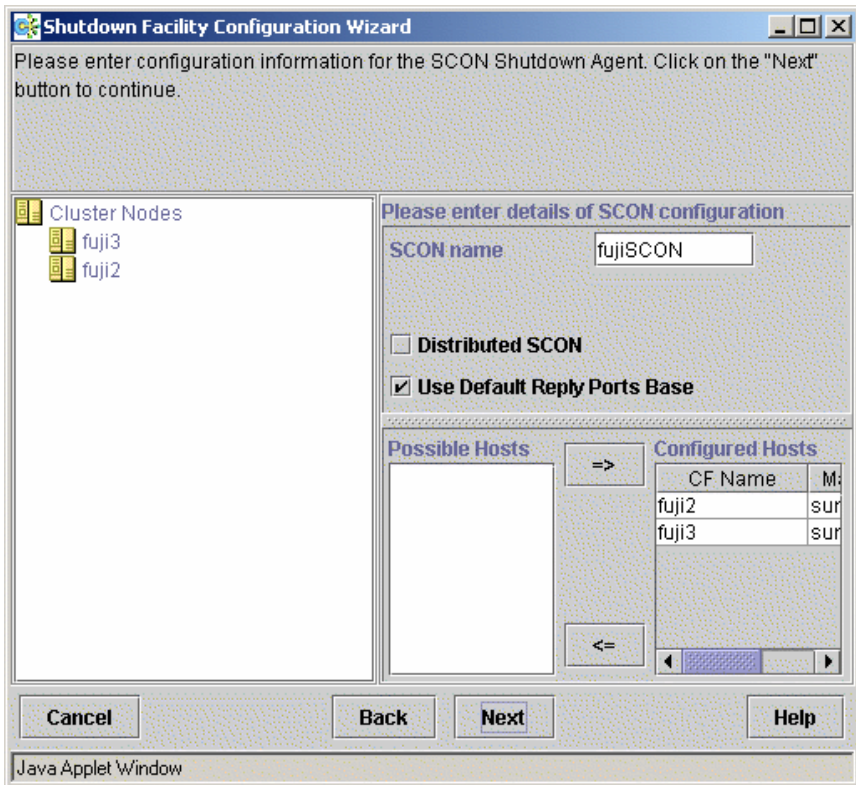


Figure 62: Details for SCON Shutdown Agent

You can click *Distributed SCON* to configure distributed SCON (see Figure 63).

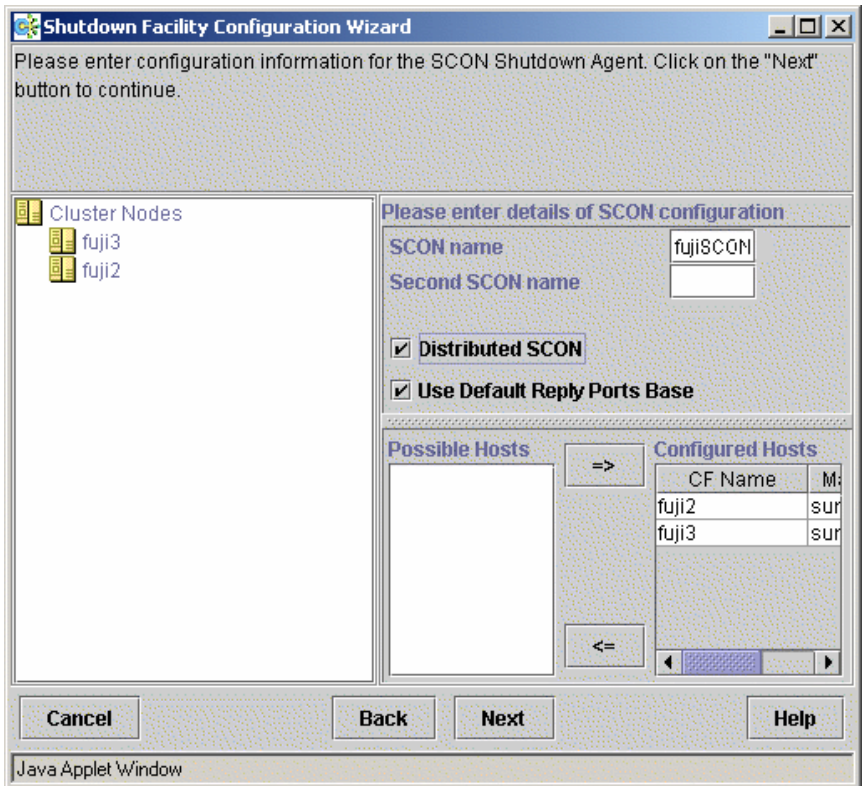


Figure 63: Configuring the SCON Shutdown Agent

If you choose RCCU and uncheck the *Use defaults* check box, the screen shown in Figure 64 appears. Enter the details for each cluster node, namely *RCCU name*, *RCCU-tty*, *Control Port*, *Console port* and *password*. Then click the *Next* button.

Here we used the values `rccu2`, `tty1`, `8010`, and `23` for `fujij2` and `rccu2`, `tty1`, `8010`, and `23` for `fujij3`.

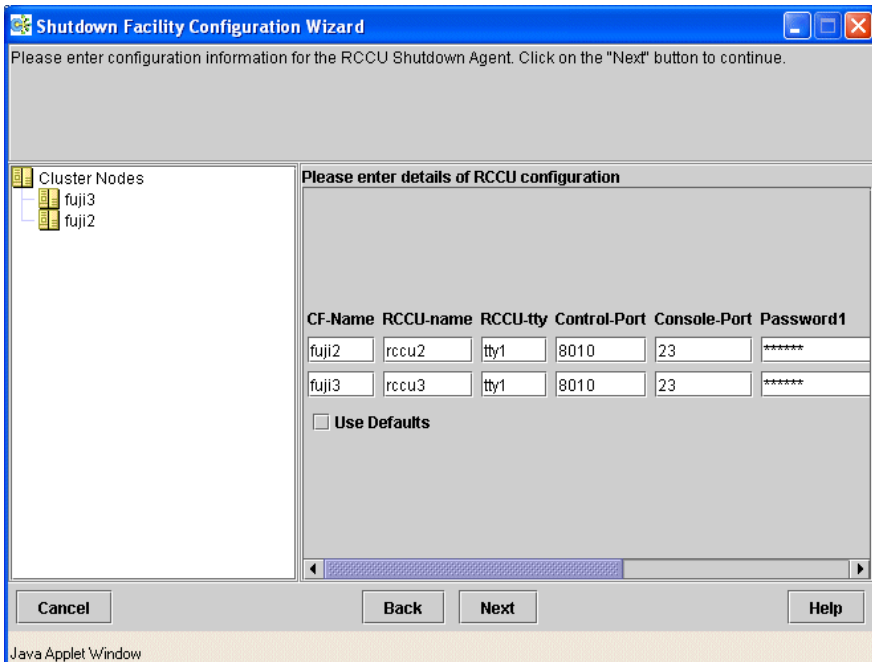


Figure 64: Configuring RCCU

If *Use Defaults* is checked, the default values are used (see Figure 65).

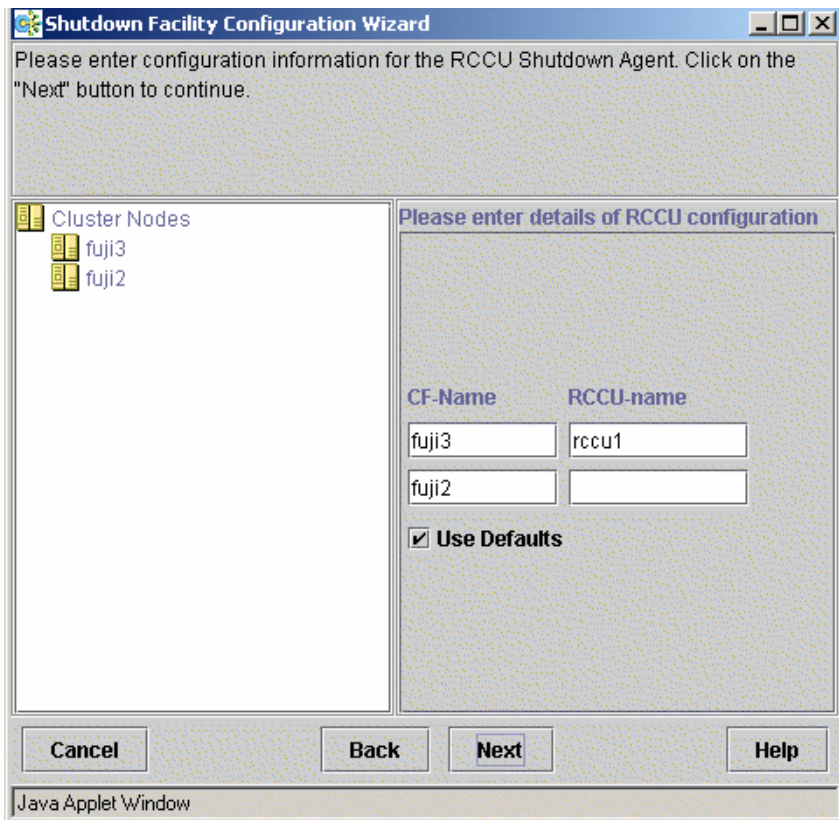


Figure 65: RCCU default values

The default values for *RCCU-name*, *RCCU-tty* and *Control-Port* are `rccu1`, `tty1`, and `8018`, respectively.

Figure 66 is the screen in which to enter the *NPS Shutdown Agent* details. Enter NPS name, password, and choose the action. You can choose the value *cycle* or leave off for action. Then click *Next*.

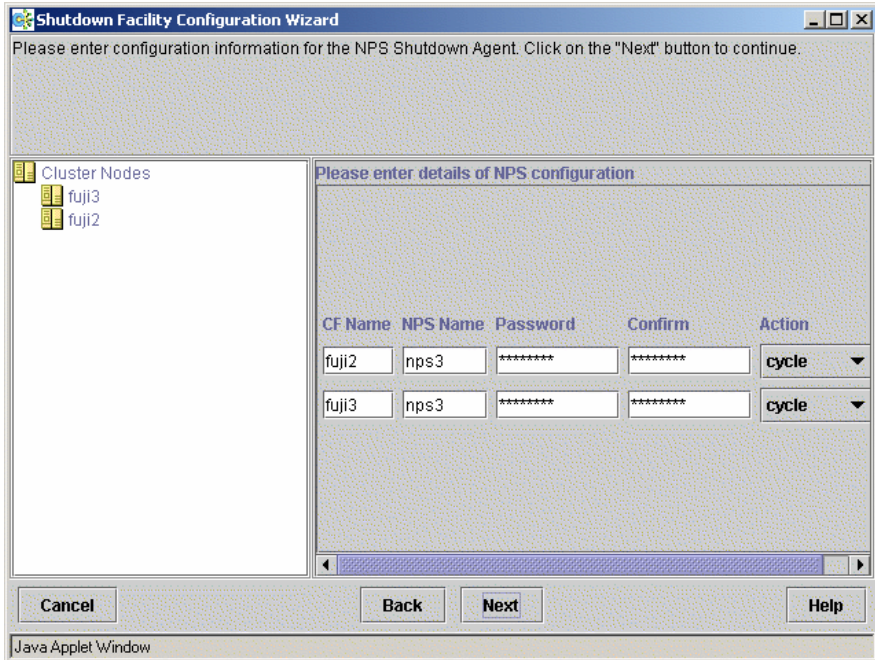


Figure 66: Configuring the NPS Shutdown Agent

The action is, by default, *cycle*, which means that the node is power cycled after shutdown.

If you choose RPS, the screen shown in Figure 67 appears. Enter the details for each of the cluster nodes; namely, the IP address of the RPS unit, *User*, *Password*, and *Action*. Then click the *Next* button.

Shutdown Facility Configuration Wizard

Please enter configuration information for the RPS Shutdown Agent. Click on the "Next" button to continue.

Cluster Nodes

- fuji3
- fuji2

Please enter details of RPS configuration

CF-Name	Action	No of Units
fuji3	cycle	1
fuji2	cycle	1

IP of Unit	User	Password	Confirm	Plugs
				<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4
				<input type="checkbox"/> 1 <input type="checkbox"/> 2 <input type="checkbox"/> 3 <input type="checkbox"/> 4

Cancel Back Next Help

Java Applet Window

Figure 67: Configuring the RPS Shutdown Agent

You can continue to *Add*, *Delete*, or *Edit* the SAs as shown in Figure 68.

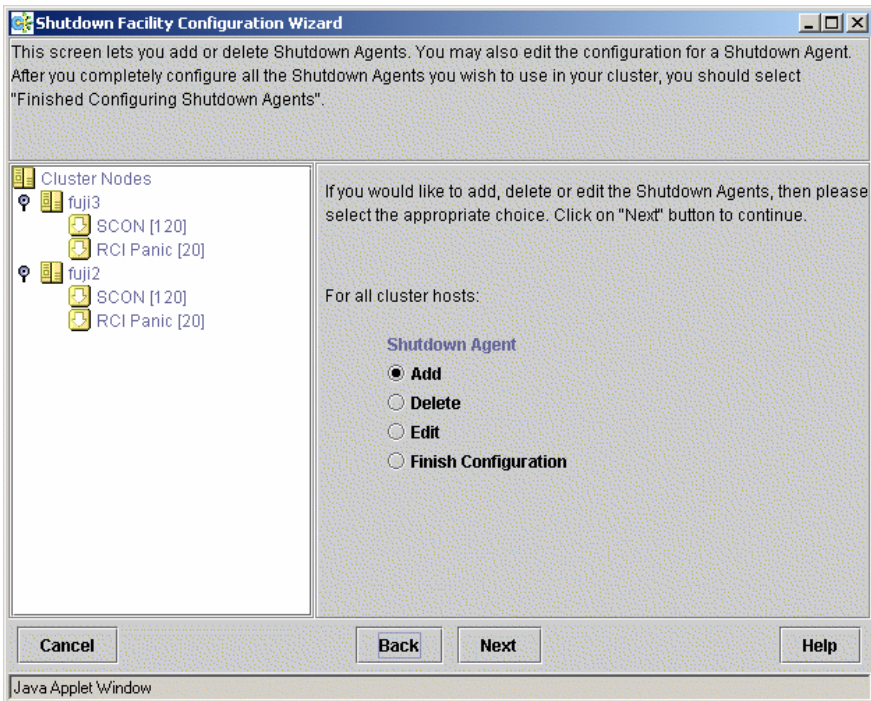


Figure 68: Add, Delete, Edit Shutdown Agent

If you have finished, select *Finish Configuration* and click on *Next* (see Figure 69).



Figure 69: Finishing configuration

Next use the *UP* or *DOWN* buttons to arrange the order of the SAs (see Figure 70). The SA on the top of the list is the primary SA and will be invoked first if SF needs to eliminate a node. Click on *DEFAULT* to use the recommended order for the SAs. Click on *Next*.

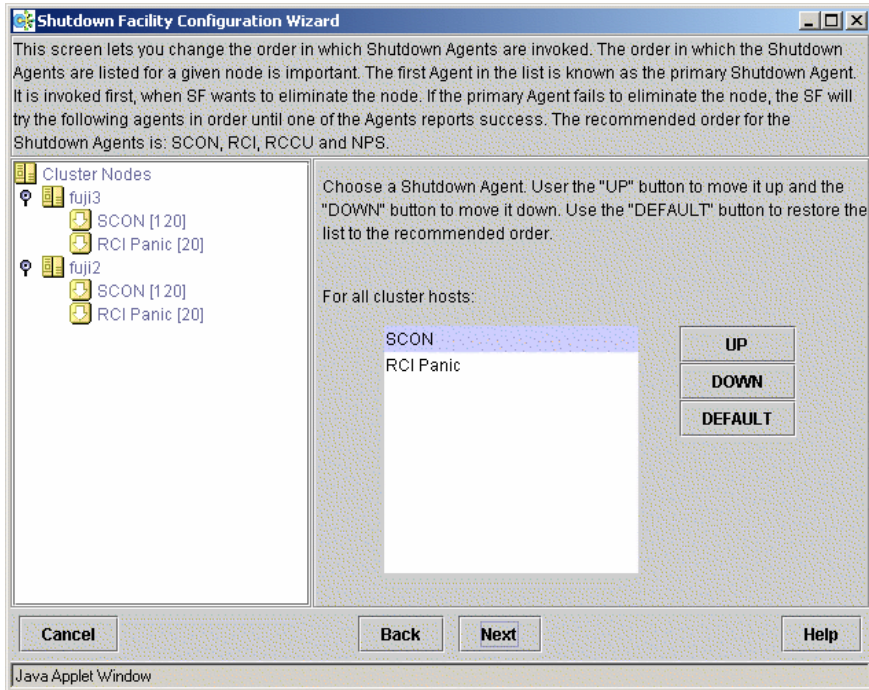


Figure 70: Order of the Shutdown Agents

The following screen lets you enter the timeout values for the configured SAs for each node (see Figure 71). Enter timeout values for all nodes and for each SA or click on the *Use Defaults* button. Select *Next* to go to the next screen.

The screenshot shows a Java Applet Window titled "Shutdown Facility Configuration Wizard". The main text reads: "This screen lets you enter the timeout values for the configured Shutdown Agents for each of the hosts." On the left, a tree view shows "Cluster Nodes" with two nodes: "fuji3" and "fuji2". Each node has two sub-items: "SCON [1 20]" and "RCI Panic [20]". The main area contains a table for entering timeout values:

CF Host Name	Shutdown Agent	Timeout(Sec)
fuji3	SCON	120
	RCI Panic	20
fuji2	SCON	120
	RCI Panic	20

Below the table is a "Use Defaults" button. At the bottom of the window are "Cancel", "Back", "Next", and "Help" buttons. The status bar at the bottom left says "Java Applet Window".

Figure 71: Shutdown Agent time-out values

The screen for entering node weights and administrative IP addresses appears (see Figure 72). Node weights should be an integer value greater than 0. You can select the *Admin IP* from the list of choices or enter your own. Enter node weights and *Admin IP* addresses for all CF nodes.

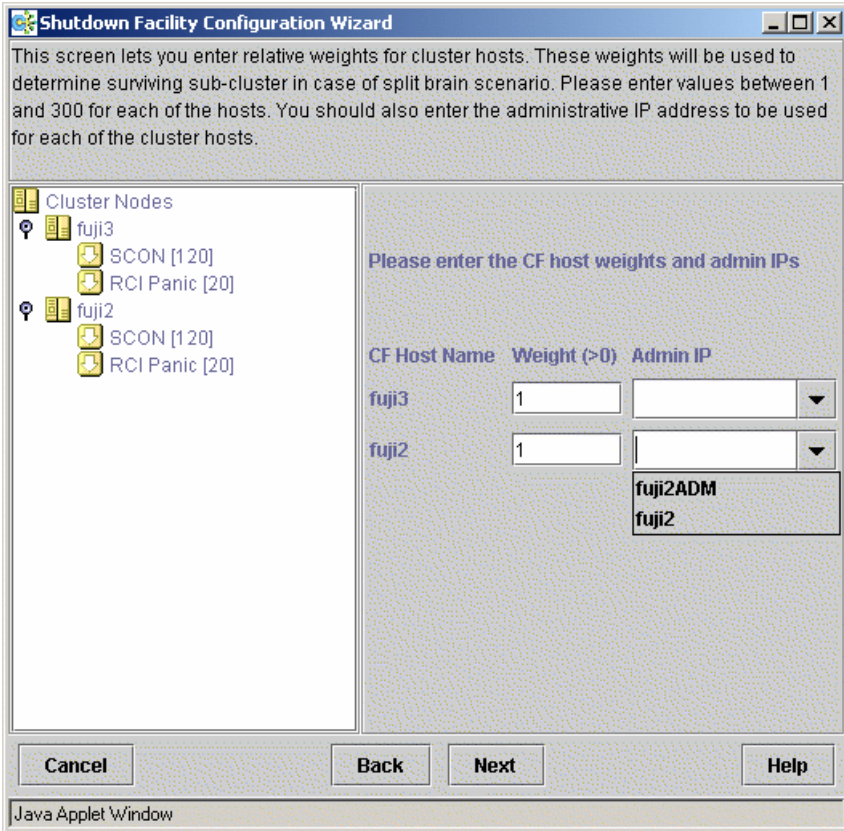


Figure 72: Entering host weights and admin IPs

For our cluster we will give each node an equal node weight of 1. The node weight and the `RMS userApplication` weight are used by SF to decide which subcluster should survive in the event of a network partition. The node weights are set in this screen. `RMS userApplication` weights, if used, are set in the RMS Wizards.

Set the *Admin IP* fields to the CF node's interface on the Administrative LAN. By convention, these IP interfaces are named *nodeADM*. although this is not mandatory. If you don't have an Administrative LAN, then enter the address to the public LAN. Click on *Next*.

The list of configuration files, created or edited, by the Wizard are shown in Figure 73. Click *Next* to save the configuration files or click *Back* to change the configuration.

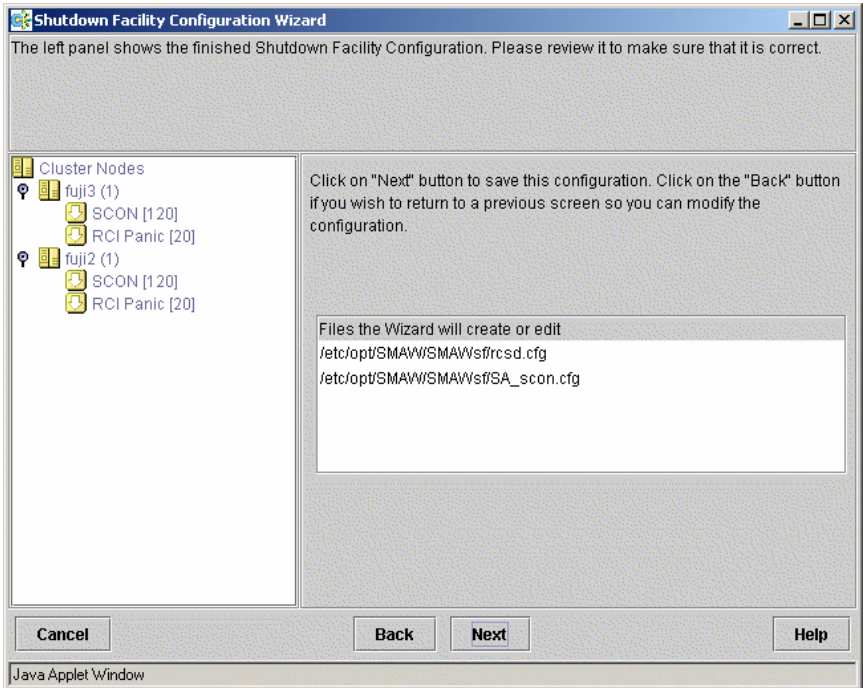


Figure 73: SF configuration files

Choose *Yes* in the confirmation popup to save the configuration (see Figure 74).



Figure 74: Saving SF configuration

The configuration status is shown in Figure 75. You can also use the *Tools* pulldown menu, and choose *Show Status* in the *Shutdown Facility* selection.

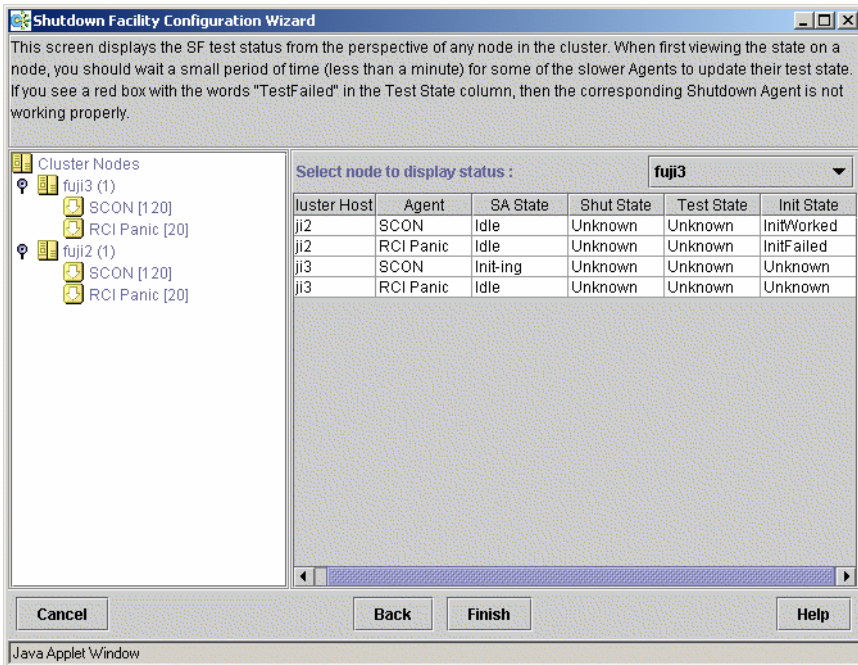


Figure 75: Status of Shutdown Agents

In the case of SAs, which have been configured but do not exist, the *Test State* will show as *Test Failed* in red.

SF has a test mechanism built into it. SF periodically has each SA verify that it can shut down cluster nodes. The SA does this by going through all the steps to shut down a node, except the very last one which would actually cause the node to go down. It then reports if the test was successful. This test is run for each node that a particular agent is configured to potentially shut down.

The table in Figure 75 shows, among other things, the results of these tests. The three columns *Cluster Host*, *Agent*, and *Test State*, when taken together in a single row, represent a test result. If the words *Test Failed* appear in red in the *Test State* column, it means that the agent found a problem when testing to see if it could shut down the node listed in the *Cluster Host* column. This indicates some sort of problem with the software, hardware, or networking resources used by that agent.

To exit the Wizard, click *Yes* at the popup, as shown in Figure 76. Before exiting, you may choose to re-edit the SF configuration.

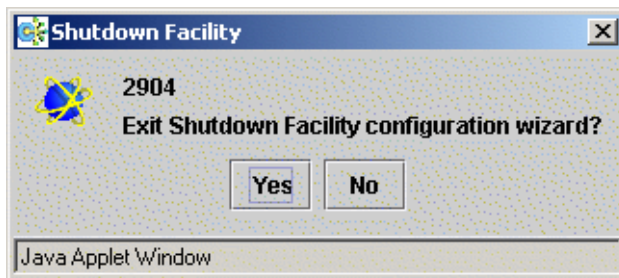


Figure 76: Exiting SF configuration wizard

8.4.2 Configuration via CLI

This section describes the setup and configuration via Command Line Interface (CLI).

i Note that the format of the configuration file is presented for information purposes only. The preferred method of configuring the shutdown facility and all SAs is to use the Cluster Admin GUI (refer to the Section “Configuring the Shutdown Facility”).

8.4.2.1 Shutdown daemon

To configure the Shutdown Daemon (SD), you will need to modify the file `/etc/opt/SMAW/SMAWsf/rcsd.cfg` on every node in the cluster.

A file, `rcsd.cfg.template`, is provided under the `/etc/opt/SMAW/SMAWsf` directory, which is a sample configuration file for the Shutdown Daemon using fictitious nodes and agents.



It is important that the `rcsd.cfg` file is identical on all cluster nodes; care should be taken in administration to ensure that this is true.

An example configuration for SD (which is created by editing the sample `rcsd.cfg.template`) follows:

```
#This file is generated by Shutdown Facility Configuration Wizard
#Generation Time : Sat Feb 22 10:32:06 PST 2003
fuj13,weight=1,admIP=fuj13ADM:agent=SA_scon,timeout=120:agent=SA_pprcip,timeout=
20:agent=SA_pprcir,timeout=20
fuj12,weight=1,admIP=fuj12ADM:agent=SA_scon,timeout=120:agent=SA_pprcip,timeout=
20:agent=SA_pprcir,timeout=20
```

The configuration file must be created in the `/etc/opt/SMAW/SMAWsf` directory and must use `rcsd.cfg` as the file name.

The format of the configuration file is as follows:

```
cluster-node1 [,weight=w1][,admIP=admIP1]:agent=SA1,
timeout=t1[:agent=SA2, timeout=T2]...
cluster-node2 [,weight=w2][,admIP=admIP2]:agent=SA1,
timeout=t1[:agent=SA2, timeout=T2]...
...
```

- *cluster-nodeN* is the *cfname* of a node within the cluster.
- *agent* and *timeout* are reserved words.
- *SA*N** is the command name of a SA.
- *t*N** is the maximum time in seconds that are allowed for the associated SA to run before assuming failure.
- *w*N** is the node weight.
- *admIP*N** is the admin interface on the Administrative LAN on this cluster node.

The order of the SAs in the configuration file should be such that the first SA in the list is the preferred SA. If this preferred SA is issued a shutdown request and if its response indicates a failure to shut down, the secondary SA is issued the shutdown request. This request/response is repeated until either an SA responds with a successful shutdown, or all SAs have been tried. If no SA is able to successfully shut down a cluster node, then operator intervention is required and the node is left in the `LEFTCLUSTER` state.

The location of the log file will be `/var/opt/SMAWsf/log/rcsd.log`.

8.4.2.2 Shutdown Agents

This section contains information on how to configure the SAs with CLI.

SCON

The configuration of the SA_scon SA involves creating a configuration file (`SA_scon.cfg`) in the correct format. The file must be located in:

`/etc/opt/SMAW/SMAWsf/SA_scon.cfg`

There exists a template file for use as an example (`SA_scon.cfg.template`) which resides in the `/etc/opt/SMAW/SMAWsf` directory.

The format of the `SA_scon.cfg` file is as follows:

`single-console-names` *Scon1* [*Scon2*] [...]

[`reply-ports-base` *number*]

`cluster-host` *cfname* *node-type*

- `single-console-names`, `reply-ports-base` and `cluster-host` are reserved words and must be in lower-case letters.
- *Scon1* is the IP name of the cluster console, *Scon2*, and ... are the names of additional cluster consoles for use in a distributed cluster console configuration.
- *number* is a port number used by SMAWRscon to reply to shutdown requests. The default value for this is 2137 and is used such that if you have four cluster nodes then the ports used on the all cluster nodes are 2137, 2138, 2139 and 2140. Note that setting `reply-ports-base` is optional.
- *cfname* is the CF name of a cluster node and *node-type* is the output of `uname -m` for that named cluster node. There must be one cluster-node line for each node in the cluster.
- *node type* for the named cluster node is the output from the following command:

```
# uname -m
```

The `SA_scon.cfg` file is as follows:

```
single-console-names fujiSCON1 fujiSCON2
cluster-host fuji1 sun4us
cluster-host fuji2 sun4us
cluster-host fuji3 sun4us
cluster-host fuji4 sun4us
```

SCON log file

`/var/opt/SMAWsf/log/SA_scon.log`

RCCU

To configure RCCU, you will need to create the following file:

`/etc/opt/SMAW/SMAWsf/SA_rccu.cfg`

A sample configuration file can be found in the following directory:

`/etc/opt/SMAW/SMAWsf/SA_rccu.cfg.template`

The configuration file `SA_rccu.cfg` contains lines that are in one of two formats: a line defining an attribute and value pair, or a line defining a cluster node setup.

- Lines defining attribute/value pairs

Attributes are similar to *global variables*, as they are values that are not modifiable for each RCCU unit, or each cluster node. Each line contains two fields: *Attribute-name Attribute-value*

The only currently supported attribute/value pair is the following:

```
Initial-connect-attempts positive_integer
```

This sets the number of connection retries for the first RCCU unit. The default value for the number of connection retries is 12.

- Lines defining a cluster node setup

Each line contains the following fields:

```
cfname RCCU-name RCCU-tty Control-port Console-port Password1 Password2
```

The fields are defined as follows:

- *cfname*: The CF name of the cluster node. This is the name assigned to the node when the cluster is created.

Use `cftool -l` to determine the CF name of each cluster node as follows:

```
# cftool -l
```

- *RCCU-name*: The IP name of the RCCU unit.
- *RCCU-tty*: The `tty` port on the *RCCU-name* to which *cfname* is connected.
- *Control-port*: The port to which one would telnet to the RCCU unit to access the RCCU control interface. When shipped from the factory, the default value for this port is 8010.
- *Console-port*: The port to which one would telnet to the RCCU unit to access the console line for the node *cfname*. When shipped from the factory, the default value for this port is 23.
- *Password1*: The password used by the user on the RCCU unit.
- *Password2*: The password used by the user admin on the RCCU unit.

RCCU log file

`/var/opt/SMAWsf/log/SA_rccu.log`



The permissions of the `SA_rccu.cfg` file are read/write by root only. This is to protect the password of a user and the admin user on the RCCU unit.

NPS

To configure NPS, you will need to create the following file:

`/etc/opt/SMAW/SMAWsf/SA_wtinps.cfg`

A sample configuration file can be found in the following directory:

`/etc/opt/SMAW/SMAWsf/SA_wtinps.cfg.template`

The configuration file `SA_wtinps.cfg` contains lines that are in one of two formats. A line defining an attribute and value pair or a line defining a plug set up:

- Lines defining attribute value pairs

Attributes are similar to global variables, as they are values that are not modifiable for each NPS unit, or each cluster node. Each line contains two fields:

Attribute-name Attribute-value

The currently supported attribute/value pairs are as follows:

`Initial-connect-attempts` *positive integer*

This sets the number of connect retries until the first connection to an NPS unit is made. The default value for the numbers of connect retries is 12.

- Lines defining a plug set up

Each line contains four fields:

Plug-ID IP-name Password Action

The four fields are:

- *Plug-ID*: The Plug-ID of the WTI NPS unit, which should correspond to a cluster node. The *CF_name* of the cluster node must be used here.
- *IP-name*: The IP name of the WTI NPS unit.
- *Password*: The password to access the WTI NPS unit.
- *Action*: The action may either be `cycle` or `leave-off`.



The *Plug-ID* defined in the `SA_wtinps.cfg` file must be defined on the WTI NPS unit.



The permissions of the `SA_wtinps.cfg` file are read/write by `root` only. This is to protect the password to the WTI NPS unit.

NPS log file

`/var/opt/SMAwsf/log/SA_wtinps.log`



NPS is not supported in all regions. Please check with your sales representative to see if the NPS is supported in your area.

An example of configuring the NPS SA is as follows:

```
# Configuration for Shutdown Agent for the WTI NPS
# Each line of the file has the format:
#
#Attribute-name Attribute-value
# - or -
#Plug-ID IP-name-of-WTI-box password {cycle|leave-off}
#
# Sample:
# initial-connect-attempts 12
# fuji2wtinps1.mycompany.comwtipwdcycle
# fuji3wtinps1.mycompany.comwtipwdleave-off
# fuji4wtinps2.mycompany.comnewpwdcycle
```



```
# fuji5wtinps2.mycompany.comnewpwdleave-off
#
# Note:
#The Plug-ID's that are specified here must be
#configured on the named WTI NPS unit.
#
# Note:
#The permissions on the file should be read/write
#only for root. This is to protect the password
#of the WTI NPS unit.
#
fuji2 nps6 mypassword cycle
fuji3 nps6 mypassword cycle
```

RPS

To configure RPS, you will need to create the following file:

```
/etc/opt/SMAW/SMAWsf/SA_rps.cfg
```

A sample configuration file can be found at the following location:

```
/etc/opt/SMAW/SMAWsf/SA_rps.cfg.template
```

The configuration file `SA_rps.cfg` contains lines with four fields (and some subfields) on each line. Each line defines a node in the cluster than can be powered-off (leaving it off) or powered-off and then on again. The fields are:

- *cfname*: The name of the node in the CF cluster. With redundant power supply, there may be more than one RPS necessary to power off one node. In this case, more than one entry with the same name will be needed.
- *Access-Information*: The access information is of the following format:
ip-address-of-unit[:*port*:*user*:*password*]
- The fields for *port*, *user*, and *password* can be missing, but not the corresponding colon. If a field (other than *port*) is missing, it must have a default value configured in the `rsb` software. The software `SMAWrsb` must be of version 1.2A0000 or later. The correct value for port is auto detected. It should always be omitted.
- *Index*: The index must be the index of the plug, which corresponds to the given Cluster-Node (the name of the node in the CF cluster).
- *Action*: The action may either be `cycle` or `leave-off`. If it is `cycle`, it will be powered on again after power off. If it is `leave-off`, a manual action is required to turn the system back on.



The permissions of the `SA_rps.cfg` file are read/write by root only.

RPS log file

`/var/opt/SMAWsf/log/SA_rps.log`

An example of configuring the RPS SA is as follows:

```
fuji2 172.0.111.221::root:rpspwd 1 cycle
fuji3 172.0.111.222::root:rpspwd 2 cycle
fuji4 172.0.111.223::root:rpspwd 3 leave-off
fuji5 172.0.111.224::root:rpspwd 4 leave-off
```

8.5 SF facility administration

This section provides information on administering SF. SF can be administered with the CLI or Cluster Admin. It is recommended to use Cluster Admin.

8.5.1 Starting and stopping SF

This section describes the following administrative procedures for starting and stopping SF:

- Manually via the CLI
- Automatically via the `rc` script interface

8.5.1.1 Starting and stopping SF manually

SF may be manually started or stopped by using the `sdtool(1M)` command. The `sdtool(1M)` command has the following options:

```
sdtool [-bcCsSrel] [-k CF-node-name] [-d off|on]
```

Refer to the Chapter “Manual pages” for more information on CLI commands.

8.5.1.2 Starting and stopping SF automatically

SF can be started automatically using the `S64rcfs` RC-script available under the `/etc/rc2.d` directory. The `rc` start/stop script for SF is installed as `/etc/init.d/RC_sf`.

8.6 Logging

Whenever there is a recurring problem where the cause cannot be easily detected, turn on the debugger with the `sdtool -d on` command. This will dump the debugging information into the `/var/opt/SMAWsf/log/rscd.log`, which will provide additional information to find the cause of the problem.

Note that the `rscd` log file does not contain logging information from any SA. Refer to the SA specific log files for logging information from a specific SA.

9 System console

This chapter discusses the `SCON` product functionality and configuration. The `SCON` product is installed on the cluster console.

This chapter discusses the following:

- The Section “Overview” discusses the role of the cluster console and the hardware platforms.
- The Section “Topologies” discusses the two distinct topologies imparting different configuration activities for the `SCON` product.
- The Section “Network considerations” notes the network configuration of both a single cluster console and distributed cluster console configuration.
- The Section “Configuration on the cluster console” discusses the steps necessary for the configuration on the cluster console.
- The Section “Updating a configuration on the cluster console” discusses updating the cluster console configuration after the addition or the removal of the cluster nodes.
- The Section “Configuration on the cluster nodes” discusses the recommended method of configuring the `SA_scon`, the Shutdown Agent, and the Shutdown Facility.
- The Section “Using the cluster console” explains how to access the consoles of individual cluster nodes.

9.1 Overview

This section discusses the `SCON` product functionality and configuration. The `SCON` product is installed on the cluster console.

9.1.1 Role of the cluster console

In `PRIMECLUSTER`, a cluster console is used to replace the consoles for standalone systems. This cluster console is used to provide a single point of control for all cluster nodes. In addition to providing administrative access, a cluster console runs the `SMAWRscon` software which performs needed node elimination tasks when required.

In most installations of PRIMECLUSTER a single cluster console can be used, but in some instances multiple cluster consoles must be configured in order to provide adequate administrative access to cluster nodes. The instances where multiple cluster consoles are needed are:

- When the cluster uses two or more PRIMEPOWER 800, 900, 1000, 1500, 2000, or 2500 cabinets which do not share a common system management console.
- When cluster nodes are separated by a large distance (more than what the cluster administrator deems to be reasonable) such that it would be unreasonable for them to share a common cluster console. This may be the case when the cluster nodes are placed far apart in order to provide a disaster recovery capability.

When two or more cluster consoles are used in a cluster it is called a distributed cluster console configuration. The pre-installation and installation steps for both the single cluster console and distributed cluster console are identical while the configuration step differs between the two.

9.1.2 Platforms

The cluster console is a generic term describing one of several hardware platforms on which the SCON product can run. The selection of a cluster console platform is in turn dependant on the platform of the cluster nodes:

- PRIMEPOWER 100, 200, 400, 600, 650, and 850 clusters nodes:

A cluster console is optional. If a cluster console is desired, use one of the following:

- RCA unit and a PRIMESTATION
- RCCU unit and a PRIMESTATION

- PRIMEPOWER 800, 900, 1000, 1500, 2000, 2500 clusters nodes

A cluster console is optional. If a cluster console is desired, it must be the System Management Console already present for the node.

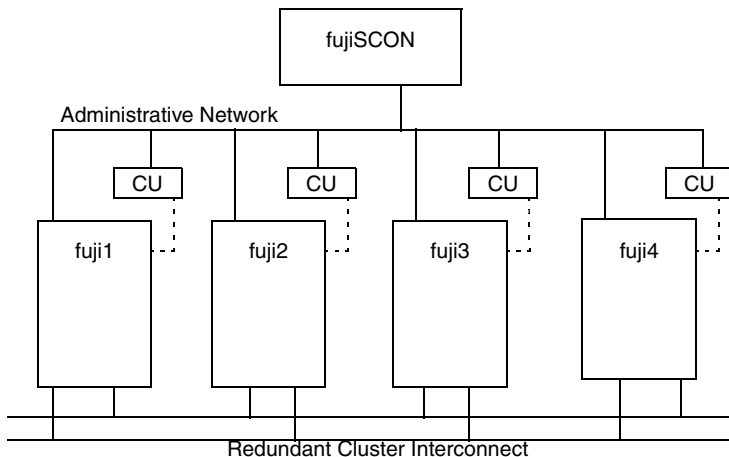
9.2 Topologies

The cluster console can be configured in two distinct topologies imparting different configuration activities for the SCON product. This section discusses the two topologies.

In both topologies, the console lines of the cluster nodes are accessible from the cluster console(s) via a serial-line-to-network converter unit. This unit may be one of several types supported in PRIMEPOWER clusters such as the RCA (Remote Console Access) or RCCU (Remote Console Control Unit). The SCON product does not differentiate between the units and as such their setup is not addressed in this manual. For information regarding specifics of these units, refer to your customer support center.

9.2.1 Single cluster console

A single cluster console configuration is one in which the console lines for all cluster nodes are accessible from one central cluster console as depicted in Figure 77. Note that the conversion unit (CU) in the diagram represents a generic conversion unit, which is responsible for converting serial-line to network access and represents either the RCA or RCCU units.



- - - - - Console Lines

Figure 77: Single cluster console

This single cluster console runs the SMAWRscon software which is responsible for performing the node elimination tasks for all nodes in the cluster. When configuring the single cluster console, all cluster nodes will be known to it and at runtime all cluster nodes will forward shutdown requests to it. SCON is responsible for node elimination tasks when the SA_scon Shutdown Agent is used.

9.2.2 Distributed cluster console

A distributed cluster console configuration is one in which there is more than one cluster console and each cluster console has access to a selected subset of the console lines for the cluster nodes. Note that the console line for each cluster node may only be accessed by one cluster console. A distributed cluster console configuration is depicted in Figure 78. Note that the CU in Figure 78 represents a generic conversion unit which is responsible for converting serial-line to network access and represents either the RCA or RCCU units.

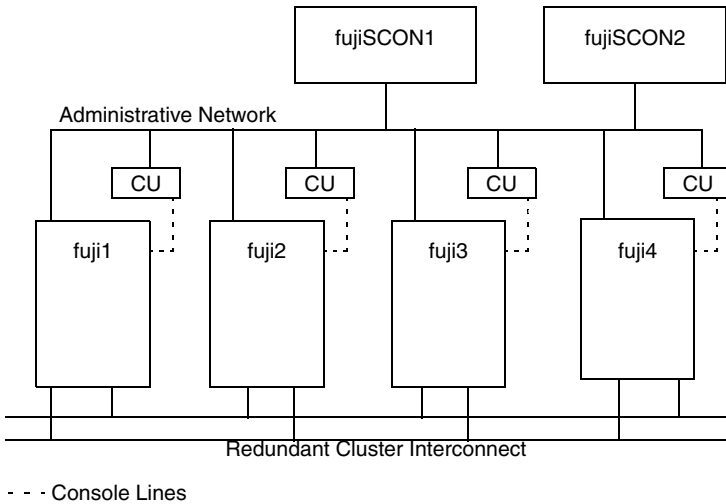


Figure 78: Distributed cluster console

In our example, fujiSCON1 controls access to fuji1 and fuji2 and fujiSCON2 controls access to fuji3 and fuji4. When configuring the SCON product on fujiSCON1 only fuji1 and fuji2 will be known by it, similarly on fujiSCON2 the SCON product will only know of fuji3 and fuji4.

At runtime, all shutdown requests are sent to all cluster consoles and the cluster console responsible for the node being shut down performs the work and responds to the request.

9.3 Network considerations

There are several things to note in regards to the network configuration of both a single cluster console and distributed cluster console configuration:

- The cluster console(s) are not on the cluster interconnect.
- All CUs, cluster consoles, and cluster nodes are on an administrative network.
- The administrative network is physically separate from the public network(s).

9.4 Configuration on the cluster console

The configuration on the cluster console consists of several steps:

- Updating the `/etc/hosts` file
- Running the `Configure` script
- Optionally editing the `rmshosts` file

9.4.1 Updating the `/etc/hosts` file

The cluster console must know the IP address associated with the CF name of each cluster node. In most cases the CF name of the cluster node is the same as the `uname -n` of the cluster node, but in other cases the cluster administrator has chosen a separate CF name for each cluster node that does not match the `uname -n`.

For each cluster node, using the editor of your choice, add an entry to the `/etc/hosts` file for each CF name so that the cluster console can communicate with the cluster node. The CF name must be used because the Shutdown Facility on each cluster node and the cluster console communicate using only CF names.



Note that when working with a distributed cluster console configuration, all cluster consoles must have an entry for each cluster node, regardless of which cluster console administers which sub-set of cluster nodes.

As an example, referring to our sample FUJI cluster (refer to the PRIMECLUSTER *Installation Guide (Solaris)*, “Cluster site planning worksheet”), the CF name of the cluster nodes are `fujij2` and `fujij3` which happen to match the public IP names of their nodes. Since the cluster console (`fujijSCON`) is on the administration network and on the public network then `fujijSCON` can directly contact the cluster nodes by using the CF names because they happen to match the public IP names of the nodes. So in our sample cluster, no extra `/etc/hosts` work will need to be done.

This setup may not always be the case because the administrator may have chosen that the cluster console will not be accessible on the public network, or the CF names do not match the public IP names. In either of these cases then aliases would have to be set up in the `/etc/hosts` file so that the cluster console can contact the cluster nodes using the CF name of the cluster node. Assume that the sample FUJI cluster chose CF names of `fujij2cf` and `fujij3cf` (instead of `fujij2` and `fujij3`), then entries in the `/etc/hosts` file would have to be made that look like:

```
172.25.200.4 fujij2ADM fujij2cf
172.25.200.5 fujij3ADM fujij3cf
```

9.4.2 Running the Configure script

The configuration of the SCON product is slightly different depending on the platform of the cluster nodes.

If the cluster consists of a PRIMEPOWER 800, 900, 1000, 1500, 2000, or 2500 node, the script will derive the partition information from the partition tables on the management console. It will place the correct entries into the `/etc/uucp/Systems` and `/etc/uucp/Devices` files and install symbolic link under `/dev`.

If the cluster consists of PRIMEPOWER 100, 200, 400, 600, 650, or 850 nodes, then the entries in the `/etc/uucp/Systems` and `/etc/uucp/Devices` files are already present. They were created when performing the setup of the cluster console.

Enter the following to run the `Configure` script:

```
# /opt/MAW/MAWRscon/bin/Configure
```

i Note that running the `Configure` script when working with a distributed cluster console will only show the sub-set of cluster nodes that are administered by the local cluster console. The sub-set of cluster nodes administered by other cluster consoles will not appear in the output of the `Configure` script. This is true regardless of the platform type of the cluster nodes.

The `Configure` script will ask several questions regarding the cluster console configuration, typically one can use the default response which is selected by using a carriage return.

9.4.3 Editing the `rmshosts` file

The `/opt/SMAW/SMAWRscon/etc/rmshosts` file contains the list of cluster nodes that are configured on the local cluster console. The order in which the nodes appear in the file are treated as a priority list in the event of a split-cluster.

If you want to change the priority of cluster nodes, you can reorder them. When reordering the node names, ensure that all node names are spelled correctly and that all nodes in the cluster are included in the file. The priority is taken from here only when the default weights for the cluster nodes are used.

After editing, or overwriting, the `rmshosts` file all processes associated with the SCON product must be restarted. This can be done by either rebooting the cluster console or by using the `ps` command to find all related processes and issuing them a `SIGKILL`:

```
# kill -KILL `ps -elf | grep scon | grep -v grep | awk -e '{print $4}'`
```

9.4.4 Additional steps for distributed cluster console

The SCON product arbitrates between sub-sets of cluster nodes in a distributed cluster console configuration. In order for this to occur correctly, the list of cluster nodes in the `rmshosts` file on all cluster consoles must be a complete list of all cluster nodes and all cluster nodes must appear in the same order.

Update the `rmshosts` file by editing the `/opt/SMAW/SMAWRscon/etc/rmshosts` file and adding a line with the CF name of all cluster nodes that are not listed.

9.4.5 rmshosts method file

The entries in this file determine whether the SCON does split-cluster processing before eliminating a node. By default a `no` entry of the form `cfname uucp no` causes split-cluster processing before eliminating a node, and a `yes` entry does not allow split-cluster processing to be done.



This file needs to be edited only if you are using other Shutdown Agents along with SCON.

Change the entries of the following form:

```
cfname uucp no
```

to

```
cfname uucp yes
```



Make sure that the number and names of cluster nodes are consistent across `rmshosts` and the `rmshosts.method` file. In the case of distributed console, they should be consistent across all console nodes.

9.5 Updating a configuration on the cluster console

Once a cluster is configured with a cluster console, if cluster nodes are added or removed the cluster console configuration must be updated to reflect the new cluster. Modifying the cluster console configuration will be different, depending on the platform of the cluster nodes:

- PRIMEPOWER 100, 200, 400, 600, 650, and 850 clusters:
 - Perform the needed setup of the cluster console hardware as defined. See instructions specific to the cluster console hardware at your site.
 - Re-run the `Configure` script.
- PRIMEPOWER 800, 900, 1000, 1500, 2000, 2500 clusters:
 - Remove all entries for that refer to partitions (having `SconK` as part of their tags) from the `/etc/uucp/Systems` and `/etc/uucp/Devices` files. For configurations that use CF names different from unames, remove the comments inserted earlier by the `Configure` script.
 - Re-run the `Configure` script.

9.6 Configuration on the cluster nodes

The recommended method of configuring the `SA_scon` and the Shutdown Facility is to use the Cluster Admin GUI.

9.6.1 Manually configuring the SCON SA

Information on manual configuration is presented here for those who choose to do so.

This section contains other information in addition to `SA_scon` Shutdown Agent, and the Shutdown Facility configuration. Please be sure to review all sections and apply those that are relevant to your cluster.

9.6.2 Configuration of the Shutdown Facility

In order for the Shutdown Facility to begin using the `SA_scon` Shutdown Agent and the Shutdown Facility must be configured properly. Please refer to Section “Configuring the Shutdown Facility” for more information.

9.6.3 Other configuration of the cluster nodes



Note that this section describes work that only needs to be done on cluster nodes that are PRIMEPOWER 100, 200, 400, 600, 650, or 850 nodes. This configuration work must not be done on the 800, 900, 1000, 1500, 2000, and 2500 models.

In addition to the configuration of the `SA_scon` Shutdown Agent and Shutdown Facility there may be additional configuration work needed on the cluster nodes in order to make them work with the SCON product.

9.6.3.1 Redirecting console input/output

Most likely the console input and output have already been redirected as part of the hardware setup of the cluster console. This information is provided as a backup.

Use the `eeeprom` command to modify the `input-device`, `output-device`, and `ttya-mode` settings on the nodes boot prom as follows:

```
# eeeprom input-device=ttya
# eeeprom output-device=ttya
# eeeprom ttya-mode=9600,8,n,1,-
```

9.6.3.2 Booting with `kadb`

Ensure that the cluster nodes boot using `kadb` by using the `eeeprom` command to set the boot file to `kadb`. The command is as follows:

```
# eeeprom boot-file=kadb
```

Restrictions

PRIMEPOWER nodes only reboot automatically after a panic if the setting of the `eeeprom` variable `boot-file` is not `kadb`. The `SCON` kill on PRIMEPOWER 200, 400, 600, 650, 850 nodes requires the `kadb` setting. An automatic reboot after panic is not possible on those nodes if the elimination via panic is supposed to be a fall back elimination method after a failing `SCON` elimination.

Setting the alternate keyboard abort sequence

Edit the `/etc/default/kbd` file and ensure that the line defining the keyboard abort sequence is uncommented and set to the alternate abort sequence. The line should look exactly like the following:

```
KEYBOARD_ABORT=alternate
```

9.7 Using the cluster console

This section explains how to access the consoles of individual cluster nodes.



Note that this function is only available on clusters of PRIMEPOWER 100, 200, 400, 600, 650, and 850 nodes. The console access for 800, 900, 1000, 1500, 2000, and 2500 nodes are handled through the system management software.

9.7.1 Without XSCON

The `SCON` `Configure` script automatically starts the `SMAWRscon` software running on the cluster console. Since this software is already running, all the administrator needs to do in order to get a console window for each cluster node is to use the `xco` utility to start a console window as follows:

```
# /opt/SMAW/SMAWRscon/bin/xco cfname
```

cfname is the CF name of a cluster node.

9.7.2 With XSCON

The console window can be accessed using the `SMAWxscn` software by setting the `XSCON_CU` environment variable in the administrators environment. It must be set to: `/opt/SMAW/SMAWRscon/bin/scn.scr`. As an example in `korn` shell:

```
# export XSCON_CU=/opt/SMAW/SMAWRscon/bin/scn.scr
```

The `xco` utility will use the `SCON` command to open windows in this environment.

10 CF over IP

This chapter discusses CF over IP and how it is configured.

This chapter discusses the following:

- The Section “Overview” introduces CF over IP and describes its use.
- The Section “Configuring CF over IP” details how to configure CF over IP.

10.1 Overview

i All IP configuration must be done prior to using CF over IP. The devices must be initialized with a unique IP address and a broadcast mask. IP must be configured to use these devices. If the configuration is not done, `cfconfig(1M)` will fail to load CF, and CF will not start.

CF communications are based on the use of interconnects. An interconnect is a communications medium which can carry CF's link-level traffic between the CF nodes. A properly configured interconnect will have connections to all of the nodes in the cluster through some type of device. This is illustrated in Figure 79.

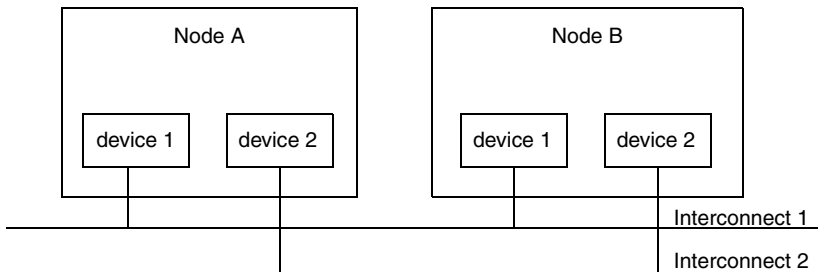


Figure 79: Conceptual view of CF interconnects

When CF is used over Ethernet, Ethernet devices are used as the interfaces to the interconnects. The interconnects themselves are typically Ethernet hubs or switches. An example of this is shown in Figure 80.

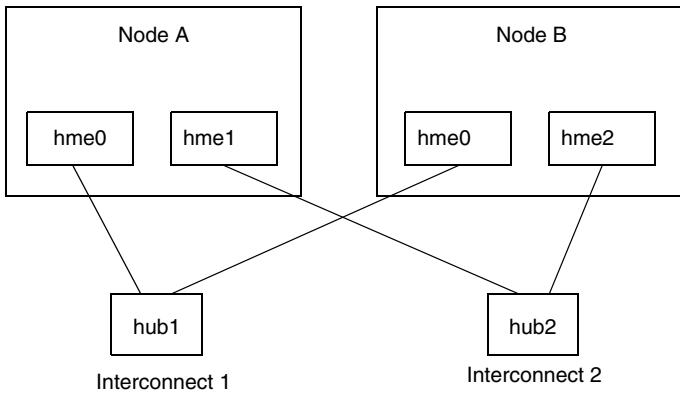


Figure 80: CF with Ethernet interconnects

When CF is run over IP, IP interfaces are the devices used to connect to the interconnect. The interconnect is an IP subnetwork. Multiple IP subnetworks may be used for the sake of redundancy. Figure 81 shows a CF over IP configuration.

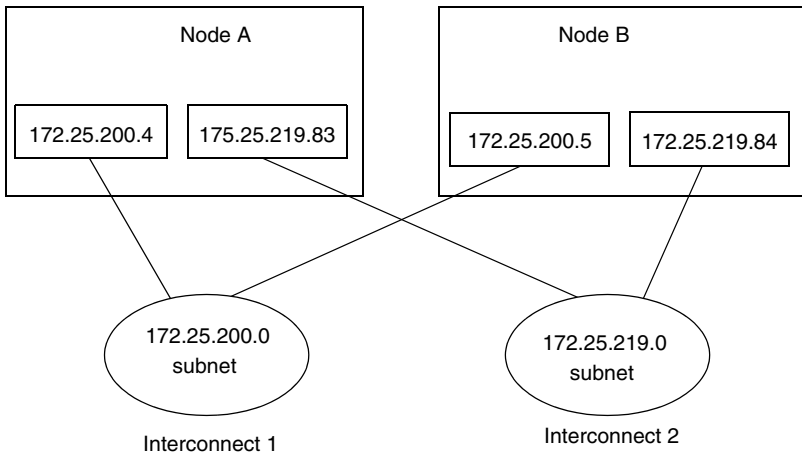


Figure 81: CF with IP interconnects

It is also possible to use mixed configurations in which CF is run over both Ethernet devices and IP subnetworks.

When using CF over IP, you should make sure that each node in the cluster has an IP interface on each subnetwork used as an interconnect. You should also make sure that all the interfaces for a particular subnetwork use the same IP broadcast address and the same netmask on all cluster nodes. This is particularly important since CF depends on an IP broadcast on each subnet to do its initial cluster join processing.



The current version does not allow CF to reach nodes that are on different subnets.



Caution

When selecting a subnetwork to use for CF, you should use a private subnetwork that only cluster nodes can access. CF security is based on access to its interconnects. Any node that can access an interconnect can join the cluster and acquire root privileges on any cluster node. When CF over IP is used, this means that any node on the subnetworks used by CF must be trusted. You should not use the public interface to a cluster node for CF over IP traffic unless you trust every node on your public network.

10.2 Configuring CF over IP

To configure CF over IP, you should do the following:

- Designate which subnetworks you want to use for CF over IP. Up to four subnetworks can be used.
- Make sure that each node that is to be in the cluster has IP interfaces properly configured for each subnetwork. Make sure the IP broadcast and netmasks are correct and consistent on all nodes for the subnetworks.
- Make sure that all of these IP interfaces are up and running.
- Run the CF Wizard in Cluster Admin.

The CF Wizard has a screen which allows CF over IP to be configured. The Wizard will probe all the nodes that will be in the cluster, find out what IP interfaces are available on each, and then offer them as choices in the CF over IP screen. It will also try to group the choices for each node by subnetworks. See Section “CF, CIP, and CIM configuration” in Chapter “Cluster Foundation” for details.

CF uses special IP devices to keep track of CF over IP configuration. There are four of these devices named as follows:

```
/dev/ip0  
/dev/ip1  
/dev/ip2  
/dev/ip3
```

These devices do not actually correspond to any device files under `/dev` in Solaris. Instead, they are just place holders for CF over IP configuration information within the CF product. Any of these devices can have an IP address and broadcast address assigned by the `cfconfig(1M)` command (or by Cluster Admin which invokes the `cfconfig(1M)` command in the Wizard).

If you run `cfconfig(1M)` by hand, you may specify any of these devices to indicate you want to run CF over IP. The IP device should be followed by an IP address and broadcast address of an interface on the local node. The addresses must be in internet dotted-decimal notation. For example, to configure CF on Node A in Figure 81, the `cfconfig(1M)` command would be as follows:

```
# cfconfig -S A clustername /dev/ip0 \  
172.25.200.4 172.25.200.255 /dev/ip1 172.25.219.83
```

It really does not matter which IP device you use. The above command could equally have used `/dev/ip2` and `/dev/ip3`.



The `cfconfig(1M)` command does not do any checks to make sure that the IP addresses are valid.

The IP devices chosen in the configuration will appear in other commands such as `cftool -d` and `cftool -r`.

IP interfaces will not show up in CF pings using `cftool -p` unless they are configured for use with CF and the CF driver is loaded.



`cftool -d` shows a relative speed number for each device, which is used to establish priority for the message send. If the configured device is IP, the relative speed 100 is used. This is the desired priority for the logical IP device. If a Gigabit Ethernet hardware device is also configured, it will have priority.

11 Diagnostics and troubleshooting

This chapter provides help for troubleshooting and problem resolution for PRIMECLUSTER Cluster Foundation. This chapter will help identify the causes of problems and possible solutions. If a problem is in another component of the PRIMECLUSTER suite, the reader will be referred to the appropriate manual. This chapter assumes that the installation and verification of the cluster have been completed as described in the PRIMECLUSTER *Installation Guide (Solaris)*.

This chapter discusses the following:

- The Section “Beginning the process” discusses collecting information used in the troubleshooting process.
- The Section “Symptoms and solutions” is a list of common symptoms and the solutions to the problems.
- The Section “Collecting troubleshooting information” gives steps and procedures for collecting troubleshooting information.

11.1 Beginning the process

Start the troubleshooting process by gathering information to help identify the causes of problems. You can use the CF log viewer facility from the Cluster Admin GUI, look for messages on the console, or look for messages in the `/var/adm/messages` file. You could use the `cftool(1M)` command for checking states, configuration information. To use the CF log viewer click on the *Tools* pull-down menu and select *View Syslog messages*. The log messages are displayed. You may search the logs using a date/time filter or scan for messages based on severity levels. To search based on date/time, use the date/time filter and press the *Filter* button. To search based on severity levels, click on the *Severity* button and select the desired severity level. You can use keyword also to search the log. To detach the CF log viewer window, click on the *Detach* button; click on the *Attach* button to attach it again.

Collect information as follows:

- Look for messages on the console that contain the identifier CF.
- Look for messages in `/var/adm/messages`. You might have to look in multiple files (`/var/adm/messages.M`).

- Use `cftool`.
 - `cftool -l`: Check local node state
 - `cftool -d`: Check device configuration
 - `cftool -n`: Check cluster node states
 - `cftool -r`: Check the route status

Error log messages from CF are always placed in the `/var/adm/messages` file; some messages may be replicated on the console. Other device drivers and system software may only print errors on the console. To have a complete understanding of the errors on a system, both console and error log messages should be examined. The Section “Alphabetical list of messages” contains messages that can be found in the `/var/adm/messages` file. This list of messages gives a description of the cause of the error. This information is a good starting point for further diagnosis.

All of the parts of the system put error messages in this file or on the console and it is important to look at all of the messages, not just those from the PRIMECLUSTER suite. The following is an example of a CF error message from the

`/var/adm/messages` file:

```
Nov  9 08:51:45 fuji2 unix: LOG3.0973788705 1080024 1008 4
0 1.0 cf:ens CF: Icf Error: (service err_type
route_src route_dst). (0 0 0 0 0 0 0 2 0 0 0 5 0 0 0 5)
```

The first 80 bytes are the `log3` prefix as in the following:

```
Nov  9 08:51:45 fuji2 unix: LOG3.0973788705 1080024 1008 4
0 1.0 cf:ens
```

This part of the message is a standard prefix on each CF message in the log file that gives the date and time, the node name, and `log3` specific information. Only the date, time, and node name are important in this context. The remainder is the error message from CF as in the following:

```
CF: Icf Error: (service err_type route_src route_dst). (0 0 0 0
0 0 0 0 2 0 0 0 5 0 0 0 5)
```

This message is from the `cf:ens` service (that is, the Cluster Foundation, Event Notification Service) and the error is `CF: Icf Error`. This error is described in Chapter “CF messages and codes”, Section “Alphabetical list of messages” as signifying a missing heartbeat and/or a route down. This gives us direction to look into the cluster interconnect further. A larger piece of the `/var/adm/messages` file shows as follows:

```
fuji2# tail /var/adm/messages
```

```
Nov 9 08:51:45 fuji2 unix: SUNW,pci-gem1: Link Down - cable problem?
Nov 9 08:51:45 fuji2 unix: SUNW,pci-gem0: Link Down - cable problem?
Nov 9 08:51:45 fuji2 unix: LOG3.0973788705 1080024 1008 4 0 1.0
cf:ens CF: Icf Error: (service err_type route_src route_dst). (0 0 0 0
0 0 0 0 2 0 0 0 5 0 0 0 5)
Nov 9 08:51:46 fuji2 unix: SUNW,pci-gem0: Link Down - cable problem?
Nov 9 08:51:48 fuji2 last message repeated 1 time
Nov 9 08:51:48 fuji2 unix: LOG3.0973788708 1080024 1008 4 0 1.0
cf:ens CF: Icf Error: (service err_type route_src route_dst). (0 0 0 0
0 0 0 0 2 0 0 0 4 0 0 0 4)
Nov 9 08:51:50 fuji2 unix: SUNW,pci-gem0: Link Down - cable problem?
Nov 9 08:51:52 fuji2 last message repeated 1 time
Nov 9 08:51:53 fuji2 unix: LOG3.0973788713 1080024 1008 4 0 1.0
cf:ens CF: Icf Error: (service err_type route_src route_dst). (0 0 0 0
0 0 0 0 2 0 0 0 4 0 0 0 4)
Nov 9 08:51:53 fuji2 unix: LOG3.0973788713 1080024 1015 5 0 1.0
cf:ens CF: Node fuji2 Left Cluster POKE. (0 0 2)
Nov 9 08:51:53 fuji2 unix: Current Nodee Status = 0
```

Here we see that there are error messages from the Ethernet controller indicating that the link is down, possibly because of a cable problem. This is the clue we need to solve this problem; the Ethernet used for the interconnect has failed for some reason. The investigation in this case should shift to the cables and hubs to insure that they are all powered up and securely connected.

Several options for the command `cftool` are listed above as sources for information. Some examples are as follows:

```
fuji2# cftool -l
```

Node	Number	State	Os	Cpu
fuji2	2	UP	Solaris	Sparc

This shows that the local node has joined a cluster as node number 2 and is currently UP. This is the normal state when the cluster is operational. Another possible response is as follows:

```
fuji2# cftool -l
```

Node	Number	State	Os
fuji2	--	COMINGUP	--

This indicates that the CF driver is loaded and that the node is attempting to join a cluster. If the node stays in this state for more than a few minutes, then something is wrong and we need to examine the `/var/adm/messages` file. In this case, we see the following:

```
fuji2# tail /var/adm/messages
```

```

May 30 17:36:39 fuji2 unix: pseudo-device: fcp0
May 30 17:36:39 fuji2 unix: fcp0 is /pseudo/fcp@0
May 30 17:36:53 fuji2 unix: LOG3.0991269413 1080024 1007 5
0 1.0 cf:eventlog CF: (TRACE): JoinServer:
Startup.
May 30 17:36:53 fuji2 unix: LOG3.0991269413 1080024 1009 5
0 1.0 cf:eventlog CF: Giving UP Mastering
(Cluster already Running).
May 30 17:36:53 fuji2 unix: LOG3.0991269413 1080024 1006 4
0 1.0 cf:eventlog CF: fuji4: busy: local node not
DOWN: retrying.

```

We see that this node is in the LEFTCLUSTER state on another node (fuji4). To resolve this condition, see Chapter “GUI administration” for a description of the LEFTCLUSTER state and the instructions for resolving the state.

The next option to `cftool` shows the device states as follows:

```
fuji2# cftool -d
```

Number	Device	Type	Speed	Mtu	State	Configured	Address
1	/dev/hme0	4	100	1432	UP	YES	00.80.17.28.21.a6
2	/dev/hme3	4	100	1432	UP	YES	08.00.20.ae.33.ef
3	/dev/hme4	4	100	1432	UP	YES	08.00.20.b7.75.8f
4	/dev/ge0	4	1000	1432	UP	YES	08.00.20.b2.1b.a2
5	/dev/ge1	4	1000	1432	UP	YES	08.00.20.b2.1b.b5

Here we can see that there are two interconnects configured for the cluster (the lines with YES in the Configured column). This information shows the names of the devices and the device numbers for use in further troubleshooting steps.

The `cftool -n` command displays the states of all the nodes in the cluster. The node must be a member of a cluster and UP in the `cftool -l` output before this command will succeed as shown in the following:

```
fuji2# cftool -n
```

Node	Number	State	Os	Cpu
fuji2	1	UP	Solaris	Sparc
fuji3	2	UP	Solaris	Sparc

This indicates that the cluster consists of two nodes `fuji2` and `fuji3`, both of which are UP. If the node has not joined a cluster, the command will wait until the join succeeds.

`cftool -r` lists the routes and the current status of the routes as shown in the following example:

```
fuji2# cftool -r
```


Node	Number	Srcdev	Dstdev	Type	State	Destaddr
fuji2	1	4	4	4	UP	08.00.20.b2.1b.cc
fuji2	1	5	5	4	UP	08.00.20.b2.1b.94
fuji3	2	4	4	4	UP	08.00.20.b2.1b.a2
fuji3	2	5	5	4	UP	08.00.20.b2.1b.b5

This shows that all of the routes are UP. If a route shows a DOWN state, then the step above where we examined the error log should have found an error message associated with the device. At least the CF error noting the route is down should occur in the error log. If there is not an associated error from the device driver, then the diagnosis steps are covered below.

The last route to a node is never marked DOWN, it stays in the UP state so that the software can continue to try to access the node. If a node has left the cluster or gone down, there will still be an entry for the node in the route table and one of the routes will still show as UP. Only the `cftool -n` output shows the state of the nodes as shown in the following:

```
fuji2# cftool -r
```

Node	Number	Srcdev	Dstdev	Type	State	Destaddr
fuji2	2	3	2	4	UP	08.00.20.bd.5e.a1
fuji3	1	3	3	4	UP	08.00.20.bd.60.e4

```
fuji2# cftool -n
```

Node	Number	State	Os	Cpu
fuji2	2	UP	Solaris	Sparc
fuji3	1	LEFTCLUSTER	Solaris	Sparc

11.2 Symptoms and solutions

The previous section discussed the collection of data. This section discusses symptoms and gives guidance for troubleshooting and resolving the problems. The problems dealt with in this section are divided into two categories: problems with joining a cluster and problems with routes, either partial or complete loss of routes. The solutions given here are either to correct configuration problems or to correct interconnect problems. Problems outside of these categories or solutions to problems outside of this range of solutions are beyond the scope of this manual and are either covered in another product's manual or require

technical support from your customer service representative. Samples from the error log (`/var/adm/messages`) have the `log3` header stripped from them in this section.

11.2.1 Join-related problems

Join problems occur when a node is attempting to become a part of a cluster. The problems covered here are for a node that has previously successfully joined a cluster. If this is the first time that a node is joining a cluster, the *PRIMECLUSTER Installation Guide (Solaris)* section on verification covers the issues of initial startup. If this node has previously been a part of the cluster and is now failing to rejoin the cluster, here are some initial steps in identifying the problem.

First, look in the error log and at the console messages for any clue to the problem. Have the Ethernet drivers reported any errors? Any other unusual errors? If there are errors in other parts of the system, the first step is to correct those errors. Once the other errors are corrected, or if there were no errors in other parts of the system, proceed as follows.

Is the CF device driver loaded? The device driver puts a message in the log file when it loads and the `cftool -l` command will indicate the state of the driver. The logfile message looks as follows:

```
CF: (TRACE): JoinServer: Startup.
```

`cftool -l` prints the state of the node as follows:

```
fuji2# cftool -l
```

```
Node      Number State      0s
fuji2 --   COMINGUP  --
```

This indicates the driver is loaded and the node is trying to join a cluster. If the errorlog message above does not appear in the logfile or the `cftool -l` command fails, then the device driver is not loading. If there is no indication in the `/var/adm/messages` file or on the console why the CF device driver is not loading, it could be that the CF kernel binaries or commands are corrupted, and you might need uninstall and reinstall CF. Before any further steps can be taken, the device driver must be loaded.

After the CF device driver is loaded, it attempts to join a cluster as indicated by the message “CF: (TRACE): JoinServer: Startup.”. The join server will attempt to contact another node on the configured interconnects. If one or more other nodes have already started a cluster, this node will attempt to join that cluster. The following message in the error log indicates that this has occurred:

```
CF: Giving UP Mastering (Cluster already Running).
```

If this message does not appear in the error log, then the node did not see any other node communicating on the configured interconnects and it will start a cluster of its own. The following two messages will indicate that a node has formed its own cluster:

```
CF: Local Node fuji2 Created Cluster FUJI. (#0000 1)
CF: Node fuji2 Joined Cluster FUJI. (#0000 1)
```

At this point, we have verified that the CF device driver is loading and the node is attempting to join a cluster. In the following list, problems are described with corrective actions. Find the problem description that most closely matches the symptoms of the node being investigated and follow the steps outlined there.



Note that the `log3` prefix is stripped from all of the error message text displayed below. Messages in the error log will appear as follows:

```
Mar 10 09:47:55 fuji2 unix: LOG3.0952710475 1080024 1014 4
0 1.0 cf:ens
CF: Local node is missing a route from node: fuji3
```

However they are shown here as follows:

```
CF: Local node is missing a route from node: fuji3
```

Join problems

Problem:

The node does not join an existing cluster, it forms a cluster of its own.

Diagnosis:

The error log shows the following messages:

```
CF: (TRACE): JoinServer: Startup.
CF: Local Node fuji4 Created Cluster FUJI. (#0000 1)
CF: Node fuji2 Joined Cluster FUJI. (#0000 1)
```

This indicates that the CF devices are all operating normally and suggests that the problem is occurring some place in the interconnect. The first step is to determine if the node can see the other nodes in the cluster over the interconnect. Use `cftool` to send an echo request to all the nodes of the cluster:

```
fuji2# cftool -e
```

```
Localdev Srcdev Address Cluster Node Number Joinstate
3 2 08.00.20.bd.5e.a1 FUJI fuji2 2 6
3 3 08.00.20.bd.60.ff FUJI fuji3 1 6
```

This shows that node `fuji3` sees node `fuji2` using interconnect device 3 (Localdev) on `fuji3` and device 2 (Srcdev) on `fuji2`. If the `cftool -e` shows only the node itself then look under the Interconnect Problems heading for the problem “The node only sees itself on the configured interconnects.” If some or all of the expected cluster nodes appear in the list, attempt to rejoin the cluster by unloading the CF driver and then reloading the driver as follows:

```
fuji2# cfconfig -u
```

```
fuji2# cfconfig -l
```



There is no output from either of these commands, only error messages in the error log.

If this attempt to join the cluster succeeds, then look under the Problem: “The node intermittently fails to join the cluster.” If the node did not join the cluster then proceed with the problem below “The node does not join the cluster and some or all nodes respond to `cftool -e`.”

Problem:

The node does not join the cluster and some or all nodes respond to `cftool -e`.

Diagnosis:

At this point, we know that the CF device is loading properly and that this node can communicate to at least one other node in the cluster. We should suspect at this point that the interconnect is missing messages. One way to test this hypothesis is to repeatedly send echo requests and see if the result changes over time as in the following example:

```
fuji2# cftool -e
```

```
Localdev Srcdev Address Cluster Node Number Joinstate
3 2 08.00.20.ae.33.ef FUJI fuji1 3 6
3 2 08.00.20.bd.5e.a1 FUJI fuji2 2 6
3 3 08.00.20.bd.60.ff FUJI fuji3 1 6
```

fuji2# cftool -e

```
Localdev Srcdev Address Cluster Node Number Joinstate
3 2 08.00.20.ae.33.ef FUJI fuji1 3 6
3 2 08.00.20.bd.5e.a1 FUJI fuji2 2 6
3 3 08.00.20.bd.60.ff FUJI fuji3 1 6
3 3 08.00.20.bd.60.e4 FUJI fuji4 1 6
```

fuji2# cftool -e

```
Localdev Srcdev Address Cluster Node Number Joinstate
3 2 08.00.20.ae.33.ef FUJI fuji1 3 6
3 2 08.00.20.bd.5e.a1 FUJI fuji2 2 6
3 3 08.00.20.bd.60.ff FUJI fuji3 1 6
```

fuji2# cftool -e

```
Localdev Srcdev Address Cluster Node Number Joinstate
3 2 08.00.20.ae.33.ef FUJI fuji1 3 6
3 2 08.00.20.bd.5e.a1 FUJI fuji2 2 6
3 3 08.00.20.bd.60.ff FUJI fuji3 1 6
3 3 08.00.20.bd.60.e4 FUJI fuji4 1 6
```

fuji2# cftool -e

```
Localdev Srcdev Address Cluster Node Number Joinstate
3 2 08.00.20.ae.33.ef FUJI fuji1 3 6
3 2 08.00.20.bd.5e.a1 FUJI fuji2 2 6
3 3 08.00.20.bd.60.ff FUJI fuji3 1 6
3 3 08.00.20.bd.60.e4 FUJI fuji4 1 6
```

fuji2# cftool -e

```
Localdev Srcdev Address Cluster Node Number Joinstate
3 2 08.00.20.ae.33.ef FUJI fuji1 3 6
3 2 08.00.20.bd.5e.a1 FUJI fuji2 2 6
3 3 08.00.20.bd.60.ff FUJI fuji3 1 6
3 3 08.00.20.bd.60.e4 FUJI fuji4 1 6
```

Notice that the node `fujii4` does not show up in each of the echo requests. This indicates that the connection to the node `fujii4` is having errors. Because only this node is exhibiting the symptoms, we focus on that node. First, we need to examine the node to see if the Ethernet utilities on that node show any errors. If we log on to `fujii4` and look at the network devices, we see the following:

Number	Device	Type	Speed	Mtu	State	Configured	Address
1	/dev/hme0	4	100	1432	UP	NO	00.80.17.28.2c.fb
2	/dev/hme1	4	100	1432	UP	NO	00.80.17.28.2d.b8
3	/dev/hme2	4	100	1432	UP	YES	08.00.20.bd.60.e4

The `netstat(1M)` utility in Solaris reports information about the network interfaces. The first attempt will show the following:

```
fujii4# netstat -i
```

Name	Mtu	Net/Dest	Address	Ipkts	Ierrs	Opkts	Oerrs	Collis	Queue
lo0	8232	loopback	localhost	65	0	65	0	0	0
hme0	1500	fujii4	fujii4	764055	8	9175	0	0	0
hme1	1500	fujii4-priv	fujii4-priv	2279991	0	2156309	0	7318	0

Notice that the `hme2` interface is not shown in this report. This is because Solaris does not report on interconnects that are not configured for TCP/IP. To temporarily make Solaris report on the `hme2` interface, enter the `ifconfig plumb` command as follows:

```
fujii4# ifconfig hme2 plumb
```

```
fujii4# netstat -i
```

Name	Mtu	Net/Dest	Address	Ipkts	Ierrs	Opkts	Oerrs	Collis	Queue
lo0	8232	loopback	localhost	65	0	65	0	0	0
hme0	1500	fujii4	fujii4	765105	8	9380	0	0	0
hme1	1500	fujii4-priv	fujii4-priv	2282613	0	2158931	0	7319	0
hme2	1500	default	0.0.0.0	752	100	417	0	0	0

Here we can see that the `hme2` interface has 100 input errors (`Ierrs`) from 752 input packet (`Ipkts`). This means that one in seven packets had an error; this rate is too high for PRIMECLUSTER to use successfully. This also explains why `fujii4` sometimes responded to the echo request from `fujii2` and sometimes did not.



It is always safe to “plumb” the interconnect. This will not interfere with the operation of PRIMECLUSTER.

To resolve these errors further, we can look at the undocumented `-k` option to the Solaris `netstat` command as follows:

```
fujii4# netstat -k hme2
```

```
hme2:
ipackets 245295 ierrors 2183 opackets 250486 oerrors 0 collisions 0
defer 0 framing 830 crc 1353 sqe 0 code_violations 38 len_errors 0
ifspeed 100 buff 0 oflo 0 uflo 0 missed 0 tx_late_collisions 0
retry_error 0 first_collisions 0 nocarrier 0 inits 15 nocanput 0
allocbfail 0 runt 0 jabber 0 babble 0 tmd_error 0 tx_late_error 0
rx_late_error 0 slv_parity_error 0 tx_parity_error 0 rx_parity_error 0
slv_error_ack 0 tx_error_ack 0 rx_error_ack 0 tx_tag_error 0
rx_tag_error 0 eop_error 0 no_tmbs 0 no_tbufs 0 no_rbufs 0
rx_late_collisions 0 rbytes 22563388 obytes 22729418 multircv 0 multixmt 0
brdcstrcv 472 brdcstxmt 36 norcvbuf 0 noxmtbuf 0 phy_failures 0
```

Most of this information is only useful to specialists for problem resolution. The two statistics here that are of interest are the `crc` and `framing` errors. These two error types add up to exactly the number reported in `ierrors`. Further resolution of this problem consists of trying each of the following steps:

- Ensure the Ethernet cable is securely inserted at each end.
- Try repeated `cftool -e` and look at the `netstat -i`. If the results of the `cftool` are always the same and the input errors are gone or greatly reduced, the problem is solved.
- Replace the Ethernet cable.
- Try a different port in the Ethernet hub or switch or replace the hub or switch, or temporarily use a cross-connect cable.
- Replace the Ethernet adapter in the node.

If none of these steps resolves the problem, then your support personnel will have to further diagnose the problem.

Problem:

The following console message appears on node `fujii2` while node `fujii3` is trying to join the cluster with node `fujii2`:

```
Mar 10 09:47:55 fujii2 unix: LOG3.0952710475 1080024 1014 4
0 1.0 cf:ens CF: Local node is missing a route from
node: fujii3
Mar 10 09:47:55 fujii2 unix: LOG3.0952710475 1080024 1014 4
0 1.0 cf:ens CF: missing route on local device: /
dev/hme3
Mar 10 09:47:55 fujii2 unix: LOG3.0952710475 1080024 1014 4
0 1.0 cf:ens CF: Node fujii3 Joined Cluster FUJI. (0
1 0)
```

Diagnosis:

Look in /var/adm/messages on node fuji2.

Same message as on console.

No console messages on node fuji3.

Look in /var/adm/messages on node fuji3:

fuji2# cftool -d

Number	Device	Type	Speed	Mtu	State	Configured	Address
1	/dev/hme0	4	100	1432	UP	NO	08.00.06.0d.9f.c5
2	/dev/hme1	4	100	1432	UP	YES	00.a0.c9.f0.15.c3
3	/dev/hme2	4	100	1432	UP	YES	00.a0.c9.f0.14.fe
4	/dev/hme3	4	100	1432	UP	NO	00.a0.c9.f0.14.fd

fuji3# cftool -d

Number	Device	Type	Speed	Mtu	State	Configured	Address
1	/dev/hme0	4	100	1432	UP	NO	08.00.06.0d.9f.c5
2	/dev/hme1	4	100	1432	UP	YES	00.a0.c9.f0.15.c3
3	/dev/hme2	4	100	1432	UP	YES	00.a0.c9.f0.14.fe
4	/dev/hme3	4	100	1432	UP	YES	00.a0.c9.f0.14.fd

/dev/hme3 is not configured on node fuji2

Mar 10 11:00:28 fuji2 unix:WARNING:hme3:no MII link detected

Mar 10 11:00:31 fuji2 unix:LOG3.0952714831 1080024 1008 4 0

1.0cf:ens

CF:Icf Error:(service err_type route_src route_dst).(0 0 0 0 2

0 0 0 3 0 0 0

3 0 0 0)

Mar 10 11:00:53 fuji2 unix:NOTICE:hme3:100 Mbps full-duplex link up

Mar 10 11:01:11 fuji2 unix:LOG3.0952714871 1080024 1007 5 0

1.0cf:ens

CF (TRACE):Icf:Route UP:node src dest.(0 2 0 0 0 3 0 0 0 3 0 0

0)

The hme3 device or interconnect temporarily failed.

fuji2# cftool -n

Node	Number	State	Os	Cpu
fuji2	1	LEFTCLUSTER	Solaris	Sparc
fuji3	2	UP	Solaris	Sparc

Problem:

/dev/hme3 is not configured on node fuji2.


```
Mar 10 11:00:28 fuji2 unix: WARNING: hme3: no MII link detected
Mar 10 11:00:53 fuji2 unix: NOTICE: hme3: 100 Mbps full-duplex
link up
```

Diagnosis:

Look in /var/adm/messages **on node** fuji2:

```
Mar 10 11:00:28 fuji2 unix: WARNING: hme3: no MII link detected
Mar 10 11:00:31 fuji2 unix: LOG3.0952714831 1080024 1008 4
0 1.0cf:ens CF: Icf Error: (service err_type
route_src route_dst). (0 0 0 0 0 2 0 0 0 3 0 0 0 3 0 0 0)
Mar 10 11:00:53 fuji2 unix: NOTICE: hme3: 100 Mbps full-duplex
link up
Mar 10 11:01:11 fuji2 unix: LOG3.0952714871 1080024 1007 5
0 1.0cf:ens CF (TRACE): Icf: Route UP: node src
dest. (0 2 0 0 0 3 0 0 0 3 0 0 0)
```

Problem:

The hme3 device or interconnect temporarily failed. It could be the NIC on either of the cluster nodes or a cable or hub problem.

Node in LEFTCLUSTER state

IF SF is not configured, and node fuji2 panicked and has rebooted. The following console message appears on node fuji2:

```
Mar 10 11:23:41 fuji2 unix: LOG3.0952716221 1080024 1012 4
0 1.0
cf:ens CF: fuji2: busy: local node not down: retrying.
```

Diagnosis:

Look in /var/adm/messages **on node** fuji2:

```
Mar 10 11:23:41 fuji2 unix: LOG3.0952716221 1080024 1007 5
0 1.0 cf:ens CF (TRACE): JoinServer: Startup.
Mar 10 11:23:41 fuji2 unix: LOG3.0952716221 1080024 1009 5
0 1.0 cf:ens CF: Giving UP Mastering (Cluster
already Running).
Mar 10 11:23:41 fuji2 unix: LOG3.0952716221 1080024 1012 4
0 1.0 cf:ens CF: Join postponed, server fuji3 is
busy.
```

... last message repeats.

No new messages on console or in /var/adm/messages **on** fuji2:

```
fuji2: cftool -n
```

Node	Number	State	Os	Cpu
fuji2	1	LEFTCLUSTER	Solaris	Sparc
fuji3	2	UP	Solaris	Sparc

Identified problem:

Node fuji2 has left the cluster and has not been declared DOWN.

Fix:

To fix this problem, enter the following command:

```
# cftool -k
```

This option will declare a node down. Declaring an operational node down can result in catastrophic consequences, including loss of data in the worst case. If you do not wish to declare a node down, quit this program now.

```
Enter node number: 1
Enter name for node #1: fuji2
cftool(down): declaring node #1 (fuji2) down
cftool(down): node fuji2 is down
```

The following console messages then appear on node fuji2:

```
Mar 10 11:34:21 fuji2 unix: LOG3.0952716861 1080024 1005 5
0 1.0
cf:ens CF: MYCLUSTER: fuji2 is Down. (0 1 0)
Mar 10 11:34:29 fuji2 unix: LOG3.0952716869 1080024 1004 5
0 1.0
cf:ens CF: Node fuji2 Joined Cluster MYCLUSTER. (0 1 0)
```

The following console message appears on node fuji2:

```
Mar 10 11:32:37 fuji2 unix: LOG3.0952716757 1080024 1004 5
0 1.0
cf:ens CF: Node fuji2 Joined Cluster MYCLUSTER. (0 1 0)
```

11.3 Collecting troubleshooting information

If a failure occurs in the PRIMECLUSTER system, collect the following information required for investigations from all cluster nodes. Then, contact your local customer support.

1. Obtain the following PRIMECLUSTER investigation information:

- Use `fjsnap` to collect information required for error investigations.
- Retrieve the system dump.
- Collect the Java Console on the clients.




Refer to the Java console documentation in the *Web-Based Admin View Operation Guide*.

- Collect screen shots on the clients.

Refer to the screen hard copy documentation in the *Web-Based Admin View Operation Guide*.

2. In case of application failures, collect such investigation material.

3. If the problem is reproducible, then include a description on how it can be reproduced.

-  It is essential that you collect the debugging information described in this section. Without this information, it may not be possible for customer support to debug and fix your problem.
-  Be sure to gather debugging information from all nodes in the cluster. It is very important to get this information (especially the `fjsnap` data) as soon as possible after the problem occurs. If too much time passes, then essential debugging information may be lost.
-  If a node is panicked, execute `sync` in OBP mode and take a system dump.

11.3.1 Executing the `fjsnap` command

The `fjsnap` command is a Solaris system information tool provided with the Enhanced Support Facility `FJSVsnap` package. In the event of a failure in the PRIMECLUSTER system, the necessary error information can be collected to pinpoint the cause.

Execute the `fjsnap` command as follows:

1. Log in as root.
2. Execute one of the following `fjsnap` commands:

```
# /opt/FJSVsnap/bin/fjsnap -h output
```

```
# /opt/FJSVsnap/bin/fjsnap -a output
```

- As `-a` collects all detailed information, the data is very large. When `-h` is specified, only cluster control information is collected.
- In *output*, specify the special file name or output file name (for example, `/dev/rmt/0`) of the output medium to which the error information collected with the `fjsnap` command is written.

For details about the `fjsnap` command, see the `README` file included in the `FJSVsnap` package.



When to run `fjsnap`:

- If an error message appears during normal operation, execute `fjsnap` immediately to collect investigation material.
- If the necessary investigation material cannot be collected because of a hang, shut down the system, and start the system in single mode. Execute the `fjsnap` command to collect information.
- If the system has rebooted automatically to multi-user mode, then execute the `fjsnap` command to collect information.

11.3.2 System dump

If the system dump is collected while the node is in panicked, retrieve the system dump as investigation material. The system dump is saved as a file during the node's startup process. The default destination directory is `/var/crash/node_name`.

11.3.3 SCF dump

You need to collect the System Control Facility (SCF) dump if one of the following messages is output:

- 7003 An error was detected in RCI.
 (node:*nodename* address:*address* status:*status*)
- 7004 The RCI monitoring agent has been stopped due to an
 RCI address error.
 (node:*nodename* address:*address*)

A message from the SCF driver



Refer to the *Enhanced Support Facility User's Guide* for details on SCF driver messages.

The RAS monitoring daemon, which is notified of a failure from SCF, stores SCF dump in the `/var/opt/FJSVhwr/scf.dump` file. You need to collect SCF dump by executing the following commands:

```
# cd /var/opt
```

```
# tar cf /tmp/scf.dump.tar ./FJSVhwr
```

12 CF messages and codes

This chapter is a printed version of information that can be found on the PRIMECLUSTER CD.

This chapter discusses the following:

- The Section “cfconfig messages” discusses the `cfconfig` command and its error messages.
- The Section “cipconfig messages” describes the `cipconfig` command and its messages.
- The Section “cftool messages” details the `cftool` command and its messages.
- The Section “rcqconfig messages” discusses the `rcqconfig` command and its messages.
- The Section “rcquery messages” describes the `rcquery` command and its messages.
- The Section “CF runtime messages” discusses CF runtime messages.
- The Section “CF Reason Code table” lists CF reason codes.
- The Section “Error messages for different systems” provides a pointer for accessing error messages for different systems.
- The Section “Solaris/Linux ERRNO table” lists error messages for Solaris and Linux by number.
- The Section “Resource Database messages” explains the Resource Database messages.
- The Section “Shutdown Facility” lists messages, causes, and actions.
- The Section “Monitoring Agent messages” details the MA messages.

The following lexicographic conventions are used in this chapter:

- Messages that will be generated on `stdout` or `stderr` are shown on the first line(s).
- Explanatory text is given after the message.
- Messages that will be generated in the `system-log` file and may optionally appear on the console are listed after the explanation.

- Message text tokens shown in a italic font style are placeholders for substituted text.
- Many messages include a token of the form #0407, which always denotes a hexadecimal reason code. Section “CF Reason Code table” has a complete list of these codes.

12.1 cfconfig messages

The `cfconfig` command will generate an error message on `stderr` if an error occurs. Additional messages giving more detailed information about this error may be generated by the support routines in the `libcf` library. However, these additional messages will only be written to the system log file, and will not appear on `stdout` or `stderr`.

Refer to the `cfconfig` manual page for an explanation of the command options and the associated functionality. The `cfconfig` manual page also describes the format of all non-error related command output.

12.1.1 Usage message

A usage message will be generated if:

- Multiple `cfconfig` options are specified (all options are mutually exclusive).
- An invalid `cfconfig` option is specified.
- No `cfconfig` option is specified.
- The `-h` option is specified.

Usage:

```
cfconfig [-d|-G|-g|-h|-L|-l|-S nodename clustername device [device [...]] |-s clustername device [device [...]]|-u]
  -d delete configuration
  -g get configuration
  -G get configuration including address information
  -h help
  -L fast load (use configured devicelist)
  -l load
  -S set configuration (including nodename)
  -s set configuration
  -u unload
```


A device can either be a network device or an IP device like `/dev/ip[0-3]` followed by the IP-Address and Broadcast Address number.

12.1.2 Error messages

cfconfig -l

```
cfconfig: cannot load: #0423: generic: permission denied
```

The CF startup routine has failed. This error message usually indicates that an unprivileged user has attempted to start CF. You must have administrative privileges to start, stop, and configure CF. An additional error message, for this case, will also be generated in the system-log file:

```
OSDU_start: failed to open /dev/cf (EACCES)
```

```
cfconfig: cannot load: #041f: generic: no such file or directory
```

```
cfconfig: check that configuration has been specified
```

The CF startup routine has failed. This error message usually indicates that the CF configuration file `/etc/default/cluster` cannot be found. Additional error messages, for this case, may also be generated in the system-log file:

```
OSDU_getconfig: failed to open config file (errno)
```

```
OSDU_getconfig: failed to stat config file (errno)
```

```
cfconfig: cannot load: #0405: generic: no such device/resource  
cfconfig: check if configuration entries match node's device list
```

The CF startup routine has failed. This error message usually indicates that the CF configuration file does not match the physical hardware (network interfaces) installed in/on the node.

```
cfconfig: cannot load: #04xx: generic: reason_text
```

The CF startup routine has failed. One cause of an error message of this pattern is that the CF cluster configuration file has been damaged or is missing. If you think this is the case, delete and then re-specify your cluster configuration information, and try the command again. If the same error persists, see below.

Additional error messages, for this case, will also be generated in the system-log file:

```
OSDU_getconfig: corrupted config file
```

```
OSDU_getconfig: failed to open config file (errno)
```

```
OSDU_getconfig: failed to stat config file (errno)
OSDU_getconfig: read failed (errno)
```

Another cause of an error message of this pattern is that the CF driver and/or other kernel components may have somehow been damaged. Remove and then re-install the CF package. If this does not resolve the problem, contact your customer support representative. Additional error messages, for this case, will also be generated in the system-log file:

```
OSDU_getconfig: malloc failed
OSDU_getstatus: mconn status ioctl failed (errno)
OSDU_nodename: malloc failed
OSDU_nodename: uname failed (errno)
OSDU_start: failed to get configuration
OSDU_start: failed to get nodename
OSDU_start: failed to kick off join
OSDU_start: failed to open /dev/cf (errno)
OSDU_start: failed to open /dev/mconn (errno)
OSDU_start: failed to select devices
OSDU_start: failed to set clustername
OSDU_start: failed to set nodename
OSDU_start: icf_devices_init failed
OSDU_start: icf_devices_setup failed
OSDU_start: IOC_SOSD_DEVSELECTED ioctl failed
OSDU_start: netinit failed
```

If the device driver for any of the network interfaces to be used by CF responds in an unexpected way to DLPI messages, additional message output (in the system-log) may occur, with no associated command error message. These messages may be considered as warnings, unless a desired network interface cannot be configured as a cluster interconnect. These messages are:

```
d1_attach: DL_ACCESS error
d1_attach: DL_ATTACH_REQ putmsg failed (errno)
d1_attach: DL_BADPPA error
d1_attach: DL_OUTSTATE error
d1_attach: DL_SYSERR error
d1_attach: getmsg for DL_ATTACH response failed (errno)
d1_attach: unknown error
d1_attach: unknown error hexvalue
d1_bind: DL_ACCESS error
d1_bind: DL_BADADDR error
d1_bind: DL_BIND_REQ putmsg failed (errno)
d1_bind: DL_BOUND error
d1_bind: DL_INITFAILED error
d1_bind: DL_NOADDR error
d1_bind: DL_NOAUTO error
d1_bind: DL_NOTESTAUTO error
```

```

dl_bind: DL_NOTINIT error
dl_bind: DL_NOXIDAUTO error
dl_bind: DL_OUTSTATE error
dl_bind: DL_SYSERR error
dl_bind: DL_UNSUPPORTED error
dl_bind: getmsg for DL_BIND response failed (errno)
dl_bind: unknown error
dl_bind: unknown error hexvalue
dl_info: DL_INFO_REQ putmsg failed (errno)
dl_info: getmsg for DL_INFO_ACK failed (errno)

```

It is also possible that while CF is examining the kernel device tree, looking for eligible network interfaces, that a device or streams responds in an unexpected way. This may trigger additional message output in the system-log, with no associated command error message. These messages may be considered as warnings, unless a desired network interface cannot be configured as a cluster interconnect. These messages are:

```

get_net_dev: cannot determine driver name of nodename device
get_net_dev: cannot determine instance number of nodename
device
get_net_dev: device table overflow - ignoring /dev/drivernameN
get_net_dev: dl_attach failed: /dev/drivernameN
get_net_dev: dl_bind failed: /dev/drivernameN
get_net_dev: dl_info failed: /dev/drivername
get_net_dev: failed to open device: /dev/drivername (errno)
get_net_dev: not an ethernet device: /dev/drivername
get_net_dev: not DL_STYLE2 device: /dev/drivername
icf_devices_init: cannot determine instance number of drivername
device
icf_devices_init: device table overflow - ignoring /dev/scin
icf_devices_init: di_init failed
icf_devices_init: di_prom_init failed
icf_devices_init: dl_bind failed: /dev/scin
icf_devices_init: failed to open device: /dev/scin (errno)
icf_devices_init: no devices found
icf_devices_select: devname device not found
icf_devices_select: fstat of mclx device failed: /devices/
pseudo/icfn - devname (errno)
icf_devices_select: mcl_select_dev failed: /devices/pseudo/
icfn - devname (errno)
icf_devices_select: open of mclx device failed: /devices/
pseudo/icfn - devname (errno)
icf_devices_setup: calloc failed: devname
icf_devices_setup: failed to create mclx dev: /devices/pseudo/
icfn - devname (errno)
icf_devices_setup: failed to open /dev/kstat (errno)

```

```

icf_devices_setup: failed to open mclx device: /devices/
pseudo/icfn - devname (errno)
icf_devices_setup: failed to stat mclx device: /dev/mclx
(errno)
icf_devices_setup: failed to stat mclx device: /devices/
pseudo/icfn - devname (errno)
icf_devices_setup: I_LIST failed: devname
(errno)icf_devices_setup: I_LIST 0 failed: devname (errno)
icf_devices_setup: I_PLINK failed: /devices/pseudo/icfn -
devname (errno)
icf_devices_setup: I_POP failed: devname (errno)
icf_devices_setup: I_PUSH failed: devname (errno)
icf_devices_setup: mcl_set_device_id failed: /devices/pseudo/
icfn - devname (errno)
icf_devices_setup: mclx_get_device_info failed: /devices/
pseudo/icfn - devname (errno)
icf_devices_setup: mclx device already linked: /devices/
pseudo/icfn - devname (errno)
icf_devices_setup: mclx not a device
mcl_select_device: MC1_IOC_SEL_DEV ioctl failed (errno)
mcl_set_device_id: MC1_IOC_SET_ID ioctl failed (errno)
mclx_get_device_info: MC1X_IOC_GET_INFO ioctl failed (errno)

```

cfconfig -u

```

cfconfig: cannot unload: #0406: generic: resource is busy
cfconfig: check if dependent service-layer module(s) active

```

The CF shutdown routine has failed. This error message is generated if a PRIMECLUSTER Layered Service still has a CF resource active/allocated. RMS, SIS, OPS, CIP, etc., need to be stopped before CF can be unloaded. An additional error message, for this case, will also be generated in the system-log file:

```

OSDU_stop: failed to unload cf_drv
cfconfig: cannot unload: #0423: generic: permission denied

```

The CF shutdown routine has failed. This error message usually indicates that an unprivileged user has attempted to stop CF. You must have administrative privileges to start, stop, and configure CF. An additional error message, for this case, will also be generated in the system-log file:

```

OSDU_stop: failed to open /dev/cf (EACCES)
cfconfig: cannot unload: #04xx: generic: reason_text

```

The cause of an error message of this pattern is that the CF driver and/or other kernel components may have somehow been damaged. Remove and then re-install the CF package. If this does not resolve the problem, contact your customer support representative. Additional error messages, for this case, will also be generated in the system-log file:

```
mclx_get_device_info: MCLX_IOC_GET_INFO ioctl failed (errno)
OSDU_stop: disable unload failed
OSDU_stop: enable unload failed
OSDU_stop: failed to open /dev/cf (errno)
OSDU_stop: failed to open mclx device: /devices/pseudo/icfn
(errno)
OSDU_stop: failed to unlink mclx device: /devices/pseudo/icfn
(errno)
OSDU_stop: failed to unload cf_drv
OSDU_stop: failed to unload mcl module
OSDU_stop: failed to unload mclx driver
OSDU_stop: mclx_get_device_info failed: /devices/pseudo/icfn
```

cfconfig -s

cfconfig -S

```
cfconfig: specified nodename: bad length: #407: generic: invalid
parameter
```

This usually indicates that nodename is too long. The maximum length is 31 characters.

```
cfconfig: invalid nodename: #407: generic: invalid parameter
```

This indicates that nodename contains one or more non-printable characters.

```
cfconfig: node already configured: #0406: generic: resource is
busy
```

This error message usually indicates that there is an existing CF configuration. To change the configuration of a node, you must first delete (`cfconfig -d`) any pre-existing configuration. Also, you must have administrative privileges to start, stop, and configure CF. A rare cause of this error would be that the CF driver and/or other kernel components have somehow been damaged. If you believe this is the case, remove and then re-install the CF package. If this does not resolve the problem, contact your customer support representative. Additional error messages may also be generated in the system-log file:

```
OSDU_getconfig: corrupted config file
OSDU_getconfig: failed to open config file (errno)
OSDU_getconfig: failed to stat config file (errno)
```

```
OSDU_getconfig: malloc failed
OSDU_getconfig: read failed (errno)
cfconfig: too many devices specified: #0407: generic: invalid
parameter
```

Too many devices have been specified on the command line. The current limit is set to 255.

```
cfconfig: clustername cannot be a device: #0407: generic:
invalid parameter
```

This error message indicates that “clustername,” is a CF-eligible device. This usually means that the clustername has accidentally been omitted.

```
cfconfig: invalid clustername: #0407: generic: invalid parameter
```

This error message indicates that clustername is a CF-eligible device.

```
cfconfig: duplicate device names specified: #0407: generic:
invalid parameter
```

This error message indicates that duplicate device names have been specified on the command line. This is usually a typographical error, and it is not permitted to submit a device name more than once.

```
cfconfig: device [device [...]]: #0405: generic: no such device/
resource
```

This error message indicates that the specified device names are not CF-eligible devices. Only those devices displayed by `cftool -d` are CF-eligible devices.

```
cfconfig: cannot open mconn: #04xx: generic: reason_text
```

This message should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

```
cfconfig: cannot set configuration:
#04xx: generic: reason_text
```

This message can occur if concurrent `cfconfig -s` or `cfconfig -S` commands are being run. Otherwise, it should not occur unless the CF driver and/or other kernel components have somehow been damaged. If this is the case, remove and then re-install the CF package. If the problem persists, contact your customer support representative. Additional error messages may also be generated in the system-log file:

```
OSDU_setconfig: config file exists
OSDU_setconfig: failed to create config file (errno)
OSDU_setconfig: write failed (errno)
```

```
cfconfig: cannot get new configuration: #04xx: generic:
reason_text
```

This message indicates that the saved configuration cannot be read back. This may occur if concurrent `cfconfig -s` or `cfconfig -S` commands are being run, or if disk hardware errors are reported. Otherwise, it should not occur unless the CF driver and/or other kernel components have somehow been damaged. If this is the case, remove and then re-install the CF package. If the problem persists, contact your customer support representative. Additional error messages may also be generated in the system-log file:

```
OSDU_getconfig: corrupted config file
OSDU_getconfig: failed to open config file (errno)
OSDU_getconfig: failed to stat config file (errno)
OSDU_getconfig: malloc failed
OSDU_getconfig: read failed (errno)
cfconfig: cannot load: #04
xx: generic: reason_text
```

This error message indicates that the device discovery portion of the CF startup routine has failed. (See error messages associated with `cfconfig -l` above).

cfconfig -g

```
cfconfig: cannot get configuration: #04xx: generic: reason_text
```

This message indicates that the CF configuration cannot be read. This may occur if concurrent `cfconfig` commands are being run, or if disk hardware errors are reported. Otherwise, it should not occur unless the CF driver and/or other kernel components have somehow been damaged. If this is the case, remove and then re-install the CF package. If the problem persists, contact your customer support representative.

Additional error messages may also be generated in the system-log file:

```
OSDU_getconfig: corrupted config file
OSDU_getconfig: failed to open config file (errno)
OSDU_getconfig: failed to stat config file (errno)
OSDU_getconfig: malloc failed
OSDU_getconfig: read failed (errno)
```

cfconfig -d

```
cfconfig: cannot get joinstate: #0407: generic: invalid
parameter
```

This error message usually indicates that the CF driver and/or other kernel components have somehow been damaged. remove and then re-install the CF package. If this does not resolve the problem, contact your customer support representative.

cfconfig: cannot delete configuration: #0406: generic:
resource is busy

This error message is generated if CF is still active (i.e., if CF resource(s) are active/allocated). The configuration node may not be deleted while it is an active cluster member.

cfconfig: cannot delete configuration: #04xx: generic:
reason_text

You must have administrative privileges to start, stop, and configure CF. A rare cause of this error would be that the CF driver and/or other kernel components have somehow been damaged. If you believe this is the case, remove and then re-install the CF package. If this does not resolve the problem, contact your customer support representative. An additional error message will also be generated in the system-log file:

OSDU_delconfig: failed to delete config file (*errno*)

12.2 cipconfig messages

The `cipconfig` command will generate an error message on `stderr` if an error occurs. Additional error messages giving more detailed information about the error may be generated by the support routines of the `libcf` library. However, these additional messages will only be written to the system-log file, and will not appear on `stdout` or `stderr`.

Refer to the `cipconfig` manual page for an explanation of the command options and associated functionality. The `cipconfig` manual page also describes the format of all non-error related command output.

12.2.1 Usage message

A usage message will be generated if:

- Multiple `cipconfig` options are specified (all options are mutually exclusive).
- An invalid `cipconfig` option is specified.
- No `cipconfig` option is specified.
- The `-h` option is specified.


```
usage: cipconfig [-l|-u|-h]
        -l start/load
        -u stop/unload
        -h help
```

12.2.2 Error messages

cipconfig -l

cipconfig: could not start CIP – detected a problem with CF.
 cipconfig: cannot open mconn: #04xx: generic: *reason_text*

These messages should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

cipconfig: cannot setup cip: #04xx: generic: *reason_text*

The cip startup routine(s) have failed. There may be problems with the configuration file. Additional error messages will be generated in the system-log file:

```
OSDU_cip_start: cip kickoff failed (errno)
OSDU_cip_start: dl_attach failed: devpathn
OSDU_cip_start: dl_bind failed: devpathn
OSDU_cip_start: dl_info failed: devpath
OSDU_cip_start: failed to open device: /dev/cip (errno)
OSDU_cip_start: failed to open device: devpath (errno)
OSDU_cip_start: I_PLINK failed: devpath (errno)
OSDU_cip_start: POPing module failed: errno
OSDU_cip_start: ppa n is not valid: devpath
OSDU_cip_start: setup controller/speed failed: devpath
(errno)
```

If the device driver for any of the network interfaces used by CIP responds in an unexpected way to DLPI messages, additional message output may occur:

```
dl_info: DL_INFO_REQ putmsg failed (errno)
dl_info: getmsg for DL_INFO_ACK failed (errno)
dl_attach: DL_ACCESS error
dl_attach: DL_ATTACH_REQ putmsg failed (errno)
dl_attach: DL_BADPPA error
dl_attach: DL_OUTSTATE error
dl_attach: DL_SYSERR error
dl_attach: getmsg for DL_ATTACH response failed (errno)
dl_attach: unknown error
dl_attach: unknown error hexvalue
```

```
dl_bind: DL_ACCESS error
dl_bind: DL_BADADDR error
dl_bind: DL_BIND_REQ putmsg failed (errno)
dl_bind: DL_BOUND error
dl_bind: DL_INITFAILED error
dl_bind: DL_NOADDR error
dl_bind: DL_NOAUTO error
dl_bind: DL_NOTESTAUTO error
dl_bind: DL_NOTINIT error
dl_bind: DL_NOXIDAUTO error
dl_bind: DL_OUTSTATE error
dl_bind: DL_SYSERR error
dl_bind: DL_UNSUPPORTED error
dl_bind: getmsg for DL_BIND response failed (errno)
dl_bind: unknown error
dl_bind: unknown error hexvalue
```

If these messages appear and they do not seem to be associated with problems in your CIP configuration file, contact your customer support representative.

cipconfig -u

```
cipconfig: cannot unload cip: #04xx: generic: reason_text
The CIP shutdown routine has failed. Usually this mean that another
PRIMECLUSTER Layered Service has a CIP interface open (active). It
must be stopped first. Additional error messages may be generated in
the system-log file:
OSDU_cip_stop: failed to unload cip driver
OSDU_cip_stop: failed to open device: /dev/cip (errno)
```

12.3 cftool messages

The `cftool` command will generate an error message on `stderr` if an error condition is detected. Additional messages, giving more detailed information about this error, may be generated by the support routines of the `libcf` library. Note that these additional error messages will only be written to the system-log file, and will not appear on `stdout` or `stderr`.

Refer to the `cftool` manual page for an explanation of the command options and the associated functionality. The `cftool` manual page also describes the format of all non-error related command output.

12.3.1 Usage message

A usage message will be generated if:

- Conflicting `cftool` options are specified (some options are mutually exclusive).
- An invalid `cftool` option is specified.
- No `cftool` option is specified.
- The `-h` option is specified.

```
usage: cftool [-c][-l][-n][-r][-d][-v][-p][-e][-i nodename][-A
cluster][-T timeout][-F][-C count][-I nodename][-E xx.xx.xx.xx.xx.xx][-
P][-m][-u][-k][-q][-h]
```

```
-c                clustername
-l                local nodeinfo
-n                nodeinfo
-r                routes
-d                devinfo
-v                version
-p                ping
-e                echo
-i                icf stats for nodename
-m                mac stats
-u                clear all stats
-k                set node status to down
-q                quiet mode
-h                help
-F                flush ping queue. Be careful, please
-T timeout      millisecond ping timeout
-I                raw ping test by node name
-P                raw ping
-A cluster      ping all interfaces in one cluster
-E xx.xx.xx.xx.xx.xx  raw ping by 48-bit physical address
-C count       stop after sending count raw ping messages
```

A device can either be a network device or an IP device like `/dev/ip[0-3]` followed by IP address and broadcast address.

12.3.2 Error messages

cftool: CF not yet initialized

cftool -c

cftool: failed to get cluster name: *#xxxx*: service: *reason_text*

This message should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

cftool -d

cftool: cannot open mconn: *#04xx*: generic: *reason_text*

This message should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

cftool -e

cftool: cannot open mconn: *#04xx*: generic: *reason_text*

This message should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

cftool -i nodename

cftool: *nodename*: No such node

cftool: cannot get node details: *#xxxx*: service: *reason_text*

Either of these messages indicates that the specified nodename is not an active cluster node at this time.

cftool: cannot open mconn: *#04xx*: generic: *reason_text*

This message should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

cftool -k

cftool(down): illegal node number

This message indicates that the specified node number is non-numeric or is out of allowable range (1–64).

```
cftool(down): not executing on active cluster node
```

This message is generated if the command is executed either on a node that is not an active cluster node or on the specified LEFTCLUSTER node itself.

```
cftool(down): cannot declare node down: #0426: generic:
invalid node name
```

```
cftool(down): cannot declare node down: #0427: generic:
invalid node number
```

```
cftool(down): cannot declare node down: #0428: generic: node
is not in LEFTCLUSTER state
```

One of these messages will be generated if the supplied information does not match an existing cluster node in LEFTCLUSTER state.

```
cftool(down): cannot declare node down: #xxx: service:
reason_text
```

Other variations of this message should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

cftool -l

```
cftool: cannot get nodename: #04xx: generic: reason_text
```

```
cftool: cannot get the state of the local node: #04xx: generic:
reason_text
```

These messages should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

cftool -m

```
cftool: cannot open mconn: #04xx: generic: reason_text
```

```
cftool: cannot get icf mac statistics: #04xx: generic: reason_text
```

These messages should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

cftool -n

```
cftool: cannot get node id: #xxx: service: reason_text  
cftool: cannot get node details: #xxx: service: reason_text
```

This message should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

cftool -p

```
cftool: cannot open mconn: #04xx: generic: reason_text
```

This message should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

cftool -r

```
cftool: cannot get node details: #xxx: service: reason_text
```

These messages should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

cftool -u

```
cftool: cannot open mconn: #04xx: generic: reason_text  
cftool: clear icf statistics: #04xx: generic: reason_text
```

These messages should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

cftool -v

```
cftool: cannot open mconn: #04xx: generic: reason_text  
cftool: unexpected error retrieving version: #04xx: generic:  
reason_text
```

These messages should not occur unless the CF driver and/or other kernel components are damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

12.4 rcqconfig messages

The `rcqconfig` command will generate an error message on standard error if an error condition is detected. Additional messages, giving more detailed information about this error, may be generated by the support routines of the `libcf` library. Please note that these additional error messages will only be written to the system-log file during `cfconfig -l`, and will not appear on standard out or standard error.

Refer to the `rcqconfig` manual page for an explanation of the command options and the associated functionality.

12.4.1 Usage message

A usage message will be generated if:

- Conflicting `rcqconfig` options are specified (some options are mutually exclusive).
- An invalid `rcqconfig` option is specified.
- The `'-h'` option is specified.

```
usage: rcqconfig [ -g | -h ] or
       rcqconfig -s or
       rcqconfig [ -v ] [ -c ]
           [ -a Add-node-1 ... Add-node-n ]
           [ -x Ignore-node-1 ... Ignore-node-n ]
           [ -d Delete-node-1 ... Delete-node-n ]
           [ -m quorum-method-1 ... quorum-method-n ]
```

12.4.2 Error messages

```
rcqconfig -a node-1 node-2 ... node-n
-g and -a cannot exist together.
```

This error message usually indicates that get configuration option (`-g`) cannot be specified with this option (`-a`). Refer to the manual pages for the correct syntax definition.

```
Nodename is not valid nodename.
```

This error message usually indicates that the length of the node is less than 1 or greater than 31 bytes. Refer to the manual pages for the correct syntax definition.

```
rcqconfig : failed to start
```

The following errors will also be reported in standard error if rcqconfig fail to start.

```
rcqconfig failed to configure qsm since quorum node set is empty.
```

This error message usually indicates that the quorum configuration does not exist. Refer to the manual pages for rcqconfig(1M) for the correct syntax to configure the quorum nodes.

```
cfreg_start_transaction: `#2813: cfreg daemon not present`
```

The rcqconfig routine has failed. This error message usually indicates that the synchronization daemon is not running on the node. The cause of error messages of this pattern may be that the cfreg daemon has died and the previous error messages in the system log or console will indicate why the daemon died. Restart the daemon using cfregd -r. If it fails again, the error messages associated with it will indicate the problem. The data in the registry is most likely corrupted. If the problem persists, contact your customer service support representative.

```
cfreg_start_transaction: `#2815: registry is busy`
```

The rcqconfig routine has failed. This error message usually indicates that the daemon is not in synchronized state or if the transaction has been started by another application. This messages should not occur. The cause of error messages of this pattern is that the registries are not in consistent state. If the problem persists, unload the cluster by using cfconfig -u and reload the cluster by using cfconfig -l. If the problem still persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_start_transaction: `#2810: an active transaction exists`
```

The rcqconfig routine has failed. This error message usually indicates that the application has already started a transaction. If the cluster is stable, the cause of error messages of this pattern is that different changes may be done concurrently from multiple nodes. Therefore, it might take longer time to commit. Retry the command again. If the problem persists, the cluster might not be in a stable state. The error messages in the log will indicate the problem. If this is the case, unload the cluster by using cfconfig -u and reload the cluster by using cfconfig -l. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

Too many nodename are defined for quorum. Max node = 64

This error message usually indicates that if the number of node specified are more than 64 for which the quorum is to be configured. The following errors will also be reported in standard error if there are too many nodename defined:

```
cfreg_get: `#2809: specified transaction invalid`
```

The `rcqconfig` routine has failed. This error message usually indicates that the information supplied to get the specified data from the registry is not valid (e.g. transaction aborted due to time period expiring or synchronization daemon termination, etc.). This messages should not occur. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_get: `#2819: data or key buffer too small`
```

The `rcqconfig` routine has failed. This error message usually indicates that the specified size of the data buffer is too small to hold the entire data for the entry. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

Cannot add node node that is not up.

This error message usually indicates that the user is trying to add a node whose state is not up in the NSM node space. Try to bring up the down node or remove the node from the list which quorum is to be configured.

Cannot proceed. Quorum node set is empty.

This error message usually indicates that if no node is specified to this option or there is no configured node prior to this call. The following errors will also be reported in standard error if quorum node set is empty:

The following errors will also be reported in standard error if `rcqconfig` fails to start:

```
cfreg_put: `#2809: specified transaction invalid`
```

The `rcqconfig` routine has failed. This error message usually indicates that the information supplied to get the specified data from the registry is not valid (e.g. transaction aborted due to time period expiring or synchronization daemon termination, etc.). This messages should not occur. Try to unload the cluster by

using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_put: `#2820: registry entry data too large`
```

The `rcqconfig` routine has failed. This error message usually indicates that the specified size data is larger than 28K. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
rcqconfig -s  
stopping quorum space methods `#0408: unsuccessful`
```

The `rcqconfig` routine has failed. This error message usually indicates that there is no method specified.

```
rcqconfig -x ignore_node=1 ... ignore_node=n  
-g and -x cannot exist together.
```

This error message usually indicates that get configuration option (`-g`) cannot be specified with this option (`-x`). Refer to the manual pages for the correct syntax definition.

```
Nodename is not valid nodename.
```

This error message usually indicates that the length of the node is less than 1 or greater than 31 bytes.

```
rcqconfig : failed to start
```

The following errors will also be reported in standard error if `rcqconfig` fails to start:

```
cfreg_start_transaction: `#2813: cfreg daemon not present`
```

The `rcqconfig` routine has failed. This error message usually indicates that the synchronization daemon is not running on the node. The cause of error messages of this pattern may be that the `cfreg` daemon has died and the previous error messages in the system log or console will indicate why the daemon died. Restart the daemon using `cfregd -r`. If it fails again, the error messages associated with it will indicate the problem. The data in the registry is most likely corrupted. If the problem persists, contact your customer service support representative.

```
cfreg_start_transaction: `#2815: registry is busy`
```

The `rcqconfig` routine has failed. This error message usually indicates that the daemon is not in synchronized state or if the transaction has been started by another application. This messages should not occur. If the problem persists, unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem still persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_start_transaction: `#2810: an active transaction exists`
```

The `rcqconfig` routine has failed. This error message usually indicates that the application has already started a transaction. If the cluster is stable, the cause of error messages of this pattern is that different changes may be done concurrently from multiple nodes. Therefore, it might take longer time to commit. Retry the command again. If the problem persists, the cluster might not be in a stable state. The error messages in the log will indicate the problem. If this is the case, unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
Too many ignore node names are defined for quorum. Max node = 64
```

This error message usually indicates that if the number of ignore nodes specified are more than 64. The following errors will also be reported in standard error if the ignore node names exceed 64.

```
cfreg_get: `#2809: specified transaction invalid`
```

The `rcqconfig` routine has failed. This error message usually indicates that the information supplied to get the specified data from the registry is not valid (e.g. transaction aborted due to time period expiring or synchronization daemon termination, etc.). This messages should not occur. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_get: `#2804: entry with specified key does not exist`
```

The `rcqconfig` routine has failed. This error message usually indicates that the specified entry does not exist. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_get: `#2819: data or key buffer too small`
```

The `rcqconfig` routine has failed. This error message usually indicates that the specified size of the data buffer is too small to hold the entire data for the entry. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
Can not add node node that is not up.
```

This error message usually indicates that the user is trying to add a node whose state is not up in the NSM node space. Try to bring up the down node or remove the node from the list which quorum is to be configured.

```
Can not proceed. Quorum node set is empty.
```

This error message usually indicates that if no node is specified to this option or there is no configured node prior to this call. The following errors will also be reported in standard error if Quorum node set is empty:

```
cfreg_put: `#2809: specified transaction invalid`
```

The `rcqconfig` routine has failed. This error message usually indicates that the information supplied to get the specified data from the registry is not valid (e.g. transaction aborted due to time period expiring or synchronization daemon termination, etc.). This messages should not occur. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_put: `#2820: registry entry data too large`
```

The `rcqconfig` routine has failed. This error message usually indicates that the event information (data being passed to the kernel) to be used for other sub-systems) is larger than 32K. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_put: `#2807: data file format is corrupted`
```

The `rcqconfig` routine has failed. This error message usually indicates that the registry data file format has been corrupted. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using

`cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cms_post_event: `#0c01: event information is too large`
```

The `rcqconfig` routine has failed. This error message usually indicates that the event information (data being passed to the kernel) to be used for other sub-systems) is larger than 32K. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
rcqconfig -m method_name-1 ... method_name -n
```

`-g` and `-m` cannot exist together.

This error message usually indicates that get configuration option (`-g`) cannot be specified with this option (`-x`). Refer to the manual pages for the correct syntax definition.

```
Methodname is not valid method name.
```

This error message usually indicates that the length of the node is less than 1 or greater than 31 bytes. Refer to the manual pages for the correct syntax definition.

```
rcqconfig : failed to start
```

The following errors will also be reported in standard error if `rcqconfig` fail to start:

```
cfreg_start_transaction: `#2813: cfreg daemon not present`
```

The `rcqconfig` routine has failed. This error message usually indicates that the synchronization daemon is not running on the node. The cause of error messages of this pattern may be that the `cfreg` daemon has died and the previous error messages in the system log or console will indicate why the daemon died. Restart the daemon using `cfregd -r`. If it fails again, the error messages associated with it will indicate the problem. The data in the registry is most likely corrupted. If the problem persists, contact your customer service support representative.

```
cfreg_start_transaction: `#2815: registry is busy`
```

The `rcqconfig` routine has failed. This error message usually indicates that the daemon is not in synchronized state or if the transaction has been started by another application. This message should not occur. The cause of error

messages of this pattern is that the registries are not in consistent state. If the problem persists, unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem still persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_start_transaction: `#2810: an active transaction exists`
```

The `rcqconfig` routine has failed. This error message usually indicates that the application has already started a transaction. If the cluster is stable, the cause of error messages of this pattern is that different changes may be done concurrently from multiple nodes. Therefore, it might take longer time to commit. Retry the command again. If the problem persists, the cluster might not be in a stable state. If this is the case, unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
Too many method names are defined for quorum. Max method = 8
```

This error message usually indicates that if the number of methods specified are more than 8. The following errors will also be reported in standard error if Quorum method names exceed the limit:

```
cfreg_get: `#2809: specified transaction invalid`
```

The `rcqconfig` routine has failed. This error message usually indicates that the information supplied to get the specified data from the registry is not valid (e.g. transaction aborted due to time period expiring or synchronization daemon termination, etc.). This message should not occur. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_get: `#2804: entry with specified key does not exist`
```

The `rcqconfig` routine has failed. This error message usually indicates that the specified entry does not exist. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_get: `#2819: data or key buffer too small`
```

The `rcqconfig` routine has failed. This error message usually indicates that the specified size of the data buffer is too small to hold the entire data for the entry. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_put: `#2809: specified transaction invalid`
```

The `rcqconfig` routine has failed. This error message usually indicates that the information supplied to get the specified data from the registry is not valid (e.g. transaction aborted due to time period expiring or synchronization daemon termination, etc.). This messages should not occur. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_put: `#2820: registry entry data too large`
```

The `rcqconfig` routine has failed. This error message usually indicates that the event information (data being passed to the kernel) to be used for other sub-systems) is larger than 32K. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_put: `#2807: data file format is corrupted`
```

The `rcqconfig` routine has failed. This error message usually indicates that the registry data file format has been corrupted. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cms_post_event: `#0c01: event information is too large`
```

The `rcqconfig` routine has failed. This error message usually indicates that the event information (data being passed to the kernel) to be used for other sub-systems) is larger than 32K. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the

cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
rcqconfig -d node-1 node-2 ... node-n
```

`-g` and `-d` cannot exist together.

This error message usually indicates that get configuration option (`-g`) cannot be specified with this option (`-d`). Refer to the manual pages for the correct syntax definition.

```
Nodename is not valid nodename.
```

This error message usually indicates that the length of the node is less than 1 or greater than 31 bytes. Refer to the manual pages for the correct syntax definition.

```
rcqconfig : failed to start
```

The following errors will also be reported in standard error if `rcqconfig` fails to start:

```
cfreg_start_transaction: `#2813: cfreg daemon not present`
```

The `rcqconfig` routine has failed. This error message usually indicates that the synchronization daemon is not running on the node. The cause of error messages of this pattern may be that the `cfreg` daemon has died and the previous error messages in the system log or console will indicate why the daemon died. Restart the daemon using `cfregd -r`. If it fails again, the error messages associated with it will indicate the problem. The data in the registry is most likely corrupted. If the problem persists, contact your customer service support representative.

```
cfreg_start_transaction: `#2815: registry is busy`
```

The `rcqconfig` routine has failed. This error message usually indicates that the daemon is not in synchronized state or if the transaction has been started by another application. This messages should not occur. The cause of error messages of this pattern is that the registries are not in consistent state. If the problem persists, unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem still persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_start_transaction: `#2810: an active transaction exists`
```


The `rcqconfig` routine has failed. This error message usually indicates that the application has already started a transaction. If the cluster is stable, the cause of error messages of this pattern is that different changes may be done concurrently from multiple nodes. Therefore, it might take longer time to commit. Retry the command again. If the problem persists, the cluster might not be in a stable state. If this is the case, unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
Too many nodename are defined for quorum. Max node = 64
```

This error message usually indicates that if the number of node specified are more than 64 for which the quorum is to be configured. The following errors will also be reported in standard error if nodename defined exceed the maximum limit:

```
cfreg_get: `#2809: specified transaction invalid`
```

The `rcqconfig` routine has failed. This error message usually indicates that the information supplied to get the specified data from the registry is not valid (e.g. transaction aborted due to time period expiring or synchronization daemon termination, etc.). This message should not occur. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_get: `#2804: entry with specified key does not exist`
```

The `rcqconfig` routine has failed. This error message usually indicates that the specified entry does not exist. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_get: `#2819: data or key buffer too small`
```

The `rcqconfig` routine has failed. This error message usually indicates that the specified size of the data buffer is too small to hold the entire data for the entry. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_put: `#2809: specified transaction invalid`
```

The `rcqconfig` routine has failed. This error message usually indicates that the information supplied to get the specified data from the registry is not valid (e.g. transaction aborted due to time period expiring or synchronization daemon termination, etc.). This message should not occur. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_put: `#2820: registry entry data too large`
```

The `rcqconfig` routine has failed. This error message usually indicates that the specified size data is larger than 28K. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cfreg_put: `#2807: data file format is corrupted`
```

The `rcqconfig` routine has failed. This error message usually indicates that the registry data file format has been corrupted. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

```
cms_post_event: `#0c01: event information is too large`
```

The `rcqconfig` routine has failed. This error message usually indicates that the event information (data being passed to the kernel) to be used for other sub-systems) is larger than 32K. The cause of error messages of this pattern is that the memory image may have somehow been damaged. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

12.5 rcquery messages

The `rcquery` command will generate an error message on `stderr` if an error condition is detected. Additional messages, giving more detailed information about this error, may be generated by the support routines of the `libcf` library. Please note that these additional error messages will only be written to the system-log file, and will not appear on `stdout` or `stderr`.

Refer to the `rcquery` manual page for an explanation of the command options and the associated functionality.

12.5.1 Usage message

A usage message will be generated if:

- An invalid `rcquery` option is specified.
- The `-h` option is specified.

```
Usage:    rcquery [ -v ] [ -l ] [-h]
          -v verbose
          -l loop
          -h help
```

12.5.2 Error messages

```
rcquery -v -l
failed to register user event
`# 0c0b: user level ENS event memory limit overflow`
```

The `rcquery` routine has failed. It usually indicates that either the total amount of memory allocated or the amount of memory allocated for use on a per-open basis exceed the limit. Try to unload the cluster by using `cfconfig -u` and reload the cluster by using `cfconfig -l`. If the problem persists, remove and then re-install the CF package. If this does not resolve the problem, contact your customer service support representative.

12.6 CF runtime messages

All CF runtime messages include an 80-byte ASCII `log3` prefix, which includes a timestamp, component number, error type, severity, version, product name, and structure id. This header is not included in the message descriptions that follow.

All of the following messages are sent to the system-log file, and ‘node up’ and ‘node down’ messages are also sent to the console.

There are some common tokens (shown in bold italic font) substituted into the error and warning messages that follow. If necessary, any not covered by this global explanation will be explained in the text associated with the specific message text.

- *clustername* — The name of the cluster to which the node belongs (or is joining). It is specified in the cluster configuration (see `cfconfig -s`).
- *err_type* — Identifies the type of ICF error reported. There are three types of errors:
 1. Debug (none in released product)
 2. Heartbeat missing
 3. Service error (usually, “route down”)
- *nodename* — The name by which a node is known within a cluster (usually derived from `uname -n`).
- *nodenum* — A unique number assigned to each and every node within a cluster.
- *route_dst* — The ICF route number (at the remote node) associated with a specific route.
- *route_src* — The ICF route number (on the local node) associated with a route. An ICF route is the logical connection established between two nodes over a cluster interconnect.
- *servername* — The nodename of the node acting as a join server for the local (client) node that is attempting to join the cluster.
- *service* — Denotes the ICF registered service number. There are currently over 30 registered ICF services.

This first set of messages are “special” in that they deal with the CF driver basic initialization and de-initialization:

CF: cf_attach Error: invalid command. (#0425 bad_cmd)
CF: cf_attach Error: invalid instance. (#0425 cf_instance instance)
CF: cf_attach Error: phase 1 init failure. (#reason_code)
CF: cf_attach Error: phase 2 init failure. (#reason_code)
CF: cf_attach Error: unable to create cf minor.
CF: cf_detach Error: invalid instance. (#0425 cf_instance instance)

These messages are associated with a CF initialization failure. They should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

12.6.1 Alphabetical list of messages

CF: carp_broadcast_version: Failed to announce version
cip_version

This message will occur if CIP fails to initialize successfully, indicating some sort of mismatch between CIP and CF. This message should not occur unless the CF driver and/or other kernel components have somehow been damaged. Remove and then re-install the CF package. If the problem persists, contact your customer support representative.

CF: carp_event: bad nodeid (#0000 nodenum)

This message is generated by CIP when a bad nodenumber is received.

CF: cip: Failed to register ens EVENT_CIP

This message is generated when CIP initialization cannot register for the event EVENT_CIP.

CF: cip: Failed to register ens EVENT_NODE_LEFTCLUSTER

This message is generated when CIP initialization cannot register for the event EVENT_NODE_LEFTCLUSTER.

CF: cip: Failed to register icf channel ICF_SVC_CIP_CTL

This message is generated when CIP initialization cannot register with ICF for the service ICF_SVC_CIP_CTL.

CF: cip: message SYNC_CIP_VERSION is too short

This message is generated when CIP receives a garbled message.

CF: `ens_nicf_input Error: unknown msg type received. (#0000 msgtype)`

This message is generated by ENS when a garbled message is received from ICF. The message is dropped.

CF: Giving UP Mastering (Cluster already Running).

This message is generated when a node detects a join server and joins an existing cluster, rather than forming a new one. No action is necessary.

CF: Giving UP Mastering (some other Node has Higher ID).

This message is generated when a node volunteers to be a join server, but detects an eligible join server with a higher id. No action is necessary.

CF: Icf Error: (service err_type route_src route_dst). (#0000 service err-type route_src route_dst)

This message is generated when ICF detects an error. It is most common to see this message in missing heartbeat and route down situations.

CF: Join client *nodename* timed out. (#0000 nodenum)

This message is generated on a node acting as a join server, when the client node does not respond in time.

CF: Join Error: Invalid configuration: multiple devs on same LAN.

This message is generated when a node is attempting to join or form a cluster. Multiple network interconnects cannot be attached to the same LAN segment.

CF: Join Error: Invalid configuration: asymmetric cluster.

This message is generated when a node is joining a cluster that has a active node that does not support asymmetric clustering, and has configured an incompatible (asymmetric) set of cluster interconnects.

CF: Join postponed: received packets out of sequence from *servername*.

This message is generated when a node is attempting to join a cluster, but is having difficulty communicating with the node acting as the join server. Both nodes will attempt to restart the join process.

CF: Join postponed, server *servername* is busy.

This message is generated when a node is attempting to join a cluster, but the join server is busy with another client node. (Only one join may be active in/on the cluster at a time.) Another reason for this message to

be generated is that the client node is currently in LEFTCLUSTER state. A node cannot re-join a cluster, unless its state is DOWN. (See the `cftool -k` manual page.)

CF: Join timed out, server *servername* did not send node number: retrying.

CF: Join timed out, server *servername* did not send nsm map: retrying.

CF: Join timed out, server *servername* did not send welcome message.

These messages are generated when a node is attempting to join a cluster, but is having difficulty communicating with the node acting as the join server. The join client node will attempt to continue the join process.

CF: Local node is missing a route from node: *nodename*

CF: missing route on local device: *devicename*

These messages are generated when an asymmetric join has occurred in a cluster, and the local node is missing a route to the new node. The *nodename* and *devicename* of the associated cluster interconnect are displayed, in case this is not the desired result.

CF: Local Node *nodename* Created Cluster *clustername*. (#0000 *nodenum*)

This message is generated when a node forms a new cluster.

CF: Local Node *nodename* Left Cluster *clustername*.

This message is generated when a node leaves a cluster.

CF: No join servers found.

This message is generated when a node cannot detect any nodes willing to act as join servers.

CF: Node *nodename* Joined Cluster *clustername*. (#0000 *nodenum*)

This message is generated when a node joins an existing cluster.

CF: Node *nodename* Left Cluster *clustername*. (#0000 *nodenum*)

This message is generated when a node leaves a cluster.

CF: Received out of sequence packets from join client: *nodename*

This message is generated when a node, acting as a join server, is having difficulty communicating with the client node. Both nodes will attempt to restart the join process.

CF: Starting Services.

This message is generated by CF as it is starting.

CF: Stopping Services.

This message is generated by CF as it is stopping.

CF: User level event memory overflow: Event dropped (#0000 *eventid*)

This message is generated when an ENS user event is received, but there is no memory for the event to be queued.

CF: *clustername: nodename* is Down. (#0000 *nodenum*)

This message is generated when a node has left the cluster in an orderly manner (i.e., `cfconfig -u`).

CF: *nodename* Error: local node has no route to node: join aborted.

This message is generated when a node is attempting to join a cluster, but detects that there is no route to one or more nodes that are already members of the cluster.

CF: *nodename* Error: no echo response from node: join aborted.

This message is generated when a node is attempting to join a cluster, but is having difficulty communicating with all the nodes in the cluster.

CF: *servername*: busy: cluster join in progress: retrying

CF: *servername*: busy: local node not DOWN: retrying

CF: *servername*: busy mastering: retrying

CF: *servername*: busy serving another client: retrying

CF: *servername*: local node's status is UP: retrying

CF: *servername*: new node number not available: join aborted

These messages are generated when a node is attempting to join a cluster, but the join server is busy with another client node. (Only one join may be active in/on the cluster at a time.) Another reason for this message to be generated is that the client node is currently in LEFTCLUSTER state. A node cannot re-join a cluster, unless its state is DOWN. (See the `cftool -k` manual page.)

CF (TRACE): *cip*: Announcing version *cip_version*

This message is generated when a CIP initialization is complete.

CF (TRACE): *EnsEV*: Shutdown

This message is generated when the ENS event-daemon shuts down.

CF (TRACE): *EnsND*: Shutdown

This message is generated when the ENS `node_down`-daemon shuts down.

CF (TRACE): Icf: Route UP: node src dest (#0000 *nodenum route_src route_dst*)

This message is generated when an ICF route is (re-)activated.

CF (TRACE): JoinServer: Stop

This message is generated when the join server mechanism is deactivated.

CF (TRACE): JoinServer: Startup

This message is generated when the join daemon is started.

CF (TRACE): JoinServer: ShutDown

This message is generated when an active join daemon shuts down.

CF (TRACE): Load: Complete

This message is generated when CF initialization is complete.

12.7 CF Reason Code table

Code	Reason	Service	Text
0401	REASON_SUCCESS		Operation was successful
	/* generic error codes */		
0401	REASON_NOERR	generic	Request not completed
0402	REASON_ALERTED	generic	Interrupted call
0403	REASON_TIMEOUT	generic	Timedout call
0404	REASON_NO_MEMORY	generic	Out of memory
0405	REASON_NO_SUCH_DEVICE	generic	No such device/ resource
0406	REASON_DEVICE_BUSY	generic	Resource is busy
0407	REASON_INVALID_PARAMETER	generic	Invalid parameter
0408	REASON_UNSUCCESSFUL	generic	Unsuccessful
0409	REASON_ADDRESS_ALREADY_EXISTS	generic	Address already exists
040a	REASON_BAD_ADDRESS	generic	Bad memory address
040b	REASON_INSUFFICIENT_RESOURCES	generic	Insufficient resources
040c	REASON_BUFFER_OVERFLOW	generic	Buffer overflow

Code	Reason	Service	Text
040d	REASON_INVALID_OWNER	generic	Invalid owner
040e	REASON_INVALID_HANDLE	generic	Invalid handle
040f	REASON_DUPNAME	generic	Duplicate name
0410	REASON_USAGE	generic	Usage
0411	REASON_NODATA	generic	No data
0412	REASON_NOT_INITIALIZED	generic	Driver not initialized
0413	REASON_UNLOADING	generic	Driver unloading
0414	REASON_REASSEMBLY_DOWN	generic	Sender died while sending data
0415	REASON_WENT_DOWN	generic	Destination node went down
0416	REASON_TRANSMIT_TIMEOUT	generic	Data transmission timeout
0417	REASON_BAD_PORT	generic	Bad destination port
0418	REASON_BAD_DEST	generic	Bad destination
0419	REASON_YANK	generic	Message transmission flushed
041a	REASON_SVC_BUSY	generic	SVC has pending transmissions
041b	REASON_SVC_UNREGISTER	generic	SVC has been unregistered
041c	REASON_INVALID_VERSION	generic	Invalid version
041d	REASON_NOT_SUPPORTED	generic	Function not supported
041e	REASON_EPERM	generic	Not super-user
041f	REASON_ENOENT	generic	No such file or directory
0420	REASON_EINTR	generic	Interrupted system call
0421	REASON_EIO	generic	I/O error
0422	REASON_ENXIO	generic	No such device or address (I/O req)
0423	REASON_EACCES	generic	Permission denied

Code	Reason	Service	Text
0424	REASON_EEXIST	generic	File exists
0425	REASON_DDI_FAILURE	generic	Error in DDI/DKI routine
0426	REASON_INVALID_NODENAME	generic	Invalid node name
0427	REASON_INVALID_NODENUMBER	generic	Invalid node number
0428	REASON_NODE_NOT_LEFTC	generic	Node is not in LEFTCLUSTER state
0429	REASON_CORRUPT_CONFIG	generic	Corrupt/invalid cluster config
042a	REASON_FLUSH	generic	Messages transmission flushed
042b	REASON_MAX_ENTRY	generic	Maximum entries reached
042c	REASON_NO_CONFIGURATION	generic	No configuration exists
	<i>/* mrpc reasons */</i>		
0801	REASON_MRPC_CLT_SVCUNAVAIL	mrpc	Service not registered on Client
0802	REASON_MRPC_SRV_SVCUNAVAIL	mrpc	Service not registered on Server
0803	REASON_MRPC_CLT_PROCVAIL	mrpc	Service Procedure not avail on Clt
0804	REASON_MRPC_SRV_PROCVAIL	mrpc	Service Procedure not avail on Srv
0805	REASON_MRPC_INARGTOOLONG	mrpc	Input argument size too big
0806	REASON_MRPC_OUTARGTOOLONG	mrpc	Output argument size too big
0807	REASON_MRPC_RETARGOVERFLOW	mrpc	Return argument size overflow
0808	REASON_MRPC_VERSIONMISMATCH	mrpc	Version mismatch
0809	REASON_MRPC_ICF_FAILURE	mrpc	ICF send failed
080a	REASON_MRPC_INTR	mrpc	Interrupted RPC

Code	Reason	Service	Text
080b	REASON_MRPC_RECURSIVE	mrpc	Illegal recursive call
080c	REASON_MRPC_SVC_EXIST	mrpc	Service already registered
	<i>/* ens reasons */</i>		
0c01	REASON_ENS_INFOTOOBIG	ens	Event information is too large
0c02	REASON_ENS_TOOSOON	ens	Attempt to post event before <code>ens_init</code>
0c03	REASON_ENS_NODEST	ens	Remote or local not specified in <code>howto</code>
0c04	REASON_ENS_DAEMONNOTIFY	ens	Invalid event posting by event daemon
0c05	REASON_ENS_NOICF	ens	Attempt to post remote before ICF config
0c06	REASON_ENS_OLDACKVERS	ens	Old version kernel has acked event
0c07	REASON_ENS_IMPLICITACK	ens	Event handler did not obtain ack handle
0c08	REASON_ENS_ACKNOTREQ	ens	Event acknowledgment not required
0c09	REASON_ENS_NOTEVHANDLER	ens	Obtainer of ack handle not event handler
0c0a	REASON_ENS_NOACKHANDLE	ens	Cannot locate event ack handle
0c0b	REASON_ENS_MEMLIMIT	ens	User level ENS event memory limit overflow
0c0c	REASON_ENS_DUPREG	ens	Duplicate event registration
0c0d	REASON_ENS_REGNOTFOUND	ens	Event registration not found
0c0e	REASON_ENS_INFOTOOSMALL	ens	Event information size too small

Code	Reason	Service	Text
0c0f	REASON_ENS_BADFAILNODE	ens	Node cannot post LEFTCLUSTER or NODE DOWN for itself
	<i>/* nsm */</i>		
1001	REASON_NSM_BADVERSION	nsm	Data structure version mismatch
1002	REASON_NSM_NONODES	nsm	No nodes have been specified
1003	REASON_NSM_TOOMANYNODES	nsm	Too many nodes have been specified
1004	REASON_NSM_BADNODEID	nsm	Node ID out of node name space range
1005	REASON_NSM_BADNETALEN	nsm	Invalid network address length
1006	REASON_NSM_ICFCREATE	nsm	Failure trying to create ICF node
1007	REASON_NSM_ICFDELETE	nsm	Failure trying to delete ICF node
1008	REASON_NSM_BADSTARTNODE	nsm	Invalid starting node specified
1009	REASON_NSM_BADINFOLEN	nsm	Invalid event information length
100a	REASON_NSM_BADCNODEID	nsm	Control node out of name space range
100b	REASON_NSM_BADCNSTATUS	nsm	Control node status invalid
100c	REASON_NSM_BADANODEID	nsm	Invalid node ID for node being added
100d	REASON_NSM_ADDNODEUP	nsm	Node being added is already operational
100e	REASON_NSM_NONODE	nsm	Node does not exist in the node name space

Code	Reason	Service	Text
100f	REASON_NSM_NODEFAILURE	nsm	A node has been declared dead
1010	REASON_NSM_NODETIMEOUT	nsm	Heartbeat timeout has expired for a node
1011	REASON_NSM_BADOUTSIZE	nsm	Invalid value for MRPC outsize
1012	REASON_NSM_BADINSIZE	nsm	Invalid value for MRPC insize
1013	REASON_NSM_BADNDNOTIFY	nsm	Failure to post NODE DOWN event
1014	REASON_NSM_VERSIONERR	nsm	nsetinfo versioning error
	<i>/* mrpc */</i>		
1401	REASON_ICF_MRPC_SZSM	icfmrpc	Output argument size too small
1402	REASON_ICF_MRPC_BADNDNUM	icfmrpc	Node does not exist
1403	REASON_ICF_MRPC_BADADDR	icfmrpc	mesh address does not exist
	<i>/* user events */</i>		
1801	REASON_UEV_ALREADYOPEN	uev	Process already has event device open
1802	REASON_UEV_TOOMANYEVENTS	uev	Too many user events initialized
1803	REASON_UEV_BADHANDLE	uev	Invalid user event handle specified
1804	REASON_UEV_NOTOPEN	uev	Process does not have event device open
1805	REASON_UEV_REGISTERED	uev	Duplicate user event registration
	<i>/* node group */</i>		

Code	Reason	Service	Text
1c01	REASON_NG_DEF_SYNTAX	ng	Bad definition syntax
1c02	REASON_NG_DUPNAME	ng	Name exists already
1c03	REASON_NG_EXIST	ng	Group does not exist
1c04	REASON_NG_ND_EXIST	ng	Node does not exist
1c05	REASON_NG_NAMELEN	ng	Too long a node name
1c06	REASON_NG_STATE	ng	Unknown parser state
1c07	REASON_NG_NODEINFO	ng	Failed to get up-node info
1c08	REASON_NG_ITER_STALE	ng	Iterator is stale
1c09	REASON_NG_ITER_NOSPACE	ng	Iterator pool exhausted
1c0a	REASON_NG_ITER_NOENT	ng	The end of iteration
1c0b	REASON_NG_MEMBER	ng	Node is not a group member
1c0c	REASON_NG_NOENT	ng	No node is up
1c0d	REASON_NG_UNPACK	ng	Failed to unpack definition
1c0e	REASON_NG_DUPDEF	ng	Identical group definition
<i>/* distributed mount services */</i>			
2001	REASON_DMS_INVALIDDCNG	dms	Invalid client node group
2002	REASON_DMS_MNTINUSE	dms	Mount in use
2003	REASON_DMS_DEVINUSE	dms	Device in use
2004	REASON_DMS_FSCKFAILED	dms	Failover <code>fsck</code> failed
2005	REASON_DMS_MNTFAILED	dms	Failover mount failed
2006	REASON_DMS_MNTBUSY	dms	Mount is busy
2007	REASON_DMS_NOMNTPT	dms	No mount point specified
2008	REASON_DMS_NODBENT	dms	Specified mount point not found

Code	Reason	Service	Text
2009	REASON_DMS_BADSTATE	dms	Server is up or failover in progress
200a	REASON_DMS_SUBMOUNT	dms	Specified mount point is CFS submount
200b	REASON_MAX_REASON_VAL	dms	Last reason
<i>/* join */</i>			
2401	REASON_JOIN_FAILED	join	Node has failed to join cluster
2402	REASON_JOIN_DISABLED	join	Cluster join not started
2403	REASON_JOIN_SHUTDOWN	join	Join daemon shut down
<i>/* cfreg */</i>			
2801	REASON_CFREG_STOPREQUESTED	cfreg	cfreg daemon stop requested
2802	REASON_CFREG_DUPDAEMON	cfreg	cfreg daemon already running
2803	REASON_CFREG_BADCONFIG	cfreg	Internal cfreg configuration error
2804	REASON_CFREG_NOENTRY	cfreg	Entry with specified key does not exist
2805	REASON_CFREG_COMMITTED	cfreg	Specified transaction committed
2806	REASON_CFREG_NOTOPEN	cfreg	Data file not open
2807	REASON_CFREG_CORRUPTFILE	cfreg	Data file format is corrupt
2808	REASON_CFREG_NSIERR	cfreg	Internal packaging error
2809	REASON_CFREG_INVALIDTRANS	cfreg	Specified transaction invalid
280a	REASON_CFREG_ACTIVETRANS	cfreg	An active transaction exists

Code	Reason	Service	Text
280b	REASON_CFREG_NOREQUESTS	cfreg	No daemon requests available
280c	REASON_CFREG_REQOVERFLOW	cfreg	Daemon request buffer overflow
280d	REASON_CFREG_NODAEMON	cfreg	cfreg daemon not present
280e	REASON_CFREG_BADREQUEST	cfreg	Unknown daemon request
280f	REASON_CFREG_REGBUSY	cfreg	Register is busy
2810	REASON_CFREG_REGOWNED	cfreg	Registry is owned
2811	REASON_CFREG_INVALIDUPDATE	cfreg	Invalid update
2812	REASON_CFREG_INVALIDKEY	cfreg	Invalid registry key
2813	REASON_CFREG_OVERFLOW	cfreg	Data or key buffer too small
2814	REASON_CFREG_TOOBIG	cfreg	Registry entry data too large
<i>/* cflog Message Catalogs */</i>			
2c01	REASON_CFLOG_NOCAT	cflog	cflog could not open message catalog
<i>/* qsm Message Catalogs */</i>			
3001	REASON_QSM_DUPMETHODNAME	qsm	Duplicate quorum method name
3002	REASON_QSM_TRYAGAIN	qsm	Need to try again later
3003	REASON_QSM_BUSY	qsm	Method has been registered already
3004	REASON_QSM_IDLE	qsm	Method has not been registered
3005	REASON_QSM_STOP	qsm	qsm stop requested
<i>/* sens */</i>			

Code	Reason	Service	Text
3401	REASON_SENS_BADSEQ	sens	Invalid sequence number
3402	REASON_SENS_TOOsoon	sens	SENS not initialized
3403	REASON_SENS_DUPACK	sens	Duplicate registration for completion ack
3404	REASON_SENS_NOREG	sens	Registration does not exist
3405	REASON_SENS_BADMAP	sens	Node missing from node map
3406	REASON_SENS_NOUREG	sens	User event registration does not exist
3407	REASON_SENS_NOUEVENT	sens	Event not received
<i>/* CFRS */</i>			
3801	REASON_CFRS_BADFCPSRCCONF	cfrs	cfcf not configured on source node
3802	REASON_CFRS_BADFCPDSTCONF	cfrs	cfcf not configured on destination node
3803	REASON_CFRS_BADEXECsrccONF	cfrs	cfsh not configured on source node
3804	REASON_CFRS_BADEXECdSTCONF	cfrs	cfsh not configured on execution node
3805	REASON_CFRS_BADDSTPATH	cfrs	Invalid destination file path
3806	REASON_CFRS_DSTPATHTOOLONG	cfrs	Destination file path too long
3807	REASON_CFRS_SRCACCESSERR	cfrs	Cannot access source file
3808	REASON_CFRS_SRCNOTREG	cfrs	Source file is not regular file
3809	REASON_CFRS_SRCREADERR	cfrs	Source file read error
380a	REASON_CFRS_NOCMD	cfrs	No command string specified

Code	Reason	Service	Text
380b	REASON_CFRS_CMDTOOLONG	cfrs	Command string too long
380c	REASON_CFRS_OUTPUTWRTErr	cfrs	Command output write error
380d	REASON_CFRS_NSIERROR	cfrs	Internal CFRS NSI error
380e	REASON_CFRS_DSTABORTEEXEC	cfrs	Execution aborted on execution node
380f	REASON_CFRS_INVALIDIOCTL	cfrs	Invalid ioctl call
3810	REASON_CFRS_BADDSTNODE	cfrs	Destination node not in cluster
3811	REASON_CFRS_BADROPHANDLE	cfrs	Bad remote operation handle
3812	REASON_CFRS_SRCEXECABORTED	cfrs	Remote exec aborted on source node
3813	REASON_CFRS_RESPOUTTOOSMALL	cfrs	Response output buffer too small
3814	REASON_CFRS_MRPCOUTSIZE	cfrs	Unexpected MRPC outsize error
3815	REASON_CFRS_DSTNODELEFT	cfrs	Destination node has left the cluster
3816	REASON_CFRS_DSTDAEMONDOWN	cfrs	cfregd on destination node down
3817	REASON_CFRS_DSTSTATERR	cfrs	Failure to stat dst file
3818	REASON_CFRS_DSTNOTREG	cfrs	Existing dstpath not regular file
3819	REASON_CFRS_DSTTMPOPENERR	cfrs	Cannot open tmp file on dst node
381a	REASON_CFRS_DSTTMPCHOWNERR	cfrs	Cannot chown tmp file on dst node
381b	REASON_CFRS_DSTTMPCHMODERR	cfrs	Cannot chmod tmp file on dst node

Code	Reason	Service	Text
381c	REASON_CFRS_DSTTMPWRITEERR	cfrs	tmp file write error on dst node
381d	REASON_CFRS_DSTTMPCLOSEERR	cfrs	tmp file close error on dst node
381e	REASON_CFRS_DSTRENAMEERR	cfrs	Failed to rename existing dstpath
381f	REASON_CFRS_TMPRENAMEERR	cfrs	Failed to tmp file to dstpath
3820	REASON_CFRS_DUPIFC	cfrs	Duplicate remote operation handle error
3821	REASON_CFRS_STALESUBFCREQ	cfrs	Stale remote operation handle error
3822	REASON_CFRS_BADSPAWN	cfrs	Failure to spawn exec cmd on dstnode
	<i>/* CFSF */</i>		
4001	REASON_CFSF_PENDING	cfsf	Invalid node down request with pending ICF failure
4002	REASON_MAX_REASON_VAL		Last reason

12.8 Error messages for different systems

Refer to the file `/usr/include/sys/errno.h` for the meaning of an ERRNO for a particular system.

12.9 Solaris/Linux ERRNO table

Solaris No	Linux No	Name	Description
1	1	EPERM	Operation not permitted / not super-user Typically this error indicates an attempt to modify a file in some way forbidden except to its owner or the super-user. It is also returned for attempts by ordinary users to do things allowed only to the super-user.
2	2	ENOENT	No such file or directory A file name is specified and the file should exist but doesn't, or one of the directories in a path name does not exist.
3	3	ESRCH	No such process, LWP, or thread No process can be found in the system that corresponds to the specified PID, LWPID_t, or thread_t.
4	4	EINTR	Interrupted system call An asynchronous signal (such as interrupt or quit), which the user has elected to catch, occurred during a system service function. If execution is resumed after processing the signal, it will appear as if the interrupted function call returned this error condition. In a multithreaded application, EINTR may be returned whenever another thread or LWP calls fork.(2)
5	5	EIO	I/O error Some physical I/O error has occurred. This error may in some cases occur on a call following the one to which it actually applies.

Solaris No	Linux No	Name	Description
6	6	ENXIO	No such device or address I/O on a special file refers to a sub-device which does not exist, or exists beyond the limit of the device. It may also occur when, for example, a tape drive is not on-line or no disk pack is loaded on a drive.
7	7	E2BIG	Arg list too long An argument list longer than ARG_MAX bytes is presented to a member of the exec family of functions (see <code>exec(2)</code>). The argument list limit is the sum of the size of the argument list plus the size of the environment's exported shell variables.
8	8	ENOEXEC	Exec format error A request is made to execute a file which, although it has the appropriate permissions, does not start with a valid format (see <code>a.out(4)</code>).
9	9	EBADF	Bad file number Either a file descriptor refers to no open file, or a <code>read(2)</code> (respectively, <code>write(2)</code>) request is made to a file that is open only for writing (respectively, reading).
10	10	ECHILD	No child processes A <code>wait(2)</code> function was executed by a process that had no existing or unwaited-for child processes.

Solaris No	Linux No	Name	Description
11	11	EAGAIN	Try again / no more processes or no more LWPs For example, the <code>fork(2)</code> function failed because the system's process table is full or the user is not allowed to create any more processes, or a call failed because of insufficient memory or swap space.
12	12	ENOMEM	Out of memory / not enough space During execution of <code>brk()</code> or <code>sbrk()</code> (see <code>brk(2)</code>), or one of the <code>exec</code> family of functions, a program asks for more space than the system is able to supply. This is not a temporary condition; the maximum size is a system parameter. On some architectures, the error may also occur if the arrangement of text, data, and stack segments requires too many segmentation registers, or if there is not enough swap space during the <code>fork(2)</code> function. If this error occurs on a resource associated with Remote File Sharing (RFS), it indicates a memory depletion which may be temporary, dependent on system activity at the time the call was invoked.
13	13	EACCES	Permission denied An attempt was made to access a file in a way forbidden by the protection system.

Solaris No	Linux No	Name	Description
14	14	EFAULT	Bad address The system encountered a hardware fault in attempting to use an argument of a routine. For example, <code>errno</code> potentially may be set to <code>EFAULT</code> any time a routine that takes a pointer argument is passed an invalid address, if the system can detect the condition. Because systems will differ in their ability to reliably detect a bad address, on some implementations passing a bad address to a routine will result in undefined behavior.
15	15	ENOTBLK	Block device required A non-block device or file was mentioned where a block device was required (for example, in a call to the <code>mount(2)</code> function).
16	16	EBUSY	Device or resource busy An attempt was made to mount a device that was already mounted or an attempt was made to unmount a device on which there is an active file (open file, current directory, mounted-on file, active text segment). It will also occur if an attempt is made to enable accounting when it is already enabled. The device or resource is currently unavailable. <code>EBUSY</code> is also used by mutexes, semaphores, condition variables, and read-write locks, to indicate that a lock is held, and by the processor control function <code>P_ONLINE</code> .
17	17	EEXIST	File exists An existing file was mentioned in an inappropriate context (for example, call to the <code>link(2)</code> function).

Solaris No	Linux No	Name	Description
18	18	EXDEV	Cross-device link A hard link to a file on another device was attempted.
19	19	ENODEV	No such device An attempt was made to apply an inappropriate operation to a device (for example, read a write-only device).
20	20	ENOTDIR	Not a directory A non-directory was specified where a directory is required (for example, in a path prefix or as an argument to the <code>chdir(2)</code> function).
21	21	EISDIR	Is a directory An attempt was made to write on a directory.
22	22	EINVAL	Invalid argument An invalid argument was specified (for example, unmounting a non-mounted device), mentioning an undefined signal in a call to the <code>signal(3C)</code> or <code>kill(2)</code> function.
23	23	ENFILE	File table overflow The system file table is full (that is, <code>SYS_OPEN</code> files are open, and temporarily no more files can be opened).
24	24	EMFILE	Too many open files No process may have more than <code>OPEN_MAX</code> file descriptors open at a time.
25	25	ENOTTY	Not a TTY - inappropriate ioctl for device A call was made to the <code>ioctl(2)</code> function specifying a file that is not a special character device.

Solaris No	Linux No	Name	Description
26	26	ETXTBSY	Text file busy (obsolete) An attempt was made to execute a pure-procedure program that is currently open for writing. Also an attempt to open for writing or to remove a pure-procedure program that is being executed.
27	27	EFBIG	File too large The size of the file exceeded the limit specified by resource <code>RLIMIT_FSIZE</code> ; the file size exceeds the maximum supported by the file system; or the file size exceeds the offset maximum of the file descriptor.
28	28	ENOSPC	No space left on device While writing an ordinary file or creating a directory entry, there is no free space left on the device. In the <code>fcntl(2)</code> function, the setting or removing of record locks on a file cannot be accomplished because there are no more record entries left on the system.
29	29	ESPIPE	Illegal seek A call to the <code>lseek(2)</code> function was issued to a pipe.
30	30	EROFS	Read-only file system An attempt to modify a file or directory was made on a device mounted read-only.
31	31	EMLINK	Too many links An attempt to make more than the maximum number of links, <code>LINK_MAX</code> , to a file.

Solaris No	Linux No	Name	Description
32	32	EPIPE	Broken pipe A write on a pipe for which there is no process to read the data. This condition normally generates a signal; the error is returned if the signal is ignored.
33	33	EDOM	Math argument out of domain of function The argument of a function in the math package (3M) is out of the domain of the function.
34	34	ERANGE	Math result not representable The value of a function in the math package (3M) is not representable within node precision.
35	42	ENOMSG	No message of desired type An attempt was made to receive a message of a type that does not exist on the specified message queue (see <code>msgrcv(2)</code>).
36	43	EIDRM	Identifier removed This error is returned to processes that resume execution due to the removal of an identifier from the file system's name space (see <code>msgctl(2)</code> , <code>semctl(2)</code> , and <code>shmctl(2)</code>).
37	44	ECHRNG	Channel number out of range
38	45	EL2NSYNC	Level 2 not synchronized
39	46	EL3HLT	Level 3 halted
40	47	EL3RST	Level 3 reset
41	48	ELNRNG	Link number out of range
42	49	EUNATCH	Protocol driver not attached
43	50	ENOCSI	No CSI structure available
44	51	EL2HLT	Level 2 halted

Solaris No	Linux No	Name	Description
45	35	EDEADLK	Resource deadlock condition A deadlock situation was detected and avoided. This error pertains to file and record locking, and also applies to mutexes, semaphores, condition variables, and read-write locks.
46	37	ENOLCK	No record locks available There are no more locks available. The system lock table is full (see <code>fcntl(2)</code>).
47	125	ECANCELED	Operation canceled The associated asynchronous operation was canceled before completion.
48	95	ENOTSUP	Not supported This version of the system does not support this feature. Future versions of the system may provide support.
49	122	EDQUOT	Disc quota exceeded A <code>write(2)</code> to an ordinary file, the creation of a directory or symbolic link, or the creation of a directory entry failed because the user's quota of disk blocks was exhausted, or the allocation of an inode for a newly created file failed because the user's quota of inodes was exhausted.
50	52	EBADE	Invalid exchange
51	53	EBADR	Invalid request descriptor
52	54	EXFULL	Exchange full
53	55	ENOANO	No anode
54	56	EBADRQC	Invalid request code
55	57	EBADSLT	Invalid slot
56	35	EDEADLOCK	File locking deadlock error
57	59	EBFONT	Bad font file format

Solaris No	Linux No	Name	Description
58	-	EOWNERDEAD	Process died with the lock
59	-	ENOTRECOVERABLE	Lock is not recoverable
60	60	ENOSTR	Device not a stream A <code>putmsg(2)</code> or <code>getmsg(2)</code> call was attempted on a file descriptor that is not a STREAMS device.
61	61	ENODATA	No data available No data (for no-delay I/O).
62	62	ETIME	Timer expired The timer set for a STREAMS <code>ioctl(2)</code> call has expired. The cause of this error is device-specific and could indicate either a hardware or software failure, or perhaps a timeout value that is too short for the specific operation. The status of the <code>ioctl()</code> operation is indeterminate. This is also returned in the case of <code>_lwp_cond_timedwait(2)</code> or <code>cond_timedwait(2)</code> .
63	63	ENOSR	Out of stream resources During a STREAMS <code>open(2)</code> call, either no STREAMS queues or no STREAMS head data structures were available. This is a temporary condition; one may recover from it if other processes release resources.
64	64	ENONET	Node is not on the network This error is Remote File Sharing (RFS) specific. It occurs when users try to advertise, unadvertise, mount, or unmount remote resources while the node has not done the proper startup to connect to the network.
65	65	ENOPKG	Package not installed This error occurs when users attempt to use a call from a package which has not been installed.

Solaris No	Linux No	Name	Description
66	66	EREMOTE	Object is remote This error is RFS-specific. It occurs when users try to advertise a resource which is not on the local node, or try to mount/unmount a device (or pathname) that is on a remote node.
67	67	ENOLINK	Link has been severed This error is RFS-specific. It occurs when the link (virtual circuit) connecting to a remote node is gone.
68	68	EADV	Advertise error This error is RFS-specific. It occurs when users try to advertise a resource which has been advertised already, or try to stop RFS while there are resources still advertised, or try to force unmount a resource when it is still advertised.
69	69	ESRMNT	Srmount error This error is RFS-specific. It occurs when an attempt is made to stop RFS while resources are still mounted by remote nodes, or when a resource is readvertised with a client list that does not include a remote node that currently has the resource mounted.
70	70	ECOMM	Communication error on send This error is RFS-specific. It occurs when the current process is waiting for a message from a remote node, and the virtual circuit fails.
71	71	EPROTO	Protocol error Some protocol error occurred. This error is device-specific, but is generally not related to a hardware failure
72	-	ELOCKUNMAPPED	Locked lock was unmapped

Solaris No	Linux No	Name	Description
74	72	EMULTIHOP	Multihop attempted This error is RFS-specific. It occurs when users try to access remote resources which are not directly accessible.
76	73	EDOTDOT	RFS specific error This error is RFS-specific. A way for the server to tell the client that a process has transferred back from mount point.
77	74	EBADMSG	Not a data message /* trying to read unreadable message */ During a read(2), getmsg(2), or ioctl(2) L_RECVFD call to a STREAMS device, something has come to the head of the queue that can not be processed. That something depends on the call: read(): control information or passed file descriptor. getmsg():passed file descriptor. ioctl():control or data information.
78	36	ENAMETOOLONG	File name too long The length of the path argument exceeds PATH_MAX, or the length of a path component exceeds NAME_MAX while _POSIX_NO_TRUNC is in effect; see limits(4).
79	75	EOVERFLOW	Value too large for defined data type
80	76	ENOTUNIQ	Name not unique on network Given log name not unique.
81	77	EBADF	File descriptor in bad state Either a file descriptor refers to no open file or a read request was made to a file that is open only for writing.

Solaris No	Linux No	Name	Description
83	79	ELIBACC	Cannot access a needed shared library Trying to exec an a.out that requires a static shared library and the static shared library does not exist or the user does not have permission to use it.
84	80	ELIBBAD	Accessing a corrupted shared library Trying to exec an a.out that requires a static shared library (to be linked in) and exec could not load the static shared library. The static shared library is probably corrupted.
85	81	ELIBSCN	.lib section in a.out corrupted Trying to exec an a.out that requires a static shared library (to be linked in) and there was erroneous data in the .lib section of the a.out. The .lib section tells exec what static shared libraries are needed. The a.out is probably corrupted.
86	82	ELIBMAX	Attempting to link in too many shared libraries Trying to exec an a.out that requires more static shared libraries than is allowed on the current configuration of the system. See NFS Administration Guide.
87	83	ELIBEXEC	Cannot exec a shared library directly Attempting to exec a shared library directly.
88	84	EILSEQ	Illegal byte sequence Illegal byte sequence when trying to handle multiple characters as a single character.
89	38	ENOSYS	Function not implemented / operation not applicable Unsupported file system operation.

Solaris No	Linux No	Name	Description
90	40	ELOOP	Symbolic link loop Number of symbolic links encountered during path name traversal exceeds MAXSYMLINKS.
91	85	ERESTART	Restartable system call Interrupted system call should be restarted.
92	86	ESTRPIPE	Streams pipe error (not externally visible) If pipe/FIFO, don't sleep in stream head.
93	39	ENOTEMPTY	Directory not empty
94	87	EUSERS	Too many users Too many users (for UFS).
95	88	ENOTSOCK	Socket operation on non-socket
96	89	EDESTADDRREQ	Destination address required A required address was omitted from an operation on a transport endpoint. Destination address required.
97	90	EMSGSIZE	Message too long A message sent on a transport provider was larger than the internal message buffer or some other network limit.
98	91	EPROTOTYPE	Protocol wrong type for socket A protocol was specified that does not support the semantics of the socket type requested.
99	92	ENOPROTOOPT	Protocol not available A bad option or level was specified when getting or setting options for a protocol.
120	93	EPROTONOSUPPORT	Protocol not supported The protocol has not been configured into the system or no implementation for it exists.

Solaris No	Linux No	Name	Description
121	94	ESOCKTNSUPPORT	Socket type not supported The support for the socket type has not been configured into the system or no implementation for it exists.
122	95	EOPNOTSUPP	Operation not supported on transport end-point For example, trying to accept a connection on a datagram transport endpoint.
123	96	EPFNOSUPPORT	Protocol family not supported The protocol family has not been configured into the system or no implementation for it exists. Used for the Internet protocols.
124	97	EAFNOSUPPORT	Address family not supported by protocol An address incompatible with the requested protocol was used.
125	98	EADDRINUSE	Address already in use User attempted to use an address already in use, and the protocol does not allow this.
126	99	EADDRNOTAVAIL	Cannot assign requested address Results from an attempt to create a transport end-point with an address not on the current node.
127	100	ENETDOWN	Network is down Operation encountered a dead network.
128	101	ENETUNREACH	Network is unreachable Operation was attempted to an unreachable network.
129	102	ENETRESET	Network dropped connection because of reset The node you were connected to crashed and rebooted.

Solaris No	Linux No	Name	Description
130	103	ECONNABORTED	Software caused connection abort A connection abort was caused internal to your node.
131	104	ECONNRESET	Connection reset by peer A connection was forcibly closed by a peer. This normally results from a loss of the connection on the remote node due to a timeout or a reboot
132	105	ENOBUFS	No buffer space available An operation on a transport endpoint or pipe was not performed because the system lacked sufficient buffer space or because a queue was full.
133	106	EISCONN	Transport endpoint is already connected A connect request was made on an already connected transport endpoint; or, a <code>sendto(3N)</code> or <code>sendmsg(3N)</code> request on a connected transport endpoint specified a destination when already connected.
134	107	ENOTCONN	Transport endpoint is not connected A request to send or receive data was disallowed because the transport endpoint is not connected and (when sending a datagram) no address was supplied.
135	117	EUCLEAN	Structure needs cleaning
137	118	ENOTNAM	Not a XENIX named type file
138	119	ENAVAIL	No XENIX semaphores available
139	120	EISNAM	Is a named type file
140	121	EREMOTEIO	Remote I/O error
141	-	EINIT	Define EINIT 141 /* Reserved for future use* /

Solaris No	Linux No	Name	Description
142	-	EREMDEV	Define EREMDEV 142 /* Error 142 */
143	108	ESHUTDOWN	Cannot send after transport endpoint shutdown A request to send data was disallowed because the transport endpoint has already been shut down.
144	109	ETOOMANYREFS	Too many references: cannot splice
145	110	ETIMEDOUT	Connection timed out A <code>connect(3N)</code> or <code>send(3N)</code> request failed because the connected party did not properly respond after a period of time; or a <code>write(2)</code> or <code>fsync(3C)</code> request failed because a file is on an NFS file system mounted with the <code>soft</code> option.
146	111	ECONNREFUSED	Connection refused No connection could be made because the target node actively refused it. This usually results from trying to connect to a service that is inactive on the remote node.
147	112	EHOSTDOWN	Node is down A transport provider operation failed because the destination node was down.
148	113	EHOSTUNREACH	No route to node A transport provider operation was attempted to an unreachable node.
149	114	EALREADY	Operation already in progress An operation was attempted on a non-blocking object that already had an operation in progress.

Solaris No	Linux No	Name	Description
150	115	EINPROGRESS	Operation now in progress An operation that takes a long time to complete (such as a <code>connect()</code>) was attempted on a non-blocking object).
151	116	ESTALE	Stale NFS file handle
-	11	EWOULDBLOCK	Operation would block
-	123	ENOMEDIUM	No medium found
-	124	EMEDIUMTYPE	Wrong medium type

12.10 Resource Database messages

This section explains the Resource Database message. The message format is described below.

Italic indicates that the output content varies depending on the message.

`FJSVcluster:severity:program:message-number:message`

<i>severity</i>	Indicates the message severity level. On the message severity level, there are four types: Stop (HALT), Information (INFORMATION), Warning (WARNING), Error (ERROR). For details, refer to the table below.
<i>program</i>	Indicates the name of the Resource Database program that output this message.
<i>message-number</i>	Indicates the message number.
<i>message</i>	Indicates the message text.

Number	Message severity level	Meaning
0000-0999	Stop (HALT)	Message indicating an abnormal termination of the function in the Resource Database is output.

Table 9: Resource Database severity levels

Number	Message severity level	Meaning
2000–3999	Information (INFORMATION)	Message providing notification of information on the Resource Database operation status is output.
4000–5999	Warning (WARNING)	Message providing notification of a minor error not leading to abnormal termination of the function in the Resource Database is output.
6000–7999	Error (ERROR)	Message providing notification of a major error leading to abnormal termination of the function in the Resource Database is output.

Table 9: Resource Database severity levels

12.10.1 HALT messages

- 0100 Cluster configuration management facility terminated abnormally.
Corrective action
 Correct the cause of abnormal termination, then restart the error-detected node.
Supplement
 The cause of abnormal termination is indicated in the previous error message.
- 0101 Initialization of cluster configuration management facility terminated abnormally.
Corrective action
 Correct the cause of abnormal termination, then restart the error-detected node.
Supplement
 The cause of abnormal termination is indicated in the previous error message.
- 0102 A failure occurred in the server. It will be terminated.
Corrective action
 Follow the corrective action of the error message that was displayed right before this 0102 message

12.10.2 INFO messages

- 2100 The resource data base has already been set.
(detail:*code1-code2*)
- 2200 Cluster configuration management facility initialization started.
- 2201 Cluster configuration management facility initialization completed.
- 2202 Cluster configuration management facility exit processing started.
- 2203 Cluster configuration management facility exit processing completed.
- 2204 Cluster event control facility started.
- 2205 Cluster event control facility stopped.
- 3200 Cluster resource management facility initialization started.
- 3201 Cluster resource management facility initialization completed.
- 3202 Cluster resource management facility exit processing completed.
- 3203 Resource activation processing started.
- 3204 Resource activation processing completed.
- 3205 Resource deactivation processing started.
- 3206 Resource deactivation processing completed.

12.10.3 WARNING messages

4250 The line switching unit cannot be found because FJSVclswu is not installed.

Supplement

Devices other than the line switching unit register an automatic resource.

5200 There is a possibility that the resource controller does not start. (ident:*ident* command:*command*, ...)

Supplement

Notification of the completion of startup has not yet been posted from the resource controller.*indent* indicates a resource controller identifier while *command* indicates the startup script of the resource controller.

12.10.4 ERROR messages

???? Message not found!!

Corrective action

The text of the message corresponding to the message number is not available. Copy this message and contact your local customer support.

6000 An internal error occurred.(function:*function*
detail:*code1-code2-code3-code4*)

Corrective action

An internal error occurred in the program.

Record this message and collect information for an investigation.

Then, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).

function, code1, code2, code3, code4 indicates information required for error investigation.

6001 Insufficient memory. (detail:*code1-code2*)

Corrective action

Memory resources are insufficient to operate the Resource Database.

code1, code2 indicates information required for error investigation.

Record this message. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”). Review the estimating of memory resources.

If this error cannot be corrected by this operator response, contact your local customer support.

6002 Insufficient disk or system resources. (detail:code1-code2)

Corrective action

This failure might be attributed to the followings:

- The disk space is insufficient
- There are incorrect settings in the kernel parameter

Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).

Check that there is enough free disk space required for PRIMECLUSTER operation. If the disk space is insufficient, you need to reserve some free area and reboot the node. For the required disk space, refer to the PRIMECLUSTER *Installation Guide*.

If you still have this problem after going through the above instruction, confirm that the kernel parameter is correctly set. Modify the settings if necessary and reboot the node. Nevertheless, the above instructions are not helpful, contact your customer service representative. **code1** and **code2** indicate information required for troubleshooting.

6003 Error in option specification. (option:option)

Corrective action

Specify the correct option, then re-execute the processing. *option* indicates an option.

6004 No system administrator authority.

Corrective action

Re-execute the processing with the system administrator authority.

6005 Insufficient shared memory. (detail:code1-code2)

Corrective action

Shared memory resources are insufficient for the Resource Database to operate.

Record this message. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).

Refer to the Section “Kernel parameters for Resource Database” to review the estimate of shared memory resources (kernel parameters), Reboot the nodes that have any kernel parameters that have been changed.

If this error cannot be corrected by this operator response, contact your local customer support.

code1, *code2* indicates information required for error investigation.

- 6006 The required option *option* must be specified.
Corrective action
Specify the correct option, then re-execute the processing.
option indicates an option.
- 6007 One of the required options *option* must be specified.
Corrective action
Specify the correct option, then re-execute the processing.
option indicates an option.
- 6008 If option *option1* is specified, option *option2* is required.
Corrective action
If the option indicated by *option1* is specified, the option indicated by *option2* is required. Specify the correct option, then re-execute the processing.
- 6009 If option *option1* is specified, option *option2* cannot be specified.
Corrective action
If the option indicated by *option1* is specified, the option indicated by *option2* cannot be specified. Specify the correct option, then re-execute the processing.
- 6010 If any one of the options *option1* is specified, option *option2* cannot be specified.
Corrective action
If either option indicated by *option1* is specified, the option indicated by *option2* cannot be specified. Specify the correct option, then re-execute the processing.
- 6021 The *option* option(s) must be specified in the following order: *order*
Corrective action
Specify option *options* sequentially in the order of *order*. Then, retry execution.
option indicates those options that are specified in the wrong order, while *order* indicates the correct order of specification.

- 6025 The value of option *option* must be specified from *value1* to *value2*
Corrective action
Specify the value of the *option* in option within the range between *value1* and *value2*, and then re-execute.
option indicates the specified option while *value1*, *value2* indicate values.
- 6200 Cluster configuration management facility: configuration database mismatch. (name:*name* node:*node*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section “Collecting troubleshooting information”).
Collect the investigation information in all nodes, then reactivate the faulty node.
name indicates a database name in which a mismatch occurred, while *node* indicates a node in which an error occurred.
- 6201 Cluster configuration management facility: internal error. (node:*node* code:*code*)
Corrective action
There might be an error in the system if the kernel parameter `/etc/system (4)` is not properly set up when the cluster was installed. Check if the setup is correct (refer to Section “Kernel parameters for Resource Database”). If incorrect, reset the value of `/etc/system(4)`, and then restart the system.
If there's still any problem regardless of the fact that the value of `/etc/system (4)` is larger than the required by Resource Database, and the same value is shown when checked by a `sysdef(1M)` command, take down the message, collect information for investigation, and then contact your local customer support (refer to the Section “Collecting troubleshooting information”).
Collect the investigation information in all nodes, then reactivate the faulty node.
node indicates a node in which an error occurred while *code* indicates the code for the detailed processing performed for the error.

- 6202 Cluster event control facility: internal error.
(detail:*code1-code2*)
Corrective action
Record this message and collect information for an investigation.
Then, contact your local customer support. Collect information
required for troubleshooting (refer to the Section “Collecting trouble-
shooting information”).
code1, code2 indicates information required for error investigation.
- 6203 Cluster configuration management facility: communi-
cation path disconnected.
Corrective action
Check the state of other nodes and path of a private LAN.
- 6204 Cluster configuration management facility has not been
started.
Corrective action
Record this message and collect information for an investigation.
Then, contact your local customer support. Collect information
required for troubleshooting (refer to the Section “Collecting trouble-
shooting information”).
- 6206 Cluster configuration management facility: error in
definitions used by *target* command.
Corrective action
Record this message and collect information for an investigation.
Then, contact your local customer support. Collect information
required for troubleshooting (refer to the Section “Collecting trouble-
shooting information”).
target indicates a command name.
- 6207 Cluster domain contains one or more inactive nodes.
Corrective action
Activate the node in the stopped state.
- 6208 Access denied (*target*).
Corrective action
Record this message and collect information for an investigation.
Then, contact your local customer support. Collect information
required for troubleshooting (refer to the Section “Collecting trouble-
shooting information”).
target indicates a command name.

- 6209 The specified file or cluster configuration database does not exist (*target*).
- Corrective action**
Record this message and collect information for an investigation. Then, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
target indicates a file name or a cluster configuration database name.
- 6210 The specified cluster configuration database is being used (*table*).
- Corrective action**
Record this message and collect information for an investigation. Then, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
table indicates a cluster configuration database name.
- 6211 A table with the same name exists (*table*).
- Corrective action**
Record this message and collect information for an investigation. Then, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
table indicates a cluster configuration database name.
- 6212 The specified configuration change procedure is already registered (*proc*).
- Corrective action**
Record this message and collect information for an investigation. Then, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
proc indicates a configuration change procedure name.
- 6213 The cluster configuration database contains duplicate information.
- Corrective action**
Record this message and collect information for an investigation. Then, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).

- 6214 Cluster configuration management facility: configuration database update terminated abnormally (*target*).
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
Collect the investigation information in all nodes, then reactivate all nodes.
target indicates a cluster configuration database name.
- 6215 Cannot exceed the maximum number of nodes.
Corrective action
Since a hot extension is required for an additional node that exceeds the maximum number of configuration nodes that is allowed with Resource Database, review the cluster system configuration so that the number of nodes becomes equal to or less than the maximum number of composing nodes.
- 6216 Cluster configuration management facility: configuration database mismatch occurred because another node ran out of memory. (name:*name* node:*node*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
After collecting data for all nodes, stop the node and start it again.
name indicates a database in which a mismatch occurred and *node* indicates a node for which a memory shortfall occurred.

6217 Cluster configuration management facility: configuration database mismatch occurred because another node ran out of disk or system resources. (name:*name* node:*node*)

Corrective action

Record this message and collect information for an investigation. Then, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).

Reexamine the estimate for the disk resources and system resources (kernel parameter) (refer to the Section “Kernel parameters for Resource Database”). When the kernel parameter is changed for a given node, restart that node. If this error cannot be corrected by this operator response, contact your local customer support. After collecting data for all nodes, stop and then restart the nodes.

name indicates a database in which a mismatch occurred and *node* indicates the node in which insufficient disk resources or system resources occurred.

6218 An error occurred during distribution of file to the stopped node. (name:*name* node:*node* errno:*errno*)

Corrective action

File cannot be distributed to the stopped node from the erroneous node. Be sure to start the stopped node before the active node stops. It is not necessary to re-execute the command.

name indicates the file name that was distributed when an failure occurred, *node* indicates the node in which a failure occurred, and *errno* indicates the error number when a failure occurred.

6219 The cluster configuration management facility cannot recognize the activating node. (detail:*code1-code2*)

Corrective action

Confirm that there is no failures in Cluster Foundation (CF) or cluster interconnect. If a failure occurs in CF, take the corrective action of the CF message. If a failure occurs in cluster interconnect, check that NIC is connected to the network.

If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).

code1 and *code2* indicate information required for troubleshooting.

- 6220 The communication failed between nodes or processes in the cluster configuration management facility. (detail:*code1-code2*)
Corrective action
Confirm that there is no failures in cluster interconnect. If a failure occurs in cluster interconnect, check that NIC is connected to the network.
If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
code1 and *code2* indicate information required for troubleshooting.
- 6221 Invalid kernel parameter used by cluster configuration database. (detail:*code1-code2*)
Corrective action
The kernel parameter used for the Resource Database is not correctly set up. Modify the settings, referring to Section “Kernel parameters for Resource Database”, and reboot the node.
If you still have this problem after going through the above instruction, contact your local customer support.
code1 and *code2* indicate information required for troubleshooting.

6222 network service used by the cluster configuration management facility is not available.(detail:*code1-code2*)

Corrective action

Confirm the `/etc/inet/services` file is linked to the `/etc/services` file. If not, you need to create a symbolic link to the `/etc/services` file. When setup process is done, confirm the following network services are set up in the `/etc/inet/services` file. If any of the followings should fail to be set up, you need to add the missing.

```
dcmcom 9331/tcp# FJSVcldbm package
dcmSync 9379/tcp# FJSVcldbm package
dcm1ck 9378/tcp# FJSVcldbm package
dcmfcp 9377/tcp# FJSVcldbm package
dcmMst 9375/tcp# FJSVcldbm package
dcmEvm 9376/tcp# FJSVcldbm package
```

If this process is successfully done, confirm that the services of the `/etc/nsswitch.conf` file are defined as `services: files nisplus`. If not, you need to define them and reboot the node.

```
services: files nisplus
```

If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).

code1 and *code2* indicate information required for troubleshooting.

6223 A failure occurred in the specified command. (command: *command*, detail:*code1-code2*)

Corrective action

Confirm that you can run the program specified as an option of the `clxec(1M)` command.If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).

code1 and *code2* indicate information required for troubleshooting.

- 6226 The kernel parameter setup is not sufficient to operate the cluster control facility. (detail:*code*)
Corrective action
The kernel parameter used for the Resource Database is not correctly setup. Modify the settings, referring to the Section “Kernel parameters for Resource Database”, and reboot the node.
Then, execute the `clinitreset(1M)` command, reboot the node, and initialize the Resource Database again. Confirm that you can run the program specified as an option of the `clexec(1M)` command.
If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
code indicates a parameter type and its current value.
- 6250 Cannot run this command because FJSVclswu is not installed.
Corrective action
Install the FJSVclswu package before executing the command. Refer to the PRIMECLUSTER *Installation Guide* for further details.
- 6300 Failed in setting the resource data base (insufficient user authority).
Corrective action
No CIP is set up in the Cluster Foundation. Reset CIP, and execute again after rebooting all nodes. Refer to the Section “CF, CIP, and CIM configuration” for the setup method.
If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
code1 and *code2* represents information for investigation.
- 6301 The resource data base has already been set (insufficient user authority).
Corrective action
The setup for Resource Database is not necessary. If you need to reset the setup, execute the `clinitreset(1M)` command on all nodes, initialize the Resource Database, and then reboot all nodes. For details, refer to the manual of the `clinitreset(1M)` command.
code1 and *code2* represents information for investigation.

- 6302 Failed to create a backup of the resource database information. (detail:*code1-code2*)
Corrective action
The disk space might be insufficient. You need to reserve 1 MB or more of free disk space, and back up the Resource Database information again.
If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
code1 and *code2* indicate information required for troubleshooting.
- 6303 Failed restoration of the resource database information.(detail:*code1-code2*)
Corrective action
The disk space might be insufficient. You need to reserve 1 MB or more of free disk space, and restore the Resource Database information again.
If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
code1 and *code2* indicate information required for troubleshooting.
- 6600 Cannot manipulate the specified resource. (insufficient user authority)
Corrective action
Re-execute the specified resource with registered user authority.
- 6601 Cannot delete the specified resource. (resource:
resource rid:rid)
Corrective action
Specify the resource correctly, and then re-execute it.
resource indicates the resource name of the specified resource. *rid* indicates the resource ID of the specified resource.
- 6602 The specified resource does not exist. (detail:*code1-code2*)
Corrective action
Specify the correct resource, then re-execute the processing.
code1, code2 indicates information required for error investigation.

- 6603 The specified file does not exist.
Corrective action
Specify the correct file, then re-execute the processing.
- 6604 The specified resource class does not exist.
Corrective action
Specify the correct resource class, and then re-execute the processing.
A specifiable resource class is a file name itself that is under `/etc/opt/FJsvcluster/classes`. Confirm that there is no error in the character strings that have been specified as the resource class.
- 6606 Operation cannot be performed on the specified resource because the corresponding cluster service is not in the stopped state. (detail:*code1-code2*)
Corrective action
Stop the cluster service, then re-execute the processing.
code1, code2 indicates information required for error investigation.
- 6607 The specified node cannot be found.
Corrective action
Specify the node correctly. Then, execute again.
- 6608 Operation disabled because the resource information of the specified resource is being updated. (detail:*code1-code2*)
Corrective action
Re-execute the processing.
code1, code2 indicates information required for error investigation.
- 6611 The specified resource has already been registered. (detail:*code1-code2*)
Corrective action
If this message appears when the resource is registered, it indicates that the specified resource has been already registered. There is no need to register it again.
If this message appears when changing a display name, specify a display name that is not available because the specified display name has already been registered.
code1, code2 indicates information required for error investigation.

- 6614 Cluster configuration management facility: internal error. (detail:*code1-code2*)
Corrective action
Record this message, and contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
code1, code2 indicates information required for error investigation.
- 6615 The cluster configuration management facility is not running. (detail:*code1-code2*)
Corrective action
Reactivate the Resource Database by restarting the node. If the message is redisplayed, record this message and collect related information for investigation. Then, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
code1, code2 indicates information required for error investigation.
- 6616 Cluster configuration management facility: error in the communication routine. (detail:*code1-code2*)
Corrective action
Record this message, and contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
code1, code2 indicates information required for error investigation.
- 6653 Operation cannot be performed on the specified resource.
Corrective action
userApplication in which the specified resource is registered is not in the Deact state. You need to bring this UserApplication Deact.
- 6661 Cluster control is not running. (detail:*code*)
Corrective action
Confirm that the Resource Database is running by executing the `clgettree(1)` command. If not, reboot the node. If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
code indicates information required for troubleshooting.

- 6665 The directory was specified incorrectly.
Corrective action
Specify the correct directory.
- 6668 Cannot run this command in single-user mode.
Corrective action
Boot the node in multi-user mode.
- 6675 Cannot run this command because *product_name* has already been set up.
Corrective action
Cancel the setting of the Resource Database *product_name*. Refer to appropriate manual for *product_name*.
- 6680 The specified directory does not exist.
Corrective action
Specify the existent directory.
- 6900 Automatic resource registration processing terminated abnormally. (detail: *reason*)
Corrective action
There might be incorrect settings in the shared disk definition file that was specified by the `-f` option of the `clautoconfig(1M)` command. Check the following. For details about the shared disk definition file, refer to the “Register shared disk units” of “PRIMECLUSTER *Global Disk Services Configuration and Administration Guide*.”
- The resource key name, the device name, and the node identifier name are specified in each line.
 - The resource key name begins with `shd`.
 - The device name begins with `/dev/`.
 - The node that has the specified node identifier name exists. You can check by executing the `clgettree(1)` command.
- Modify the shared disk definition file if necessary, and then execute the `clautoconfig(1M)` command.
- If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
- reason* indicates the command that was abnormally terminated or the returned value.

- 6901 Automatic resource registration processing is aborted due to one or more of the stopping nodes in the cluster domain.
Corrective action
Start all nodes and perform automatic resource registration.
- 6902 Automatic resource registration processing is aborted due to cluster domain configuration manager not running.
Corrective action
Cancel the automatic resource registration processing since the configuration of Resource Database is not working. Take down this message and collect the information needed for an investigation. Then, contact your local customer support (refer to the Section “Collecting troubleshooting information”). Failures may be recovered by restarting all nodes after collecting investigation information.
- 6903 Failed to create logical path. (*node dev1 dev2*)
Corrective action
Contact your local customer support to confirm that a logical path can be created in the share disk unit.
If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
node indicates an identification name of the node where the logical path failed to be created. *dev1* indicates the logical path (mp1b2048), and *dev2* indicates a tangible path (c1t0d0 and c2t0d0) corresponding to the logical path.
- 6904 Fail to register resource. (detail: *reason*)
Corrective action
Failed to register resource during the automatic registration processing. This might happen when the disk resource and system resource are not properly set up. Check the system setting of kernel parameter, disk size, etc.
If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
reason indicates the reason why a direction was invalidated.

6905 Automatic resource registration processing is aborted due to mismatch instance number of logical device between nodes.

Corrective action

This message appears when the logical path of the multi-path disk is created before registering the automatic resource.

If this message appears during registering the automatic resource after adding on disks and nodes, the registration command might fail to access the logical path of the multi-path disk and check the instance number. This happens in the following conditions:

- The same logical path name is created on multiple nodes
- This path cannot be accessed from all nodes

The PRIMECLUSTER automatic resource registration has a feature to provide a same environment to all applications. If the instance number (indicates 2048 of mp1b2048) of the logical path in the same disk is different between nodes, this message appears, and the automatic resource registration process is aborted. You need to check the logical path of all nodes. Recreate the logical path if necessary. The instance number should be the same. Then, register the automatic resource again.

If the cause is the failure of accessing the logical path of the multi-path disk, there might be a failure in the disk, or the disk is disconnected to the node.

Take the corrective action and register the automatic resource again. If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).

6906 Automatic resource registration processing is aborted due to mismatch setting of disk device path between nodes.

Corrective action

This failure might be due to one of the following incorrect settings:

- Among the nodes connected to the same shared disk, the package of the multi-path disk control is not installed on all nodes.
- The detection mode of the shared disk is different between nodes.
- The number of paths to the shared disk is different between nodes.

Take the corrective action and register the automatic resource again. If you still have this problem after going through the above instruction, contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).

6907 Automatic resource registration processing is aborted due to mismatch construction of disk device between nodes.

Corrective action

When the same shared disk was mistakenly connected to other cluster system, the volume label might have been overridden. Check the disk configuration. If there's no problem with the configuration, collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).

6910 It must be restart the specified node to execute automatic resource registration. (node: *node_name*...)

Corrective action

The nodes constituting the cluster system must be restarted. Restart the nodes constituting the cluster system. After that, perform the necessary resource registration again.

node_name indicates a node identifier for which a restart is necessary. If multiple nodes are displayed with *node_name*, these node identifiers are delimited with commas. If *node_name* is A11, restart all the nodes constituting the cluster system.

- 6911 It must be matched device number information in all nodes of the cluster system executing automatic resource registration. (dev: *dev_name...*)
Corrective action
Take down this message, and contact your local customer support. The support engineer will take care of matching transaction for the information on the disk device.
dev_name represents information for investigation.
- 7500 Cluster resource management facility: internal error. (function:*function* detail:*code1-code2*)
Corrective action
Record this message, and contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
contact your local customer support (refer to the Section “Collecting troubleshooting information”).
function, code1, code2 indicates information required for error investigation.
- 7501 Cluster resource management facility: insufficient memory. (function:*function* detail:*code1*)
Corrective action
Check the memory resource allocation estimate. For the memory required by Resource Database, refer to the PRIMECLUSTER *Installation Guide*. If this error cannot be corrected by this operator response, record this message, and contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
function, code1 indicates information required for error investigation.
- 7502 Cluster resource management facility: insufficient disk or system resources. (function:*function* detail:*code1*)
Corrective action
Referring to Section “Kernel parameters for Resource Database”, review the estimate of the disk resource and system resource (kernel parameter). If the kernel parameters have been changed, reboot the node for which the kernel parameters have been changed. If this error cannot be corrected by this operator response, record this message, and contact your local customer support. Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
function, code1 indicates information required for error investigation.

- 7503 The event cannot be notified because of an abnormal communication. (type:*type* rid:*rid* detail:*code1*)
Corrective action
Record this message, and contact your local customer support.
Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
After this event is generated, restart all the nodes within a cluster domain.
type,rid indicates event information and *code1* indicates information for investigation.
- 7504 The event notification is stopped because of an abnormal communication. (type:*type* rid:*rid* detail:*code1*)
Corrective action
Record this message, and contact your local customer support.
Collect information required for troubleshooting (refer to the Section “Collecting troubleshooting information”).
After this event is generated, restart all the nodes within a cluster domain.
type, rid indicates event information and *code1* indicates information for investigation.
- 7505 The node (*node*) is stopped because event cannot be notified by abnormal communication. (type:*type* rid:*rid* detail:*code1*)
Corrective action
Record this message and collect information for an investigation.
Then, contact your local customer support. Start the stopped node in a single user mode to collect investigation information (refer to the Section “Collecting troubleshooting information”).
node indicates the node identifier of the node to be stopped, *type, rid* the event information, and *code1* the information for investigation.
- 7506 The node (*node*) is forcibly stopped because event cannot be notified by abnormal communication. (type:*type* rid:*rid* detail:*code1*)
Corrective action
Record this message and collect information for an investigation.
Then, contact your local customer support. Start the forcibly stopped node in a single user mode to collect the investigation information (refer to the Section “Collecting troubleshooting information”).
node indicates the node identifier of the node to be stopped, *type, rid* the event information, and *code1* the information for investigation.

- 7507 Resource activation processing cannot be executed because of an abnormal communication. (resource:*resource* rid:*rid* detail:*code1*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support. For details about collecting investigation information (refer to the Section “Collecting troubleshooting information”).
After this phenomena occurs, restart the node to which the resource (*resource*) belongs. *resource* indicates the resource name for which activation processing was disabled, *rid* the resource ID, and *code1* the information for investigation.
- 7508 Resource (*resource1* resource ID:*rid1*, ...) activation processing is stopped because of an abnormal communication.
(resource:*resource2* rid:*rid2* detail:*code1*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support. For details about collecting investigation information (refer to the Section “Collecting troubleshooting information”).
After this phenomena occurs, restart the node to which the resource (*resource2*) belongs.
resource2 indicates the resource name for which activation processing was not performed, *rid2* the resource ID, *resource1* the resource name for which activation processing is not performed, *rid1* the resource ID, and *code1* the information for investigation.
- 7509 Resource deactivation processing cannot be executed because of an abnormal communication. (resource:*resource* rid:*rid* detail:*code1*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section “Collecting troubleshooting information”).
After this phenomena occurs, restart the node to which the resource (*resource*) belongs.
resource indicates the resource name for which deactivation processing was not performed, *rid* the resource ID, and *code1* the information for investigation.

- 7510 Resource (*resource1* resource ID:*rid1*, ...) deactivation processing is aborted because of an abnormal communication. (resource:*resource2* rid:*rid2* detail:*code1*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section "Collecting troubleshooting information").
After this phenomena occurs, restart the node to which the resource (*resource2*) belongs.
resource2 indicates the resource name for which deactivation processing was not performed, *rid2* the resource ID, *resource1* the resource name for which deactivation processing is not performed, *rid1* the resource ID, and *code1* the information for investigation.
- 7511 An error occurred by the event processing of the resource controller. (type:*type* rid:*rid* pclass:*pclass* prid:*prid* detail:*code1*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section "Collecting troubleshooting information").
After this phenomena occurs, restart the node in which the message was displayed.
type,rid indicates the event information, *pclass, prid* indicates resource controller information, and *code1* the information for investigation.
- 7512 The event notification is stopped because an error occurred in the resource controller. (type:*type* rid:*rid* pclass:*pclass* prid:*prid* detail:*code1*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section "Collecting troubleshooting information").
After this phenomena occurs, restart the node in which the message was displayed.
type, rid indicates the event information, *pclass, prid* indicates resource controller information, and *code1* the information for investigation.

- 7513 The node(*node*) is stopped because an error occurred in the resource controller. (type:*type* rid:*rid* pclass:*pclass* prid:*prid* detail:*code1*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section “Collecting troubleshooting information”).
Start up the stopped node in a single user mode to collect investigation information.
node indicates the node identifier of the node to be stopped, *type*,*rid* the event information, *pclass*, *prid* the resource controller information, and *code1* the information for investigation.
- 7514 The node (*node*) is forcibly stopped because an error occurred in the resource controller. (type:*type* rid:*rid* pclass:*pclass* prid:*prid* detail:*code1*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section “Collecting troubleshooting information”).
Start up the forcibly stopped node in a single user mode to collect investigation information.
node indicates the node identifier of the node to be forcibly stopped, *type*, *rid* the event information, *pclass*, *prid* the resource controller information, and *code1* the information for investigation.
- 7515 An error occurred by the resource activation processing (resource:*resource* rid:*rid* detail:*code1*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section “Collecting troubleshooting information”).
After this phenomena occurs, restart the node to which the resource (*resource*) belongs. An error occurs in the resource activation processing and activation of the resource (*resource*) cannot be performed.
resource indicates the resource name in which an error occurred in the activation processing, *rid* the resource ID, and *code1* the information for investigation.

- 7516 An error occurred by the resource deactivation processing. (resource:*resource* rid:*rid* detail:*code1*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section “Collecting troubleshooting information”).
After this phenomena occurs, restart the node to which the resource (*resource*) belongs. An error occurs in the resource deactivation processing and deactivation of the resource (*resource*) cannot be performed.
resource indicates the resource name in which an error occurred in the activation processing, *rid* the resource ID, and *code1* the information for investigation.
- 7517 Resource (*resource1* resource ID:*rid1*, ...) activation processing is stopped because an error occurred by the resource activation processing. (resource:*resource2* rid:*rid2* detail:*code1*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section “Collecting troubleshooting information”).
After this phenomena occurs, restart the node to which the resource (*resource2*) belongs.
Resource2 indicates the resource name in which an error occurred in the activation processing, *rid2* the resource ID, *resource1* the resource name in which activation processing is not performed, *rid1* the resource ID, and *code1* the information for investigation.

- 7518 Resource (*resource1* resource ID:*rid1*, ...) deactivation processing is aborted because an error occurred by the resource deactivation processing. (resource:*resource2* rid:*rid2* detail:*code1*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section “Collecting troubleshooting information”).
After this phenomena occurs, restart the node to which the resource (*resource2*) belongs.
resource2 indicates the resource name in which deactivation processing was disabled, *rid2* the resource ID, *resource1* the resource name in which deactivation processing is not performed, *rid1* the resource ID, and *code1* the information for investigation.
- 7519 Cluster resource management facility: error in exit processing. (node:*node* function:*function* detail:*code1*)
Corrective action
Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section “Collecting troubleshooting information”).
node indicates the node in which an error occurred and *function*, *code1* the information for investigation.
- 7520 The specified resource (resource ID:*rid*) does not exist or be not able to set the dependence relation.
Corrective action
Specify the correct resource, then re-execute the processing.
rid indicates a resource ID of the specified resource.
- 7521 The specified resource (class:*rclass* resource:*rname*) does not exist or be not able to set the dependence relation.
Corrective action
Specify the correct resource, then re-execute the processing.
rname indicates the specified resource name and *rclass* the class name.
- 7522 It is necessary to specify the resource which belongs to the same node.
Corrective action
The resource belonging to other node is specified. Specify a resource that belongs to the same node and re-execute it.

- 7535 An error occurred by the resource activation processing. The resource controller does not exist. (*resource* resource ID:*rid*)
Corrective action
As the resource controller is not available in the resource processing, resource (*resource*) activation was not performed. Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section “Collecting troubleshooting information”).
resource indicates the resource name for which activation processing was disabled, and *rid* a resource ID.
- 7536 An error occurred by the resource deactivation processing. The resource controller does not exist. (*resource* resource ID:*rid*)
Corrective action
As the resource controller is not available in the resource deactivation processing, resource (*resource*) deactivation was not performed. Record this message and collect information for an investigation. Then, contact your local customer support (refer to the Section “Collecting troubleshooting information”).
resource indicates the resource name for which deactivation processing could not be performed, and *rid* the resource ID.
- 7537 Command cannot be executed during resource activation processing.
Corrective action
After activation processing of the resource completes, re-execute it. Resource activation processing completion can be confirmed with 3204 message that is displayed on the console of the node to which the resource belongs.
- 7538 Command cannot be executed during resource deactivation processing.
Corrective action
After deactivation processing of the resource completes, re-execute it. Resource deactivation processing completion can be confirmed with 3206 message that is displayed on the console of the node to which the resource belongs.

- 7539 Resource activation processing timed out. (code:*code*
detail:*detail*)
Corrective action
Record this message and collect information for an investigation.
Then, contact your local customer support (refer to the Section
“Collecting troubleshooting information”).
- 7540 Resource deactivation processing timed out. (code:*code*
detail:*detail*)
Corrective action
Record this message and collect information for an investigation.
Then, contact your local customer support (refer to the Section
“Collecting troubleshooting information”).
- 7541 Setting related to dependence failed.
Corrective action
After confirming the specified resource, re-execute it.
- 7542 Resource activation processing cannot be executed
because node (*node*) is stopping.
Corrective action
As the node *node* to which the resource to be activated belongs is
stopped, the resource activation processing cannot be performed.
After starting up the node to which resource to be activated belongs,
re-execute it again.
node indicates the node identifier of the node where the connection is
broken.

- 7543 Resource deactivation processing cannot be executed because node (*node*) is stopping.
Corrective action
As the node *node* to which the resource to be deactivated belongs is stopped, the resource deactivation processing cannot be performed. After starting up the node to which resource to be deactivated belongs, re-execute it again.
node indicates the node identifier of the node where the connection is broken.
- 7545 Resource activation processing failed.
Corrective action
Refer to the measures in the error message displayed between activation processing start message (3203) and completion message (3204), which are displayed when this command is executed.
- 7546 Resource deactivation processing failed.
Corrective action
Refer to the measures in the error message displayed between deactivation processing start message (3205) and completion message (3206), which are displayed when this command is executed.

12.11 Shutdown Facility

- (SMAWsf, 10, 2) : %s of %s failed, errno %d
Cause: Internal problem.
Action: Check if there are related error messages following.
If yes, take action from there.
Otherwise, call support.
- (SMAWsf, 10, 3) : Unknown command from `sd_tool`, command %d
Cause: Using illegal `sdtool` command line.
Action: Choose the correct argument when `sdtool` is invoked.

- (SMAWsf, 10, 4) : Failed to open CLI response pipe for PID %d, errno %d
Cause: `rcsd` daemon could not open the pipe to response to `sdtool`.
Action: Call support.
- (SMAWsf, 10, 6) : Failed to create a signal handler for SIGCHLD
Cause: Internal problem.
Action: Call support.
- (SMAWsf, 10, 7) : The shutdown agent %s has exceeded its configured timeout, pid %d terminated
Cause: The shutdown agent does not return in 'timeout' seconds, which is configured in `rcsd.cfg`.
Action: If increasing timeout does not help, most likely shutdown agent does not work. check the shutdown agent log and call support.
- (SMAWsf, 10, 8) : A shutdown request has come in during a test cycle, test of %s pid %d terminated
Cause: `sdtool -k` was invoked while `rcsd` was running a shutdown agent testing.
Action: No harm. Just ignore it.
- (SMAWsf, 10, 9) : A request to reconfigure came in during a shutdown cycle, this request was ignored
Cause: When `rcsd` is eliminating a node, reconfiguration (`sdtool -r`) is not allowed.
Action: Try again after the node elimination is done.
- (SMAWsf, 10, 10) : Could not correctly read the `rcsd.cfg` file.
Cause: either `rcsd.cfg` file does not exist or the syntax in `rcsd.log` is not correct.
Action: Create `rcsd.cfg` file or fix the syntax.
- (SMAWsf, 10, 11) : %s in file %s around line %d
Cause: The syntax is not correct in `rcsd.log`
Action: fix the syntax.

- (SMAWsf, 10, 12) : A request to exit `ncsd` came in during a shutdown cycle, this request was ignored
- Cause: When `ncsd` is eliminating a node, bringing the `ncsd` daemon (`sdtool -e`) is not allowed.
- Action: Try again after the killing node elimination is done.
- (SMAWsf, 10, 15) : SA %s to %s host %s failed
- Cause: The shutdown agent failed to do
initialization | testing | shutdown |
un-initialization
the node
- Action: Check the shutdown agent log and call support.
- (SMAWsf, 10, 17) : Failed to open lock file
- Cause: internal problem
- Action: Call support.
- (SMAWsf, 10, 19) : Failed to unlink/create/open CLI Pipe
- Cause: internal problem
- Action: Call support.
- (SMAWsf, 10, 20) : Illegal catalog open parameter
- Cause: internal problem.
- Action: Call support.
- (SMAWsf 10, 30) : Pthread failed: %s : errcode %d %s
- Cause: Internal problem. POSIX thread failed.
- Action: Call support.
- (SMAWsf, 10, 31) : Pthread failed: %s : errcode %d %s
- Cause: Internal problem. `ncsd` was restarted.
- Action: Call support.
- (SMAWsf, 10, 34) : Host %S MA exec: %s failed, errno %d
- Cause: Failed to execute monitor agent's API.
- Action: Call support.

- (SMAWsf, 10, 36) : Failed to cancel %s, thread f %s %s of host %s
Cause: POSIX thread was not cancellable.
Action: Call support.
- (SMAWsf, 10, 38) : Host %s, MA %s, MAH//get/state() failed
Cause: failed to call monitor agent's API
MAHostGetState ().
Action: Call support.
- (SMAWsf, 10, 101) : Malloc failed during %s
Cause: Not enough memory.
Action: Increase virtual memory size (ulimit -v) or
increase system memory. Call support if the
problem still exists.
- (SMAWsf, 30, 2) : Usage: sdtool {-d[on | off] | -s | -S | -r |
-b | -c | -e | -k node-name | -w weight-
factor | -n node-factor }
Cause: Illegal argument/command line usage.
Action: Use the correct argument
- (SMAWsf, 30, 3) : unlink failed on RCDSD response pipe %s, errno %d
Cause: Can not remove the old pipe file.
Action: Call support.
- (SMAWsf, 30, 4) : mkfifo failed on RCDSD response pipe %s, errno %d
Cause: Could not create the pipe for rcsd.
Action: Call support.
- (SMAWsf, 30, 5) : open failed on RCDSD response pipe %s, errno %d
Cause: Could not open the pipe for rcsd.
Action: Call support.
- (SMAWsf, 30, 6) : open failed on rcsdin pipe %s, errno %d
Cause: Could not open communication pipe from sdtool
to rcsd.
Action: Call support.

- (SMAWsf, 30, 7) : write failed on rcsdin pipe %s, errno %d
Cause: Could not pass command from sdttool to rcsd.
Action: Call support.
- (SMAWsf, 30, 8) : select failed, errno %d
Cause: sdttool could not get information from rcsd.
Action: Call support.
- (SMAWsf, 30, 9) : read failed, errno %d
Cause: sdttool failed to read data from rcsd daemon.
Action: Call support.
- (SMAWsf, 30, 10) : RCSD returned an error for this command, error is %d
Cause: rcsd failed to execute the command from sdttool.
Action: Check if there are related error messages following.
If yes, take action from there.
Otherwise, call support.
- (SMAWsf, 30, 12) : A shutdown is in progress for the machine %s, try again later
Cause: rcsd daemon is currently eliminating the machine. The current request is not ignored.
Action: Try again later.
- (SMAWsf, 30, 13) : The RCSD is not running
Cause: The command failed because rcsd daemon is not running.
Action: Start up rcsd daemon (sdttool -b) then try the command again.
- (SMAWsf, 30, 14) : RCSD is exiting. Command is not allowed
Cause: rcsd daemon is in the stage of shutting down. The command is not allowed.
Action: Try the command after rcsd daemon is started up.

- (SMAWsf, 30, 15) : Failed to get %s product information
Cause: Most likely the product is not installed properly.
Action: Reinstall the product.
- (SMAWsf, 30, 16) : Illegal catlog open parameter
Cause: Failed to open log file.
Action: Call support.
- (SMAWsf, 30, 17) : Could not execlp (RCSD). Errno = %d
Cause: Most likely the rcsd binary does not exist.
Action: Reinstall the package.
- (SMAWsf, 50, 3) : The SF-CF initialization failed, status %d
Cause: Most likely, CF is not configured and/or is not loaded.
Action: Configure and load CF.
- (SMAWsf, 50, 4) : The SF-CF event processing failed, status %d
Cause: Internal problem.
Action: Call support.
- (SMAWsf, 50, 6) : The SF-CF has failed to locate host %s
Cause: The nodename in the `rcsd.cfg` is not a known CF name.
Action: Use the CF name (`cftool -n`) in `rcsd.cfg`.
- (SMAWsf, 50, 9) : The SF-CF failed to declare %s down, status %d
Cause: Internal problem.
Action: Call support.
- (SMAWsf, 50, 11) : Failed to open CFSF device, reason (%d)%s
Cause: Could not open CFSF device
Action: Call support.
- (SMAWsf, 50, 12) : `h_cfsf_get_leftcluster()` failed. reason (%d) %s
Cause: Failed to call `cfsf_get_leftcluster`.
Action: Call support.

- (SMAWsf, 50, 13) : Node id %d ICF communication failure detected
Cause: CF layer has detected lost heartbeat
Action: *ncsd* will take action.
- (SMAWsf, 50, 14) : Host %s ICF communications failure detected
Cause: *ncsd* was notified the node has lost heartbeat
Action: *ncsd* take action to eliminate the node.
- (SMAWsf, 50, 20) : Failed to cancel thread of the %s-monitor
Cause: Failed to cancel thread
Action: Call support.
- (SMAWsf, 50, 21) : Failed to do %s, reason (%d)%s
Cause: Failed to call some internal functions.
Action: Call support.
- (SMAWsf, 50, 22) : Failed to get nodeid for host %s. reason (%d)%s
Cause: Not able to get the cluster node id for the node.
Action: Call support.

12.12 Monitoring Agent messages

This section lists the messages output from the Monitoring Agents. The message format is as follows.



Italics indicate that the output varies, depending on the message.

FJSVcluster:severity:program:message-number:message

<i>severity</i>	Message severity level. The levels of severity are as follows: Information (INFORMATION), Warning (WARNING), Error (ERROR). For details, refer to the table below.
<i>program</i>	Name of the program that output this message. The monitoring agent is output as DEV.
<i>message-number</i>	Message number.
<i>details</i>	Detailed classification code.

Number	Severity	Meaning
2000-3999	Information	Message providing information about the monitoring agent state.
4000-5999	Warning	Message warning about an insignificant error that does not cause the abnormal termination of the monitoring agent.
????, 6000-7999	Error	Message indicating that a significant error has occurred that caused the abnormal termination of the monitoring agent.

12.12.1 INFO message

3040	The console monitoring agent has been started. (node: <i>nodename</i>)
3041	The console monitoring agent has been stopped. (node: <i>nodename</i>)
3042	The RCI monitoring agent has been started.
3043	The RCI monitoring agent has been stopped.
3044	The console monitoring agent took over monitoring Node <i>targetnode</i> .
3045	The console monitoring agent cancelled to monitor Node <i>targetnode</i>

12.12.2 WARNING message

5001 The RCI address has been changed. (node:*nodename*
 address:*address*)

Corrective action:

The RCI address is changed while the RCI monitoring agent is running. *nodename* indicates a name of the node where the RCI address is changed. *address* indicates the changed RCI address. Check if the RCI address is correctly set up on the node.

12.12.3 ERROR message



When the error messages described in this section are output, investigate the `/var/adm/messages` file and check if another error message is output before this message. If this occurs, follow the corrective action of the other error message.

???? Message not found!!

Corrective action:

The text of the message corresponding to the message number is not available. Copy this message and contact field support.

6000 An internal error occurred.
 (function:*function* detail:*code1-code2-code3-code4*)

Corrective action:

Collect required information to contact field support. Refer to the Chapter “Diagnostics and troubleshooting” for collecting information.

6004 No system administrator authority.

Corrective action:

Execute using system administrator access privileges.

7003 An error was detected in RCI.
 (node:*nodename* address:*address* status:*status*)

Corrective action:

There is an RCI transmission failure between *nodename* indicated in the message and a node where the message was output. RCI might not be properly connected, or there might be a system failure. Check if the RCI cable is connected. If the RCI cable is disconnected, execute the `/etc/opt/FJSVcluster/bin/clrcimonctl restart` command on the node where the error message was output, and restart the RCI monitoring agent. Then, execute the `/opt/SAW/bin/sdtool -r` command to restart the Shutdown Facility (SF).

If the RCI cable is connected, there might be a hardware failure such as the RCI cable or System Control Facility (SCF). Collect required information and SCF dump to contact field support.

Refer to the Chapter “Diagnostics and troubleshooting” for collecting information.

When the hardware failure is recovered, execute the `/etc/opt/FJSVcluster/bin/clrcimonctl restart` command to restart the RCI monitoring agent and the `/opt/SAW/bin/sdtool -r` command to restart SF.

```
7004          The RCI monitoring agent has been stopped due to an
              RCI address error.
              (node:nodename address:address)
```

Corrective action:

The RCI address of other node is changed while the RCI monitoring agent is running. Collect required information and SCF dump to contact field support.

Refer to the Chapter “Diagnostics and troubleshooting” for collecting information and on SCF dump.

The field support engineer confirms if the RCI address of *nodename* indicated in the message is correctly set up. To check the previous RCI address, execute the `/opt/FJSVmadm/sbin/setrci stat` command on an arbitrary node.

If the RCI address is incorrect, set up the address again referring to the instruction for field support engineers.

Execute the `/etc/opt/FJSVcluster/bin/clrcimonctl restart` command to restart the RCI monitoring agent and the `/opt/SAW/bin/sdtool -r` command to restart the Shutdown Facility (SF) where the error message was output.

```
7040          The console was disconnected.
              (node:nodename portno:portnumber detail:code)
```

Corrective action:

The RCCU is disconnected. Check the following:

- The RCCU is powered.
- The normal lamp of the HUB connected to the RCCU is on.
- The LAN cable connectors are connected to the RCCU and HUB.

If any of above fails, execute the `/opt/SMAW/bin/sdtool -r` command on the node where the error message was output, and restart the Shutdown Facility.

If you still have a problem with connection, there might be a network failure or a failure of hardware such as RCCU, HUB and related cables. Contact field support.

If the above corrective action does not work, collect required information to contact field support. Refer to the Chapter “Diagnostics and troubleshooting” for collecting information.

```
7042          Connection to the console is refused.  
              (node:nodename portno:portnumber detail:code)
```

Corrective action:

Connection to the console cannot be established during the console monitoring agent startup. Check the followings:

- The RCCU is powered.
- The normal lamp of the HUB connected to the RCCU is on.
- The LAN cable connectors are connected to the RCCU and HUB.

If any of above fails, execute the `/opt/SMAW/bin/sdtool -r` command on the node where the error message was output, and restart the Shutdown Facility.

If you still have a problem with connection, there might be a network failure or a failure of hardware such as RCCU, HUB and related cables. Contact field support.

If the above corrective action does not work, collect required information to contact field support. Refer to the Chapter “Diagnostics and troubleshooting” for collecting information.

```
7200          The configuration file of the console monitoring  
              agent does not exist. (file:filename)
```

Corrective action:

1. Download the configuration file displayed in miscellaneous information using ftp from other nodes.
2. Store this file in the original directory.
3. Set up the same access permission mode of this file as other nodes.
4. Restart the system.

If all the nodes do not have this configuration file, collect required information to contact field support. Refer to the Chapter “Diagnostics and troubleshooting” for collecting information.

7201 The configuration file of the RCI monitoring agent does not exist. (file:*filename*)

Corrective action:

1. Download the configuration file displayed in miscellaneous information using ftp from other nodes.
2. Store this file in the original directory.
3. Set up the same access permission mode of this file as other nodes.
4. Restart the system.

If all the nodes do not have this configuration file, collect required information to contact field support. Refer to the Chapter “Diagnostics and troubleshooting” for collecting information.

7202 The configuration file of the console monitoring agent has an incorrect format. (file:*filename*)

Corrective action:

There's an incorrect format of the configuration file in the console monitoring agent.

If the configuration file name displayed in miscellaneous information is SA_rccu.cfg, reconfigure the Shutdown Facility by invoking the configuration wizard. Then, confirm if the RCCU name is correct.

If the above corrective action does not work, or the configuration file name is other than SA_rccu.cfg, collect required information to contact field support. Refer to the Chapter “Diagnostics and troubleshooting” for collecting information.

13 Manual pages

This chapter lists the online manual pages for CCBR, CF, CFS, CIP, Monitoring Agent, PAS, RCVM, Resource Database, RMS, RMS Wizards, SCON, SF, SIS, and Web-Based Admin View.

To display a manual page, type the following command:

```
$ man man_page_name
```

13.1 CCBR

System administration

`cfbackup`

save the cluster configuration information for a PRIMECLUSTER node

`cfrestore`

restore saved cluster configuration formation on a PRIMECLUSTER node

13.2 CF

System administration

`cfconfig`

configure or unconfigure a node for a PRIMECLUSTER cluster

`cfset`

apply or modify `/etc/default/cluster.config` entries into the CF module

`cftool`

print node communications status for a node or the cluster

13.3 CFS

- fsck_rcfs
file system consistency check and interactive repair
- mount_rcfs
mount RCFS file systems
- rcfs_fumount
force unmount RCFS mounted file system
- rcfs_list
list status of RCFS mounted file systems
- rcfs_switch
manual switchover or failover of a RCFS file system
- ngadmin
node group administration utility
- cfsmntd
cfs mount daemon for RCFS

13.4 CIP

System administration

- cipconfig
start or stop CIP 2.0
- ciptool
retrieve CIP information about local and remote nodes in the cluster

File format

- cip.cf
CIP configuration file format

13.5 Monitoring Agent

System administration

`clrcimonctl`

Start, stop or restart of the RCI monitoring agent daemon, and display of daemon presence

`clrcumonctl`

Start, stop or restart of the console monitoring agent daemon, and display of daemon presence

13.6 PAS

System administration

`mipcstat`

MIPC statistics

`clmstat`

CLM statistics

13.7 RCVM



Applies to transitioning users of existing Fujitsu Siemens products only.

System administration

`dkconfig`

virtual disk configuration utility

`dkmigrate`

virtual disk migration utility

`vdisk`

virtual disk driver

`dkmirror`

mirror disk administrative utility

File format

dktab
virtual disk configuration file

13.8 Resource Database



To display a Resource Database manual page, add `/etc/opt/FJSVcluster/man` to the environment variable `MANPATH`.

System administration

`clautoconfig`
execute of the automatic resource registration

`clbackuprdb`
save the resource database

`clexec`
execute the remote command

`cldeldevice`
delete resource registered by automatic resource registration

`clinitreset`
reset the resource database

`clrestorerdb`
restore the resource database

`clsetparam`
display and change the resource database operational environment

`clsetup`
set up the resource database

`clstarttrsc`
resource activation

`clstoptrsc`
resource deactivation

`clsyncfile`
distribute a file between cluster nodes

User command

`clgettree`
display the tree information of the resource database

13.9 RMS

System administration

`hvassert`
assert (test for) an RMS resource state

`hvcn`
start the RMS configuration monitor

`hvconfig`
display or save the RMS configuration file

`hvdisp`
display RMS resource information

`hvdist`
distribute RMS configuration files

`hvdump`
collect debugging information about RMS

`hvgdmake`
compile an RMS custom detector

`hvlogclean`
clean RMS log files

`hvrclev`
change default RMS start run level

`hvshut`
shut down RMS

`hvswitch`
switch control of an RMS user application resource to another node

`hvthrottle`
prevent multiple RMS scripts from running simultaneously

`hvutil`
manipulate availability of an RMS resource

File formats

hvenv.local
RMS local environment configuration file

13.10 RMS Wizards

RMS Wizards and RMS Application Wizards

RMS Wizards are documented as html pages in the SMAWRhvd0 package on the CD-ROM. After installing this package, the documentation is available in the following directory:
/usr/opt/reliant/htdocs.solaris/wizards.en

13.11 SCON

scon
start the cluster console software

13.12 SF

System administration

rcsd
Shutdown Daemon of the Shutdown Facility

rcsd.cfg
configuration file for the Shutdown Daemon

SA_rccu.cfg
configuration file for RCCU Shutdown Agent

SA_rps.cfg
configuration file for a Remote Power Switch Shutdown Agent

SA_scon.cfg
configuration file for SCON Shutdown Agent

SA_pprci.cfg
configuration file for RCI Shutdown Agent (PRIMEPOWER only)

- SA_sspint.cfg
configuration file for Sun E10000 Shutdown Agent
- SA_sunF.cfg
configuration file for sunF system controller Shutdown Agent
- SA_wtinps.cfg
configuration file for WTI NPS Shutdown Agent
- sdtool
interface tool for the Shutdown Daemon

13.13 SIS

System administration

- dtcpadmin
start the SIS administration utility
- dtcpd
start the SIS daemon for configuring VIPs
- dtcpstat
status information about SIS

13.14 Web-Based Admin View

System administration

- fjsvwvbs
stop Web-Based Admin View
- fjsvwvcnf
start, stop, or restart the web server for Web-Based Admin View
- wvCntl
start, stop, or get debugging information for Web-Based Admin View
- wvGetparam
display Web-Based Admin View's environment variable

wvSetparam

set Web-Based Admin View environment variable

wvstat

display the operating status of Web-Based Admin View

Glossary

AC

See *Access Client*.

Access Client

GFS kernel module on each node that communicates with the Meta Data Server and provides simultaneous access to a shared file system.

Administrative LAN

In PRIMECLUSTER configurations, an Administrative LAN is a private local area network (LAN) on which machines such as the System Console and Cluster Console reside. Because normal users do not have access to the Administrative LAN, it provides an extra level of security. The use of an Administrative LAN is optional.

See also *public LAN*.

API

See *Application Program Interface*.

application (RMS)

A resource categorized as a `userApplication` used to group resources into a logical collection.

Application Program Interface

A shared boundary between a service provider and the application that uses that service.

application template (RMS)

A predefined group of object definition value choices used by RMS Application Wizards to create object definitions for a specific type of application.

Application Wizards

See *RMS Application Wizards*.

attribute (RMS)

The part of an object definition that specifies how the base monitor acts and reacts for a particular object type during normal operations.

automatic switchover (RMS)

The procedure by which RMS automatically switches control of a userApplication over to another node after specified conditions are detected.

See also *directed switchover (RMS)*, *failover (RMS, SIS)*, *switchover (RMS)*, *symmetrical switchover (RMS)*.

availability

Availability describes the need of most enterprises to operate applications via the Internet 24 hours a day, 7 days a week. The relationship of the actual to the planned usage time determines the availability of a system.

base cluster foundation (CF)

This PRIMECLUSTER module resides on top of the basic OS and provides internal interfaces for the CF (Cluster Foundation) functions that the PRIMECLUSTER services use in the layer above.

See also *Cluster Foundation*.

base monitor (RMS)

The RMS module that maintains the availability of resources. The base monitor is supported by daemons and detectors. Each node being monitored has its own copy of the base monitor.

Cache Fusion

The improved interprocess communication interface in Oracle 9i that allows logical disk blocks (buffers) to be cached in the local memory of each node. Thus, instead of having to flush a block to disk when an update is required, the block can be copied to another node by passing a message on the interconnect, thereby removing the physical I/O overhead.

CCBR

See *Cluster Configuration Backup and Restore*.

Cluster Configuration Backup and Restore

CCBR provides a simple method to save the current PRIMECLUSTER configuration information of a cluster node. It also provides a method to restore the configuration information.

CF

See *Cluster Foundation*.

child (RMS)

A resource defined in the configuration file that has at least one parent. A child can have multiple parents, and can either have children itself (making it also a parent) or no children (making it a leaf object).

See also *resource (RMS)*, *object (RMS)*, *parent (RMS)*.

cluster

A set of computers that work together as a single computing source. Specifically, a cluster performs a distributed form of parallel computing.

See also *RMS configuration*.

Cluster Foundation

The set of PRIMECLUSTER modules that provides basic clustering communication services.

See also *base cluster foundation (CF)*.

cluster interconnect (CF)

The set of private network connections used exclusively for PRIMECLUSTER communications.

Cluster Join Services (CF)

This PRIMECLUSTER module handles the forming of a new cluster and the addition of nodes.

concatenated virtual disk

Concatenated virtual disks consist of two or more pieces on one or more disk drives. They correspond to the sum of their parts. Unlike simple virtual disks where the disk is subdivided into small pieces, the individual disks or partitions are combined to form a single large logical disk. (Applies to transitioning users of existing Fujitsu Siemens products only.)

See also *mirror virtual disk*, *simple virtual disk*, *striped virtual disk*, *virtual disk*.

configuration file (RMS)

The RMS configuration file that defines the monitored resources and establishes the interdependencies between them. The default name of this file is `config.us`.

Glossary

console

See *single console*.

custom detector (RMS)

See *detector (RMS)*.

custom type (RMS)

See *generic type (RMS)*.

daemon

A continuous process that performs a specific function repeatedly.

database node (SIS)

Nodes that maintain the configuration, dynamic data, and statistics in a SIS configuration.

See also *gateway node (SIS)*, *service node (SIS)*, *Scalable Internet Services (SIS)*.

detector (RMS)

A process that monitors the state of a specific object type and reports a change in the resource state to the base monitor.

directed switchover (RMS)

The RMS procedure by which an administrator switches control of a userApplication over to another node.

See also *automatic switchover (RMS)*, *failover (RMS, SIS)*, *switchover (RMS)*, *symmetrical switchover (RMS)*.

DOWN (CF)

A node state that indicates that the node is unavailable (marked as down). A LEFTCLUSTER node must be marked as DOWN before it can rejoin a cluster.

See also *UP (CF)*, *LEFTCLUSTER (CF)*, *node state (CF)*.

ENS (CF)

See *Event Notification Services (CF)*.

environment variables (RMS)

Variables or parameters that are defined globally.

error detection (RMS)

The process of detecting an error. For RMS, this includes initiating a log entry, sending a message to a log file, or making an appropriate recovery response.

Event Notification Services (CF)

This PRIMECLUSTER module provides an atomic-broadcast facility for events.

failover (RMS, SIS)

With SIS, this process switches a failed node to a backup node. With RMS, this process is known as switchover.

See also *automatic switchover (RMS)*, *directed switchover (RMS)*, *switchover (RMS)*, *symmetrical switchover (RMS)*.

gateway node (SIS)

Gateway nodes have an external network interface. All incoming packets are received by this node and forwarded to the selected service node, depending on the scheduling algorithm for the service.

See also *service node (SIS)*, *database node (SIS)*, *Scalable Internet Services (SIS)*.

GDS

See *Global Disk Services*.

GFS

See *Global File Services*.

GLS

See *Global Link Services*.

Global Disk Services

This optional product provides volume management that improves the availability and manageability of information stored on the disk unit of the Storage Area Network (SAN).

Global File Services

This optional product provides direct, simultaneous accessing of the file system on the shared storage unit from two or more nodes within a cluster.

Global Link Services

This PRIMECLUSTER optional module provides network high availability solutions by multiplying a network route.

generic type (RMS)

An object type which has generic properties. A generic type is used to customize RMS for monitoring resources that cannot be assigned to one of the supplied object types.

See also *object type (RMS)*.

graph (RMS)

See *system graph (RMS)*.

graphical user interface

A computer interface with windows, icons, toolbars, and pull-down menus that is designed to be simpler to use than the command-line interface.

GUI

See *graphical user interface*.

high availability

This concept applies to the use of redundant resources to avoid single points of failure.

interconnect (CF)

See *cluster interconnect (CF)*.

Internet Protocol address

A numeric address that can be assigned to computers or applications.

See also *IP aliasing*.

Internode Communications facility

This module is the network transport layer for all PRIMECLUSTER internode communications. It interfaces by means of OS-dependent code to the network I/O subsystem and guarantees delivery of messages queued for transmission to the destination node in the same sequential order unless the destination node fails.

IP address

See *Internet Protocol address*.

IP aliasing

This enables several IP addresses (aliases) to be allocated to one physical network interface. With IP aliasing, the user can continue communicating with the same IP address, even though the application is now running on another node.

See also *Internet Protocol address*.

JOIN (CF)

See *Cluster Join Services (CF)*.

keyword

A word that has special meaning in a programming language. For example, in the configuration file, the keyword `object` identifies the kind of definition that follows.

leaf object (RMS)

A bottom object in a system graph. In the configuration file, this object definition is at the beginning of the file. A leaf object does not have children.

LEFTCLUSTER (CF)

A node state that indicates that the node cannot communicate with other nodes in the cluster. That is, the node has left the cluster. The reason for the intermediate `LEFTCLUSTER` state is to avoid the network partition problem.

See also *UP (CF)*, *DOWN (CF)*, *network partition (CF)*, *node state (CF)*.

link (RMS)

Designates a child or parent relationship between specific resources.

local area network

See *public LAN*.

local node

The node from which a command or process is initiated.

See also *remote node*, *node*.

log file

The file that contains a record of significant system events or messages. The base monitor, wizards, and detectors can have their own log files.

MDS

See *Meta Data Server*.

message

A set of data transmitted from one software process to another process, device, or file.

message queue

A designated memory area which acts as a holding place for messages.

Meta Data Server

GFS daemon that centrally manages the control information of a file system (meta-data).

mirrored disks

A set of disks that contain the same data. If one disk fails, the remaining disks of the set are still available, preventing an interruption in data availability. (Applies to transitioning users of existing Fujitsu Siemens products only.)

See also *mirrored pieces*.

mirrored pieces

Physical pieces that together comprise a mirrored virtual disk. These pieces include mirrored disks and data disks. (Applies to transitioning users of existing Fujitsu Siemens products only.)

See also *mirrored disks*.

mirror virtual disk

Mirror virtual disks consist of two or more physical devices, and all output operations are performed simultaneously on all of the devices. (Applies to transitioning users of existing Fujitsu Siemens products only.)

See also *concatenated virtual disk*, *simple virtual disk*, *striped virtual disk*, *virtual disk*.

mount point

The point in the directory tree where a file system is attached.

multihosting

Multiple controllers simultaneously accessing a set of disk drives. (Applies to transitioning users of existing Fujitsu Siemens products only.)

native operating system

The part of an operating system that is always active and translates system calls into activities.

network partition (CF)

This condition exists when two or more nodes in a cluster cannot communicate over the interconnect; however, with applications still running, the nodes can continue to read and write to a shared device, compromising data integrity.

node

A host which is a member of a cluster. A computer node is the same as a computer.

node state (CF)

Every node in a cluster maintains a local state for every other node in that cluster. The node state of every node in the cluster must be either UP, DOWN, or LEFTCLUSTER.

See also *UP (CF)*, *DOWN (CF)*, *LEFTCLUSTER (CF)*.

object (RMS)

In the configuration file or a system graph, this is a representation of a physical or virtual resource.

See also *leaf object (RMS)*, *object definition (RMS)*, *object type (RMS)*.

object definition (RMS)

An entry in the configuration file that identifies a resource to be monitored by RMS. Attributes included in the definition specify properties of the corresponding resource. The keyword associated with an object definition is `object`.

See also *attribute (RMS)*, *object type (RMS)*.

object type (RMS)

A category of similar resources monitored as a group, such as disk drives. Each object type has specific properties, or attributes, which limit or define what monitoring or action can occur. When a resource is associated with a particular object type, attributes associated with that object type are applied to the resource.

See also *generic type (RMS)*.

online maintenance

The capability of adding, removing, replacing, or recovering devices without shutting or powering off the node.

operating system dependent (CF)

This module provides an interface between the native operating system and the abstract, OS-independent interface that all PRIMECLUSTER modules depend upon.

OPS

See *Oracle Parallel Server*.

Oracle Parallel Server

Oracle Parallel Server allows access to all data in a database to users and applications in a clustered or MPP (massively parallel processing) platform.

OSD (CF)

See *operating system dependent (CF)*.

parent (RMS)

An object in the configuration file or system graph that has at least one child.

See also *child (RMS)*, *configuration file (RMS)*, *system graph (RMS)*.

primary node (RMS)

The default node on which a user application comes online when RMS is started. This is always the nodename of the first child listed in the `userApplication` object definition.

private network addresses

Private network addresses are a reserved range of IP addresses specified by the Internet Assigned Numbers Authority. They may be used internally by any organization but, because different organizations can use the same addresses, they should never be made visible to the public internet.

private resource (RMS)

A resource accessible only by a single node and not accessible to other RMS nodes.

See also *resource (RMS)*, *shared resource*.

queue

See *message queue*.

PRIMECLUSTER services (CF)

Service modules that provide services and internal interfaces for clustered applications.

redundancy

This is the capability of one object to assume the resource load of any other object in a cluster, and the capability of RAID hardware and/or RAID software to replicate data stored on secondary storage devices.

public LAN

The local area network (LAN) by which normal users access a machine.

See also *Administrative LAN*.

Reliant Monitor Services (RMS)

The package that maintains high availability of user-specified resources by providing monitoring and switchover capabilities.

remote node

A node that is accessed through a telecommunications line or LAN.

See also *local node*.

remote node

See *remote node*.

reporting message (RMS)

A message that a detector uses to report the state of a particular resource to the base monitor.

resource (RMS)

A hardware or software element (private or shared) that provides a function, such as a mirrored disk, mirrored disk pieces, or a database server. A local resource is monitored only by the local node.

See also *private resource (RMS)*, *shared resource*.

resource definition (RMS)

See *object definition (RMS)*.

resource label (RMS)

The name of the resource as displayed in a system graph.

resource state (RMS)

Current state of a resource.

RMS

See *Reliant Monitor Services (RMS)*.

RMS Application Wizards

RMS Application Wizards add new menu items to the RMS Wizard Tools for a specific application.

See also *RMS Wizard Tools, Reliant Monitor Services (RMS)*.

RMS commands

Commands that enable RMS resources to be administered from the command line.

RMS configuration

A configuration made up of two or more nodes connected to shared resources. Each node has its own copy of operating system and RMS software, as well as its own applications.

RMS Wizard Tools

A software package composed of various configuration and administration tools used to create and manage applications in an RMS configuration.

See also *RMS Application Wizards, Reliant Monitor Services (RMS)*.

SAN

See *Storage Area Network*.

Scalable Internet Services (SIS)

Scalable Internet Services is a TCP connection load balancer, and dynamically balances network access loads across cluster nodes while maintaining normal client/server sessions for each connection.

scalability

The ability of a computing system to dynamically handle any increase in work load. Scalability is especially important for Internet-based applications where growth caused by Internet usage presents a scalable challenge.

SCON

See *single console*.

script (RMS)

A shell program executed by the base monitor in response to a state transition in a resource. The script may cause the state of a resource to change.

service node (SIS)

Service nodes provide one or more TCP services (such as FTP, Telnet, and HTTP) and receive client requests forwarded by the gateway nodes.

See also *database node (SIS)*, *gateway node (SIS)*, *Scalable Internet Services (SIS)*.

shared resource

A resource, such as a disk drive, that is accessible to more than one node.

See also *private resource (RMS)*, *resource (RMS)*.

simple virtual disk

Simple virtual disks define either an area within a physical disk partition or an entire partition. (Applies to transitioning users of existing Fujitsu Siemens products only.)

See also *concatenated virtual disk*, *striped virtual disk*, *virtual disk*.

single console

The workstation that acts as the single point of administration for nodes being monitored by RMS. The single console software, SCON, is run from the single console.

SIS

See *Scalable Internet Services (SIS)*.

state

See *resource state (RMS)*.

Storage Area Network

The high-speed network that connects multiple, external storage units and storage units with multiple computers. The connections are generally fiber channels.

striped virtual disk

Striped virtual disks consist of two or more pieces. These can be physical partitions or further virtual disks (typically a mirror disk). Sequential I/O operations on the virtual disk can be converted to I/O operations on two or more physical disks. This corresponds to RAID Level 0 (RAID0). (Applies to transitioning users of existing Fujitsu Siemens products only.)

See also *concatenated virtual disk*, *mirror virtual disk*, *simple virtual disk*, *virtual disk*.

switchover (RMS)

The process by which RMS switches control of a userApplication over from one monitored node to another.

See also *automatic switchover (RMS)*, *directed switchover (RMS)*, *failover (RMS, SIS)*, *symmetrical switchover (RMS)*.

symmetrical switchover (RMS)

This means that every RMS node is able to take on resources from any other RMS node.

See also *automatic switchover (RMS)*, *directed switchover (RMS)*, *failover (RMS, SIS)*, *switchover (RMS)*.

system graph (RMS)

A visual representation (a map) of monitored resources used to develop or interpret the configuration file.

See also *configuration file (RMS)*.

template

See *application template (RMS)*.

type

See *object type (RMS)*.

UP (CF)

A node state that indicates that the node can communicate with other nodes in the cluster.

See also *DOWN (CF)*, *LEFTCLUSTER (CF)*, *node state (CF)*.

virtual disk

With virtual disks, a pseudo device driver is inserted between the highest level of the Solaris logical Input/Output (I/O) system and the physical device driver. This pseudo device driver then maps all logical I/O requests on physical disks. (Applies to transitioning users of existing Fujitsu Siemens products only.)

See also *concatenated virtual disk*, *mirror virtual disk*, *simple virtual disk*, *striped virtual disk*.

Web-Based Admin View

This is a common base to utilize the Graphic User Interface of PRIMECLUSTER. This interface is in Java.

wizard (RMS)

An interactive software tool that creates a specific type of application using pretested object definitions. An enabler is a type of wizard.

Abbreviations

AC

Access Client

API

application program interface

bm

base monitor

CCBR

Cluster Configuration Backup/Restore

CF

Cluster Foundation

CIM

Cluster Integrity Monitor

CIP

Cluster Interconnect Protocol

CLI

command-line interface

CRM

Cluster Resource Management

DLPI

Data Link Provider Interface

ENS

Event Notification Services

GDS

Global Disk Services

GFS

Global File Services

Abbreviations

GLS	Global Link Services
GUI	graphical user interface
HA	high availability
ICF	Internode Communication Facility
I/O	input/output
JOIN	cluster join services module
LAN	local area network
MDS	Meta Data Server
MIB	Management Information Base
NIC	network interface card
NSM	Node State Monitor
OPS	Oracle Parallel Server
OSD	operating system dependant
PAS	Parallel Application Services

PRIMECLUSTER SF

PRIMECLUSTER Shutdown Facility

RCI

Remote Cabinet Interface

RMS

Reliant Monitor Services

SA

Shutdown Agent

SAN

Storage Area Network

SCON

single console software

SD

Shutdown Daemon

SF

Shutdown Facility

SIS

Scalable Internet Services

VIP

Virtual Interface Provider

Figures

Figure 1:	CIP diagram	11
Figure 2:	CF over IP diagram	12
Figure 3:	Login pop-up	17
Figure 4:	Main Web-Based Admin View screen after login	18
Figure 5:	Global Cluster Services screen in Web-Based Admin View	19
Figure 6:	Initial connection pop-up	19
Figure 7:	CF is unconfigured and unloaded	20
Figure 8:	CF loaded but not configured	21
Figure 9:	Scanning for clusters	22
Figure 10:	Creating or joining a cluster	23
Figure 11:	Selecting cluster nodes and the cluster name	24
Figure 12:	Edit CF node names	26
Figure 13:	CF loads and pings	27
Figure 14:	CF topology and connection table	28
Figure 15:	CF over IP screen	30
Figure 16:	CIP Wizard screen	31
Figure 17:	CIM configuration screen	33
Figure 18:	Summary screen	35
Figure 19:	Configuration processing screen	36
Figure 20:	Configuration completion pop-up	36
Figure 21:	Configuration screen after completion	37
Figure 22:	Main CF screen	38
Figure 23:	Cluster resource diagram	58
Figure 24:	Adding a new node	68

Figures

Figure 25:	Invoking the Cluster Admin GUI	76
Figure 26:	Top menu	77
Figure 27:	Cluster menu	78
Figure 28:	CF main screen	79
Figure 29:	CF node information	81
Figure 30:	CF topology table	83
Figure 31:	Starting CF	84
Figure 32:	Stop CF	85
Figure 33:	Stopping CF	86
Figure 34:	CF log viewer	88
Figure 35:	Detached CF log viewer	89
Figure 36:	Search based on date/time	90
Figure 37:	Search based on keyword	91
Figure 38:	Search based on severity	92
Figure 39:	ICF statistics	94
Figure 40:	MAC statistics	95
Figure 41:	Selecting a node for node to node statistics	96
Figure 42:	Node to Node statistics	97
Figure 43:	CIM options	98
Figure 44:	Add to CIM	99
Figure 45:	Unconfigure CF	100
Figure 46:	CIM Override	101
Figure 47:	CIM Override confirmation	102
Figure 48:	Three-node cluster with working connections	104
Figure 49:	Three-node cluster where connection is lost	104
Figure 50:	Node C placed in the kernel debugger too long	107

Figure 51:	Four-node cluster with cluster partition	108
Figure 52:	A three-node cluster with three full interconnects	115
Figure 53:	Broken Ethernet connection for hme1 on Node A	116
Figure 54:	Cluster with no full interconnects	117
Figure 55:	Opening the SF Configuration Wizard	131
Figure 56:	Selecting the mode of SF configuration	132
Figure 57:	Easy mode of SF configuration	133
Figure 58:	Detailed mode of SF configuration	134
Figure 59:	Choice of common configuration for all nodes	135
Figure 60:	Individual configuration for Cluster Nodes	136
Figure 61:	Choose Shutdown Agent to be added	137
Figure 62:	Details for SCON Shutdown Agent	138
Figure 63:	Configuring the SCON Shutdown Agent	139
Figure 64:	Configuring RCCU	140
Figure 65:	RCCU default values	141
Figure 66:	Configuring the NPS Shutdown Agent	142
Figure 67:	Configuring the RPS Shutdown Agent	143
Figure 68:	Add, Delete, Edit Shutdown Agent	144
Figure 69:	Finishing configuration	145
Figure 70:	Order of the Shutdown Agents	146
Figure 71:	Shutdown Agent time-out values	147
Figure 72:	Entering host weights and admin IPs	148
Figure 73:	SF configuration files	149
Figure 74:	Saving SF configuration	150
Figure 75:	Status of Shutdown Agents	150
Figure 76:	Exiting SF configuration wizard	151

Figures

Figure 77:	Single cluster console	163
Figure 78:	Distributed cluster console	164
Figure 79:	Conceptual view of CF interconnects	173
Figure 80:	CF with Ethernet interconnects	174
Figure 81:	CF with IP interconnects	174

Tables

Table 1: Local states	80
Table 2: Remote states	80
Table 3: CF log viewer severity levels	92
Table 4: Basic layout for the CF topology table	113
Table 5: Topology table with check boxes shown	114
Table 6: Topology table for 3 full interconnects	116
Table 7: Topology table with broken Ethernet connection	117
Table 8: Topology table with no full interconnects	118
Table 9: Resource Database severity levels	257

Index

.rhosts 16
/etc/cip.cf 59
/etc/hosts
 CF names 166
 CIP configuration 10
 CIP symbolic name 39
 CIP Wizard 32
 configuring cluster console 165
 updating 165
/etc/rmshosts 167
/etc/system 56
/etc/uucp/Devices 166
/etc/uucp/Systems 166
/mydir/rdb.tar.Z 69
/tmp 69
/usr/sbin/shutdown 72, 73

A

adding
 new node 59
 nodes 24
 to CIM 99
administrative access 162
alternate abort sequence 170
automatic resource registration 64
awk script 63

B

backing up
 configuration 40
 Resource Database 68
booting with kadb 170
broadcast messages 12
broken interconnects 103

C

CCBR
 See Cluster Configuration Backup
 and Restore
CCBRHOME directory 43

CF

See also Cluster Foundation

CF commands
 cfconfig 301
 cfset 301
 cftool 301
CF driver 20
CF over IP 11, 30, 173
 broadcast mask 173
 CF Wizard 175
 cftool -d 176
 configure 175
 devices 176
 mixed configurations 174
 scenarios 12
 unique IP address 173
CF Registry
 cfregd 49
 user-level daemon 49
CF Remote Services 34
CF Wizard
 bringing up 21
 CF driver 112
 CF over IP 30, 175
 displaying interconnects 30
 edit node names 26
 error message 37
 new cluster 23
 new node on existing cluster 112
 scanning for clusters 22
 summary screen 35
CF/CIP Wizard, starting 10
cfbackup 40
cfconfig 203
cfconfig -L 111
cfconfig -l 111
cfconfig messages 196
CFCP 14
cfcp 15, 34
CFReg 52
cfrestore 40

- cfset 13, 301
 - CFCP 14
 - CFSH 14
 - CLUSTER_TIMEOUT 14
 - maximum entries 14
 - options 14
 - tune timeout 15
- CFSH 14
- cfsh 34
- cfsmntd 302
- cftool 208
- cftool -d 176
- cftool messages 206
- cftool -n 103
- CIM
 - See* Cluster Integrity Monitor
- CIP
 - See* Cluster Internet Protocol
- CIP commands
 - cip.cf 302
 - cipconfig 302
 - ciptool 302
- CIP Wizard
 - /etc/hosts 32
 - CIP interface 32
 - CIP names 33
 - Cluster Admin 10
 - configuration file 32
 - numbering 32
 - screen 31
 - starting 10
- cip.cf 38, 39
- cipconfig messages 204
- clautoconfig 61
- clbackuprdb 69, 304
- clgettree 60, 61, 64, 70, 72, 305
 - output 60
 - verify configuration 61
- CLI
 - See* Command Line Interface
- clinitreset 60, 61, 72, 304
- clrestorerdb 73, 304
- clroot 17
- clsetparam 66, 304
- clsetup 60, 61, 70, 71, 72, 304
- clstartsrc 304
- clstoprsc 304
- cluster
 - additional node 56
 - avoiding single point of failure 9
 - CF states 80
 - CIP traffic 9
 - data file 49
 - interfaces 8
 - name 7
 - node in consistent state 50
 - number of interconnects 9
 - partition 107
- Cluster Admin
 - administration 75
 - CF over IP 30
 - cluster menu 76
 - configuring cluster nodes 169
 - invoking 76
 - login screen 19
 - main screen 38
 - Shutdown Facility 131
 - starting 19, 76
 - starting CF 83
 - stopping CF 83
 - top menu 76
- Cluster Configuration Backup and Restore 40
 - ccbr.conf 42
 - CCBRHOME directory 43
 - cfbackup 40
 - cfrestore 40
 - configuration file 42, 44
 - OS files 45
 - root files 45
- cluster console, single
 - See* SCON
- cluster consoles 161
 - configuration 165
 - distinct 163
 - distributed 162, 164
 - IP name 153
 - multiple 162

- redirecting input/output 169
- role of 161
- updating configuration 168
- using 170
- xco utility 171
- XSCON_CU variable 171
- Cluster Foundation
 - administration 75
 - configuration 7
 - connection table 28
 - dependency scripts 86
 - device driver 182
 - devices 113
 - driver load time 112
 - interface 8
 - IP interfaces 8
 - loading driver 20
 - log viewer 87
 - main table 80
 - name 153
 - names 165, 166
 - node information 81
 - node name 8, 59
 - quorum set 34
 - Reason Code table 229
 - remote services 34
 - runtime messages 224
 - security 15
 - topology table 28, 83, 111
 - unconfigure 100
- Cluster Integrity Monitor 50
 - adding a node 98
 - CF quorum set 34
 - cfcf 34
 - cfsh 34
 - configuration screen 34
 - node state 50
 - Node State Management 50
 - options 98
 - override 101
 - override confirmation 102
 - quorum state 51
 - rcqconfig 51
- Cluster Internet Protocol
 - /etc/cip.cf 59
 - /etc/hosts 10, 39
 - CF Wizard 59
 - cfname 39
 - configuration 9
 - configuration error 70
 - configuration file 38
 - configuration reset 72
 - configuration verification 71
 - defining 9
 - file format 38
 - interfaces 9, 60
 - IP information 39
 - name 59, 60
 - ping command 60
 - properly configured 59
 - stopping 39
 - subnetwork 60
 - syntax 39
- CLUSTER_TIMEOUT 14
- collecting information 191
- COMING UP state 86
- Command Line Interface
 - configuring RCCU 154
 - configuring SA 153
 - configuring SCON 153
 - configuring with 151
 - SD 151
- configuration
 - changing 51
 - hardware 69
 - restore 67
 - updating on cluster console 168
 - verify 72
- Configuration Wizard, invoking 131
- Configure script
 - cluster console 165
 - running 166
- configuring
 - CF 10
 - CF driver 21
 - CF over IP 175

Index

- CIM 50
- CIP 9, 10, 31, 38
- CIP with CF Wizard 59
- cluster console 165
- NPS 125
- NPS shutdown agent 142
- RCI 122
- resource database 59
- SA_scon 169
- SCON 123
- SF 169
 - with CLI 151
- connection table 29
- contents, manual 1
- conversion unit 163
- corrupt data 105
- creating
 - cluster, example 16
 - new cluster 23
- D**
- data, corrupt 105
- debugging 159
- default values, Solaris kernel 56
- defining virtual CIP interfaces 9
- devices
 - displayed 111
 - Ethernet 114
 - unconnected 28
- diagnostics 177
- diskinfo file 62
- display statistics 93
- displayed devices 111
- distributed cluster consoles 162, 164
- dkconfig 303
- dkmigrate 303
- dkmirror 303
- dktab 304
- documentation 2
- DOWN state 86, 104, 105
- E**
- editing
 - /etc/hosts file 165

- CF node names 26
 - cip.cf file 38
 - cluster.config file 13
 - diskinfo file 63
 - existing configuration 134
 - kbd file 170
 - rcsd.cfg.template 152
 - rmshosts file 167
 - rmshosts method file 168
 - SCON 167
 - SF configuration 151
- EMC power devices 62
- EMC Symmetrix 62
- ERRNO table 241
- ERROR messages
 - MA 296
 - Resource Database 261
- error messages 208
 - different systems 240
 - rcqconfig 211
 - rcquery 223
- errors, CIP configuration 70
- Ethernet
 - CF over IP 173
 - devices 114
 - Gigabit 176
- excluded device list 62
- F**
- fjsnap command 191
- fjsvvws 307
- fjsvwvcnf 307
- fsck_rdfs 302
- full interconnect 28, 113
- G**
- Gigabit Ethernet 176
- GUI
 - See* Cluster Admin
- H**
- HALT messages 258

I

ICF statistics 94
ifconfig 39
INFO messages
 MA 295
 Resource Database 259
init command 103
interconnects
 CF 8
 CF over IP 173
 Ethernet 114
 full 28
 IP 30
 IP subnetwork 174
 number of 9
 partial 28
 topology table 113
interfaces 8
 CIP 11
 Cluster Internet Protocol 60
Internet Protocol address 165
 CIP interface 32
 RCCU 124
INVALID state 86
IP address
 See Internet Protocol address
IP interfaces 8
IP name, CIP interface 32
IP over CF 11
IP subnetwork 174

J

join problems 183
joining a running cluster 65

K

kadb
 booting with 170
 restrictions 170
kbd 170
kernel parameters 56
keyword, search based on 91

L

Largest Sub-cluster Survival 128
LEFTCLUSTER state 103, 106, 108
 cluster partition 107
 defined 315
 description 104
 displaying 103
 in kernel debugger too long 106
 lost communications 105
 panic/hung node 106
 purpose 105
 recovering from 106
 shutdown agent 105
 troubleshooting 189
LOADED state 86
loading
 CF driver 20, 111
 CF driver differences 111
 CF driver with CF Wizard 27
 CF duration 27
local file systems, mount 73
local states 80
login
 password 17
 screen 19
low latency 8

M**MA**

See Monitoring Agents

MA commands

 clrcumonctl 303
 clrcimonctl 303
MAC statistics 95
main CF table 80
manual contents 1
manual pages
 display 301
 listing 301
marking down nodes 86
messages
 alphabetical 225
 CF 224
 cfconfig 196

- cftool 206
- cipconfig 204
- error 208
- HALT 258
- MA 294
- rcqconfig 211
- rcquery 223
- SF 288
- mirror virtual disks 316
- Monitoring Agent messages 294
- Monitoring Agents 119
- mount_rdfs 302
- mountall 73
- Multi-path automatic generation 63
- Multi-Path Disk Control Load Balance 63
- multiple cluster consoles 162

N

names

- /etc/hosts file 165
- attribute 154
- CCBR 42
- CCBRHOME directory 43
- CF 82
- CF cluster 113
- CF node 39
- cname 10, 70, 157, 171
- CIP 71
- cluster 7, 24, 81
- configuration file 7
- connections table 29
- determining CF name 155
- IP 32, 156
- node 157
- plug-ins 41
- RCCU 155
- rmshosts file 167
- symbolic node 39
- tuple entries 13
- user 17, 76
- Web-Based Admin View 8
- with asterisk 112

network considerations 165

- network outages 8
- Network Power Switch 121
 - configuration 125
 - configuring SA 142
 - setup 125
- ngadmin 302
- Node State Management 50
- Node to Node statistics 97
- nodes
 - adding 24
 - adding a new 67
 - details 81
 - in kernel debugger 103
 - joining a running cluster 65
 - marking down 86
 - other configuration 169
 - panicked 103
 - shut down 86

NPS

See Network Power Switch

O

OS files 45

P

- panicked nodes 103
- partial interconnects 28, 113
- PAS commands
 - clmtest 303
 - mipcstat 303
- passwords 17, 76
- plumb-up state 64
- privileged user ID 17
- pseudo device driver 323
- public IP names 166
- public networks, security 15

Q

- quorum
 - CF 34
 - CIM override 101
 - reconfiguring 51
 - state 51

R

- RAID 322
- rc scripts 112
- RC_sf 158
- rc2.d directory 158
- RCA
 - See* Remote Console Access
- RCCU
 - See* Remote Console Control Unit
- rcfs_fumount 302
- rcfs_list 302
- rcfs_switch 302
- RCI
 - See* Remote Cabinet Interface 50
- rcqconfig 50, 51
- rcqconfig messages 211
- rcquery messages 223
- RC-script 158
- rcsd 306
- rcsd log 159
- rcsd.cfg 152, 306
- Reason Code table 229
- reboot command 103
- rebooting
 - after cfrestore command 41
 - clusterwide 49
 - RCI restrictions 122
 - reboot command 103
 - shut down CF 103
- reconfiguring Resource Database 69
- redirecting console input/output 169
- registering hardware 61
- Remote Cabinet Interface 50, 121
 - configuration 122
 - hardware setup 122
 - log file 123
 - node elimination 122
 - restrictions 122
 - SA 122
 - setup 122
 - shutdown mechanism 122
- Remote Console Access 163
- Remote Console Control Unit 121
 - attribute value pairs 154
 - configuring with CLI 154
 - defining cluster node 154
 - IP address 124
 - log file 124
 - SA_rccu 123
 - setup 124
 - topologies 163
- Remote Power Switch 121
 - SA_rps 124
 - setup 124
- remote states 80
- reserved words, SCON 153
- Resource Database 59
 - adding new node 67
 - backing up 68
 - clgettree 61
 - clsetup 70
 - configure on new node 71
 - initializing 67
 - kernel parameters 56
 - new node 67
 - plumb-up state 64
 - reconfiguring 67, 69
 - registering hardware 61, 64
 - restoring 72, 73
 - start up synchronization 65
 - StartingWaitTime 66
- restoring Resource Database 72, 73
- RFC 1918 9
- RMS commands
 - hvasert 305
 - hvcn 305
 - hvconfig 305
 - hvdsp 305
 - hvdist 305
 - hvdump 305
 - hvenv.local 306
 - hvgdmake 305
 - hvgclean 305
 - hvsht 305
 - hvswitch 305
 - hvthrottle 305
 - hvtutil 305
- RMS Wizard Tools 129

Index

- rmshosts file 165, 167
- root 17
- root files 45
- RPS
 - See* Remote Power Switch
- S**
- SA
 - See* Shutdown Agents
 - SA specific log files 159
 - SA_rccu.cfg 306
 - SA_rps.cfg 306
 - SA_scon 123
 - SA_scon Shutdown Agent 169
 - SA_scon.cfg 153, 306
 - SA_scon.cfg template 153
 - SA_sspint.cfg 307
 - SA_sunF.cfg 307
 - SA_wtinps.cfg 307
- SCON 121, 162
 - algorithm 127
 - arbitration 167
 - configuration 123
 - port number 153
 - reserved words 153
 - SA_scon 123
 - set up 123
 - Shutdown Facility 127
 - topologies 164
- scon command 306
- scon.scr 171
- SD
 - See* Shutdown Daemon
- sdtool 307
- sdtool command 119, 158
- sdtool -d on command 159
- search
 - keyword 91
 - severity levels 92
 - time filter 90
- security
 - CF 15
 - public network 15
- selecting devices 114
- seminfo_semmni 56
- seminfo_semmns 56
- seminfo_semmnu 56
- serial-line-to-network converter 163
- setting up
 - RCCU 124
 - RCI 122
 - RPS 124
 - SCON 123
- SF
 - See* Shutdown Facility
- SF Wizard 7
 - opening 131
 - starting 38
- shminfo_shmma 56
- shminfo_shmmax 56, 57
- shminfo_shmmni 56
- shminfo_shmseg 56
- shutdown 73
- Shutdown Agents 119, 123
 - configuring with CLI 153
 - SA_scon 169
 - with LEFTCLUSTER 105
- shutdown command 103
- Shutdown Daemon 119
 - configuration file 152
 - configuration file format 152
 - configuring with CLI 151
 - rcsd.cfd 152
- Shutdown Facility 7, 119
 - configuring 169
 - internal algorithm 127
 - messages 288
 - node weight 129
 - RMS Wizard Tools 129
 - SCON 123, 127
 - selecting configuration 132
 - split-brain handling 125
 - starting and stopping 158
 - starting automatically 158
 - starting manually 158
 - stopping automatically 158
 - stopping manually 158
 - weight assignment 126

- shutdown requests 165
 - ShutdownPriority attribute 126
 - SIGKILL 167
 - simple virtual disks 321
 - single cluster console
 - See* SCON
 - single user mode 69
 - SIS commands
 - dtcpadmin 307
 - dtcpd 307
 - dtcpdbg 307
 - SMAWcf 40
 - SMAWRscon 164
 - Configure script 171
 - port number 153
 - SCON 164
 - software 161
 - starting 171
 - SMAWsf directory 152
 - Solaris/Linux ERRNO table 241
 - special priority interfaces 9
 - Specific Application Survival 129
 - Specific Hardware Survival 129
 - split-brain 125
 - LSS 128
 - SAS 129
 - SHS 129
 - start up synchronization 65
 - new node 67
 - StartingWaitTime 72
 - starting
 - CF 83, 84
 - CF Wizard 21
 - Cluster Admin 10
 - GUI 19
 - SF Wizard 38, 131
 - Web-Based Admin View 16
 - StartingWaitTime 65, 67
 - default value 66
 - value 66
 - states
 - COMING UP 86
 - DOWN 104, 105
 - INVALID 86
 - LEFTCLUSTER 103, 106, 108
 - LOADED 86
 - table of 104
 - UP 104
 - statistics, display CF 93
 - stopping
 - CF 83, 84
 - CF, third-party products 86
 - CIP 39
 - SD 121
 - SF automatically 158
 - SF manually 158
 - valid CF states 86
 - subnet mask, CIP interface 32
 - synchronization phase 65
 - synchronization, start up 65
 - syslog window 88
 - system dump 191, 192
- T**
- table of states 104
 - third-party product, shut down 86
 - time filter, search 90
 - timeout, tune 15
 - timestamp 66
 - topology table 111
 - basic layout 113
 - CF 28, 83
 - CF cluster name 113
 - CF driver 112
 - displayed devices 111
 - displaying 82
 - examples 115
 - flexibility 29
 - interconnects 113
 - selecting devices 114
 - troubleshooting 177
 - beginning 177
 - collecting information 191
 - diagnostics 177
 - join-related problems 182
 - symptoms and solutions 181
 - tunable parameters 13
 - tune timeout 15

Index

tuple entries

name 13

value 13

U

uname 165

unconfigure CF 100

unconnected devices 28, 113

UNLOADED state 86

UP state 104

updating CFReg 52

usage messages

cfconfig 196

cftool 207

cipconfig 204

rcqconfig 211

rcquery 223

user ID 17

user name 17

using the cluster console 170

V

VCMDB 62

vdisk 303

virtual disks

mirror 316

simple 321

W

WARNING messages

MA 296

Resource Database 260

Web-Based Admin View

node list 20

starting 16

wvCntl 307

wvGetparam 307

wvroot 17

wvSetparam 308

wvstat 308

X

xco utility 171

xsco utility 171

XSCON_CU variable 171

Fujitsu Siemens Computers GmbH
User Documentation
33094 Paderborn
Germany

Comments
Suggestions
Corrections

Fax: (++49) 700 / 372 00001

email: manuals@fujitsu-siemens.com
<http://manuals.mchp.siemens.de>

Submitted by

Comments on PRIMECLUSTER™
Cluster Foundation (CF) (Solaris)



Fujitsu Siemens Computers GmbH
User Documentation
33094 Paderborn
Germany

Comments
Suggestions
Corrections

Fax: (++49) 700 / 372 00001

email: manuals@fujitsu-siemens.com
<http://manuals.mchp.siemens.de>

Submitted by

Comments on PRIMECLUSTER™
Cluster Foundation (CF) (Solaris)



