PRIMECLUSTER

Installation and Administration Guide 4.2

Oracle(R) Solaris

Preface

This manual serves as your starting point for using PRIMECLUSTER. It explains the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system. Since the PRIMECLUSTER system comprises multiple features, there are several other manuals besides this one for each feature. However, by reading this manual first, you will be able to perform the series of operations because this manual refers readers to other manuals that contain feature-specific information that is necessary for the operations.

This manual also provides a functional overview of products that are supported by the PRIMECLUSTER system and describes operation procedures.

This manual only covers the basic operation of PRIMECLUSTER. For operations using different hardware and software configurations, see "Related Documentation".

The table below shows the operation flow from PRIMECLUSTER installation to the start of operation management and indicates the reference location in this manual for each operation.

Flow from PRIMECLUSTER system installation to operation management

	PRIMECLUSTER system operation flow	Reference location in this manual
1	Understanding the flow of PRIMECLUSTER system building and designing the PRIMECLUSTER system	Part 1 Planning
2	Installing the PRIMECLUSTER system	Part 2 Installation
3	Monitoring the operation status of the PRIMECLUSTER system	Part 3 Operations
4	Changing the PRIMECLUSTER system configuration after system operation has been started	Part 4 System Configuration Modifications
5	Backing up and restoring the operation environment of the PRIMECLUSTER system	Part 5 Maintenance
6	Installing products for the PRIMECLUSTER system	Part 6 PRIMECLUSTER Products

For detailed procedural explanations, refer to the reference manuals that are indicated in the target location of each part.

Target Readers

This manual is intended for all users who use PRIMECLUSTER 4.2 and perform cluster system installation and operation management. It is also intended for programmers who develop applications that operate on PRIMECLUSTER.

Configuration of This Documentation

This manual consists of six parts, appendixes, and a glossary. The contents of each part are described below.

Part 1 Planning

> Audience: System administrators who build PRIMECLUSTER systems
> Contents: This part describes the overall workflow from installation to operation of the PRIMECLUSTER system.

Part 2 Installation

> Audience: System administrators who build PRIMECLUSTER systems
> Contents: This part describes operations for software installation, cluster building, and application building.

Part 3 Operations

> Audience: System administrators who manage system operations.
> Contents: This part describes operation methods for operations such as monitoring the PRIMECLUSTER system and investigating failure causes.

Part 4 System Configuration Modification

> Audience: System administrators who build PRIMECLUSTER systems
> Contents: This part describes necessary work items for additions, modifications, and deletions to the PRIMECLUSTER system configuration.

Part 5 Maintenance

Audience: System administrators who build PRIMECLUSTER systems
Contents: This part describes the necessary work items for maintaining the operation environment of the PRIMECLUSTER system.

Part 6 PRIMECLUSTER Products

Audience: Users who operate PRIMECLUSTER products on PRIMECLUSTER systems
Contents: This part describes the versions and levels of products that are supported by the PRIMECLUSTER system and provides a functional overview of those products.

Appendix A PRIMECLUSTER System Design Worksheets

Audience: Users who design PRIMECLUSTER systems
Contents: This appendix contains the PRIMECLUSTER System Design Worksheet.

Appendix B Manual Pages

Audience: All users who use PRIMECLUSTER systems
Contents: This appendix describes the online manual pages that are used by the individual features of the PRIMECLUSTER system.

Appendix C Troubleshooting

Audience: All users who use PRIMECLUSTER systems
Contents: This appendix describes corrective actions for problems that may occur in the PRIMECLUSTER system. It also explains how to collect data when requesting a problem investigation.

Appendix D Messages

Audience: All users who use PRIMECLUSTER systems
Contents: This appendix describes the contents and actions for messages output by the userApplication Configuration Wizard GUI and messages output by procedure resources in PRIMECLUSTER systems. For other messages, this manual indicates the reference locations in other manuals.

Appendix E Using SynfinityCluster Products in PRIMECLUSTER

Audience: All users who used the SynfinityCluster system
Contents: This appendix describes feature differences and terminology differences between SynfinityCluster and PRIMECLUSTER.

Appendix F Registering, Changing, and Deleting State Transition Procedure Resources for SynfinityCluster Compatibility

Audience: All users who use SynfinityCluster-compatible resources
Contents: This appendix describes procedures for registering, changing, and deleting procedure resources when the cluster applications use procedure resources.

Appendix G Registering SynfinityCluster-compatible Resources

Audience: All users who use SynfinityCluster-compatible resources
Contents: This appendix describes the procedure for registering SynfinityCluster-compatible resources.

Appendix H Using PRIMECLUSTER in Oracle VM Server for SPARC Environments

Audience: All users who use Oracle VM Server for SPARC environments in PRIMECLUSTER systems
Contents: This appendix explains the setup and configuration procedures for using Oracle VM Server for SPARC environments in PRIMECLUSTER systems.

Appendix I Release information

Audience: All users who use PRIMECLUSTER systems
Contents: This appendix describes the main contents for change of this manual.

Glossary

Audience: All users who use PRIMECLUSTER systems
Contents: This section explains terms used to describe the PRIMECLUSTER system.

Related Documentation

Refer to the following manuals as necessary when setting up the cluster:

- *PRIMECLUSTER Concepts Guide*

- *PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*

- *PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide*

- *PRIMECLUSTER Reliant Monitor Services (RMS) Troubleshooting Guide*

- *PRIMECLUSTER Global Disk Services Configuration and Administration Guide*

- *PRIMECLUSTER Global File Services Configuration and Administration Guide*

- *PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function*

- *PRIMECLUSTER Global Link Services Configuration and Administration Guide: Multipath Function*

- *PRIMECLUSTER Web-Based Admin View Operation Guide*

- *PRIMECLUSTER Scalable Internet Services (SIS) Configuration and Administration Guide*
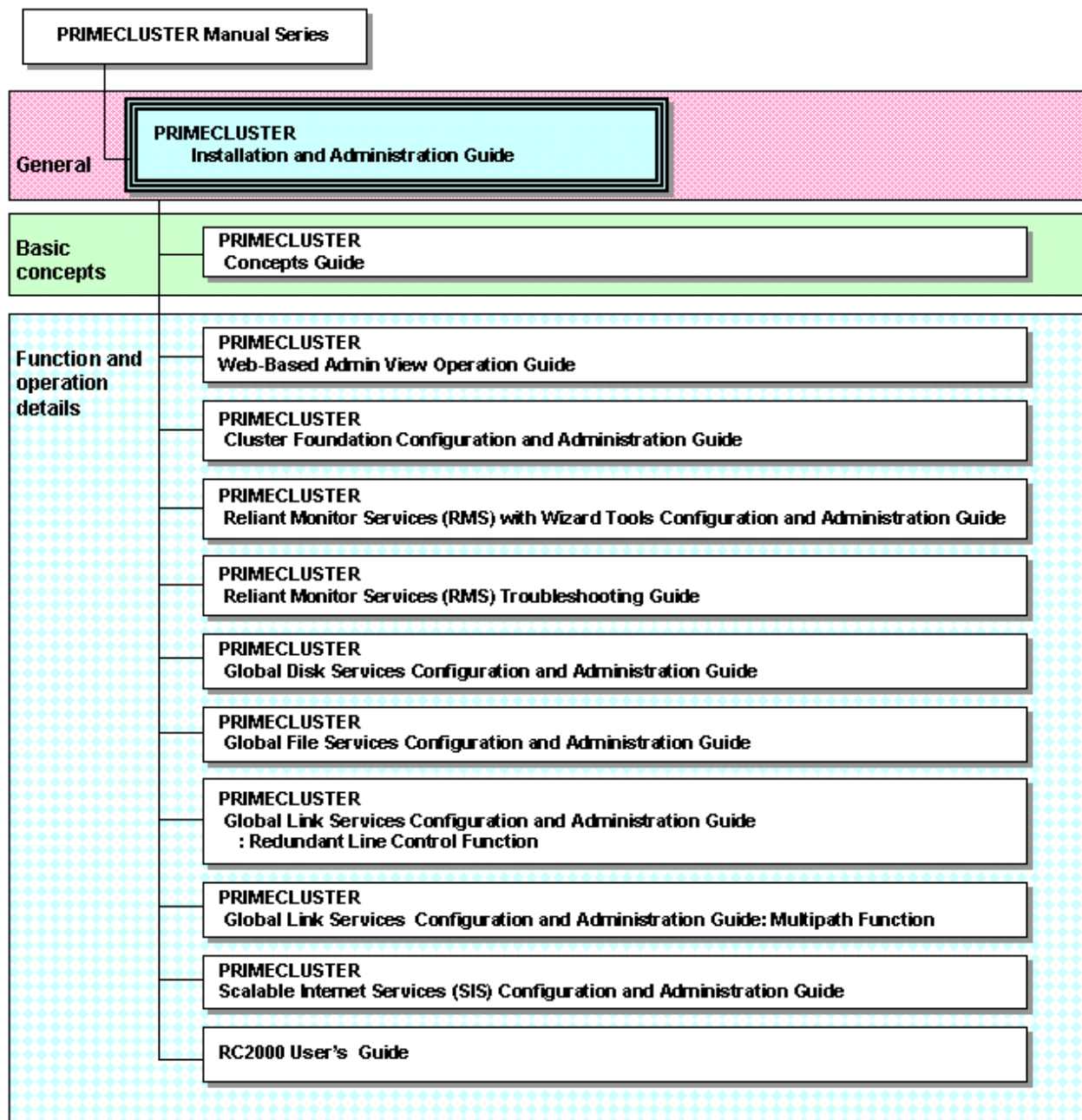
- *RC2000 User's Guide*

Note

· ·

The PRIMECLUSTER documentations include the following documentation in addition to those listed above:

- PRIMECLUSTER Installation Guide

  This printed installation guide is included with each PRIMECLUSTER product.

· ·

Manual Series

Manual Printing

If you want to print a manual, use the PDF file found on the CD-ROM for the PRIMECLUSTER product. The correspondences between the PDF file names and manuals are described in the PRIMECLUSTER Installation Guide that comes with the product.

You will need Adobe Reader to read and print the PDF file. Use Adobe Reader Version 5.0 or higher.

Online Manuals

To allow users to view the online manuals, use the Cluster management server to register each user name to one of the user groups (wvroot, clroot, cladmin, or clmon).

For information on user group registration procedures and user group definitions, see "4.3.1 Assigning Users to Manage the Cluster"

Conventions

Notation

Prompts

Command line examples that require system administrator (or root) rights to execute are preceded by the system administrator prompt, the hash sign (#). Entries that do not require system administrator rights are preceded by a dollar sign ($).

Manual page section numbers

References to the UNIX operating system commands are followed by their manual page section numbers in parentheses - for example, cp(1)

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.

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.
Italic	Variables that you must replace with an actual value.
Bold	Items in a command line that you must type exactly as shown.

Example 1

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

```
root:x:0:1:0000-Admin(0000):/:
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):/:
```

Example 2

To use the cat(1) command to display the contents of a file, enter the following command line:

```
$ cat file
```

Notation symbols

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

P **Point**

. . . .

Contains important information about the subject at hand.

. . . .

Note

. . . .

Describes an item to be noted.

. . . .

Example

. . . .

Describes operation using an example.

. . . .

Information

....

Describes reference information.

....

 See

....

Provides the names of manuals to be referenced.

....

Abbreviations

- Microsoft(R) Windows(R) 98 operating system and Microsoft(R) Windows(R) 98 Second Edition operating system are abbreviated as Windows(R) 98.

- Microsoft(R) Windows NT(R) Server operating system Version 4.0, and Microsoft(R) Windows NT(R) Server Workstation System Version 4.0 are abbreviated as Windows NT(R).

- Microsoft(R) Windows(R) 2000 operating system is abbreviated as Windows(R) 2000.

- Microsoft(R) Windows(R) Millennium Edition is abbreviated as Windows(R) Me.

- Microsoft(R) Windows(R) XP operating system is abbreviated as Windows(R) XP.

- Windows(R) 98, Windows NT(R), Windows(R) 2000, Windows(R) Me, Windows(R) XP, and Windows(R) Vista are generically referred to as Microsoft(R) Windows.

- Oracle Solaris might be described as Solaris, Solaris Operating System, or Solaris OS.

- If "Solaris X" is indicated in the reference manual name of the Oracle Solaris manual, replace "Solaris X" with "Oracle Solaris 9 (Solaris 9)" or the "Oracle Solaris 10 (Solaris 10)".

Trademarks

UNIX is a registered trademark of Open Group in the United States and other countries.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Netscape and Netscape Navigator are trademarks or registered trademarks of Netscape Communications Corporation in the United States and other countries.

Microsoft Internet Explorer is a registered trademark of Microsoft Corporation in the United States and other countries.

Microsoft, Windows, Windows NT, and Windows NT Server are registered trademarks of Microsoft Corporation in the United States and other countries.

CORBA, OMG, and ORB are registered trademarks of Object Management Group (OMG).

NetWorker is a registered trademark or trademark of EMC Corporation in the United States and other countries.

Symfoware is a registered trademark of Fujitsu Limited.

PRIMECLUSTER is a trademark of Fujitsu Limited.

Other product names are product names, trademarks, or registered trademarks of these companies.

Copyright (c) 1986-1997 Sun Microsystems, Inc.

Copyright (c) 1983-1989 Portions may be derived from Berkeley BSD system, licensed from the University of California.

(C) 1990-2000 Legato Systems, Inc. All Rights Reserved.

Requests

- No part of this documentation may be reproduced or copied without permission of FUJITSU LIMITED.

- The contents of this documentation may be revised without prior notice.

Date of publication and edition

Copyright notice

Editing record

Additions and changes	Section	Manual code/Edition
Added descriptions for ILOM Reset.	5.1.2 Configuring the Shutdown Facility 5.1.2.2.3 Specifying the Timeout Value 5.1.2.4.2 Setting by CLI A.6 Cluster Configuration Worksheet A.10.1 Cluster Configuration Worksheet	J2S1-6771-05ENZ0(01) J2S1-6771-05ENZ2(01)
Added notes.	5.1.2 Configuring the Shutdown Facility	
Changed descriptions in <How to calculate a timeout value>.	5.1.2.2.3 Specifying the Timeout Value	
Added notes on configuring the shutdown facility.	5.1.2 Configuring the Shutdown Facility	J2S1-6771-05ENZ0(02) J2S1-6771-05ENZ2(02)
Added notes on setting up cluster applications.	6.6 Setting Up Cluster Applications	
Added descriptions for maintenance mode.	7.2.2.6 Entering maintenance mode for Cluster Application	
Added notes on changing cluster applications.	8.1 Changing the Cluster Configuration	
Added notes on changing resources.	8.1.1 Changing the Cluster Application Configuration	
Added procedures to delete and change hardware.	8.10 Deleting and Changing Hardware	
Added procedures to add hardware.	9.3 Adding Hardware	
Changed messages. (3052, 3053, 7050 to 7056)	D.2 Information Messages D.4 Error Messages	
Added procedures to change the startup priority of a state transition procedure.	F.2 Changing a Procedure Resource	
Changed "LDoms (Logical Domains)" to "Oracle VM Server for SPARC."	1.2 Installation Appendix H Using PRIMECLUSTER in Oracle VM Server for SPARC Environments Glossary	5.1 edition
Added SPARC T3 series.	5.1 Initial Cluster Setup	
Added descriptions for series provided in other than Japan.	5.1.2 Configuring the Shutdown Facility 5.1.2.2.2 Using the Shutdown Configuration Wizard A.5.1 CF Configuration	
Added explanations to "See."	8.7 Changing PRIMECLUSTER Operation Attributes	

Additions and changes	Section	Manual code/Edition
Added explanations for the survival priority.	5.1.2.3.2 Setting by CLI 5.1.2.4.2 Setting by CLI	5.2 edition
Added Note when Oracle Solaris Zones of an exclusive IP zone exist on the system.	3.1 PRIMECLUSTER Installation	5.3 edition
Added Note when making the administrative LAN, used in the shutdown facility, redundant by GLS.	5.1.2 Configuring the Shutdown Facility	
Added Note when the connection to the XSCF is SSH connection.	8.2 Changing an IP Address on the Public LAN	

Contents

Part 1   Planning

Part 1 describes the workflow from PRIMECLUSTER design to installation and operation management. Users who are installing a PRIMECLUSTER system for the first time need to read this part.

Chapter 1 Build Flow

This chapter describes the workflow for building a PRIMECLUSTER system. To build a PRIMECLUSTER system, follow the procedure described below.

Figure 1.1 Flow of building a PRIMECLUSTER system
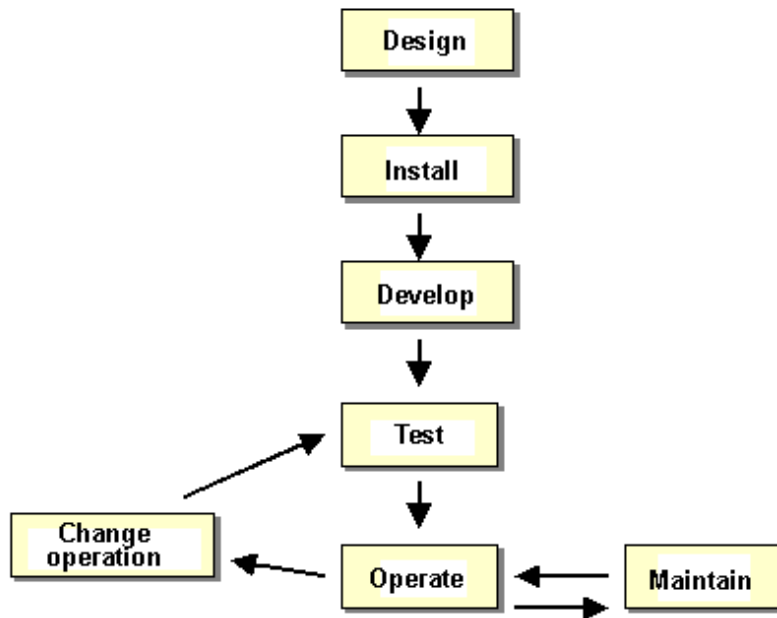
1.1 Planning

Before building a PRIMECLUSTER system, you must first design the system.

Designing a PRIMECLUSTER system

1. Select the PRIMECLUSTER products.

   Select the PRIMECLUSTER products required for the system you want to build.

   For details, see "2.1 PRIMECLUSTER Product Selection".

2. **Design the system.**

   Determine the operation environment for building the PRIMECLUSTER system. This includes selecting the applications to be used and determining the required hardware resources, such as the number of hosts, networks, and disk size.

   For details, see "2.2 System Design".

3. **Determine the cluster system operation mode.**

   Determine the number of nodes and the operation mode of the cluster system.

   For details, see "2.3 Determining the Cluster System Operation Mode".

4. **Determine the operation mode for using Web-Based Admin View.**

   Determine the operation mode for running Web-Based Admin View. Web-Based Admin View can manage up to 16 nodes.

   For details, see "2.4 Determining the Web-Based Admin View Operation Mode".

📝 **Note**

. . . .

In multi-node (3 or more nodes) operation, the 3-tier model, in which the cluster management server is installed separately, is recommended.

. . . .

5. **Determine the cluster applications.**

   Determine the number of cluster applications. Also determine which nodes are to be used for each application.

6. **Determine the resources required for each cluster application.**

   Determine the resources required for each cluster application.

   Determine the switchover network type (IP address takeover, MAC address takeover, or node name takeover) and the takeover address.

   - Determine whether a user-defined RMS configuration script is to be used. Determine whether there are other items (for example, line switching units) to be used as resources.

   - Determine the cluster interconnect paths and quantity. Two or more cluster interconnects are recommended.

   - For a disk device, determine which nodes will be sharing the device, whether the device is to be used as a RAW device (database system), whether the device is to be used as a file system (general files), and whether the device is to be grouped.

7. **Determine the failover range of the cluster application.**

   Determine the trigger for cluster application failover.

   For details, see "2.5 Setting the Failover Timing of a Cluster Application."

📘 **See**

. . . .

For details on designing the system, see "Chapter 2 Site Preparation".

. . . .

1.2 Installation

After completing the design of the PRIMECLUSTER system and determining the configuration of the PRIMECLUSTER system to be built, install the PRIMECLUSTER system.

Since the work will be performed based on the PRIMECLUSTER system design worksheet that was created, check that all items on the design worksheet have been entered.

Install the PRIMECLUSTER system by performing the following procedure in sequence from (1).

Perform the operations described in the dotted line sections if the system design matches the described conditions.

If you are installing applications after you install the PRIMECLUSTER system, go back to the operations from the Application environment setup to the Application installation.

The screens to be used differ according to the operation. The work procedures to be performed with GUI from Web-Based Admin View and the work procedures to be performed with CLI and CUI from console screens are shown in separate boxes.

Figure 1.2 Flow of PRIMECLUSTER system installation
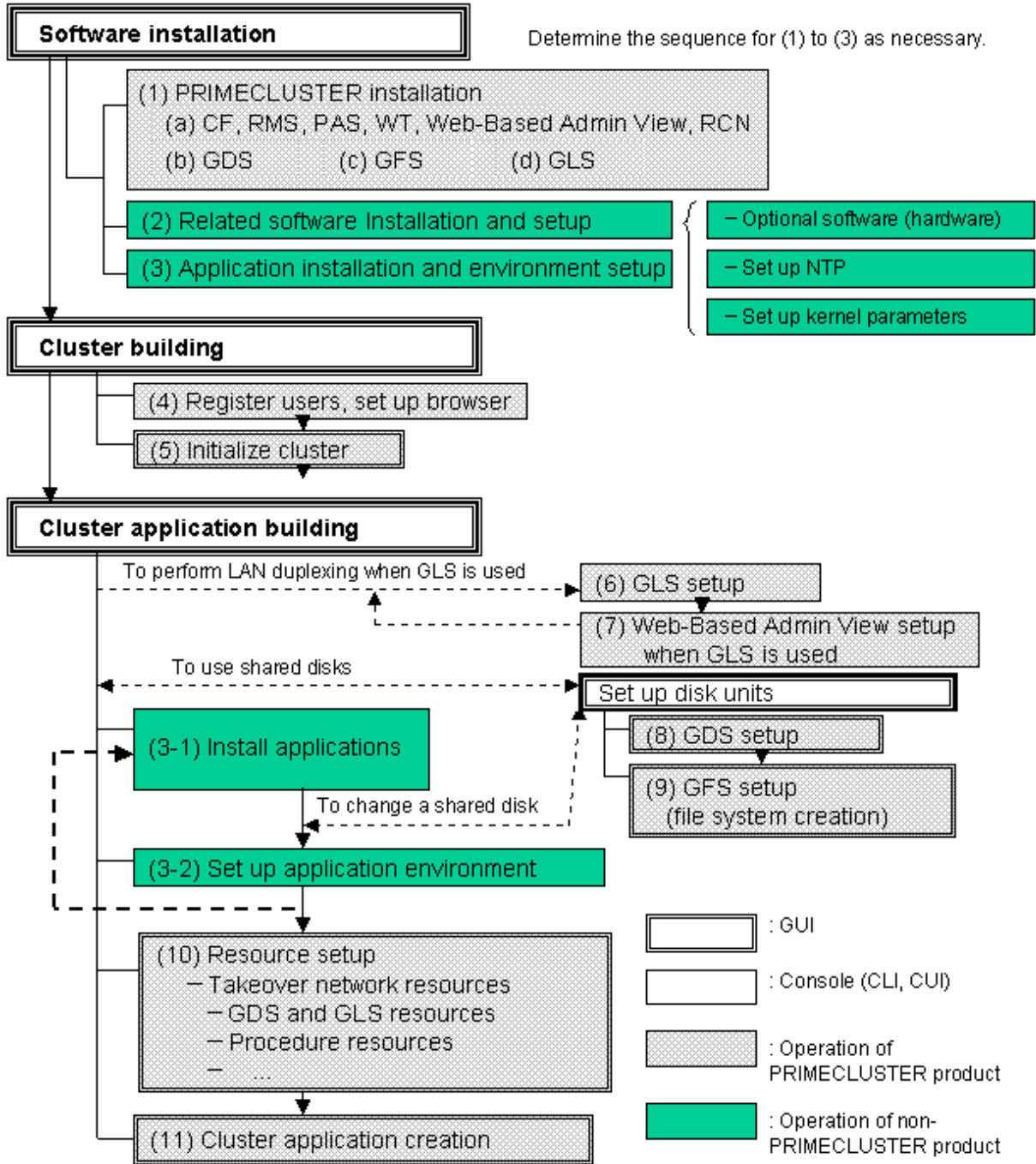
The abbreviations in the flowchart for PRIMECLUSTER system installation are explained below.

CF: Cluster Foundation

RMS: Reliant Monitor Services

PAS: Parallel Application Services

WT: Wizard Tools

RCN: RC2000

GDS: Global Disk Services

GFS: Global File Services

GLS: Global Link Services

For detailed information on each item, refer as necessary to the corresponding manual reference section mentioned in the table below.

Table 1.1 Installation procedure and manual reference sections

Work items	Manual reference section	Required/ optional
System design	Chapter 2 Site Preparation	Required
Software installation		
(1) PRIMECLUSTER installation	3.1 PRIMECLUSTER Installation	Required
(2) Related software Installation and setup	3.2 Installation and Setup of Related Software	Optional
(3) Application installation and environment setup	3.3 Installation and Environment Setup of Applications	Optional
Cluster building		
(4) User registration/browser setup	Chapter 4 Preparation Prior to Building a Cluster	Required
(5) Initial cluster setup	5.1 Initial Cluster Setup	Required
Cluster application building		
(6) GLS setup	6.1 Initial GLS Setup	Optional (required to use GLS)
(7) Web-Based Admin View setup when GLS is used		
(8) GDS setup	6.2 Initial GDS Setup	Optional (required to use GDS)
(9) GFS setup (file system creation)	6.3 Initial GFS Setup	Optional (required to use GFS)
(10) Resource setup	6.6.1 Setting Up Resources	Optional
(11) Cluster application creation	6.6 Setting Up Cluster Applications	Required
Cluster configuration in Oracle VM Server for SPARC environments		
(12) Cluster systems in Oracle VM Server for SPARC environments	H.1 Cluster Systems in Oracle VM Server for SPARC Environments	Optional

1.3 Development

To configure a user application to be monitored by PRIMECLUSTER, you need to create an RMS configuration script.

- Online script

    This script executes a process that sets the resources to Online or Standby.

- Offline script

    This script executes a process that sets the resources to Offline.

To check the state of a user application, the following RMS configuration script must be developed.

- Check script

    This script checks the state of the resource.

See
. . . .

For details on the Online/Offline script and the Check script settings, see "6.5 Setting Up Online/Offline Scripts."
. . . .

1.4 Test

Purpose

When you build a cluster system using PRIMECLUSTER, you need to confirm before starting production operations that the entire system will operate normally and cluster applications will continue to run in the event of failures.

For 1:1 standby operation, the PRIMECLUSTER system takes an operation mode like the one shown in the figure below.

The PRIMECLUSTER system switches to different operation modes according to the state transitions shown in the figure below. To check that the system operates normally, you must test all operation modes and each state transition that switches to an operation mode.

Figure 1.3 State transitions of the PRIMECLUSTER system
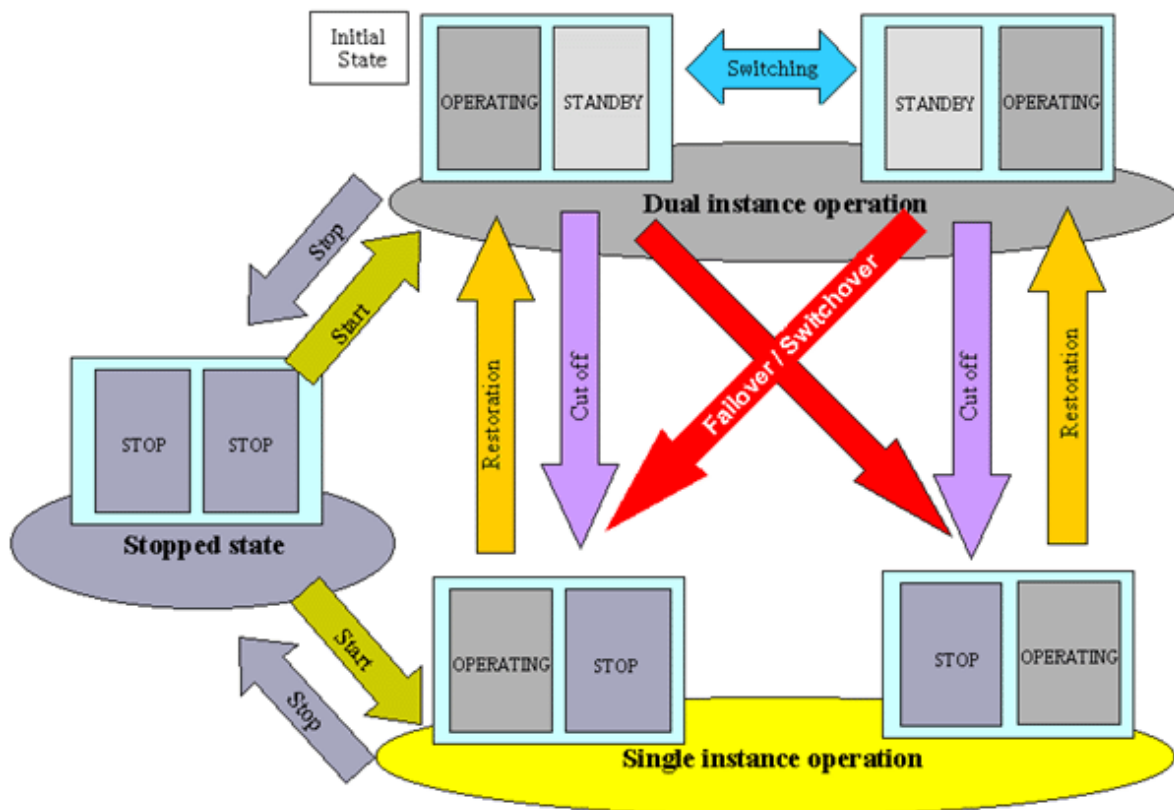
PRIMECLUSTER System State

	Description
Dual instance operation	A cluster application is running, and it can switch to the other instance in the event of a failure (failover). Two types of the dual instance operation are OPERATING and STANDBY. Even if an error occurs while the system is operating, the standby system takes over ongoing operations as an operating system. This operation ensures the availability of the cluster application even after failover.
Single instance operation	A cluster application is running, but failover is disabled. Two types of the single instance operation are OPERATING and STOP. Since the standby system is not supported in this operation, a cluster application cannot switch to other instance in the event of a failure. So, ongoing operations are disrupted.
Stopped state	A cluster application is stopped.

The above-mentioned "OPERATING", "STANDBY", and "STOP" are defined by the state of RMS and cluster application as follows;

	RMS state	Cluster application state	Remark
OPERATING	Operating	Online	-
STANDBY	Operating	Offline or Standby	-
STOP	Stopped	Unknown *	SysNode is Offline

* RMS determines the cluster application state. When RMS is stopped, the cluster application state is unknown.

Main tests for PRIMECLUSTER system operation

Startup test

Conduct a startup test and confirm the following:

- View the Cluster Admin screen of Web-Based Admin View, and check that the cluster system starts as designed when the startup operation is executed.

- If an RMS configuration script was created, check that the commands written in the script are executed properly as follows.

  - For a command that outputs a message when it is executed, check that a message indicating that the command was executed properly is displayed on the console.

  - Check that the command has been executed properly by executing the "ps(1)" command.

- A new cluster application is not started automatically during the PRIMECLUSTER system startup. To start the cluster application automatically, you must set "AutoStartUp" for that cluster application. The AutoStartUp setting must be specified as a userApplication attribute when the application is created. For details, see "6.6.2 Creating Cluster Applications."

Clear fault

If a failure occurs in a cluster application, the state of that application changes to Faulted.

To build and run this application in a cluster system again, you need to execute "Clear Fault" and clear the Faulted state.

Conduct a clear-fault test and confirm the following:

- Check that the Faulted state of a failed application can be cleared without disrupting ongoing operations.

- If an RMS configuration script was created, check that the commands written in the script are executed properly as follows.

  - For a command that outputs a message when it is executed, check that a message indicating that the command was executed properly is displayed on the console.

  - Check that the command has been executed properly by executing the "ps(1)" command.

Switchover

Conduct a failover or switchover test and confirm the following:

- Check that failover is triggered by the following event:

  - When an application failure occurs

- Check that switchover is triggered by the following events:

  - When the OPERATING node is shut down

  - When an application is terminated by the exit operation

  - When an OPERATING cluster application is stopped

- Check that failover or switchover is normally done for the followings:

  - Disk switchover

    Check that the disk can be accessed from the OPERATING node.

    For a switchover disk, you need to check whether a file system is mounted on the disk by executing the "df(1M)" command.

- If the Cmdline resources are to be used, check that the commands written in the Start and Stop scripts for the Cmdline resources are executed properly.

  - For a command that outputs a message when it is executed, check that a message indicating that the command was executed properly is displayed on the console.

  - Check that the command has been executed properly by executing the "ps(1)" command.

- If IP address takeover is set, check that the process takes place normally by executing the "ifconfig(1M)" command.

- If MAC address takeover is set, check that the process takes place normally by executing the "ifconfig(1M)" command.

- If node name takeover is set, check that both the OPERATING and STANDBY nodes have the same node name after network takeover.

- If a line switching unit is set up, check that line switching takes place correctly.

- Check that an application is switched to other node.

You need to know the operation downtime in the event of a failure, so measure the switching time for each failure detection cause and check the recovery time.

Replacement test

Conduct a replacement and confirm the followings:

- Check that the OPERATING and STANDBY instances of the OPERATING business application occur normally when the cluster application replacement is executed. Check the followings:

  - If disk switchover is to be used, check that the disk can be accessed from the OPERATING node but not from the STANDBY node.

    For a switchover disk, you need to check whether a file system is mounted on the disk by executing the "df(1M)" command.

  - If Cmdline resources are to be used, check that the commands written in the Start and Stop scripts for the Cmdline resources are executed properly.

    - For a command that outputs a message when it is executed, check that a message indicating that the command was executed properly is displayed on the console.

    - Check that the command has been executed properly by executing the "ps(1)" command.

  - If IP address takeover is to be used, check that IP address takeover takes place normally.

  - Check that an application is switched to other node.

Stop

Conduct a stop test and confirm the followings:

- Check that an OPERATING work process can be stopped normally by the stop operation.

- Check that work processes can be started by restarting all nodes simultaneously.

- If Cmdline resources are to be used, check that the commands written in the Start and Stop scripts for the Cmdline resources are executed properly.

  - For a command that outputs a message when it is executed, check that a message indicating that the command was executed properly is displayed on the console.

  - Check that the command has been executed properly by executing the "ps(1)" command.

Work process continuity

Conduct work process continuity and confirm the followings:

- Generating some state transitions in a cluster system, check that the application operates normally without triggering inconsistencies in the application data in the event of a failure.

- For systems in which work processes are built as server/client systems, check that while a state transition is generated in the cluster system, work process services can continue to be used by clients, according to the specifications.

1.5 Operation and Maintenance

After confirming that work processes can be continued no matter what state the cluster system lapses into, you can begin actual operations.

 Note
. . . .

The cluster system can continue work processes even if a failure occurs. However, work processes cannot be continued if another failure occurs during single node operation before the first failure is corrected. To enhance reliability, you need to eliminate the cause of the failure immediately and recover the dual node operation.
. . . .

1.6 Operation Mode Change

Change the operation of the PRIMECLUSTER system when it becomes necessary to change the system configuration while the PRIMECLUSTER system is operating. The system configuration must be changed, for example, when a cluster application is added.

Chapter 2 Site Preparation

You must plan the items listed below before building the PRIMECLUSTER system.

Planning items

- PRIMECLUSTER product selection

- System design

- Cluster system operation mode

- Operation mode for using Web-Based Admin View

- Cluster applications and resources to be used by the cluster applications

P Point

An overview of each PRIMECLUSTER product is described in the *"PRIMECLUSTER Concepts Guide."* Be sure to read the guide before designing the PRIMECLUSTER system.

2.1 PRIMECLUSTER Product Selection

The sequence for selecting PRIMECLUSTER products is as follows:

1. Select the products to be used.

   Select necessary PRIMECLUSTER products according to your environment.

   For details, see "2.1.1 Product Selection".

2. Select the functions to be used.

   Check if the products provide the functions you need.

   For details, see "2.1.2 Function Selection".

2.1.1 Product Selection

The product sets described below have been prepared for PRIMECLUSTER. Select the necessary products according to how the system will be used.

- **PRIMECLUSTER Enterprise Edition (EE)**

   All-in-one cluster providing the switching (HA) cluster and parallel database.

   This product is used for scalable operations, such as RAC and Symfoware.

- **PRIMECLUSTER HA Server (HA)**

   Switchover-type cluster system that features HA (switchover) cluster functions, volume management functions, system functions, and network multiplexing functions.

The following table shows the components (modules) that are included in each product.

Components		Products	
Names	**Features**	**EE**	**HA**
Cluster Foundation (CF)	Refers to the basic function that is required for user applications or other PRIMECLUSTER services to manage or communicate within the cluster.	Y	Y
Reliant Monitor Services (RMS)	Refers to the software monitoring function that is used to realize high-availability (HA)	Y	Y

Components		Products	
Names	**Features**	**EE**	**HA**
	of the application that is to be executed within the cluster.		
Wizard Tools	Refers to the function that is used to create an application that is to be controlled with RMS.	Y	Y
Web-Based Admin View	Refers to the function for realizing PRIMECLUSTER operations and monitoring with the GUI (management view).	Y	Y
One Shot Installer	One-time batch installer for OS and PRIMECLUSTER products	Y	Y
System Information Output Tool	Collects system information required for failure investigation	Y	Y
Cluster Configuration Backup/Restore	Backup and restoration for cluster configuration	Y	Y
Global Link Services (GLS)	Provides highly reliable transmission routes by setting up redundant network.	Y	Y
Global File Services (GFS)	Refers to the function that is used to realize simultaneous access to the shared file system from multiple nodes to which the shared disk device is connected.	Y	Y
Global Disk Services (GDS)	Refers to the volume management function that is used to improve the availability and manageability of the data stored on the disk device.	Y	Y
Netcompo BASE	Online trace function (for GLS)	Y	Y
Scalable Internet Services (SIS)	Network load balancing	Y	-
Parallel Application Services (PAS)	Refers to the function that enables the high-performance and high-speed communication with the parallel databases.	Y	-

2.1.2 Function Selection

Check if the products provide the necessary functions, using the following documents:

- PRIMECLUSTER basic functions

  For information on the basic functions, see "3.3 PRIMECLUSTER modules" in the *"PRIMECLUSTER Concepts Guide."*

- Other feature

  In addition to the PRIMECLUSTER basic functions, the following function is also provided:

  - Process monitoring function

    This function monitors the live state of applications and other processes.
    For details, see "6.6.1.7.1 What Is the Process Monitoring Function?."

2.2 System Design

Use the "PRIMECLUSTER System Design Worksheet" to design the system.

The installation of the PRIMECLUSTER system is based on the completed "Appendix A PRIMECLUSTER System Design Worksheets."

2.3 Determining the Cluster System Operation Mode

To ensure that the cluster system to be built operates efficiently, you need to determine the number of nodes and an appropriate operation mode.

PRIMECLUSTER allows you to configure multiple cluster applications. The operation mode is determined depending on how you set up the cluster applications in the cluster system.

The main operation modes are listed below:

Classification	Operation mode	Number of cluster applications	Number of nodes
Standby operation	1:1 standby	1	2
	Mutual standby	2 to (number of nodes)	2 to (number of supported nodes)
	N:1 standby	2 to (number of nodes - 1)	3 to (number of supported nodes)
	Cascade	1 to (number of nodes)	3 to (number of supported nodes)
	Priority transfer	2 to (number of nodes - 1)	3 to (number of supported nodes)
	N:M standby	2 to (number of nodes - 2)	4 to (number of supported nodes)
Scalable operation	Scalable	1 to (number of nodes)	1 to (number of supported nodes)
	High-availability scalable operation	1 to (number of nodes)	2 to (number of supported nodes)

Note

....

Note that failover does not occur when a node is switched off. When designing the system, consider using an uninterruptible power supply (USP).

....

2.3.1 Standby Operation

Standby operation has the following operation modes.

The topologies for standby operation include hot-standby and cold-standby operation.

Hot-standby operation enables preliminary operation so that the operating state can be established immediately on the standby node. In hot-standby operation, the state of the cluster application running on the operating node will be Online, while that of the cluster application

on the standby node will be Standby. To perform hot-standby operation, hot-standby must be supported by the PRIMECLUSTER product to be used, the ISV application, and the user applications.

Cold-standby operation does not allow the preliminary operation needed to establish the operating state immediately on the standby node. In cold-standby operation, the state of the cluster application on the operating node will be Online, while that of the standby node will be Offline.

· ·

1:1 standby

Definition

- It is an operation mode in which a cluster system consists of 2 nodes. One is operating, and the other is standby. When a failure occurs in the operating node, a cluster application switches to the standby node. This does not disrupt ongoing operation.

Advantage

- This operation mode ensures the availability of the cluster application even after failover.

Disadvantage

- Nodes cannot be used efficiently because of a redundant configuration.

Failover image
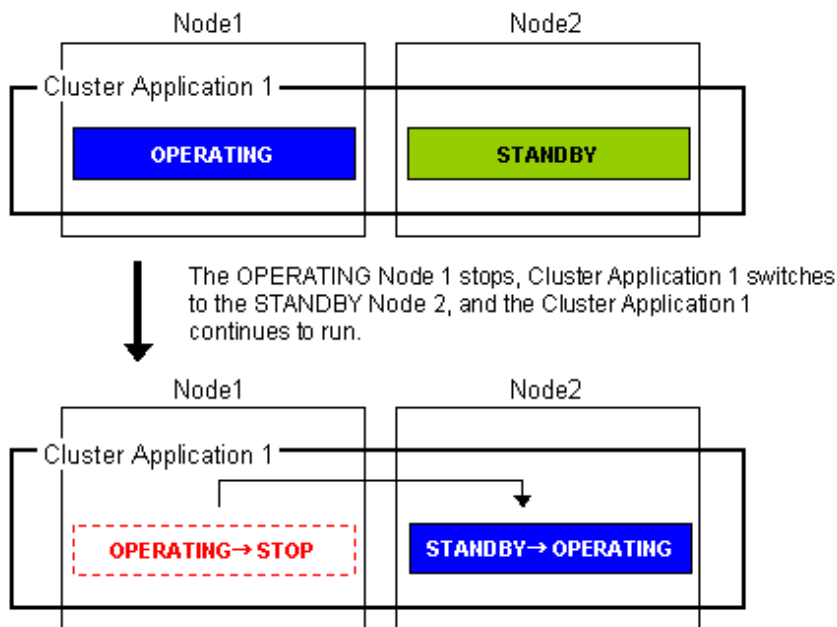
Mutual standby

Definition

- It is an operation mode in which a cluster system consists of 2 or more nodes. Normally, 2 nodes are used in this operation mode. Each node has one operating and one standby application. The operating cluster application has its own standby in each other's node.

Advantages

- Since all nodes are used for cluster application operation, the nodes can be used efficiently.

Disadvantages

- If failover occurs for any of the cluster applications, the performance of the cluster applications may drop because two or more cluster applications will be operating in the failover node. For this operation mode, you need to estimate adequate resources.

Failover image
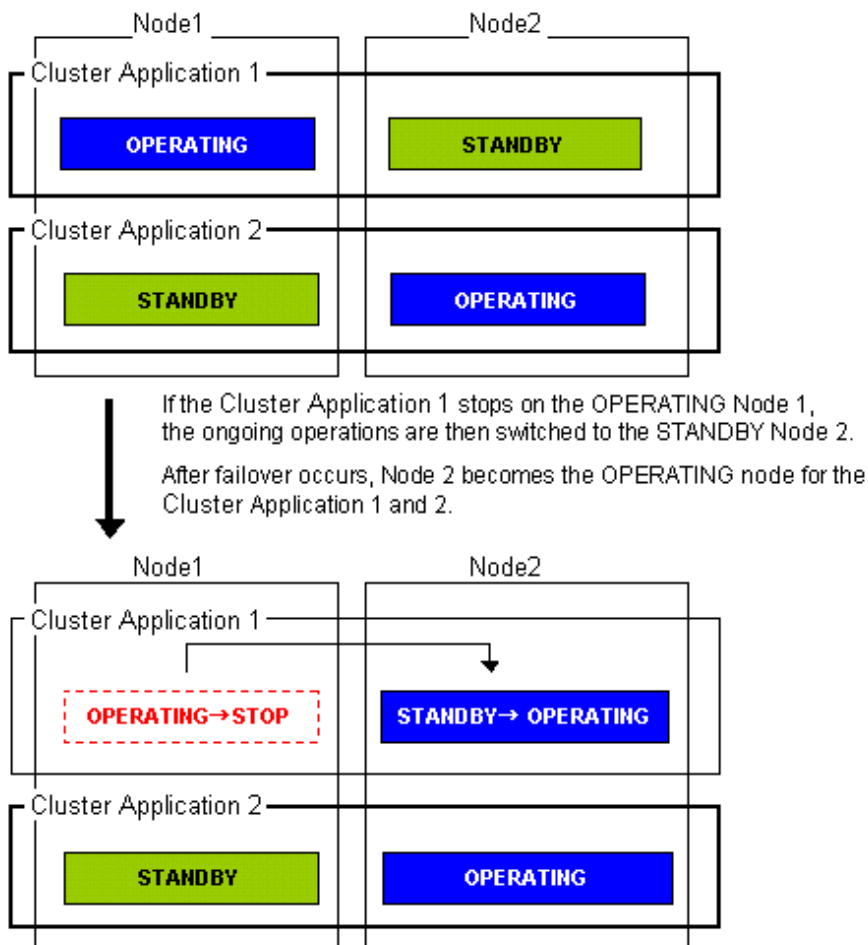
N:1 standby

Definition

- It is an operation mode in which a cluster system consists of 3 or more nodes. One is standby, and the others are operating. When a failure occurs in one of the operating nodes, a cluster application switches to the standby node. If a failure occurs in two or more operating nodes at the same time, the cluster applications switch to the standby node.

Advantages

- This operation mode ensures the availability of the cluster application even after failover.

- Since one node serves as the STANDBY node for multiple applications, the STANDBY cost can be reduced when the number of applications is large.

Disadvantages

- If failover occurs for multiple applications, the performance of the applications is reduced because multiple applications will be operating in one node.

Failover image

Node1   Node2   Node3

Cluster Application 1

OPERATING   STANDBY

Cluster application2

STANDBY   OPERATING

The OPERATING Node 1 stops, Cluster Application 1 switches to the STANDBY Node 2, and the Cluster Application 1 continues to run.

Node1   Node2   Node3

Cluster Application 1

OPERATING→STOP   STANDBY→ OPERATING

Cluster application2
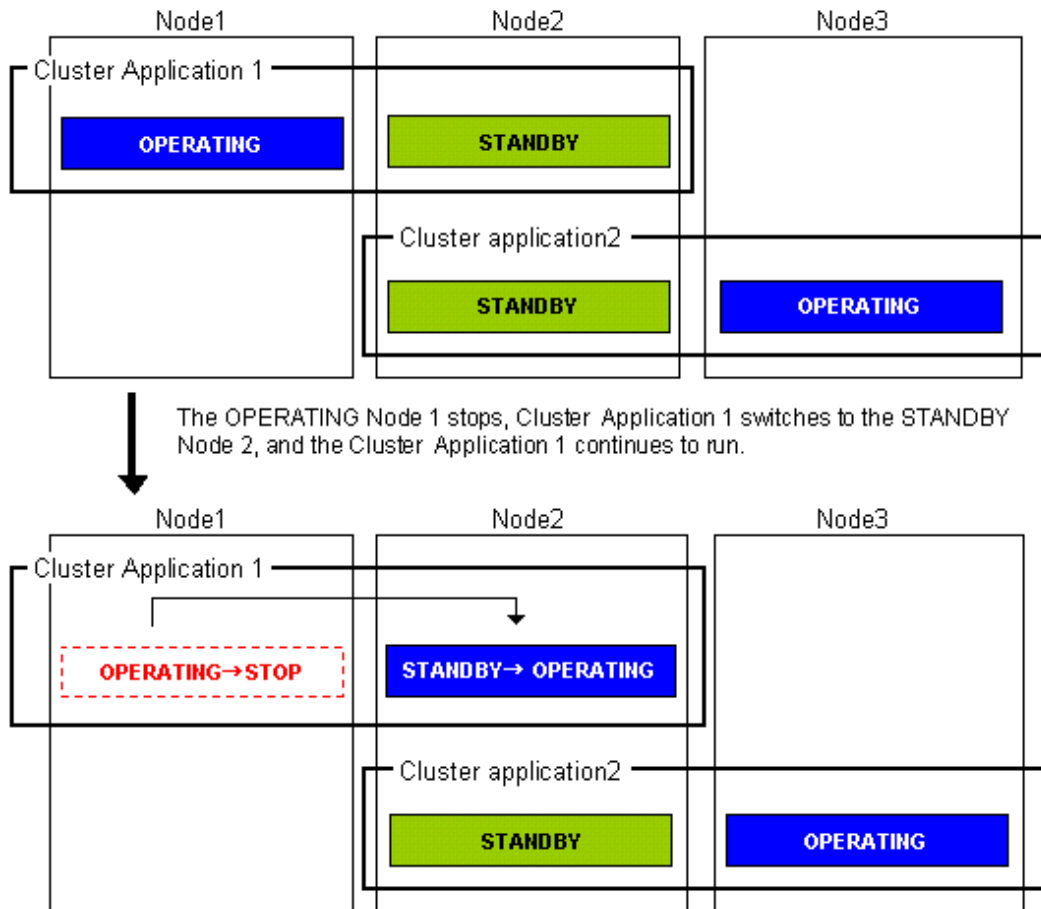
STANDBY   OPERATING

Cascade (using one cluster application)

Definition

- It is an operation mode in which a cluster system consists of 3 or more nodes: one is operating, and the others are standby. When a failure occurs in the operating node, a cluster application switches to one of the standby nodes. When a failover is even failed, this application switches to other standby node.

Advantages

- Even after one node is stopped, the redundant configuration of the cluster application can be maintained by using other nodes. The availability is guaranteed during system maintenance.

- This operation mode ensures the availability of cluster applications even after failover.

Disadvantages

- As the system has a redundant configuration, nodes cannot normally be used efficiently.

Failover image

In this example, the nodes are defined in the sequence Node 1, Node 2, and Node 3 starting from the node with the highest cluster application priority. These nodes are defined when the cluster application is set up.

Node1	Node2	Node3
Cluster Application 1		
OPERATING	STANDBY1	STANDBY2

If the OPERATING Node 1 stops, Cluster Application 1 switches to the STANDBY node with the highest application priority (Node 2 in the above diagram) . The Cluster Application 1 continues to run.

Node1	Node2	Node3
Cluster Application 1		
OPERATING→ STOP	STANDBY 1→ OPERATING	STANDBY2

If the OPERATING Node 2 stops, Cluster Application 1 switches to the STANDBY node with the highest application priority (Node 3 in the above diagram) . The Cluster Application 1 continues to run.

Node1	Node2	Node3
Cluster Application 1		
STOP	OPERATING → STOP	STANDBY2→ OPERATING

Priority transfer (application of N:1 standby)

Definition

- A single node functions as STANDBY for multiple cluster applications. For the other nodes, a single cluster application functions as OPERATING for every node of the other nodes while the other multiple cluster applications function as STOP.

- This topology uses the exclusivity function between cascade and cluster applications.

Advantages

- On that node on which a single cluster application is OPERATING, the other cluster applications do not become either OPERATING or STANDBY. Therefore, the throughput of that cluster application is guaranteed even after failover occurs.

- Because failback of the cluster application is not necessary during the restoration of a cluster application, a job can also be continued during the restoration.

- Since a single node is used as STANDBY exclusively for multiple cluster applications, the cost incurred for standby can be saved when there are many cluster applications.

Disadvantage

- Since a single node is used as STANDBY of multiple cluster applications, availability decreases when there are many cluster applications.

- If a failover occurs due to the occurrence of an error on a single node, the availability decreases because no standby node is available until the completion of the maintenance work.

Failover image

If a failure occurs in Cluster Application 1 on Node 1, failover occurs and Cluster Application 1 switches to STANDBY Node 3 to on which it continues operating.

When Cluster Application 1 enters the operating state after failover to Node 3, Cluster Application 2 stops on Node 3.

After restoring the failed cluster application on Node 1, the operator must switch each cluster application to STANDBY.

N:M standby (application of N:1 standby)

Definition

- Multiple nodes exclusively for standby server act as STANDBY for multiple cluster applications.

- For the nodes other than those that are used exclusively for standby, each of the nodes is used for a single OPERATING cluster application and multiple STOP cluster applications.

- This topology uses the exclusivity function between cascade and cluster applications.

Advantages

- On that node on which a single cluster application is in the OPERATING state, other cluster applications do not either become OPERATING or STANDBY. Therefore, the throughput of that cluster application is guaranteed even after failover occurs.

- Because failback of the cluster application is unnecessary during the restoration of a cluster application, a job can be also continued during the restoration.

- Since multiple nodes serve as STANDBY for multiple cluster applications, availability can be maintained even when there are many cluster applications is many.

Disadvantage

- Because a redundant configuration is used, the node usage efficiency is poor during normal operation.

Failover image

If a failure occurs in Cluster Application 2 on Node 2 before the recovery of a failure in Cluster Application 1, failover occurs and Cluster Application 2 switches to Node 4, where Cluster Application 1, that has an exclusive relationship, is not OPERATING. At this time, STANDBY of Cluster Application 1 that has an exclusive relationship stops.

	Node 1	Node 2	Node 3	Node 4	
Cluster Application 1	FAILURE	STOP	OPERATING	STANDBY→ STOP	Exclusive relationship between cluster applications
Cluster Application 2	STOP	OPERATING→ FAILURE	STOP	STANDBY→O PERATING	

2.3.2 Scalable Operation

This section explains the topologies used for scalable operation:

Scalable

Definition

- A cluster system consists of two or more operating nodes, and all the nodes are used for online cluster applications. This operation mode is suitable for parallel jobs that use the I/O load balancing and load sharing on a parallel database.

Advantages

- Throughput can be improved for a cluster application by expanding nodes.

- Degenerated operation is possible even if part of the cluster applications stops.

Disadvantage

- If part of the cluster applications stops, throughput of the cluster applications cannot be guaranteed because degenerated operation is assumed.

Failover image

	Node 1	Node 2	Node 3
Cluster Application	OPERATING	OPERATING	OPERATING

If even one of the running cluster applications stops, the cluster applications are degenerated to continue operation.

	Node 1	Node 2	Node 3
Cluster Application	OPERATING→ STOP	OPERATING	OPERATING

High-availability scalable operation

Definition

- Refers to the topology in which standby operation is configured for each cluster application that constitutes scalable operation. Suitable for a parallel database for which scalability and availability are required, as well as parallel job execution for which load share/load balance is used.

- Standby operation that constitutes scalable operation can be combined with 1:1 standby and N:1 standby, with priority transfer.

Advantages

- Even if failover occurs in one of the cluster applications that constitute scalable operation, the throughput of all the cluster applications can be maintained by using a redundant configuration.

- Throughput of the cluster applications can be improved with node expansion.

- Degenerated operation is possible even if part of the cluster applications stops.

Disadvantage

- Since a redundant configuration is used, the node usage efficiency is poor during normal operation.

Failover image

The following illustrates failover when two 1:1 standby operations are combined to enable scalable operation.

| | : scalable operation | | : standby operation |
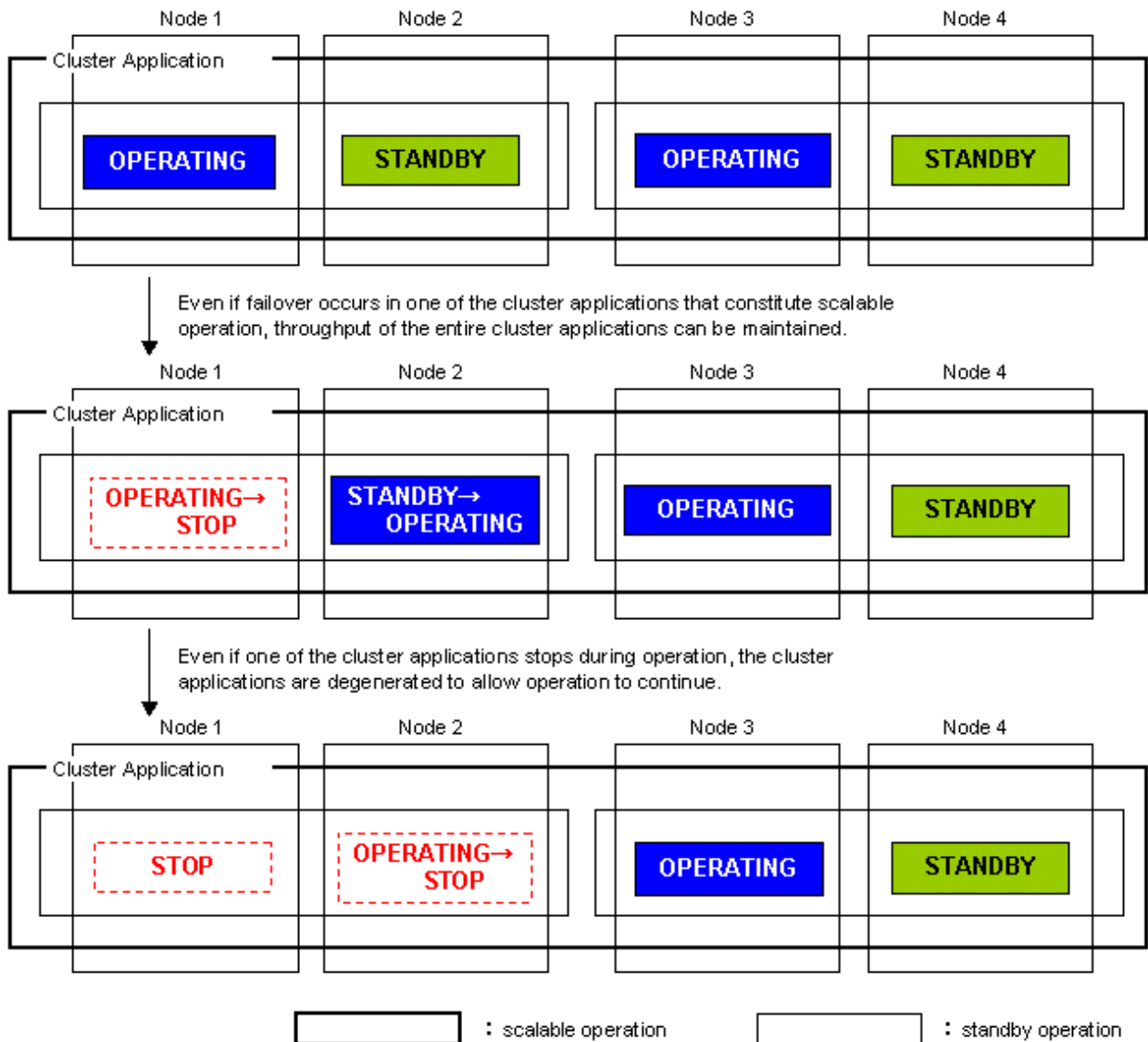
Note

··

High-availability scalable operation can be used in combination with some PRIMECLUSTER-related products. For information on the related products, see the "Table 12.1 PRIMECLUSTER product list."

··

2.4 Determining the Web-Based Admin View Operation Mode

Determine the operation mode of Web-Based Admin View according to your PRIMECLUSTER configuration. Consider system performance and RC2000 that controls the OS console.

This section describes operation modes and typical models of PRIMECLUSTER systems that run Web-Based Admin View, and provides a guideline for adopting models.

See

··

- For information on the operation modes of Web-Based Admin View, see "1.2 Web-Based Admin View topology" in the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

- RC2000 supports integrated management of the operating system consoles of multiple nodes. It is used to log into the console of a node to install software on the node or modify configuration. For details on the RC2000, see the *"RC2000 User's Manual."*

Note

When a cluster contains 3 or more nodes, the 3-tier model, in which the cluster management server is not in the cluster, is recommended.

Roles of individual nodes

Web-Based Admin View adopts a logical 3-tier architecture, which consists of clients, a cluster management server, and cluster nodes. It has the following functions.

Clients

A client is a computer with which a user manages operations. Basically, the computer is a personal computer that uses a Web browser or Solaris Workstation.

Management server

The cluster management server manages cluster operation and features web server functions. The server can be as a cluster node. The cluster management server can be duplexed. In this case the system will have a two-server configuration, consisting of a primary management server and a secondary management server.

Set up both primary and secondary management servers for redundancy.

You can dynamically move the secondary management server depending on the operation mode. The cluster management servers run Solaris.

Cluster nodes

Cluster nodes construct the PRIMECLUSTER system. The cluster nodes run Solaris.

Logical 3-tier architecture and operation models

Web-Based Admin View adopts a logical 3-tier architecture consisting of clients, management servers, and monitored nodes. Physically, the system can adopt a 2-tier architecture.

Typical operation modes that run Web-Based Admin View are introduced below.

2-tier model

In the 2-tier model, the cluster management server and the cluster node are used together on the same machine, and the client is on a machine other than the nodes and the management severs.

This model is used for a configuration where the number of nodes is relatively small.

Normally, this model is used for an operation mode such as 1:1 standby or mutual standby in a 2-node configuration.

If the RC2000 is used with this model, you need to adopt local operation.

In the local operation, RC2000 supports integrated management of the OS consoles of multiple nodes without linking to the cluster management server.

See

For details on the RC2000 operation mode, see the *"RC2000 User's Guide."*

This model supports 2 types of topology, which are described below.

Topology where a network is shared

In this topology, the public LAN and the LAN that is connected to the management client are used together. You can adopt this topology if the network users and network range are limited for security. This is the default Web-Based Admin View configuration after PRIMECLUSTER installation.
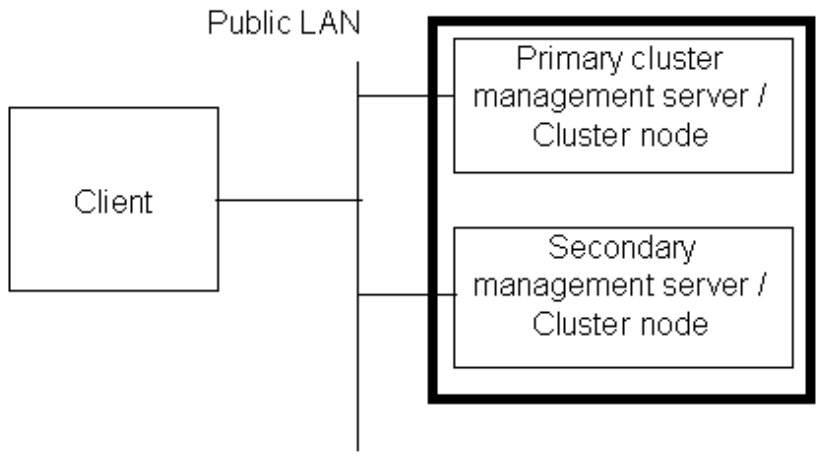
The cluster system is working in [   ]

Topology where separate LANs are used

In this topology, the public LAN and the LAN that is connected to the management client are separate. When using a management client from a public network, this topology is recommended for security. After the PRIMECLUSTER installation is done, you will need to modify the Web-Based Admin View configuration.

Specify IP addresses used for a cluster node and a client respectively. For details, see "5.1.1 Setting Up CF and CIP."
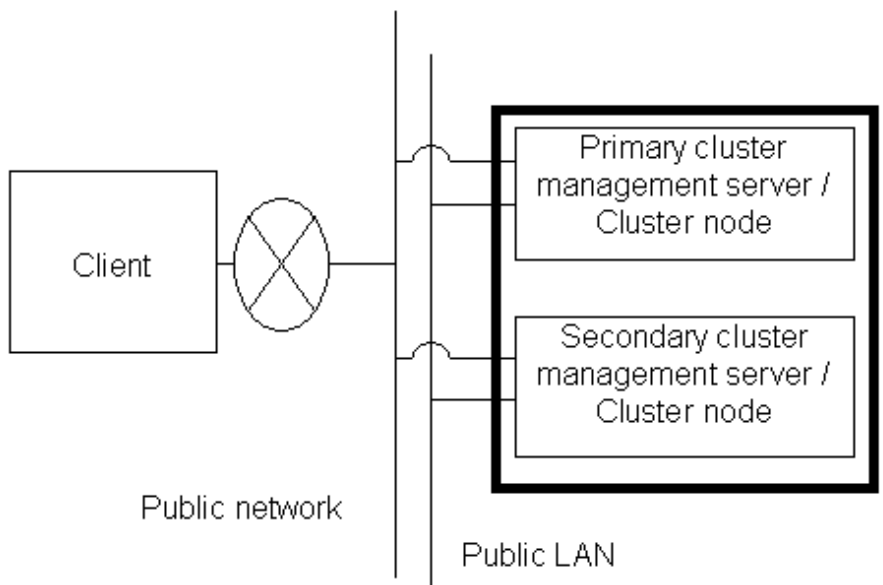
The cluster system is working in [   ]

3-tier model

In the 3-tier model, clients, cluster management servers, and cluster nodes are set up separately.

This model is adopted for configurations where the number of nodes is relatively large.

Normally, this model is used for integrated management of the PRIMECLUSTER system. You can also use this mode when you do not want to impose the load of running the management server on the cluster node or when you want to perform the integrated management of the PRIMECLUSTER system.

To maintain operation continuity and availability, Fujitsu recommends this 3-tire model for a configuration with 3 or more nodes.

When the RC2000 is used with this model, you need to adopt server/client operation.

In the server/client operation, RC2000 supports integrated management of the OS consoles of multiple nodes, and can be operated on the cluster management server as well as on the client.

See

....
For details on the RC2000 operation modes, see the *"RC2000 User's Manual."*
....

This model supports 2 types of topology, which are described below.

Topology where a network is shared

In this topology, the public LAN and the LAN that is connected to the management client are the same. You can adopt this topology if the network users and network range are limited for security. This is the default Web-Based Admin View configuration after PRIMECLUSTER installation.
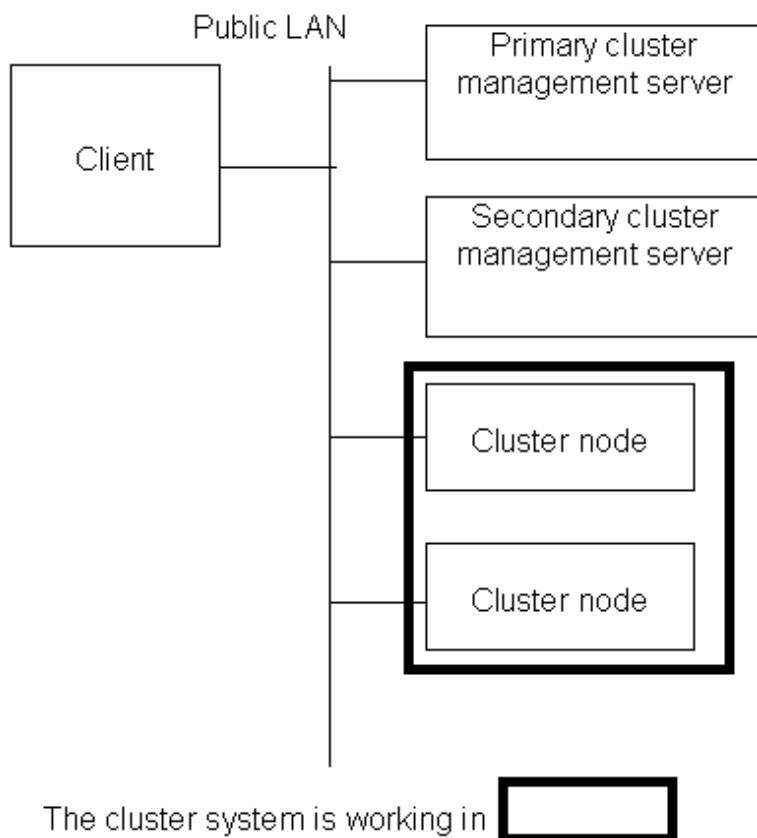
Topology where separate LANs are used

In this topology, the public LAN and the LAN that is connected to the management client are separate. When using a management client from a public network, this topology is recommended for security. After PRIMECLUSTER installation is done, you will need to modify the Web-Based Admin View configuration.

Specify IP addresses used for a cluster node and a client respectively. For details, see "5.1.1 Setting Up CF and CIP."

The cluster system is working in [　　　]
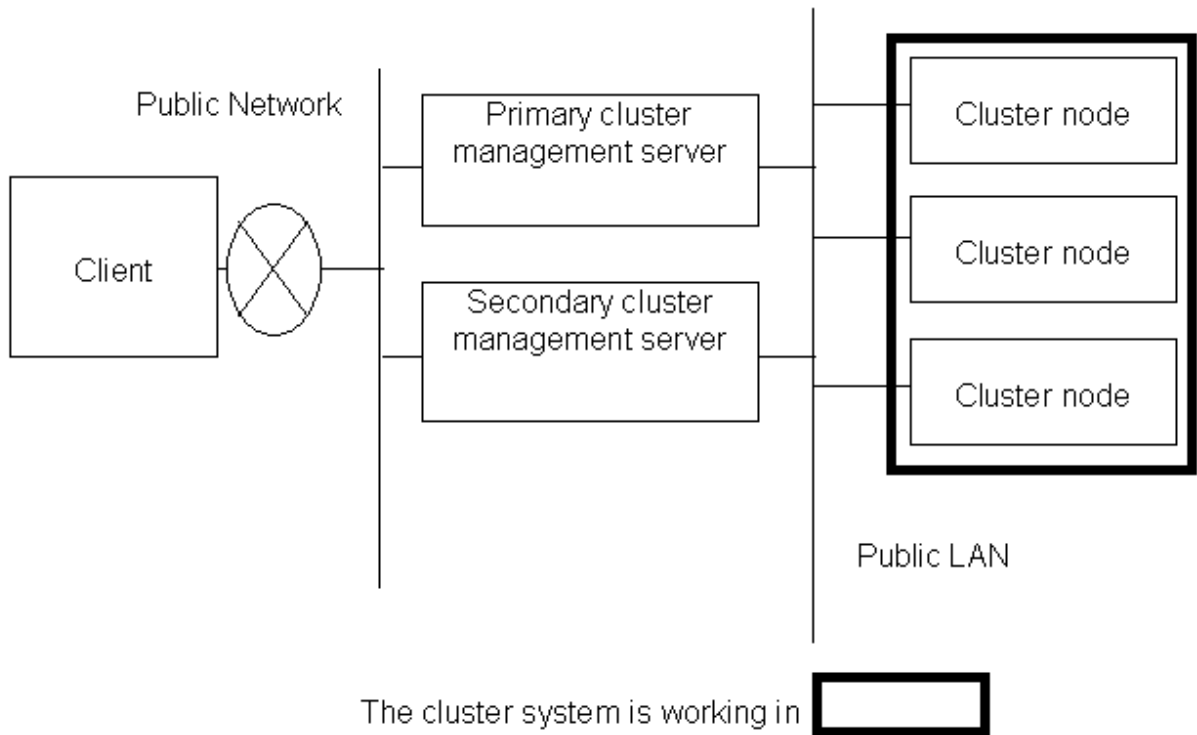
2.5 Setting the Failover Timing of a Cluster Application

Determine the trigger for cluster application failover. You can choose from the followings:

1. The cluster application does not automatically switch to other host.

2. The cluster application switches to other host in the event of a node failure or a resource failure.

3. The cluster application switches to other host in the event of a node failure, a resource failure, or RMS shutdown.

🔖 See

The failover timing is set in "6.6.2 Creating Cluster Applications."

Part 2 Installation

This part describes procedures for installing the PRIMECLUSTER system and running Web-Based Admin View.

The operations include the procedures up to installing a new PRIMECLUSTER system.

For procedures on changing the PRIMECLUSTER system configuration after the system is installed, see "Chapter 8 Changing the Operation Configuration".

Chapter 3 Software Installation

Install Solaris and other software products related to PRIMECLUSTER on each node.

This chapter describes the following topics:

- PRIMECLUSTER installation (overview)

- Installation and setup of related software

- Application installation and environment setup

3.1 PRIMECLUSTER Installation

There are 2 methods with which PRIMECLUSTER Enterprise Edition, HA server, Parallel Server, or Scalability Server can be installed, and those are One Shot Installer and Installation script:

- One Shot Installer

  The One Shot Installer tool is used to set up Solaris custom JumpStart. Using this JumpStart, you can also install Solaris software, PRIMECLUSTER, and related software products on multiple nodes at the same time.

- Installation script

  The installation script is also called the CLI Installer. It is used to install PRIMECLUSTER on a system in which Solaris software and related Fujitsu software have been installed. This method is also used for the installation of the cluster management server.

![See icon] **See**

....
For details on the installation procedures, see the *"PRIMECLUSTER Installation Guide."*
....

![Note icon] **Note**

....
When using Oracle Solaris Zones environments

When Oracle Solaris Zones of an exclusive IP zone exist on the system, perform the settings until "5.1.1 Setting Up CF and CIP" before booting a non-global zone.

If you execute a boot command of Oracle Solaris Zones while the settings of CF and CIP have not been done, the command will terminate abnormally.

[Output example]

```
# zoneadm -z <zone_name> boot
zone '<zone_name>': WARNING: unable to add network interface '<NIC_name>': link busy
zone '<zone_name>': failed to add network device
zoneadm: zone '<zone_name>': call to zoneadmd failed
```

....

3.2 Installation and Setup of Related Software

After installation of the software products related to PRIMECLUSTER is done, you need to set up the OS and hardware for installing and operating PRIMECLUSTER.

Set up the followings if necessary.
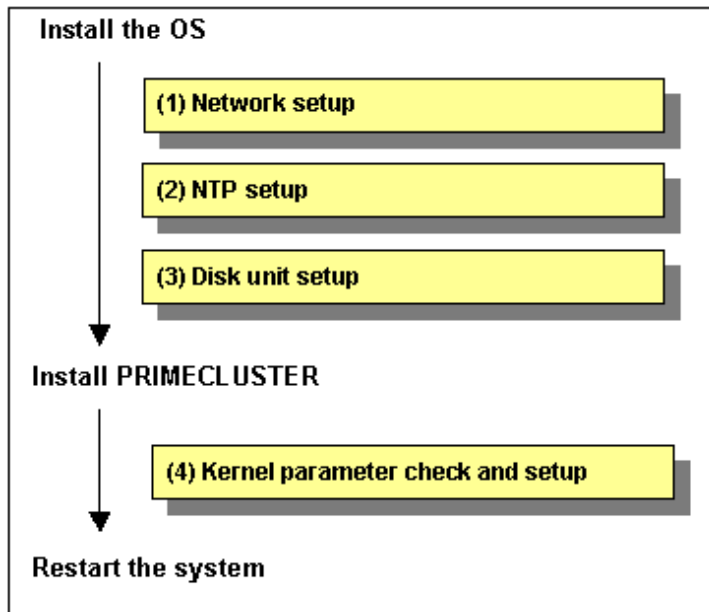
- Network setup

  If a network adapter for a public LAN has been installed, you need to configure the IP addresses after installing the OS.

  Perform this setup after installing the operating system.

  For details, see "3.2.1 Setting Up the Network".

- NTP setup

  This setup synchronizes the time in all of the nodes of the cluster system. This NTP setup needs to be done before installing PRIMECLUSTER.

- Disk unit setup

  When using a shared disk unit, you need to install and set up the related software product. This disk unit setup needs to be done before installing PRIMECLUSTER.

  For details, see "3.2.2 Setting Up Disk Units."

- Kernel parameter check and setup

  When operating the software products related to PRIMECLUSTER, you need to adjust kernel parameters according to the hardware configuration and applications. This kernel parameter setup should be done before operating PRIMECLUSTER.

  For details, see "3.2.3 Checking the Kernel Parameters."

3.2.1 Setting Up the Network

If a network adapter for a public LAN has been installed, the IP address setup is required.

You need to create the "/etc/hostname.*interface*" file for the interface.

See
. . . .
- For details, see the "Solaris X AnswerBook2 System Administrator Collection."

- For information on changing the public LAN that the PRIMECLUSTER system uses, see "8.2 Changing an IP Address on the Public LAN".
. . . .

. . . .

Web-Based Admin View automatically sets up an interface that was assigned the IP address of the host name corresponding to the node on which PRIMECLUSTER was installed. This interface will be used as a transmission path between cluster nodes and cluster management server, and between cluster management servers and clients.

. . . .

3.2.2 Setting Up Disk Units

Installation and Setup of Related Software

Install and set up the software products required for using shared disk units. See the installation guides and manuals of each software product.

 Point
. . . .

- If the disk unit uses a Fibre Channel interface, a Fibre Channel driver for the card needs to be installed and set up. For installation details, see "Installation Procedures" in *"FUJITSU PCI Fibre Channel x.x Guide for Solaris(TM) Environment."*

- If the disk array unit is used with a multipath configuration, refer to the following manuals for instructions on installing and setting up the software.

    - When the ETERNUS multipath driver (GRMPD) is used

        - "*ETERNUS Multipath Driver User's Guide*"

    - When the multipath disk control load balance option (MPLB) is used

        - *"INSTALLATION GUIDE - Multipath Disk Control load balance option x.x for Solaris(TM) Environment"*

        - *"Multipath Disk Control Load Balance option x.x Guide"*

    - When multipath disk control (MPHD) is used

        - "INSTALLATION GUIDE - Multipath Disk Control x.x for Solaris(TM) Environment "

        - *"Multipath Disk Control x,x Guide"*

- If an ETERNUS2000, ETERNUS3000, ETERNUS4000, ETERNUS6000, or ETERNUS8000 disk array is used with a single-path configuration, you need to install and setup a hard disk driver (HDDV).
  For details on installing HDDV, see the *"Hard Disk Driver x.x Guide."*

. . . .

Preparation for automatic shared disk configuration

Multipath disk

Before executing automatic shared disk configuration, format the logical units and assign labels to the units.

 See
. . . .

- For information on the procedures, refer to the manual ("Multipath Disk Control x,x Guide," "Multipath Disk Control load balance option," or "ETERNUS Multipath Driver User's Guide") of the software product to be used.

- For information on automatic configuration of shared disks, see "5.1.3 Initial Setup of the Cluster Resource Management Facility" and "6.2.1 Automatic Configuration of Shared Disks."

. . . .

 Note
. . . .

At this stage, do not create logical paths for MPHD, MPLB, and GRMPD.

. . . .

If a shared disk is automatically configured in a cluster system where MPLB or ETERNUS multipath driver is installed, and the disk supports the load balancing, the MPLB logical path will automatically be generated, then the resource will also be registered.

If you want to make the MPHD logical path automatically generate and the MPHD resource register by performing automatic configuration of a shared disk, it is necessary to set the parameter beforehand using the clsetacfparam(1M) command.

See

For details on the parameters to be set with the clsetacfparam(1M) command, see "5.3 Setting Up Shared Disk Connection Confirmation."

If the disk is not a multipath disk

Before executing automatic shared disk configuration, you need to assign a label to the disk by executing the "format(1M)" command.

See

For information on automatic configuration of shared disks, see "5.1.3 Initial Setup of the Cluster Resource Management Facility" and "6.2.1 Automatic Configuration of Shared Disks."

3.2.3 Checking the Kernel Parameters

You need to edit the values for the kernel parameters according to the environment.

Target node:

All nodes in which PRIMECLUSTER is to be installed

The kernel parameters differ according to the products and components to be used.

Check the Kernel Parameter Worksheet, and edit the value if necessary.

See

For information on the kernel parameters, see the "A.5 Kernel Parameter Worksheet".

Note

To enable the kernel parameter, it is necessary to reboot the system after the parameter setting.

3.3 Installation and Environment Setup of Applications

Install software products to be operated on the PRIMECLUSTER system and configure the environment as necessary.

To bring about application switchover in the event of a failure, you need to register the resources of software application to RMS. RMS will monitor these resources. For details, see "Chapter 6 Building Cluster Applications".

See

- For information on products supported by the PRIMECLUSTER system, see "Table 12.1 PRIMECLUSTER product list".

- For details on installing applications, see the manuals and installation guides for the individual applications.

Chapter 4 Preparation Prior to Building a Cluster

This chapter explains the preparation work that is required prior to building a cluster, such as starting up the Web-Based Admin View screen.

See

. . . . .

As preparation for building the cluster, check the operation environment. See "3 Operating Environment" in the "PRIMECLUSTER Installation Guide."

. . . . .

(1) Checking the Cluster Worksheet

↓

(2) Setting Up the Operating System Console

↓

(3) Preparations for Starting the Web-Based Admin View Screen

— Assigning Users to Manage the Cluster

— Preparing the Client Environment

— Initial Setup of Web-Based Admin View

— Setting Up the Browser

— Setting Up the Java Plug-in

↓

(4) Starting the Web-Based Admin View Screen
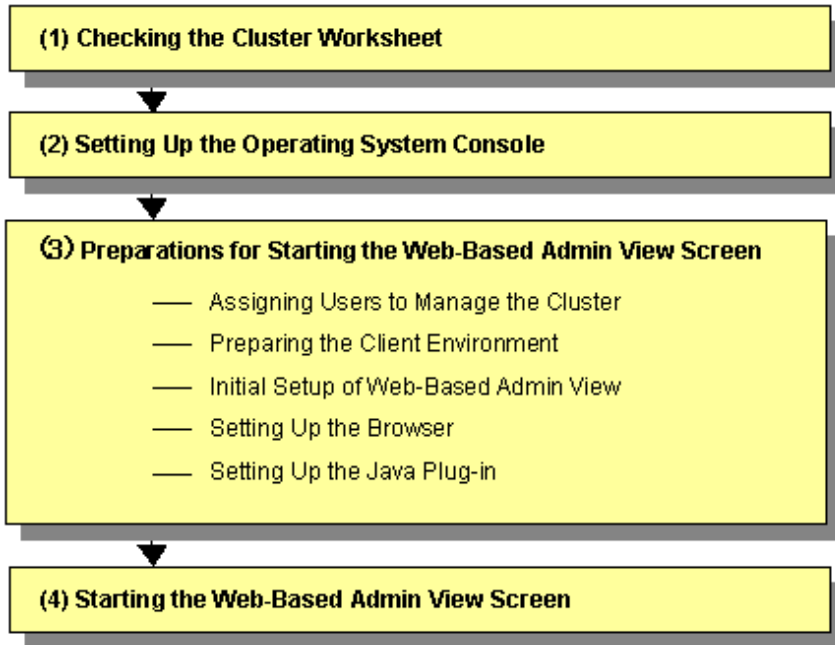
Table 4.1 Operation procedure and manual reference location for starting the Web-Based Admin View screen

	Work item	Execution Node	Required/ Optional	Manual reference location*
(1)	4.1 Checking the Cluster Worksheet	-	Required	This manual: Appendix A
(2)	4.2 Setting Up the Operating System Console	All nodes	Optional	
(3)	4.3.1 Assigning Users to Manage the Cluster	Cluster node	Required	
	4.3.2 Preparing the Client Environment	Client	Required	**WEB** "3.1.2 Prerequisite client environment"
	4.3.3 Initial Setup of Web-Based Admin View	Cluster node	Required	
	4.3.4 Setting Up the Browser	Client	Required	**WEB** "3.1.3.1 Preparing the Web browser"
	4.3.5 Setting Up the Java Plug-in	Client	Required	**WEB** "3.1.3.2 Required for the Web Browser Environment"
(4)	4.4 Starting the Web-Based Admin View Screen	Client	Required	**WEB** "3.2 Screen startup"

* The PRIMECLUSTER manual name is abbreviated as follows:

4.1 Checking the Cluster Worksheet

Make certain of filling the worksheet. If there is missing information, you should specify values and complete the sheet.

See

For information on the worksheet, see "Appendix A PRIMECLUSTER System Design Worksheets".

4.2 Setting Up the Operating System Console

In the PRIMECLUSTER system, it is recommended that XON/XOFF flow control of the operating system console be disenabled.

In the environment immediately after the Solaris is installed, XON/XOFF flow control is enabled for console messages that are output from the operating system.

Although this setting allows you to stop screen display by pressing the [*Ctrl*] + [*S*], it might affect the system by disturbing ongoing operations in the event of a console failure.

To avoid this, you need to disable XON/XOFF flow control for output from the server to the console terminal.

Change the settings in the server and the terminal (remote console connection unit) according to the procedure described below.

Operation Procedure:

1. Setting up the server

    Take the following steps on all servers in the cluster system.

    1. Log in using a system administrator access privilege.

    2. Edit the "/etc/ttydefs" file.

        Change all lines that begin with "console:" to the following:

        ```
        console:9600 hupcl opost onlcr -ixon:9600::console
        ```

    3. Restart the machine.

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

2. Setting the remote console connection unit

    If a remote console connection unit (RCCU) is connected to each server in the cluster system, perform the following.

    1. Log into the RCCU.

        ```
        % telnet <RCCU-IP-address> 8010    <-- Enter "telnet RCCU-IP-address 8010"
                                                and press the [Return] key
        Trying <RCCU-IP-address>...
        Connected to <RCCU-IP-address>.
        Escape character is '^]'.

        login :somebody <-- Enter "somebody" and press the [Return] key
        passwd: <-- Press the [Return] key
        cs >su <-- Enter "su" and press the [Return] key
        passwd: <-- Press the [Return] key
        ```

    2. Edit the rs232c.def file.

        ```
        # edit rs232c.def <-- Enter "edit rs232c.def" and press the [Return] key

        0001 #
        Press the [N] key 5 times.
        ```

```
            0002 baud 9600
            0003 parity none
            0004 stop 1
            0005 bitchar 8
            0006 flow xon
            Press the [C] key.
            0006> flow xon
            0006< flow rxon <-- Enter "flow rxon" and press the [Return] key
            0006 flow rxon
            Press the [E] key.
            # write <-- Enter "write" and press the [Return] key
            writing configuration to flash ROM . done.
```

   3. Reboot the RCCU.

```
            # reboot              <-- Enter "reboot" and press the [Return] key
            Do you really want to reboot [y/n] ? y   <-- Enter "y" and
                                                         press the [Return] key
```

3. Confirmation of modified settings

Log in again to the console of each server. While executing the command such as "ls -l", press [Ctrl] + [S] and check that the output does not stop.
If the data output stops, review the settings for Step 1 and Step 2 again.

4.3 Preparations for Starting the Web-Based Admin View Screen

Take the following steps for starting the GUI screen of Web-Based Admin View:

1. Environment setup

You can set up the following in any order:

- 4.3.1 Assigning Users to Manage the Cluster

- 4.3.2 Preparing the Client Environment

- 4.3.3 Initial Setup of Web-Based Admin View

2. Web environment preparation

You need to set up the browser first.

- 4.3.4 Setting Up the Browser

- 4.3.5 Setting Up the Java Plug-in

4.3.1 Assigning Users to Manage the Cluster

Web-Based Admin View restricts access to specific operation management GUIs by using user groups in the management server.

The table below shows the groups used for operation management GUIs of PRIMECLUSTER.

Table 4.2 Operation management GUIs of Web-Based Admin View and authorized user groups

GUI name	UNIX user group name	Privileges
All GUIs	wvroot	Root authority. This group can execute all operations.
Cluster Admin	clroot	Root authority. This group can specify settings, execute management commands, and display information.
	cladmin	Administrator authority. This group cannot specify settings. It can execute management commands and display information.

GUI name	UNIX user group name	Privileges
	clmon	User authority. This group cannot specify settings and cannot execute management commands. It can only display information.
GDS (Global Disk Services)	sdxroot	Root authority. This group can use the GDS management view.

The groups for the operation management GUIs are defined as shown in the above table.

wvroot is a special user group, and is used for Web-Based Admin View and GUIs. Users belonging to this group are granted the highest access privileges for Web-Based Admin View and all kinds of operation management GUIs.

The system administrator can grant different access privileges to users according to the products that the users need to use.

For example, a user who belongs to the "clroot" group but not to "sdxroot" is granted all access privileges when opening the Cluster Admin screen but no access privileges when opening the Global Disk Services (GDS) GUIs.

You need to create the above UNIX groups for both the primary and secondary management servers as necessary, and assign users to these groups. The Web-Based Admin View group membership should maintain consistency among all management servers associated with a specific cluster system. To set a user group, execute the "useradd(1M)" command as follows:

```
# useradd -g wvroot usrname
```

The root user is granted the highest access privilege regardless of which group the root user belongs to.

For details about user groups, see "3.1.1 User group determination" in the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

4.3.2 Preparing the Client Environment

Prepare hardware, operating systems, and Web browsers of the clients supported by Web-Based Admin View.

PC client

- CPU

  Equivalent to Pentium II 233 MHz or faster

- Memory

  128 MB or more installed

- Operating system

  Windows(R) 98, Windows NT(R), Windows(R) 2000, Windows(R) Me, Windows(R) XP, Windows(R) Vista

 Note
. . . .
There are restrictions on the Japanese display environment. For details, see "Notes" below.
. . . .

- Web browser

  Netscape(R) Communicator 4.5 to 4.7X, Netscape(R) 6.2, Netscape(R) 7.X, Microsoft(R) Internet Explorer 5.0 or later (Java Plug-in is required.)
  (Netscape(R) 7 is not supported.)

- Java Plug-in

  Java(TM) 2 Runtime Environment Standard Edition Version 1.3.1
  Java(TM) 2 Runtime Environment Standard Edition Version 1.4.2
  Java(TM) 2 Platform Standard Edition Runtime Environment Version 5.0

- Display unit

  Unit that supports 800 x 600 pixels or more and can display High Color (16 bits, 65,536 colors).

Solaris Workstation

- Memory

  128 MB or more installed (256 MB or more are recommended if multiple browsers are to be started or if an individual operation management screen is to be started more than once concurrently)

- OS

  Solaris 9 or Solaris 10

- Web browser

  Netscape Communicator 4.72 - 4.7X, Netscape 6.2, Netscape 7.X
  Mozilla Web browser (Java bundled in Solaris)

- Java Plug-in

  10.1.2.2,REV=1999.12.06.10.02 or higher
  20.1.2.2,REV=1999.10.14.18.04 or higher
  Java(TM) 2 Runtime Environment Standard Edition Version 1.3.1
  Java(TM) 2 Runtime Environment Standard Edition Version 1.4.2

📌 Note
....

If you use Solaris10 as a client OS, be sure to use the Mozilla Web browser and Java(TM) 2 Runtime Environment, Standard Edition Version 1.4.2.
....

The System Management Console that comes with PRIMEPOWER 800/900/1000/1500/2000/2500 can also be used as a client. However, the above OS must be used.

📌 Note
....

- For Japanese to be displayed in Web-Based Admin View, the operating system environment must satisfy the following conditions:

  - Japanese Windows(R) is installed, and the language environment being used by Windows(R) is set to Japanese.

  - The Japanese package is installed in the Solaris Workstation, and the language environment to be used is set to Japanese.

- For details on the Java Plug-in settings, see "3.1.3.2 Required for the Web Browser Environment" in *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

- For details on the operation environments of the Web-Based Admin View management server and the monitoring nodes, see "2.1 Prerequisite server environment" in the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

- If you are using a 3-layer configuration and operating Web-Based Admin View separately from the cluster system, the operating system of the server in which Web-Based Admin View and the GUI packages are installed must be Solaris 9.

- In clients that are supported by Web-Based Admin View, the following color depth problems may exist.

  - **Solaris**

    When Sun's PGX32 PCI graphics card is set to 8-bit color, the GUI screen of Web-Based Admin View is sometimes not displayed correctly. This problem is present if the password field and other GUI fields are displayed as a single-color blue block when the GUI screen of Web-Based Admin View is displayed. If the problem exists, use the following commands to check the color depth of the display:

    $ `xwininfo`

    The color depth of the window is displayed in the Depth output parameter.

    If the GUI screen of Web-Based Admin View is not displayed correctly and the color depth is 8 bits, you can use the pgxconfig command to set the color depth to 24 bits (Use the "24" bit color depth setting, not the "8 &24" bit setting.)

    # pgxconfig -i

For information on using the pgxconfig command, see the Sun manual *"Sun PGX32 PCI Graphics Card Installation Guide."*

- **Windows(R)**

When the GUI screen of Web-Based Admin View is displayed with Windows(R), the 16-bit or 32-bit color mode must be used. The 24-bit color mode does not properly support an alpha component (transparency) that is required for RMS graphs. If 24-bit color is used, RMS graphs may not be displayed properly.

Whether problems will occur in 24-bit color mode depends on the graphics card and video driver being used.

4.3.3 Initial Setup of Web-Based Admin View

4.3.3.1 Initial setup of the operation management server

When using Web-Based Admin View for the first time, you need to initialize the management server on each node. Take the following steps in the order listed below.

🔷 **Note**

You do not have to initialize the server on the node on which PRIMECLUSTER was installed using the One Shot Installer tool

Operation Procedure:

1. Stop Web-Based Admin View on all the management servers and nodes.

```
# /etc/init.d/fjsvwvcnf stop
# /etc/init.d/fjsvwvbs stop
```

2. Set the IP addresses of the primary management server and secondary management server.
   Execute the following command on all the management servers and nodes, referring to the example.

```
# /etc/opt/FJSVwvbs/etc/bin/wvSetparam primary-server <primary-management-server-
IP-address>
# /etc/opt/FJSVwvbs/etc/bin/wvSetparam secondary-server <secondary-management-
server-IP-address>
```

Example: Set the "primary management server IP" and "secondary management server IP" of the "operation management mode" found on the Cluster Configuration Worksheet.

```
# /etc/opt/FJSVwvbs/etc/bin/wvSetparam primary-server 10.34.214.181
# /etc/opt/FJSVwvbs/etc/bin/wvSetparam secondary-server 10.34.214.182
```

3. Restart Web-Based Admin View on all the management servers and nodes.

```
# /etc/opt/FJSVwvbs/etc/bin/wvCntl restart
# /etc/init.d/fjsvwvcnf start
```

🔷 **See**

Web-Based Admin View has some different operation management modes. For further details, see "1.2.2 System topology" and "Chapter 7 Web-Based Admin View setup modification" in the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

4.3.3.2 Confirming Web-Based Admin View Startup

This section describes the procedure for confirming whether Web-Based Admin View has been started.

Confirmation Procedure

Check that all node information is output by executing the "wvstat(1M)" command on the connected management server.

Example:

In a two-node configuration consisting of node1(10.20.30.40) and node2(10.20.30.41), node1 is the primary management server and node2 is the secondary management server.

```
# /etc/opt/FJSVwvbs/etc/bin/wvstat
primaryServer  10.20.30.40 node1  http=10.20.30.40   Run  1d+7h36m
primaryServer    Sessions: 0
primaryServer    Nodes: 2
        10.20.30.40 node1   SunOS-5.8        1d+7h36m
        10.20.30.41 node2   SunOS-5.8        1d+7h36m
secondaryServer 10.20.30.41 node2  http=10.20.30.41   Run  1d+7h36m
secondaryServer Sessions: 0
secondaryServer Nodes: 2
        10.20.30.41 node2   SunOS-5.8        1d+7h36m
        10.20.30.40 node1   SunOS-5.8        1d+7h36m
#
```

If the information is not properly displayed, Web-Based Admin View has not been started or there may be an error in the Web-Based Admin View settings. Restart Web-Based Admin View and execute the operation again. If node information is still not displayed, refer to the *"PRIMECLUSTER Web-Based Admin View Operation Guide"* and check the parameter settings.

For details on the wvstat(1M) command, see the manual page.

4.3.3.3 Setting the Web-Based Admin View Language

The language environment in which Web-Based Admin View operates is set to English as default. Even though the client has a Japanese environment, the text of cluster resource management facility messages that are sent from the cluster node is displayed in English.

If you want to display the messages in Japanese, take the following steps to set up environment variables of Web-Based Admin View. You need to set up the variables using a system administrator access privilege on all nodes and the cluster management servers that construct a cluster system.

This operation must be executed with the system administrator authority for all cluster nodes and the cluster management server that make up the cluster system.

Table 4.3 Environment variable for the operation language of Web-Based Admin View

Attribute	Variable	Possible values	Meaning
sys	Lang	C, ja	Language environment in which Web-Based Admin View operates. C: Operates in English. ja: Operates in Japanese. If this variable is not set, Web-Based Admin View operates in the English environment.

Operation Procedure:

1. Stop Web-Based Admin View on all the management servers and nodes.

```
# /etc/init.d/fjsvwvcnf stop
# /etc/init.d/fjsvwvbs stop
```

2. Add the environment variable to the definition file (/etc/opt/FJSVwvbs/etc/webview.cnf) of Web-Based Admin View, and set the language.
   Execute the following command on all the management servers and nodes, referring to the example.

```
# /etc/opt/FJSVwvbs/etc/bin/wvSetparam -add <attribute> <environment-variable> <set-value>
```

Example: Add the environment variable and set the operation language to Japanese.

```
# /etc/opt/FJSVwvbs/etc/bin/wvSetparam -add sys lang ja
```

3. Restart Web-Based Admin View on all the management servers and nodes.

```
# /etc/opt/FJSVwvbs/etc/bin/wvCntl restart
# /etc/init.d/fjsvwvcnf start
```

Note

- For Web-Based Admin View to display messages in Japanese, the language environment of the personal computers and Solaris workstations that are being used as clients must be set to Japanese. If a client has an English environment, the message contents turn into garbled characters by the above setting change.

- To change the environment variable again after it is added by the above procedure, execute the following command:

```
# /etc/opt/FJSVwvbs/etc/bin/wvSetparam lang <setting_value>
```

For details on the command, see "4.5.3 Environment variable modification" in the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

4.3.4 Setting Up the Browser

Set up a Web browser on the clients.

See

See "3.1.3.1 Preparing the Web browser" in the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

4.3.5 Setting Up the Java Plug-in

Install the Java Plug-in on the clients.

See

For details on the supported Java Plug-in versions, see "4.3.2 Preparing the Client Environment." For instructions on setting up the Java Plug-in, see "3.1.3.2 Required for the Web Browser Environment" in the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

4.4 Starting the Web-Based Admin View Screen

After completing all the preparations, start the Web-Based Admin View GUI screen.

Operation Procedure:

1. Start the Web browser in the client.

2. Specify the URL in the following format, and access the cluster management server:

   http://<*host-name*>:<*port-number*>/Plugin.cgi

   <host-name>

   IP address or host name (**httpip**) that clients use to access the primary or secondary management server.

The default value of **httpip** is the IP address that is assigned to the node name that is output when "uname -n" is executed.

<port-number>

Specify "8081."

If the port number has been changed, specify the up-to-date number.
For instructions on changing the http port number, see "7.2.1 http port number" in the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

🔷 Note
. . . .

- If the Web-Based Admin View screen cannot be started when the host name is specified in *<host-name>*, specify the IP address directly that corresponds to the host name.

- Note that the access method may be different depending on the operation management product. To use operation management products that have different access methods at the same time, see "3.3.4 Concurrent use of operation management products with different access methods" in the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

- For information on the IP address or host name (httpip) used by clients, see the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

. . . .

3. When the Web-Based Admin View is started, the following window appears.

   Enter a user name and password that have been registered to the management server, and click *OK*.
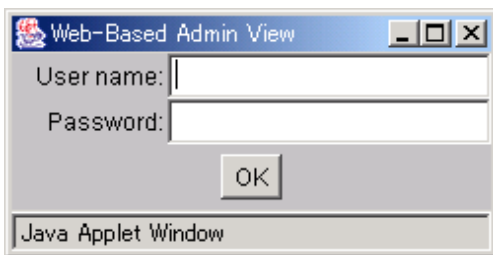
🔷 Note
. . . .

You cannot close this window by clicking "x" at the top right corner.

. . . .

4. When authentication ends, you can use Web-Based Admin View.

🔷 Note
. . . .

- After starting the Web-Based Admin View screen, do not change the page by pressing the *Forward/Next*, *Back*, or *Reread/Refresh* buttons.

- The screen of some browsers may hang.

   - If the browser is terminated, restart the browser and display Web-Based Admin View.

   - Reread the URL or restart the browser if the browser does not work (no response).

   - The operation can be continued after the Web-Based Admin View is started by moving the mouse on the browser if the page switches to a black screen.

- If the URL of the Java Plug-in is read with Netscape Communicator or Netscape Navigator, a message stating "Click here to get the plugin" may appear. In this case, the Java Plug-in may not be installed or an older Java Plug-in version that is not supported might have been installed. See the *"PRIMECLUSTER Web-Based Admin View Operation Guide"* and install the Java Plug-in. Note that if you click the "Click here to get the plugin" icon, a dialog box indicating "Plug-in is not read" appears. Select *Cancel*.

- If "Welcome to Web-Based Admin View" does not appear after you read the URL of the Java Plug-in with Internet Explorer, an appropriate Java Plug-in may not be installed. Confirm that an appropriate Java Plug-in is installed by using "Add/Remove Programs" in the control panel. If the Java Plug-in is not installed or if an older Java Plug-in version that is not supported is installed, see the *"PRIMECLUSTER Web-Based Admin View Operation Guide"* and install the Java Plug-in. Also, if the "security warning" dialog box appears, and prompts you to specify whether the "Java Plug-in" is to be installed and executed, select *No*.

- If the secondary cluster management server is set to operate dynamically, there is a function that connects automatically to the primary or secondary management server that is operating at that time even if the URL of a specific monitoring node is specified. For details, see "7.4 Secondary management server automatic migration" in the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

- If repeated errors occur during the authentication of Step 3, the message 0016 may be displayed and you may not be able to log in. For the action to take if this happens, see "Symptom 17" in "Appendix B Troubleshooting" of the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

- If Netscape 6.X is used, URL input may become disabled. This condition is a Netscape problem and Netscape is currently working to correct the problem.
  If this condition occurs, minimize the Netscape window and then return it to its original size, or restart the Netscape Web browser.

- If some problems occur while you are using Web-Based Admin View, see "Appendix A Message" and "Appendix B Troubleshooting" of the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

4.5 Web-Based Admin View Screen

When you start Web-Based Admin View, the Web-Based Admin View screen is displayed.

The left area of the Web-Based Admin View screen displays the currently supported functions as **operation menus**.

4.5.1 Operation Menu Functions

Web-Based Admin View screen supports the functions shown below.

See "Menu Outline."

Figure 4.1 Web-Based Admin View screen
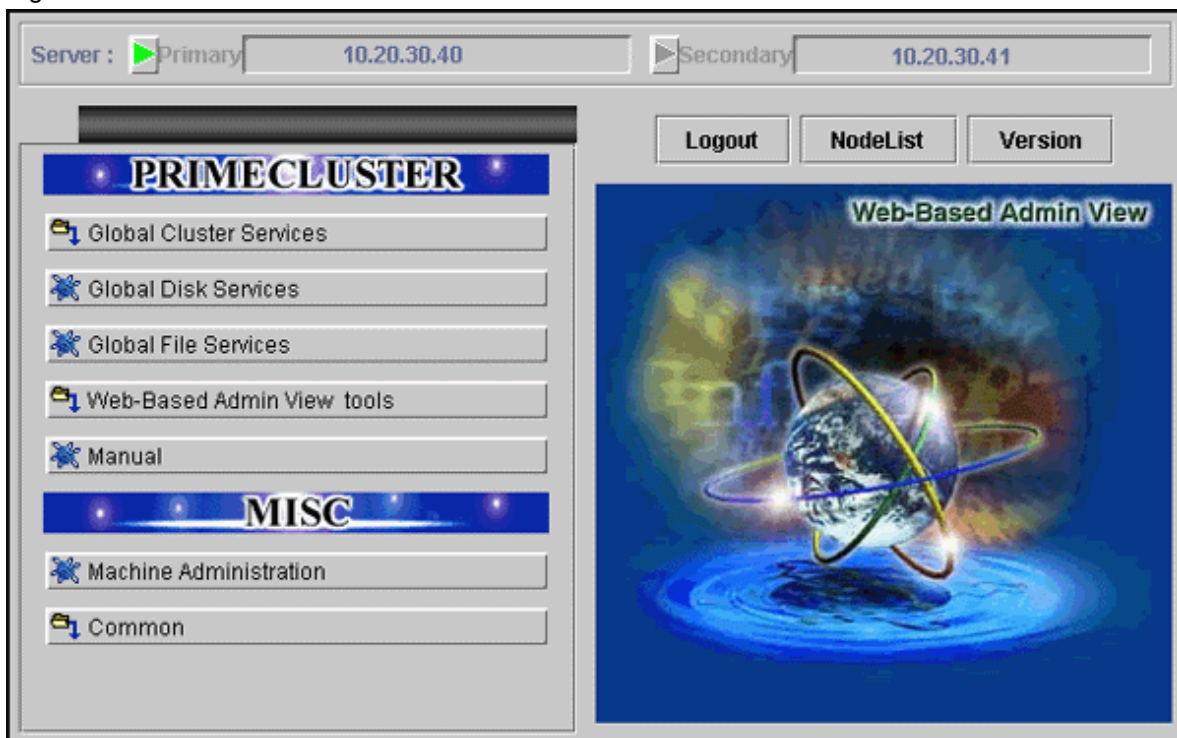
Menu Outline

The operation menus are divided into the following two categories:

1. **PRIMECLUSTER category**

   Management screens and manuals of operation management products that are provided by PRIMECLUSTER

2. **MISC category**

   Management screens and manuals of operation management products that are presented by a software product other than PRIMECLUSTER, such as the Enhanced Support Facility (ESF)

The following operations are enabled for each category:

- **PRIMECLUSTER**

  - **Operation management product name**

    You can operate the screen of the operation management product.

    For details, see the manual provided with each operation management product.

    - Global Cluster Services (CF, CRM, RMS, SIS)

    - Global Disk Services (PRIMECLUSTER GDS)

    - Global File Services (PRIMECLUSTER GFS)

    If the 3-tier model is used and an RC2000 is set, the following menu is displayed in addition to those listed above:

    - RC2000

  - **Web-Based Admin View tools**

    These tools display the Web-Based Admin View log and allow you to set the operation environment. For details, see "Part 3 Web-Based Admin View tools menu" in the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

  - **Manual**

    The PRIMECLUSTER online manual is displayed.

- **MISC**

  Buttons for starting the management screens of installed operation management products other than the PRIMECLUSTER products are displayed. In this system, this menu category is displayed because Enhanced Support Facility (ESF) is installed.

  For details, see the manual provided with each operation management product.

  - **Operation management product name**

    You can operate the management screens of installed operation management products other than the PRIMECLUSTER products.

    - Machine Administration

    - MultiPathDisk view

  - **Common**

    You can refer to manuals that are available as online manuals. For details, see the *"PRIMECLUSTER Web-Based Admin View Operation Guide."*

Operation menu transition diagram

Shown below are the transitions from the top screen of Web-Based Admin View to the other screens.
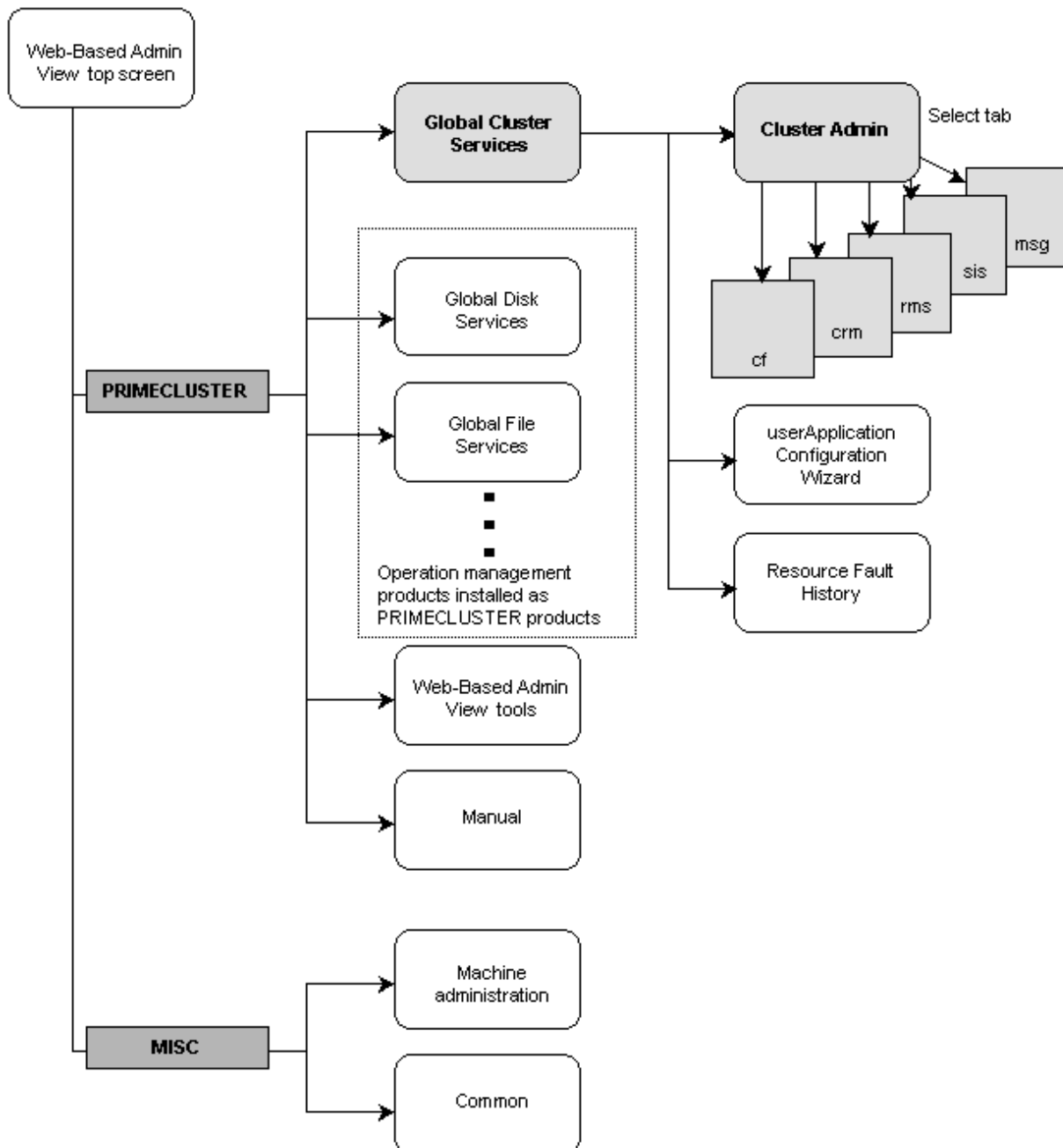
On the Cluster Admin screen, you can switch the window by clicking the following tab:

- cf: Cluster Foundation

- crm: Cluster Resource Management

- rms: Reliant Monitor Services

- sis: Scalable Internet Services

- msg: Message

The following sections describe the screens found after the Global Cluster Services menu.

4.5.2 Global Cluster Services Menu Functions

Display procedure

Web-Based Admin View screen -> *Global Cluster Services*

Exit procedure

To return to the Web-Based Admin View screen, click the arrow next to *the Global Cluster Services*.

Figure 4.2 Web-Based Admin View screen (Global Cluster Services menu)
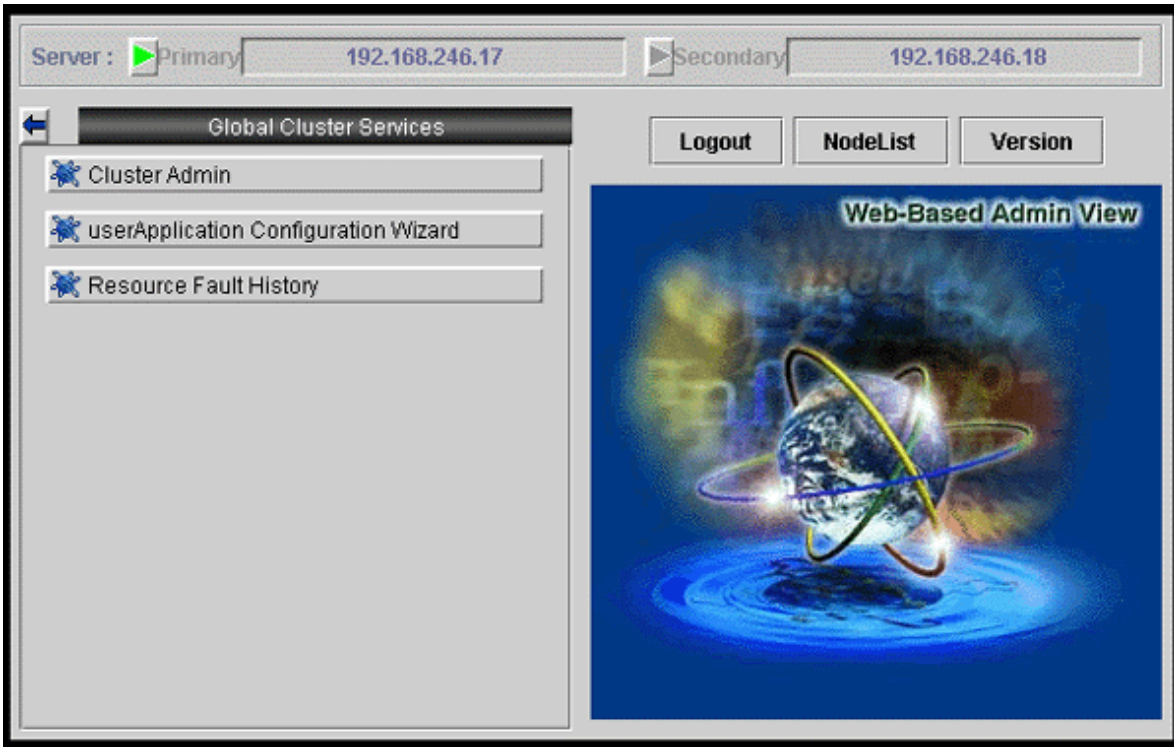
Overview of the Global Cluster Services menu

- **Cluster Admin**

  This function allows you to monitor the status of the PRIMECLUSTER system and operate the system.

- **userApplication Configuration Wizard**

  This function allows you to create cluster applications.

- **Resource Fault History**

  This function allows you to display the resource fault history. For details, see "C.3.2 Resource Fault History."

4.5.3 Cluster Admin Functions

Display procedure

Web-Based Admin View screen -> Select *Global Cluster Services*. -> Select *Cluster Admin*. -> Node selection screen -> Select the node.

Exit procedure

Select the *Exit* in the *File* menu. -> Confirmation screen -> Select the *Yes*. -> *Global Cluster Services* menu

Cluster Admin supports the functions described below.

The manual reference locations are indicated in "Overview of Cluster Admin."

Figure 4.3 Web-Based Admin View screen (Cluster Admin)
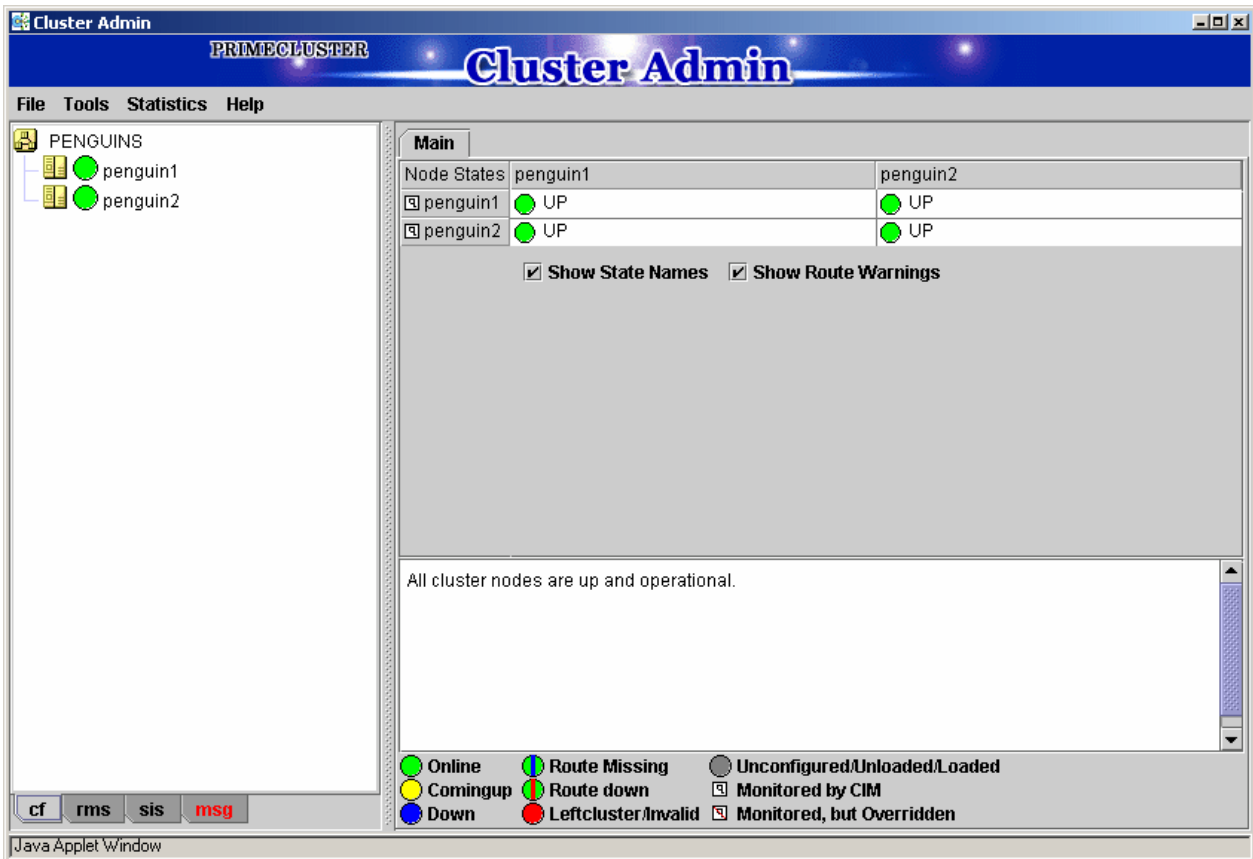
Overview of Cluster Admin

- **cf (CF: Cluster Foundation)**

  This function allows you to manage, build, monitor, and diagnose the cluster.

  Reference location: "5.1 Initial Cluster Setup", "Chapter 7 Operations"

- **crm (CRM: Cluster resource management facility)**

  This function manages the resource database, which contains information about the hardware devices (including shared disks, line switching units, and network interface cards).

  Reference location: "5.1.3 Initial Setup of the Cluster Resource Management Facility", "Chapter 7 Operations"

- **rms (RMS: Reliant Monitor Services)**

  This function monitors the status of the cluster system and manages applications and resources.

  Reference location: "Chapter 7 Operations"

- **sis (SIS: Scalable Internet Services)**

  This function enables PRIMECLUSTER to act as a scalable, reliable, and easily managed network system.

  Reference location: PRIMECLUSTER Scalable Internet Services (SIS) Configuration and Administration Guide

- **msg (Message)**

  Cluster control messages are displayed.

  Reference location: "Chapter 7 Operations"

4.5.4 userApplication Configuration Wizard Functions

Display procedure

Web-Based Admin View screen -> *Global Cluster Services* -> *userApplication Configuration Wizard*

Exit procedure

Click *the End* button. -> Global Cluster Services menu

userApplication Configuration Wizard supports the functions described below.

The manual reference locations are indicated in "Overview of userApplication Configuration Wizard."

Figure 4.4 Web-Based Admin View screen (userApplication Configuration Wizard)
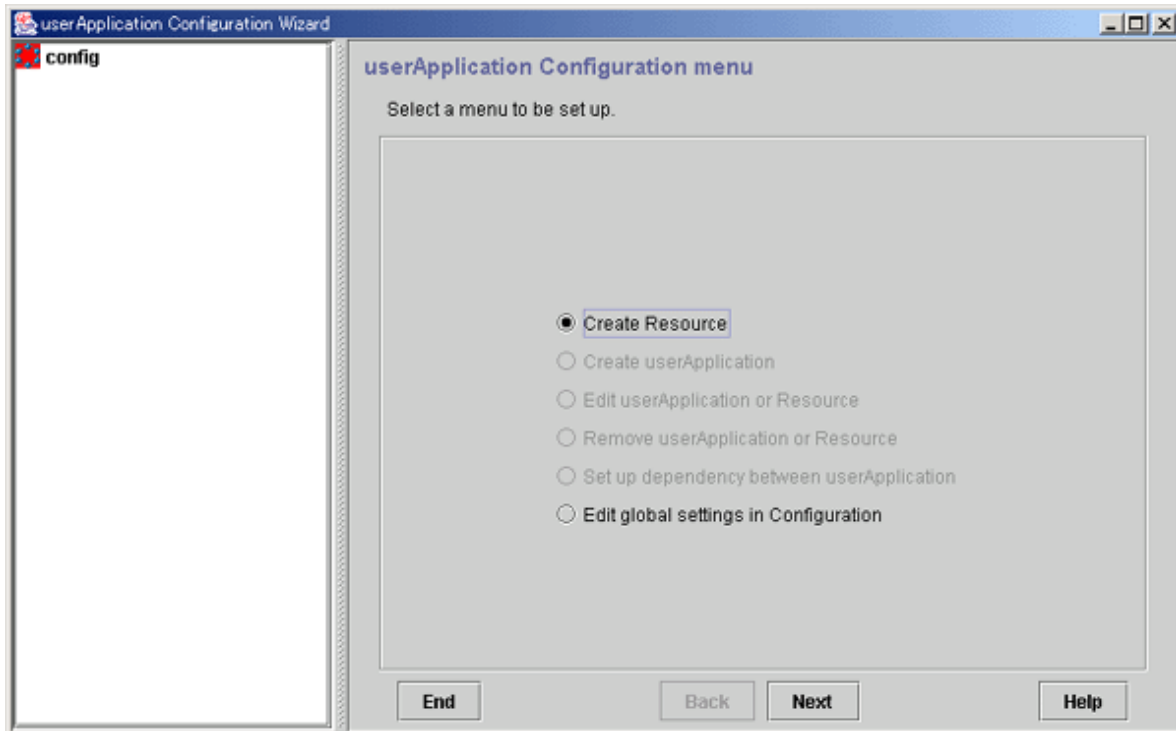
Overview of userApplication Configuration Wizard

- **Create Resource**

  This function creates a new resource.

  Reference location: "6.6.1 Setting Up Resources"

- **Create userApplication**

  This function creates cluster applications.

  Reference location: "6.6.2 Creating Cluster Applications"

- **Edit userApplication or Resource**

  This function edits the attributes of cluster applications or resources.

  Reference locations: "8.1.2 Changing the Operation Attributes of a Cluster Application" and "8.1.3.2 Changing the Attributes Used by a Resource or a Resource Interface"

- **Remove userApplication or Resource**

  This function removes cluster applications or resources.

  Reference location: "8.8 Deleting a Cluster Application"

- **Set up dependency between userApplication**

  This function sets up dependency between cluster applications.

  Reference location: "6.6.3 Setting Up Dependency Relationships Between Cluster Applications"

- **Edit global settings in Configuration**

This function edits global settings in configuration.

Reference location: "6.6.4 Editing global settings in Configuration"

4.6 Exiting the Web-Based Admin View Screen

To exit the Web-Based Admin View screen, follow the procedure below.

Logging out of the screen

To log out of the Web-Based Admin View screen, follow the procedure below.

1. Close all screens if the management screen of the operation management product is displayed.

2. When only the Web-Based Admin View screen is displayed, select the *Logout*.

Exiting the screen

To exit the Web-Based Admin View screen, follow the procedure below.

1. Log out from the Web-Based Admin View screen according to "Logging out of the screen" described above.

2. The login screen will be displayed. To exit the Web-Based Admin View screen, execute one of the following operations while the login screen is still displayed:

   - Terminate the Web browser.

   - Specify another URL in the Web browser to switch the screen. (Enter a new URL or specify a bookmark.)

   - Select the *Back* button of the browser.

Note

- To terminate the Web browser, select the *Close* in the *File* menu, or click the "x" at the top right corner of the screen.
  Note that if you are using Netscape, selecting the *Exit* in the *File* menu will terminate all browser screens.

- At the login screen, clicking the "x" at the top right corner of the screen will not terminate the screen.

- The login screen will remain temporarily after exiting the browser.

Chapter 5 Building a Cluster

The procedure for building a PRIMECLUSTER cluster is shown below.

(1) Initial Cluster Setup

— Setting Up CF and CIP

— Setting Up the Shutdown Facility

— Initial Setup of the Cluster Resource Management Facility

↓

(2) Setting Up Power Supply Linkage

↓

(3) Setting Up Shared Disk Connection Confirmation

↓

(4) Setting Up Fault Resource Identification and Operator Intervention Request
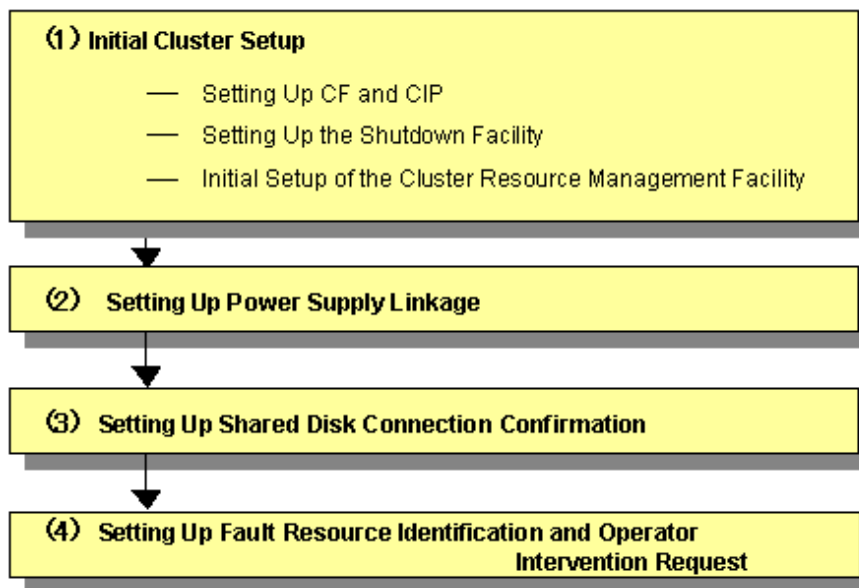
Table 5.1 Cluster building procedure and manual reference locations

	Work item	Execution Node	Required/ Optional	Manual reference location*
(1)	5.1 Initial Cluster Setup			
	5.1.1 Setting Up CF and CIP	All nodes	Required	**CF** "2.1 CF, CIP and CIM configuration"
	5.1.2 Configuring the Shutdown Facility	All nodes	Required	**CF** "8 Shutdown Facility"
	5.1.3 Initial Setup of the Cluster Resource Management Facility	All nodes	Required	**CF** "4.3 Resource database configuration"
(2)	5.2 Setting Up Power Supply Linkage	All nodes	Optional	"Machine Administration Guide"
(3)	5.3 Setting Up Shared Disk Connection Confirmation	All nodes	Required	
(4)	5.4 Setting Up Fault Resource Identification and Operator Intervention Request	1 node	Required	

* The PRIMECLUSTER manual names are abbreviated as follows:

- **CF:** *PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*

5.1 Initial Cluster Setup

This section describes the initial setup of the PRIMECLUSTER cluster.

For details on the setup methods, see the reference locations indicated in the table below.

	Contents	Manual reference location*[1]
1	Checking Console Configuration*[2]	- For a PRIMEPOWER/S-Series: 5.1.2.1.1 Checking Console Configuration

	Contents	Manual reference location*¹
		- For a SPARC Enterprise M3000, M4000, M5000, M8000, or M9000: 5.1.2.2.1 Checking Console Configuration - For a SPARC Enterprise T1000, T2000: 5.1.2.3.1 Checking Console Configuration - For a SPARC Enterprise T5120, T5220, T5140, T5240, T5440, or SPARC T3 series: 5.1.2.4.1 Checking Console Configuration
2	5.1.1 Setting Up CF and CIP (setting up cluster configuration information and IP addresses)	**CF** "2.1 CF, CIP and CIM configuration"
3	5.1.2 Configuring the Shutdown Facility	**CF** "8 Shutdown Facility"
4	5.1.3 Initial Setup of the Cluster Resource Management Facility	**CF** "4.3 Resource database configuration"

*¹ The PRIMECLUSTER manual names are abbreviated as follows:

- **CF:** *PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*

*² "1. Check console information" is the prerequisite of "3. Configure Shutdown Facility". Be aware that you need to check the console information before "2. Configure CF and CIP".

5.1.1 Setting Up CF and CIP

Set up Cluster Foundation (CF) and CIP using the CF Wizard of Cluster Admin. Enter the information described in the table below. The worksheet for the CF and CIP setup is the "A.6 Cluster Configuration Worksheet". For details on the setting, see "2.1.4 Example of creating a cluster" in the *"PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide"*.

Setup item	Description	Described in worksheet
Cluster name	Define the name of the cluster systems. Use up to 31 printable ASCII characters (except space, line feed, and tab characters) for each name. Cluster names are always processed as uppercase characters.	"Cluster name"
Cluster nodes	Select the nodes that will construct a cluster system.	"Node name (uname-n)" for "Node 1" and "Node 2"
CF node names	Set the names of the nodes that construct the cluster. Use up to 11 printable ASCII characters (except space, line feed, and tab characters) for each node name. CF node names are always processed as lowercase characters. In the setup screen, the same names as the "Cluster nodes" above are displayed by default.	"CF node name" for "Node 1" and "Node 2"
Cluster interconnects	In each node of the cluster, determine the network interface to be used in CF inter-node communication. A representative network interface is the Ethernet device. Set the network interfaces to be used for CF inter-node communication so that they are activated when the system is started.	"Path 0 NIC name" and "Path 1 NIC name" for "Node 1" and "Node 2"

Setup item	Description	Described in worksheet
	However, IP addresses must not be assigned to the interfaces.	
IP interconnects	Optional. This setup allows you to operate CF over IP.	"IP interconnect setup"
CIP subnets	Set the following items, and set the IP address used by CF: CIP subnet count Host suffix Subnet number Subnet mask	"Number of subnets," "Subnet IP," and "Net mask" for "CIP"
Usage confirmation of CF remote services	Check whether the following functions are to be enabled: Remote file copy (cfcp) Remote command execution (cfsh) With the default settings, these services are "Not selected". To use RMS, you need to select at least one function.	"Usage status of CF remote services"
Cluster Integrity Monitor (CIM) configuration	Set the nodes to be monitored by CIM.	"Node in the CF quorum set" for "Node 1" and "Node 2"

Note

- If you enable any one of the CF remote services, do not connect the following systems in the same cluster interconnect:

  - Systems that have a security problem

  - Systems in which cluster interconnects are not secured

- If you used the userApplication Configuration Wizard (GUI), the two remote services, "remote file copy" and "remote command execution," will be enabled automatically. If the cluster interconnects are not secured, comment out the following two lines in the "/etc/default/cluster.config" file for all cluster nodes after you build userApplication:

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

The results are as follows:

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

- A CF or CIP setup error can be due to one of the following causes:

  - A cluster interconnect is disconnected or incorrectly connected.

  - A network interface for a cluster interconnect is not activated.

  - An IP address is assigned to a network interface for a cluster interconnect.

- The cluster node name will be automatically used for the CF node name. The name string must consist of eleven characters or less. Change it if necessary.

- Do not add CIP nodenames manually to the /etc/hosts or /etc/inet/ipnodes file because the CF Wizard automatically updates the /etc/hosts and /etc/inet/ipnodes files on each node in the cluster.

See

....

For information on the corrective action to be applied when the setting of the cluster interconnect fails, see "11 Diagnostics and Troubleshooting" in the *"PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."*

....

5.1.2 Configuring the Shutdown Facility

This section explains the procedure for configuring the shutdown facility with the shutdown configuration wizard or CLI.

The configuration procedure for the shutdown facility varies depending on the machine type. Check the machine type of hardware and set an appropriate shutdown agent.

The following table shows the shutdown agent necessary by machine type.

Server machine type name			RCI		XSCF			RCCU	ALOM	ILOM	
			Panic	Reset	Panic	Reset	Console Break	Console Break	Console Break	Panic	Reset
PRIME POWER	200 400 600		Y	Y	-	-	-	Y	-	-	-
	250 450	XSCF (*1)	Y	Y	Y	Y	Y	-	-	-	-
		RCCU (*2)	Y	Y	-	-	-	Y	-	-	-
	650 850		Y	Y	-	-	-	Y	-	-	-
	800		Y	Y	-	-	-	-	-	-	-
	900		Y	Y	-	-	-	-	-	-	-
	1000 2000		Y	Y	-	-	-	-	-	-	-
	1500 2500		Y	Y	-	-	-	-	-	-	-
SPARC Enter prise	M3000 M4000 M5000 M8000 M9000	Japan	Fujitsu	Y	Y	Y	Y	Y	-	-	-	-
			Other than Fujitsu	-	-	Y(*6)	Y(*6)	Y(*6)	-	-	-	-
		Other than Japan		-	-	Y(*7)	Y(*7)	Y(*7)	-	-	-	-
	T1000 T2000		-	-	-	-	-	-	Y(*3)	-	-
	T5120 T5220 T5140 T5240 T5440		-	-	-	-	-	-	-	Y(*3)	Y (*3) (*4) (*5)
SPARC	T3 series		-	-	-	-	-	-	-	Y(*3)	Y (*3) (*4)
S-Series			-	-	-	-	-	Y	-	-	-

(*1) XSCF is used for the console.

- 50 -

(*2) RCCU is used for the console.

(*3) You cannot configure the shutdown configuration wizard. You configure with CLI.

(*4) When using ILOM Reset, you need to apply a patch for PRIMECLUSTER (914468-07 or later).

(*5) When using ILOM Reset, you need firmware for SPARC Enterprise server (System Firmware 7.1.6.d or later).

(*6) When using SPARC Enterprise M3000, M4000, M5000, M8000, or M9000 provided by companies other than Fujitsu in Japan, you need to apply a patch for PRIMECLUSTER (914468-05 or later).

(*7) When using SPARC Enterprise M3000, M4000, M5000, M8000, or M9000 with logos of both Fujitsu and Oracle provided in other than Japan, you need to apply a patch for PRIMECLUSTER (914468-07 or later).

📝 Note
. . . .

- When you are operating the shutdown facility by using one of the following shutdown agents, do not use the console.

    - XSCF Panic

    - XSCF Reset

    - XSCF Console Break

    - RCCU Console Break

    - ILOM Panic

    - ILOM Reset

  If you cannot avoid using the console, stop the shutdown facility of all nodes beforehand. After using the console, cut the connection with the console, start the shutdown facility of all nodes, and then check that the status is normal. For details on stop, start, and the state confirmation of the shutdown facility, see the manual page describing sdtool(1M).

- In the /etc/inet/hosts file, you must describe IP addresses and host names of the administrative LAN used by the shutdown facility for all nodes. Check that IP addresses and host names of all nodes are described.

- When you set up asynchronous RCI monitoring, you must specify the timeout interval (kernel parameter) in /etc/system for monitoring via SCF/RCI. Kernel parameters vary depending on the server type. Then check your server type so you can set the appropriate timeout interval.

- To make the administrative LAN, used in the shutdown facility, redundant by GLS, use the physical IP address takeover function of NIC switching mode.

. . . .

📑 See
. . . .

For details on the shutdown facility and the asynchronous monitoring function, refer to the following manuals:

- "3.3.1.8 PRIMECLUSTER SF" in the "*PRIMECLUSTER Concept Guide*".

- "8. Shutdown Facility" in "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*".

. . . .

5.1.2.1 For a PRIMEPOWER/S-Series

5.1.2.1.1 Checking Console Configuration

The console used is different in dependence on the machine type. Check the hardware machine type and configure information of console.

- For a PRIMEPOWER 250 or 450:

    - RCCU or XSCF

- For a PRIMEPOWER 200, 400, 600, 650 or 850:

    - RCCU

- For a S-Series:

    - RCCU

📒 Note
....

- Check the console information before initially configuring the cluster system.

- Set the IP address of RCCU or XSCF to the same segment as the administrative LAN.
....

Checking RCCU

If RCCU is used for the console, record the following information on RCCU. Note, however, that you do not need to record if you use the console with the factory settings.

- RCCU IP address or a RCCU host name registered in the "/etc/inet/hosts" file

- User name to log in the RCCU control port

- Password to log in the RCCU control port

- Password to log in the RCCU control port using superuser access privileges.

📘 See
....

For details on how to configure and check RCCU, see the instruction manual provided with RCCU.
....

Checking XSCF

Check the following settings concerning XSCF before setting the shutdown facility.

- In XSCF, the telnet must be effective in connected permission protocol type from the outside.

- The log in user account must be made for the shutdown facility, and group ID must be root.

📒 Note
....

When the connection to XSCF is a serial port connection alone, it is not supported in the shutdown facility. Please use XSCF-LAN.
....

Moreover, record the following information on XSCF.

- XSCF IP address or an XSCF host name registered in the "/etc/inet/hosts" file of the node

- Log in user account and password for shutdown facility in XSCF

    *1) When the network routing is set, IP address of XSCF need not be the same to management LAN segment of the cluster node.

📘 See
....

For information on how to configure and confirm XSCF, see the "*XSCF User's Guide*".
....

5.1.2.1.2 Using the Shutdown Configuration Wizard

The configuration procedure for the shutdown facility varies depending on the machine type. Check the hardware machine type and configure an appropriate shutdown agent.

- For a PRIMEPOWER 250, or 450:

    - RCI Panic

    - XSCF Panic

- Console Break

- RCI Reset

- XSCF Reset

- For a PRIMEPOWER 200, 400, 600, 650, or 850:

- RCI Panic

- Console Break

- RCI Reset

- For a PRIMEPOWER 800, 900, 1000, 1500, 2000, or 2500:

- RCI Panic

- RCI Reset

- For a S-Series:

- Console Break

Starting up the shutdown configuration wizard

From the CF main window of the Cluster Admin screen, select the *Tool* menu and then *Shutdown Facility -> Configuration Wizard*. The shutdown configuration wizard will start.

Information
. . . .

You can also configure the shutdown facility immediately after you complete the CF configuration with the CF wizard.

The following confirmation popup screen will appear. Click *Yes* to start the shutdown configuration wizard.
. . . .

Selecting a configuration mode

You can select either of the following two modes to configure the shutdown facility:

- *Easy configuration (recommended)*

- *Detailed configuration*

This section explains how to configure the shutdown facility using *Easy configuration (recommended).* With this mode, you can configure the PRIMECLUSTER shutdown facility according to the procedure.

Figure 5.1 Selecting the SF configuration mode
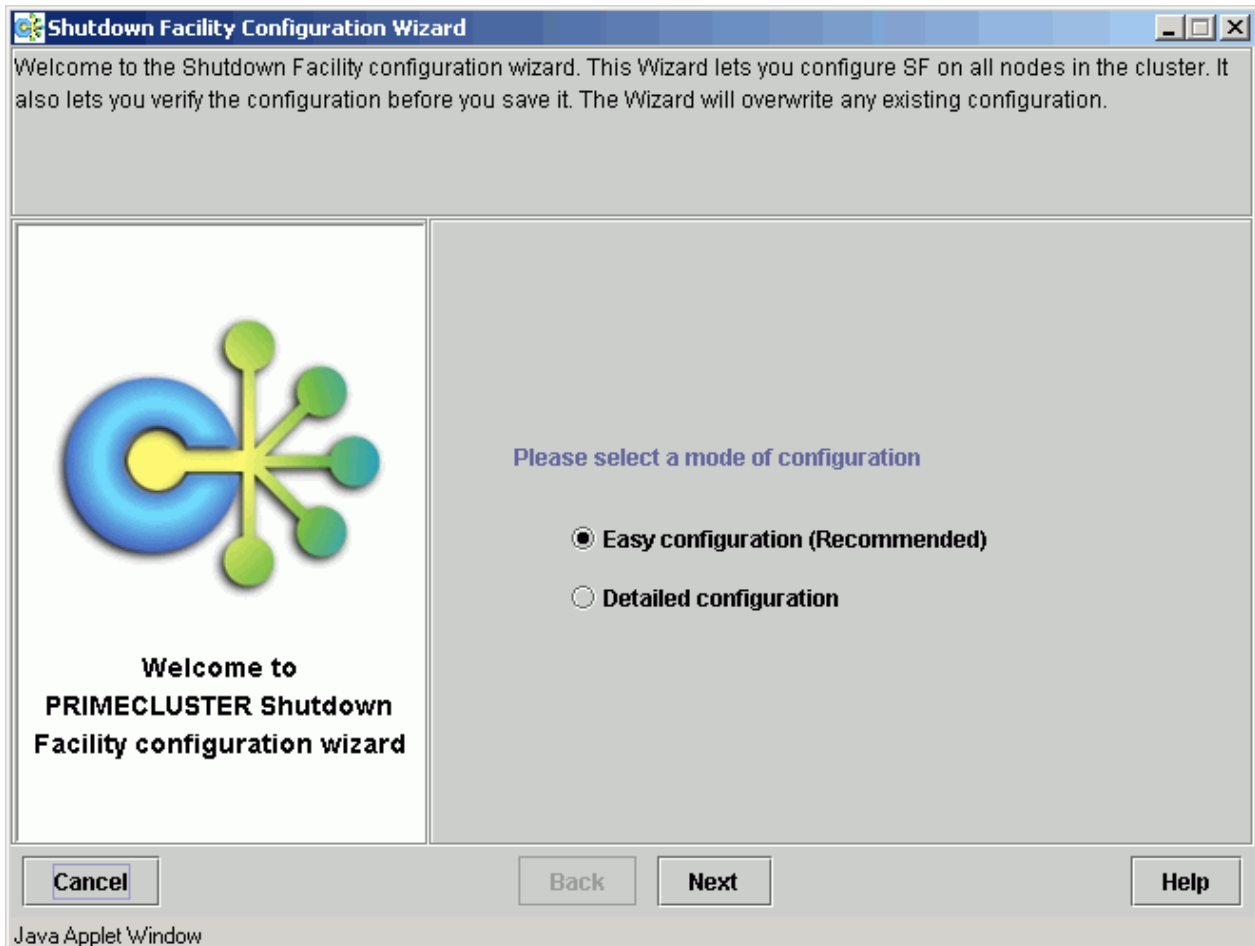
Select *Easy configuration (recommended)* and then click *Next*.

Selecting a shutdown agent

After you confirm the hardware machine type, select an appropriate shutdown agent.

Figure 5.2 Selecting a shutdown agent
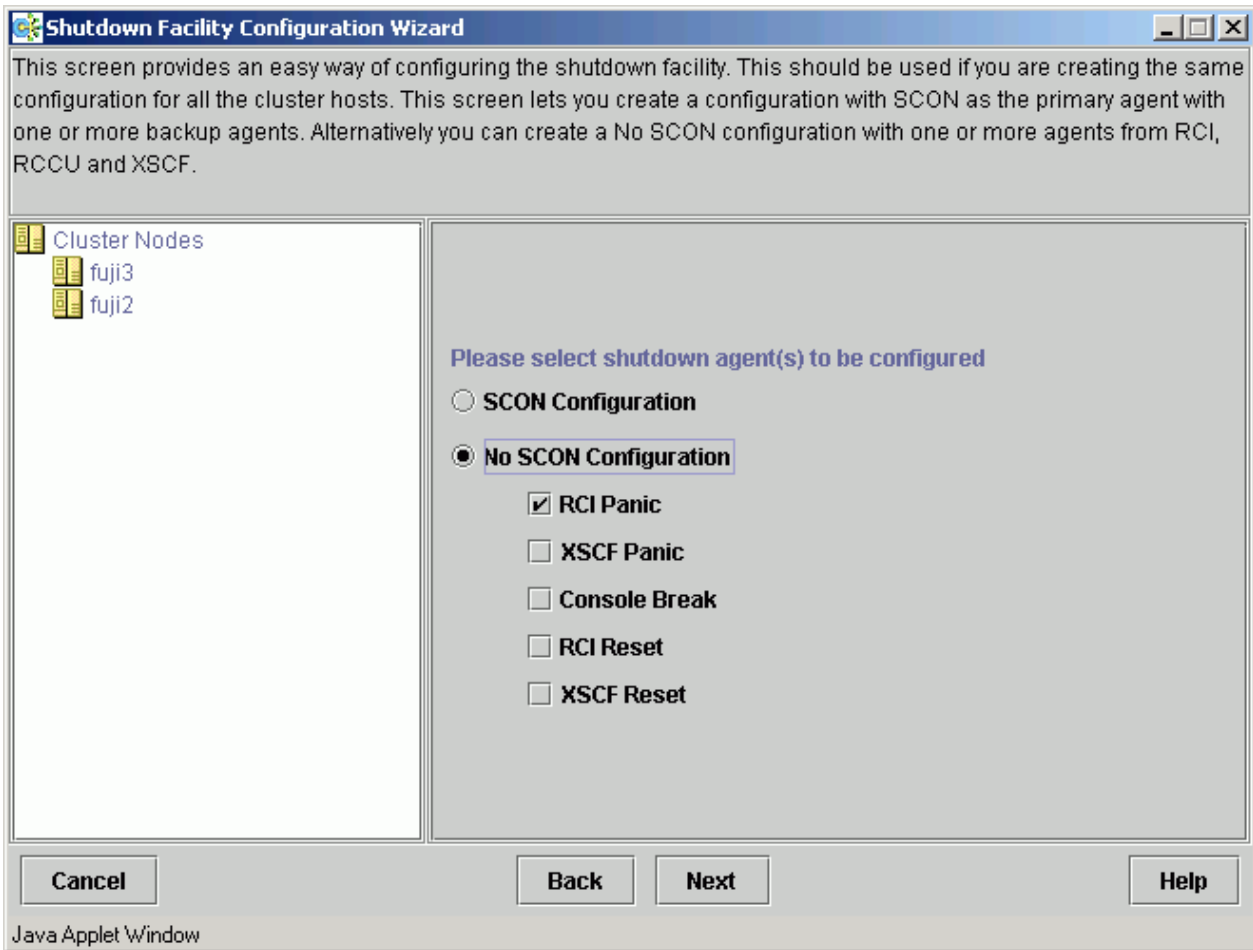
Select *No SCON Configuration*.

Then, select all the following shutdown agents depending on a hardware machine type.

- For a PRIMEPOWER 250 or 450:

  - *RCI Panic*

  - *XSCF Panic*

  - *Console Break*

  - *RCI Reset*

  - *XSCF Reset*

  * If you select *Console Break*, you can select *RCCU* or *XSCF* as the **Console Break agent**. To choose *RCCU*, you have to configure the RCCU unit in advance.

- For a PRIMEPOWER 200, 400, 600, 650, or 850:

  - *RCI Panic*

  - *Console Break*

  - *RCI Reset*

- For a PRIMEPOWER 800, 900, 1000, 1500, 2000, or 2500:

  - *RCI Panic*

  - *RCI Reset*

- For a S-Series:

    - *Console Break*

    * Select *RCCU* as *Console Break agent* after selecting *Console Break*.

Upon the completion of configuration, click *Next*.

Configuring XSCF

If you select *XSCF Panic* or *XSCF Reset* for the shutdown agent, the screen for configuring the XSCF will appear.

Enter the settings for XSCF that you recorded in "5.1.2.1.1 Checking Console Configuration".

Figure 5.3 Configuring XSCF

XSCF-name

　　Enter the IP address of XSCF or the host name of XSCF that is registered in the /etc/inet/hosts file.

User-Name

　　Enter a user name to log in to the control port.

Password

　　Enter a password to log in to the control port.

Once you have made all the necessary settings, click *Next*.

Configuring a Console Break agent

If you select *Console Break* as the shutdown agent, the screen used for selecting a Console Break agent will appear.

Figure 5.4 Selecting a Console Break agent
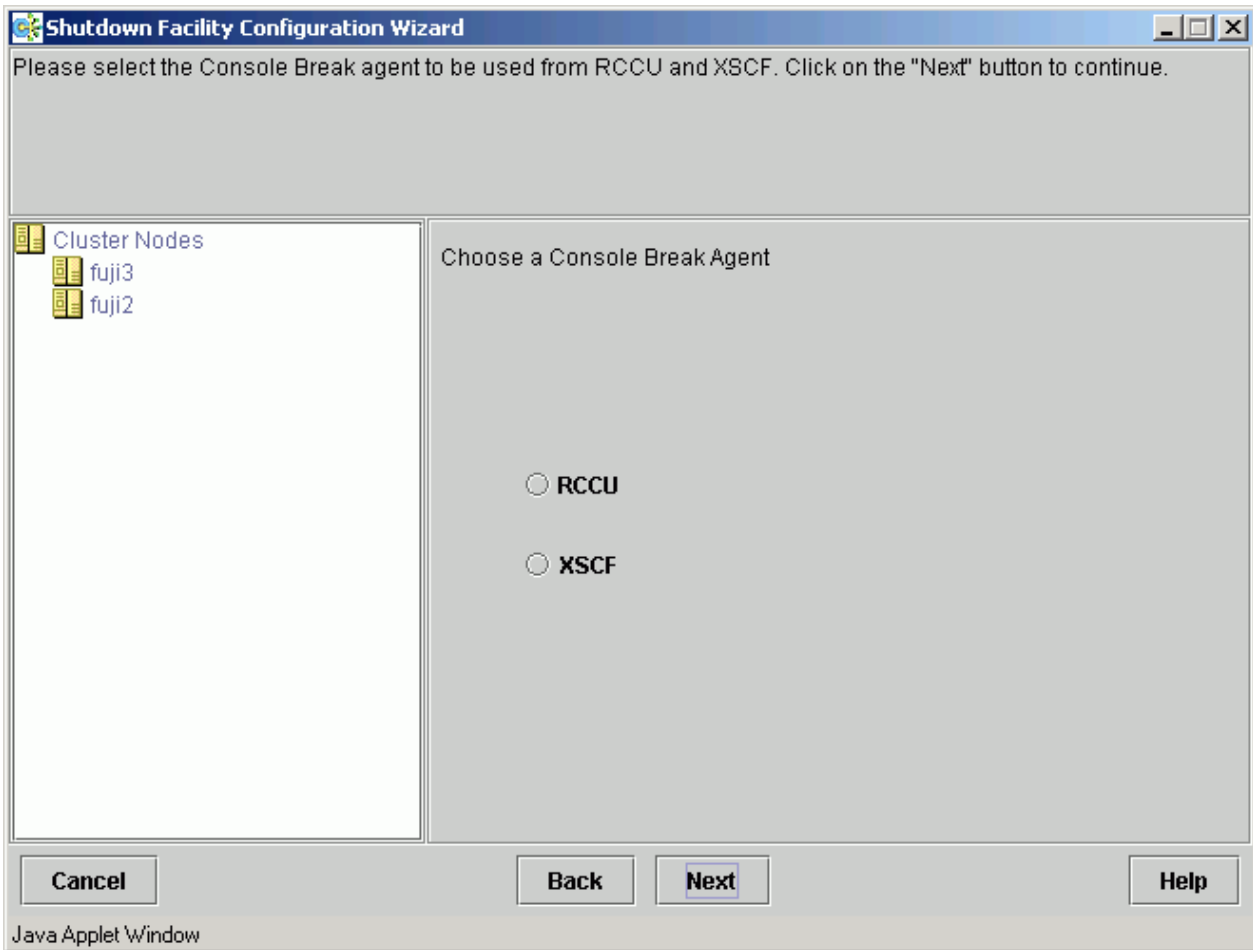
The selection to be made for the Console Break agent varies depending on the machine type to be set up. Confirm the hardware machine type and then set up an appropriate Console Break agent.

- For a PRIMEPOWER 250 or 450:

  RCCU or XSCF

- For a PRIMEPOWER 200, 400, 600, 650, or 850:

  RCCU

- For a S-Series:

  RCCU

Upon the completion of configuration, click *Next*.

Configuring RCCU

If you select RCCU as the Console Break agent, you must configure RCCU.

Enter the settings for RCCU that you recorded in "5.1.2.1.1 Checking Console Configuration".

If you wish to use RCCU with its factory settings, select *Use Defaults*.

Otherwise, uncheck *Use Defaults*, and then enter the user name, password, and superuser password.

Figure 5.5 Configuring RCCU (Use Defaults)
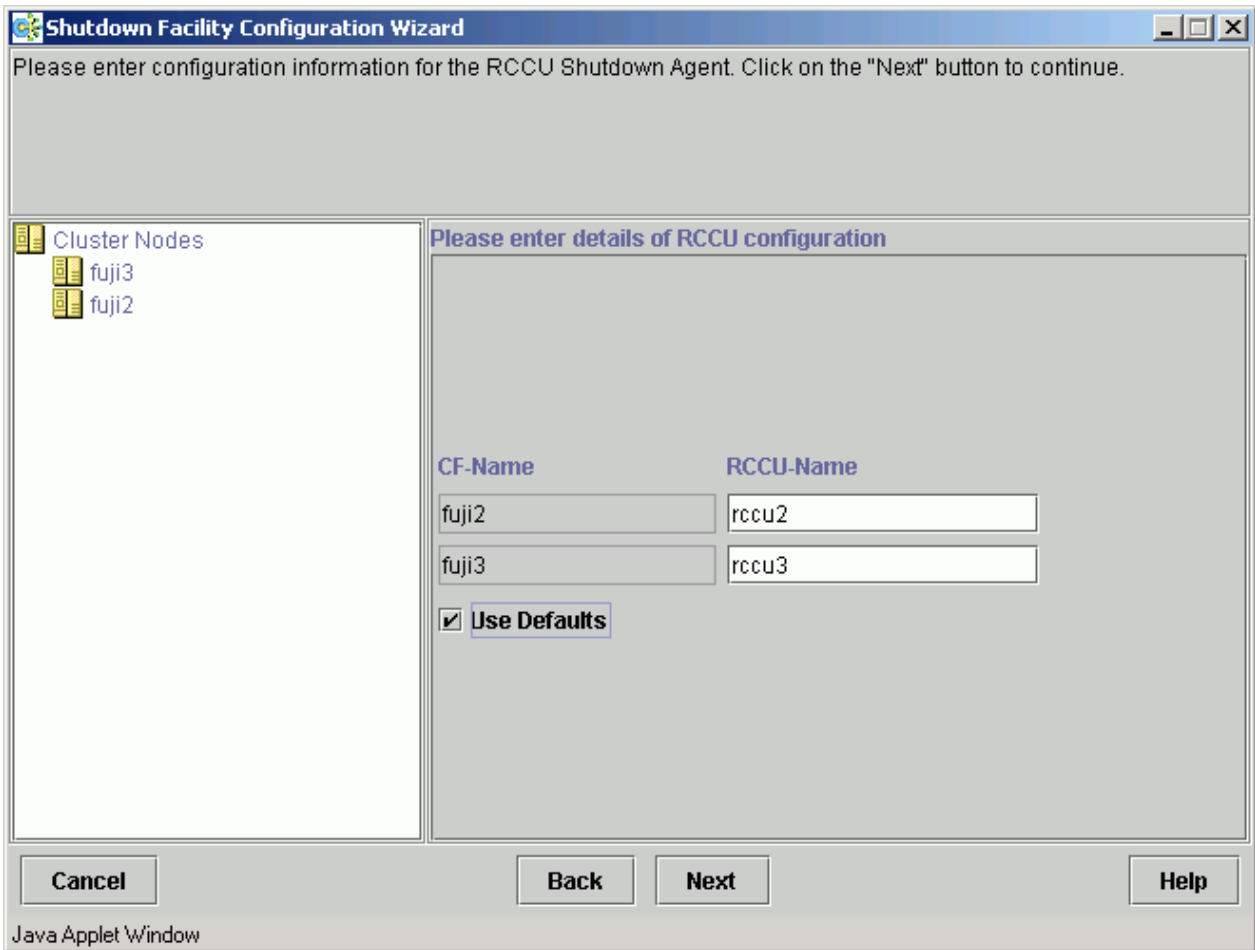
RCCU name

Enter the IP address of RCCU or the host name of RCCU that is described in the /etc/inet/hosts file.

After you have completed this configuration, click *Next*.

Figure 5.6 Configuring RCCU (Does not use default)
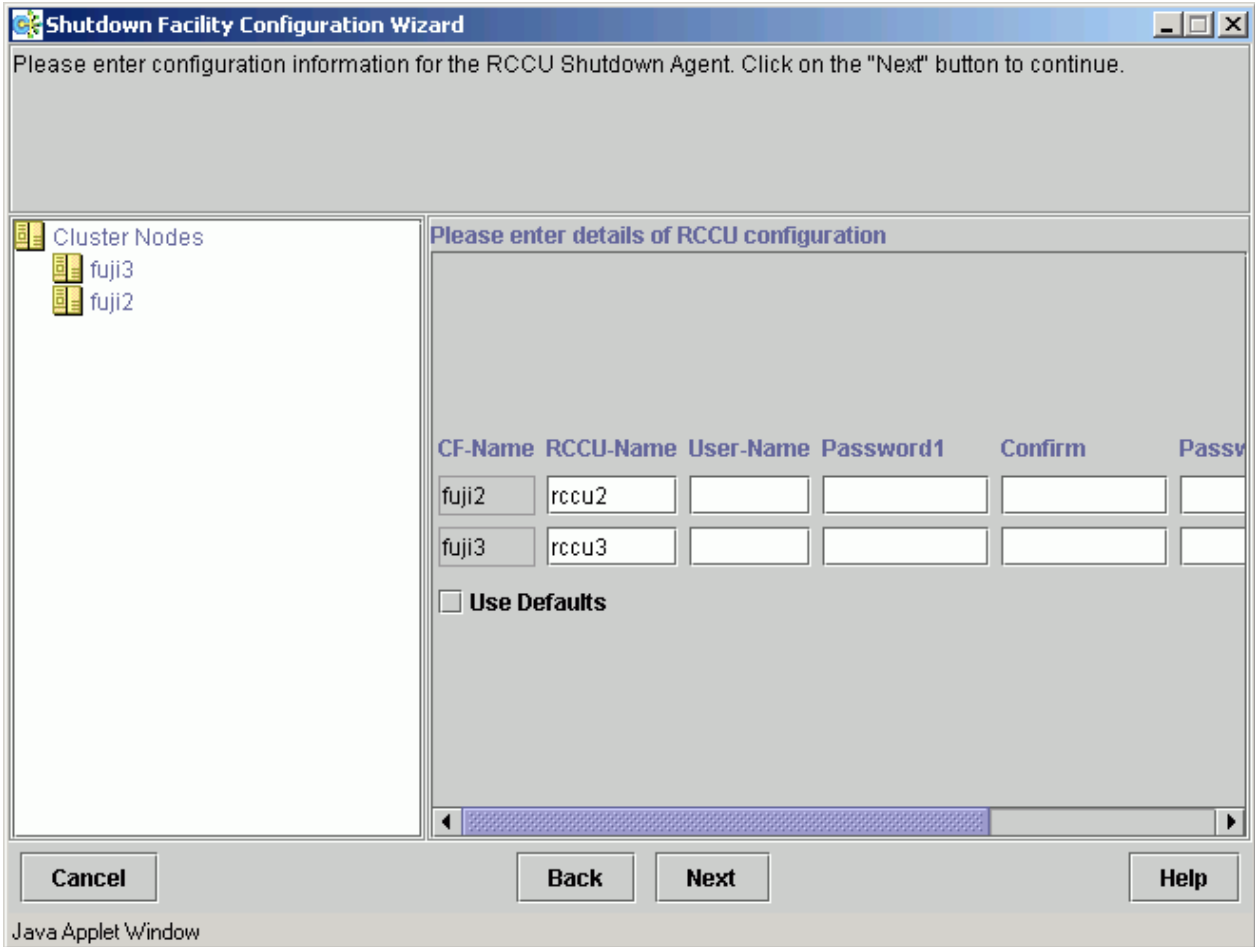
RCCU name

Enter the IP address of RCCU or the host name of RCCU that is described in the /etc/inet/hosts file.

User-Name

Enter a user name to log in to the control port of RCCU.

Password 1

Enter a password to log in to the control port of RCCU.

Confirm

Enter the password that has been set for *Password 1* for confirmation.

Password 2 (admin)

Enter a password to log in to the control port of RCCU with superuser access privileges.

Confirm

Enter the password that has been set for *Password 2 (Admin)* for confirmation.

Upon the completion of configuration, click *Next*.

Configuring Wait for PROM

🏷 Note
....
Wait for PROM is currently not supported.
Uncheck the check box if checked, and then click the *Next* button.
....

Figure 5.7 Configure Wait for PROM
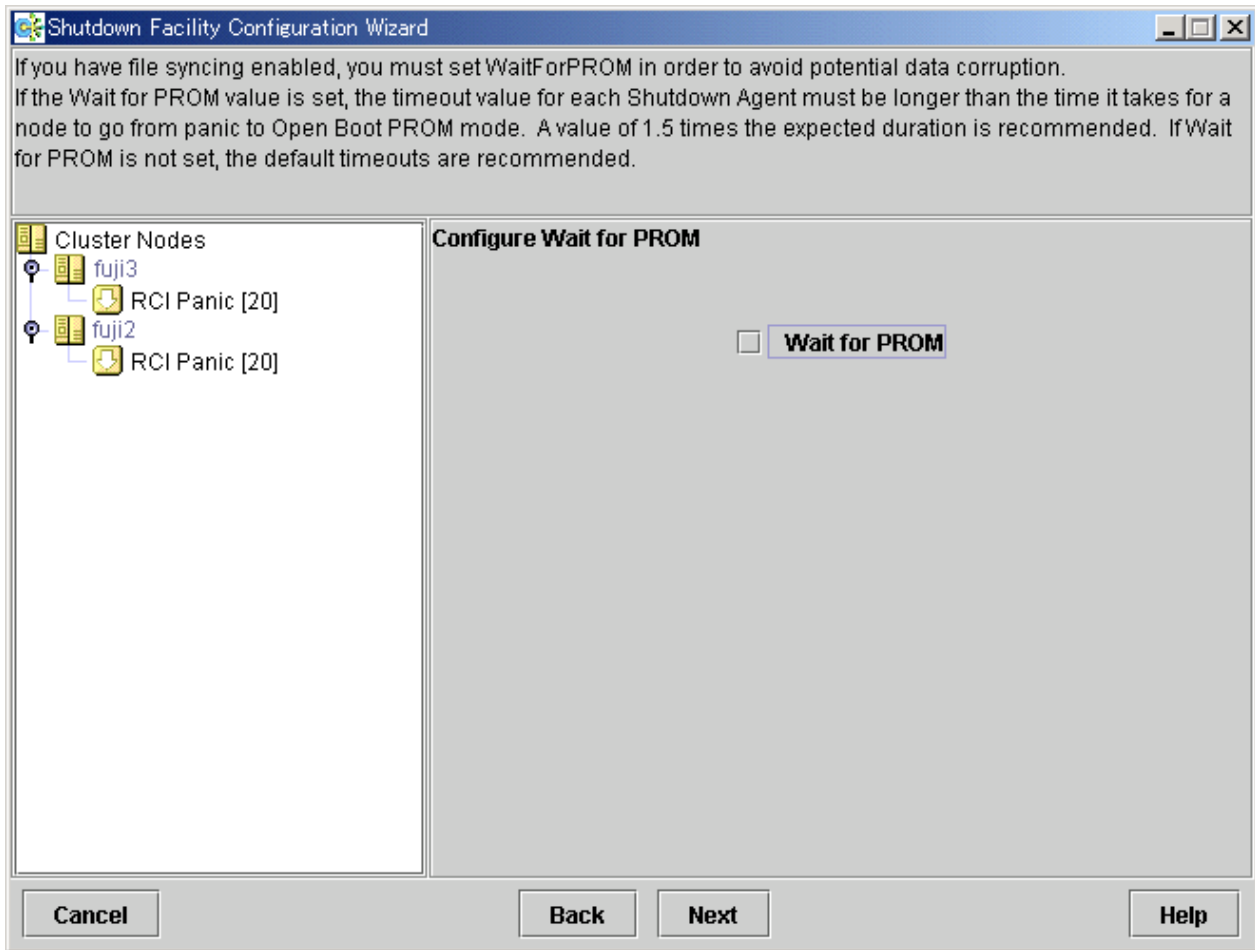
Entering node weights and administrative IP addresses

Enter the weights of the nodes and the IP addresses for the administrative LAN.

Figure 5.8 Entering node weights and administrative IP addresses
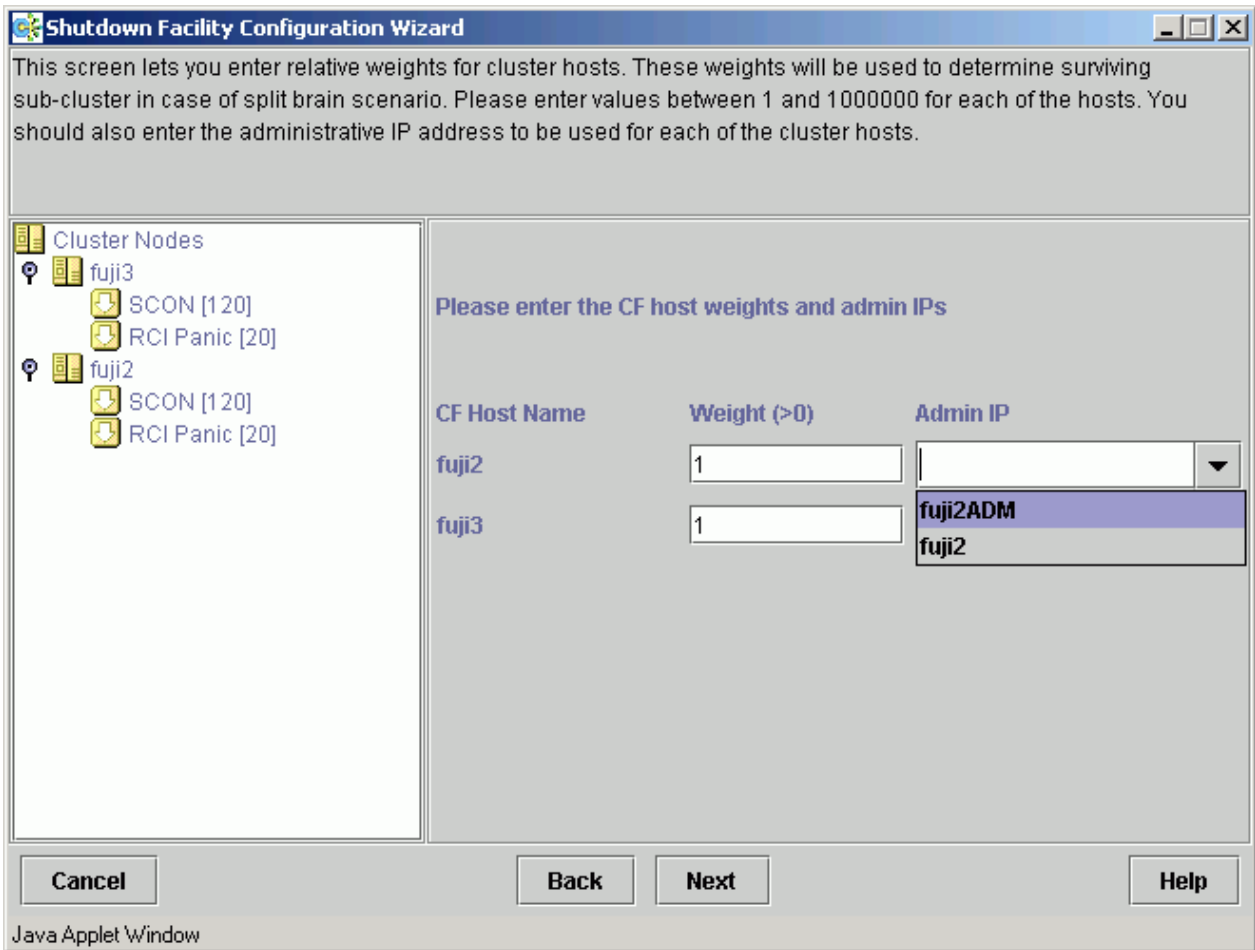
Weight

> Enter the weight of the node that constitutes the cluster. Weight is used to identify the survival priority of the node group that constitutes the cluster. Possible values for each node range from 1 to 300.
> For details on survival priority and weight, refer to the explanations below.

Admin IP

> Enter an IP address directly or click the tab to select the host name that is assigned to the administrative IP address.

Once you have completed this configuration, click *Next*.

Survival priority

> Even if a cluster partition occurs due to a failure in the cluster interconnect, all the nodes will still be able to access the user resources. For details on the cluster partition, see "2.2.2.1 Protecting data integrity" in the *"PRIMECLUSTER Concept Guide"*.
> To guarantee the consistency of the data constituting user resources, you have to determine the node groups to survive and those that are to be forcibly stopped.
> The weight assigned to each node group is referred to as a "Survival priority" under PRIMECLUSTER.
> The greater the weight of the node, the higher the survival priority. Conversely, the less the weight of the node, the lower the survival priority. If multiple node groups have the same survival priority, the node group that includes a node with the name that is first in alphabetical order will survive.

> Survival priority can be found in the following calculation:

> Survival priority = SF node weight + ShutdownPriority of userApplication

SF node weight (Weight):

> Weight of node. Default value = 1. Set this value while configuring the shutdown facility.

userApplication ShutdownPriority:

Set this attribute when userApplication is created. For details on how to change the settings, see "8.1.2 Changing the Operation Attributes of a Cluster Application".

See

. . . .

For details on the ShutdownPriority attribute of userApplication, see "6.6.5 Attributes".

. . . .

Survival scenarios

The typical scenarios that are implemented are shown below:

[Largest node group survival]

- Set the weight of all nodes to 1 (default).

- Set the attribute of ShutdownPriority of all user applications to 0 (default).

	Node group1			Node group2
	node1	node2	node3	node4
weight of node	1	1	1	1
ShutdownPriority of app1 = 0				0
ShutdownPriority of app2 = 0				0
ShutdownPriority of app3 = 0				0
Live priority	3			1

[Specific node survival]

- Set the "weight" of the node to survive to a value more than double the total weight of the other nodes.

- Set the ShutdownPriority attribute of all user applications to 0 (default).

In the following example, node1 is to survive:

	Node group1	Node group2		
	node1	node2	node3	node4
weight of node	10	1	1	1
ShutdownPriority of app1 = 0		0		
ShutdownPriority of app2 =0			0	
ShutdownPriority of app3 =0				0
Live priority	10	3		

[Specific application survival]

- Set the "weight" of all nodes to 1 (default).

- Set the ShutdownPriority attribute of the user application whose operation is to continue to a value more than double the total of the ShutdownPriority attributes of the other user applications and the weights of all nodes.

  In the following example, the node for which app1 is operating is to survive:

	Node group1	Node group2		
	node1	node2	node3	node4
weight of node	1	1	1	1
ShutdownPriority of app1 =20	20			
ShutdownPriority of app2 =1			1	
ShutdownPriority of app3 =1				1
Live priority	21	5		

Saving the configuration

Confirm and then save the configuration. In the left-hand panel of the window, those nodes that constitute the cluster are displayed, as are the shutdown agents that are configured for each node.

Figure 5.9 Saving the configuration
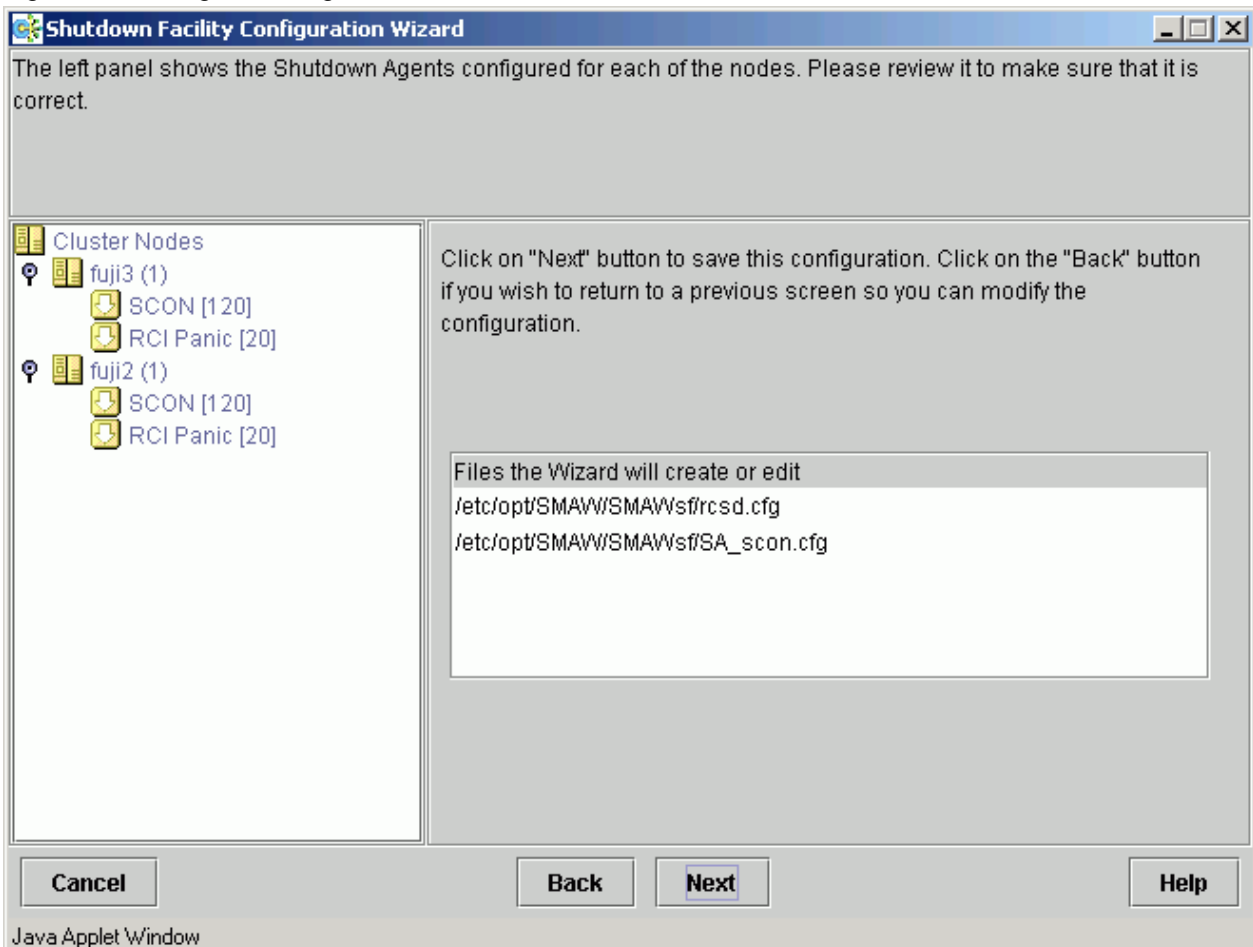
Click *Next*. A popup screen will appear for confirmation. Select *Yes* to save the setting.

Displaying the configuration of the shutdown facility

If you save the setting, a screen displaying the configuration of the shutdown facility will appear. On this screen, you can confirm the configuration of the shutdown facility on each node by selecting each node in turn.

Information

You can also view the configuration of the shutdown facility by selecting *Shutdown Facility -> Show Status* from the *Tool* menu.

Figure 5.10 Show Status

Shut State

    "Unknown" is shown during normal system operation. If an error occurs and the shutdown facility stops the relevant node successfully, "Unknown" will change to "KillWorked".

Test State

    Indicates the state in which the path to shut down the node is tested when a node error occurs. If the test of the path has not been completed, "Unknown" will be displayed. If the configured shutdown agent operates normally, "Unknown" will be changed to "TestWorked".

Init State

    Indicates the state in which the shutdown agent is initialized.

To exit the configuration wizard, click *Finish*. Click *Yes* in the confirmation popup screen that appears.

. . . .

Confirm that the shutdown facility is operating normally.

If "InitFailed" is displayed in the Initial state even when the configuration of the shutdown facility has been completed or if "Unknown" is displayed in the Test state or "TestFailed" is highlighted in red, the agent or hardware configuration may contain an error. Check the /var/adm/messages file and the console for an error message. Then, apply appropriate countermeasures as instructed the message that is output.

. . . .

. . . .

For details on how to respond to the error messages that may be output, see the following manual.

- "12.12 Monitoring Agent messages" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*".

. . . .

5.1.2.1.3 Specifying the Timeout Value

Confirm that the timeout value for each shutdown agent is as shown below. You can confirm the timeout value in the left-hand panel of the shutdown configuration wizard screen.

<How to calculate a timeout value>

- For RCCU or XSCF

    - 4 or fewer nodes

    Timeout value = 20 (seconds)

    - 5 or more nodes

    Timeout value = 6 x number of cluster nodes + 2 (seconds)

    Example) For 5 nodes: 6 x 5 + 2 = 32 (seconds)

- For RCI

    Timeout value = 20 (seconds)

If the value has not been set as explained above, set a timeout value as follows.

Setting a timeout value

In the CF main window of Cluster Admin, select *Shutdown Facility -> Configuration Wizard* from the *Tool* menu and then start the configuration wizard.

Figure 5.11 Selecting a configuration mode
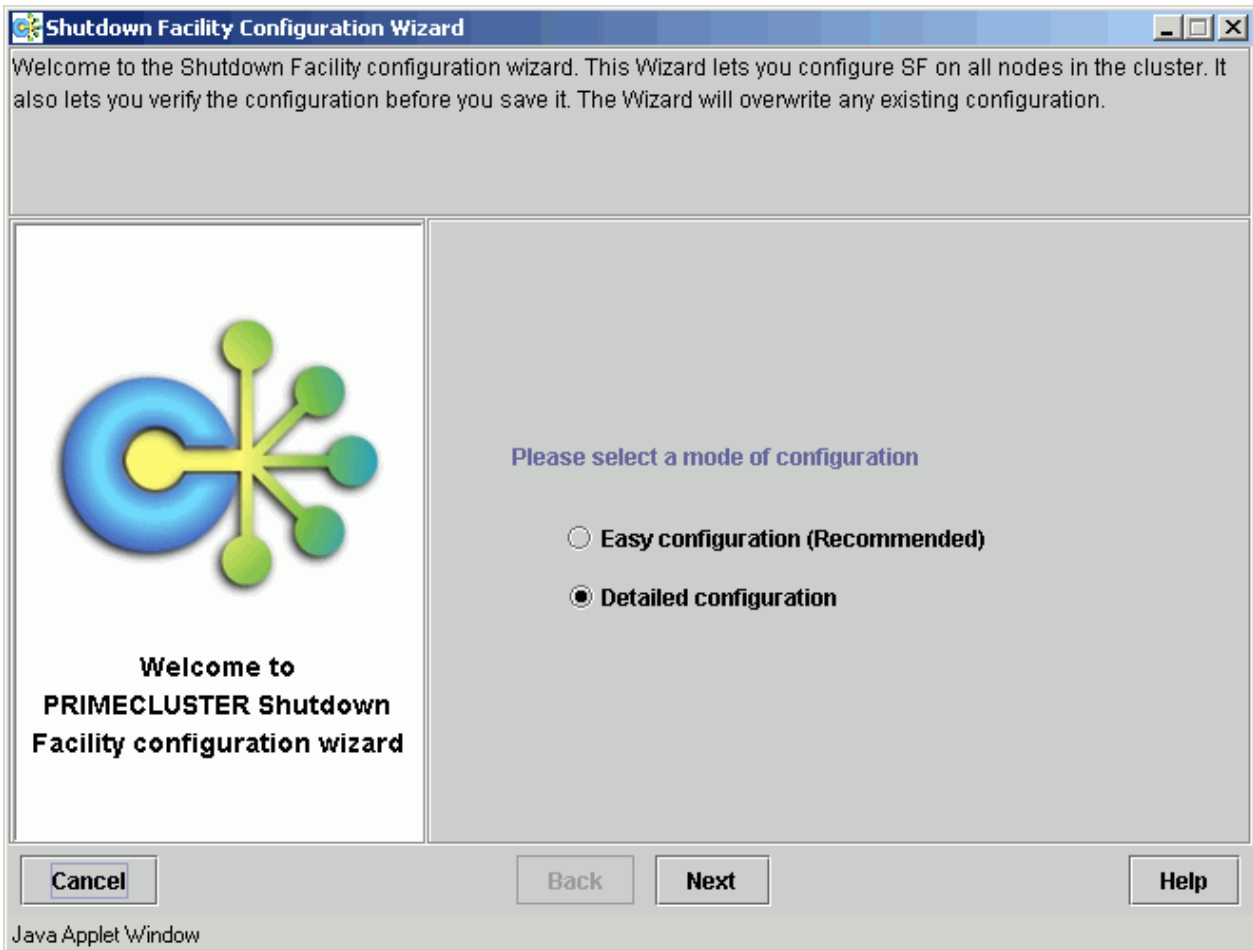
Select *Detailed configuration*, and then click *Next*.
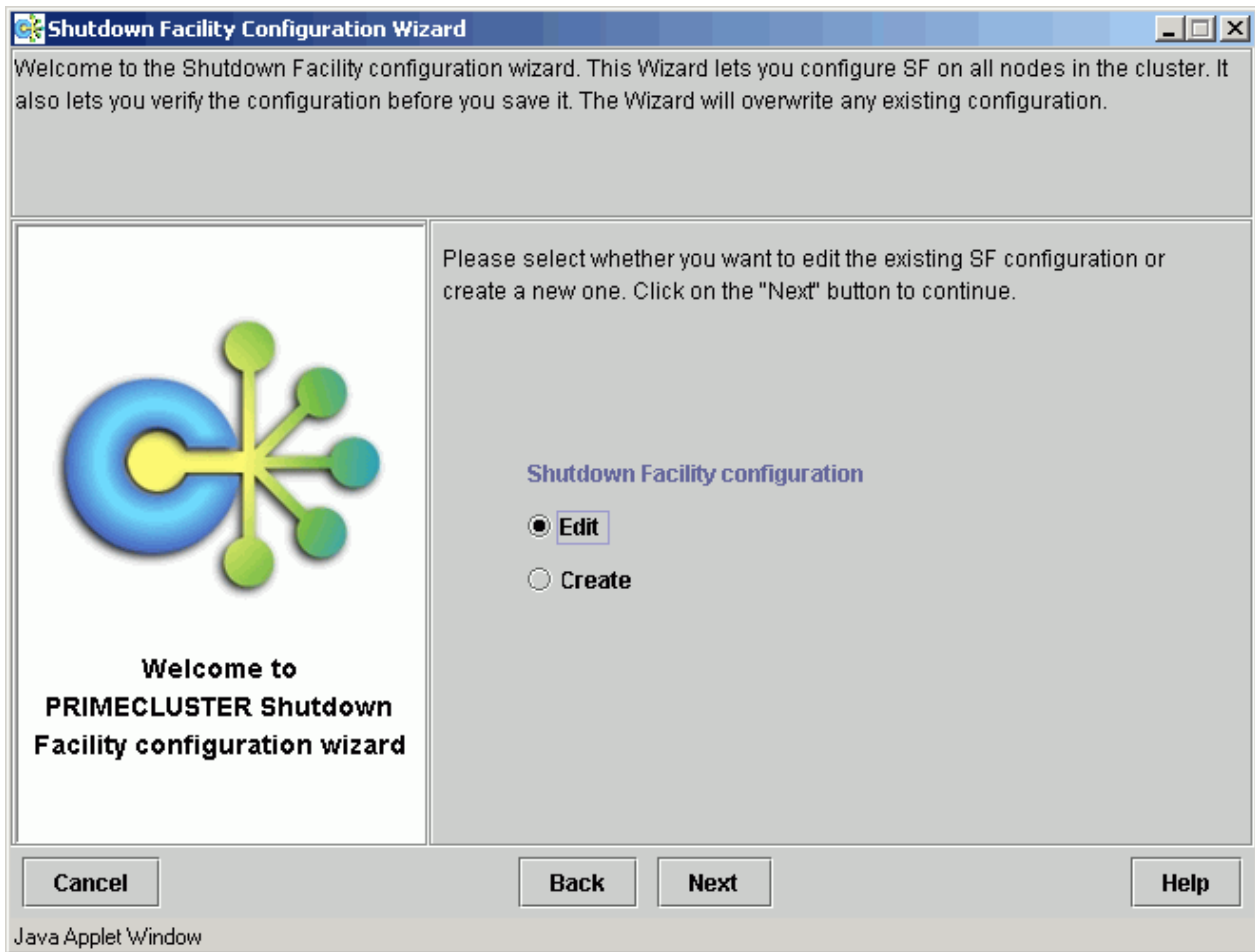
Select *Edit*, and then click *Next*.
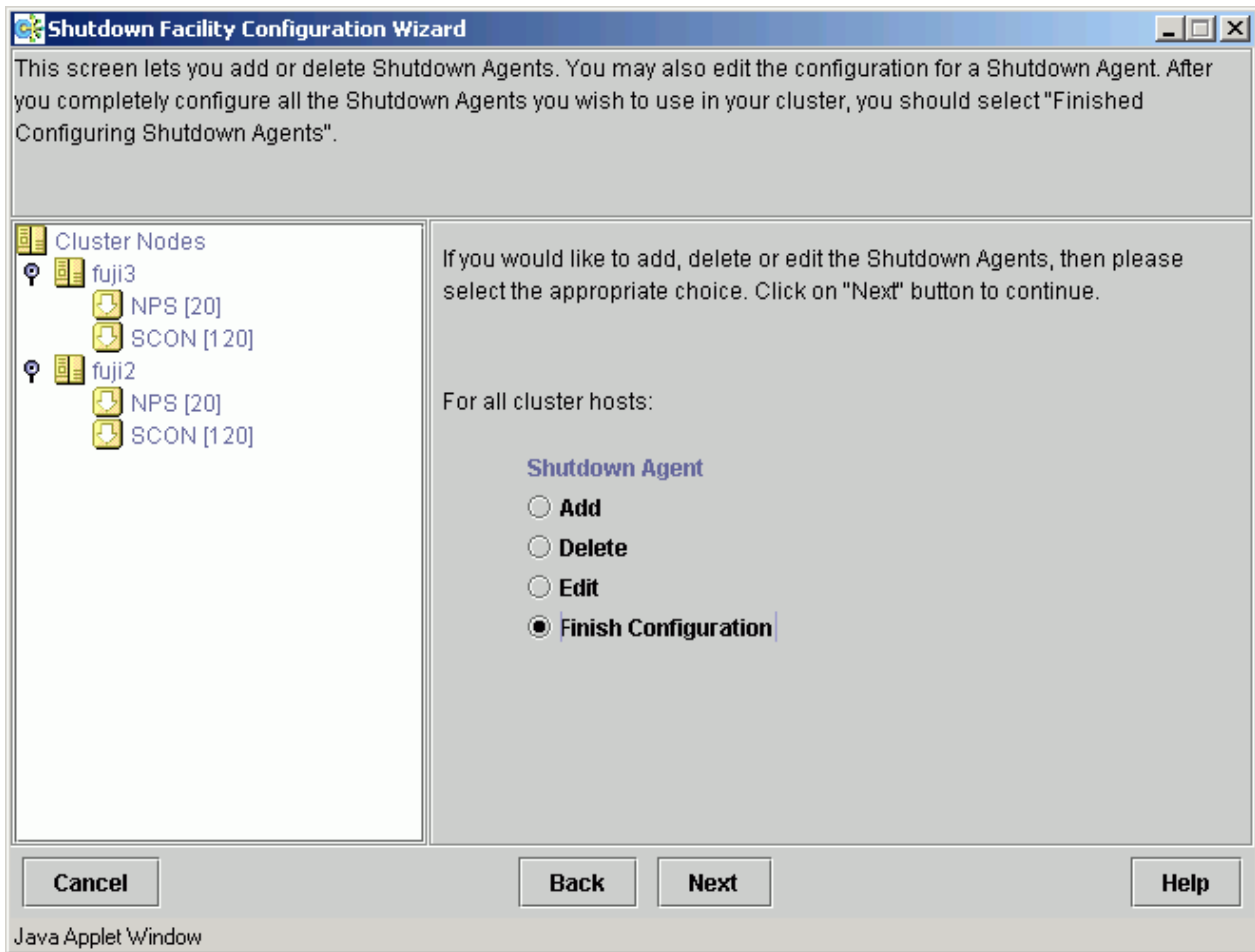
Select *Finish configuration*, and then click *Next*.

Figure 5.12 Changing the Shutdown Agent order
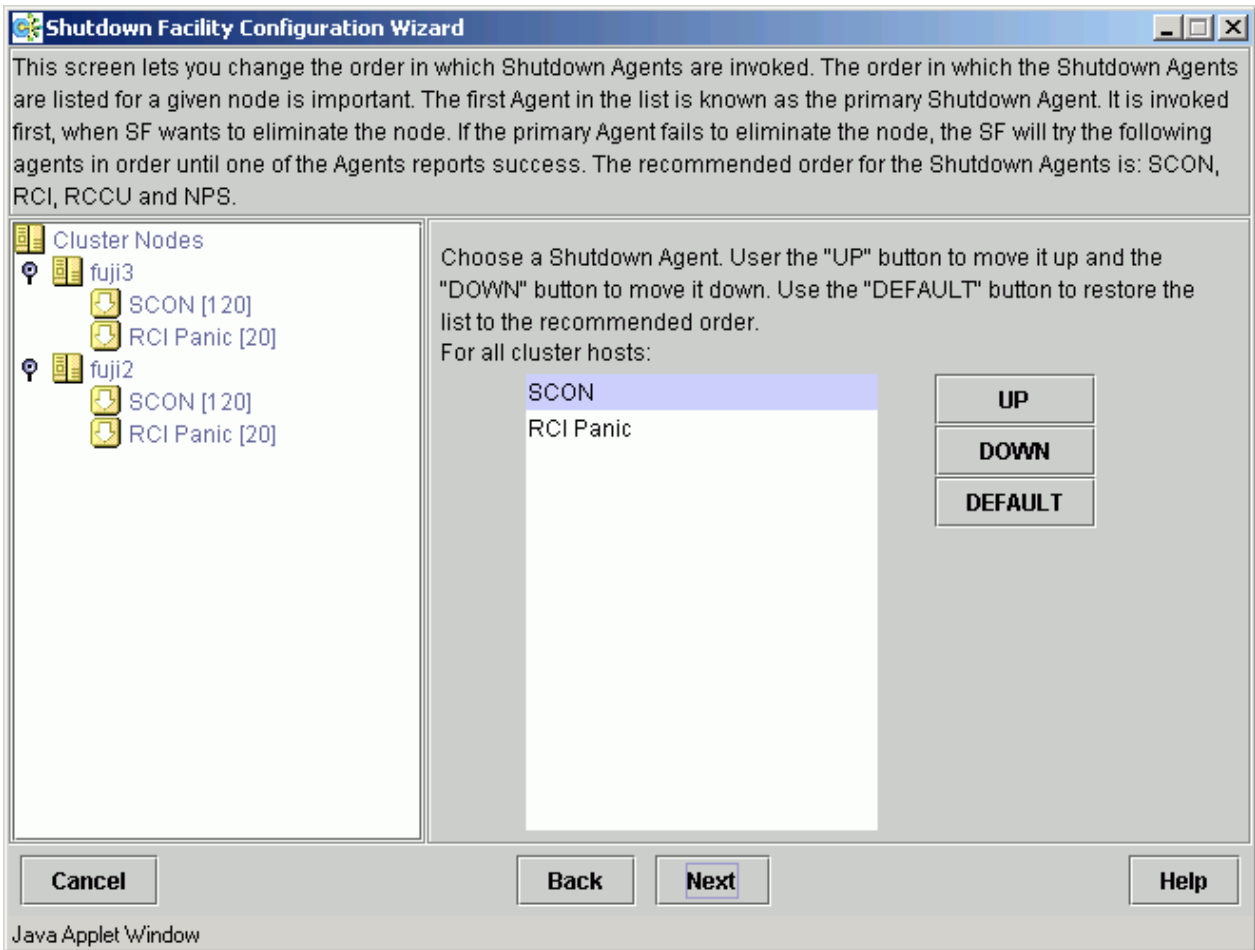
Click *Next*.

![Note icon] Note
. . . .
Do not change the order in which shutdown agents are invoked.
. . . .

Figure 5.13 Timeout value
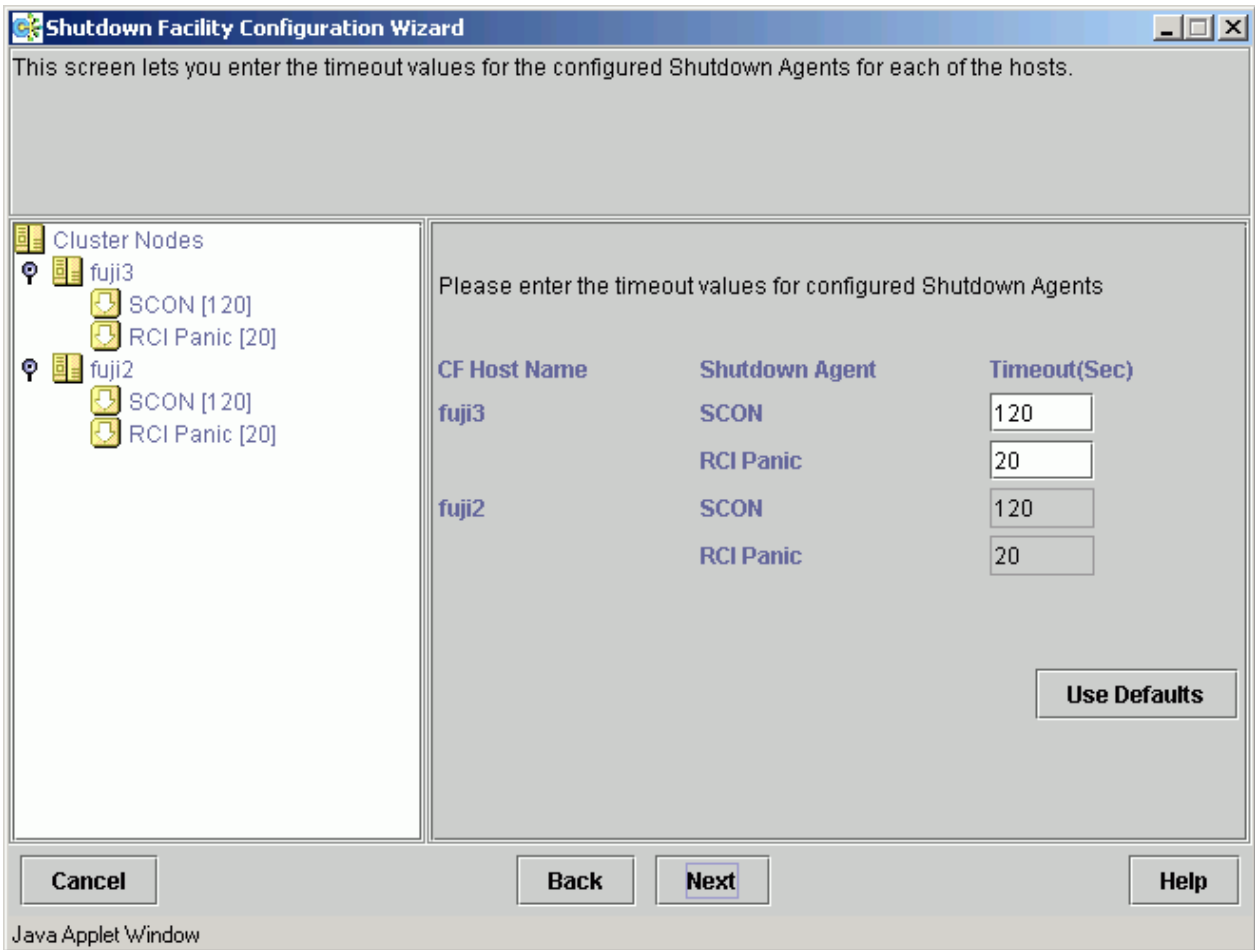
Enter the timeout values in seconds. The default value is 20 seconds.

Once you have completed the configuration, click *Next*.

The "Entering node weights and administrative IP addresses" screen will appear. Click *Next* and then save the settings.

5.1.2.2 For a SPARC Enterprise M3000, M4000, M5000, M8000, or M9000

5.1.2.2.1 Checking Console Configuration

In SPARC Enterprise M3000, M4000, M5000, M8000, and M9000, XSCF is used. The connection method to XSCF as the shutdown facility can be selected from SSH or the telnet.

Default connection is SSH.

Please confirm the following settings concerning XSCF before setting the shutdown facility.

- Commonness

    - The log in user account must be made excluding root for the shutdown facility, and the platadm authority must be given.

- At the SSH connection

    - In XSCF, SSH must be effective in connected permission protocol type from the outside.

    - User inquiries of the first SSH connection (such as generation of the RSA key) must be completed by connecting to XSCF from all the cluster nodes via SSH using the log in user account for the shutdown facility.

- At the telnet connection

    - In XSCF, telnet must be effective in connected permission protocol type from the outside.

Moreover, record the following information on XSCF.

- XSCF IP address or an XSCF host name registered in the "/etc/inet/hosts" file of the node

- Log in user account and password for shutdown facility in XSCF

  *1) When the network routing is set, IP address of XSCF need not be the same to management LAN segment of the cluster node.

5.1.2.2.2 Using the Shutdown Configuration Wizard

The required shutdown agent and the configuration procedure for the shutdown agent vary depending on the hardware machine type.

Check the following combinations of the hardware machine types and shutdown agents, and then set up an appropriate shutdown agent.

- SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 provided by companies other than Fujitsu in Japan

- SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 with logos of both Fujitsu and Oracle provided in other than Japan

    - XSCF Panic

    - Console Break

    - XSCF Reset

- SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 other than above

    - RCI Panic

    - XSCF Panic

    - Console Break

    - RCI Reset

    - XSCF Reset

Setting up the operation environment for the asynchronous monitoring

This setting is required only for the following cases:

- SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 provided by companies other than Fujitsu in Japan

- SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 with logos of both Fujitsu and Oracle provided in other than Japan

You need to set up the operation environment for the asynchronous monitoring beforehand.

Execute the following command in any one of the nodes in the cluster system.

```
# /etc/opt/FJSVcluster/bin/cldevparam -p VendorType 1
```

Execute the following command in all nodes to confirm that the setting is correct.

```
# /etc/opt/FJSVcluster/bin/cldevparam -p VendorType
1
#
```

Starting up the shutdown configuration wizard

From the CF main window of the Cluster Admin screen, select the *Tool* menu and then *Shutdown Facility -> Configuration Wizard*. The shutdown configuration wizard will start.

ℹ️ Information

You can also configure the shutdown facility immediately after you complete the CF configuration with the CF wizard.

The following confirmation popup screen will appear. Click *Yes* to start the shutdown configuration wizard.

Selecting a configuration mode

You can select either of the following two modes to configure the shutdown facility:

- *Easy configuration (recommended)*

- *Detailed configuration*

This section explains how to configure the shutdown facility using *Easy configuration (recommended)*. With this mode, you can configure the PRIMECLUSTER shutdown facility according to the procedure.

Figure 5.14 Selecting the SF configuration mode
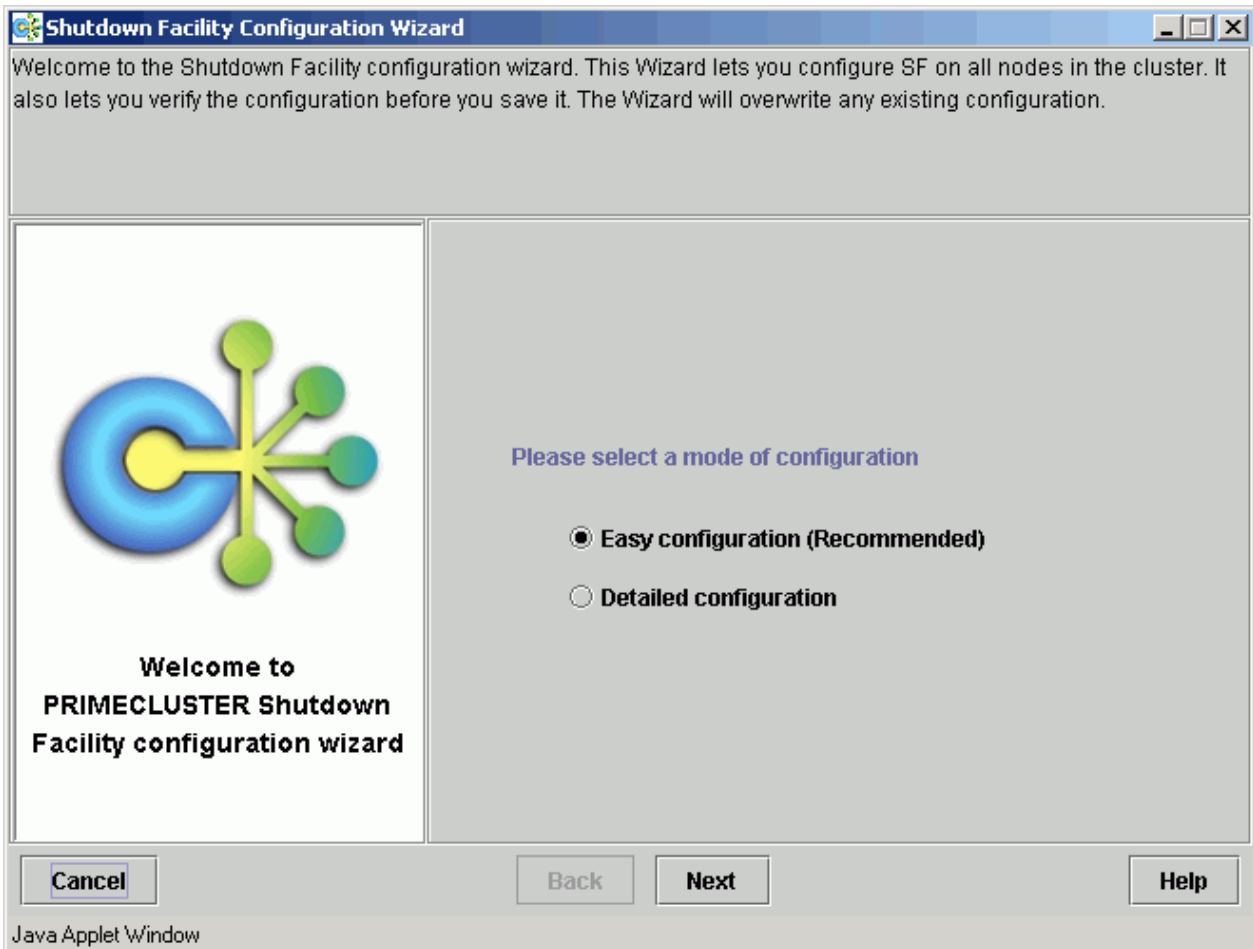
Select *Easy configuration (recommended)* and then click *Next*.

Selecting a shutdown agent

After you confirm the hardware machine type, select an appropriate shutdown agent.

Figure 5.15 Selecting a shutdown agent
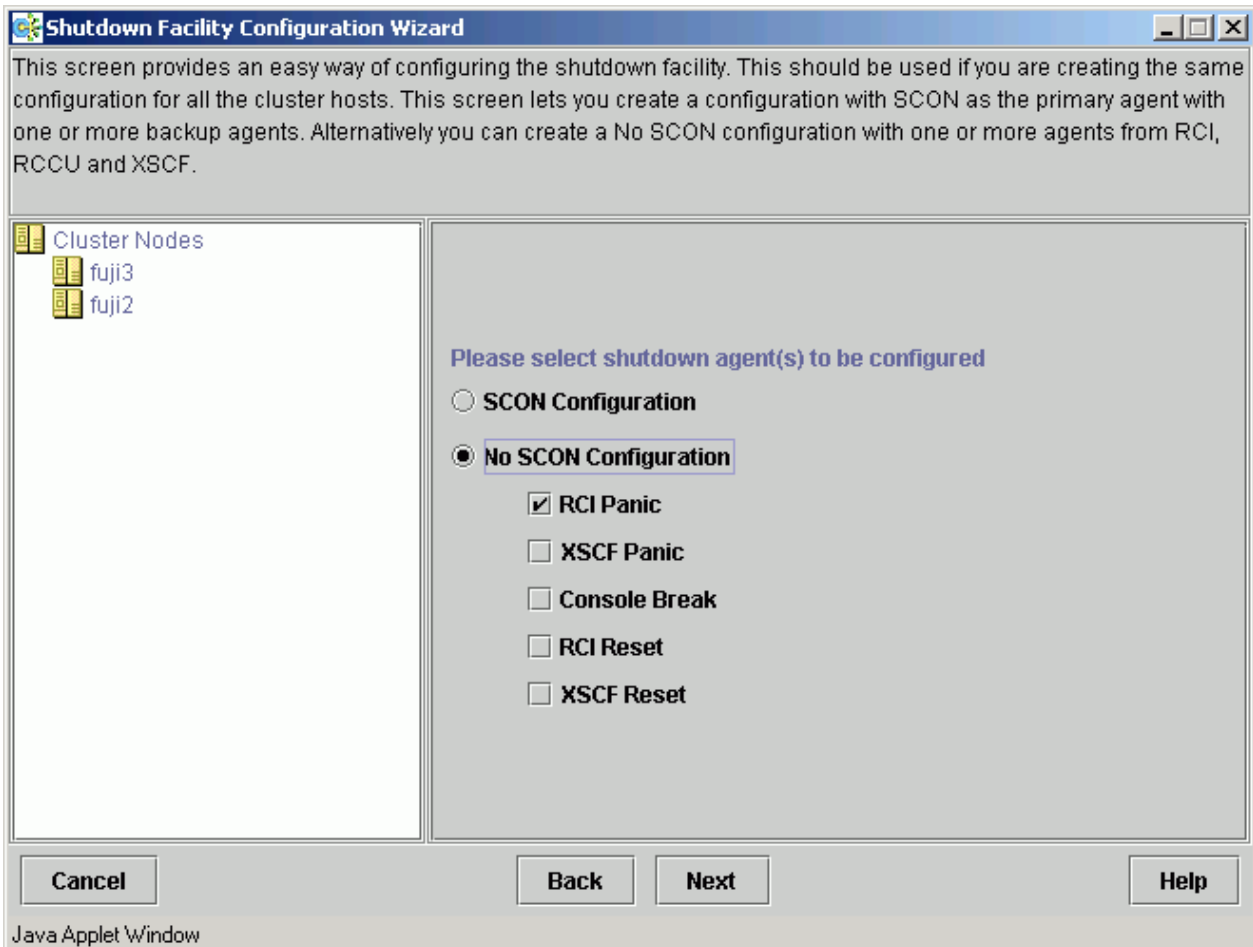
Select *No SCON Configuration*.

Then, select all the following shutdown agents depending on the hardware machine type.

- SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 provided by companies other than Fujitsu in Japan

- SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 with logos of both Fujitsu and Oracle provided in other than Japan

  - XSCF Panic

  - Console Break

  - XSCF Reset

- SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 other than above

  - RCI Panic

  - XSCF Panic

  - Console Break

  - RCI Reset

  - XSCF Reset

Upon the completion of configuration, click *Next*.

Configuring XSCF

If you select *XSCF Panic* or *XSCF Reset* for the shutdown agent, the screen for configuring the XSCF will appear.

Enter the settings for XSCF that you recorded in "5.1.2.2.1 Checking Console Configuration".

Figure 5.16 Configuring XSCF
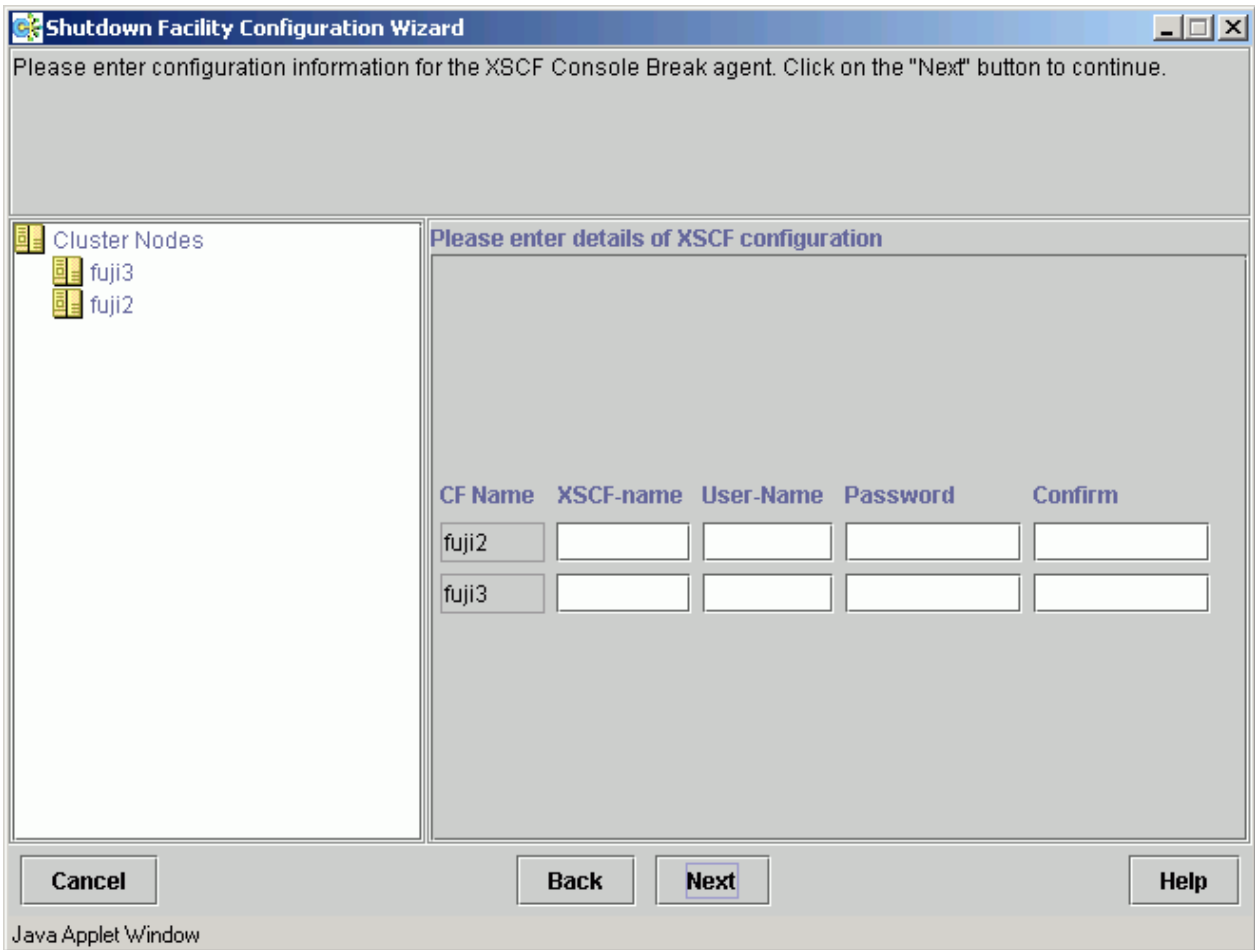
XSCF-name

　　Enter the IP address of XSCF or the host name of XSCF that is registered in the /etc/inet/hosts file.

User-Name

　　Enter a user name to log in to the control port.

Password

　　Enter a password to log in to the control port.

Upon the completion of configuration, click *Next*.

Configuring a Console Break agent

　　If you select *Console Break* as the shutdown agent, the screen used for selecting a Console Break agent will appear.

　　If you use the SPARC Enterprise M3000, M4000, M5000, M8000, or M9000, select XSCF of Console Break agent.

Figure 5.17 Selecting a Console Break agent
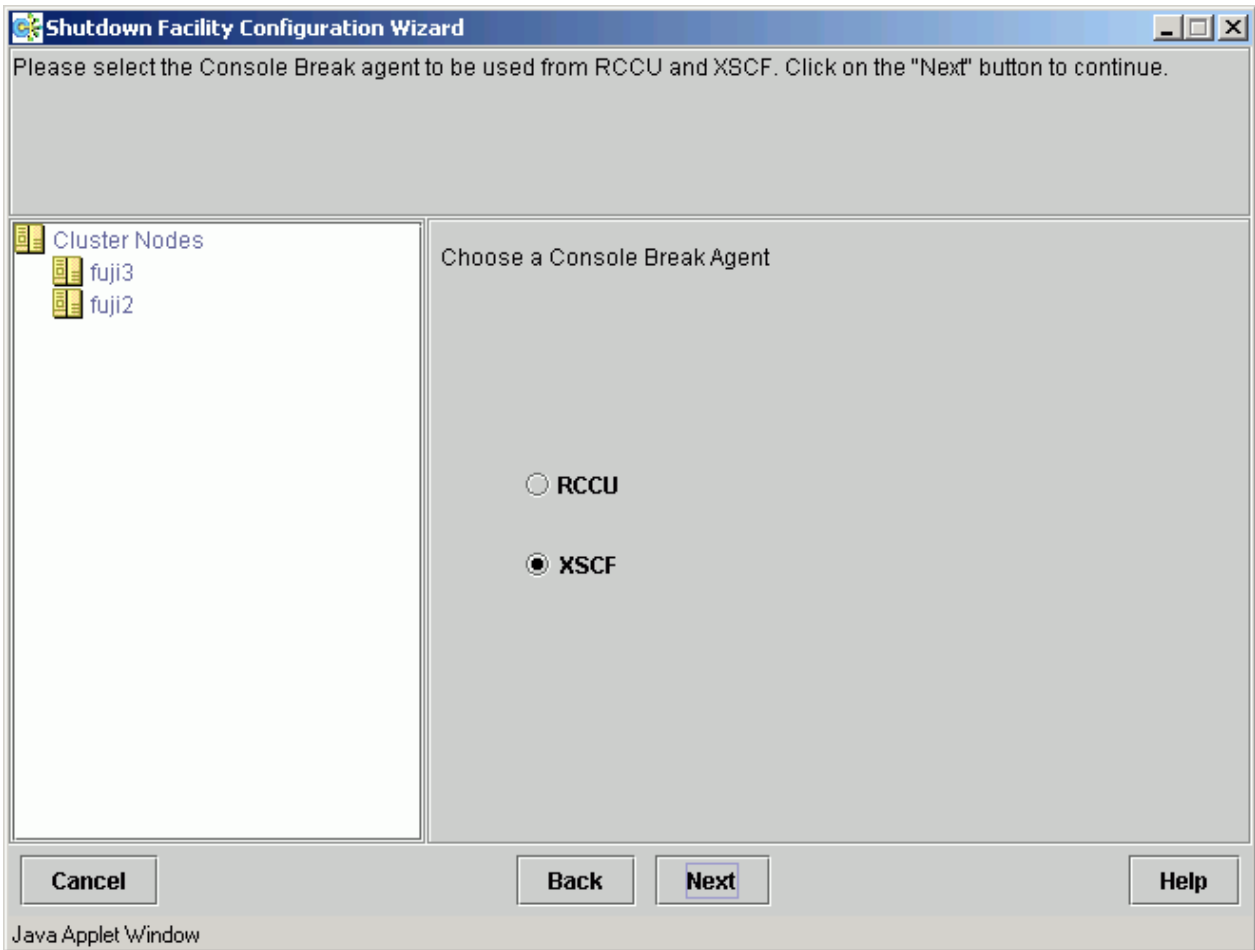
Upon the completion of configuration, click *Next*.

Configuring Wait for PROM

 Note

. . . .

Wait for PROM is currently not supported.
Uncheck the check box if checked, and then click *Next*.

. . . .

Figure 5.18 Configure Wait for PROM
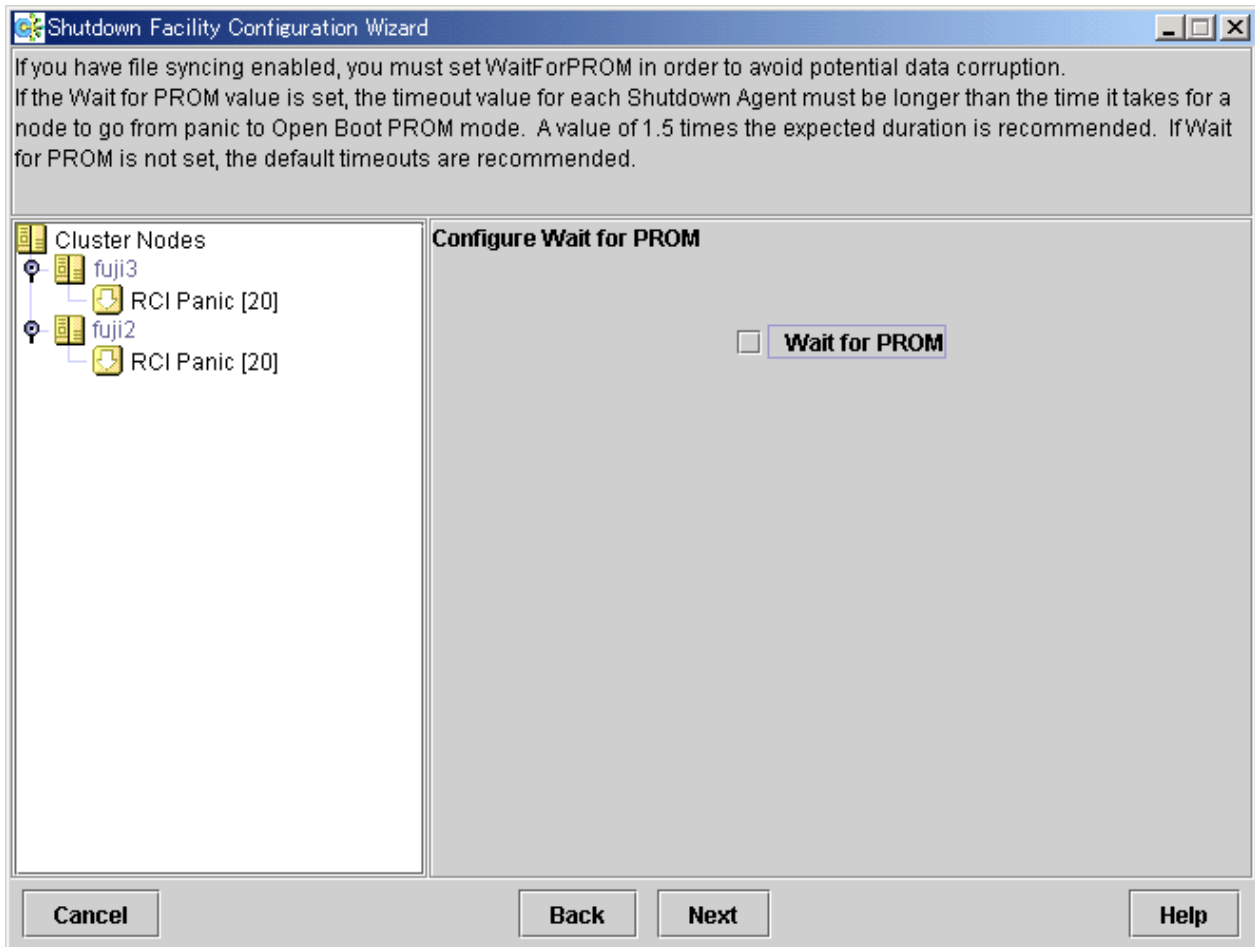
Entering node weights and administrative IP addresses

Enter the weights of the nodes and the IP addresses for the administrative LAN.

Figure 5.19 Entering node weights and administrative IP addresses
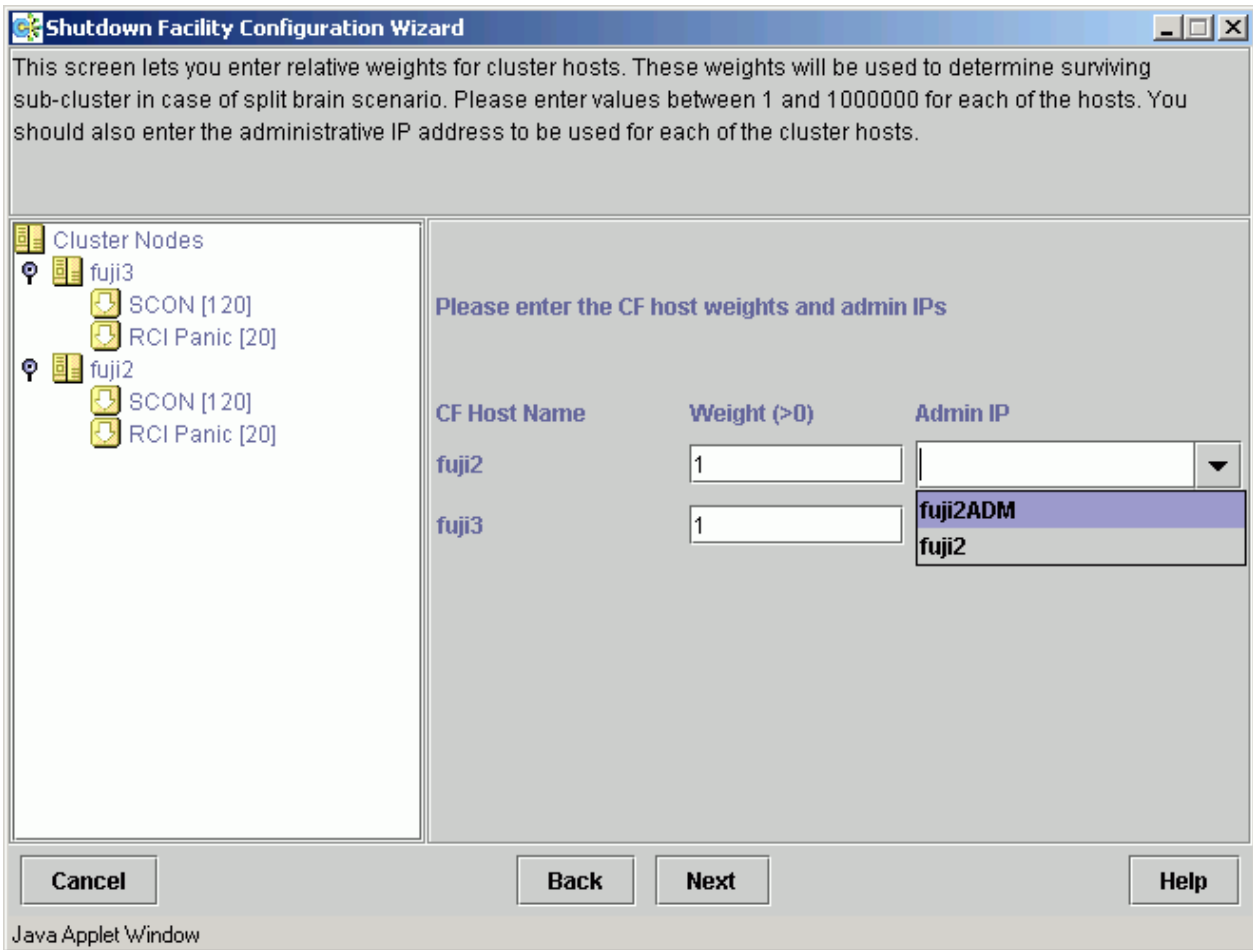
Weight

> Enter the weight of the node that constitutes the cluster. Weight is used to identify the survival priority of the node group that constitutes the cluster. Possible values for each node range from 1 to 300.
>
> For details on survival priority and weight, refer to the explanations below.

Admin IP

> Enter an IP address directly or click the tab to select the host name that is assigned to the administrative IP address.

Upon the completion of configuration, click *Next*.

Survival priority

> Even if a cluster partition occurs due to a failure in the cluster interconnect, all the nodes will still be able to access the user resources. For details on the cluster partition, see "2.2.2.1 Protecting data integrity" in the "*PRIMECLUSTER Concept Guide*".
>
> To guarantee the consistency of the data constituting user resources, you have to determine the node groups to survive and those that are to be forcibly stopped.
>
> The weight assigned to each node group is referred to as a "Survival priority" under PRIMECLUSTER.
>
> The greater the weight of the node, the higher the survival priority. Conversely, the less the weight of the node, the lower the survival priority. If multiple node groups have the same survival priority, the node group that includes a node with the name that is first in alphabetical order will survive.

> Survival priority can be found in the following calculation:

> Survival priority = SF node weight + ShutdownPriority of userApplication

SF node weight (Weight):

> Weight of node. Default value = 1. Set this value while configuring the shutdown facility.

userApplication ShutdownPriority:

Set this attribute when userApplication is created. For details on how to change the settings, see "8.1.2 Changing the Operation Attributes of a Cluster Application".

![See icon] **See**

. . . .

For details on the ShutdownPriority attribute of userApplication, see "6.6.5 Attributes".

. . . .

Survival scenarios

The typical scenarios that are implemented are shown below:

[Largest node group survival]

- Set the weight of all nodes to 1 (default).

- Set the attribute of ShutdownPriority of all user applications to 0 (default).

[Specific node survival]

- Set the "weight" of the node to survive to a value more than double the total weight of the other nodes.

- Set the ShutdownPriority attribute of all user applications to 0 (default).

In the following example, node1 is to survive:

[Specific application survival]

- Set the "weight" of all nodes to 1 (default).

- Set the ShutdownPriority attribute of the user application whose operation is to continue to a value more than double the total of the ShutdownPriority attributes of the other user applications and the weights of all nodes.

  In the following example, the node for which app1 is operating is to survive:

	Node group1	Node group2		
	node1	node2	node3	node4
weight of node	1	1	1	1
ShutdownPriority of app1 =20	20			
ShutdownPriority of app2 =1			1	
ShutdownPriority of app3 =1				1
Live priority	21	5		

Saving the configuration

Confirm and then save the configuration. In the left-hand panel of the window, those nodes that constitute the cluster are displayed, as are the shutdown agents that are configured for each node.

Figure 5.20 Saving the configuration
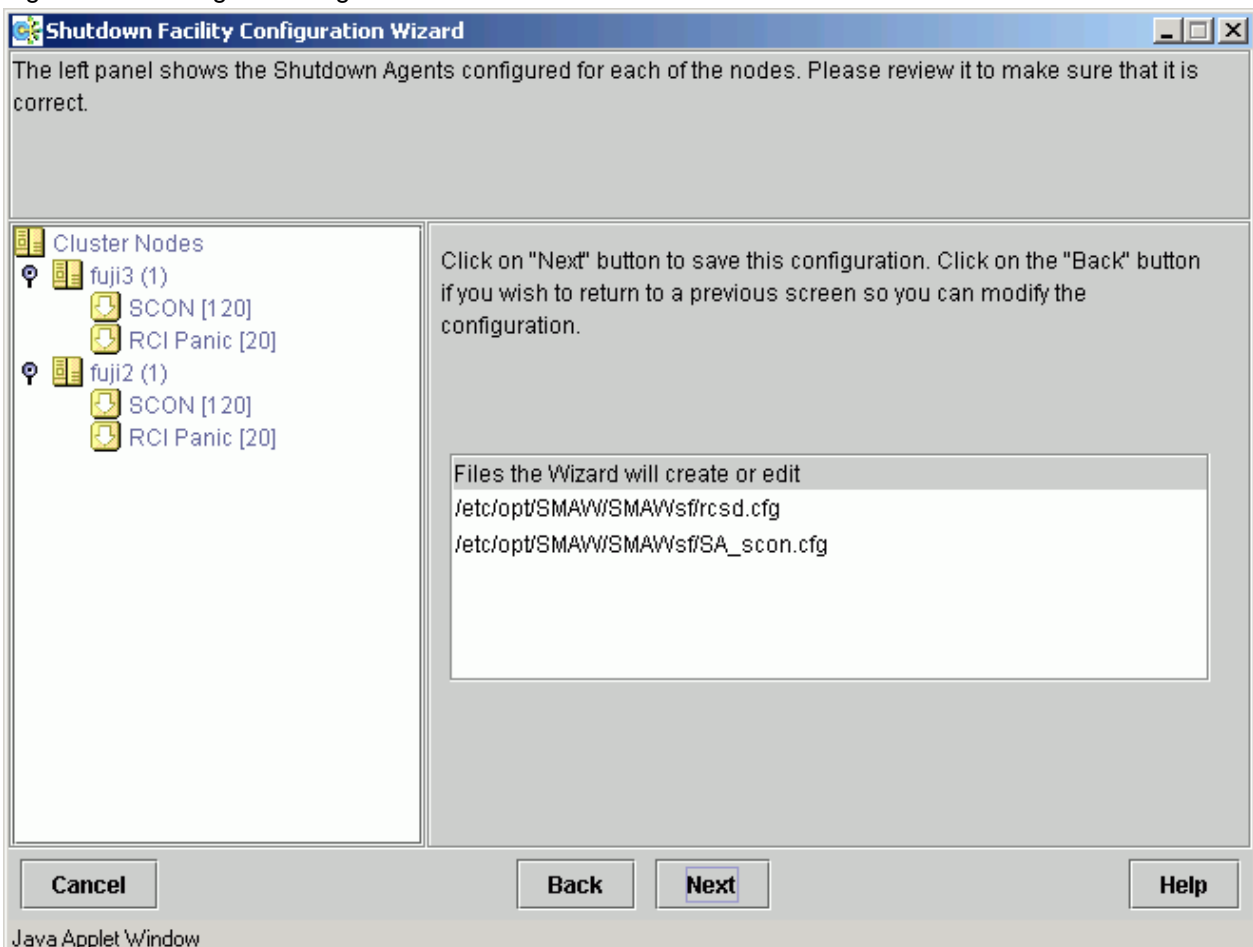
Click *Next*. A popup screen will appear for confirmation. Select *Yes* to save the setting.

Displaying the configuration of the shutdown facility

If you save the setting, a screen displaying the configuration of the shutdown facility will appear. On this screen, you can confirm the configuration of the shutdown facility on each node by selecting each node in turn.

ℹ️ Information

You can also view the configuration of the shutdown facility by selecting *Shutdown Facility -> Show Status* from the *Tool* menu.

Figure 5.21 Show Status
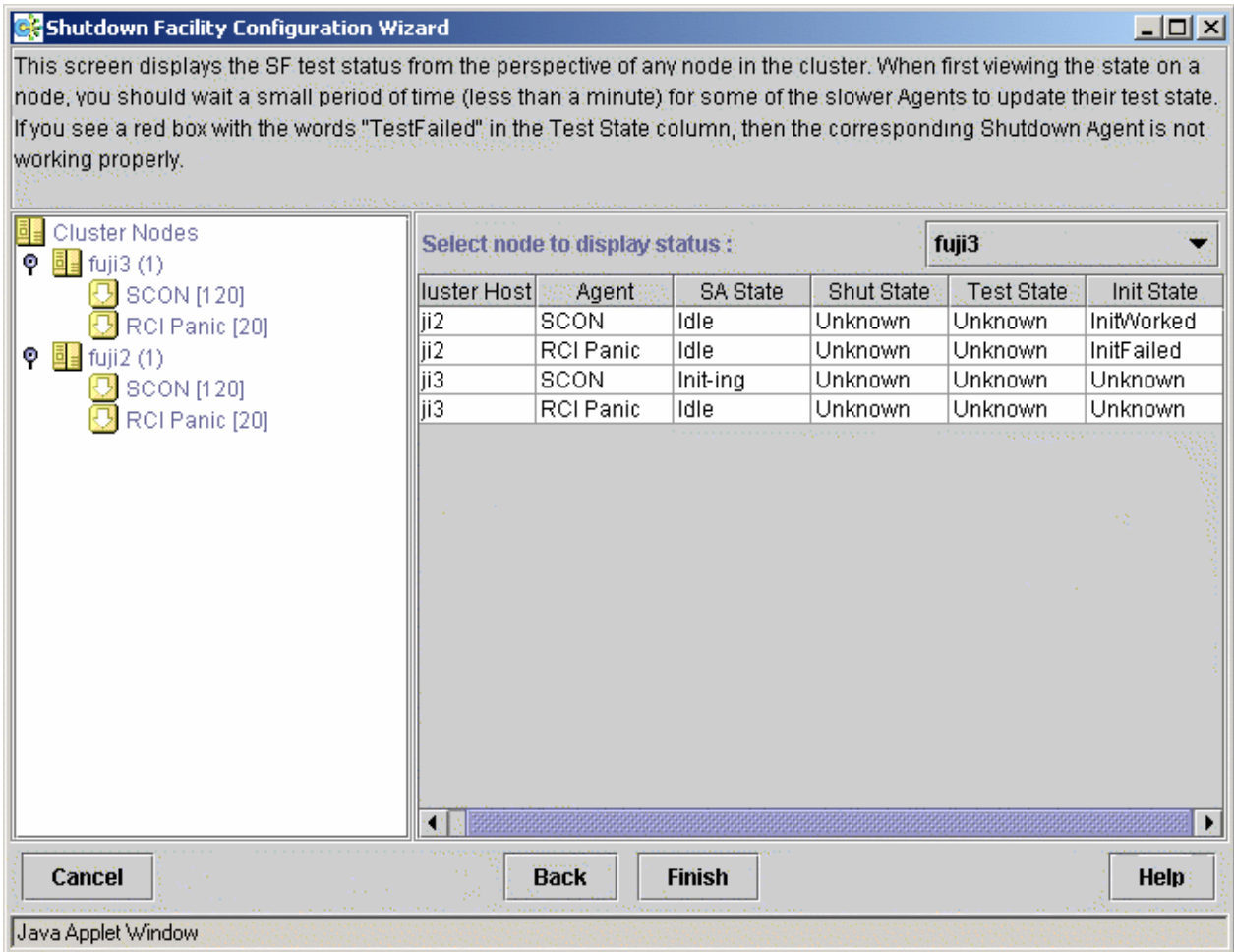
Shut State

"Unknown" is shown during normal system operation. If an error occurs and the shutdown facility stops the relevant node successfully, "Unknown" will change to "KillWorked".

Test State

Indicates the state in which the path to shut down the node is tested when a node error occurs. If the test of the path has not been completed, "Unknown" will be displayed. If the configured shutdown agent operates normally, "Unknown" will be changed to "TestWorked".

Init State

Indicates the state in which the shutdown agent is initialized.

To exit the configuration wizard, click Finish. Click Yes in the confirmation popup screen that appears.

> 📌 **Note**
>
> ....
>
> Confirm that the shutdown facility is operating normally.
>
> - If "InitFailed" is displayed in the Initial state even when the configuration of the shutdown facility has been completed or if "Unknown" is displayed in the Test state or "TestFailed" is highlighted in red, the agent or hardware configuration may contain an error. Check the /var/adm/messages file and the console for an error message. Then, apply appropriate countermeasures as instructed the message that is output.
>
> - If connection to XSCF is telnet, the test state becomes TestFailed at this point in time. Confirm that the shutdown facility is operating normally, after the "Setting of the connection method to the XSCF".
>
> ....

> 📖 **See**
>
> ....
>
> For details on how to respond to the error messages that may be output, see the following manual.
>
> - "12.12 Monitoring Agent messages" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*".
>
> ....

5.1.2.2.3 Specifying the Timeout Value

Confirm that the timeout value for each shutdown agent is as shown below. You can confirm the timeout value in the left-hand panel of the shutdown configuration wizard screen.

<How to calculate a timeout value>

- For XSCF Panic/XSCF Break

    - 4 or fewer nodes

    Timeout value = 20 (seconds)

    - 5 or more nodes

    Timeout value = 6 x number of cluster nodes + 2 (seconds)

    Example) For 5 nodes: 6 x 5 + 2 = 32 (seconds)

- For XSCF Reset

    - 4 or fewer nodes

    Timeout value = 40 (seconds)

    - 5 or more nodes

    Timeout value = 6 x number of cluster nodes + 22 (seconds)

    Example) For 5 nodes: 6 x 5 + 22 = 52 (seconds)

- For RCI

    Timeout value = 20 (seconds)

If the value has not been set as explained above, set a timeout value as follows.

Setting a timeout value

In the CF main window of Cluster Admin, select *Shutdown Facility -> Configuration Wizard* from the *Tool* menu and then start the configuration wizard.

Figure 5.22 Selecting a configuration mode
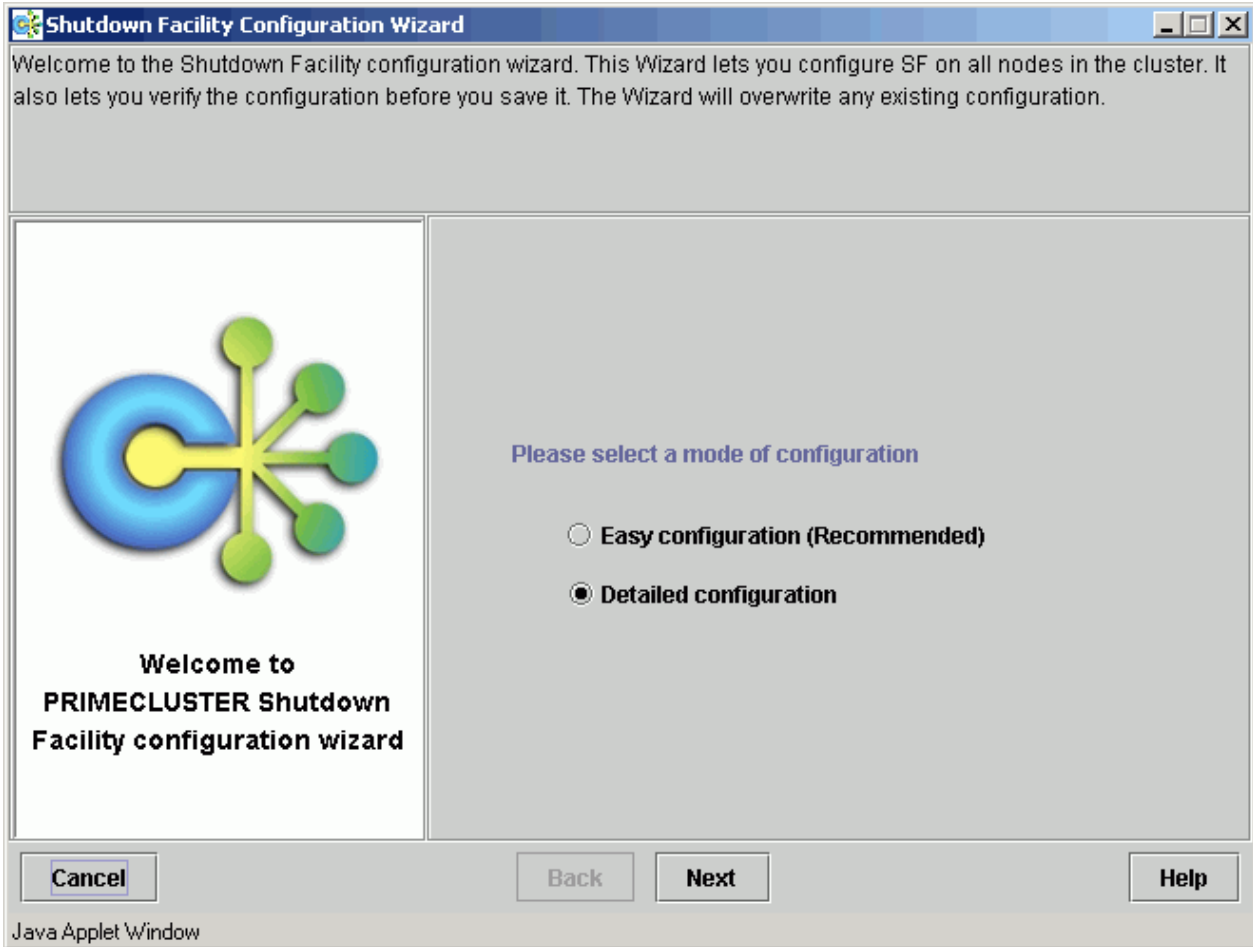
Select Detailed configuration, and then click *Next*.
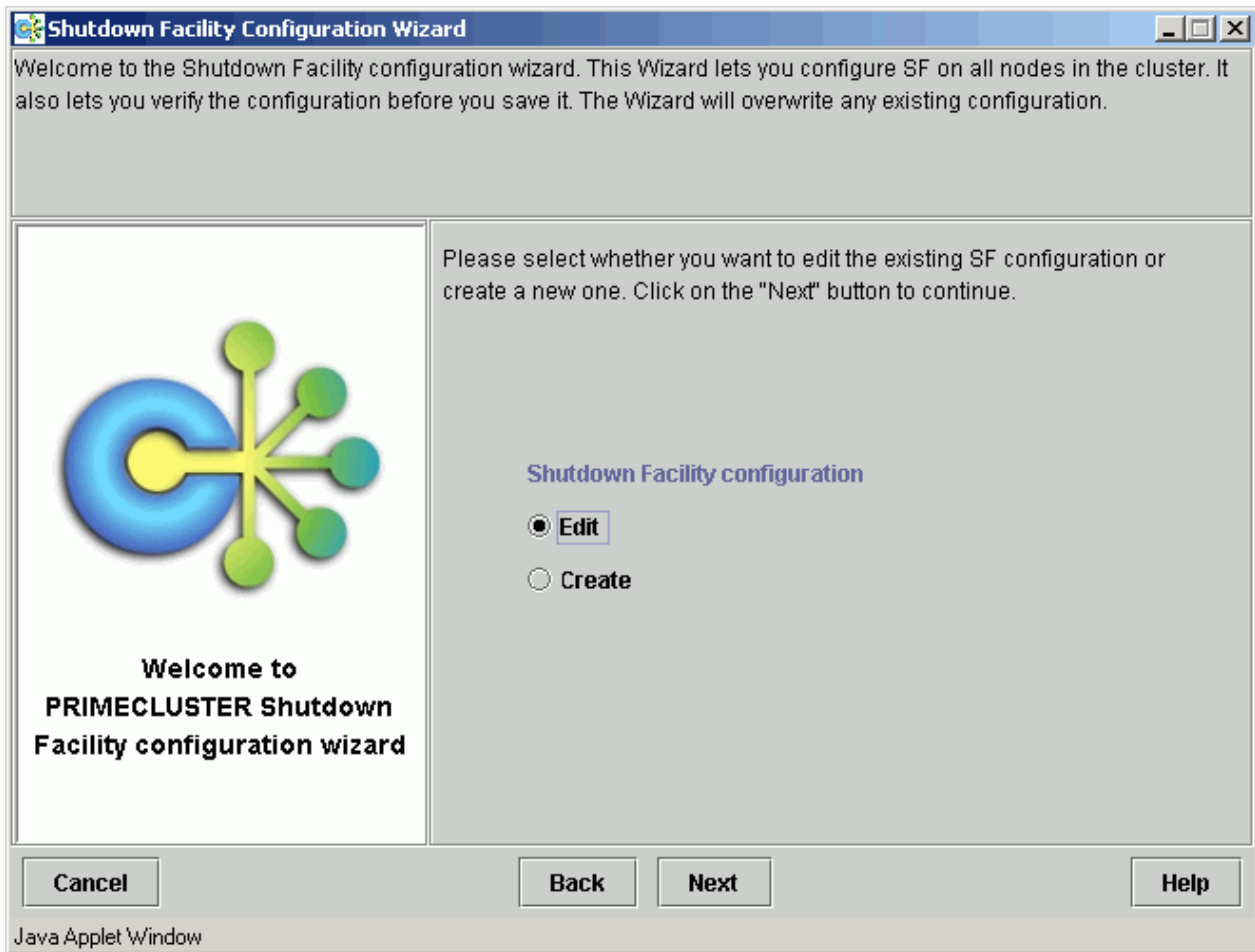
Select *Edit*, and then click *Next*.
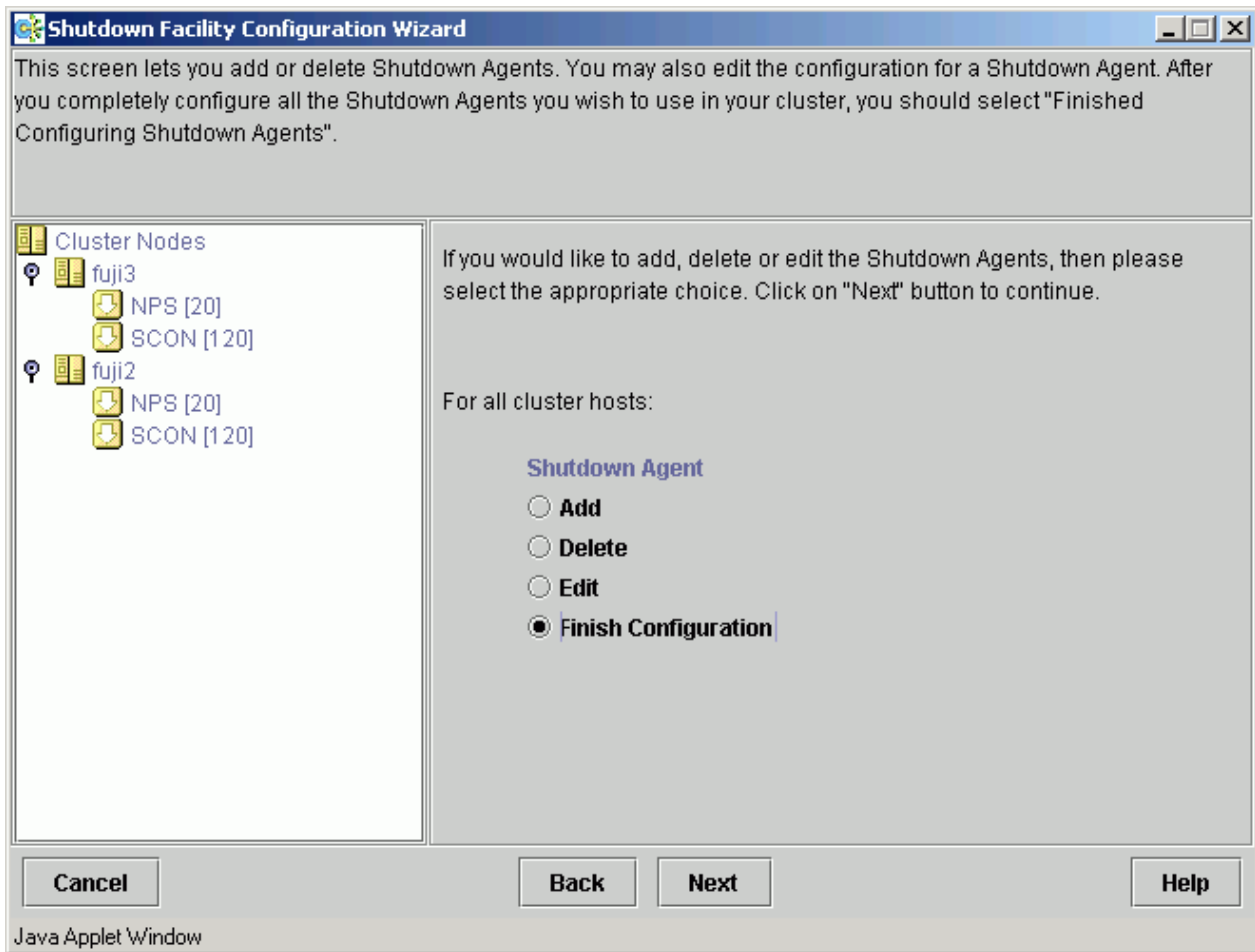
Select *Finish Configuration*, and then click *Next*.

Figure 5.23 Changing the Shutdown Agent order
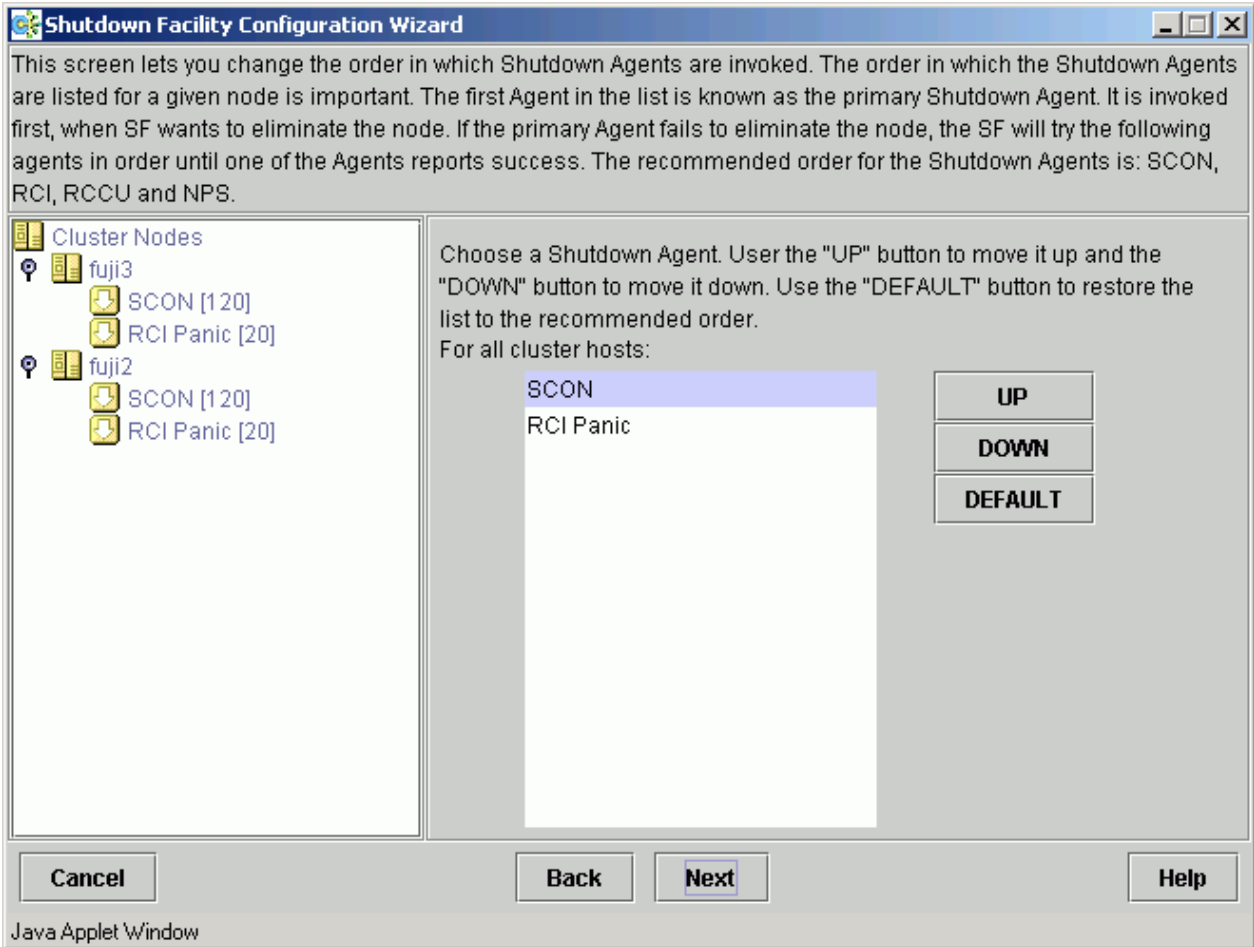
Click *Next*.

Note

Do not change the order in which shutdown agents are invoked.

Figure 5.24 Timeout value
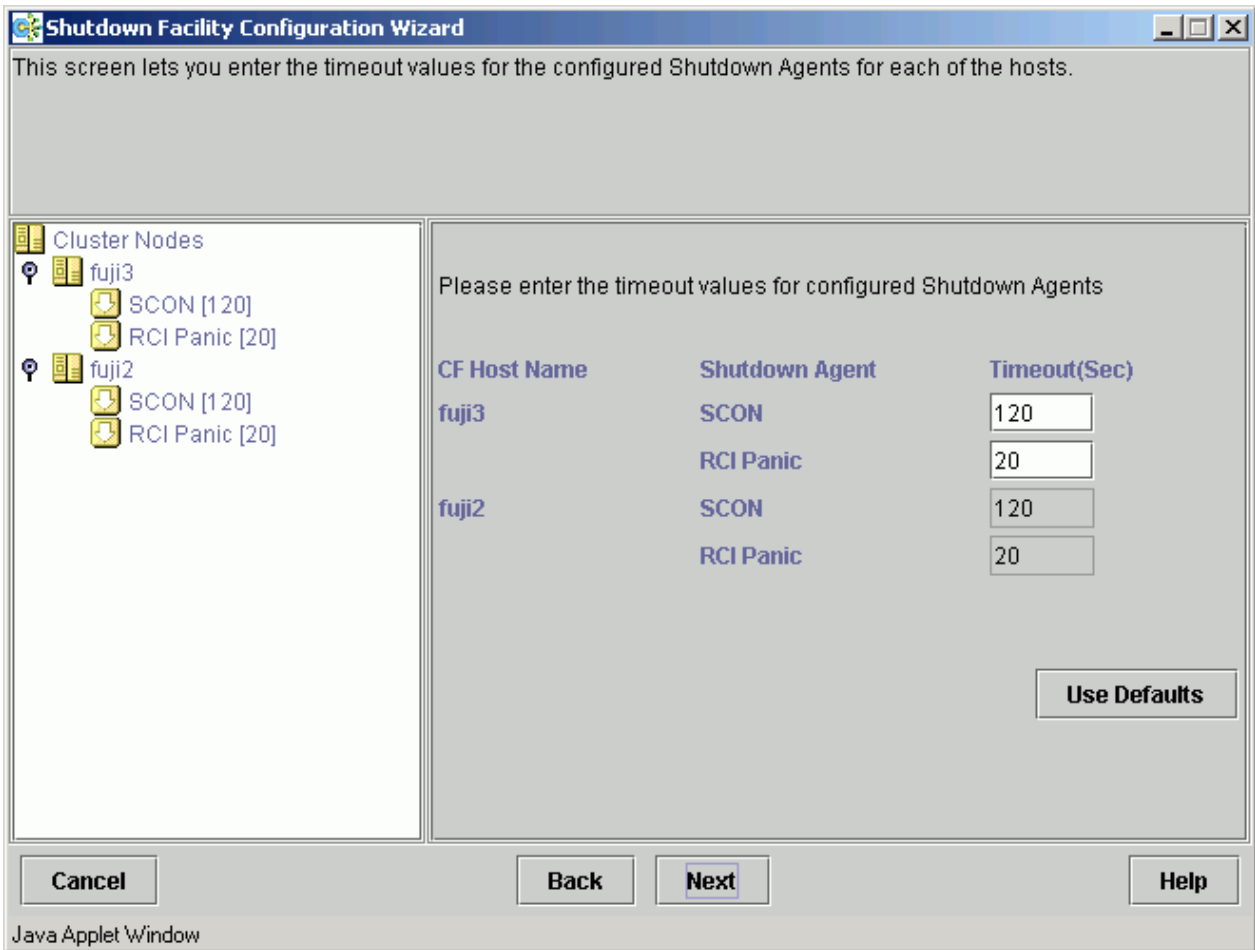
Change the timeout values to the values (in seconds) calculated by the formula in <How to calculate a timeout value>. The default value is displayed as 20 seconds.

Upon the completion of configuration, click *Next*.

The "Entering node weights and administrative IP addresses" screen will appear. Click *Next* and then save the settings.

5.1.2.2.4 Setting of the connection method to the XSCF

The default of setting of the connection method to the XSCF is SSH connection, in the SPARC Enterprise M3000, M4000, M5000, M8000, or M9000. The procedure when changing to the telnet connection is the following.

Change of the connection method

Execute the following command in all nodes to change a connection method.

```
# /etc/opt/FJSVcluster/bin/clrccusetup -m -t telnet
```

Example)

```
# /etc/opt/FJSVcluster/bin/clrccusetup -m -t telnet <RETURN>
# /etc/opt/FJSVcluster/bin/clrccusetup -l <RETURN>
Device-name cluster-host-name IP-address host-name user-name connection-type
-----------------------------------------------------------------------
xscf         fuji2              xscf2      1          xuser       telnet
xscf         fuji3              xscf3      1          xuser       telnet
```

Starting up the shutdown facility

Execute the following command in each node, and confirm the shutdown facility has started.

```
# /opt/SMAW/bin/sdtool -s
```

If the state of configuration of shutdown facility is displayed, shutdown facility is started.

If "The RCSD is not running" is displayed, shutdown facility is not started.

If shutdown facility is started, execute the following command, and restart the shutdown facility.

```
# /opt/SMAW/bin/sdtool -r
```

If shutdown facility is not started, execute the following command, and start the shutdown facility.

```
# /opt/SMAW/bin/sdtool -b
```

5.1.2.3 For a SPARC Enterprise T1000, T2000

5.1.2.3.1 Checking Console Configuration

ALOM in console can be used by SPARC Enterprise T1000 or T2000.

Confirm the following settings concerning ALOM before setting the shutdown facility.

- The log in user account must be made for the shutdown facility, and c level (console access) authority must be given.

- In connected permission protocol type from the outside, the telnet must be effective.

- The following ALOM composition variables must not have been changed from default.

    - if_emailalerts : false(default)

    - sc_clieventlevel : 2(default)

    - sc_cliprompt : sc(default)

📋 Note

- Connected permission from the outside to ALOM is default and SSH. In that case, it is not supported in the shutdown facility.

- When the connection to ALOM is a serial port connection alone, it is not supported in the shutdown facility.

Moreover, record the following information on ALOM.

- ALOM IP address(*1) or an ALOM host name registered in the "/etc/inet/hosts" file.

- User name used to log in the ALOM.

- Password used to log in the ALOM.

*1) When the network routing is set, Internet Protocol address of ALOM need not be the same to management LAN segment of the cluster node.

📖 See

For information on how to configure and confirm ALOM, see the "*Advanced Lights out Management (ALOM) CMT Guide*".

5.1.2.3.2 Setting by CLI

You cannot execute the setting of the shutdown facility of SPARC Enterprise T1000 or T2000 in the shutdown configuration Wizard.

The following confirmation popup screen will appear, immediately after you complete the CF configuration with the CF wizard. Click <*No*> to end the shutdown configuration wizard.

Set up the shutdown facility according to the following procedure:

Configuring the shutdown facility

Create the "/etc/opt/SMAW/SMAWsf/rcsd.cfg" by the following contents in all nodes.

```
CFNameX,weight=weight,admIP=myadmIP:agent=SA_sunF,timeout=timeout
CFNameX,weight=weight,admIP=myadmIP:agent=SA_sunF,timeout=timeout
```

CFNameX

> For the CF node name of the cluster host.

weight

> For the Weight of node.

myadmIP

> For the IP address of the Administrative LAN of the own node.

agent

> For the name of shutdown agent.

> You appoint "SA_sunF" of the ALOM shutdown agent, in SPARC Enterprise T1000 or T2000.

timeout

> For the timeout duration of the shutdown agent.

> For the 40 seconds in SPARC Enterprise T1000 or T2000.

Example)

```
node1,weight=1,admIP=10.20.30.100:agent=SA_sunF,timeout=40
node2,weight=1,admIP=10.20.30.200:agent=SA_sunF,timeout=40
```

Survival priority

> Even if a cluster partition occurs due to a failure in the cluster interconnect, all the nodes will still be able to access the user resources. For details on the cluster partition, see "2.2.2.1 Protecting data integrity" in the "*PRIMECLUSTER Concept Guide*".
> To guarantee the consistency of the data constituting user resources, you have to determine the node groups to survive and those that are to be forcibly stopped.
> The weight assigned to each node group is referred to as a "Survival priority" under PRIMECLUSTER.
> The greater the weight of the node, the higher the survival priority. Conversely, the less the weight of the node, the lower the survival priority. If multiple node groups have the same survival priority, the node group that includes a node with the name that is first in alphabetical order will survive.

> Survival priority can be found in the following calculation:

> Survival priority = SF node weight + ShutdownPriority of userApplication

SF node weight (Weight):

Weight of node. Default value = 1. Set this value while configuring the shutdown facility.

userApplication ShutdownPriority:

Set this attribute when userApplication is created. For details on how to change the settings, see "8.1.2 Changing the Operation Attributes of a Cluster Application".
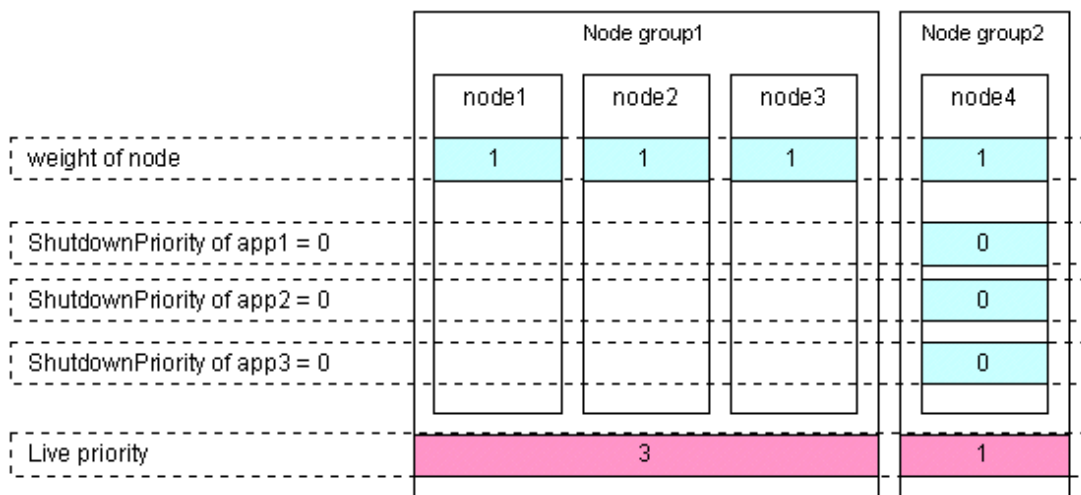
🗄 See
. . . .

For details on the ShutdownPriority attribute of userApplication, see "6.6.5 Attributes".
. . . .

Survival scenarios

The typical scenarios that are implemented are shown below:

[Largest node group survival]

- Set the weight of all nodes to 1 (default).

- Set the attribute of ShutdownPriority of all user applications to 0 (default).

	Node group1			Node group2
	node1	node2	node3	node4
weight of node	1	1	1	1
ShutdownPriority of app1 = 0				0
ShutdownPriority of app2 = 0				0
ShutdownPriority of app3 = 0				0
Live priority	3			1

[Specific node survival]

- Set the "weight" of the node to survive to a value more than double the total weight of the other nodes.

- Set the ShutdownPriority attribute of all user applications to 0 (default).

In the following example, node1 is to survive:

[Specific application survival]

- Set the "weight" of all nodes to 1 (default).

- Set the ShutdownPriority attribute of the user application whose operation is to continue to a value more than double the total of the ShutdownPriority attributes of the other user applications and the weights of all nodes.

In the following example, the node for which app1 is operating is to survive:

Configuring the ALOM shutdown agent

Enter the settings for ILOM that you recorded in "5.1.2.3.1 Checking Console Configuration".

Create the "/etc/opt/SMAW/SMAWsf/rcsd.cfg" by the following contents in all nodes.

```
SystemContorollerTag SystemControllerHostName SystemControllerLogin PWord
void void CFBAMEX
SystemContorollerTag SystemControllerHostName SystemControllerLogin PWord
void void CFBAMEX
```

SystemControllerTag

Type of system controller.
If it is the ALOM shutdown agent, set "system-controller-alom-2k".

SystemControllerHostName

> IP address of ALOM.

SystemControllerLogin

> For the admin user name that you defined at the time of ALOM setting.

PWord

> For the admin user password that you defined at the time of ALOM setting.

CFNameX

> For the CF node name of the cluster host.

Example)

```
system-controller-alom-2k 10.20.30.100 admin admin01 void void node1
system-controller-alom-2k 10.20.30.200 admin admin01 void void node2
```

Starting up the shutdown facility

Execute the following command in each node, and confirm the shutdown facility has started.

```
# /opt/SMAW/bin/sdtool -s
```

If the state of configuration of shutdown facility is displayed, shutdown facility is started.

If "The RCSD is not running" is displayed, shutdown facility is not started.

If shutdown facility is started, execute the following command, and restart the shutdown facility.

```
# /opt/SMAW/bin/sdtool -r
```

If shutdown facility is not started, execute the following command, and start the shutdown facility.

```
# /opt/SMAW/bin/sdtool -b
```

Displaying the configuration of the shutdown facility

Execute the following command in each node, and confirm the configuration of the shutdown facility.

```
# /opt/SMAW/bin/sdtool -s
Cluster Host     Agent      SA State    Shut State    Test State    Init State
------------------------------------------------------------------------------
node1            SA_sunF    Idle        Unknown       TestWorked    InitWorked
node2            SA_sunF    Idle        Unknown       TestWorked    InitWorked
```

Shut State

> "Unknown" is shown during normal system operation. If an error occurs and the shutdown facility stops the relevant node successfully, "Unknown" will change to "KillWorked".

Test State

> Indicates the state in which the path to shut down the node is tested when a node error occurs. If the test of the path has not been completed, "Unknown" will be displayed. If the configured shutdown agent operates normally, "Unknown" will be changed to "TestWorked".

Init State

> Indicates the state in which the shutdown agent is initialized.

> 📙 **Note**
>
> Confirm that the shutdown facility is operating normally.
>
> In the following cases, there is a possibility that the mistake is found in the configuration setting of the agent or hardware.
>
> - An initial state was displayed as "InitFailed" though the setting of the shutdown facility was completed.
>
> - The test was displayed as "Unknown" or "TestFailed" though the setting of the shutdown facility was completed.
>
> Confirm whether the error message is output to /var/adm/messages file and the console output screen.
>
> Afterwards, correspond according to the content of the output message.

5.1.2.4 For a SPARC Enterprise T5120, T5220, T5140, T5240, T5440, or SPARC T3 series

5.1.2.4.1 Checking Console Configuration

In SPARC Enterprise T5120, T5220, T5140, T5240, T5440, or SPARC T3 series, ILOM is used.

Check the following settings concerning ILOM before setting the shutdown facility.

- The log in user account is made for the shutdown facility, and CLI mode of that is set to the default mode (*1).

- User inquiries of the first SSH connection (such as generation of the RSA key) are completed by connecting to ILOM from all the cluster nodes via SSH using the log in user account for the shutdown facility.

If you are using ILOM 3.0, please check the following settings as well.

- The log in user account for the shutdown facility must be set to one of the following privileges:

    - If the keyswitch_state parameter is set to normal

        - Console, Reset and Host Control, Read Only (cro)

        - Operator(*2)

    - If the keyswitch_state parameter is set to locked

        - Admin (a)

        - Administrator(*2)

    If a necessary privilege is not set, TestFailed or KillFailed of shutdown agent would be occurred.

- The log in user account for the shutdown facility must not be using SSH host-based key authentication.

Moreover, record the following information on ILOM.

- ILOM IP address(*3)

- Log in user account and password for shutdown facility in ILOM

*1) You can check if CLI mode of the log in user account is set to the default mode by the following procedure.

    1. Log in CLI of ILOM.

    2. Check prompt status.
       Prompt status that is set to the default mode.
       ->
       Prompt status that is set to alom mode.
       sc>

*2) Due to compatibility of ILOM 3.0 with ILOM 2.x, this operation is also available for users with administrator or operator privileges from ILOM 2.x.

*3) When the network routing is set, IP address of ILOM need not be the same to management LAN segment of the cluster node.

5.1.2.4.2 Setting by CLI
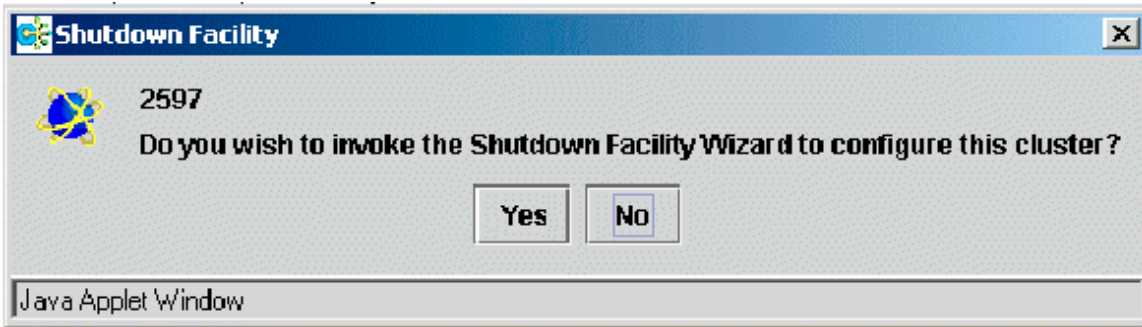
Configuring the shutdown facility

Create the "/etc/opt/SMAW/SMAWsf/rcsd.cfg" by the following contents in all nodes.

```
CFNameX,weight=weight,admIP=myadmIP:agent=SA_ilomp,timeout=timeout:agent=SA_ilomr,timeout=timeout
CFNameX,weight=weight,admIP=myadmIP:agent=SA_ilomp,timeout=timeout:agent=SA_ilomr,timeout=timeout
```

CFNameX

For the CF node name of the cluster host.

weight

For the Weight of node.

myadmIP

For the IP address of the Administrative LAN of the own node.

agent

For the name of shutdown agent.

You appoint the ILOM shutdown agent in the order of "SA_ilomp" and "SA_ilomr", in SPARC Enterprise T5120, T5220, T5140, T5240, T5440, or SPARC T3 series.

timeout

For the timeout duration of the shutdown agent.

For the 70 seconds in SPARC Enterprise T5120, T5220, T5140, T5240, T5440, or SPARC T3 series.

Example)

```
node1,weight=1,admIP=10.20.30.100:agent=SA_ilomp,timeout=70:agent=SA_ilomr,timeout=70
node2,weight=1,admIP=10.20.30.200:agent=SA_ilomp,timeout=70:agent=SA_ilomr,timeout=70
```

Survival priority

Even if a cluster partition occurs due to a failure in the cluster interconnect, all the nodes will still be able to access the user resources. For details on the cluster partition, see "2.2.2.1 Protecting data integrity" in the "*PRIMECLUSTER Concept Guide*".
To guarantee the consistency of the data constituting user resources, you have to determine the node groups to survive and those that are to be forcibly stopped.
The weight assigned to each node group is referred to as a "Survival priority" under PRIMECLUSTER.
The greater the weight of the node, the higher the survival priority. Conversely, the less the weight of the node, the lower the survival priority. If multiple node groups have the same survival priority, the node group that includes a node with the name that is first in alphabetical order will survive.

Survival priority can be found in the following calculation:

Survival priority = SF node weight + ShutdownPriority of userApplication

SF node weight (Weight):

Weight of node. Default value = 1. Set this value while configuring the shutdown facility.

userApplication ShutdownPriority:

Set this attribute when userApplication is created. For details on how to change the settings, see "8.1.2 Changing the Operation Attributes of a Cluster Application".

🔖 See
· ·

For details on the ShutdownPriority attribute of userApplication, see "6.6.5 Attributes".

· ·

Survival scenarios

The typical scenarios that are implemented are shown below:

[Largest node group survival]

- Set the weight of all nodes to 1 (default).

- Set the attribute of ShutdownPriority of all user applications to 0 (default).

	Node group1			Node group2
	node1	node2	node3	node4
weight of node	1	1	1	1
ShutdownPriority of app1 = 0				0
ShutdownPriority of app2 = 0				0
ShutdownPriority of app3 = 0				0
Live priority	3			1

[Specific node survival]

- Set the "weight" of the node to survive to a value more than double the total weight of the other nodes.

- Set the ShutdownPriority attribute of all user applications to 0 (default).

In the following example, node1 is to survive:

	Node group1	Node group2		
	node1	node2	node3	node4
weight of node	10	1	1	1
ShutdownPriority of app1 = 0		0		
ShutdownPriority of app2 =0			0	
ShutdownPriority of app3 =0				0
Live priority	10	3		

[Specific application survival]

- Set the "weight" of all nodes to 1 (default).

- Set the ShutdownPriority attribute of the user application whose operation is to continue to a value more than double the total of the ShutdownPriority attributes of the other user applications and the weights of all nodes.

In the following example, the node for which app1 is operating is to survive:

	Node group1	Node group2		
	node1	node2	node3	node4
weight of node	1	1	1	1
ShutdownPriority of app1 =20	20			
ShutdownPriority of app2 =1			1	
ShutdownPriority of app3 =1				1
Live priority	21	5		

Configuring a ILOM shutdown agent

Enter the settings for ILOM that you recorded in "5.1.2.4.1 Checking Console Configuration".

Execute the clrccusetup(1M) command in all nodes, and register the console information of local node.

Example)

```
# /etc/opt/FJSVcluster/bin/clrccusetup -a ilom 10.20.30.51 admin <RETURN>
Enter User's Password:
Re-enter User's Password:
# /etc/opt/FJSVcluster/bin/clrccusetup -l
Device-name cluster-host-name   IP-address     host-name    user-name
---------------------------------------------------------------------------
```

```
ilom          node1                    10.20.30.50     -          admin
ilom          node1                    10.20.30.51     -          admin
```

Starting up the console asynchronous monitoring

Execute the command in each node, and register the console information of local node.

```
# /etc/opt/FJSVcluster/bin/clrccumonctl
```

If "The devrccud daemon exixts." is displayed, the console asynchronous monitoring daemon is started.

If "The devrccud daemon does not exixts." is displayed, the console asynchronous monitoring daemon is not started. Execute the following command, and start the console asynchronous monitoring daemon.

```
# /etc/opt/FJSVcluster/bin/clrccumonctl start
```

Starting up the shutdown facility

Execute the following command in each node, and confirm the shutdown facility has started.

```
# /opt/SMAW/bin/sdtool -s
```

If the state of configuration of shutdown facility is displayed, shutdown facility is started.

If "The RCSD is not running" is displayed, shutdown facility is not started.

If shutdown facility is started, execute the following command, and restart the shutdown facility.

```
# /opt/SMAW/bin/sdtool -r
```

If shutdown facility is not started, execute the following command, and start the shutdown facility.

```
# /opt/SMAW/bin/sdtool -b
```

Displaying the configuration of the shutdown facility

Execute the following command in each node, and confirm the configuration of the shutdown facility.

```
# /opt/SMAW/bin/sdtool -s
Cluster Host   Agent         SA State    Shut State   Test State   Init State
------------------------------------------------------------------------------
node1          SA_ilomp.so   Idle        Unknown      TestWorked   InitWorked
node1          SA_ilomr.so   Idle        Unknown      TestWorked   InitWorked
node2          SA_ilomp.so   Idle        Unknown      TestWorked   InitWorked
node2          SA_ilomr.so   Idle        Unknown      TestWorked   InitWorked
```

Shut State

   "Unknown" is shown during normal system operation. If an error occurs and the shutdown facility stops the relevant node successfully, "Unknown" will change to "KillWorked".

Test State

   Indicates the state in which the path to shut down the node is tested when a node error occurs. If the test of the path has not been completed, "Unknown" will be displayed. If the configured shutdown agent operates normally, "Unknown" will be changed to "TestWorked".

Init State

   Indicates the state in which the shutdown agent is initialized.

![Note icon] **Note**

. . . .

Confirm that the shutdown facility is operating normally on the display result of the sdtool -s command.

- In the following cases, there is a possibility that the mistake is found in the configuration setting of the agent or hardware.

    - An initial state was displayed as "InitFailed" though the setting of the shutdown facility was completed.

    - The test was displayed as "Unknown" or "TestFailed" though the setting of the shutdown facility was completed.

  Confirm whether the error message is output to /var/adm/messages file and the console output screen.
  Afterwards, correspond according to the content of the output message.

- If you connected to ILOM, do not do three connections or more at the same time.
  If unavoidable, stop the shutdown facility of all nodes beforehand. Then, confirm the normal of shutdown facility started in all nodes after the cut connection. For details on stop, start, and the state confirmation of the shutdown facility, see the manual page describing sdtool(1M).

. . . .

5.1.3 Initial Setup of the Cluster Resource Management Facility

This section explains how to set up the resource database that the cluster resource management facility (CRM) manages.

Set up the CRM resource database according to the following procedure:

1. Initial setup

   Set up the resource database that CRM manages.

2. Automatic configure

   Register the connected hardware devices (shared disks, line switching units, and network interface cards) to the resource database that CRM manages.

![Note icon] **Note**

. . . .

You can check the NIC (network interface card) to be registered by executing the "ifconfig(1M)" command in the plumb up state.

. . . .

Set up the CRM resource database from the CRM main window. Use the CRM main window as follows:

Operation procedure

1. Select *PRIMECLUSTER -> Global Cluster Services -> Cluster Admin* in the *Web-Based Admin View operation* menu.

2. When the "Cluster Admin" screen is displayed, select the *crm* tab.
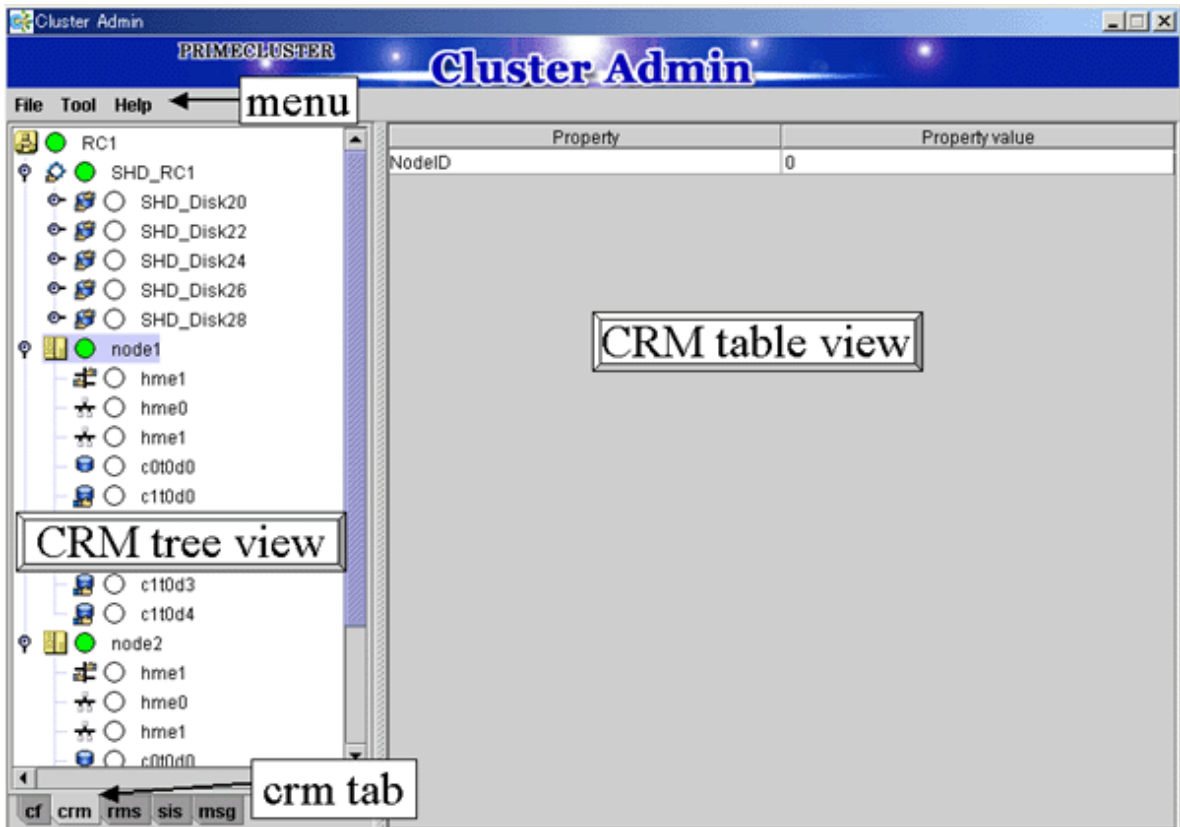
The areas shown in the screen are described below.

Menu bar

This area displays the menu. See "7.1.2.1.3 Operations".

CRM tree view

This area displays the resources registered to CRM. The resources are displayed in a tree structure.
For details on the colors and status of the icons displayed in the tree, see "7.1.2.1 Displayed Resource Types".

CRM table view

This area displays attribute information for the resource selected in the CRM tree view. For information on the displayed information, see "7.1.2.2 Detailed Resource Information".

5.1.3.1 Initial Configuration Setup

Set up the resource database that CRM manages.

When setting up the initial configuration, make sure that all nodes in the cluster have been started and that CF configuration is completed.

Operation procedure

1. Select the *Initial setup* in the *Tool* menu.

Figure 5.25 Screen for cluster resource management facility
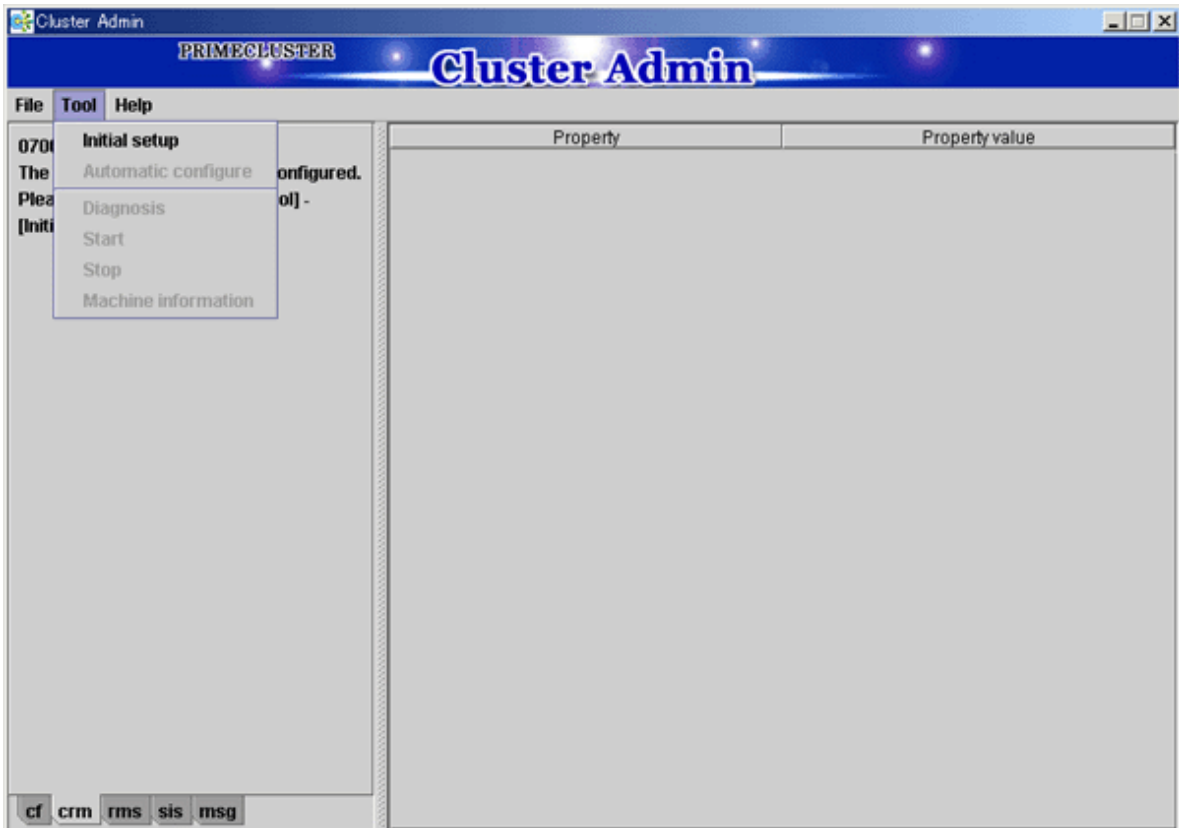
 Note

The *Initial setup* can be selected only if the resource database has not been set.

2. The screen for initial setup is displayed.

Figure 5.26 Screen for initial setup
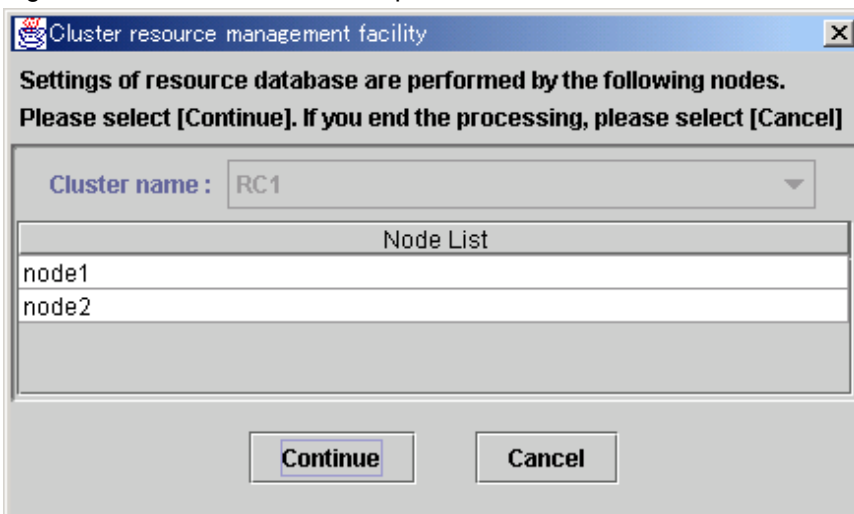
Cluster name

This area displays the names of the clusters that make up the resource database. The cluster names displayed here were defined during CF configuration.

Node List

This area displays the list of the nodes that make up the resource database.

Note

Check that the nodes that were configured in the cluster built with CF and the nodes displayed here are the same.

If the nodes do not match, check the following:

- Whether all nodes displayed by selecting the *cf* tab in the Cluster Admin screen are Up.

- Whether Web-Based Admin View is operating in all nodes.

For instructions on checking this, see "4.3.3.2 Confirming Web-Based Admin View Startup".

Continue button

Click this button to set up the resource database for the displayed cluster.

Initial setup is executed on all nodes displayed in the Node list.

Cancel button

Click this button to cancel processing and exit the screen.

3. Check the displayed contents, and click the *Continue* to start initial setup.

4. The screen below is displayed during execution of initial setup.
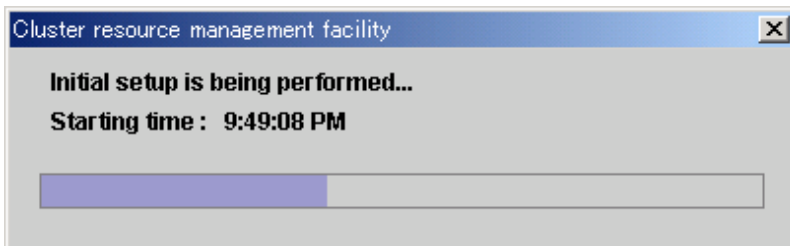
5. When initial setup ends, the following message is displayed.

If *Ok* is selected, the automatic configure screen is displayed. For details on automatic configure, see "5.1.3.2 Automatic Configure".
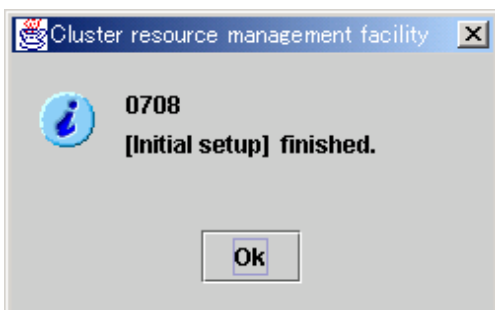
- If a message appears during operation at the CRM main window, or if a message dialog box entitled "Cluster resource management facility" appears, refer to the following:

    - "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*"

        - "12.10 Resource Database messages"

    - Appendix D in this manual

        - D.7.2 Failed Resource and Operator Intervention Messages (GUI)

        - Display format 1 in D.1 Searching for a Message

- If you want to add, delete, or rename a disk class from the *Global Disk Services* screen after executing Initial Setup from the CRM main window, close the *Cluster Admin* screen.

5.1.3.2 Automatic Configure

The automatic configure process registers the connected hardware devices (shared disks, line switching units, and network interface cards) to the resource database.

Operation Procedure

1. Display the automatic configure screen by one of the following methods:

    - Respond to the "0708 [*Initial setup*] finished." message, which is displayed after initial setup is completed.

    - Select *Tool -> Automatic configure* from the CRM tree view on the CRM main window of Cluster Admin while the cluster name displayed at the highest level is selected.

        Figure 5.27 Screen for automatic configure
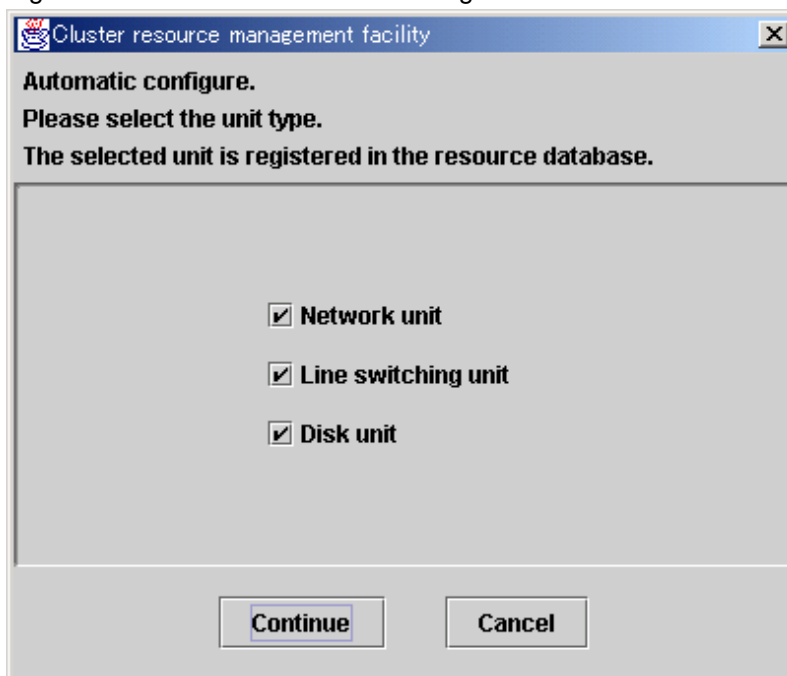
Network unit

Select this item to detect network interfaces automatically and register them to the resource database.

You must select this item to use public LAN takeover.

Line switching unit

Select this item to detect connected line switching units automatically and register them to the resource database.

You must select this item to use the line takeover function.

Disk unit

Select this item to detect connected shared disk units automatically and register them to the resource database. You must select this item to use system disk mirroring or to use Global Disk Services (GDS) and Global File Services (GFS) in shared disk unit takeover.

Note

. . . .

When you use the multipath disk control function to set shared disk units, the following messages may be displayed on the console, depending on the configuration:

```
WARNING: mphd2048: paths setup information of command is reverse against device one.

WARNING: mphd2049: paths setup information of command is reverse against device one.
```

These messages indicate that connection of multipath disk that is stored in multipath disk control and actual connection of multipath disk are not identical. But multipath disk control automatically reconfigures the multipath according to the actual connection. You can ignore these messages.

. . . .

Continue button

Click this button to detect the selected device type automatically and register it to the resource database.

Cancel button

Click this button to cancel processing and exit the screen.

2. Select the device type, and click the *Continue*.

Processing is executed according to the following flow:
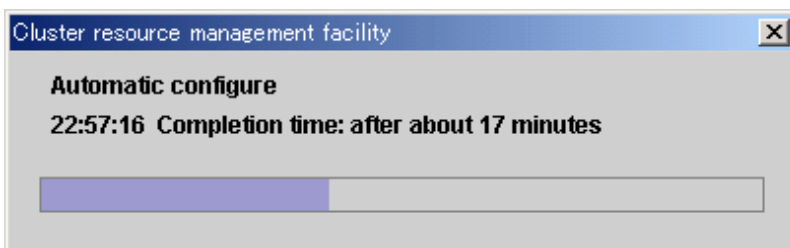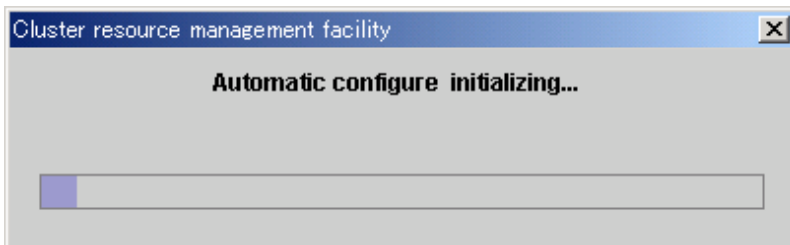
3. When automatic configure is completed, the following message is displayed:
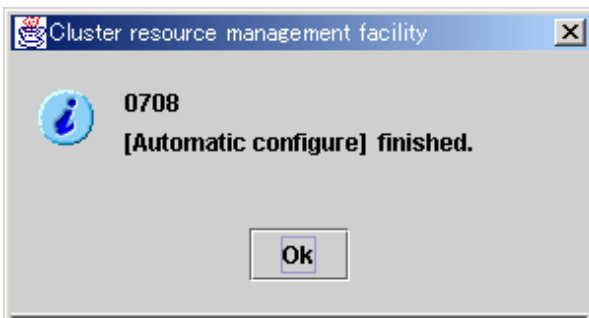
4. When the initial setup and automatic configure are completed, the following screen appears.
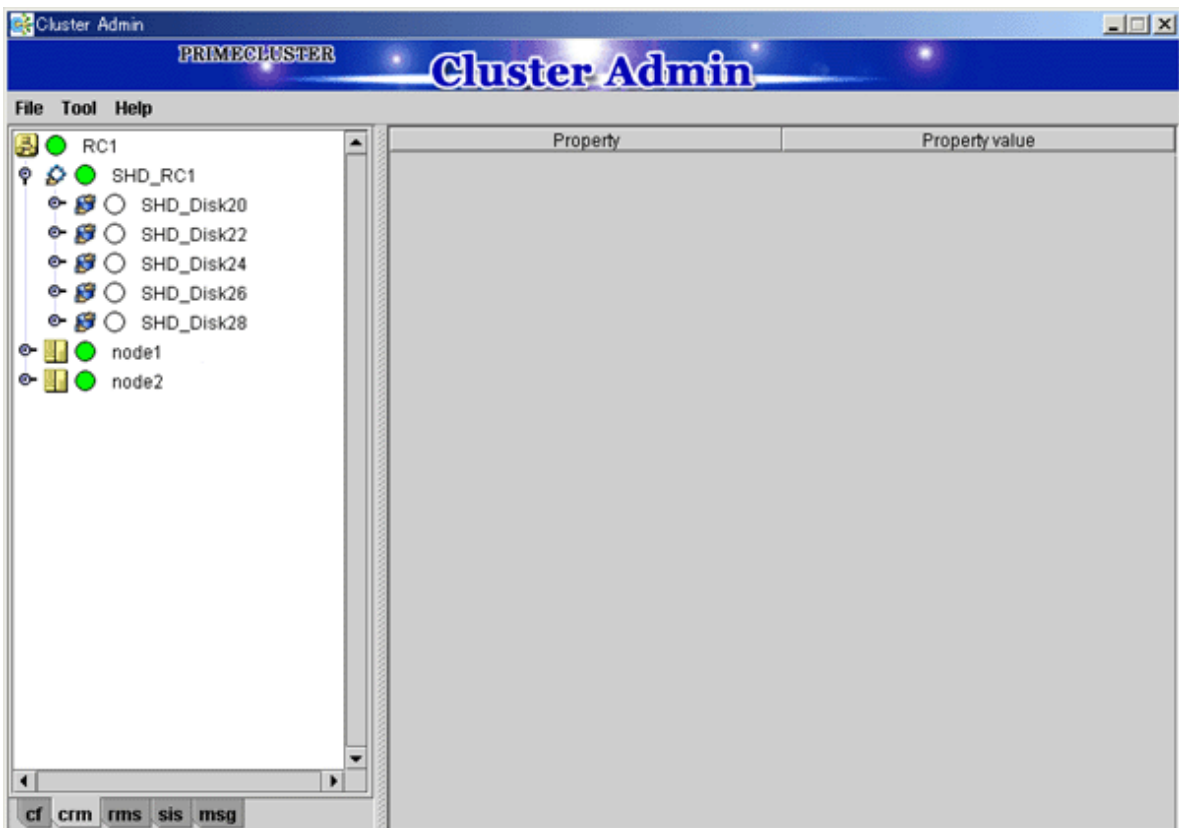
5. When automatic configuration is completed, go to the CRM main window and confirm that the resource registration is completed by checking the following.

   - Whether the number of shared disk unit resources is less than the actual device configuration.

   - Whether any disks other than shared disk unit are registered as shared disk unit.

   - Whether the number of public LAN resources is less than the actual device configuration.

   - Whether all connected line switching units are displayed.

   If the actual device configuration and the resources do not match each other as described above, automatic detection may have been disabled because of one of the following causes:

   - There is a connection path failure between a host device and a disk array unit.

   - A disk array unit is not ready.

   - A network adapter failed.

   - A network adapter driver failed.

- A line switching unit is not connected properly.

- A line switching unit is not turned on.

If the resources are not registered correctly, first review the above causes. Then in the CRM main window, select the *Automatic configure* in the *Tool* to re-register the resources.

Note

If a message is displayed in the CRM main window during operation or a message is displayed in the message dialog box entitled "Cluster resource management facility," see the following manuals for details:

- "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide"

  - "12.10 Resource Database messages"

- Appendix D in this manual

  - D.7.2 Failed Resource and Operator Intervention Messages (GUI)

  - Display format 1 in D.1 Searching for a Message

- If you want to add, delete, or rename a disk class from the *Global Disk Services* screen after executing automatic configuration from the CRM main window, close the *Cluster Admin* screen.

- The following message may be displayed on the console during automatic configuration:

```
FJSVcluster: Error: dcmevmd: 6000: An internal error occurred.
(function:_evm_res_all_callback detail:0x9060b00-0x4-0-0)
```

If this message is displayed, all the hardware devices registered to the resource database may not be displayed on the CRM tree view on the CRM main window. Restart the Cluster Admin screen and open the CRM main window again.

5.2 Setting Up Power Supply Linkage

If you want to set up power supply linkage, see the "*Machine Administration Guide*".

5.3 Setting Up Shared Disk Connection Confirmation

Shared disk connection confirmation is a function that performs the followings:

- Check that the shared disk is powered.

- Check that the cable is properly connected.

- Detect new shared disks.

Note

If the paths of a disk array unit are multiplexed, and the paths are controlled by a multipath disk control function, shared disk connection confirmation will not detect an error even if an error occurs in part of the multiplexed paths.

If an error is detected in a shared disk unit or if a new shared disk unit is detected, shared disk connection confirmation outputs a message to syslogd(1M) and the CRM main window.

This function allows you to specify whether or not RMS activation is to be suppressed in the node in which the failure was detected. If a failure is detected in a shared disk unit, the userApplication might not operate normally and user resources in that shared disk unit might become damaged. It is therefore recommended that you specify that RMS activation should be suppressed if a failure is detected. After PRIMECLUSTER is installed, shared disk connection confirmation is initially disabled by default. To change this, follow the procedures described below.

> 📒 **Note**
> ....
>
> - Set up shared disk connection confirmation after the initial setup of the Cluster Resource Management facility is completed.
>
> - To view the manual pages of each command, add "/etc/opt/FJSVcluster/man" to the MANPATH variable.
> ....

Enabling shared disk connection confirmation

Take the following steps on each node. Shared disk connection confirmation will be executed from the next node startup.

RMS activation is to be suppressed if a failure is detected

1. Add the following to the "RELIANT_INITSCRIPT" environment variable of RMS:

   ```
   /etc/opt/FJSVcluster/bin/clinitscript
   ```

   Example: Add the following to /opt/SMAW/SMAWRrms/bin/hvenv.local.

   ```
   export RELIANT_INITSCRIPT=/etc/opt/FJSVcluster/bin/clinitscript
   ```

2. Execute the "clsetacfparam(1M)" command.

   Example:

   # **/etc/opt/FJSVcluster/bin/clsetacfparam -p auto**

> 📒 **Note**
> ....
>
> If there is an error in the shared disk unit, it is detected by the registered script (/etc/opt/FJSVcluster/bin/clinitscript) and RMS stops. In this case, the following RMS message will be sent to syslogd(1M):
>
> "(INI, 12): FATAL ERROR: InitScript returned non-zero exit code 1."
> ....

RMS activation is not to be suppressed if a failure is detected

1. Execute the "clsetacfparam(1M)" command.

   Example:

   # **/etc/opt/FJSVcluster/bin/clsetacfparam -p auto**

Disabling shared disk connection confirmation

To cancel the setting previously made for "Enabling shared disk connection confirmation," perform the following procedure on all the nodes. After you next boot the node, shared disk connection will not be confirmed.

"Suppress RMS activation if an error is detected" was previously configured

1. Delete the following from the RELIANT_INITSCRIPT environment variable of RMS.

   ```
   /etc/opt/FJSVcluster/bin/clinitscript
   ```

   Example: Delete the following from /opt/SMAW/SMAWRrms/bin/hvenv.local.

   ```
   export RELIANT_INITSCRIPT=/etc/opt/FJSVcluster/bin/clinitscript
   ```

2. Execute the "clsetacfparam(1M)" command.

   Example:

   # **/etc/opt/FJSVcluster/bin/clsetacfparam -p none**

"Do not suppress RMS activation if an error is detected" was specified

1. Execute the "clsetacfparam(1M)" command.

    Example:

    ```
    # /etc/opt/FJSVcluster/bin/clsetacfparam -p none
    ```

See

. . . .

- For details on the RMS environment variables, see "10 Appendix - Environment Variables" in the "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide*".

- For details on the "clsetacfparam(1M)" command, see the manual page for clsetacfparam(1M).

. . . .

5.4 Setting Up Fault Resource Identification and Operator Intervention Request

The fault resource identification is a function that outputs a message to syslogd(1M) and Cluster Admin and a history of failed resources to Resource Fault History if a failure occurs in a resource or node that is registered to a cluster application.

After setting the initial configuration of the resource database, specify the settings for enabling fault resource identification and operator intervention request. An example of a message displayed by fault resource identification is shown below.

```
6750 A resource failure occurred.  SysNode:node1RMS userApplication:app0 Resource:apl1
```

The operator intervention request function displays a query-format message to the operator if a failed resource or a node in which RMS has not been started is found when a cluster application is started. The messages for operator intervention requests are displayed to syslogd(1M) and Cluster Admin.

```
1421 userApplication "app0" was not started automatically because all SysNodes that
make up userApplication were not started within the prescribed time.
Forcibly start userApplication in SysNode "node1RMS"? (no/yes)

Message number: 1001
Warning: When userApplication is forcibly started, the safety check becomes disabled.
If the operation is used incorrectly, data may be damaged and the consistency may be lost.
Check that userApplication to be forcibly started is not online in the cluster before executing the
forced startup.
```

See

. . . .

For details on the messages displayed by the fault resource identification function and the messages displayed by the operator intervention request function, see "D.7.2 Failed Resource and Operator Intervention Messages (GUI)" and "D.5 Operator Intervention Messages".

. . . .

This section describes procedures for operating fault resource identification and operator intervention request.

Note

. . . .

- After PRIMECLUSTER is installed, fault resource identification and operator intervention request are initially disabled. The following Cluster Admin functions are also disabled:

    - Messages for fault resource identification and operator intervention request are not displayed to Cluster Admin.

    - The list of resources that are currently affected by faults is not displayed in the Resource Fault History screen of Cluster Admin.

    - The fault history of the resources is not displayed in the Resource Fault History screen of Cluster Admin.

- To view the manual pages of each command, add "/etc/opt/FJSVcluster/man" to the MANPATH variable.

. . . .

Enabling the operation of fault resource identification and operator intervention request

1. Execute the "clsetparam(1M)" command and specify the settings for enabling the fault resource identification and operator intervention request. Execute this procedure in any node that is part of the cluster system.

   ```
   # /etc/opt/FJSVcluster/bin/clsetparam -p AppWatch ON
   ```

2. Execute the "clsetparam(1M)" command, and check that the parameters are set so that the operation of fault resource identification and operator intervention request is enabled. Execute this procedure on any node that is part of the cluster system.

   ```
   # /etc/opt/FJSVcluster/bin/clsetparam -p AppWatch
   ```

   ```
   ON
   ```

3. Restart all the operating nodes.

   If a node is stopped, the fault resource identification and operator intervention request begin operating in that node from the next node reboot.

Disabling the operation of fault resource identification and operator intervention request

To cancel the setting previously made for "Enabling the operation of fault resource identification and operator intervention request," perform the following procedure:

1. Execute the "clsetparam(1M)" command and specify the settings for disenabling the fault resource identification and operator intervention request. Execute this procedure in any node that is part of the cluster system.

   ```
   # /etc/opt/FJSVcluster/bin/clsetparam -p AppWatch OFF
   ```

2. Execute the "clsetparam(1M)" command, and check that the parameters are set so that the operation of fault resource identification and operator intervention request is disenabled. Execute this procedure on any node that is part of the cluster system.

   ```
   # /etc/opt/FJSVcluster/bin/clsetparam -p AppWatch
   ```

   ```
   OFF
   ```

3. Restart all the operating nodes.

   If a node is stopped, the fault resource identification and operator intervention request stop operating in that node from the next node reboot.

Chapter 6 Building Cluster Applications

The procedure for building a cluster application is shown below.
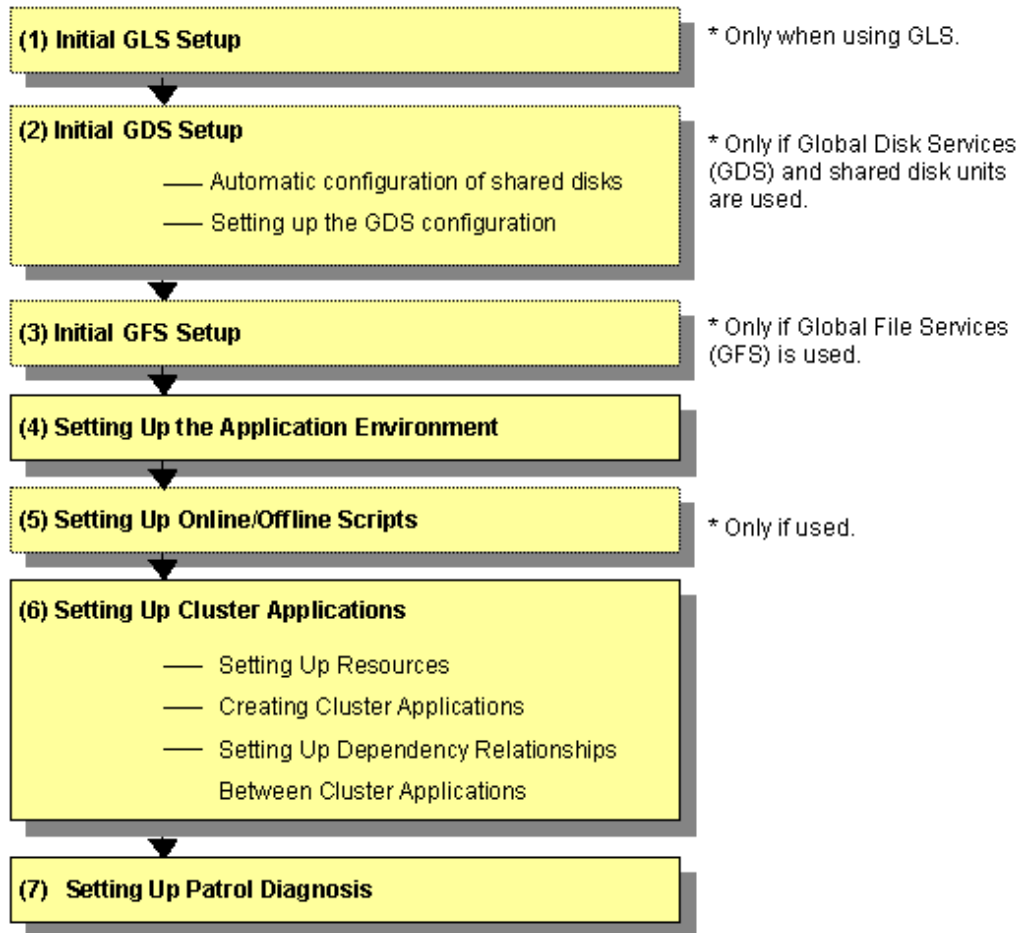
Table 6.1 Application building procedure and manual reference locations

	Work item	Execution Nodes	Required/ optional	Manual reference location*
(1)	6.1 Initial GLS Setup	All nodes	Optional (required when GLS is used)	GLSR GLSM
(2)	6.2 Initial GDS Setup	All nodes	Optional (required when GDS is used)	
	6.2.1 Automatic Configuration of Shared Disks	All nodes	Optional	**CF** "4.4.3 Automatic resource registration"
	6.2.2 GDS Configuration Setup	All nodes	Optional	**GDSG** "Chapter 5 Operations"
(3)	6.3 Initial GFS Setup	All nodes	Optional (required when GFS is used)	GFSG

	Work item	Execution Nodes	Required/ optional	Manual reference location*
(4)	6.4 Setting Up the Application Environment	All nodes	Required	Manuals for each application
(5)	6.5 Setting Up Online/Offline Scripts	All nodes	Optional	RMS "2.9 Environment variables", "10 Appendix - Environment variables"
(6)	6.6 Setting Up Cluster Applications 6.6.1 Setting Up Resources 6.6.2 Creating Cluster Applications 6.6.3 Setting Up Dependency Relationships Between Cluster Applications	All nodes	Required	"Solaris X Reference Manual Collection," "Solaris Answer Book"
(7)	6.7 Setting Up Patrol Diagnosis	All nodes	Required	

* The names of the reference PRIMECLUSTER manuals are abbreviated as follows:

- **RMS:** *PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide*

- **CF:** *PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*

- **GDSG:** *PRIMECLUSTER Global Disk Services Configuration and Administration Guide*

- **GFSG:** *PRIMECLUSTER Global File Services Configuration and Administration Guide*

- **GLSR:** *PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function*

- **GLSM:** *PRIMECLUSTER Global Link Services Configuration and Administration Guide: Multipath Function*

6.1 Initial GLS Setup

This section outlines the steps for configuring Global Link Services (GLS).

6.1.1 GLS Setup

For information on the initial GLS setup, see "Chapter 5 Operation on Cluster System" in the "*PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function*" and "Chapter 7 Administration on a Cluster System" in the "*PRIMECLUSTER Global Link Services Configuration and Administration Guide: Multipath Function.*"

This section describes how to set up "the Single system without NIC sharing of the NIC switching mode (IPv4)" that GLS (redundant line control function) provides. This procedure is described in the example below.

🚩 Note

You can use PRIMECLUSTER "takeover network" and GLS "IP address takeover" together in a same cluster system; however, you cannot configure them on the same interface. If this occurs, communication using takeover IP addresses is disenabled.

For example, if you select hme1 as the interface when you set up the PRIMECLUSTER "takeover network," do not set hme1 in the GLS environment settings (do not specify hme1 using "-t" option of the "hanetconfig create" command).

The setup values correspond to those on the "A.7.2 GLS Setup Worksheet."
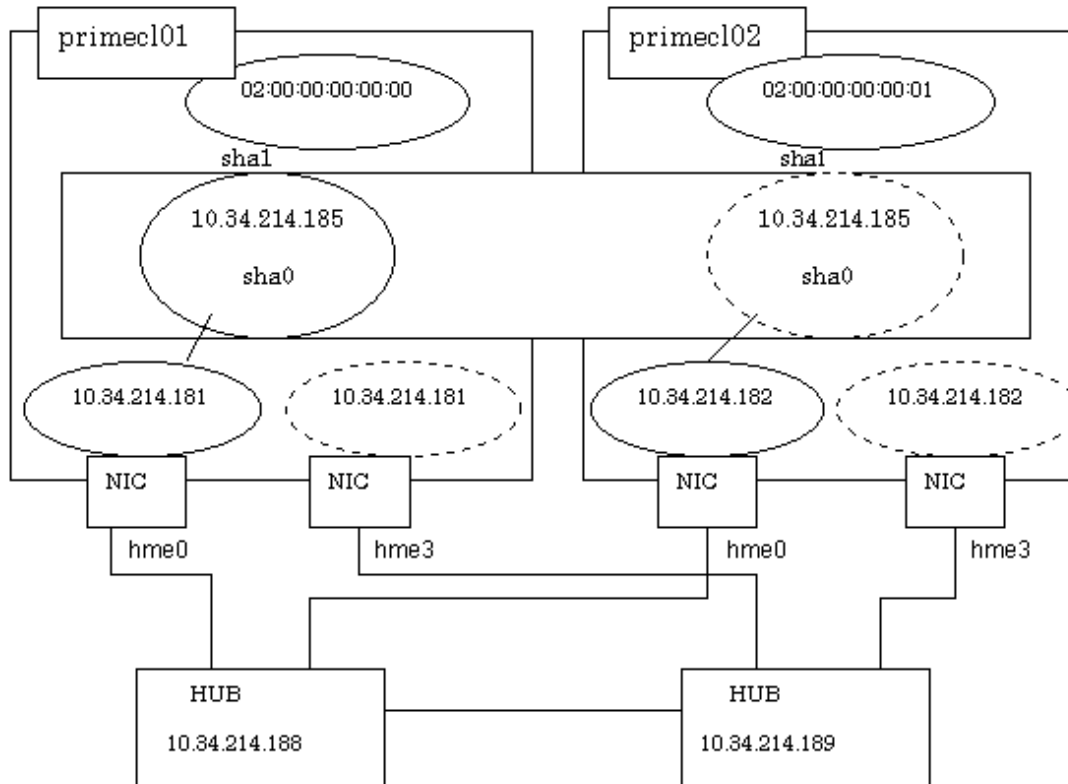
Operation Procedure:

If the OPERATING node is [HOST-primecl01]

1. Setting up the system

    1. Define the IP address and Host name in /etc/inet/hosts file.

    ```
    10.34.214.185   takeoverIP   # Virtual IP
    10.34.214.181   primecl01    # primecl01 Physical IP
    10.34.214.182   primecl02    # primecl02 Physical IP
    10.34.214.188   swhub1       # primary HUB IP
    10.34.214.189   swhub2       # secondary HUB IP
    ```

    2. Configure /etc/hostname.hme0 file as follows.

    Contents of /etc/hostname.hme0

    ```
    primecl01
    ```

2. Rebooting

    Run the following command and reboot the system. After rebooting the system, verify hme0 is enabled using ifconfig command.

    ```
    # /usr/sbin/shutdown -y -i6 -g0
    ```

3. Creating of virtual interface

    For the underlined parameter, specify the physical IP address of the node.

    ```
    # /opt/FJSVhanet/usr/sbin/hanetconfig create -n sha0 -m d -i 10.34.214.185 -e 10.34.214.181 -t
    hme0,hme3
    ```

    Check that the virtual interface has been set up correctly.

    ```
    # /opt/FJSVhanet/usr/sbin/hanetconfig print
    ```

4. Setting up the HUB monitoring function

   For the underlined parameter, specify the IP addresses of the hubs to be monitored.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetpoll create -n sha0 -p 10.34.214.188,10.34.214.189 -b off
   ```

   Check that the facility has been set up correctly.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetpoll print
   ```

5. Setting up the Standby patrol monitoring function

   For the underlined parameter, specify the MAC address using the format 02:xx:xx:xx:xx:xx. Specify the address so that the values of the two nodes are not the same.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetconfig create -n sha1 -m p -a 02:00:00:00:00:00 -t sha0
   ```

   Check that the facility has been set up correctly.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetconfig print
   ```

6. Creating of the takeover IP address (takeover virtual Interface)

   ```
   # /opt/FJSVhanet/usr/sbin/hanethvrsc create -n sha0
   ```

   Check that the registration has been done correctly.

   ```
   # /opt/FJSVhanet/usr/sbin/hanethvrsc print
   ```

If the STANDBY node is [HOST-primecl02]

1. Setting up the system

   1. Define the IP address and Host name in /etc/inet/hosts file. Defined content is same as HOST-primecl01.

   2. Configure /etc/hostname.hme0 file as follows.

      Contents of /etc/hostname.hme0

      ```
      primecl02
      ```

2. Rebooting

   Run the following command and reboot the system. After rebooting the system, verify hme0 is enabled using ifconfig command.

   ```
   # /usr/sbin/shutdown -y -i6 -g0
   ```

3. Creating of virtual interface

   For the underlined parameter, specify the physical IP address of the node.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetconfig create -n sha0 -m d -i 10.34.214.185 -e 10.34.214.182 -t
   hme0,hme3
   ```

   Check that the virtual interface has been set up correctly.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetconfig print
   ```

4. Setting up the HUB monitoring function

   For the underlined parameter, specify the IP addresses of the hubs to be monitored.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetpoll create -n sha0 -p 10.34.214.188,10.34.214.189 -b off
   ```

   Check that the facility has been set up correctly.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetpoll print
   ```

5. Setting up the Standby patrol monitoring function

```
# /opt/FJSVhanet/usr/sbin/hanetconfig create -n sha1 -m p -a 02:00:00:00:00:01 -t sha0
```

Check that the facility has been set up correctly.

```
# /opt/FJSVhanet/usr/sbin/hanetconfig print
```

6. Creating of the takeover IP address (takeover virtual Interface)

```
# /opt/FJSVhanet/usr/sbin/hanethvrsc create -n sha0
```

Check that the resources have been set up correctly.

```
# /opt/FJSVhanet/usr/sbin/hanethvrsc print
```

Post-setup processing

After the OPERATING and STANDBY node setup is done, create the Gls resources, and register them to the cluster application.

For details, see "6.6.1.4 Creating Gls Resources" and "6.6.2 Creating Cluster Applications."

See

For information on GLS (redundant line control function) and other operation modes, see "Appendix B Examples of Setting Up" in the "*PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function.*"

6.1.2 Setting Up Web-Based Admin View When GLS is Used

To use Web-Based Admin View in a network that was made redundant with GLS, you must set up Web-Based Admin View after setting up the NIC switching mode or the fast switching mode.

See

For setup details, see "2.3 Setup with GLS" in the "*PRIMECLUSTER Web-Based Admin View Operation Guide.*"

6.2 Initial GDS Setup

Take the following steps to configure Global Disk Services (GDS). If shared disk units are to be used, you are recommended to use GDS to manage the shared disk units.

Note

If you plan to add, delete, or rename a disk cluster from the Global Disk Services screen, close the Cluster Admin screen before starting the operation.

6.2.1 Automatic Configuration of Shared Disks

Before setting up GDS, you must have configured the shared disks using automatic configuration. If the disk units were already configured with automatic configuration in "5.1.3.2 Automatic Configure," skip this operation.

Note

If disk units are not set up for automatic configuration during the initial setup or if a shared disk unit was added after the initial configuration setup, execute automatic configuration of disk units before the GDS setup.

6.2.1.1 Executing Automatic Configuration

After setting up the disk units according to the procedure described in "3.2.2 Setting Up Disk Units," either execute *Automatic configure* from the Tools menu of the CRM main window in Cluster Admin, or execute the following command in any one of the nodes in the cluster system:

```
# /etc/opt/FJSVcluster/bin/clautoconfig -r
```

Note
. . . .

If disk units were specified during automatic resource registration in "5.1.3.1 Initial Configuration Setup," the resources have already been created. If you want to check whether the disk units were registered as resources, you can reexecute automatic resource registration.
. . . .

See
. . . .

For details, see the manual page for the "clautoconfig(1M)" command or "4.4.3 Automatic resource registration" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*."
. . . .

6.2.2 GDS Configuration Setup

The GDS setup operations are described below.

These setup operations can be performed in any sequence.

- 6.2.2.1 Setting Up System Disk Mirroring

   Set up system disk mirroring in both nodes, and then restart both nodes.

- 6.2.2.2 Setting Up Shared Disks

   Set up the shared disk volumes.

Note
. . . .

When one or more single nodes for which the class has already been created are converted to the cluster system, the setting of GDS is needed. For details, see "A.2.26 Changing Over from Single Nodes to a Cluster System" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*".
. . . .

6.2.2.1 Setting Up System Disk Mirroring

Take the following setup procedures to enable system disk mirroring.

The setup values correspond to the values that were specified in the "A.8.1 System Disk Mirror Setup Worksheet." In the operation procedure, the "System Disk Mirror Setup Worksheet" is abbreviated as "worksheet."

See
. . . .

For setup details, see "5.2.1 System Disk Settings" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*."
. . . .

Operation Procedure:

1. Select *Global Disk Services* at the Web-Based Admin View top screen.

   The GDS Management screen (hereinafter main screen) is displayed.
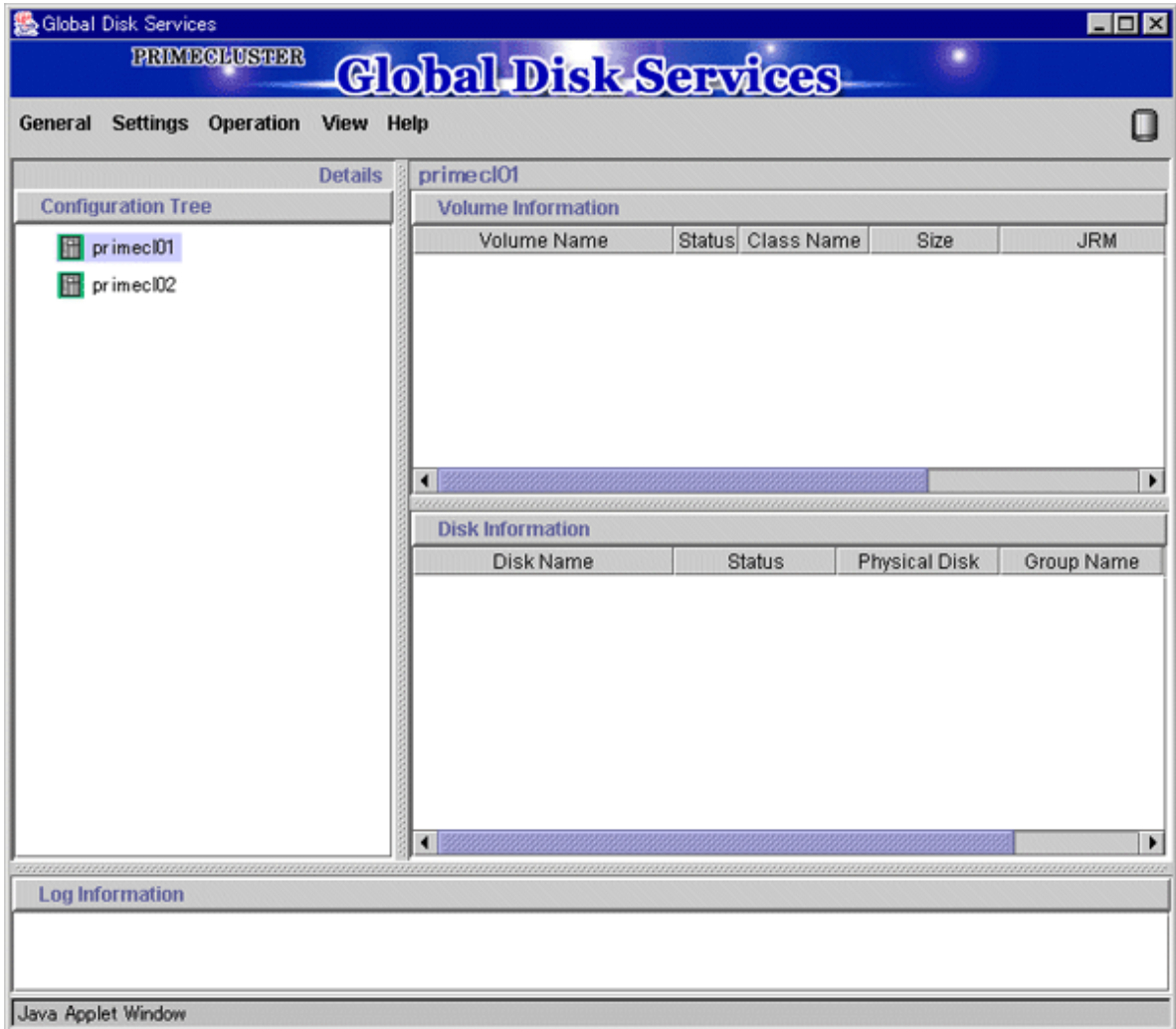
2. From the GDS configuration tree, select the node in which the system disk mirror is to be set, click the *Settings* menu, and select
System Disk Settings.
A list of disks that can be used for mirrored disks for the selected node is displayed.

Select the system disk ("Physical disk name" on the worksheet), and click *Next*.
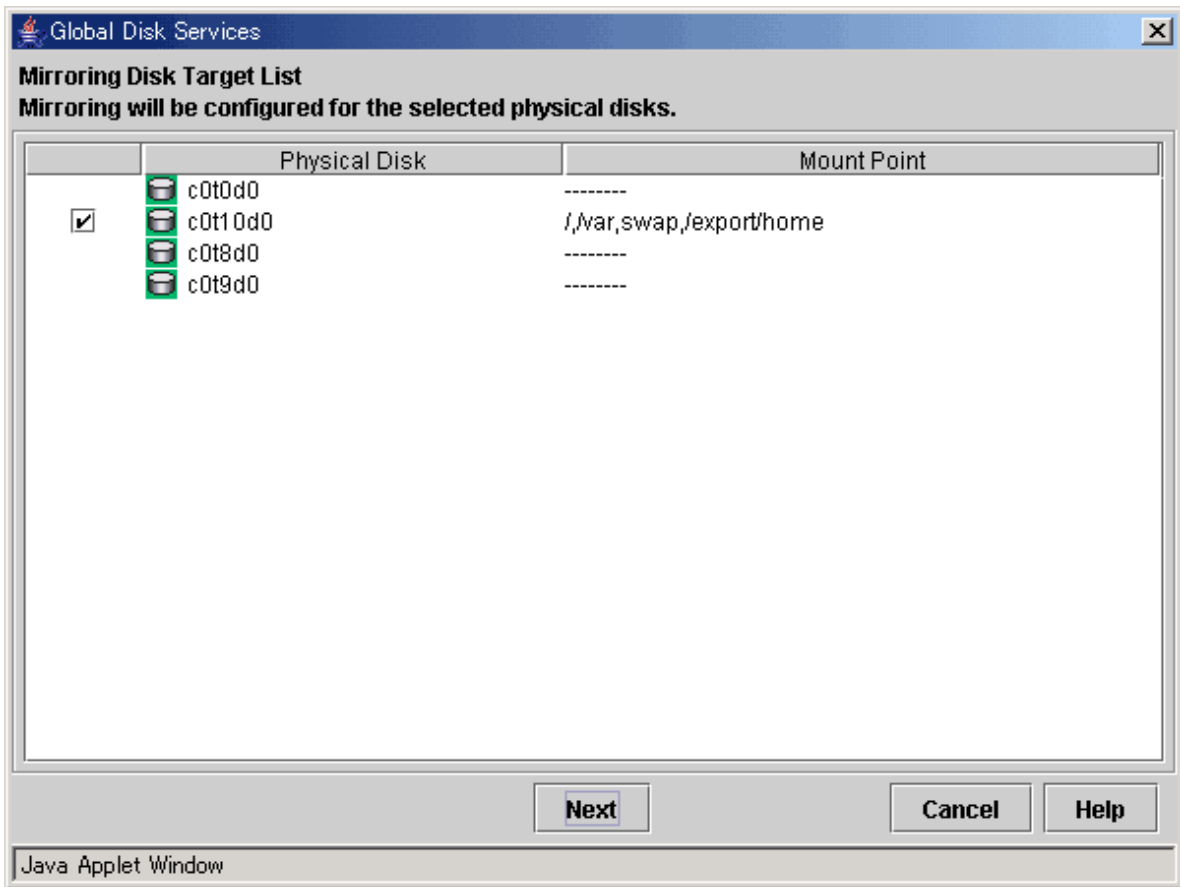
3. Specify class name of the root class.

Enter the class name ("Class name" on the worksheet) of the root class, and click *Next*.
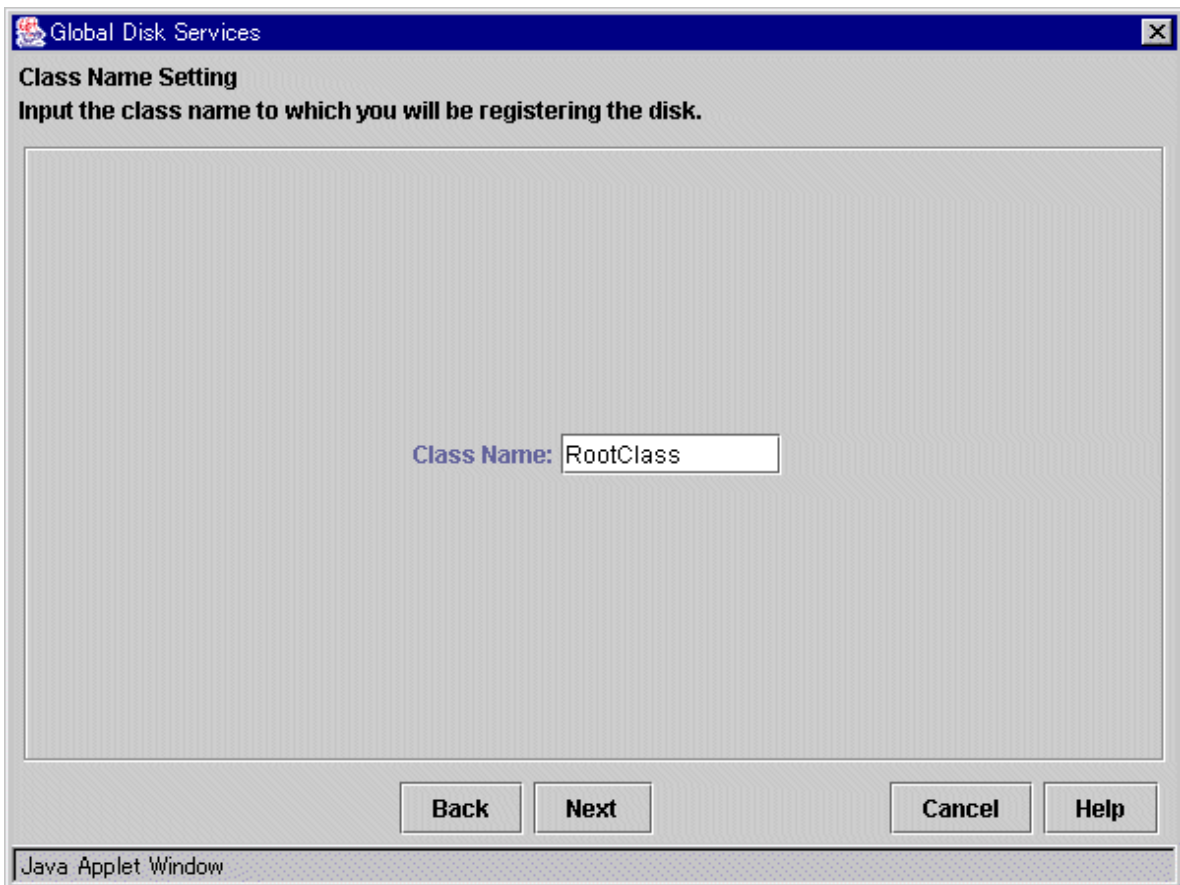
Note

Specify the class name so that the class names of the root class are not duplicated among cluster nodes.

4. Add a mirror disk to the group.

Enter the "Group Name" ("Group name" on the worksheet), then from the "Physical Disk List," select the "mirror disk" ("Mirror disk name" on the worksheet) for the system disk, and click *Add*.
Check that the "mirror disk" that was selected is displayed in "Group Configuration Disk," and then click *Next*.
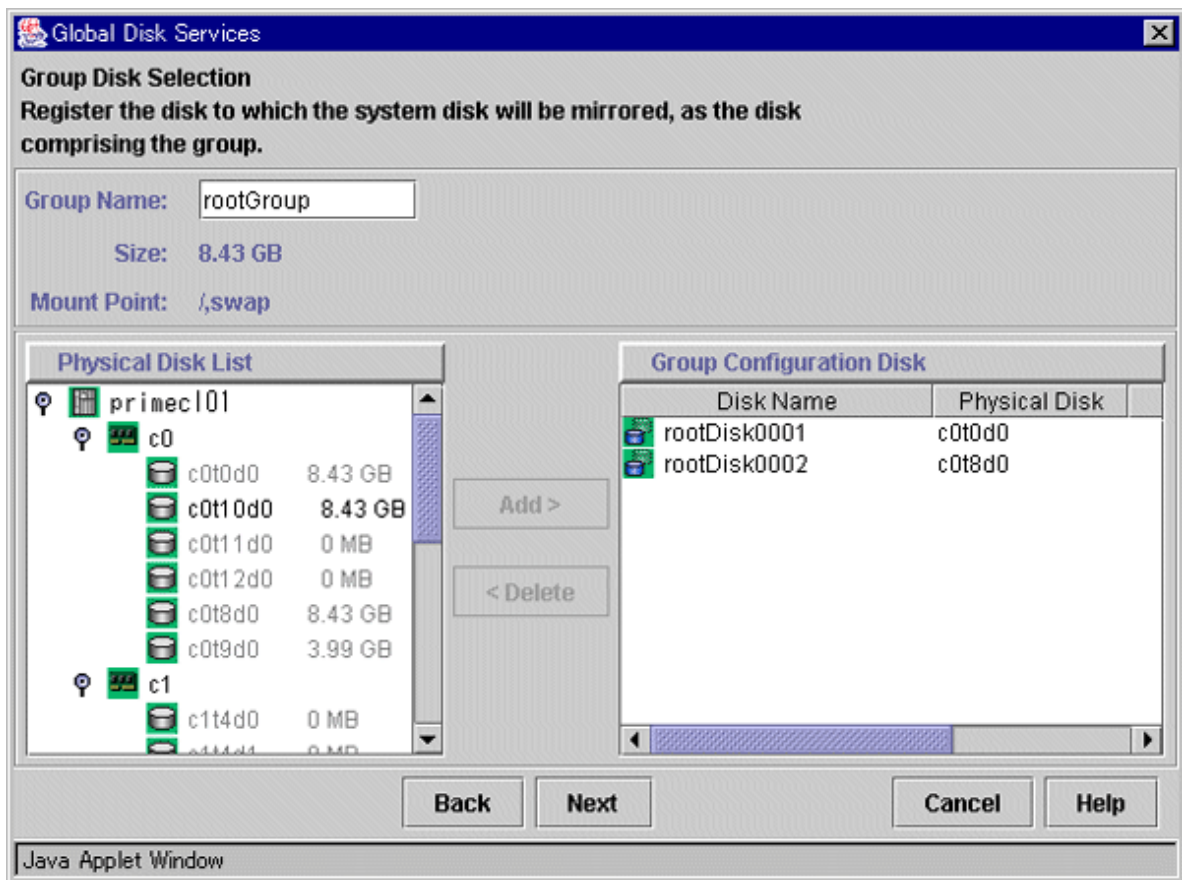
5. Select a spare disk ("Spare disk name" on the worksheet) from the "Physical Disk List," and click *Add*.
   Check that the spare disk that was selected is registered to "Spare Disk," and then click *Next*.
   If a spare disk is unnecessary, go to Step 6.
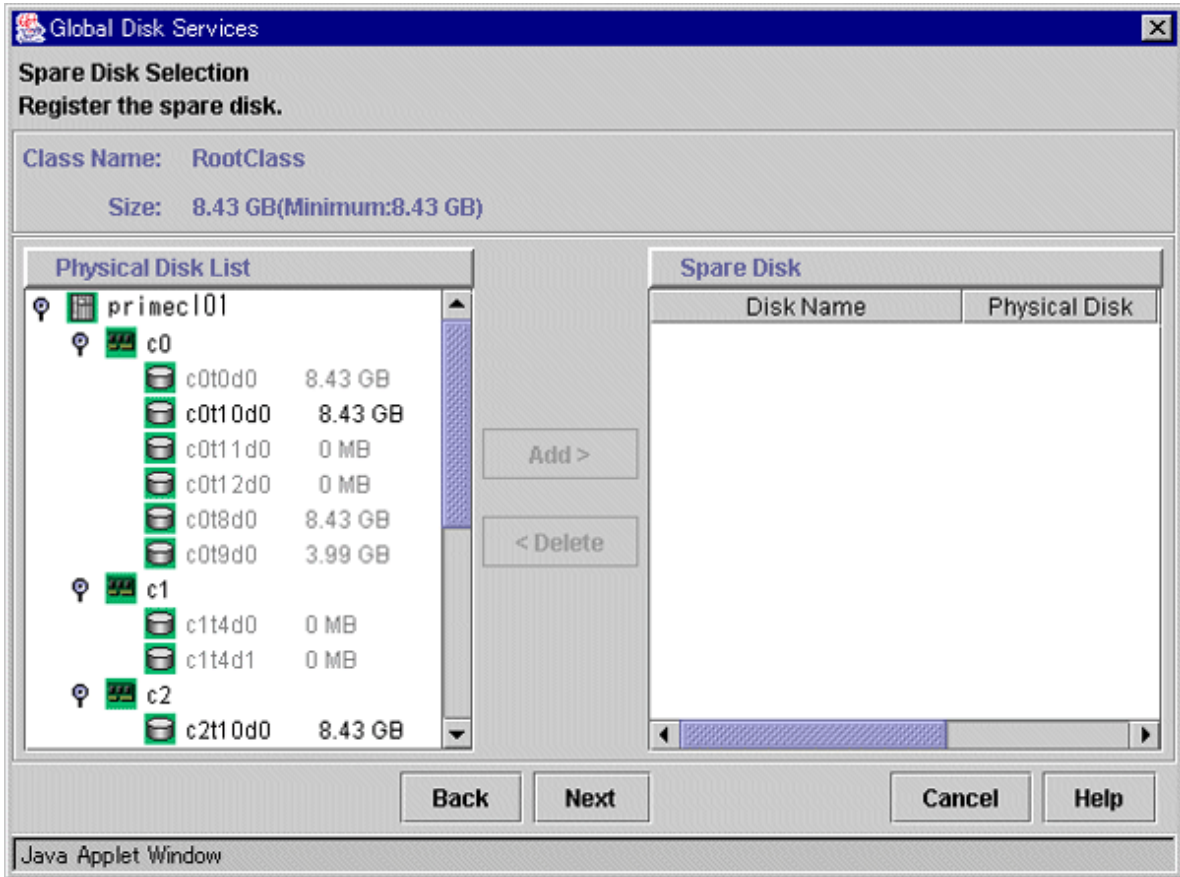
6. Check the system disk configuration.

   Check the physical disk name and the mirror disk name, and then click *Create*.
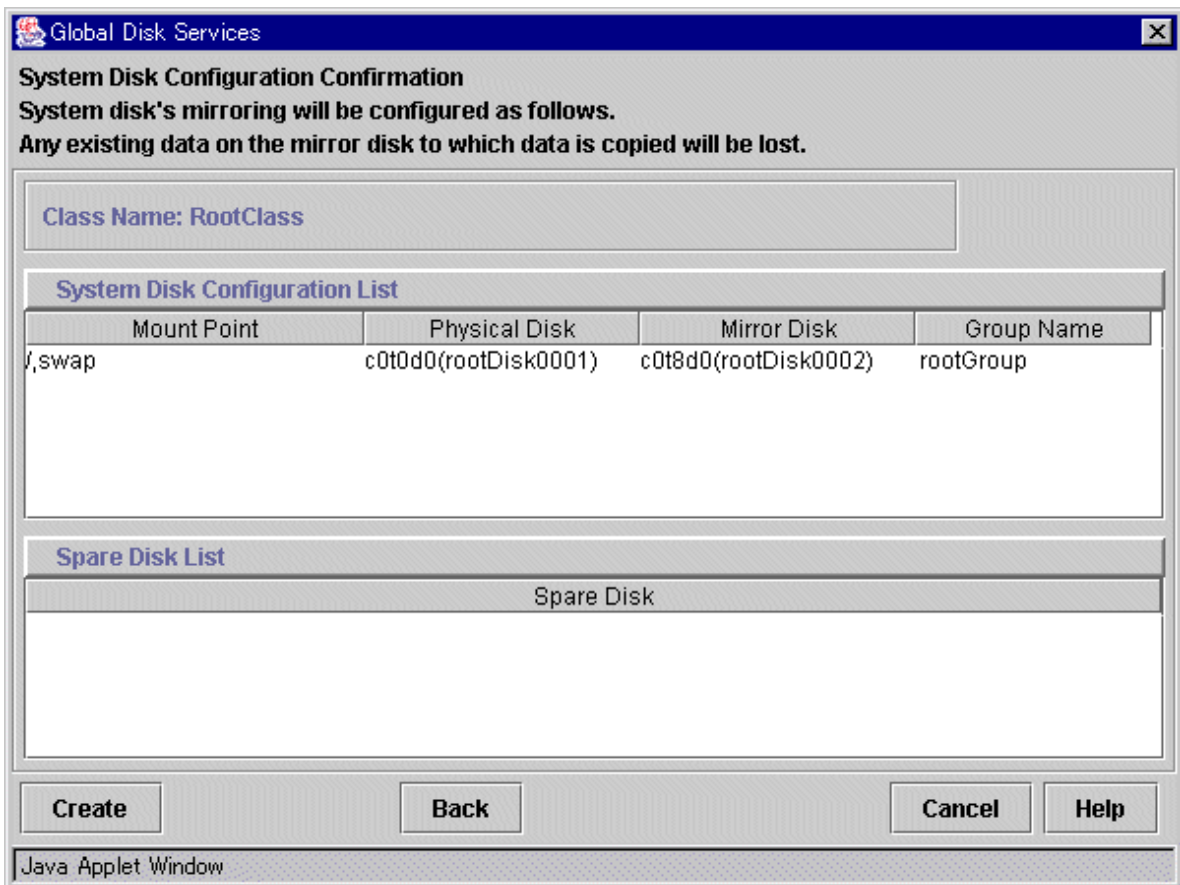
After creation of the system disk is completed, the following screen is displayed.
Check the screen contents, and then click *OK*.

Set up mirroring for the system disk of primecl02 on each node, and then, restart all the nodes.
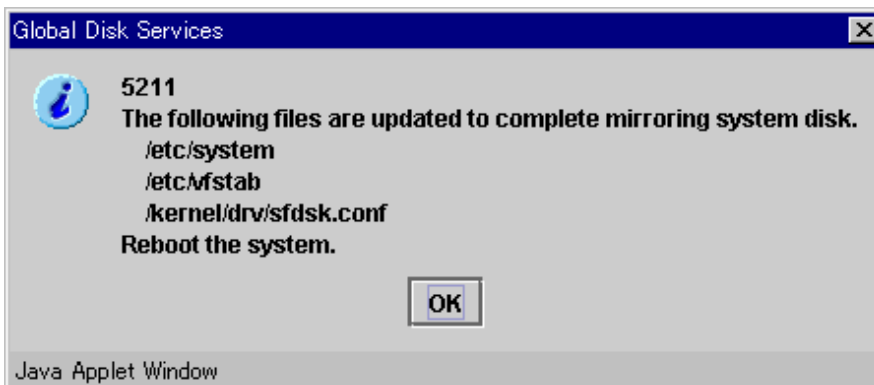
![Note icon] **Note**

. . . .

GFS partition in the system disk

If a GFS local file system is using a partition in the system disk, you need to change the following for the partition after the system disk setup is completed but before the system is restarted.

- Configuration changes in the partition information

  For information on the setup change procedure, see the "*PRIMECLUSTER Global File Services Configuration and Administration Guide*" or the manual page for the "sfxadm(1M)" command.

- /etc/vfstab file modification

  Change the device special file names from special file names of the physical slices to special file names of the logical volumes.

The GFS local file system cannot be used as /(root), /usr, /var, or /opt. Also partitions in the system disk cannot be used by a GFS shared file system. For details, see the "*PRIMECLUSTER Global File Services Configuration and Administration Guide*."

. . . .

6.2.2.2 Setting Up Shared Disks

Set up the shared disks according to the following procedures:

- Volume setup

- File system setup

Volume setup

There are four types of volumes:

    a. Single volume

    b. Mirror volume

    c. Striped volume

    d. Volume created in a concatenation group

This section separately describes the volume setup procedures for a single volume (a) and for other volumes (b, c, d). For details, see "Chapter 5 Operations" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*."

The values to be set for the individual items correspond to the values in the "GDS Configuration Worksheet."

Note

. . . .

If you plan to add, delete, or rename a disk class from the GDS Management screen (hereinafter main screen), close the Cluster Admin screen before starting the operation.

. . . .

Single volume setup

    If you are not using a single volume, this setup is unnecessary.

    Operation Procedure:

      1. Start the main screen

        Choose *Global Disk Services* on the Web-Based Admin screen.

2. Disk class creation and physical disk registration

At the main screen, select *Class Configuration* from the *Settings* menu.
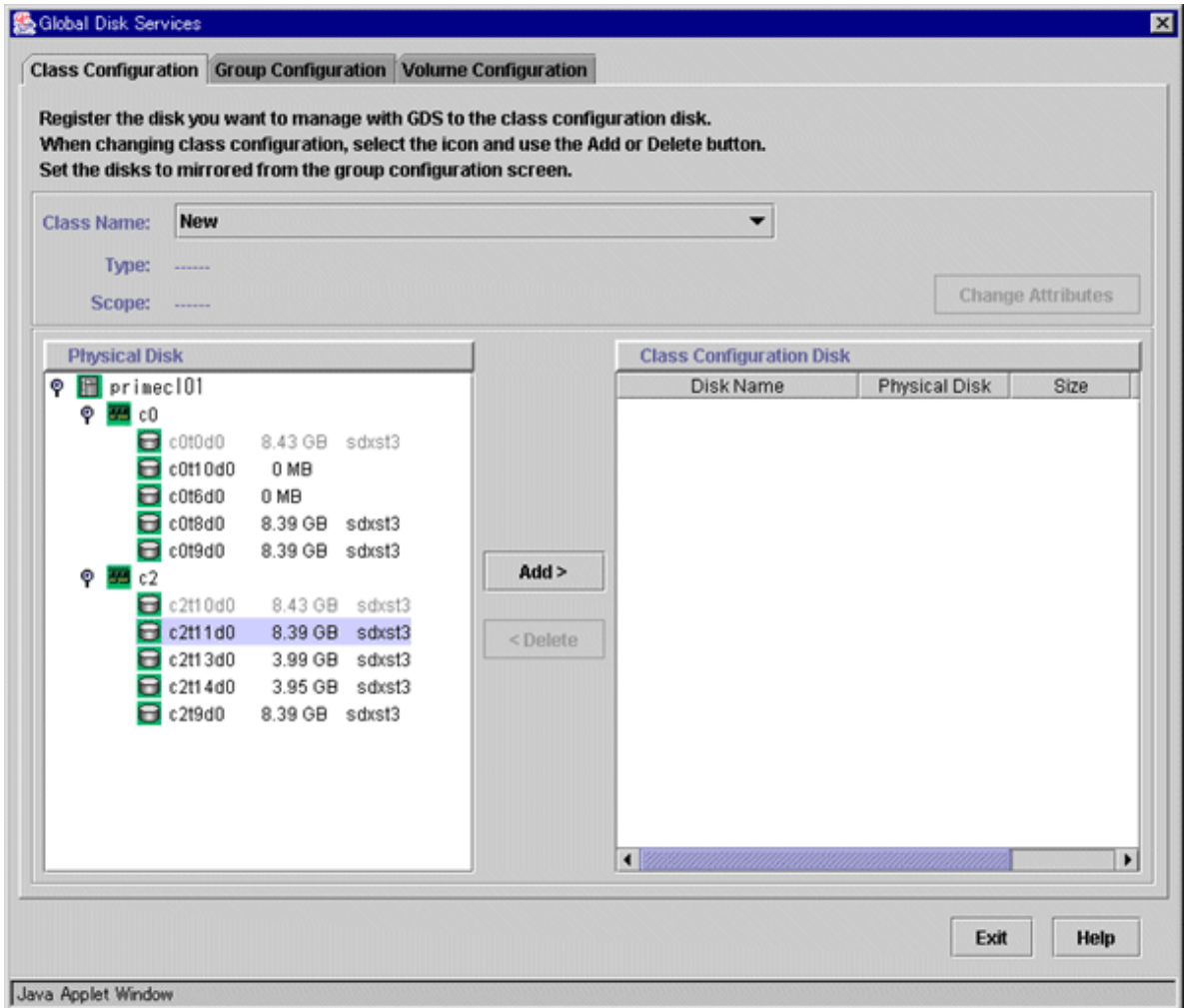
At the above screen, select the physical disk to be registered from the "*Physical Disk*" list, and then click *Add*. When *Add* is clicked, the class attribute definition screen opens. Enter the *Class name* but do not change the *Type* value (leave the value as "shared"). Then click *Exit*.

3. Disk type attribute setup

At the main screen, select the disk that was registered in Step 1 from the disk information field, and select *Operation* -> *Change Attributes* from the menu bar.
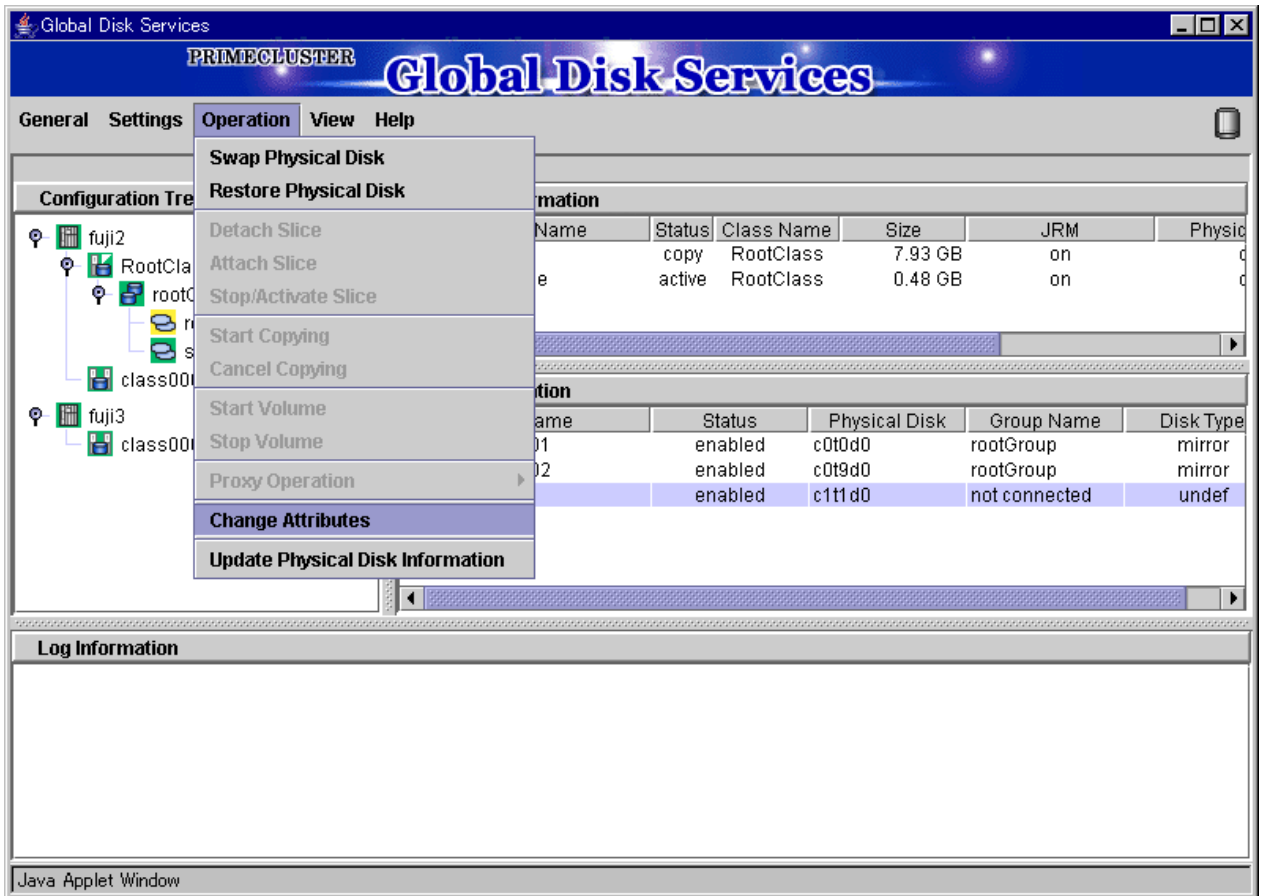
Set *Disk Type* to "single," and then click *OK*.
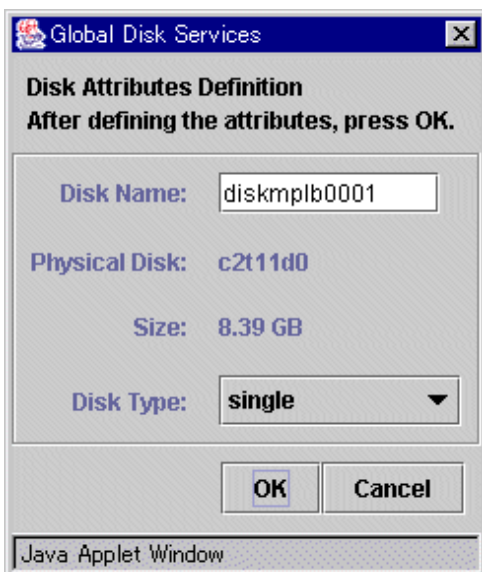
4. Volume creation

Select *Settings* -> *Volume Configuration*, and then select the disk that was registered in Step 1 from the *Group and Disk List*. Select "*Unused*" in the volume diagram, and enter the "Volume Name," the "Volume Size," and the volume attributes.

Click *Add* to enable the settings.

Check the settings, and then click *Exit*.
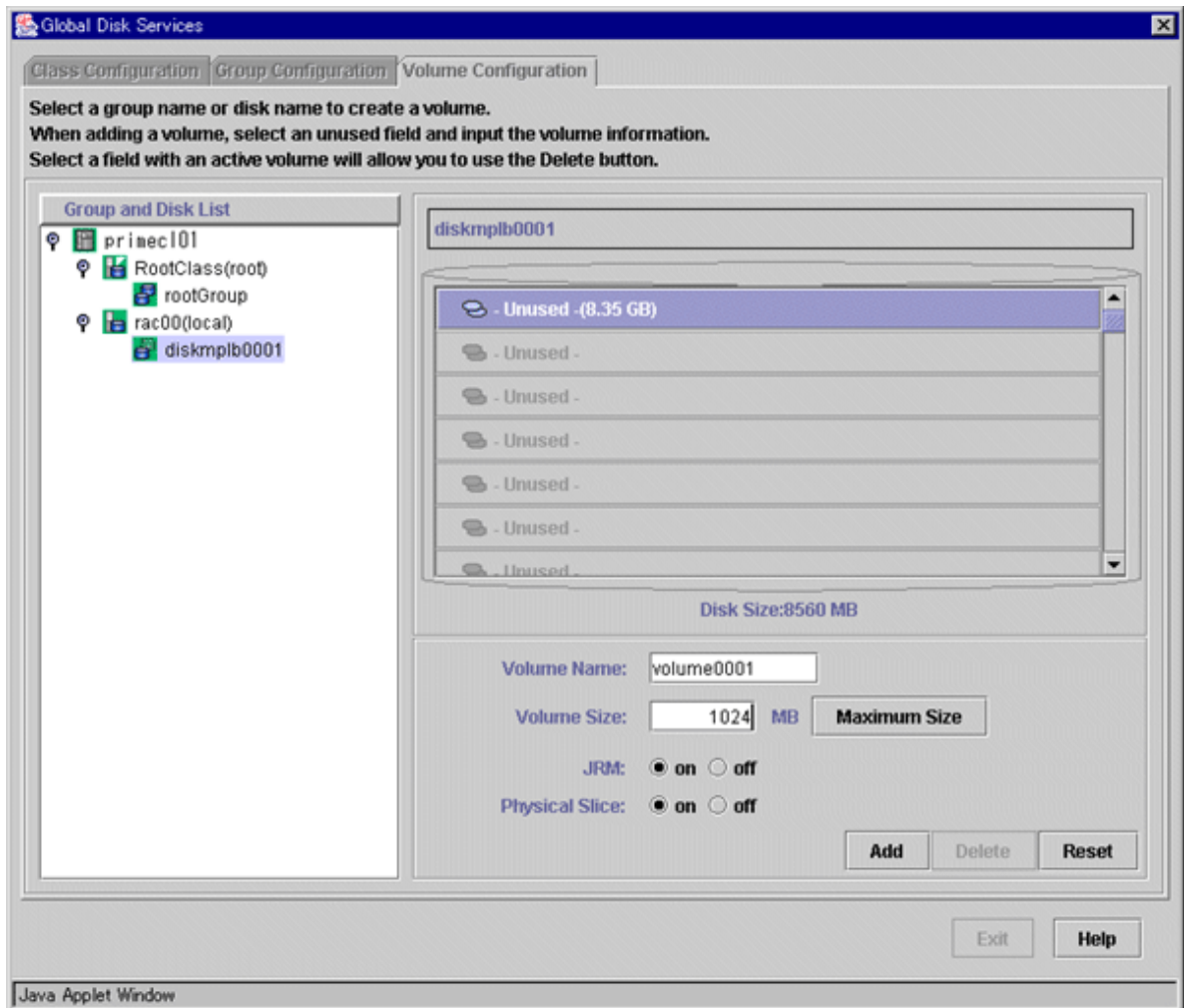
Setup for other volumes

If you are using only a single volume, this setup is unnecessary.

Operation Procedure:

1. Creating disk classes and registering physical disks

   At the main screen, select *Class Configuration* from the *Settings* menu.
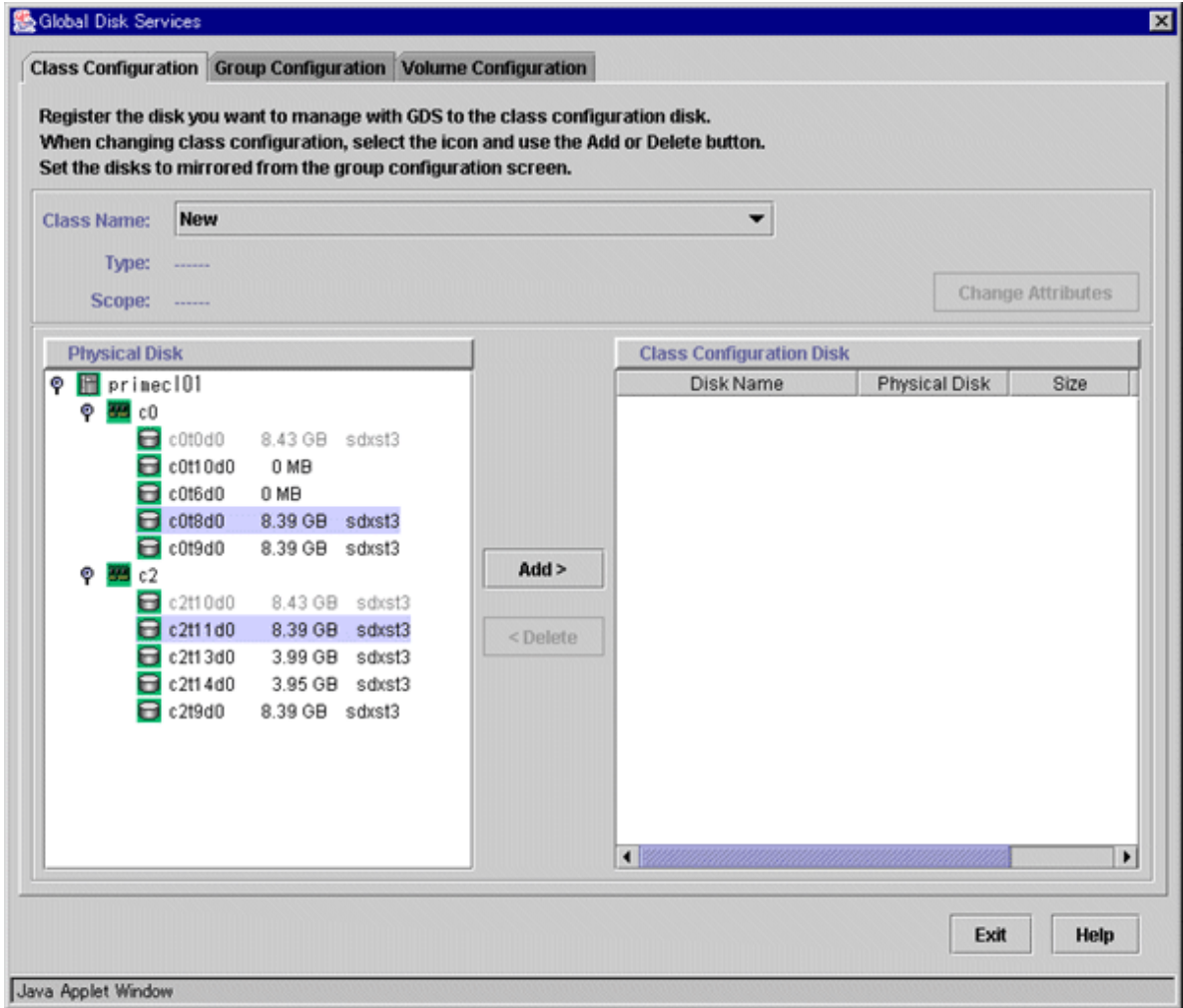
At the above screen, select the physical disk to be registered from the *Physical Disk* list, and then click *Add*. When *Add* is clicked, the class attribute definition screen opens. Enter the "Class name" but do not change the "Type" value (leave the value as "shared"). Then click *Exit*.

1. Setting up the disk group configuration

   Click the Group Configuration tab.
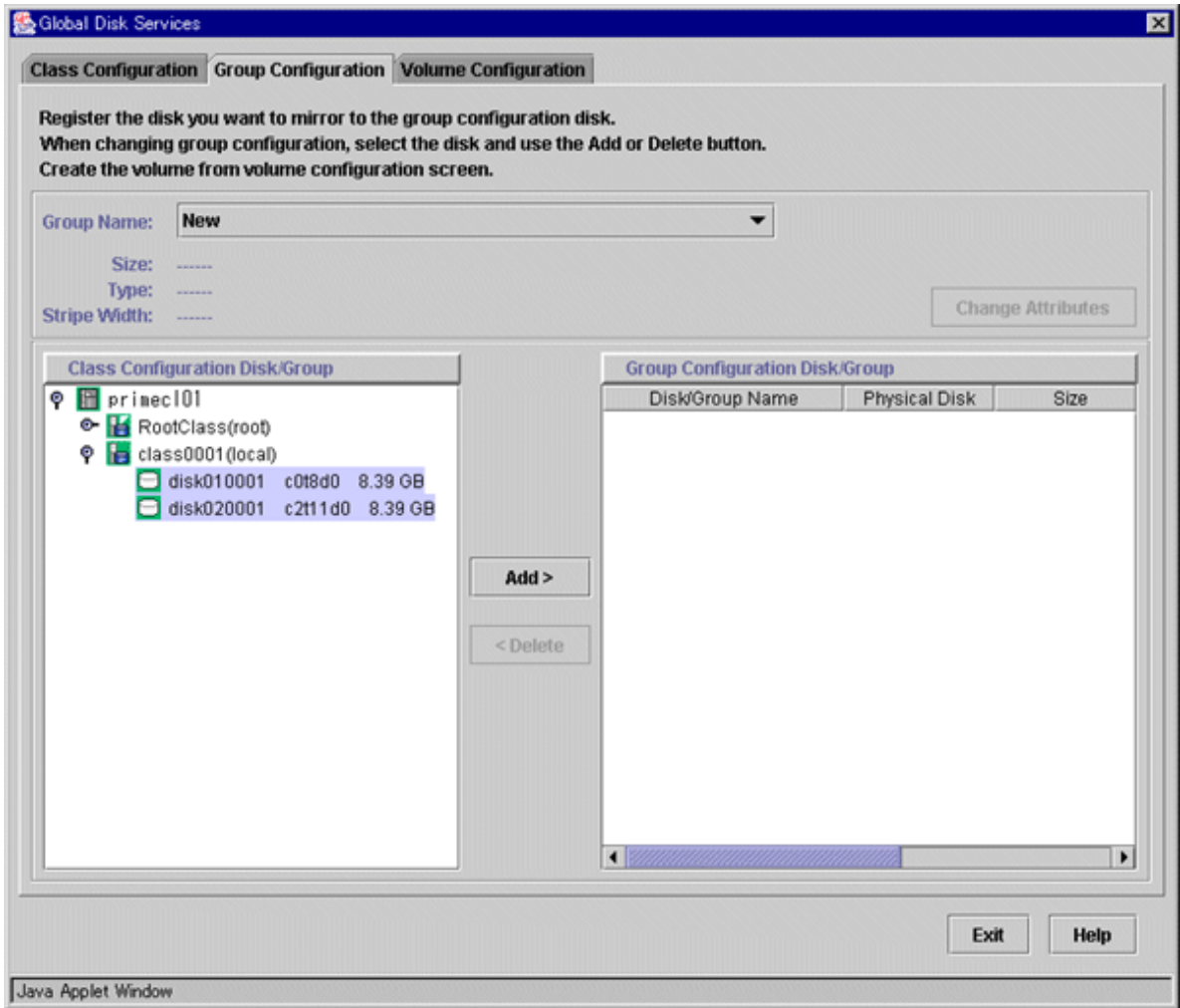
   At the above screen, select the disk to be added to the group from the *Class Configuration Disk/Group* list, and then click *Add*.

   Enter the "Group Name", "Type" and "Stripe Width" in the group attribute definition screen, and then click *OK*.

2. Creating a volume

   Click the *Volume Configuration* tab, and select the group that was created in Step 2 from the *Group and Disk List*. Select *Unused* in the volume diagram, and enter the "Volume Name," the "Volume Size," and the volume attributes.

Click *Add* to enable the settings.

Check the setup information, and then click *Exit*.
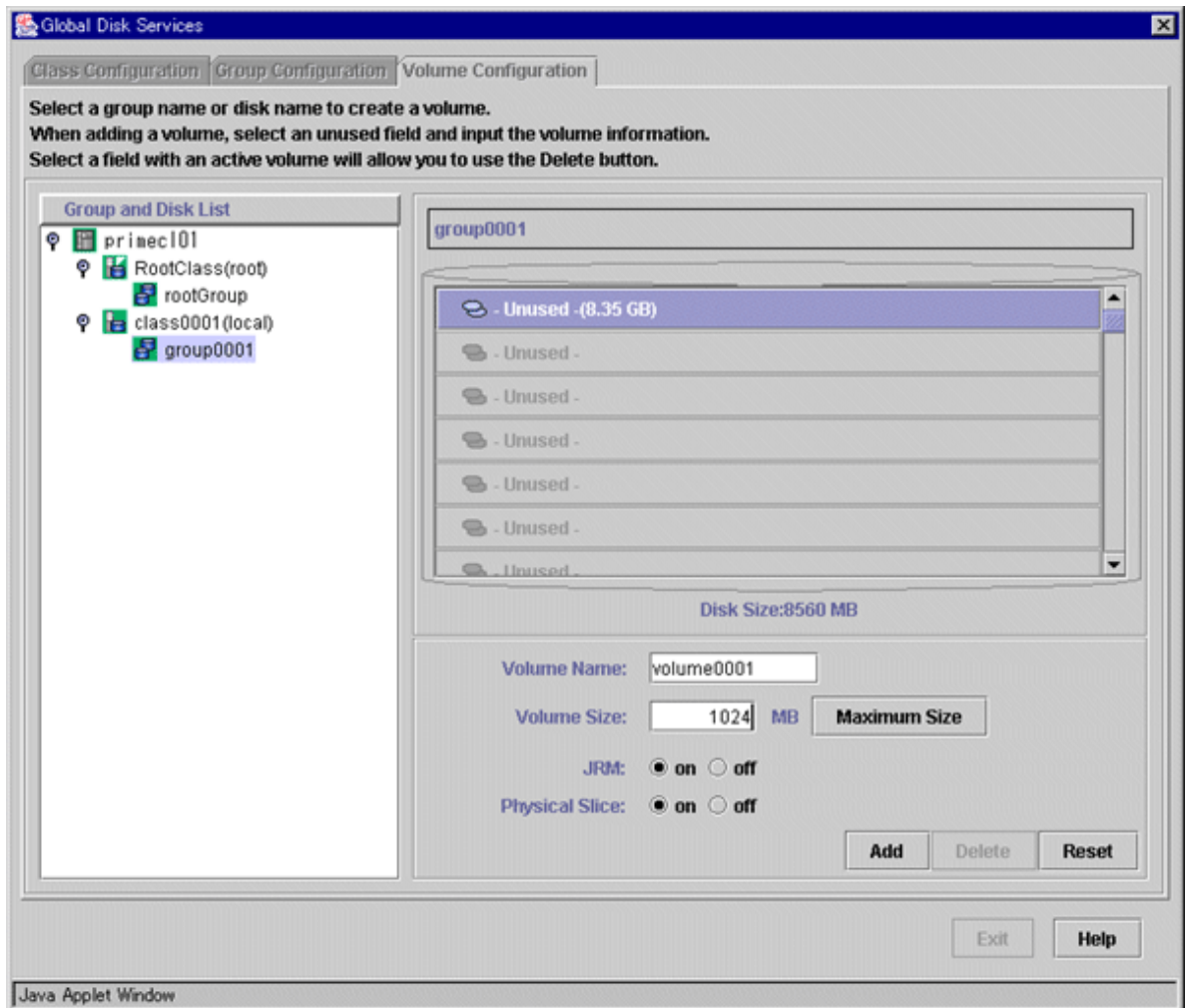
3. Checking the configuration

The disk configuration is displayed as shown below.
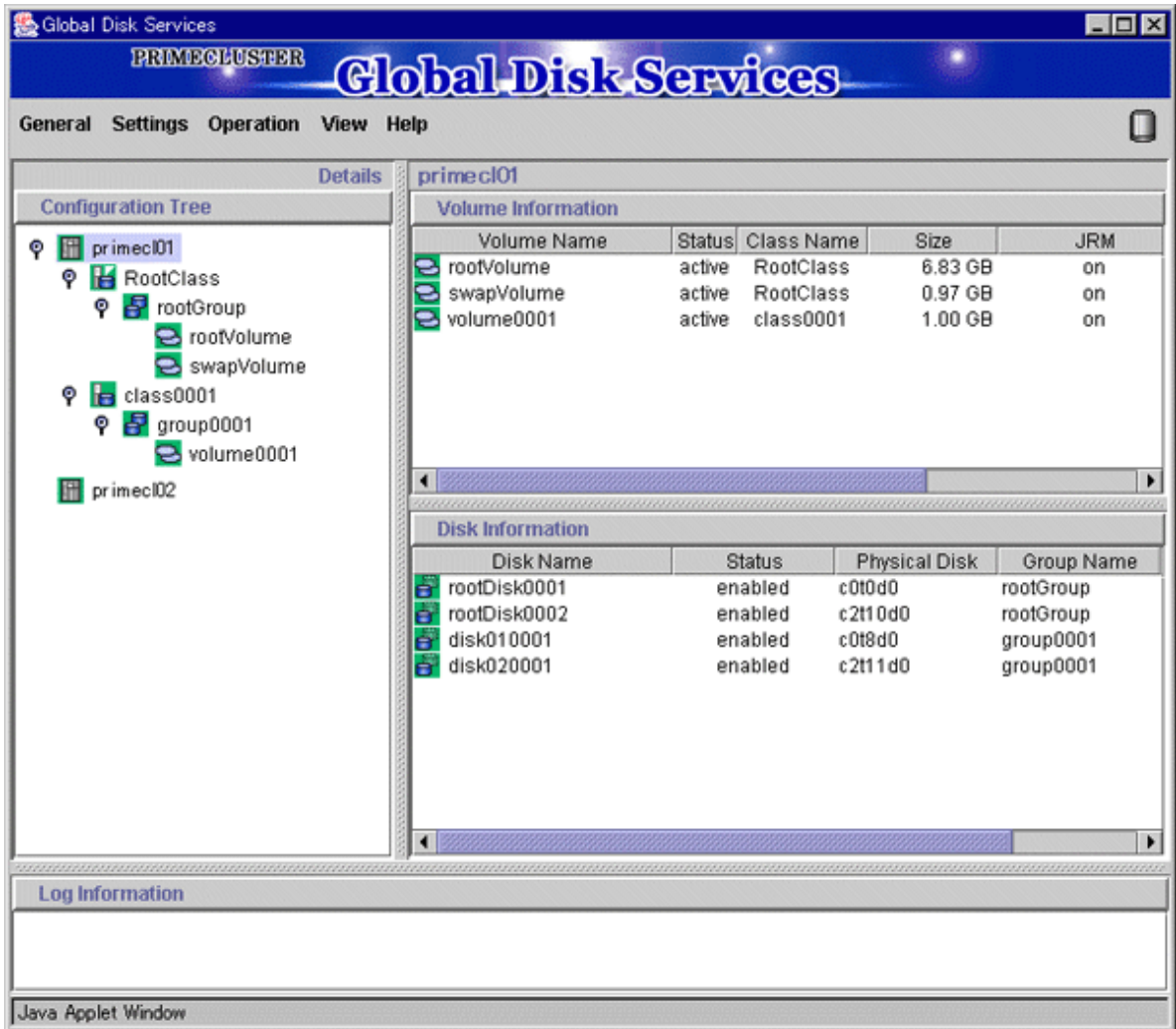
File system setup

If you are using a volume as an "ufs" file system or a GFS local file system, you can set up the file system from the file system configuration screen. For details, see "5.2.4 File System Configuration" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide.*"

Note

. . . .

If you set up a file system according to this procedure, the entries of the created file system are added to the /etc/vfstab file. To use the created file system as an Fsystem resource, you must modify the entries that were added to the "/etc/vfstab" file. For details, see "6.6.1.2 Creating Fsystem Resources."

. . . .

Operation Procedure:

1. At the main screen, select *File System Configuration* from the *Settings* menu, enter the "File System Type" of the volume and other information, and then click *Create*.
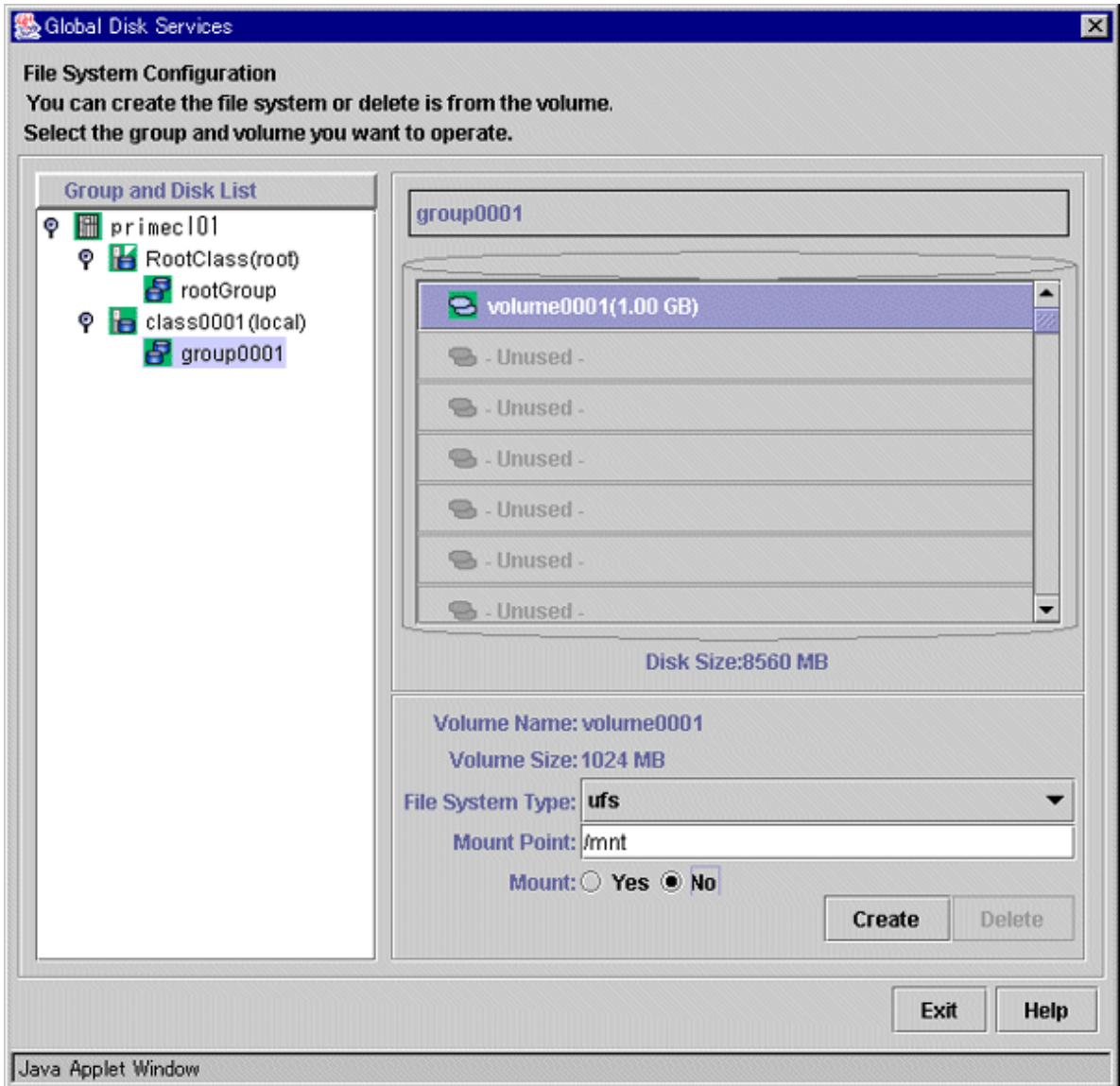
2. To exit the File System Configuration, click Exit.

6.3 Initial GFS Setup

For a Global File Services (GFS) local file system to be created on the Global Disk Services (GDS) volume, there must be a GDS volume and that volume must be active. If the volume is stopped, start the volume.

See
....
For information on the volume startup procedure, see "4.2.3 Operation" or "D.4 sdxvolume - Volume operations" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide.*"
....

To use the GFS Local File System in standby operation, set up GFS according to the flow shown below.
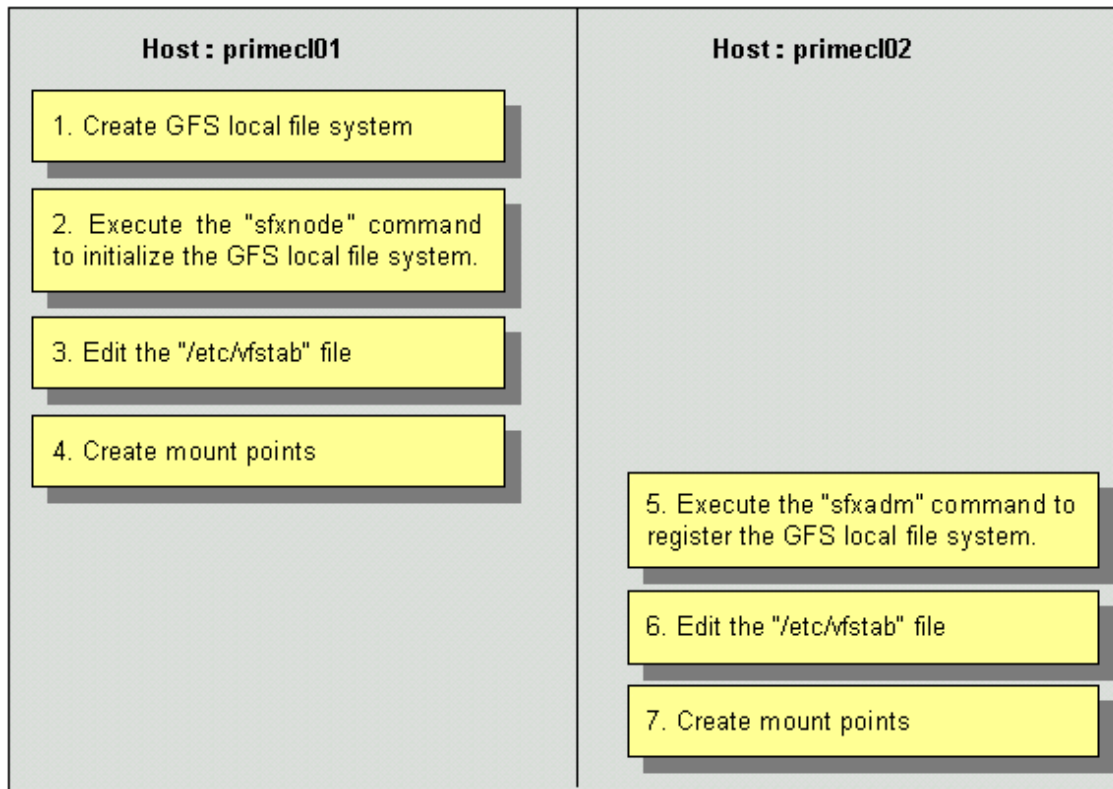
The device name and mount points that are specified here correspond to the values on the "A.7.1 GFS Local File System Setup Worksheet."

Operation Procedure:

1. Create the GFS Local File System. [Execute on the primecl01 host.]

   Create a Local File System of the file system type supported by GFS. (The file system can also be created from the GDS screen.)

   ```
   # sfxnewfs /dev/sfdsk/class0001/rdsk/volume0001
   # sfxnewfs /dev/sfdsk/class0001/rdsk/volume0002
   ```

   Get the hostid in primecl02.

   ```
   # hostid
   80f175ca
   ```

2. Execute initial setup of the GFS Local File System according to the following procedure. [Execute on the primecl01 host.]

   ```
   # sfxnode -a -n primecl02 -i 80f175ca /dev/sfdsk/class0001/rdsk/volume0001
   # sfxnode -a -n primecl02 -i 80f175ca /dev/sfdsk/class0001/rdsk/volume0002
   ```

3. Edit the /etc/vfstab file. [Execute on the primecl01 host.]

   Using a text editor like vi, add the GFS local file system entries to the /etc/vfstab file.
   Make sure that each line begins with "#RMS#".

   ```
   #RMS#/dev/sfdsk/class0001/dsk/volume0001 /dev/sfdsk/class0001/rdsk/volume0001  /disk1 sfxfs - no
   -
   #RMS#/dev/sfdsk/class0001/dsk/volume0002 /dev/sfdsk/class0001/rdsk/volume0002  /disk2 sfxfs - no
   -
   ```

4. Create the mount points. [Execute on the primecl01 host.]

   ```
   # mkdir /disk1
   # mkdir /disk2
   ```

5. Register the GFS Local File System. [Execute on the primecl02 host.]

```
# sfxadm -A /dev/sfdsk/class0001/rdsk/volume0001
# sfxadm -A /dev/sfdsk/class0001/rdsk/volume0002
```

6. Edit the /etc/vfstab file. [Execute on the primecl02 host.]

Using a text editor like vi, add the GFS local file system entries to the /etc/vfstab file. Check that each line begins with "#RMS#".

```
#RMS#/dev/sfdsk/class0001/dsk/volume0001 /dev/sfdsk/class0001/rdsk/volume0001 /disk1 sfxfs - no
-
#RMS#/dev/sfdsk/class0001/dsk/volume0002 /dev/sfdsk/class0001/rdsk/volume0002 /disk2 sfxfs - no
-
```

7. Create the mount points. [Execute on the primecl02 host.]

```
# mkdir /disk1
# mkdir /disk2
```

6.4 Setting Up the Application Environment

Configure an environment for the applications to be used in the PRIMECLUSTER system.

The environment configuration for the individual applications may involve registering resources to the PRIMECLUSTER system.

There are also products that require you to set up an environment that uses the shared disk units and takeover networks that were set in this chapter.

See
....
See the manuals for the individual applications.
....

Setting the local MAC address

Set a local MAC address in each node if the following conditions apply:

- If MAC address takeover is to be performed and network interfaces for multiple LANs are connected to the same transmission path

- If MAC address takeover is to be performed in the network interface for an expansion LAN that Web-Based Admin View does not use

Determine one local MAC address between the interfaces for takeover, and set the address in the /etc/opt/FSUNnet/mactool/macaddr.conf file of each node. If a local MAC address is set, the LAN network interface is not activated until the cluster application is started. The interface in which the local MAC address is set cannot be used for Web-Based Admin View. You cannot set up an IP address in this interface. Also, this interface cannot be connected to power control box 4.

See
....
For details, see the "Netcompo FNA-LAN User's Guide."
....

6.5 Setting Up Online/Offline Scripts

Create Online and Offline scripts to start and stop ISV applications and user applications in line with the userApplication state transition.

Set the created scripts as Cmdline resources and set those resources in userApplication.

For details, see "6.6.1.1 Creating Cmdline Resources."

- An Online script is started when userApplication is switched to Online.

- An Offline script is started when userApplication is switched to Offline.

- A Check script is used to monitor the state of the resource that is started or stopped with an Online or Offline script.

This section presents script examples and describes notes on script creation.

Sample scripts

This section shows samples of the Online and Offline scripts, which are set as Cmdline resources.

Figure 6.1 Start script/Stop script

```ksh
#!/bin/ksh
#
# Script.sample
#     Sample of Online/Offline Script
#
# Copyright(c) 2002 FUJITSU LIMITED.
# All rights reserved.
#
# $1 -c : OnlineScript
#    -u : OfflineScript

if [[ $1 = "-c" ]]; then
      # Start your application
elif [[ $1 = "-u" ]]; then
      # Stop your application
else
      # Default operation
      exit 1  # Error
fi
exit 0
```

The above script sample is both for the Start script and Stop script.

An example of Check script is shown below:

Figure 6.2 Check script

```
#!/bin/sh
#
# Script.sample.check
#    Sample of Check script
#
# Copyright(c) 2003 FUJITSU LIMITED.
# All rights reserved.
#
# Check the current state of target resource.

# If status is Online:
      exit 0
# If status is not Online:
      exit 1
```

Set up the above scripts in the Cmdline resource as shown below:

  - Start script $FULL_PATH/Script.sample -c

  - Stop script $FULL_PATH/Script.sample -u

  - Check script $FULL_PATH/Script.sample.check

For information on how to set up these scripts, see "6.6.1.1 Creating Cmdline Resources."

Notes on script creation

Hot-standby operation

To enable hot-standby operation of the Cmdline resources, the following must be prepared:

  - Online/Offline/Check scripts that support hot-standby operation.

  - The setting of attributes for the Cmdline resources

1. Create the Online, Offline, and Check scripts to support hot-standby operation. The sample scripts are shown below.

Figure 6.3 Start script/Stop script (hot-standby operation)

```sh
#!/bin/sh
#
# Script.sample
#    Sample of Online/Offline Script
#
# Copyright(c) 2003 FUJITSU LIMITED.
# All rights reserved.
#
# $1 -c : OnlineScript
#    -u : OfflineScript

if [[ $1 = "-c" ]]; then
     if [ ${HV_LAST_DET_REPORT} = "Offline" ]; then
          if [ ${HV_INTENDED_STATE} = "Standby" ]; then
                # commands for Offline -> Standby
          else
                # commands for Offline -> Online
          fi
     else
          # commands for Standby -> Online
     fi
elif [[ $1 = "-u" ]]; then
     if [ ${HV_LAST_DET_REPORT} = "Standby" ]; then
          # commands for Standby -> Offline
     else
          # commands for Online -> Offline
     fi
else
     # Default operation
     exit 1 # Error
fi
exit 0
```

The following example shows Check script that supports hot-standby operation.

Figure 6.4 Check script (hot-standby operation)

```
#!/bin/sh
#
# Script.sample.check
#
#    Sample of Check script
#
# Copyright(c) 2003 FUJITSU LIMITED.
# All rights reserved.
#

# Check the current state of target resource.

# If status is Online:
    exit 0

# If status is Standby:
    exit 4

# If status is Faulted:
    exit 2

# If status is Offline:
    exit 1
```

2. Setting attributes for the Cmdline resources

   Enable the STANDBYCAPABLE and the ALLEXITCODES attributes.

   For details, see "6.6.1.1 Creating Cmdline Resources."

Online/Offline script exit code

The state transition process of userApplication changes according to the exit code of the Online/Offline script:

0: Normal exit

The system assumes that the state transition of the Cmdline resource was processed normally, and state transition processing of the userApplication continues. If all the resources of the userApplication are processed normally, the state transition of the userApplication is also processed normally.

Other than 0: Abnormal exit

The system assumes that an error occurred during the state transition of the Cmdline resources and interrupts state transition processing of the userApplication.

Check script exit code

The state of the Cmdline resource is determined by the exit code of Check script. The exit code and the Cmdline resource are associated each other as follows:

0: Indicates the Online state.

Other than 0: Indicates the Offline state.

When ALLEXITCODES variables of the Cmdline resources are enabled, Check script will provide more detailed state of the resource. The exit code and Cmdline resource are associated each other as follows:

0: Indicates the Online state.

1: Indicates the Offline state.

2: Indicates the Faulted state.

3: Indicates the Unknown state.

4: Indicates the Standby state.

5: Indicates the Onlinewarning state.

6: Indicates the Offlinefaulted state.

For details, see "HELP."

Timeout

If script processing is not completed within the specified time, a timeout occurs, script processing is interrupted by the SIGTERM signal, and state transition ends with an error.

Default: 300 seconds

The timeout value can be specified with the TIMEOUT flag value of the Cmdline resources.

When creating the Cmdline resource, you need to set up a timeout value in "Setting up Cmdline flags" of "6.6.1.1 Creating Cmdline Resources." If a timeout occurs when a Cmdline resource is used, change the timeout value to an appropriate value according to the instructions in "8.1.3 Changing Resources."

Environment variables

When the script is executed, the environment variables shown in the table below are set.

Environment variable	Outline
HV_APPLICATION	This variable sets the userApplication name that the resource belongs to. Example) app1
HV_AUTORECOVER	This variable sets the AUTORECOVER value. For details on AUTORECOVER, see "Setting up Cmdline flags" in "6.6.1.1 Creating Cmdline Resources." 0: No is set. 1: Yes is set.
HV_FORCED_REQUEST	This variable sets a value that indicates whether or not forced failover was requested by operator intervention. 0: Forced failover was not requested. 1: Forced failover was requested.
HV_NODENAME	This variable contains the resource name. Example) Cmdline0
HV_OFFLINE_REASON	This variable sets the trigger for bringing the resource Offline. SWITCH: The resource was set to Offline because of a userApplication switchover request (hvswitch). STOP: The resource was set to Offline because of a userApplication stop request (hvutil -f) FAULT: The resource was set to Offline because of a resource fault. DEACT: The resource was set to Offline because of a userApplication deactivate request (hvutil -d) SHUT: The resource was set to Offline because of an RMS stop request (hvshut)
HV_SCRIPT_TYPE	This variable sets the type of script that was executed. Online: Online script Offline: Offline script
HV_LAST_DET_REPORT	This variable sets the state of the current resources.

Environment variable	Outline
	Online: Online state Offline: Offline state Standby: Standby state Faulted: Faulted state Warning: Warning state
HV_INTENDED_STATE	This variable sets the resource state that is expected after state transition is completed. Online: Online state Offline: Offline state Standby: Standby state Faulted: Faulted state Warning: Warning state
NODE_SCRIPTS_TIME_OUT	This variable sets the timeout duration (seconds) of the script. Example) 300

RMS also has other environment variables. For information on the other environment variables, see the files below.

File Names	Contents
hvenv	Environment variable file to be defined in the entire cluster system
hvenv.local	Environment variable file to be defined only in that node

See
. . . .

- For details on the RMS environment variable files (hvenv and hvenv.local), see "2.9 Environment Variables" in the "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide*."

- For details on the RMS environment variables, see "10 Appendix - Environment Variables" in the "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide*."

. . . .

6.6 Setting Up Cluster Applications

This section explains how to set up a cluster application by using the cluster application GUI.

The terms "userApplication" and "Resource" that appear on the setup screen refer to "cluster application" and "resource" respectively.

Information
. . . .

- After you finish setting up the cluster application, start the cluster applications. For instructions on starting the application, see "7.2.2.1 Starting a Cluster Application."

- For instructions on changing a cluster application, see "8.1 Changing the Cluster Configuration." For instructions on deleting a cluster application, see "8.8 Deleting a Cluster Application."

. . . .

Note
. . . .

- Set up the cluster application and resources based on the cluster application and resource information in the "A.9 Cluster Application Worksheets," which was created in the planning stage. If you need to change the cluster application after it is created, the worksheet is helpful.

- You cannot share one resource with multiple userApplication.

- If there are files with the same name as shown below on setup of a cluster application, you cannot set up a cluster application properly. If files with the same name exist, delete them before setting up a cluster application.

    - /tmp/hostcheck

    - /tmp/sdxinfo

    - /tmp/work

    - /tmp/workc

    - /tmp/workn

    - /tmp/macaddr

    - /tmp/wgcnf000.tmp

    - /tmp/wgcnf001.tmp

    - /tmp/wgcnf002.tmp

    - /tmp/wgcnf003.tmp

- If you use userApplication Configuration Wizard, the CF remote services (remote file copy and remove command execution) are enabled automatically. For instructions on disabling these definitions, see Notes in "5.1.1 Setting Up CF and CIP."

- Do not make any settings that cause multiple cluster system nodes to be monitored with a single primary management server of Web-Based Admin View.
  Use a single primary management server of Web-Based Admin View to monitor a single cluster system.

- If you click <Registration> on the confirm registration screen for userApplication or Resource, the dialog boxes will appear (0805 followed by 0813) to show the status of the configuration process. While these dialog boxes are displayed, do not operate the userApplication Configuration Wizard screen, which is on the background.

Setting up cluster applications by using GUIs

Operation Procedure:
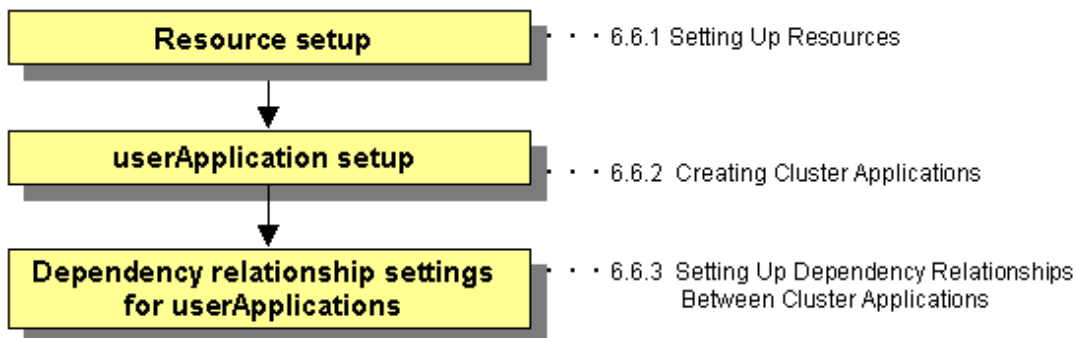
1. Resource setup

    Create all resources that make up the userApplication.

2. userApplication setup

    Use all the resources and create the userApplication.

🗒 Note

- RMS will not start if there is any remaining resource that is not part of userApplication. Delete unnecessary resources.

- To create multiple userApplication, repeat steps 1) and 2).

3. Dependency relationship settings for userApplication

   Dependency relationship settings are enabled when two or more userApplication are created.

   These settings define how each userApplication is controlled on the same cluster node.

Starting the userApplication Configuration Wizard

1. At the Web-Based Admin View top screen, select *Global Cluster Services*.

2. At the Global Cluster Services screen, select *userApplication Configuration Wizard*.

Explanation of GUI screen and icons

The userApplication Configuration Wizard screen has the configuration shown below.
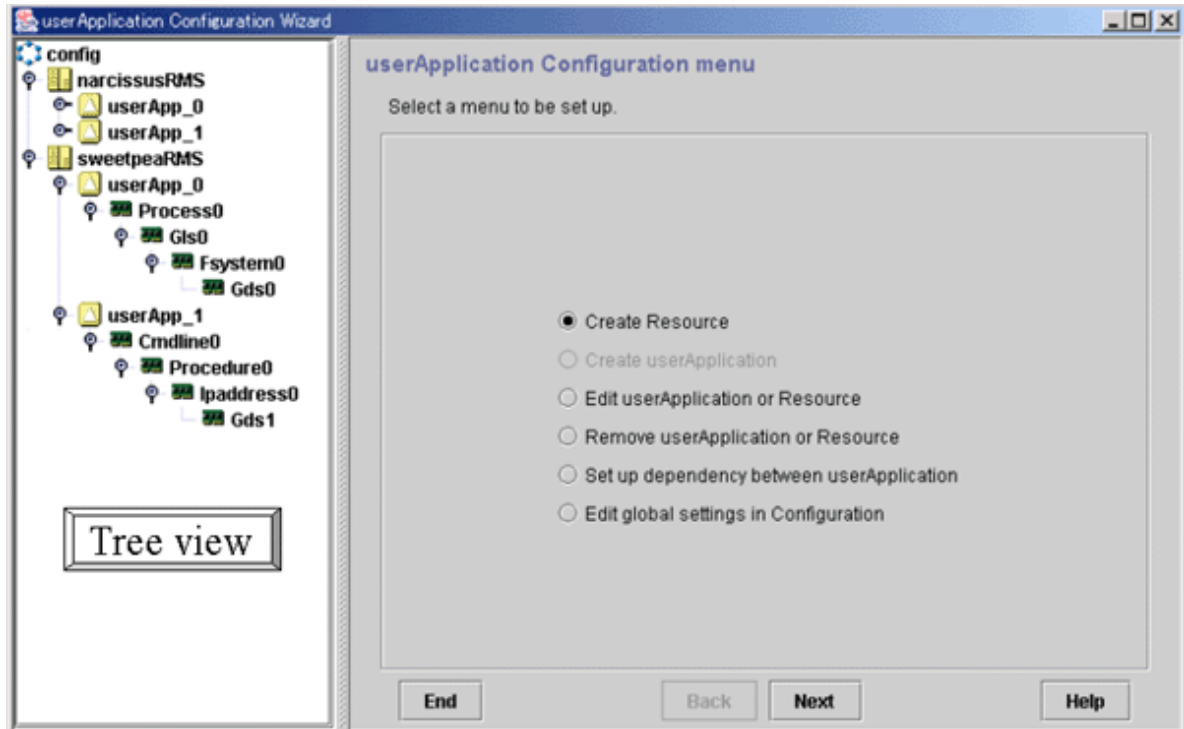
The following table shows the icons that are displayed in the tree view of the userApplication Configuration Wizard.

Icon	Description
	Indicates the configuration.
	Indicates that a discrepancy was found in the configuration information. A resource that does not belong to any userApplication or userApplication with no resources was found.
	Indicates SysNode.
	Indicates userApplication.
	Indicates userApplication that was created with the TURNKEY Wizard. The userApplication Configuration Wizard does not allow change and deletion operations for this userApplication and its Resources.
	Indicates a controller.
	Indicates a resource.

6.6.1 Setting Up Resources

This section explains how to register resources to RMS using the "userApplication Configuration Wizard" screen.

The resources include the following:

Resource type	Outline
Cmdline resources	Use this type to set a user-created script file or command as a resource. The resource is required to start or stop the ISV application or application in line with the userApplication state transition, and in the same way, to cause the state transition with the application stop.
Fsystem resources	Mounts a file system at userApplication startup.
Gds resources	Sets up a disk class in userApplication. The disk class is defined in Global Disk Services.
Gls resources	Sets up a takeover IP address in userApplication. The Ipaddress is defined in Global Link Services (network redundancy).
Takeover network resources	Sets up a takeover network in userApplication. The two types of takeover addresses are takeover IP address and takeover node name.
Procedure resources	Sets up a state transition procedure in userApplication.
Process monitoring resources	Sets up a user-created program to be monitored by the process monitoring function in userApplication.
Line switching resources	Sets up a line switching unit in userApplication. (used simultaneously with Netcompo WAN control or Netcompo BCDR)
ISV resources	Sets up an ISV (Independent Software Vendor) application in userApplication.

 Note

. . . .

CUI (hvw) cannot display or operate any configuration that was created by using resources other than those listed above.

. . . .

 Note

. . . .

Resource name

When you create a resource, first select a resource type and then enter a resource name. The resource name must be a string of **18 or fewer characters** consisting of alphanumeric characters and "_" (underscore), **starting with an uppercase alphabetic character**. For this name, do not use the same name ("config") as the Configuration name or the same name ("Cmdline" and "Fsystem" etc.) as resource type name.

Example: Cmdline0

. . . .

The flow of resource creation is shown below.

Figure 6.5 Resource creation flow
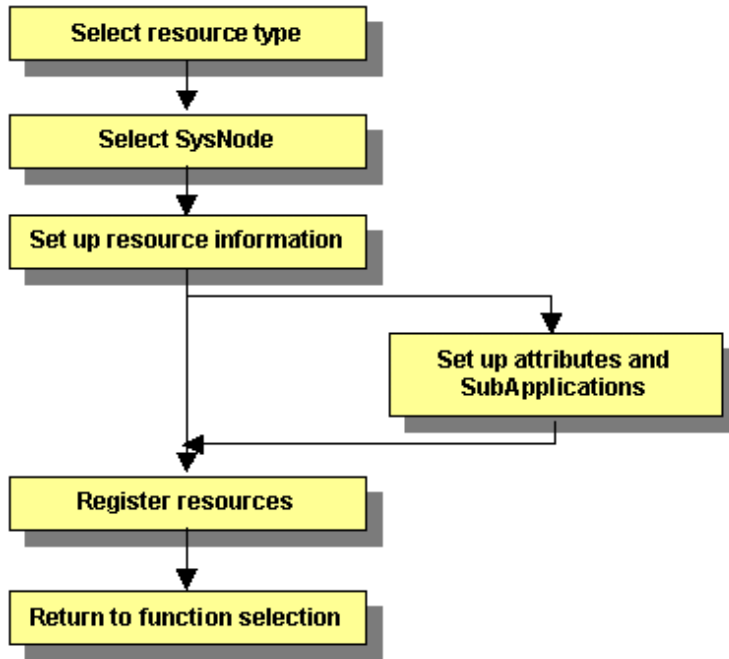
> ![Note] **Note**
> ....
> When specifying an association (priority setting) between resources of the same type, create the resource starting from the one with the higher priority.
> ....

6.6.1.1 Creating Cmdline Resources

This section describes the procedure for creating Cmdline resources.

Figure 6.6 Flow of Cmdline resource creation
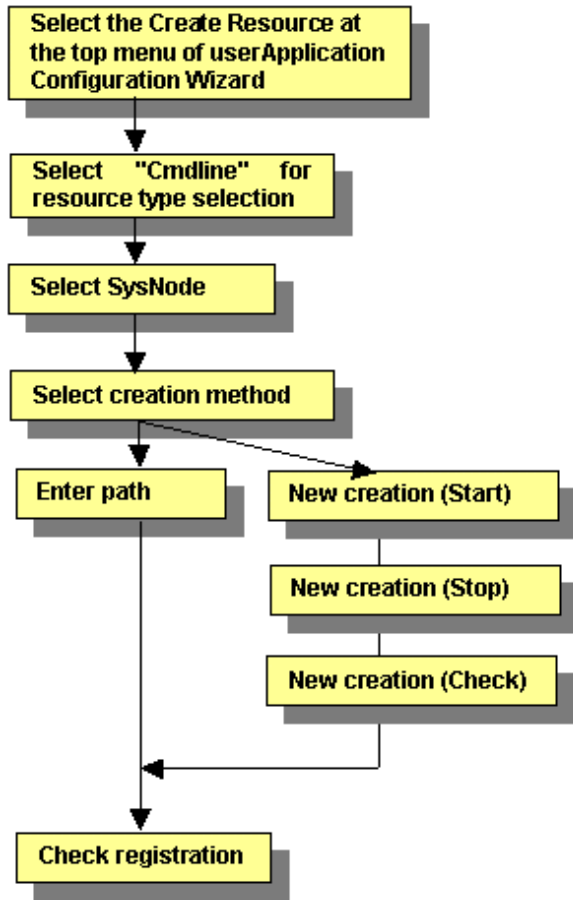
Creating resources

At the top menu of the "userApplication Configuration Wizard," select *Create Resource*.

Figure 6.7 Top menu of the userApplication Configuration Wizard
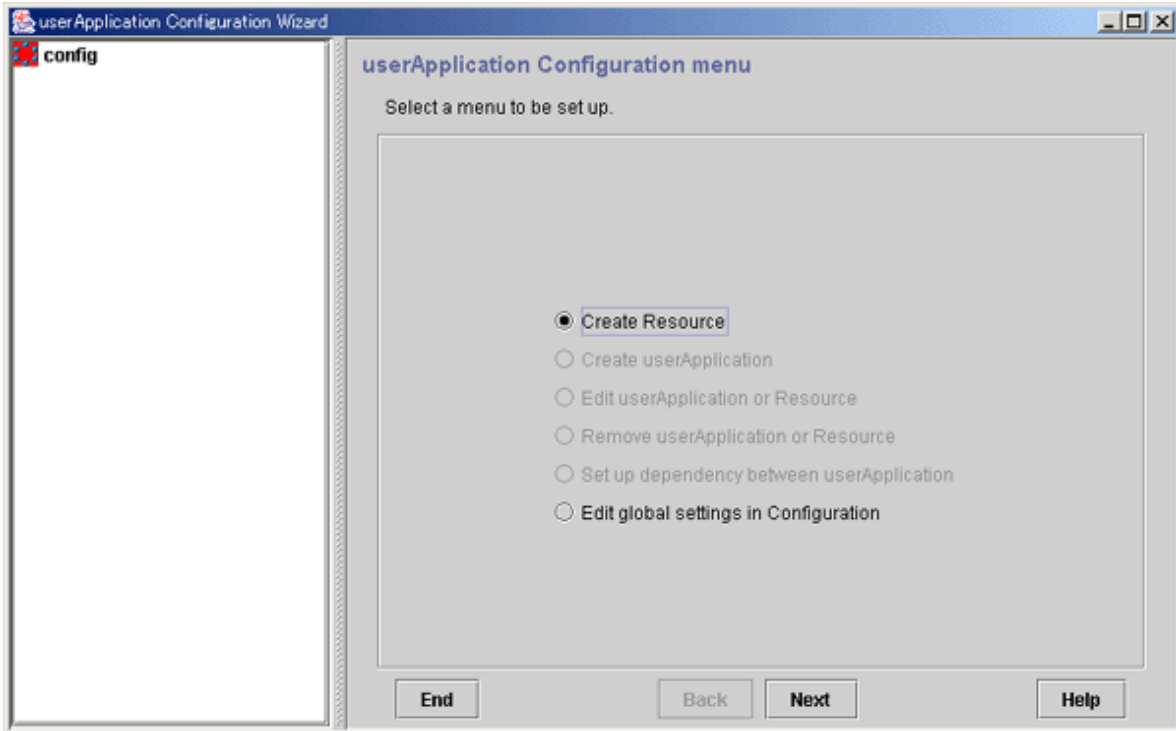
Click *Next* to go to the "Create Resource" screen.

Selecting the resource type

Select the type of resource that you want to create.

Figure 6.8 Create Resource
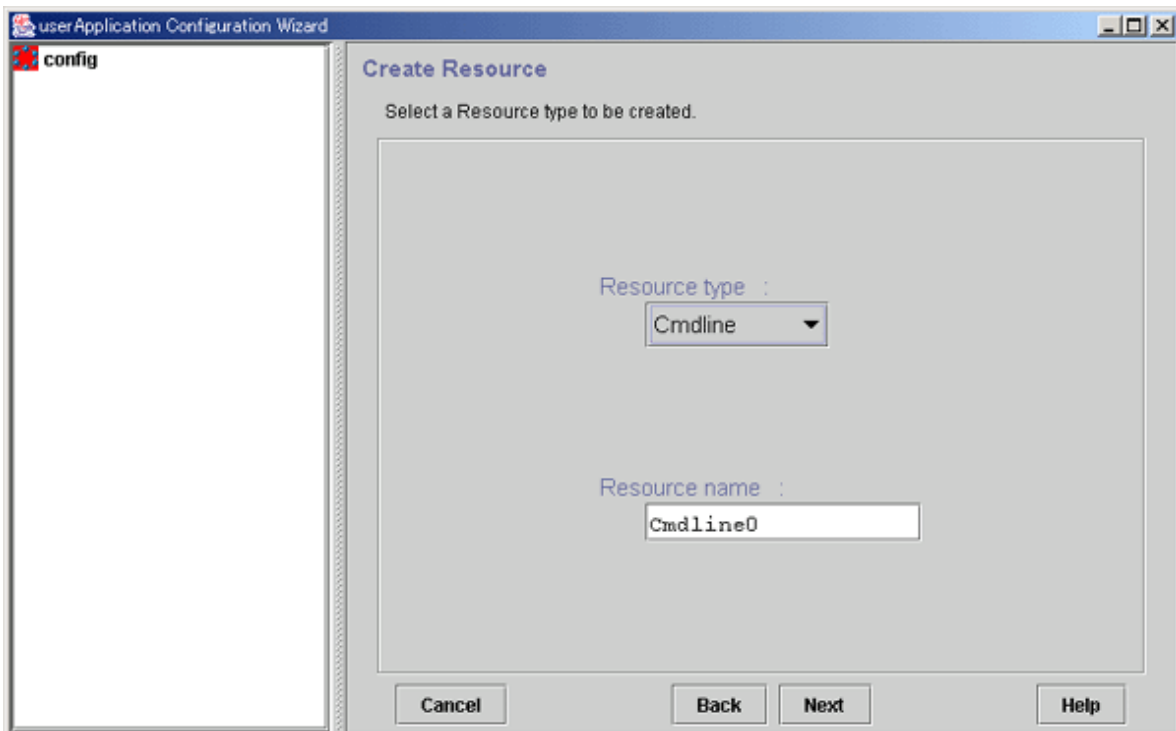
Resource Type

Select the type of resource to be created from the types described in the table below.

- **Cmdline** (Cmdline resources)

- **Fsystem** (Fsystem resources)

- **Gds** (Gds resources)

- **Gls** (Gls resources)

- **Ipaddress** (Takeover network resources)

- **Procedure** (Procedure resources)

- **Process** (Process monitoring resources)

- **SH_SWLine** (Line switching resources)

- **Oracle, Nwcl, Nwsv, Nwst** (ISV resources)

Resource Name

When you create a resource, first select a resource type and then enter a resource name. The resource name must be a string of **18 or fewer characters** consisting of alphanumeric characters and "_" (underscore), **starting with an uppercase alphabetic character**. For this name, do not use the same name ("config") as the Configuration name.

Click *Next* to go to the "Select SysNode" screen.

Selecting the SysNode

Select the SysNode in which the resource is to exist. SysNode refers to the CIP node name used by RMS.

Figure 6.9 Select SysNode
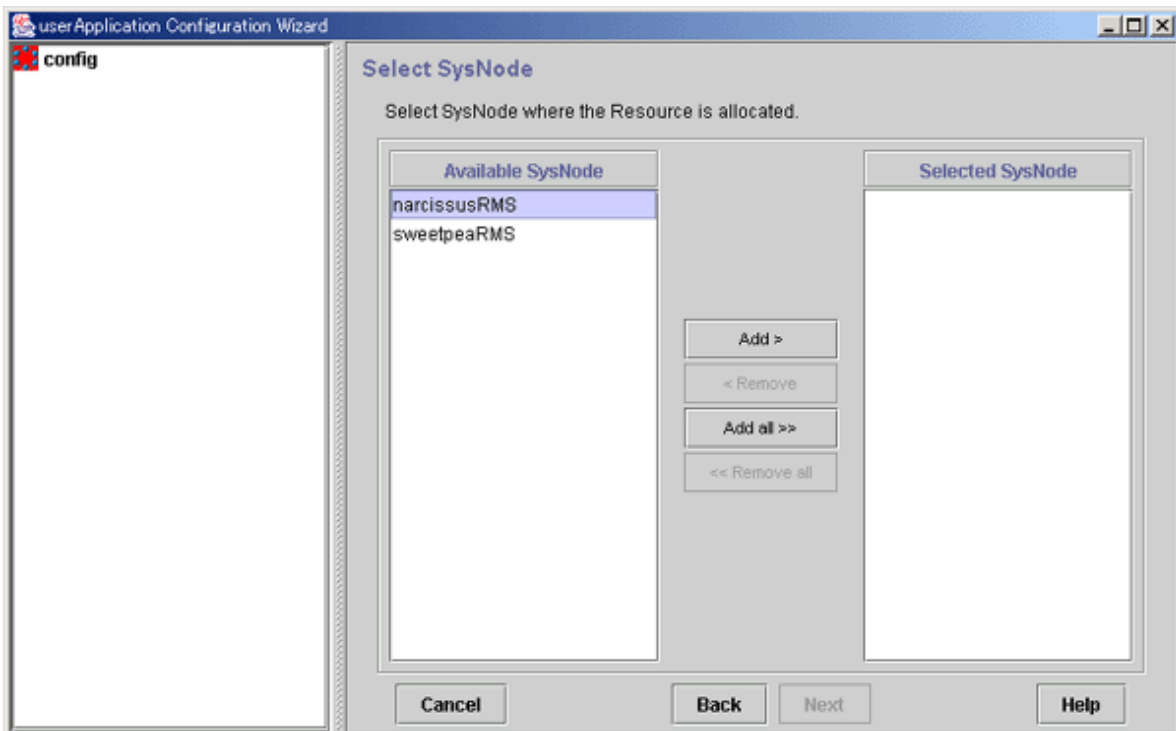
Available SysNode

Of the nodes in a cluster system, only the names accessible through Web-Based Admin View are displayed.

Selected SysNode

The SysNode name in which the resource is to exist is displayed. This information is used as filtering information for the interfaces used by each resource.

From the *Available SysNode*, select the SysNode, and then click *Add*. To add all the listed SysNodes, click *Add all*.

To delete a SysNode, select the SysNode to be deleted from *Selected SysNodes*, and then click *Remove*. To delete all listed SysNodes, click *Remove all*.

Click *Next* to go to the "Set up each resource" page.

Selecting the Cmdline creation method

Cmdline resources are general-purpose resources. By defining three scripts for starting, stopping, and monitoring the user application (program) to be monitored, you can create a resource that matches the purpose.
The scripts correspond to the following processes in RMS

Start script

It starts a program in Online processing.

Stop script

It stops a program in Offline processing.

Check script

This script is started periodically.
It notifies the user program state.

Select a method of configuring the script files and commands as the resource.

Figure 6.10 Set up command
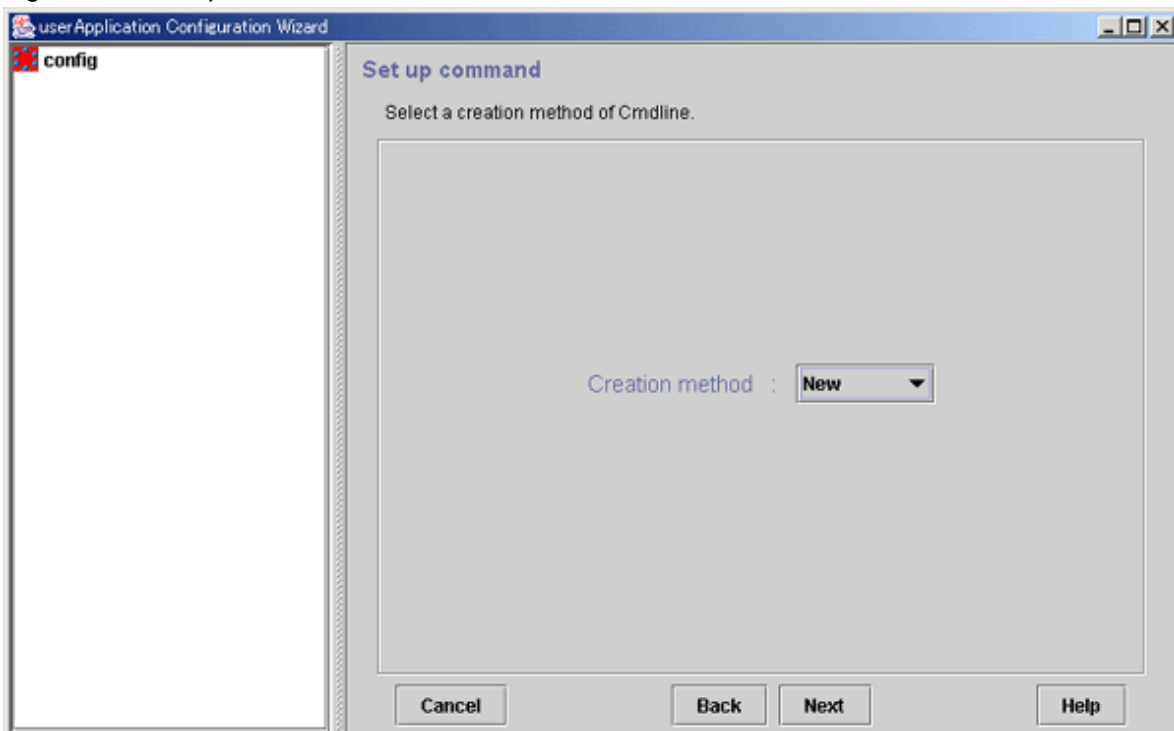
Creation methods

Select the creation method from the following:

- New

    You can create a new script file.

- Path Input

    You can specify the existing script or command.
    Click *Next*.

After completing the setup, click *Next*. This will lead you to the following operation according to the selected item:

New: Creating new script files

Path input: Creating scripts by path input

Creating new script files

![Note icon] **Note**

. . . .

When you create a script, you must carefully create a script end value, a timeout value, and environmental variable. For details, see "Notes on script creation" in "6.5 Setting Up Online/Offline Scripts."

For details on the script sample, see "Sample scripts" in "6.5 Setting Up Online/Offline Scripts."

. . . .

If you specify New for Creation method of Cmdline, the Start script creation screen appears.

Start script creation

The Cmdline contents can be input directly into the editing area. The text copy, cut, and paste operations are enabled only in the editing area.

Figure 6.11 Start script creation
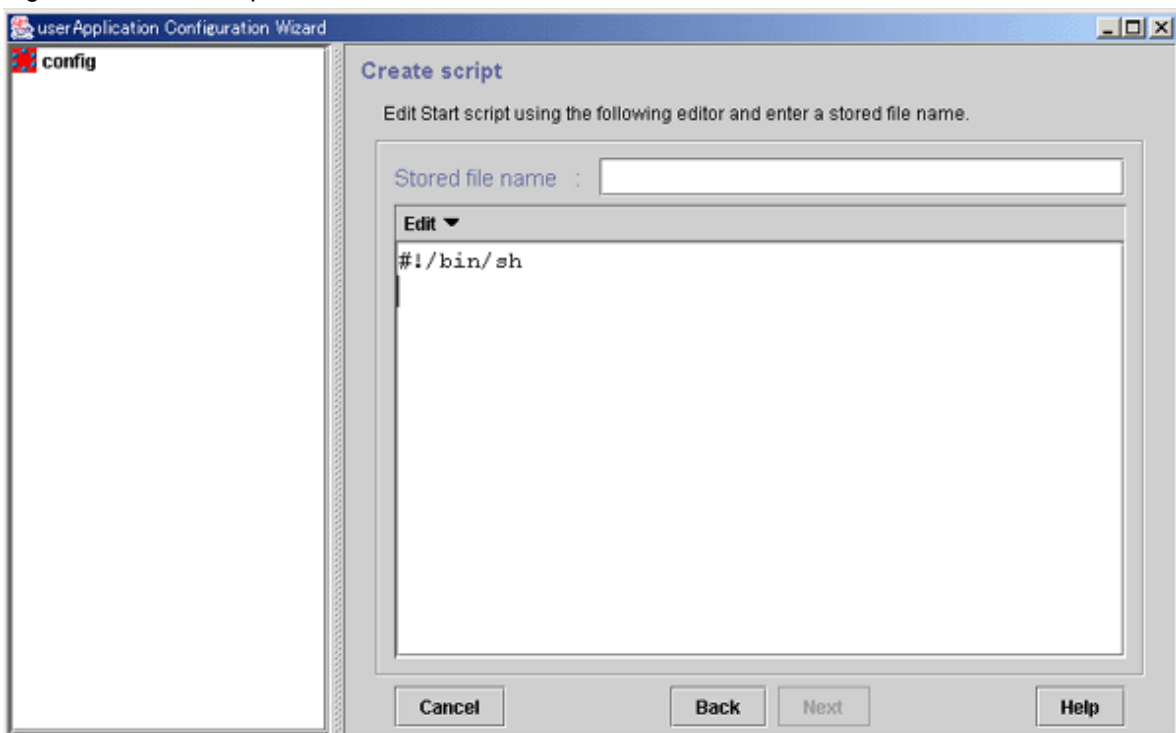
Stored file name

Specify the file name to be saved.
The following characters cannot be used in the file name or the arguments: back slash ("\"), tilde ("~"), percent sign ("%"), ampersand ("&"), and at sign ("@").
Even though processing is cancelled or the screen is closed during Cmdline creation, the script files are still stored in the "/opt/FJSVwvucw/scripts/start" directory. Therefore the scripts can be reused when the Cmdline resources are re-created.
The way in which scripts are stored in the following directories differs depending on the type of the script:

- Start script
  /opt/FJSVwvucw/scripts/start

- Stop script
  /opt/FJSVwvucw/scripts/stop

- Check script
  /opt/FJSVwvucw/scripts/check

Edit

You can edit scripts from the simplified menu. The editing operation is described below:

Item	Operation
Move cursor	You can move the cursor to any position by placing the mouse pointer at the target position and single-clicking the mouse pointer or by operating the arrow keys on the keyboard in the top, bottom, left, or right directions.
Insert text	You can input characters at the cursor position.
Delete text	To delete 1 character Press the [*Delete*] key to delete the character at the cursor position. Press the [*Backspace*] key to delete the character positioned directly before the cursor position. To delete a string with one or more characters, use the following keys: [*Delete*] key [*Backspace*] key *Edit* -> *Delete* from the menu bar
Cut text	Select a string with one or more characters in the editing area and go to *Edit* -> *Cut*.
Copy text	Select a string with one or more characters in the editing area and go to *Edit* -> *Copy*.
Paste text	Position the cursor at the position where the character string that was previously cut or copied within the editing area is to be pasted, and then select *Edit* -> *Paste*.

Click *Next* to go to "Stop script creation."

Stop script creation

Create a Stop script by using the same procedure as that for the Start script.

Figure 6.12 Stop script creation
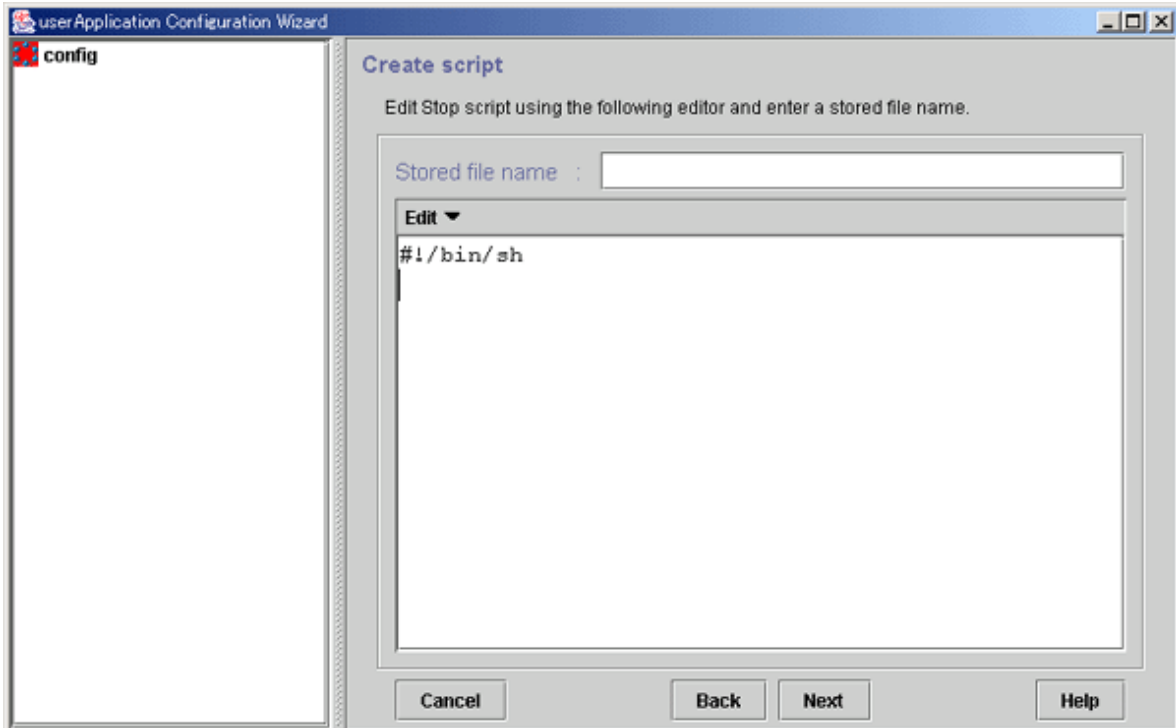
Click *Next* to go to "Check script creation."

![Note icon] **Note**

You can omit setup of this script file. Just click *Next*.

Check script creation

Create a Check script by using the same procedure as that for the Start script.

Figure 6.13 Check script creation
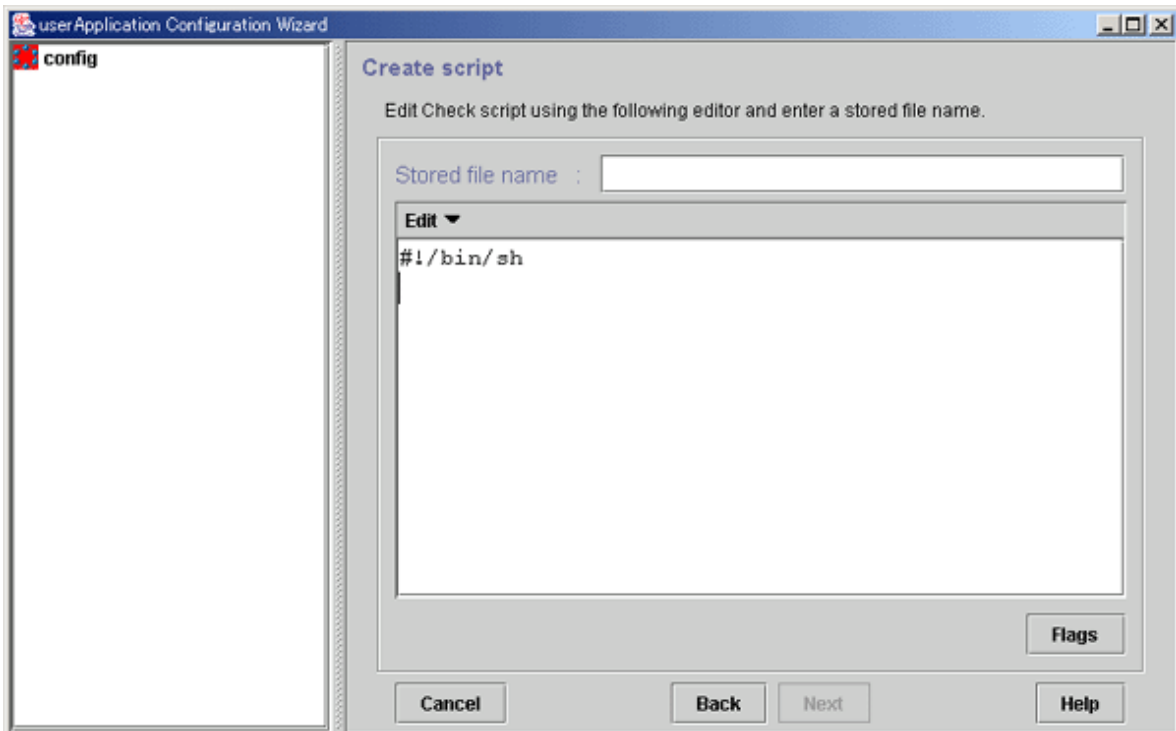
Flags button

You can set script attributes from a dialog box. See "Setting up Cmdline flags," for details.

![Note icon] Note
....
If "NULLDETECTOR" was set to "Yes" at the "Set Cmdline Flag" screen, a Check script does not need to be configured. In this case, the Check script is not configured as resource information because the Check script is not started from RMS.
....

Creating scripts by path input

If a script already exists, you can enter a path to select the script.

![Note icon] Note
....
The following characters cannot be used in the file path or the arguments: back slash ("\"), tilde ("~"), percent sign ("%"), ampersand ("&"), and at sign ("@").
....

Figure 6.14 Cmdline setup
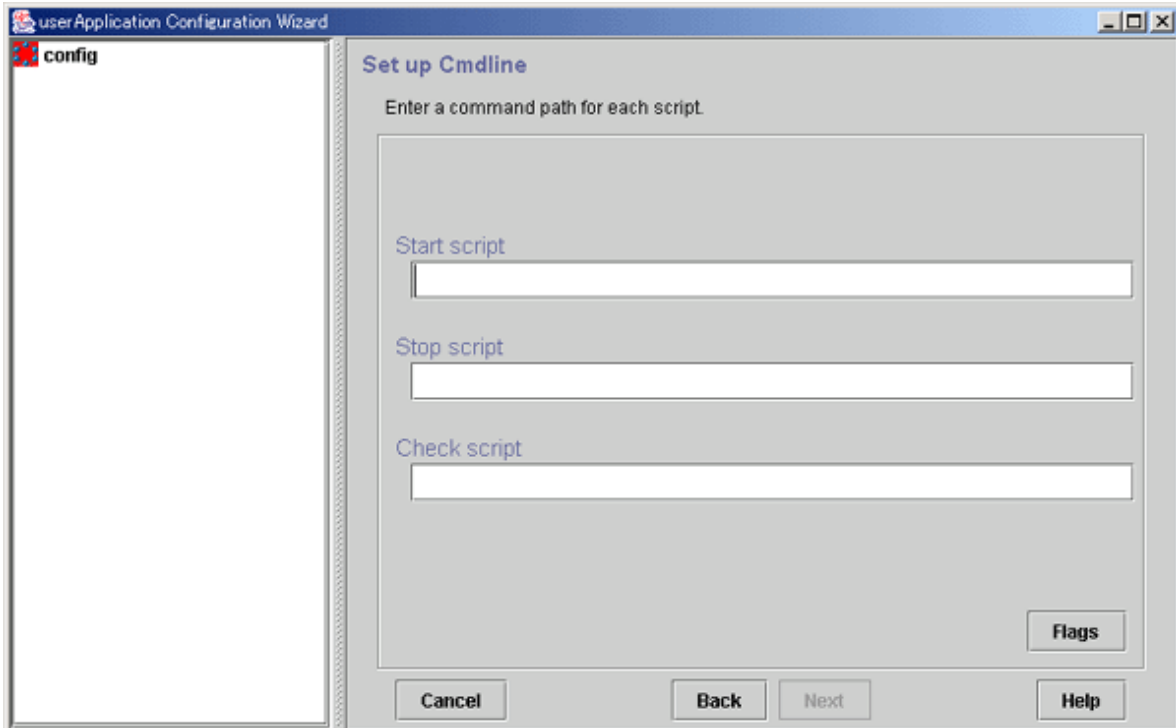
Start script

> Enter the path of the Start script. If you are specifying arguments, separate the arguments with "blanks."
> The Start script is always a required script. The script must be entered with a full path name.

Stop script

> Enter the path of the Stop script. If you are specifying arguments, separate the arguments with "blanks."
> The Stop script is not required information. The script must be entered with a full path name.

Check script

> Enter the path of the Check script. If you are specifying arguments, separate the arguments with "blanks." The script must be entered with a full path name.
> Although the Check script is always a required script, it does not have to be specified if NULLDETECTOR was set to "Yes" in the "Set Cmdline Flag" screen. Even if the script is setup, the information is deleted during resource registration because the Check script is not started from RMS.

***Flags* button**

> Set the script attributes. For information, see "Setting up Cmdline flags," which is described later.

Setting up Cmdline flags

> The method of a controlling script defined to a Cmdline resource can be adjusted by changing the flag values.

> The following default value is for Cmdline resource. Other resources have a default value individually.

Figure 6.15 Set up Cmdline flags
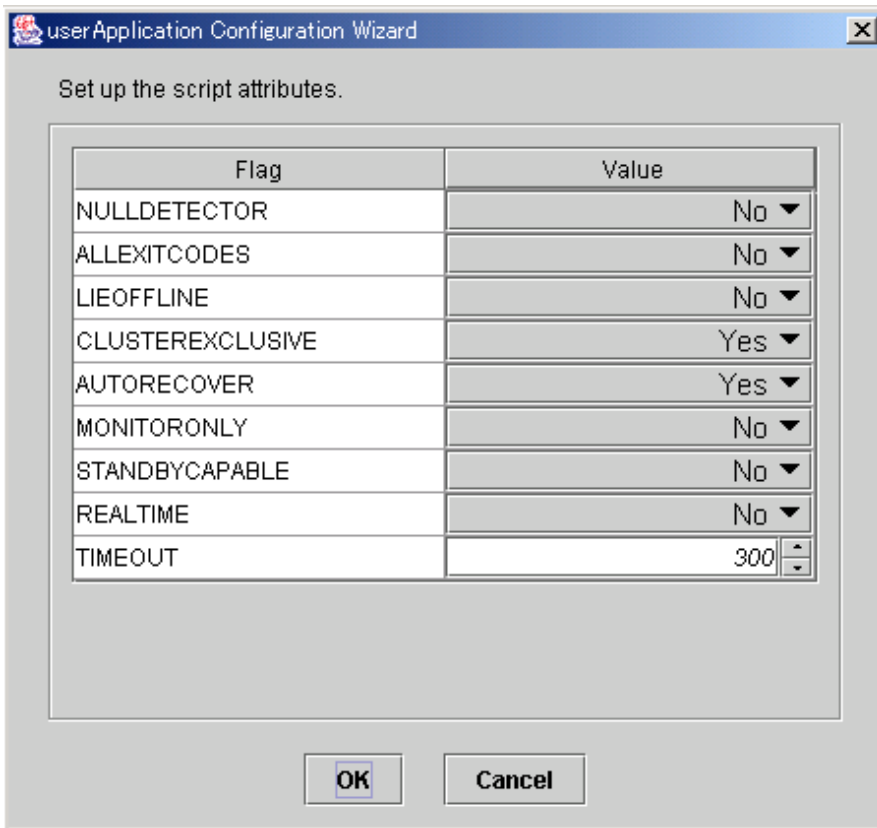
After completing the setup, click *OK* to return to the "Set up Cmdline" screen.

Flag	Outline
NULLDETECTOR	It synchronizes with the Online/Offline processing of userApplication, and executing the Online/Offline script on the Cmdline resource. It is not monitored after the Online/Offline script on the Cmdline resource processing succeeded. This flag is used to register the program that starts or stops in line with Online or Offline processing as a resource. Also, all the other flag values will be set to "No". The default value is "No."
ALLEXITCODES	If the flag is set to "No", the exit code of the Check script will be interpreted as follows: 0: Online Other than 0: Offline If the flag is set to "Yes", the exit code of the Check script will be interpreted as follows: 0: Online 1: Offline 2: Faulted 3: Unknown 4: Standby 5: Online warning 6: Offline faulted The default value is "No."
LIEOFFLINE	If this flag is set to "Yes," the resource is notified as being Offline if a Stop script is not specified. If a Stop script is specified, the failure of the script

Flag	Outline
	triggers fault processing. The default value is "No."
CLUSTEREXCLUSIVE	If this flag is set to "Yes", the resource becomes Online only on one node in a cluster system. If the resource becomes Online on two or more nodes at the same time because of a script problem, the state of userApplication to which this resource belongs becomes Inconsistent. The default value is "Yes."
AUTORECOVER	If this flag is set to "Yes," RMS tries to recover the resource if it becomes faulted. In this way, RMS tries to prevent the userApplication from being switched to another host This recovery is tried only once. The default value is "Yes."
MONITORONLY	This flag controls whether the "faulted" condition in the resource makes the userApplication Faulted. If this flag is set to yes, the resource becomes Faulted, but this does not make the userApplication Faulted. The system does not allow this value to be set to "Yes" for Cmdline resources that can be created from the GUI. If you use this flag, keep the "No" setting because RMS startup will be disabled if the flag value is set to "Yes." The default value is "No."
STANDBYCAPABLE	If this flag is set to "Yes," RMS sets the StandbyCapable attribute to 1 for this resource. For detailed information regarding this attribute, see "9 Appendix -Attributes" in "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide.*" The default value is "No."
REALTIME	If the flag is set to "No", the Check script will be started in the TS class (time-sharing scheduling) . If this flag is set to "Yes," the Check script is started with the RT class (real time mode). Note that since the operating system assigns the highest priority to processes that are started with the RT class, any bugs in the script or commands may have a large effect on system performance. The default value is "No."
TIMEOUT	This flag sets the timeout interval (seconds) for program start and stop processing. The default value is "300."

Check Cmdline registration information

You can check a list of the scripts that has been set up previously. You can also configure resource attributes by selecting the *Attributes* tab and switching the screen.

For information on the resource attributes, see "6.6.5 Attributes."

Figure 6.16 Confirmation of registration information
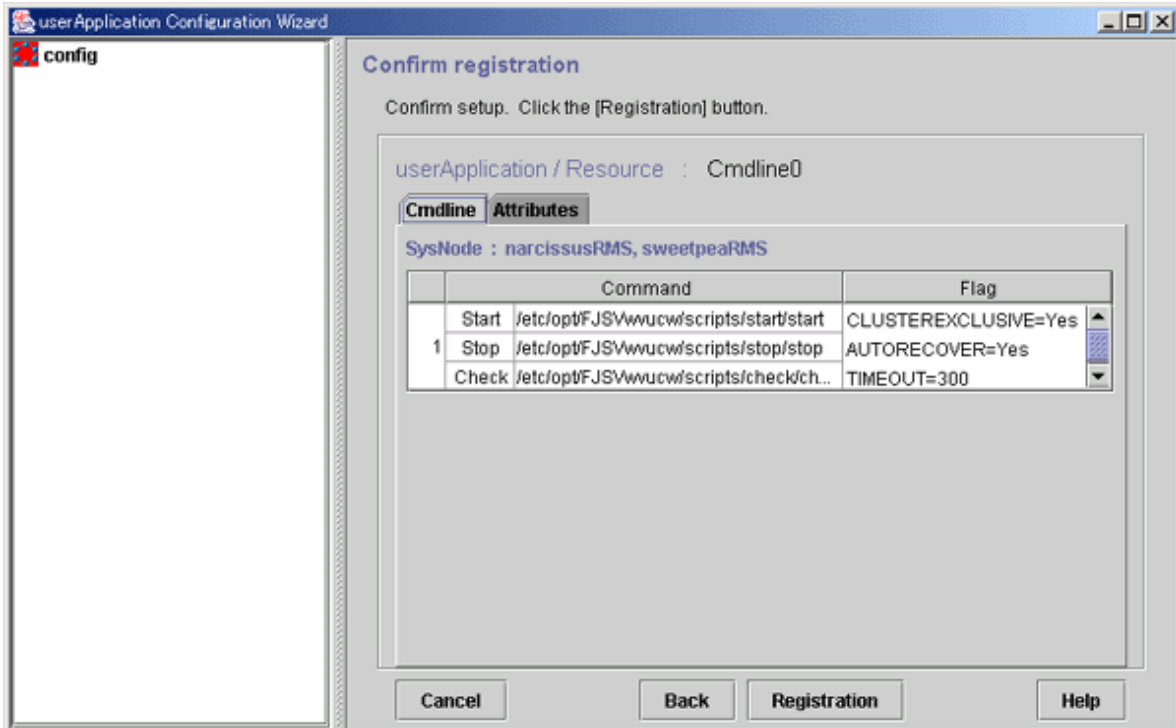
The attributes that are displayed in the "Flag" column are only those Resource attributes for which a value has been set. For example, if *No* is set for the AUTORECOVER attribute, then it means that AUTORECOVER attribute has not been set, and it is not displayed in the Flag column.
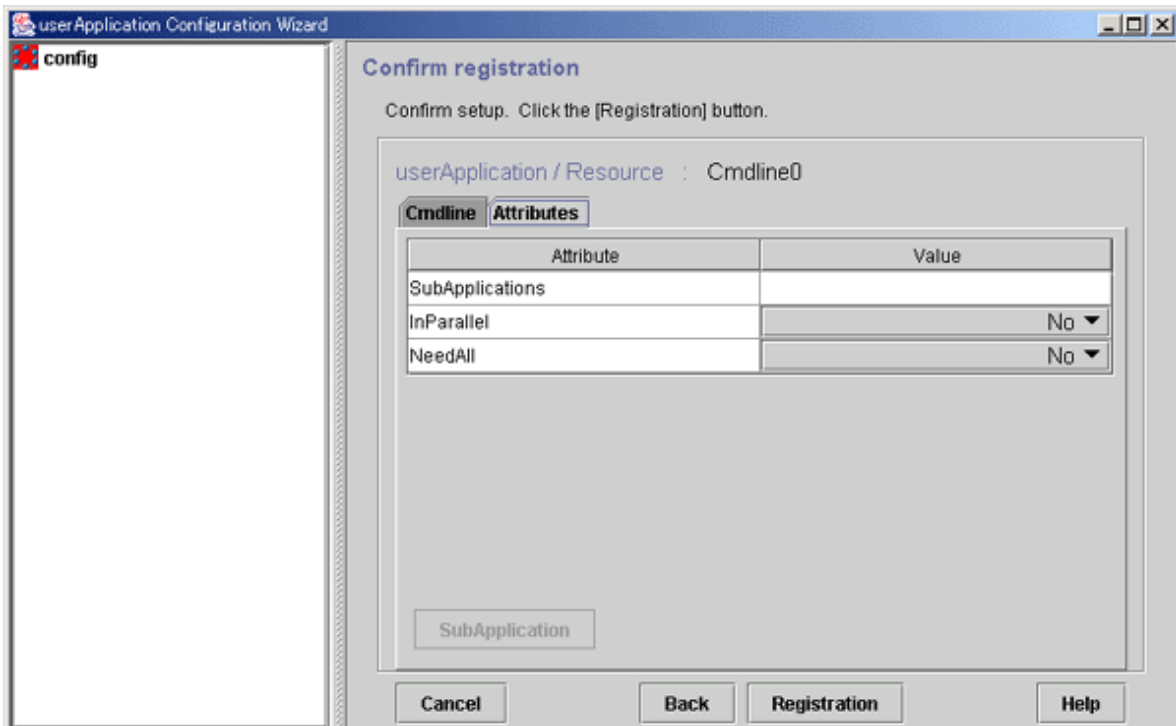
SubApplication button

Associate the other Cmdline resources or process monitoring resources with the Cmdline resource that has been created above. See "Resource association."

After checking the registration information, click *Registration*.

Resource association

Use this function to configure a startup sequence among resources of the same type.

By assigning startup priorities to resources, you can clarify the order in which the resources are to be switched to Online or Offline.
In Online processing, resources that have a high startup priority are switched to Online before resources that have a low startup priority.
Conversely in Offline processing, resources that have a low startup priority are switched to Offline before resources that have a high startup priority.

If a startup priority is not assigned or if resources have the same priority, the sequence is undefined.

Figure 6.17 Associate resources
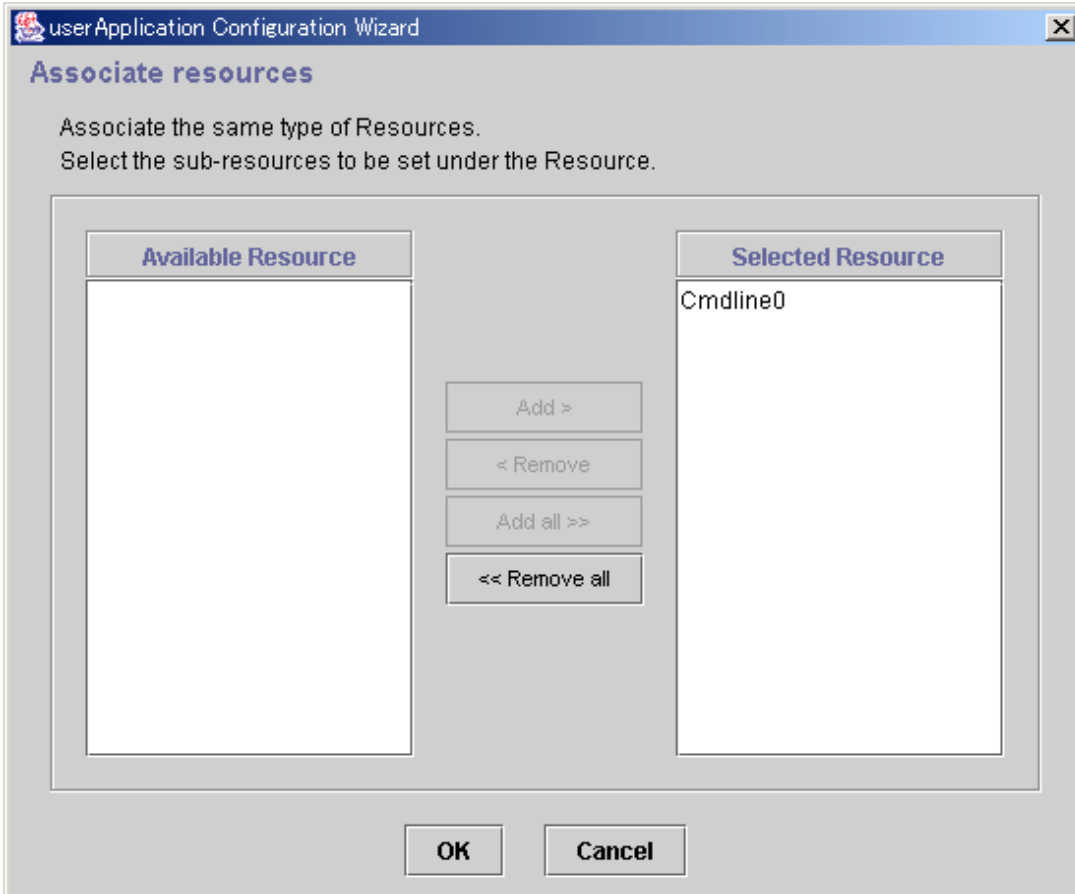
Available Resource

Available resource is referred to as the resource of the same type that can be associated, and satisfies the following conditions.

- The resource is of the same type as the resource that was called.

- The resource is not being used by another cluster application.

Note: Cmdline and process monitoring resources are considered the same type of resources.

Selected Resource

Resources to be set under the current resource being created.

From *Available Resource* select the resource to be configured under the current resource, and then click *Add*. To add all listed resources, click *Add all*. To delete a resource from under the current resource, select the resource to be deleted from the *Selected Resource*, and then click *Remove*. To delete all listed resources, click *Remove all*.

After completing the setup, click *OK*, and return to the "Confirm Registration Information" screen.

Example

In the configuration shown below, the startup priority has been set to resources of the same type.

To build this configuration, build the resources according to the procedure shown below.
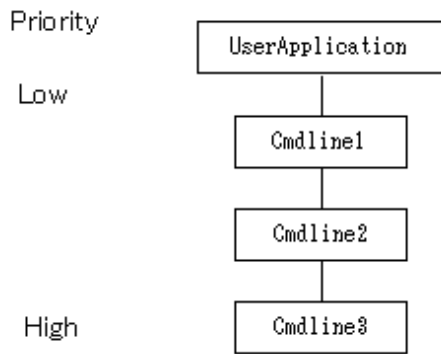
1. Create Cmdline3.

2. Create Cmdline2, and set Cmdline3 as subApplication.

3. Create Cmdline1, and set Cmdline2 as subApplication.

4. Create userApplication, and set Cmdline1 as subApplication.

The procedure for building the following configuration is shown below.
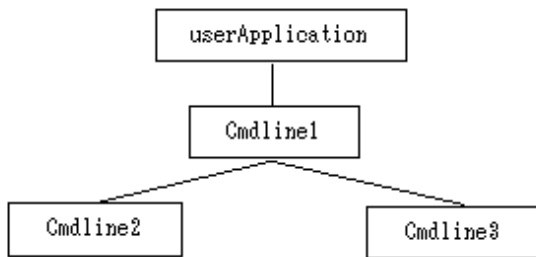
1. Create Cmdline3.

2. Create Cmdline2.

3. Create Cmdline1, and set Cmdline2 and Cmdline3 as subApplication.

4. Create userApplication, and set Cmdline1 as subApplication.

In the configuration shown below, the other resources were not set as subApplication during the creation of resource.
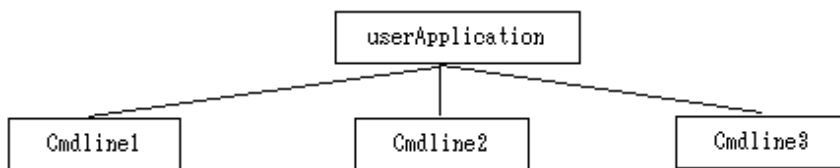
6.6.1.2 Creating Fsystem Resources

Set up an Fsystem resource if you want to mount a file system when userApplication is started.
To control multiple mount points in parent-child relationships, create the file system as one Fsystem resource.
You need to work on the following prerequisites to create an Fsystem resource.

Mount point characteristics

- Mount points are used to mount UFS and other local file systems.

- The file system can be shared as an NFS file system in the network. The file system can also be set up so that it is not shared.

- When a file system is shared in the network, the NFS Lock Failover function becomes enabled.

Notes on using the file lock facility

If you use an NFS server function, set it so that the Lock information for NFS is succeeded.

If a transfer of the Lock information for NFS is not possible, you can get multiple Lock information for the same file. Therefore, the application that uses fcnlt(2) and lockf(3C) cannot execute the exclusive control. And it becomes the factor to cause file destruction.

But the following attention is necessary on using this function. Execute the next attention whether an application uses a lock unidentified.

- On the node operation with NFS servers, do not use the NFS client function. Do not implement NFS mount.

- A file system in which NFS is shared with the operation of NFS servers must be used only from the NFS client. Do not use a file directly from an application on a node on which NFS servers are operating.

- When failover occurs, the following processing is performed. The access from the client is reserved until the NFS lock recovery processing is completed though the state of the resource becomes Online when the start of statd and lockd is completed.

    - Stopping the statd(1M) and lockd(1M) daemon

    - Starting the statd(1M) and lockd(1M) daemon

    - NFS lock recovery processing

- The mount point that sets NFS Lock Failover cannot be set to multiple Fsystem resources.

- If mount point that NFS is shared needs the NFS Lock Failover, please set those mount points in a Fsystem resource including the mount point that set NFS Lock Failover.

Concretely, it is configured as about the following.

```
userApplication0
    Fsystem0
        mountpoint0  (NFS share, NFS Lock Failover)
        mountpoing1  (NFS share)
        mountpoing2  (NFS share)
        mountpoing3  (NFS share)
```

📗 Note

....

It is not configured as about the following.

- Setting was multiple Fsystem

```
userApplication0
    Fsystem0
        mountpoint0  (NFS share, NFS Lock Failover)
        mountpoing1  (NFS share)
    Fsystem1
        mountpoint2  (NFS share, NFS Lock Failover)
        mountpoing3  (NFS share)
```

- Setting was multiple userApplication/Fsystem

```
userApplication0
    Fsystem0
        mountpoint0  (NFS share, NFS Lock Failover)
        mountpoing1  (NFS share)
userApplication1
    Fsystem1
        mountpoint2  (NFS share, NFS Lock Failover)
        mountpoing3  (NFS share)
```

....

Notes on using NFS

Protocols

NFS can use TCP and UDP protocols. After using TCP for failover and failback in sequence, recovery sometimes takes longer than with UDP. Therefore, please use the UDP protocol in PRIMECLUSTER.

Creating 31 or more mount points in Fsystem resources

The minimum Timeout value of the Fsystem resource is determined by the formula " No. of mount points in Fsystem resources x 6 seconds".

The default timeout is 180 seconds. If you set 31 or more mount points in Fsystem resources, it is necessary to change Timeout to the proper value based on "No. of mount points in Fsystem resources x 6 seconds" by selecting the *Attributes* tab in the screen shown in "Checking Fsystem registration information."

6.6.1.2.1 Prerequisites

This section describes operations that must be performed before the file system is set up with the GUI. Although not mentioned in the procedure, editor commands such as those featured by vi(1) are generally used in file editing.

Editing the /etc/vfstab file

Define the mount point in the "/etc/vfstab" file on all nodes where userApplication is configured to use Fsystem.

```
#RMS#/dev/sfdsk/class0001/dsk/volume0001 /dev/sfdsk/class0001/rdsk/volume0001 /disk1 ufs - no -
#RMS#/dev/sfdsk/class0001/dsk/volume0002 /dev/sfdsk/class0001/rdsk/volume0002 /disk2 sfxfs - no -
```

Note

- Lines must begin with "#RMS#".

- You cannot set different volumes on the same mount point.

Preparations for sharing a file system in a network(NFS)

1. Create Ipaddress or Gls resources.

   Since a takeover network is necessary when an NFS service is used as a cluster application, create an Ipaddress or a Gls resource.

   For details on the setup procedure, see "6.6.1.5 Creating Takeover Network Resources" and "6.6.1.4 Creating Gls Resources."

   ## Note

   Set this resource to the same userApplication as the Fsystem resource.

2. Carry out preparations for enabling NFS Lock Failover.

   Select the "Edit global settings in Configuration" command from the menu, and specify the settings for enabling NFS Lock Failover.

   For details on the setup procedure, see "6.6.4 Editing global settings in Configuration."

3. Set up the NFS service.

   3-1) Create a directory for dummy entries.

   Dummy entries are necessary to ensure that the NFS service is started when Solaris is started.

   Execute the command below to create a blank directory.

   ```
   # mkdir /var/opt/SMAWRrms/nfsshare
   ```

   3-2) Edit the /etc/dfs/dfstab file.

   Add the entries shown below to the file in all nodes that make up userApplication. The first entry is a dummy entry, and the subsequent entries are entries of the file system to be shared.

   ```
   share -F nfs -o ro=fuji2 -d "dummy directory for NFS sharing on PRIMECLUSTER" /var/opt/SMAWRrms/
   nfsshare
   ```

```
#RMS# share -F nfs /disk1
#RMS# share -F nfs /disk2
```

📇 Note
. . . .

- Be sure to start each line with "#RMS#" (except for the dummy entry).

- In the dummy entry, set the "o" option to ensure that network clients cannot see the entry.

  ```
  Example) share -F nfs -o ro=hostname_of_local_node
  ```
. . . .

3-3) Start the NFS service.

To start the daemon without rebooting the system, execute the following command:

```
# /etc/init.d/nfs.server start
```

If another file system is used

The disk must be formatted and the file system must be created beforehand. For information on formatting and file system commands, see the "Solaris X Reference Manual Collection." If you plan to use GFS as the file system, see "6.3 Initial GFS Setup" when you build the file system.

6.6.1.2.2 Setup Method

This section describes how to create Fsystem resources.

Figure 6.18 Flow of Fsystem resource creation
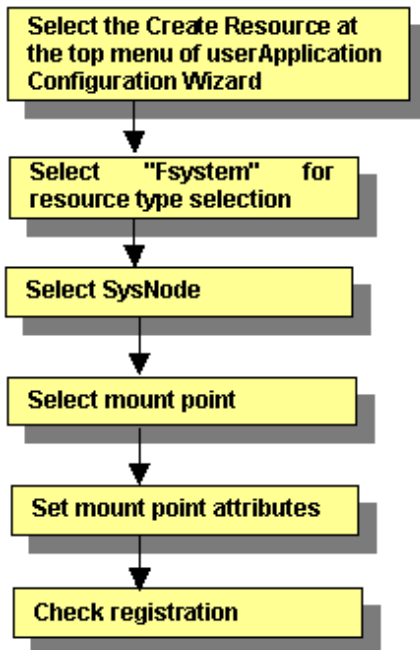
For information on the above operations up to "Select SysNode," see "6.6.1.1 Creating Cmdline Resources." This section describes the operations starting from "Select mount point."

Selecting mount points

Select the mount points.

Figure 6.19 Select mount point
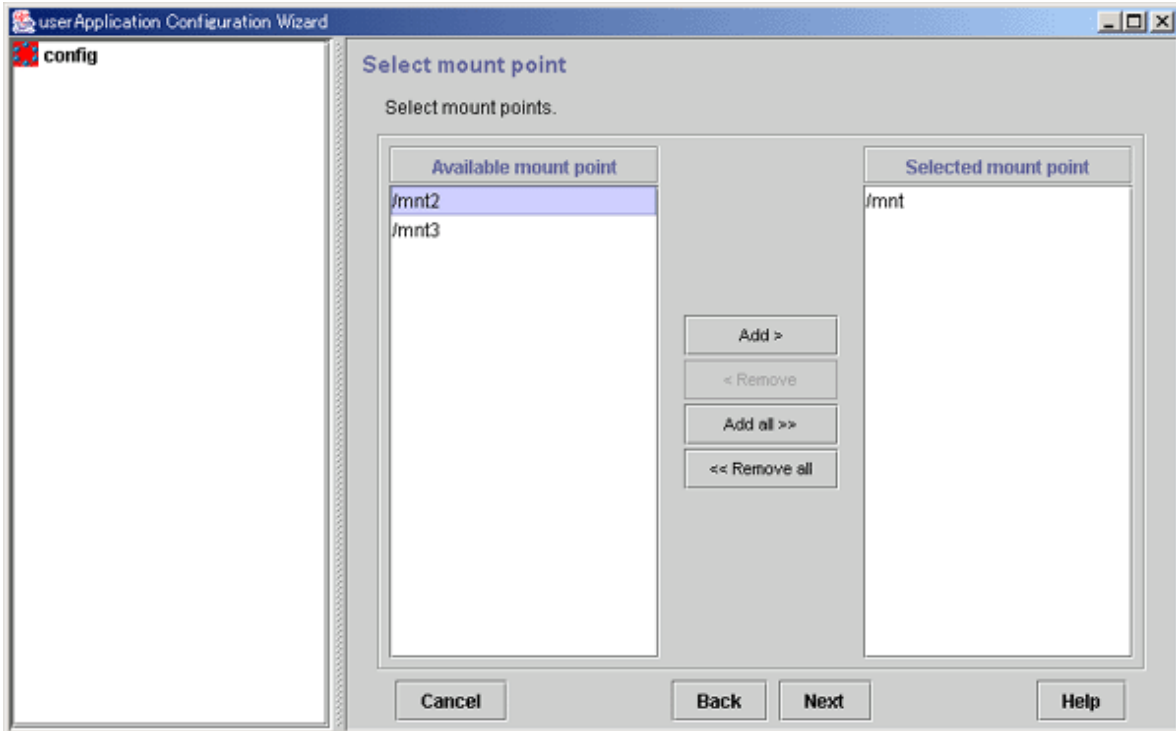
Available mount point

> Names of mount points where the file system can be mounted.
> The mount points that were configured for "/etc/vfstab" (described earlier) are displayed.

Selected mount point

> Names of mount points to be used for file system mounting.

From *Available mount point*, select the mount point at which the file system is to be mounted, and then click *Add*. To add all listed mount points, click *Add all*. To delete a mount point on which the file system is not going to be mounted, select the mount point to be deleted from *Selected mount point*, and then click *Remove*. To delete all listed mount points, click *Remove all*.

After completing the setup, click *Next* to go to "Setting mount point attributes."

Setting mount point attributes

> Set attributes for the mount points.

Figure 6.20 Select attribute
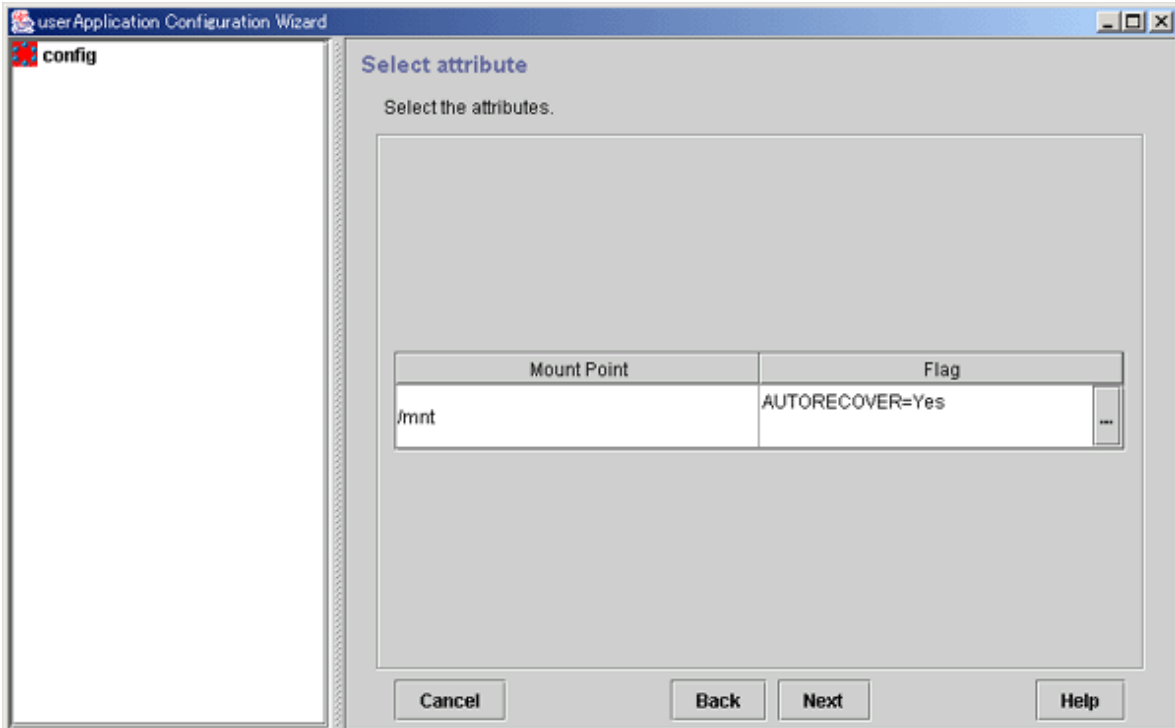
Mount point

The mount name is displayed.

Flag

The attributes that are set for the individual mount points are displayed.

[...] button

Press this button to set attributes for controlling mount point management. When this button is clicked, the Set Mount Button Attributes screen shown below is displayed.

After completing the setup, click *Next* to go to "Fsystem Registration Information Confirmation."

Attribute setup for mount points

Configure the attributes for controlling the mount points.

Figure 6.21 Attribute setup for mount points
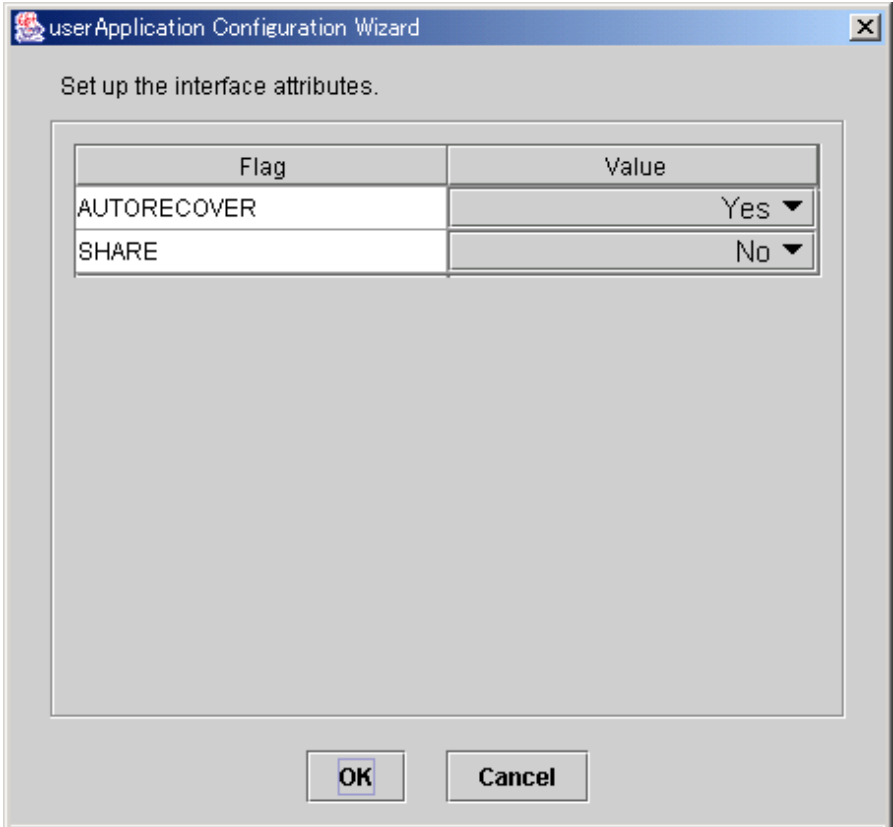
After completing the setup, click *OK* to go to the "Select Attribute" screen.

Flag	Outline
AUTORECOVER	If this flag is set to "Yes" and the specified file system is unmounted, RMS automatically attempts to remount the specified file system when it is unmounted. If this attempt fails, Fault processing is initiated. The default value is "Yes."
SHARE	If this flag is set to Yes, a mounted directory is shared (NFS). In such a case, make a specific entry to "/etc/dfs/dfstab". See "Preparations for sharing a file system in a network(NFS)" in "6.6.1.2.1 Prerequisites." The default value is "No."
NFSLOCKFAILOVER	If this flag is set to "Yes," Lock information for NFS Lock Failover is stored for this mount point. To use NFS Lock Failover, you must set this flag to "Yes" for one of the mount points. If you set this attribute to "Yes," you must also set SHARE to "Yes" at the same time. The default value is "No." For each Fsystem resource, this flag can be set for only one mount point. Even if userApplication has multiple Fsystem resources and NFS Lock Failover is to be used, set this attribute to "Yes" for only one mount point in userApplication.

Checking Fsystem registration information

Check the Fsystem registration information. You can also select the *Attributes* tab, and set resource attributes by switching the screen.

For information on the resource attributes, see "6.6.5 Attributes."

Figure 6.22 Confirmation of registration information
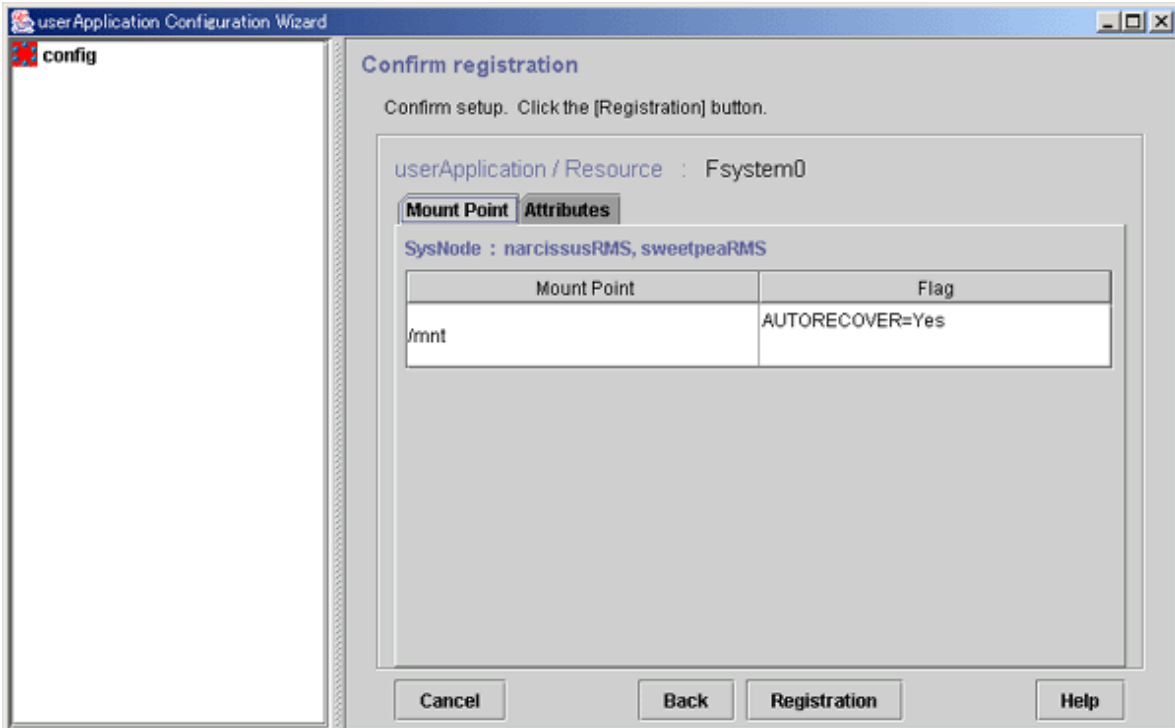
The attributes that are displayed in the "Flag" column are only those Resource attributes for which a value has been set. For example, if "No" is set for an attribute that takes a "Yes" or "No" setting, such as AUTORECOVER, that attribute is not displayed in the "Flag" column.
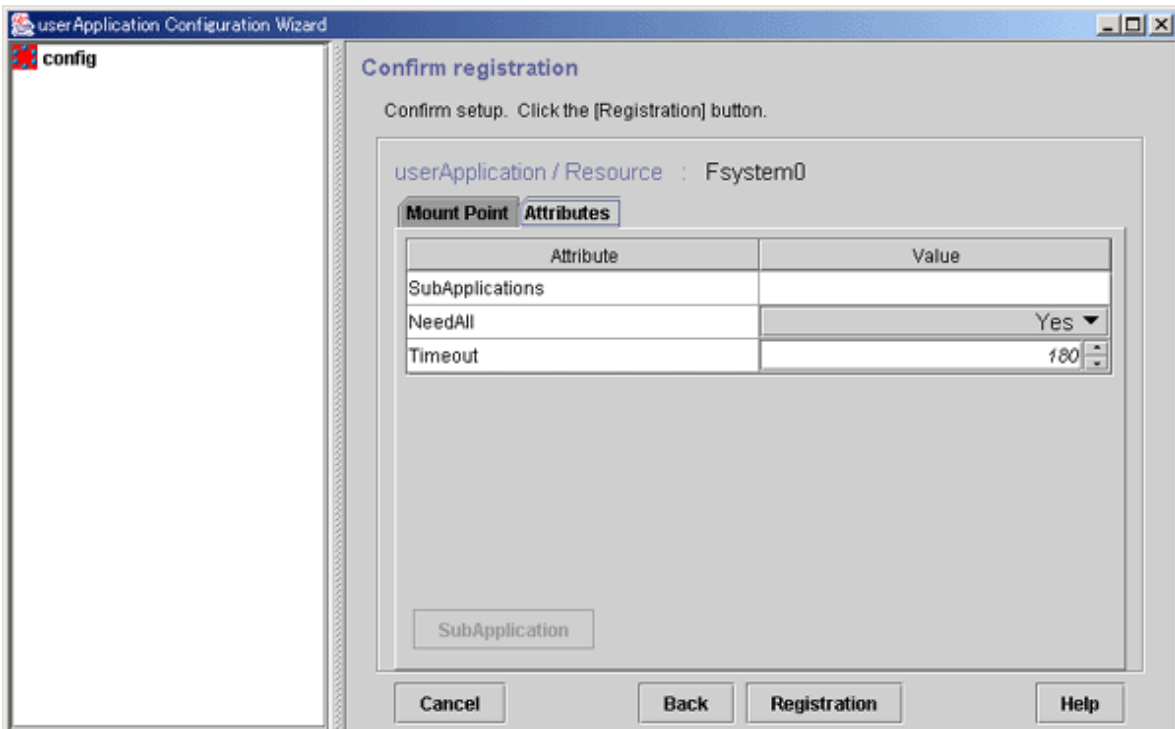
SubApplication button

This button is used for specifying other resources that the current resource depends on. For an Fsystem resource, this button is disabled. For multiple directories in parent-child relationships, configure the directories in one resource.

Check the registration contents, and then click *Registration*.

6.6.1.3 Creating Gds Resources

Set up resources for the disk classes that are defined in Global Disk Services (GDS).

The disk classes must be created before this screen is used.

> 📔 **Note**
> ....
> If a disk class is registered to a Gds resource when equivalent volume copy is being executed, the equivalent copy operation is canceled. Therefore, the data becomes invalid in some of the slices. However, recovery is not necessary because equivalent copy is executed automatically when the volume is started. If you want to recover the slice state before the volume is started, complete Gds resource setup and then start equivalent copy. For instructions on starting equivalent copy, see "5.3.4 Copying Operation" or "D.10 sdxcopy - Synchronization copying operation" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide.*"
> ....

For instructions on creating disk classes, see "6.2.2.2 Setting Up Shared Disks."

Figure 6.23 Flow of Gds resource creation
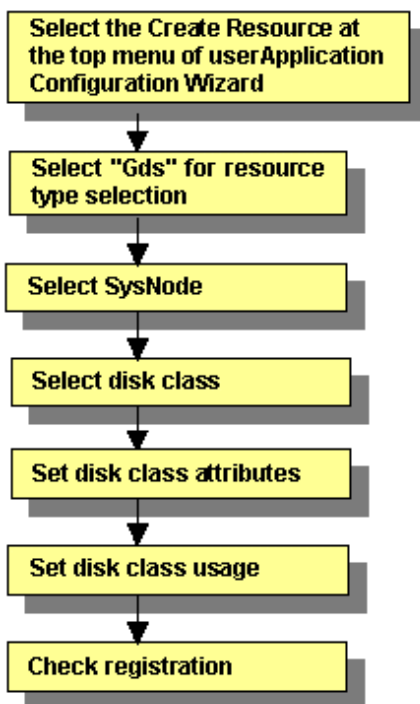
For information on the above operations up to "Select SysNode," see "6.6.1.1 Creating Cmdline Resources." This section describes the operations starting from "Select disk class."

Selecting the disk class

Select the disk class.

Figure 6.24 Select disk class
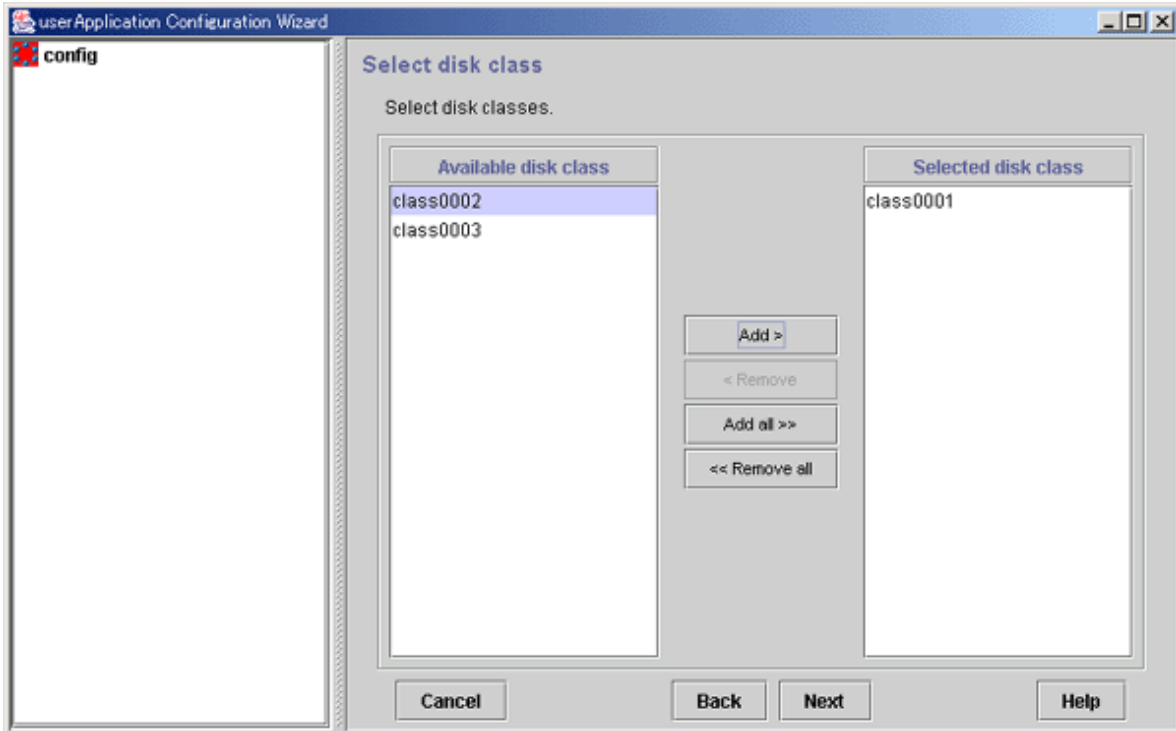
Available Disk Class

> Names of disk classes that can be set.
> The disk classes (shared disks) that are defined in GDS are displayed.

Selected Disk Class

> Names of disk classes to be set.

Select disk classes from *Available disk class*, and then click *Add*. To add all listed disk classes, click *Add all*. To delete a disk class to be set, select the disk class to be deleted from *Selected disk class*, and then click *Remove*. To delete all listed disk classes, click *Remove all*.

After completing the setup, click *Next* to go to the "Set Disk Class Attribute" screen.

Disk class attribute setup

> Set the disk class attribute.

Figure 6.25 Set up disk class attribute
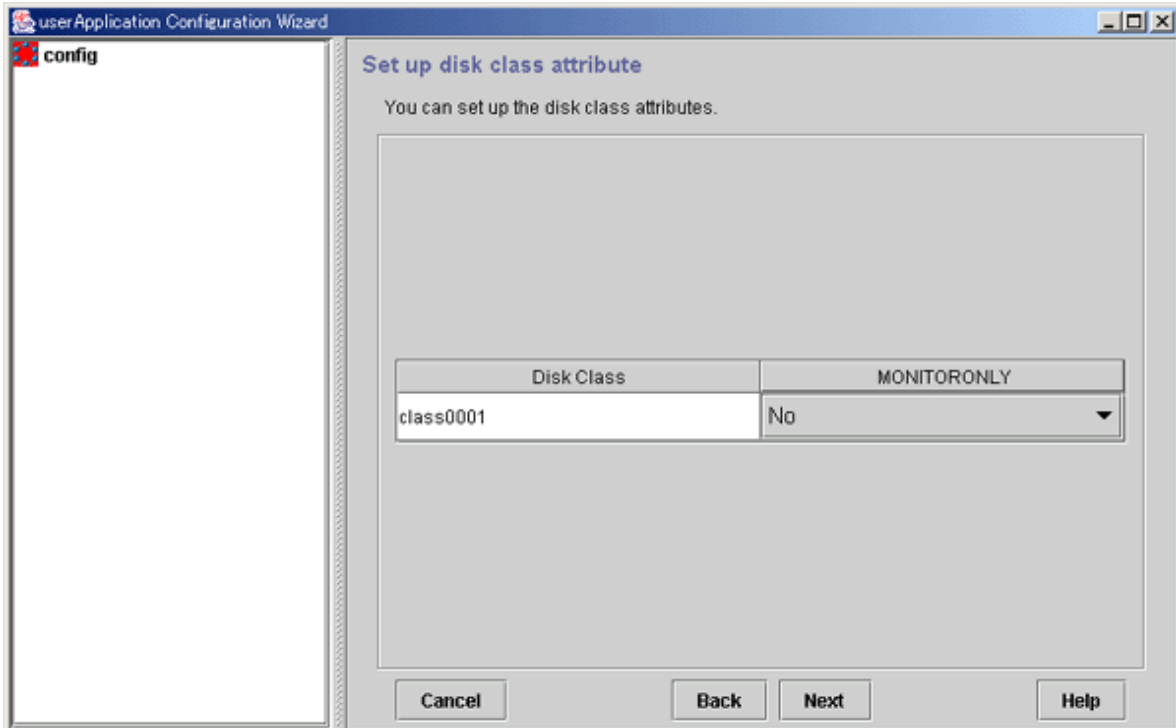
Disk Class

The disk classes (common disks) that were defined in GDS (described earlier) are displayed.

Flag	Outline
MONITORONLY	This flag determines whether disk class failures are to be reported to the userApplication. If "Yes" is set and a disk class failure occurs, the disk class is switched to faulted state but the Gds resources remain online, and userApplication failover does not occur.
	Carry out the preliminary design, including the higher applications" before determining the disk classes for which this flag is to be set to "Yes." The default value is "No."

The MONITORONLY attribute must be set to "No" for at least one disk class.

This will prevent userApplication failover in the event of all disk class failures and RMS' notifying the upper applications of the failures.

After completing the setup, click *Next* to go to the "Set up disk class use" screen.

Setting the disk class use

Configure the use of the disk class.

Figure 6.26 Set up disk class use
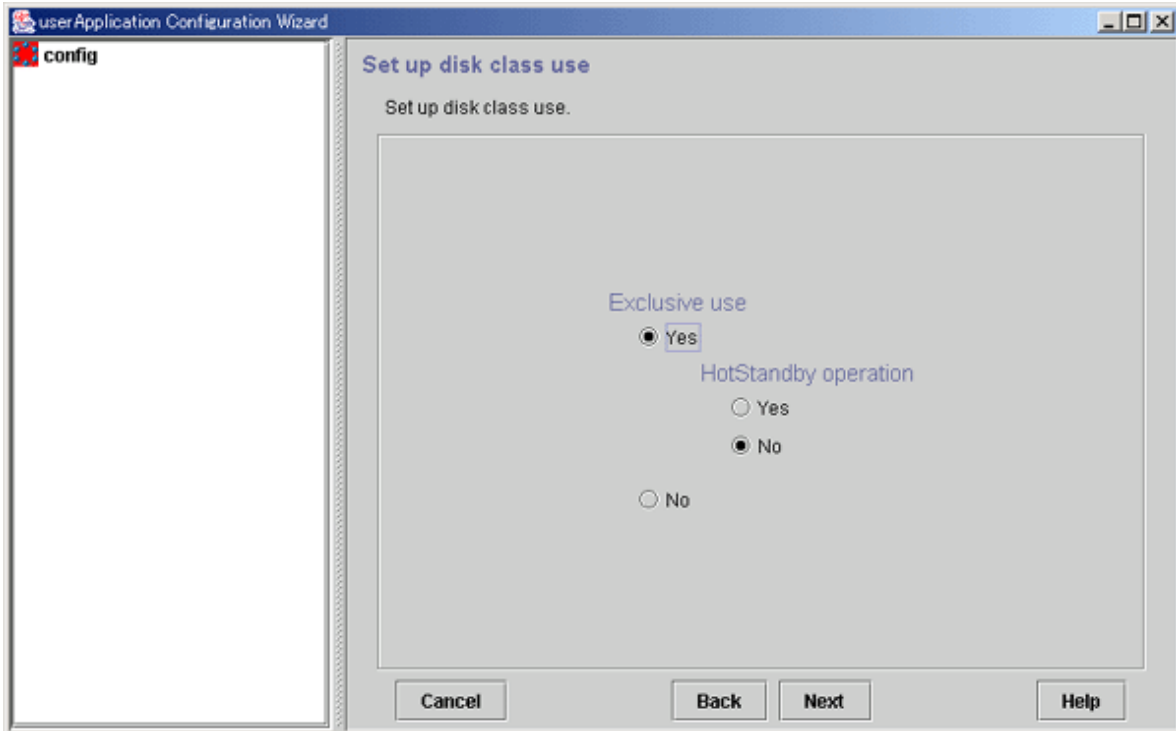
Exclusive use

Select "*Yes*" or "*No*" for *Exclusive use*. If you select "*Yes*", also select "*Yes*" or "*No*" for *HotStandby operation*.

Specify the options according to the use of the shared disk.
After completing the setup, click *Next* to go to the "Confirm Registration Information" screen.

Shared disk uses

Table 6.2 Shared disk uses and setting methods

Use	Exclusive use	HotStandby operation
Switched disk	Yes	No
Shared disk for simultaneous access	Yes	Yes
Simultaneous shared disk	No	-

Shared disks have the following features for each use:

- Switched disk

  Only the OPERATING node can use the shared disk.
  The non-OPERATING nodes cannot access the volume.

- Shared disk for simultaneous access

  All nodes can use the shared disk simultaneously.
  Select this use if applications that require disk access from a node other than the OPERATING node are set inside the same userApplication.
  The exclusive control for maintaining data integrity in the shared disk must be handled by the applications, and not by PRIMECLUSTER.

- Simultaneous shared disk

  Select this use if multiple userApplications share the disk classes. An example is when Oracle real application clusters are used.
  The exclusive control for maintaining data integrity in the shared disk must be handled by the applications, and not by PRIMECLUSTER.

> **Note**
>
> - During operation of a cluster application, try not to have applications or user processes that are not related to the cluster applications use the shared disks. If this occurs, the state transition might fail.
>
> - Do not create a home directory for user accounts on a switchover disk. The reasons are as follows:
>
>   - The file system on the switching disk is mounted only on the OPERATING node of the cluster application.
>
>   - PRIMECLUSTER forcibly terminates process that uses the file system on the switching disk by executing the "kill(1M)" command to unmount the file system in the event of cluster application failover.
>
>   - If the file system cannot be unmounted properly because of a failure of forced termination of process, PRIMECLUSTER might forcibly stop the node to prevent the file system data on the shared disk from being destroyed by double mounting.

Checking registration information for Gds resources

Check the registration information for the Gds resources. You can also set resource attributes by selecting the *Attributes* tab and switching the screen.

For information on the resource attributes, see "6.6.5 Attributes."

Figure 6.27 Confirm registration
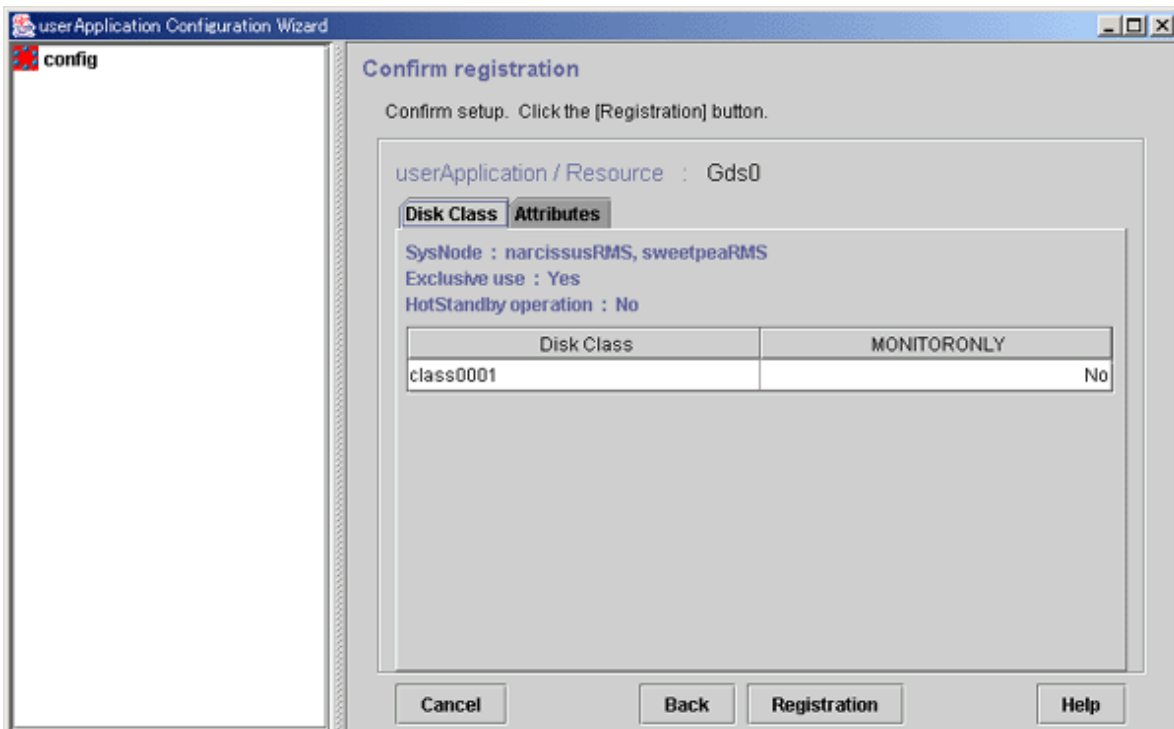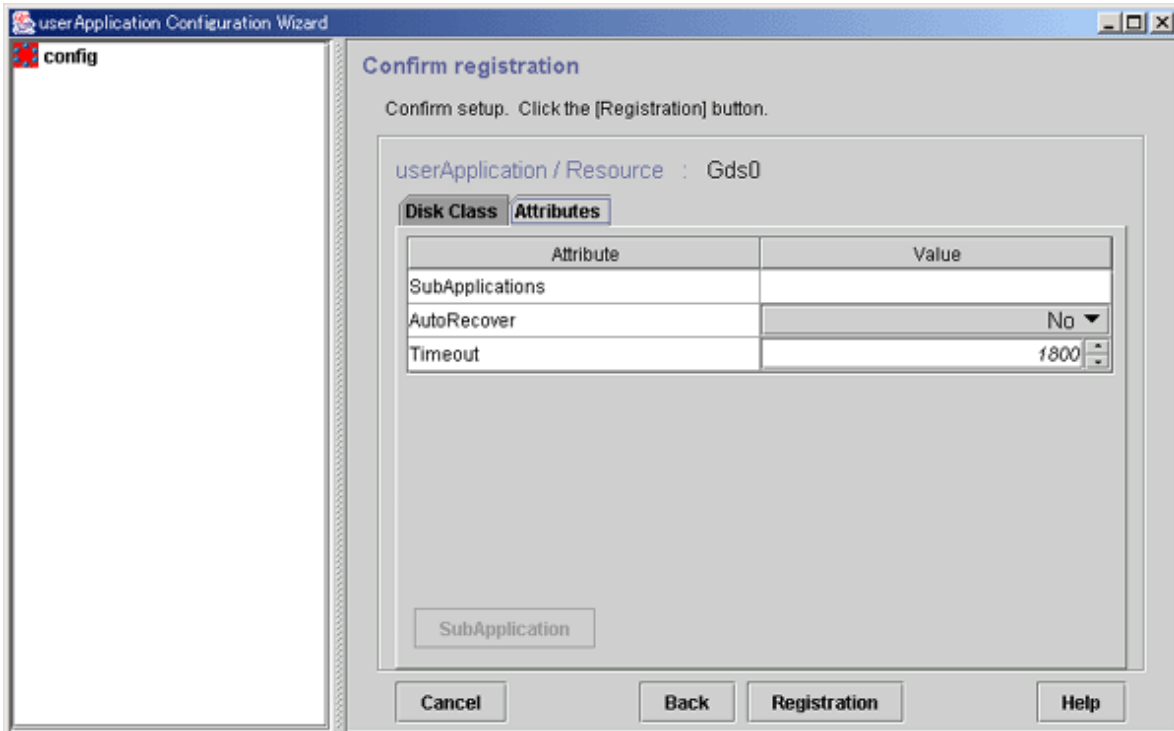
SubApplication button

Use this button to associate a previously created Gds resource under the current Gds resource. This button can be selected only if there are resources of the same type that can be associated. For setting instructions, see "Resource association" in "6.6.1.1 Creating Cmdline Resources."

Check the registration information, and then click *Registration*.

6.6.1.4 Creating Gls Resources

Configure the takeover IP addresses that are defined by the redundant line control function of Global Link Services (GLS).

Note

For details on using the multipath function of Global Link Services (GLS), see "Chapter 7 Administration on a Cluster System" in "*PRIMECLUSTER Global Link Services Configuration and Administration Guide: Multipath Function.*"

Figure 6.28 Flow of Gls (redundant line control function) resource creation
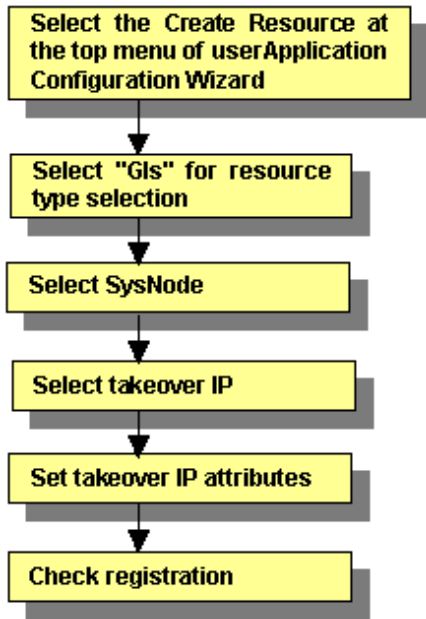

For information on the above operations up to "Select SysNode," see "6.6.1.1 Creating Cmdline Resources." This section describes the operations from "Select takeover IP."

Selecting the takeover IP address

Select the takeover IP address.

Figure 6.29 Takeover IP address selection
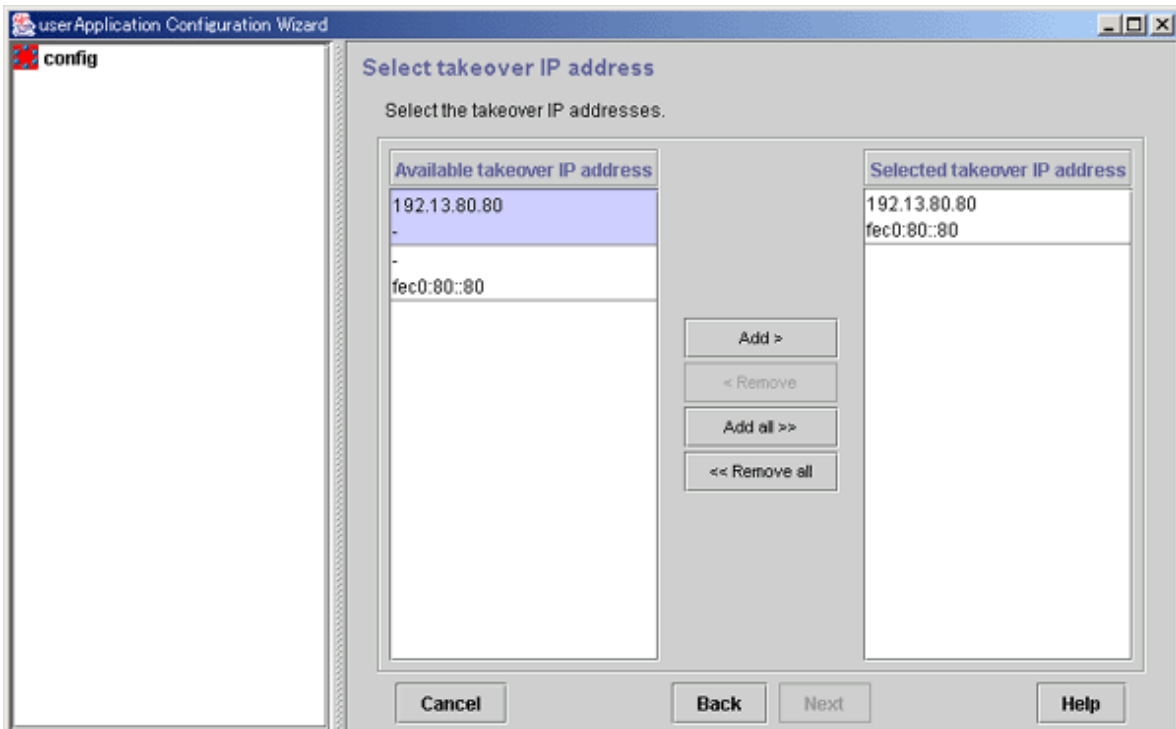
Available takeover IP address

IP addresses that can be taken over.

Selected takeover IP address

Takeover IP addresses.

From *Available takeover IP address*, select takeover IP addresses, and then click *Add*. To add all listed takeover IP addresses, click *Add all*. To delete a takeover IP address, select the takeover IP address to be deleted from *Selected takeover IP address*, and then click *Remove*. To delete all listed takeover IP address, click *Remove all*.

After completing the setup, click *Next* to go to the "Set Takeover IP Address Attribute" screen.

Setting attributes of the takeover IP address

Set the attributes of the takeover IP address.

Figure 6.30 Attribute setup for the takeover IP address
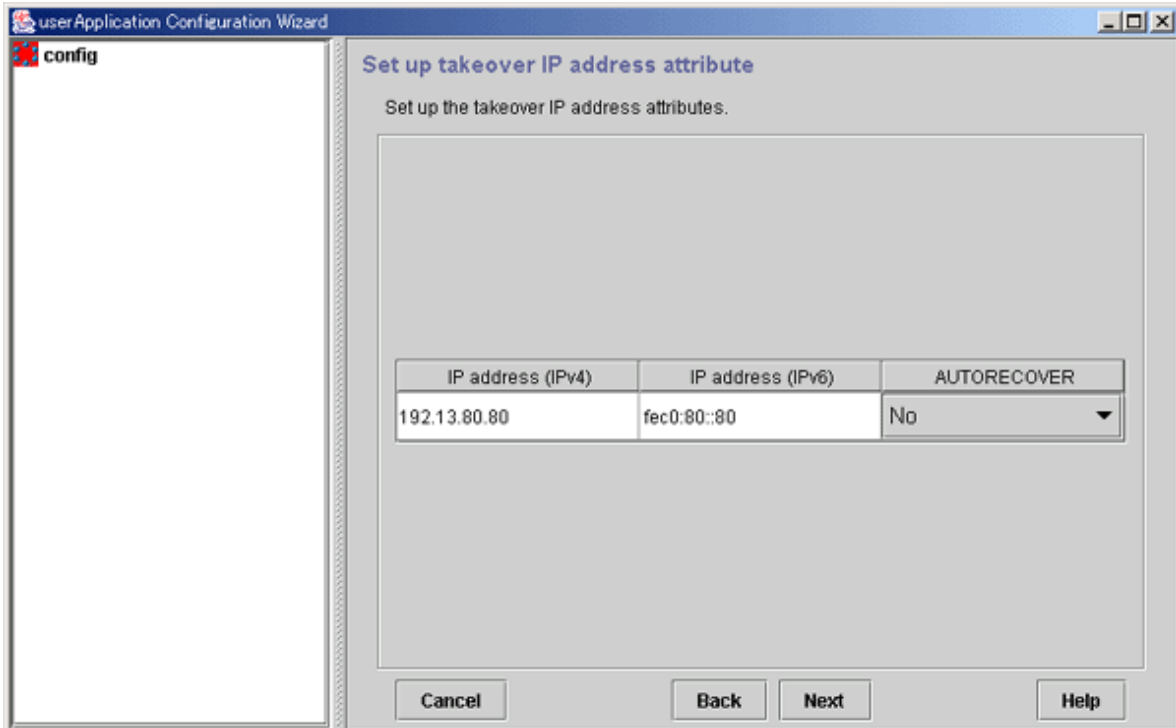
After completing the setup, click *Next* to go to "Checking Gls Resource Registration Information."

Flag	Outline
AUTORECOVER	If the flag is set to "Yes," RMS attempts to restore the faulted resource for a given amount of time to prevent userApplication being switched to the other host. The default timeout is 60 seconds. You can change it by tuning the resource attribute Timeout. If the faulted resource does not recover within a specified time, userApplication is moved to the other host. If the flag is set to "No", RMS switches userApplication to the other host on detecting the faulted resource.

Checking registration information for Gls resources

Check the registration information for the Gls resources. You can also set resource attributes by selecting the *Attributes* tab and switching the screen.

For information on the resource attributes, see "6.6.5 Attributes."

Gls resource names are displayed as GlsX (X is a number, such as 0 or 1).

To register a Gls resource to a cluster application, select this resource name from *Available Resources* on the Select Resource screen.

For details, see "6.6.2 Creating Cluster Applications."

After registering the Gls resources to the cluster application, be sure to restart the system.

Then start RMS, look at the RMS tree, and check that the Gls resources are displayed correctly. For details, see "7.1.3.1 RMS Tree."

Figure 6.31 Confirmation of registration information
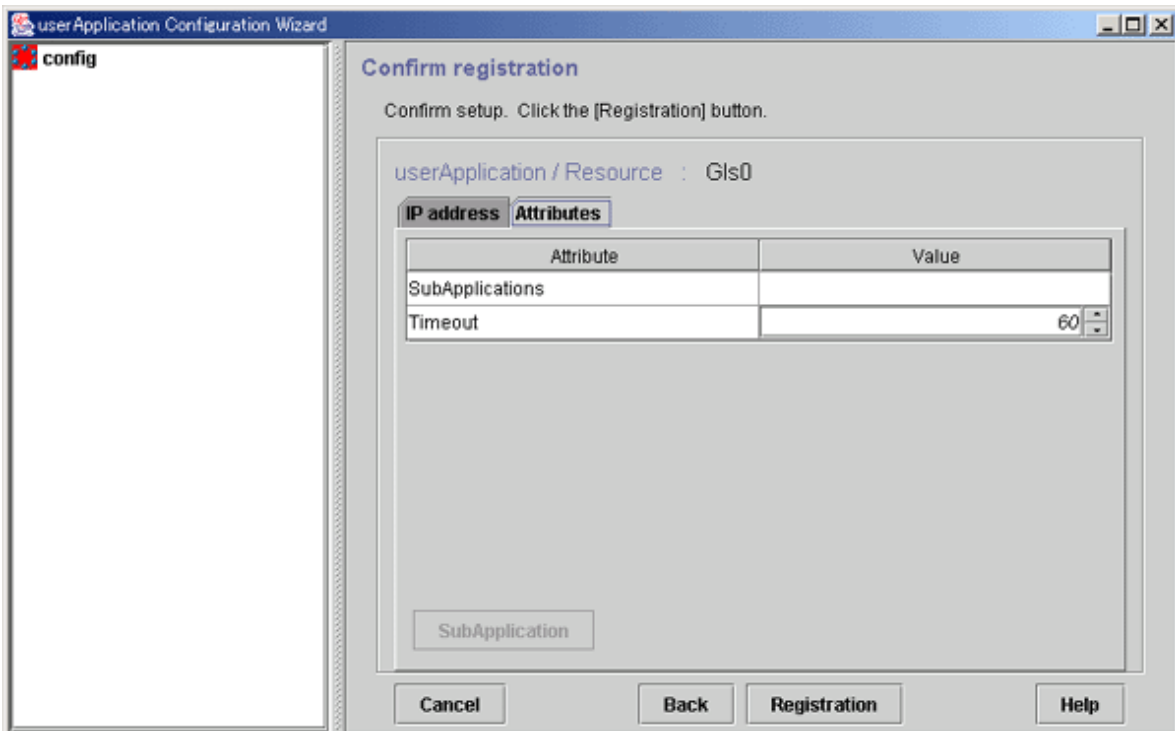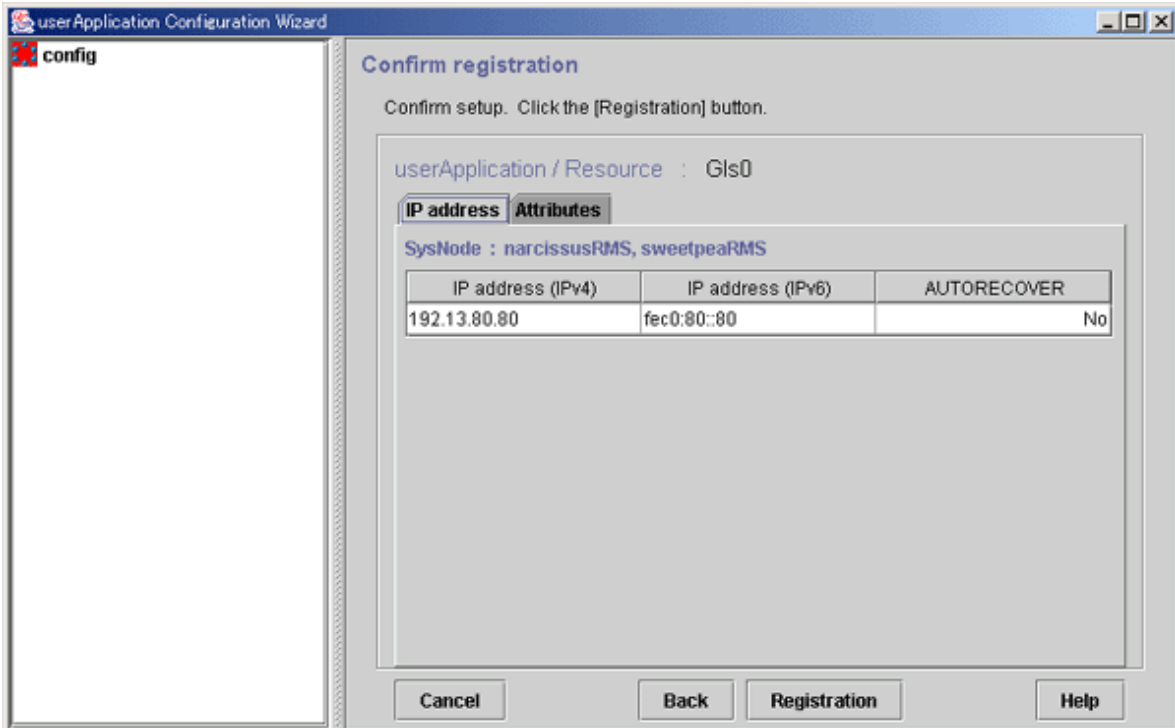
SubApplication button

Use this button to associate a previously created Gls resource under the current Gls resource. This button can be selected only if there are resources of the same type that can be associated. For setting instructions, see "Resource association" in "6.6.1.1 Creating Cmdline Resources."

After checking the registration information, click *Registration*.

6.6.1.5 Creating Takeover Network Resources

Normally to use a takeover network, use either "Gls" (Global Link Services) or "Ipaddress" as the resource type. If the availability of the takeover network is needed, use "Gls."

A takeover network must be set up if you are building a system like a client/server system that communicates with cluster services that operate in the cluster system.

Takeover networks allow communication to continue with the same network name from outside the cluster even if a cluster application that operates in the cluster system undergoes failover.

Takeover network types

The takeover network types are IP address takeover, MAC address takeover, and node name takeover.

- IP address takeover

    When switchover takes place, the defined IP address is taken over to the OPERATING node.

    This is the basic function of a takeover network.

- MAC address takeover

    The MAC address is also taken over to the OPERATING node.

- Node name takeover

    The node name* is also taken over to the OPERATING node.

    Use this type if a program that operates as a cluster application in the cluster node recognizes node names.

    * This is the same value as the host name that is obtained when uname -n is executed.

Files that are edited automatically by the GUI

If the GUI is used to execute the setup, certain files are edited as shown below.

The files differ according to the takeover network type.

```
# Start of lines added by FJSVwvucw - DO NOT DELETE OR CHANGE THIS LINE
```

```
# Mon Aug 05 21:01:43 JST 2002
```

<Specify information that is dependent on that particular file. For example, for /etc/inet/hosts, the information becomes as follows:>

```
192.168.246.100 Ipaddress01
```

```
# End of lines added by FJSVwvucw - DO NOT DELETE OR CHANGE THIS LINE
```

- IP address takeover

    /etc/inet/hosts

    /usr/opt/reliant/etc/hvipalias

- Node name takeover

    /etc/inet/hosts

    /usr/opt/reliant/etc/hvipalias

    /etc/nodename

- MAC address takeover

    Lines are not created automatically. The lines must be specified beforehand.

Note

- In node name takeover, the same node name is set to both the OPERATING node and the STANDBY node.

- 172 -

- All cluster nodes must have at least one network interface card so the card can be used by IPv4.

- IPv6 is not supported. If you want to use IPv6, use Gls.

- Although a PRIMECLUSTER "takeover network" and the "IP address takeover" function of GLS can be set up in the same cluster system, do not use them on the same interface. If the two are used on the same interface, connection using takeover IP addresses will be disenabled.
  For example, if you select hme1 as the interface to be used when you set up the PRIMECLUSTER "takeover network," do not set hme1 in the GLS environment settings (do not specify hme1 in the -t option of the "hanetconfig create" command).

6.6.1.5.1 Prerequisites

This section describes the prerequisites for MAC address takeover.
MAC address takeover is enabled concurrently with takeover of an IP address. The MAC address takeover cannot be used by itself. To use MAC address takeover, you need to configure the following three files on each cluster node.

- /etc/opt/FSUNnet/mactool/macaddr.conf

- /usr/opt/reliant/etc/hvipalias

- /etc/inet/hosts

Although not mentioned in the procedure descriptions, the commands of a text editor such as vi(1) are generally used for file editing.

Using MAC address takeover

1. Edit the "/etc/opt/FSUNnet/mactool/macaddr.conf" file.

   On all cluster nodes, add the takeover MAC address entry as shown below.

   The format is "NIC MAC_address cluster."

   ```
   hme0 02:11:22:33:44:55 cluster
   ```

   For the MAC address, specify the value to be used in the Netcompo FNA-LAN.

2. Edit the "/usr/opt/reliant/etc/hvipalias" file.

   On all cluster nodes, add entries for the target nodes of MAC address takeover as shown below.

   The format is "CF_node_name takeover_network_name NIC Netmask MAC_address."

   ```
   fuji2 Ipaddress01 hme0 0xffffff00 02:11:22:33:44:55
   ```

   ```
   fuji3 Ipaddress01 hme0 0xffffff00 02:11:22:33:44:55
   ```

   Add all cluster nodes that will be using MAC address takeover. The same information must be set in all cluster nodes.

   To view more detailed information for the "/usr/opt/reliant/etc/hvipalias" file, see the comments in /usr/opt/reliant/etc/hvipalias.

3. Edit the "/etc/inet/hosts" file.

   On all cluster nodes, add the takeover network entry as shown below.

   ```
   192.168.246.100 Ipaddress01
   ```

6.6.1.5.2 Setup Method

This section describes how to create takeover network resources.

Figure 6.32 Flow of takeover network resource creation
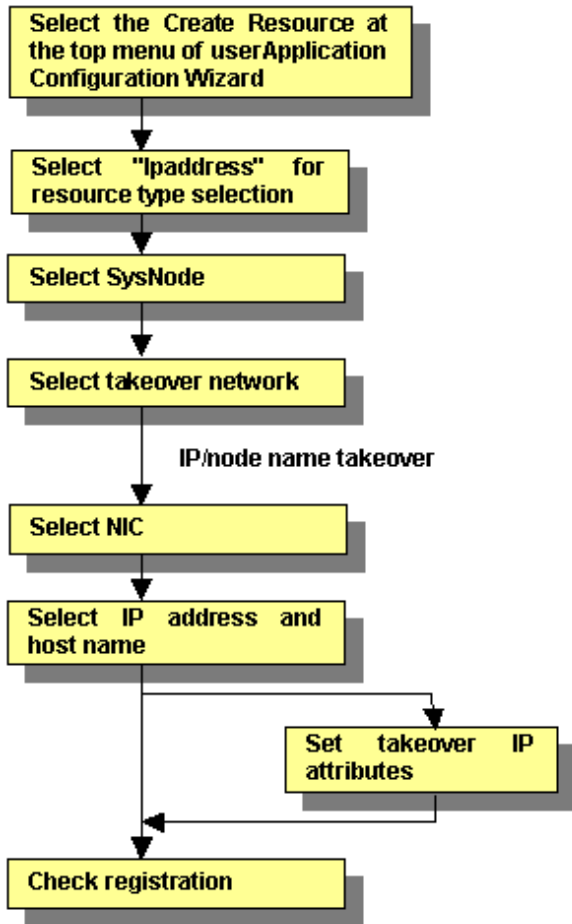
For information on the above operations up to "Select SysNode," see "6.6.1.1 Creating Cmdline Resources." This section describes the operations starting from "Select takeover network."

Selecting a takeover network

Select the takeover network type.

Figure 6.33 Network type selection
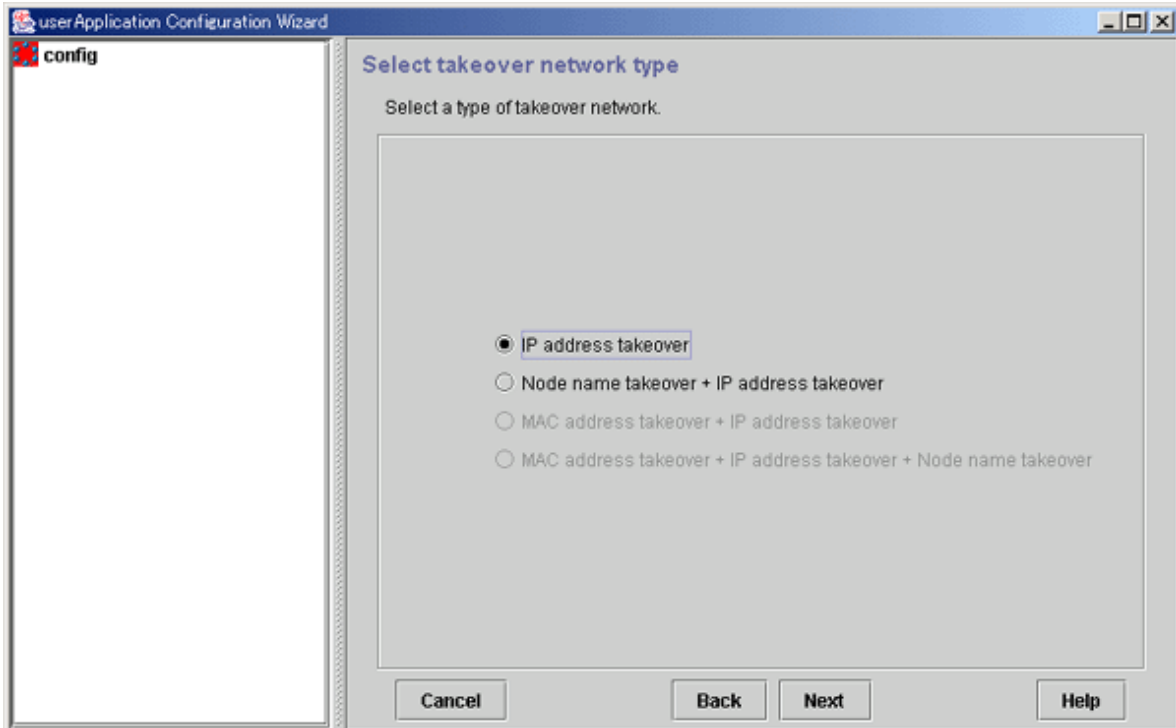
IP address takeover

> Select this item to enable IP address takeover.

Node name takeover + IP address takeover

> Select this item to enable node name takeover and IP address takeover.
> If node name takeover has already been set, you are not allowed to set twice.

In this version, "MAC address takeover + IP address takeover" and "MAC address takeover + node name takeover + IP address takeover" cannot be selected.

To set MAC address takeover, perform the operations described in "6.6.1.5.1 Prerequisites," and then select "IP address takeover" or "node name takeover + IP address takeover" and create the takeover network resources.

You can configure a network interface for each SysNode. The takeover network settings are enabled for one network interface:

- Multiple IP address takeover settings are enabled for each network interface.

- Only one node name takeover setting is enabled in a cluster system.

After completing the setup, click *Next* to go to the "Select Interface" screen.

🛈 **Note**
. . . .

- Do not configure MAC address takeover for a network interface that uses Web-Based Admin View.

- To activate a network interface card (NIC) when the system is started, you must create the /etc/hostname.*NIC-name* file. For instructions on creating and setting up the file, see "Solaris X Reference Manual Collection."

- Node name takeover changes the uname of the system. To set node name takeover for userApplication, you need to restart all nodes configuring userApplication.
  If a node name takeover resource has been deleted, check whether /etc/nodename has been properly changed in all nodes, and then restart all nodes. If /etc/nodename has not been properly changed, change the node name by executing the "setuname(1M)" command, and then restart the nodes.
  Also check whether /etc/nodename is set properly when system conversion or software upgrade is implemented. If /etc/nodename is not set properly, change the node name by executing the "setuname(1M)" command, and then restart the nodes.

. . . .

Selecting an interface

Select a network interface card (NIC).
The NIC cards registered to Cluster Resource Manager are displayed.

Figure 6.34 Select interface
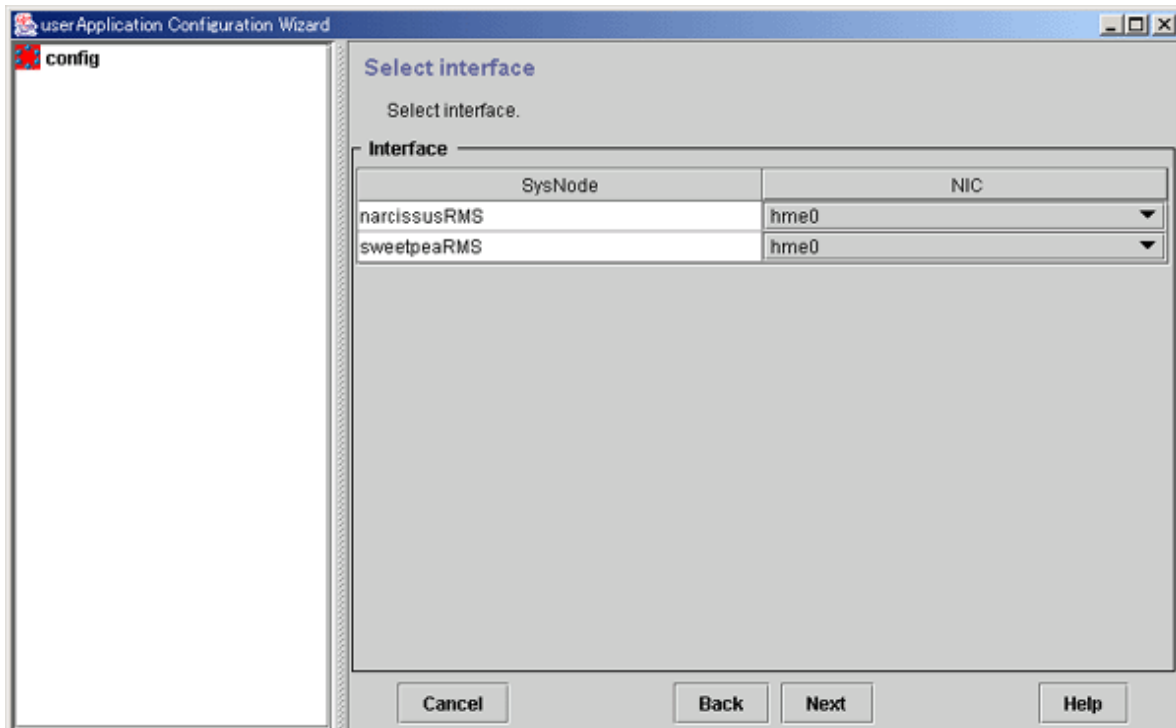
Interface

Select the network interface to be used in each SysNode.

After completing the setup, click *Next* to go to the "Select IP address and host name" screen.

Selecting or creating an IP address or host name

Select or create the takeover IP address or host name.

Figure 6.35 Select IP address and host name
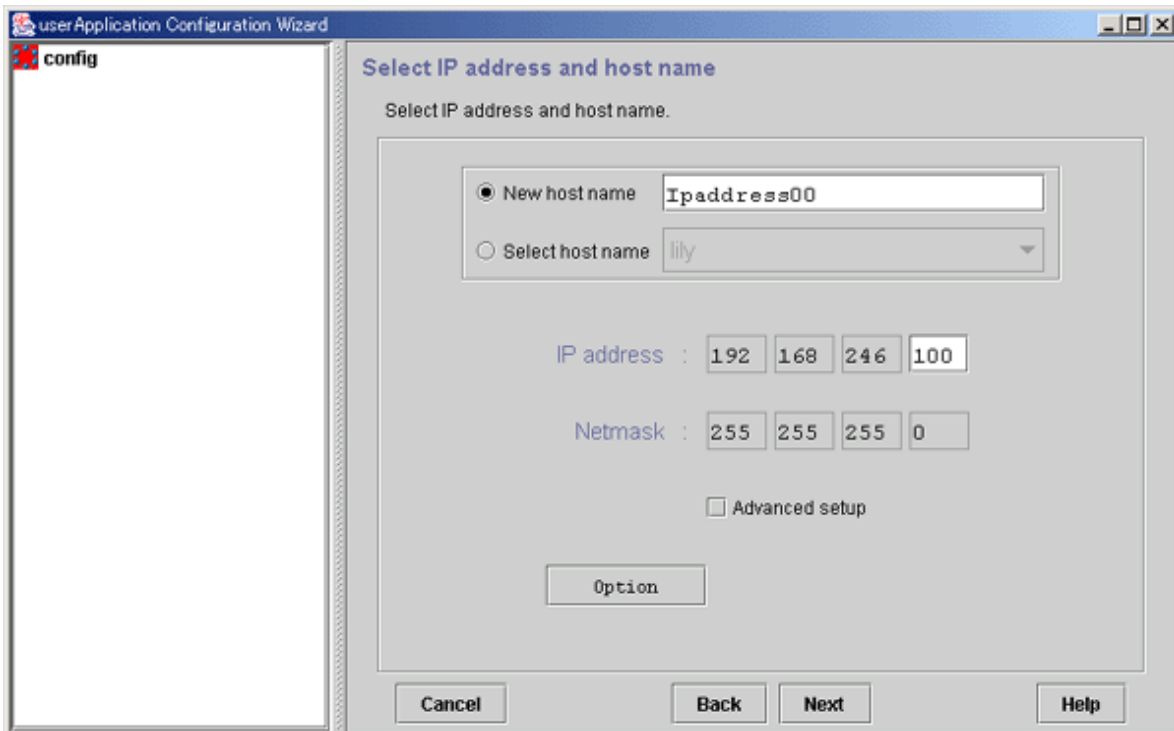
New host name

You can set a new takeover IP address and a takeover node name. The setup information is added to the "/etc/inet/hosts" or "/usr/opt/reliant/etc/hvipalias" file on all nodes configuring the cluster system.

Specify a character string of up to 14 characters that begins with an alphabet letter and consists of only alphanumeric characters.

Select host name

You can select the IP address or node name from information that has been set. If IP addresses or node names were previously configured to /etc/inet/hosts and /usr/opt/reliant/etc/hvipalias, select the IP address or node name from those settings.

IP address

Enter the takeover IP address.
Enter 0 to 255 numbers in the address input area.

Netmask

Enter the net mask value.
Enter 0 to 255 numbers in the address input area.

Advanced setup

Set up a masked net mask or change an IP address.

***Option* button**

Press this button to configure attributes for the takeover IP address. For more information, see "Setting up Takeover IP Address Attributes" below.

🔶 **Note**

. . . .

If /usr/opt/reliant/etc/hvipalias and /etc/inet/hosts files have been modified to accommodate the entries for the takeover IP address and node name takeover in advance (as described in the Prerequisites section), message 0840 appears when the Next button is clicked. This message confirms whether the existing values are to be used without modification. To use the values, select *Yes*.

If you select *No*, the existing settings in the file are deleted and the GUI re-creates the information. In this case, takeover network information can be deleted automatically from /usr/opt/reliant/etc/hvipalias and /etc/inet/hosts when the Ipaddress resource is deleted.

. . . .

Setting up Takeover IP Address Attributes

Configure the interface attributes for the takeover IP address.

Figure 6.36 Attribute setup for takeover IP address

Flag	Outline
BASE	If this flag is set to "base," the specified address is assigned to the physical interface. If this flag is set to "virtual," the specified address is assigned to the virtual interface. The default value is "virtual." **[Note]** To specify MAC address takeover in the interface attributes of the takeover IP address, change this setting from "virtual" to "base."
AUTORECOVER	If this flag is set to "Yes", RMS automatically attempts to create an interface if the specified address becomes disabled. A failure of this attempt might trigger Fault processing. The default value is "Yes."
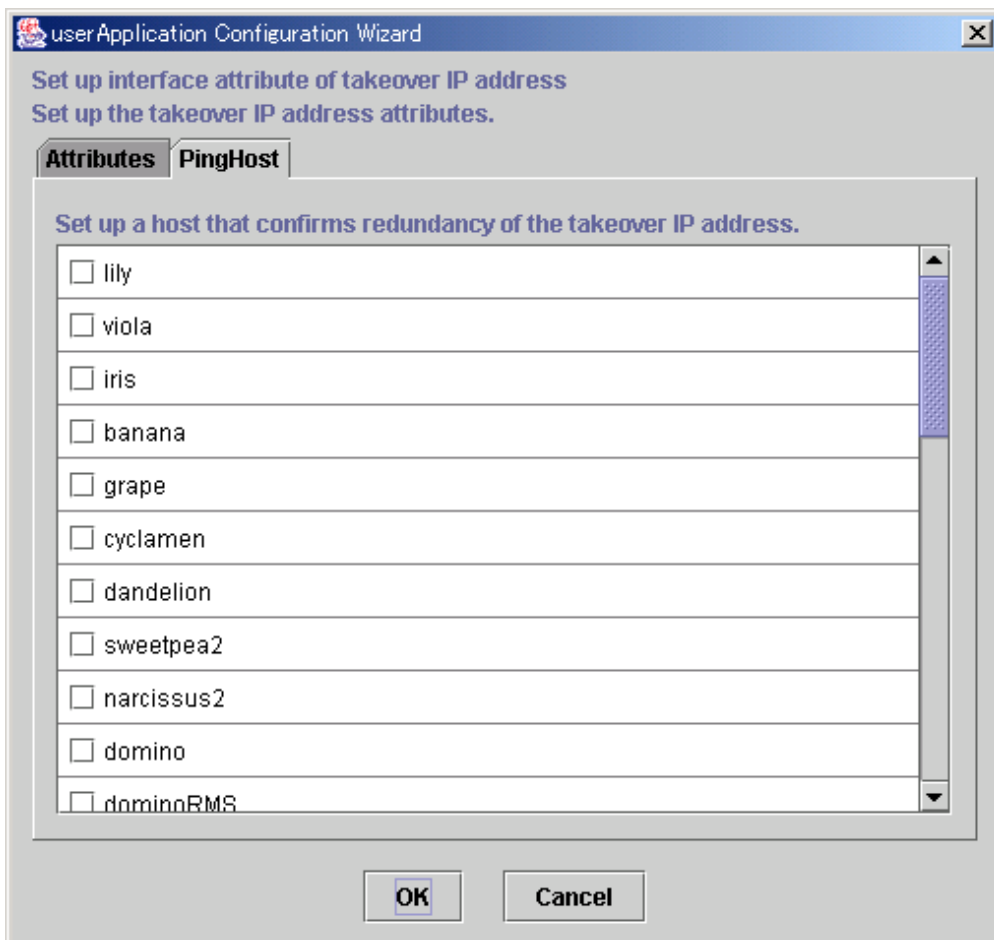
After the setup is completed, click *OK* to return to the "*Select IP address and host name*" screen.

The availability of the takeover IP address is verified by executing the "ping" command.

We recommend that you specify two or more hosts, which are not used for the cluster system, and are in the same network segment that does not use a hub or router. This prevents adverse effects from hub and router failures.

The host information to be used by PingHost must be in /etc/inet/hosts.

Checking registration information for the takeover network

Check the registration information for the takeover network. You can also set resource attributes by selecting the *Attributes* tab and switching the screen.

For information on the resource attributes, see "6.6.5 Attributes."

Figure 6.37 Confirm registration
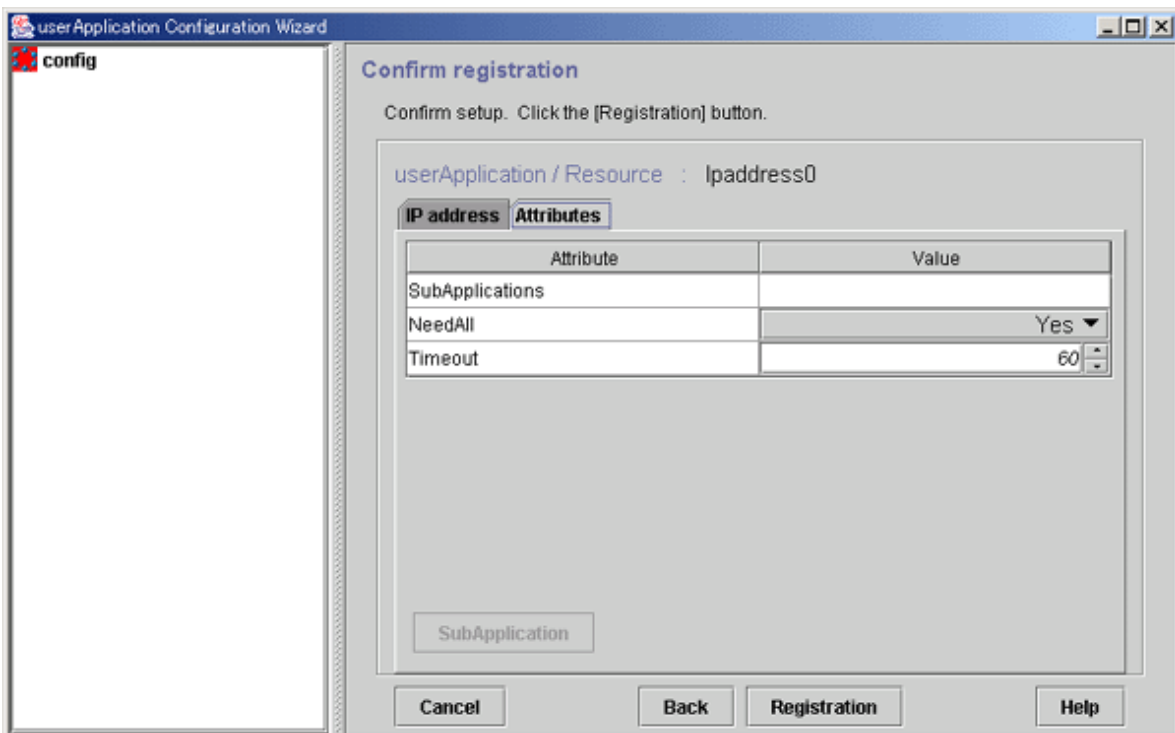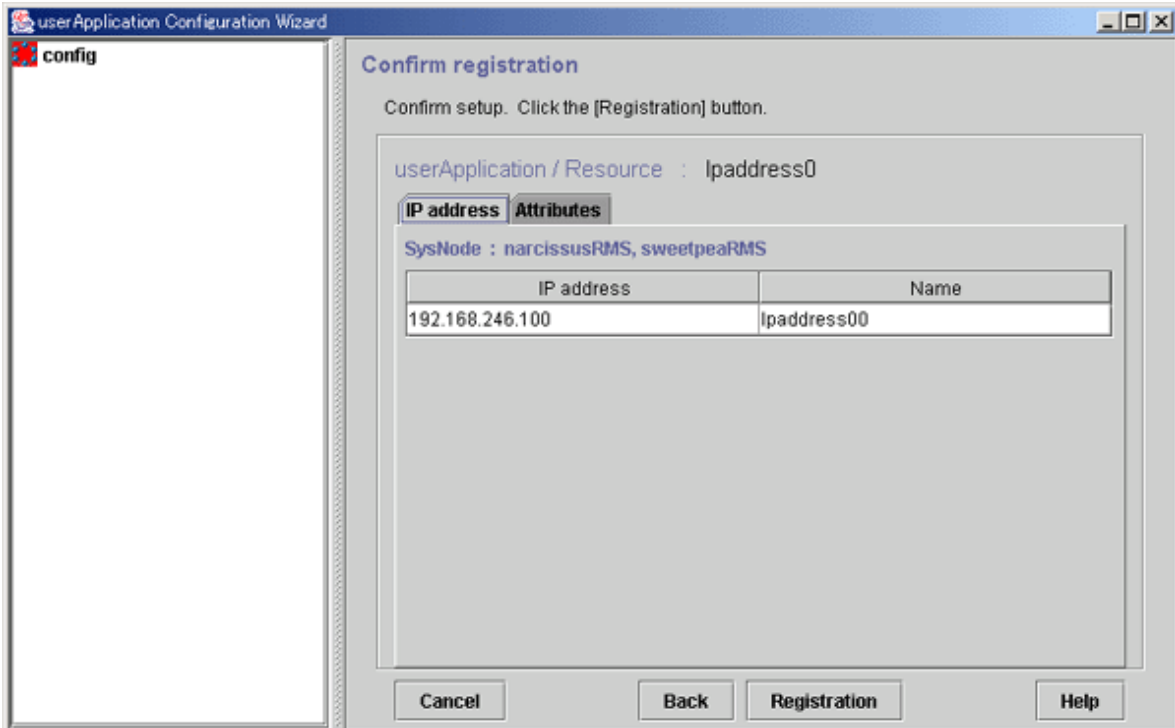
SubApplication button

Select this button to associate a previously created takeover network under the current takeover network. This button can be selected only if there are resources of the same type that can be associated. For setting instructions, see "Resource association" in "6.6.1.1 Creating Cmdline Resources."

Check the registration information, and then click *Registration*.

6.6.1.6 Creating Procedure Resources

A procedure resource must be created to migrate a SynfinityCluster product to PRIMECLUSTER. Procedure resources can be created only for those products described in "Part 6 PRIMECLUSTER Products."

6.6.1.6.1 Prerequisites

To create a procedure resource, you must first create a state transition procedure and register the procedure to the resource database.

![Note icon] Note

If a process is started without process monitoring from a state transition procedure, the hard- and software limits in the file descriptor (fd) for that process will both be "1024".

If a process started from a state transition procedure requires a file descriptor above "1024", please add a "ulimit" definition at the beginning of processing for the state transition procedure, and specify the maximum value for the file descriptor.

![Example icon] Example

For setting the maximum value for the file descriptor to "4096":

```
#!/bin/sh
ulimit -n 4096
...
```

![See icon] See

For details on how to register a state transition procedure, see "Appendix F Registering, Changing, and Deleting State Transition Procedure Resources for SynfinityCluster Compatibility."

6.6.1.6.2 Setup Instructions

Register the procedure resources that were registered to the cluster resource manager to RMS.

Figure 6.38 Flow of procedure resource creation
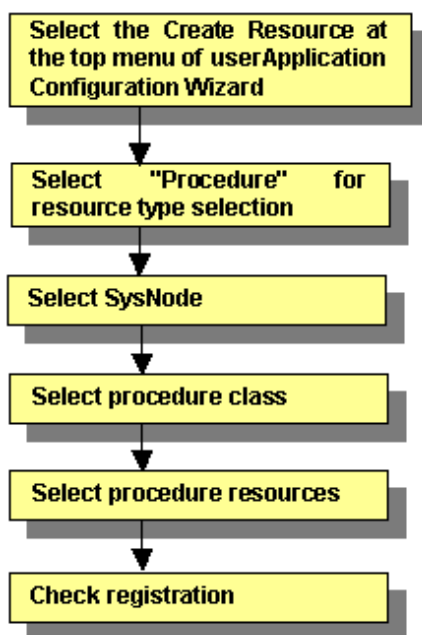
For information on the above operations up to "Select SysNode," see "6.6.1.1 Creating Cmdline Resources." This section describes the operations from *Select procedure class*.

Selecting the procedure class

Select the procedure class to be created.

Figure 6.39 Select procedure class
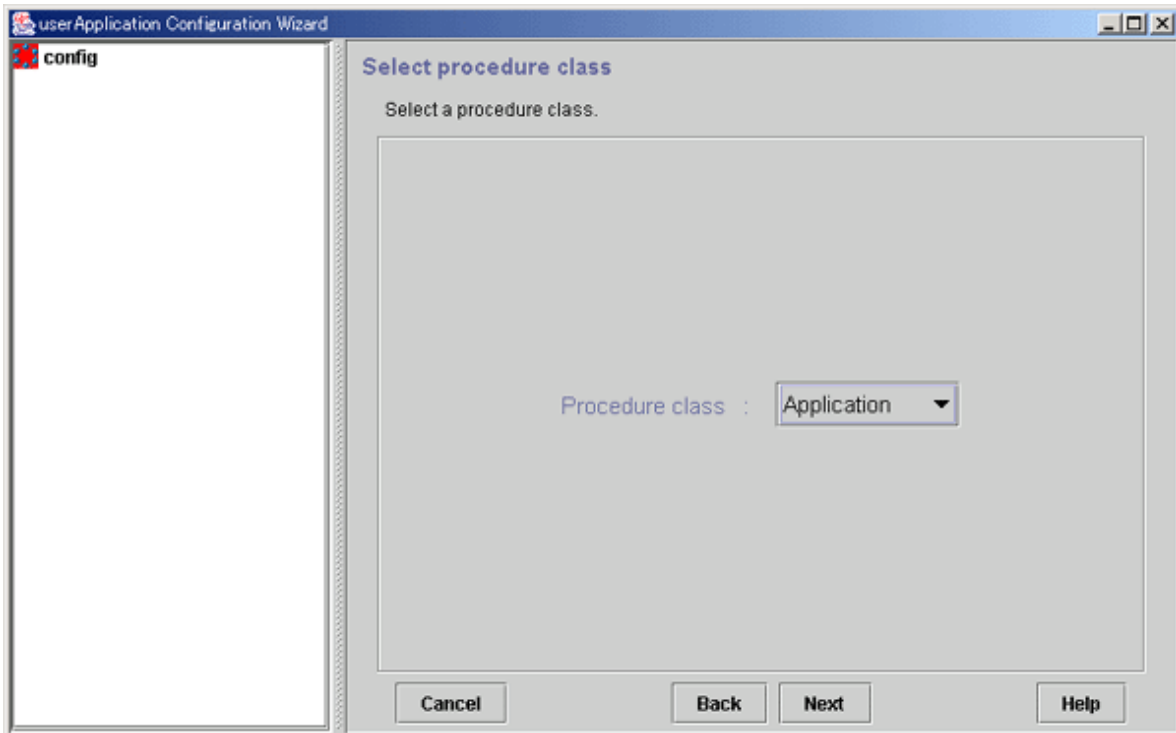
Procedure class

Select the procedure class. Only the resource classes of resources that have been registered to the cluster resource manager are displayed. For example, if only procedure resources of the Application class have been registered, only *Application* is displayed.
The class resource manager provides four standard procedure classes:

- **Application**

  This class is used to make general applications to be cluster-aware.

- **BasicApplication**

  It is a class used for DBMS.

- **SystemState2**

  It is used to make a part of OS functionality cluster-aware. This OS function should be started at "/etc/rc2.d" in a non-cluster single node. SystemState2 does not automatically start during OS startup. It is used to start a cluster application only on the OPERATING node.

- **SystemState3**

  It is used to make a part of OS functionality cluster-aware. This OS function should be started at "/etc/rc3.d" in a non-cluster single node. SystemState3 does not automatically start during OS startup. It is used to start a cluster application only on the OPERATING node.

![Note]

Note

To create an application resource that is set in one cluster application on each node, the application resource name should be the same as the application type on each node.

After completing the setup, click *Next* to go to the "Select procedure resource" screen.

Selecting the procedure resource

Select the procedure resource to be created.

Figure 6.40 Select procedure resource
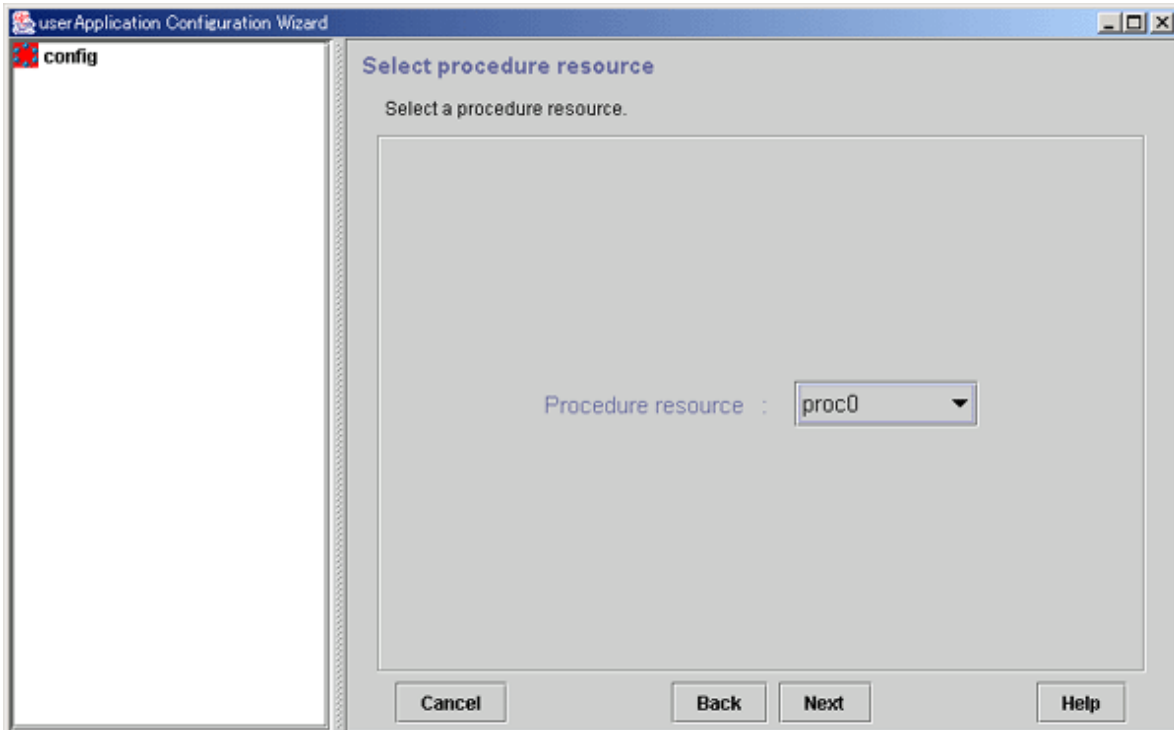
Procedure resource

Select the procedure resource to be created from the displayed list. The procedure resource name consists of up to 32 characters.

After completing the setup, click *Next* to go to the "Confirm registration" screen.

Checking the registration information for the procedure resource

Check the registration information for the procedure resource. You can also set resource attributes by selecting the *Attributes* tab and switching the screen.

For information on the resource attributes, see "6.6.5 Attributes."

Figure 6.41 Confirm registration
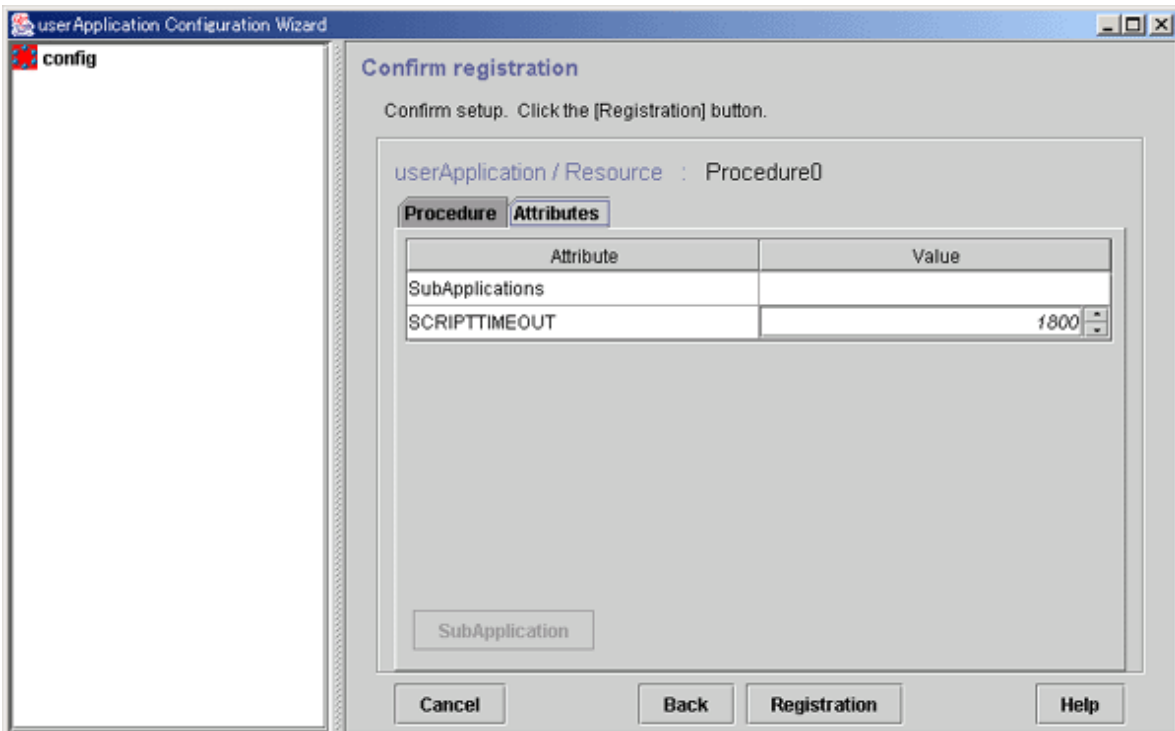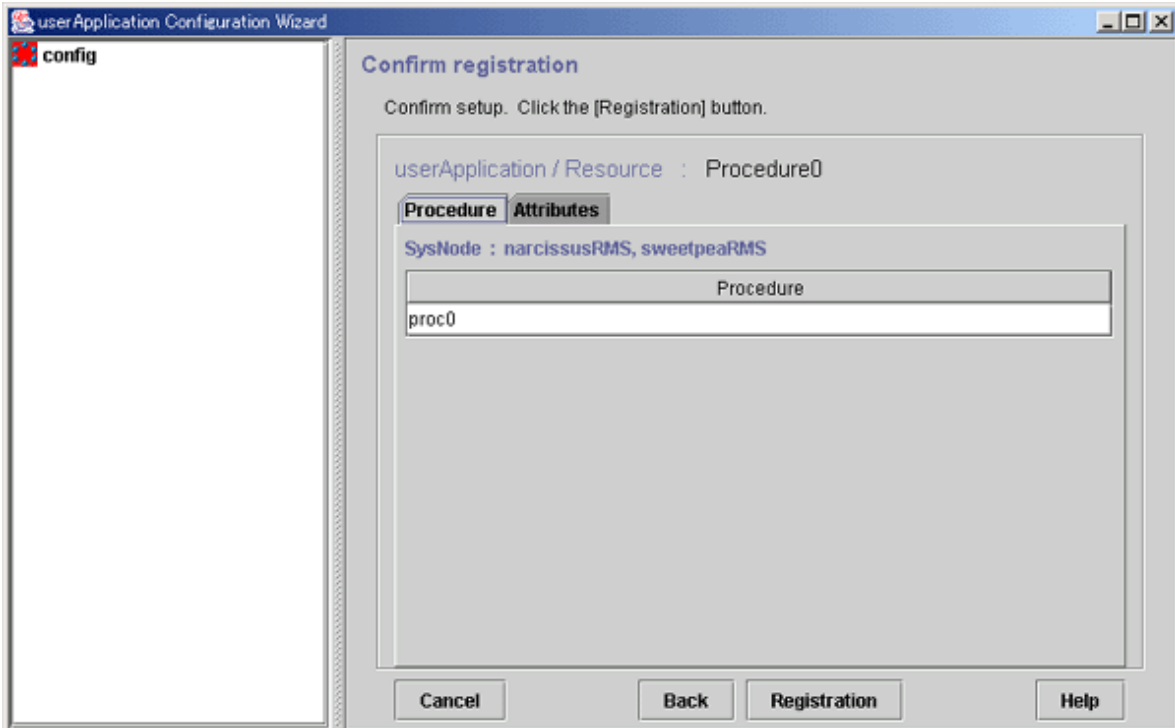
SubApplication button

This button associates the other procedure resources to the procedure resource that has been created above. For more information, see " Associate resources," which is described later.

Check the registration information, and then click *Registration*.

Resource Association

Use resource association to assign a startup order to resources of the same type.

Figure 6.42 Associate resources
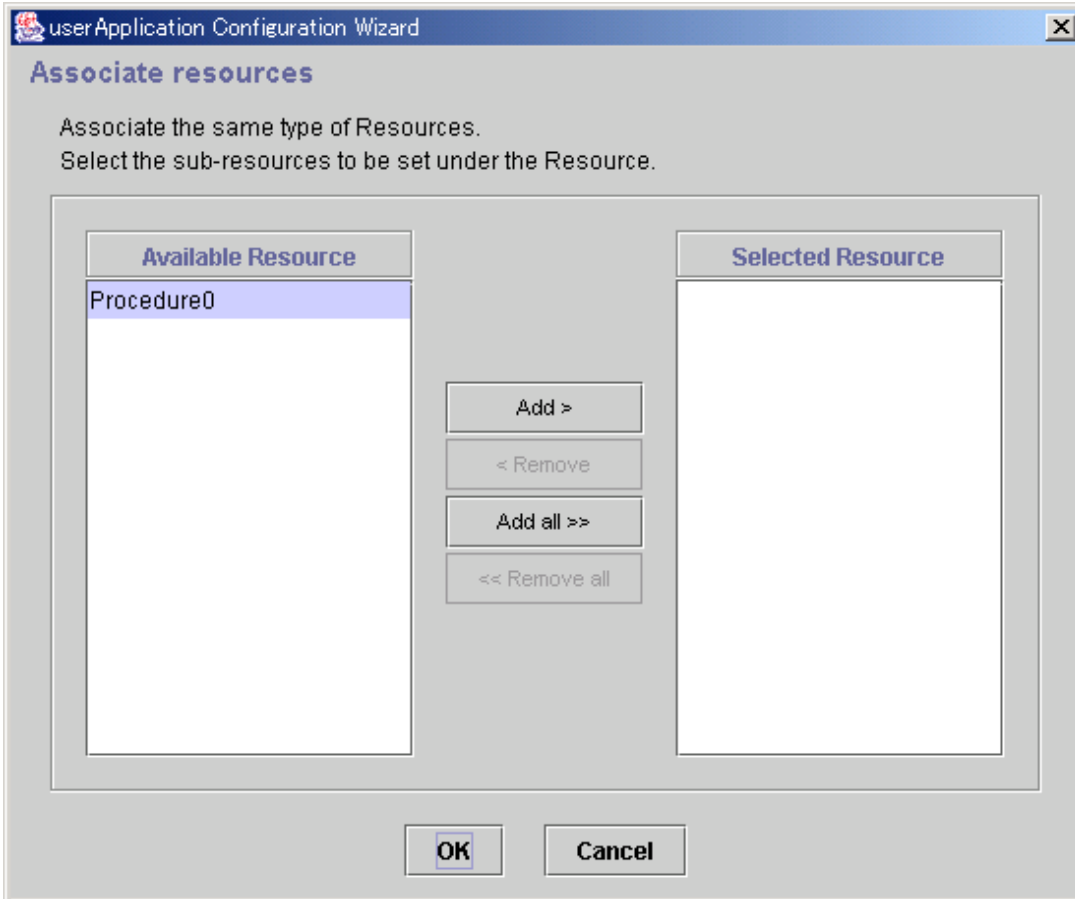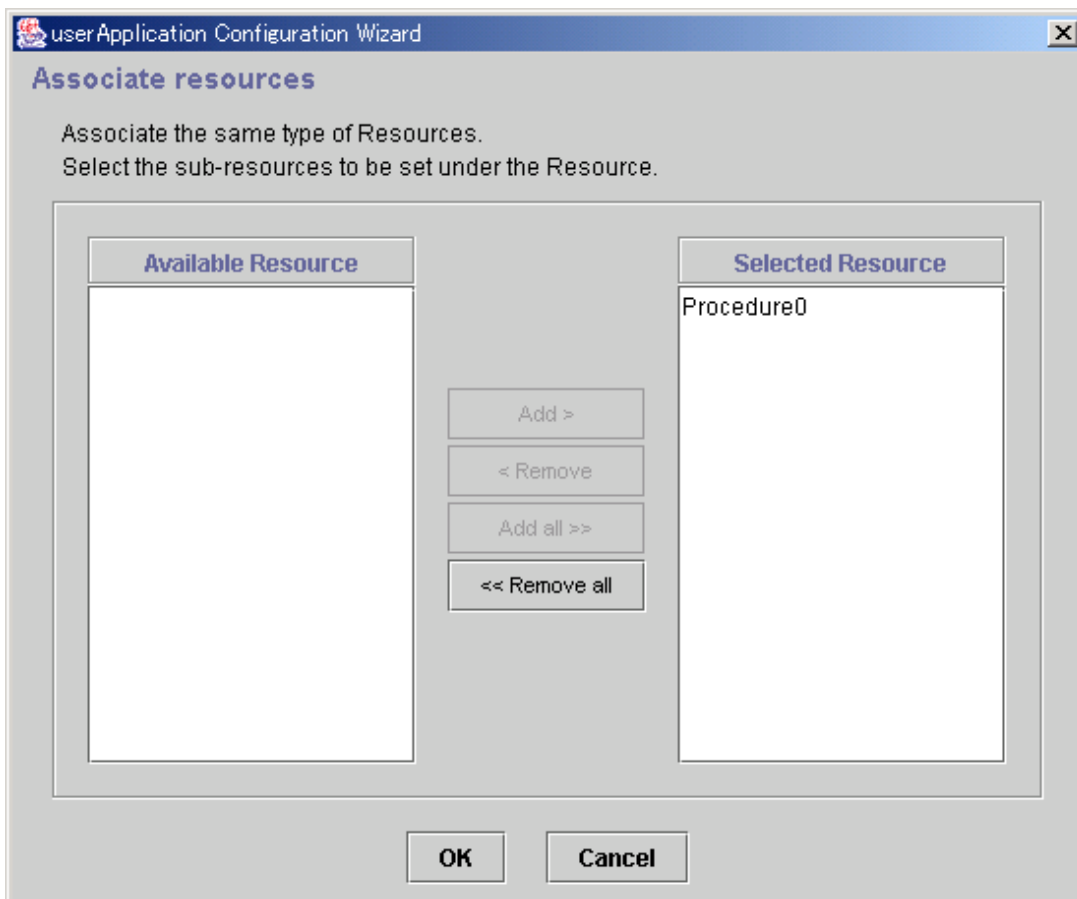
Available Resource

Available resource is referred to as the resource of the same type that can be associated, and satisfies the following conditions.

- The resource is of the same type as the resource that was called.

- The resource is not being used by another cluster application.

Selected Resource

Resources to be set under the current resource being created.

From *Available Resource*, select the resources to be configured under the current resource, and then click *Add*. To add all listed resources, click *Add all*. To remove a resource, select the resource to be removed from *Selected Resource*, and then click *Remove*. To delete all listed resources, click *Remove all*.

After completing the setup, select *OK* to return to the "*Confirm registration*" screen.

6.6.1.7 Creating Process Monitoring Resources

This section describes how to set up process monitoring resources.

Before describing the setup method, this section also outlines the process monitoring function and prerequisites for specific uses.

6.6.1.7.1 What Is the Process Monitoring Function?

Overview of the process monitoring function

The process monitoring function monitors the live state of processes. The main features are as follows:

- Changes in the live status of a process can be monitored.
  (This setup is quite easy, so the user does not need to prepare commands for monitoring the live status of a process.)

- Notifies RMS of the live state of any process immediately, and this provides high-speed switchover.

- If any process terminates abnormally because of an unexpected error, that process is automatically restarted.

A relationship diagram of the process monitoring function and RMS is shown below. The process monitoring function consists of three components: the "clmonproc" command, the Process Monitoring Daemon (prmd), and the Detector (hvdet_prmd).

- "clmonproc" command

  The "clmonproc" command is executed from the Online or Offline script. The command requests prmd to start a specified process and to stop live monitoring.

- prmd daemon

  prmd is a daemon process that starts a process and stops live monitoring according to requests received from the "clmonproc" command. If the live state of a process being monitored changes, prmd notifies hvdet prmd immediately.

- hvdet_prmd daemon

  After receiving change information on the live state of a process from prmd, the "hvdet prmd" process notifies the RMS Base Monitor (BM) of the changes.
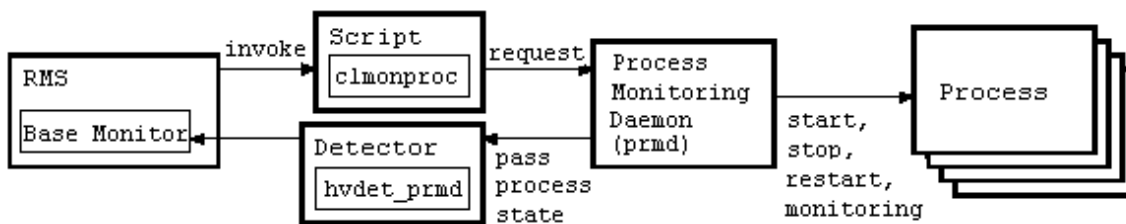
Benefits of using the process monitoring function

Described below are the benefits of using the process monitoring function.

Easy setup

Since prmd monitors whether there are any processes to be monitored, the user does not need to create a check command for each process to be monitored. The check command is used to determine whether the process to be monitored exists. Therefore with little work, the user can easily monitor the existence of processes.

High-speed detection of abnormal process termination

If the process monitoring function is not used, abnormal termination of a monitored process is detected by using a Cmdline resource to execute the "aforementioned check" command periodically. This delays detection of abnormal termination of a monitored process by execution time interval of the check command. However, if the process monitoring function is used, prmd uses signal processing to detect abnormal termination in monitored processes. This process monitoring function allows abnormal process termination to be detected at high speed compared to when check commands are executed periodically.

Automatic restart of any process that terminates abnormally

If any process terminates abnormally because of an unexpected error, the process monitoring function restarts that process automatically.

Reduction of CPU resource consumption

To shorten the time required to detect abnormal termination of a monitored process without using the process monitoring function, you must shorten the execution time interval of the check command. However, since this leads to frequent generation and execution of the check command, many CPU resources may be used up. Generally a command like the "ps" command is used as the check command. However, when a command that uses relatively more CPU resources, like the "ps" command, is used, the CPU resource consumption may become even more pronounced.

When the process monitoring function is used, prmd uses a signal process to monitor abnormal termination of the monitored process. A process that uses many CPU resources, such as one that issues a check command periodically, is not executed.

With the method that uses Cmdline resources, the number of check commands increases in proportion to the number of RMS objects because a check command is executed for each RMS object that is defined by the process to be monitored. Therefore if many check commands are executed periodically, many CPU resources may be used.

When the process monitoring function is used, it is always just one prmd that monitors the live stage of the process. Therefore, prmd does not use many CPU resources in proportion to the increase in the number of processes to be monitored.

6.6.1.7.2 Prerequisites

Normally, you do not need to work on this prerequisite for using the process monitoring function. This is required only when you want to use the processing monitoring function for the specific purpose described below.

The identification number of the detector (the state notification module for notifying RMS) used by the process monitoring function is the same as the identification number to be used by other functions.

Note

- The GUI uses "0" (default value) as the detector identification number of the process monitoring function. If a detector identification number other than "0" is already used, use that value.

- This setup is required only for changing the detector identification number from the default value "0."

- In this setup, the setting of the detector identification number must not exceed the allowed value "127."

- Configure the same detector identification number on all cluster nodes.

- Stop RMS and all running cluster applications on all nodes before changing the identification number.

Example

In the example described below, the detector identification number of the process monitoring function is changed from the initial value "0" to "2."

If the process monitoring function is being used with the initial value "0," the setting *<detector_name_of_process_monitoring_function>*.g*<identification_number>* will exist. Remove that setting.

```
# cd /opt/SMAW/SMAWRrms/bin
# ls -l  hvdet_prmd.g0
lrwxrwxrwx   1 root     other          31 Dec 20 12:21 hvdet_prmd.g0 -> /opt/SMAW/SMAWRrms/bin/
hvdet_prmd*
# rm hvdet_prmd.g0
```

Execute the "clmonsetdet" command as shown below. For the command argument, specify "2," which is the new detector identification number.

```
# /etc/opt/FJSVcluster/bin/clmonsetdet 2
```

Check whether the settings were set correctly.

```
# ls -l  hvdet_prmd.g2
lrwxrwxrwx   1 root     other          31 Dec 27 12:21 hvdet_prmd.g2 -> /opt/SMAW/SMAWRrms/bin/
hvdet_prmd*
#
```

6.6.1.7.3 Setup Instructions

This section explains how to create process monitoring resources.

Figure 6.43 Flow of process monitoring resource creation
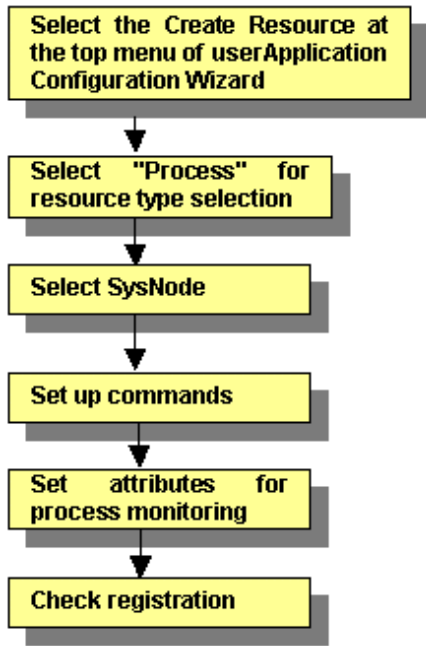
For information on the above operation up to "Select SysNode," see "6.6.1.1 Creating Cmdline Resources." This section describes the operations from "Set startup path."

Inputting the process startup command

Enter the startup path of the process to be monitored.

Figure 6.44 Set up command
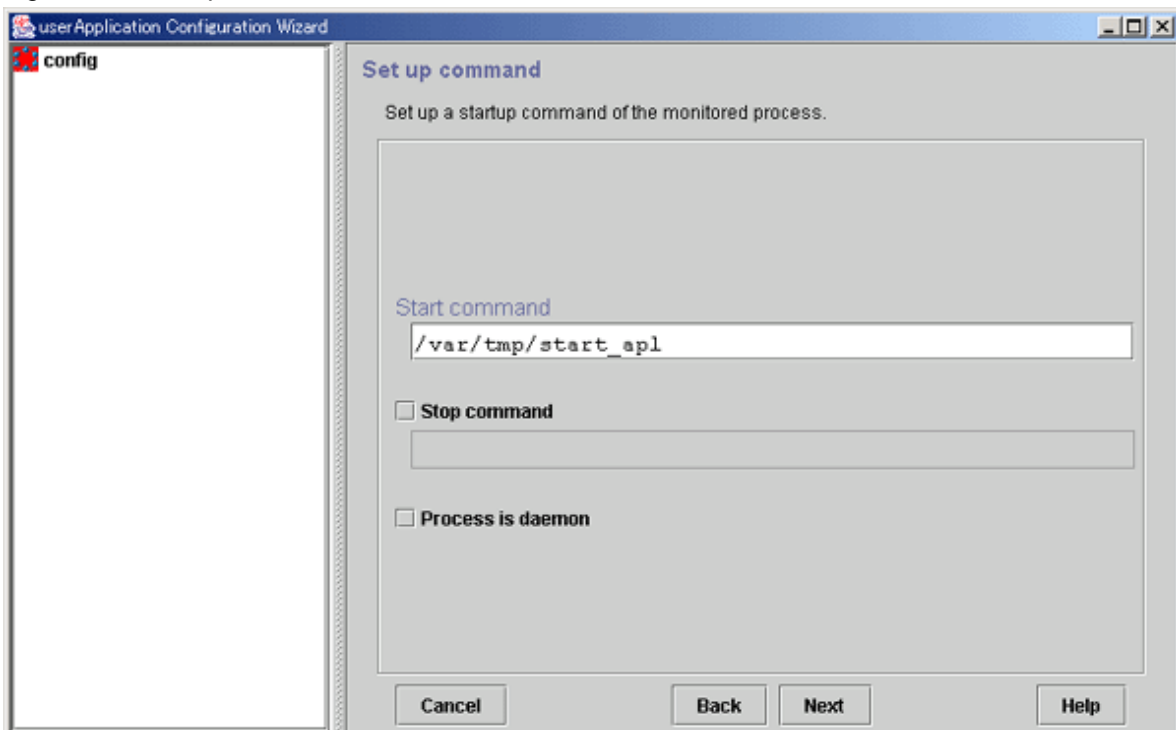
Start command

Enter the program name for starting the process to be monitored during online processing using a full path. If spaces are included in the full pathname, the pathname must be enclosed in double quotation marks (").

For example, enter the command line as follows:

```
"/var/tmp 1/start_apl"
```

The process monitoring function cannot monitor the following processes. It is necessary to specify a start command for each process monitoring resource.

- Programs that exit out of own after initiating the other programs in the background.

    Example: Shell script that exits out of own after starting three programs of "prog0", "prog1", and "prog2".

    ```
    #!/bin/sh

    prog0 &
    prog1 &
    prog2 &

    exit 0
    ```

Be aware that the process monitoring function cannot monitor child process that is generated with the start command.

Stop command

Enter the method for stopping the monitored process during Offline processing.
If there is a command for stopping the monitored process, set the checkbox to ON, and enter the command line using a full path. If spaces are included in the command line, the command line must be enclosed in double quotation marks (").

For example, enter the command line as follows:

```
"/var/tmp 1/stop_apl"
```

If the stop command is omitted, the software exit signal (hereafter referred as SIGTERM) is sent from the process monitoring facility to stop the monitored process. Some processes might not be stopped by SIGTERM. In that case, Offline processing will fail.

Process is daemon

Specify this item if the monitored process is to be operated as a daemon.
A daemon refers to a process that executes the following types of processes at startup:

- Execution in the background

- Promotion to process group leader

The former refers to child process generation when the fork system call is issued and termination of the parent process when the exit system call is issued. The latter refers to promotion to process group leader when the setpgrp system call is issued.

Note

. . . .

- The following characters cannot be used in the file name or the arguments: back slash ("\"), tilde ("~"), percent sign ("%"), ampersand ("&"), and at sign ("@").

- Single quotation marks ("'") and tabs cannot be entered.

- If a stop command is entered, a script for executing that command is created automatically. The script is stored in the following directory:

    ```
    /opt/FJSVwvucw/scripts/process_stop
    ```

. . . .

Setting attributes for the process

Enter attributes for the process to be monitored

Figure 6.45 Set up process monitoring attribute
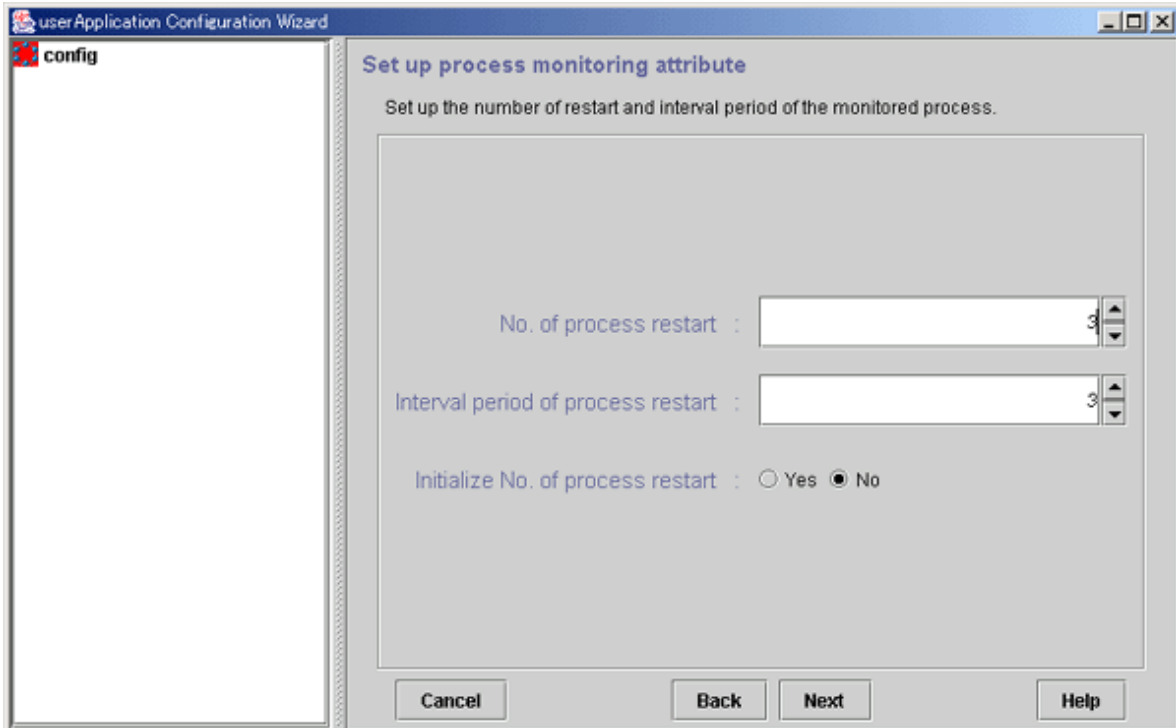
No. of process restart

Specify the number of times the monitored process is to be restarted between 0 and 99 (default is 3). If 0 is specified and the monitored process stops, the process will become Faulted.

Interval period of process restart

This is the interval from when the process monitoring facility determines that the process has stopped until the facility executes restart. The specification range is 0 to 3600 seconds (default: 3 seconds).

Initialize No. of process restart

Specify whether or not the counter that has the specified *No. of process restart* value in the process monitoring facility as its maximum value is to be initialized periodically. If you select *Yes*, the counter is initialized every "*No. of process restart* value x 60 seconds." If you select *No*, the counter is not initialized periodically.

Checking the registration information of the process monitoring resource

Check the registration information for the process monitoring resource. You can also configure resource attributes by selecting the *Attributes* tab and switching the screen.

For information on the resource attributes, see "6.6.5 Attributes."

Figure 6.46 Confirm registration
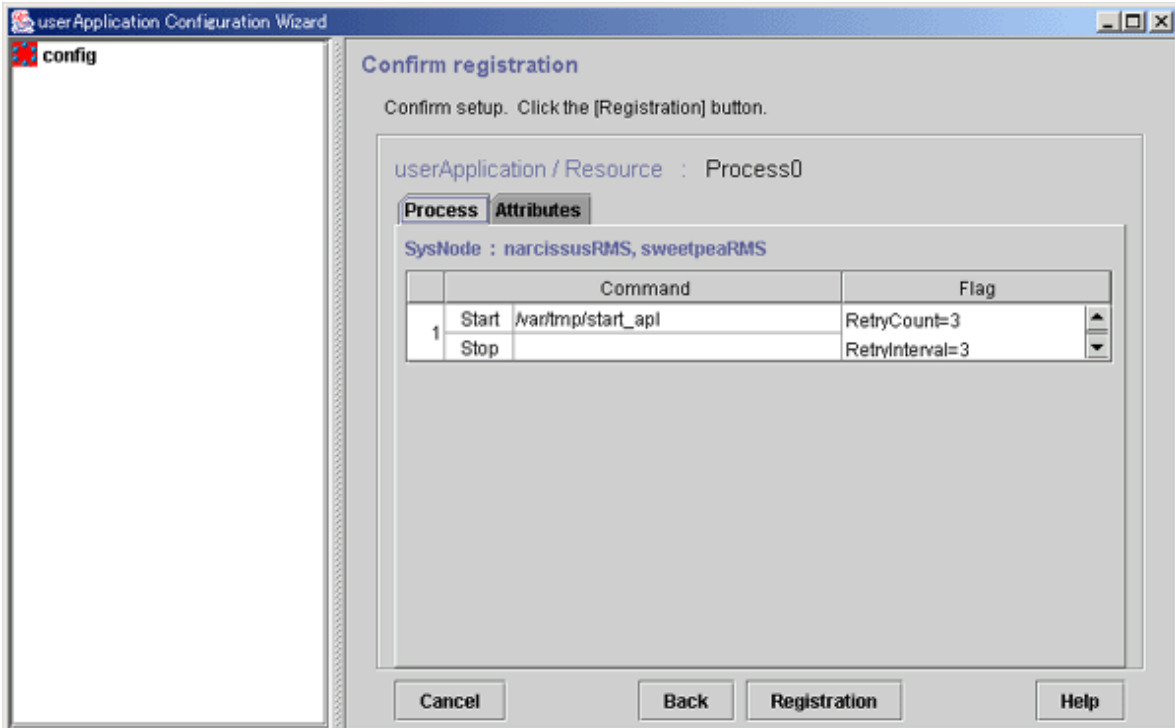
Flag

RetryCount indicates the number of times the process is to be restarted.

RetryInterval indicates the interval before starting the process.

Initialize=Yes indicates that the retry count of the process is to be initialized periodically. If *No* was specified for *Initialize No. of process restart* in the "Set up process monitoring attribute" screen, this attribute is not displayed.

Daemon=Yes indicates that the process is to be started as a daemon. If *Process is daemon* was not checked at the command setup screen, this attribute is not displayed.
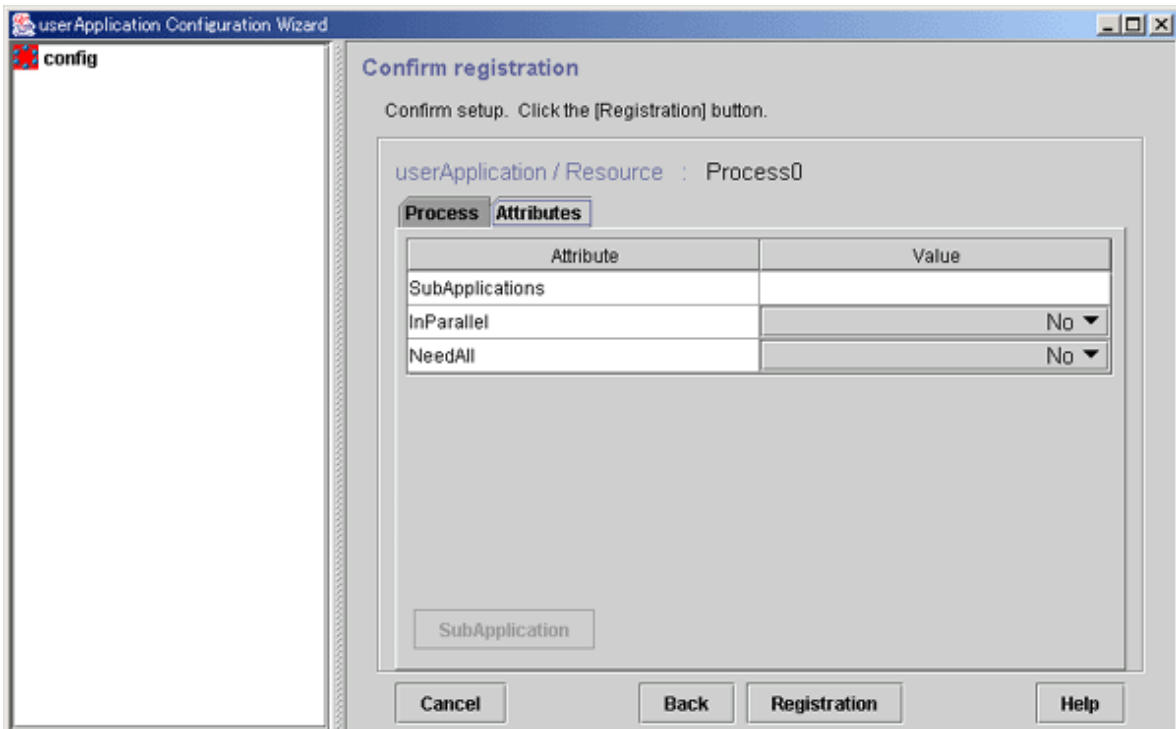
SubApplication button

This button associates the other processing monitoring resources to the Cmdline or process monitoring resource that has been created above. This button can be selected only if there are resources that can be associated. For setting instructions, see "Resource association" in "6.6.1.1 Creating Cmdline Resources."

After checking the registration information, click *Registration*.

6.6.1.8 Creating Line Switching Unit Resources

You should complete on the prerequisites before creating a line switching unit resource.

6.6.1.8.1 Prerequisites

You must always perform the setup described below before using a line switching unit resource.

Resource registration

To use a line switching unit, you must first register a switching line resource (SH_SWLine class) to the resource database.

This section describes the procedure for registering a switching line resource to the resource database.

Registration procedure flow

1. Check the resource name of the line switching unit.

2. Register the switching line resource.

3. Check the registration information.

Registration procedure

1. Confirmation of the resource name of the line switching unit resource

   Check the resource names of the line switching units that are registered to the resource database by using the "clgettree(1)" command.

```
# clgettree
Cluster 1 cluster
    Domain 2 CLUSTER
        Shared 7 SHD_CLUSTER
                SHD_DISK 21 SHD_Disk21 UNKNOWN
                        DISK 22 c5t0d0 ON node1
                        DISK 27 c4t0d0 ON node2
                SH_SWU 18 SWU2002 UNKNOWN
        Node 3 node1 ON
                Ethernet 29 hme0 ON
                DISK 19 c0t0d0 UNKNOWN
                DISK 22 c5t0d0 ON
        Node 5 node2 ON
                Ethernet 30 hme0 ON
                DISK 25 c0t0d0 UNKNOWN
                DISK 27 c4t0d0 ON
```

   In this example, the line indicating "SH_SWU" shows the resource name is "SWU2002". If you cannot confirm the resource name, see "5.1.3.2 Automatic Configure" and register the line switching units to the resource database.

See
. . . .

   - For details on the information that is output by executing the "clgettree(1)" command, see "Explanation of resources related to line switching units."

   - For details on the "clgettree(1)" command, see the manual page.
. . . .

1. Registration of switching line resources

   Execute the "claddswursc(1M)" command to add the switching line resources to the resource database.

   ```
   # claddswursc -k sh_swl_1 -s SWU2002 -0 node1 -1 node2 -m 0x3
   ```

   In this example, the line switching unit called "SWU2002," which was checked in step 2, is used to register the switching line resource called "sh_swl_1."

   Port 0 of the switching unit is connected to "node1," and port 1 to "node2."

   Since the two switching units LSU01 and LSU00 are to be used, the mask value is set to 0x3.

   ## 🔖 See
   ....
   For details on the "claddswursc(1M)" command, see the manual page.
   ....

2. Confirmation of registration information

   Execute the "clgettree(1)" command to check that the switching unit resources of the line switching unit have been registered to the resource database.

   Example) Resources of the "SWLine" class are the switching unit resources of the line switching unit.
   The resources of the "SH_SWLine" class show a sharing relationship of the switching units for the line switching unit.

   ```
           # /etc/opt/FJSVcluster/bin/clgettree
   Cluster 1 cluster
       Domain 2 CLUSTER
           Shared 7 SHD_CLUSTER
                   SHD_DISK 21 SHD_Disk21 UNKNOWN
                           DISK 22 c5t0d0 ON node1
                           DISK 27 c4t0d0 ON node2
                   SH_SWU 18 SWU2002 UNKNOWN
                           SH_SWLine 19 sh_swl_1 UNKNOWN
                                   SWLine 35 sh_swl_1P0 UNKNOWN node1
                                   SWLine 36 sh_swl_1P1 UNKNOWN node2
           Node 3 node1 ON
                   Ethernet 29 hme0 ON
                   DISK 19 c0t0d0 UNKNOWN
                   DISK 22 c5t0d0 ON
                   SWLine 35 sh_swl_1P0 UNKNOWN
           Node 5 node2 ON
                   Ethernet 30 hme0 ON
                   DISK 25 c0t0d0 UNKNOWN
                   DISK 27 c4t0d0 ON
                   SWLine 36 sh_swl_1P1 UNKNOWN
   ```

Explanation of resources related to line switching units

This section explains the information that is output by the "clgettree" command.

[Output example]

```
Cluster 1 cluster
    Domain 2 CLUSTER
        Shared 7 SHD_CLUSTER
            SHD_DISK 21 SHD_Disk21 UNKNOWN
                    DISK 22 c5t0d0 ON node1
                    DISK 27 c4t0d0 ON node2
            SH_SWU 18 SWU2002 UNKNOWN                        .....(a)
                    SH_SWLine 19 sh_swl_1 UNKNOWN            .....(b)
                        SWLine 35 sh_swl_1P0 UNKNOWN node1   .....(c)
                        SWLine 36 sh_swl_1P1 UNKNOWN node2   .....(d)
        Node 3 node1 ON
```

```
                    Ethernet 29 hme0 ON
                    DISK 19 c0t0d0 UNKNOWN
                    DISK 22 c5t0d0 ON
                    SWLine 35 sh_swl_1P0 UNKNOWN
          Node 5 node2 ON
                    Ethernet 30 hme0 ON
                    DISK 25 c0t0d0 UNKNOWN
                    DISK 27 c4t0d0 ON
                    SWLine 36 sh_swl_1P1 UNKNOWN
```

[Explanation]

(a) This line identifies a line switching unit resource.

   In the example shown in the above figure, "SWU2002" is the resource name of the resource representing the line switching unit.
   This resource is displayed when PRIMECLUSTER recognizes the line switching unit when automatic resource registration is executed.

(b) This line identifies a switching line shared resource of the line switching unit.

   In the example shown in the above figure, "sh_swl_1" is the resource name of the switching line shared resource (switching line name).
   This resource is displayed when the switching line resources of the line switching unit are registered to the resource database with the "claddswursc" command.
   As shown in the example of the above figure, this resource belongs to the "SH_SWLine" class.

(c),(d)

   These lines represent switching line resources of the line switching unit.
   In the example shown in the above figure, "sh_swl_1P0" and "sh_swl_1P1" are resource names for switching line resources.
   These resources are displayed when the switching line resources of the line switching unit are registered to the resource database with the "claddswursc" command.
   As shown in the example of the above figure, these resources belong to the "SWLine" class.

6.6.1.8.2 Setup Procedure

This section describes the procedure for setting up a line switching unit resource.

Figure 6.47 Flow of creating a line switching unit resource
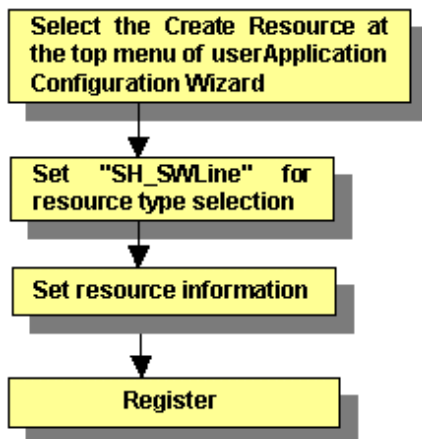
For information on the above operations up to "Selecting SH_SWLine for resource type selection," see "6.6.1.1 Creating Cmdline Resources." This section describes the operations from "Set resource information."

 Note

. . . .
The displayed Help information when you click the *Help* button includes the information for CUI only.
. . . .

Setting resource information

1. Display the top screen.

    Select *SH_SWLine* for the resource type selection. The following screen is displayed:
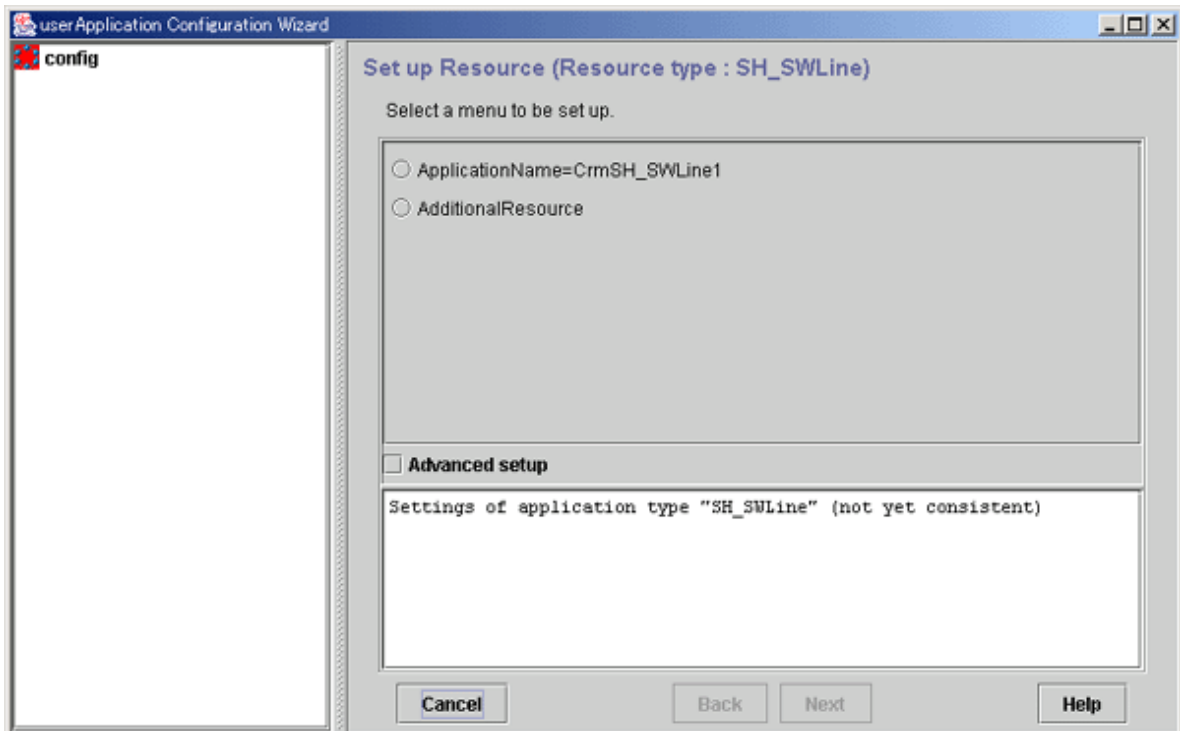
    To terminate the selection process, click the *Cancel* button.

2. Check the name of the line switching unit resource.

    name in *ApplicationName=name* is the name of the line switching unit resource to be created.

    -   If the name is correct, go to step 5.

    -   To change the name, go to step 3.

3. Select *ApplicationName=name* and click the *Next* button.

   The screen for changing the resource name is displayed.
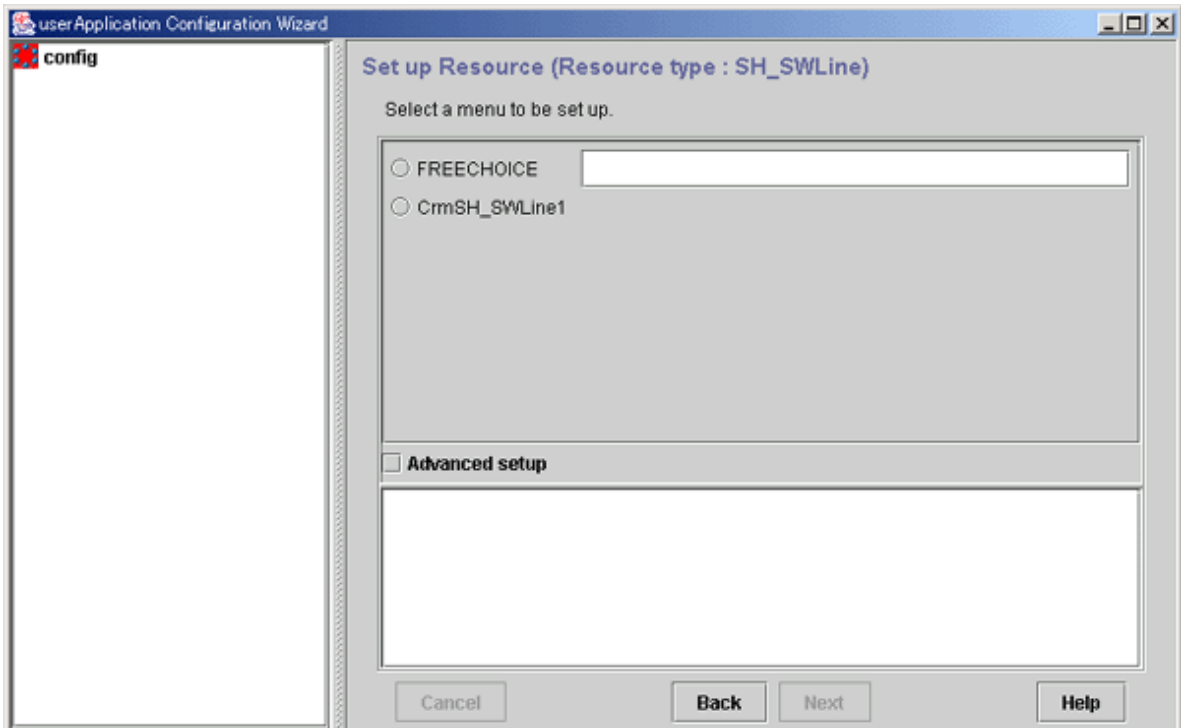
4. Click *FREECHOICE* and enter the resource name. After entering the name, click the *Next* button.

   An updated top screen containing the entered information is displayed. Go to step 2 to check the information.

5. Select the *Advanced setup* checkbox.

   A new menu will be added to the screen.
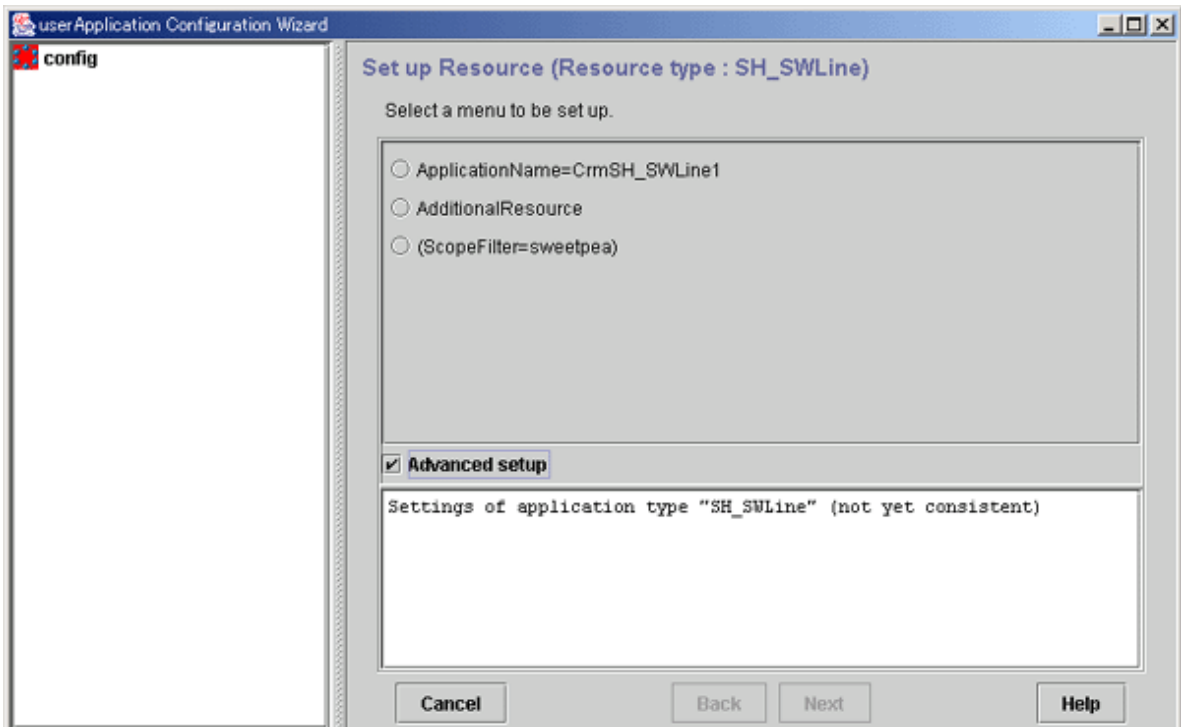
6. Check the nodes to which the line switching unit is connected.

   node-names in *ScopeFilter=node-name* shows the CF node names, which are separated with colons (":").

   - If all nodes to which the line switching unit is connected are displayed, go to step 9.

   - If there is an error, go to step 7.

Information
. . . .
Of the switching line resources (SH_SWLine class) that are registered to the resource database, those resources that are common to these nodes can be added to the line switching unit resource according to the procedure described later.
. . . .

7. Select *ScopeFilter=node-names*" and click *Next* button.

   The screen for entering the nodes to which the line switching unit is connected is displayed.
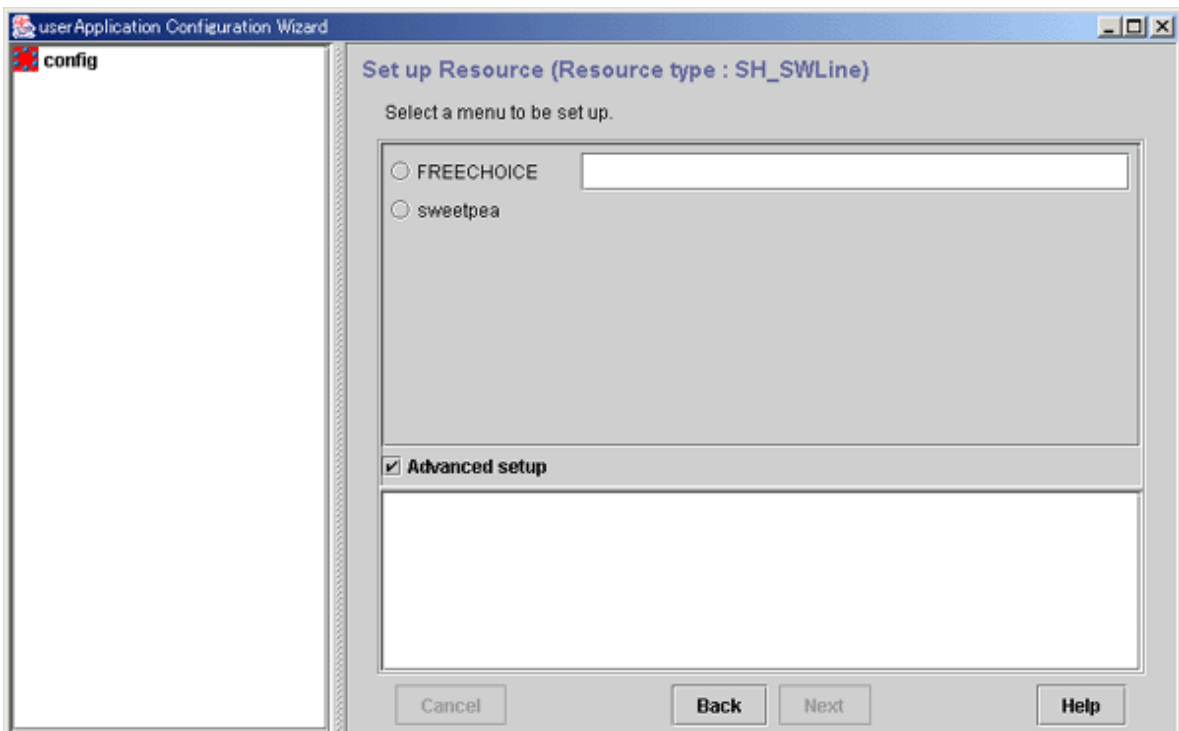
8. Click *FREECHOICE* and enter the names of the CF nodes to which the line switching unit is connected. Separate the names with colons (":"). After entering the CF node names, click the *Next* button.

   An updated top screen containing the entered information is displayed. Go to step 5 and check the information.

9. Select *AdditionalResource* and then click the *Next* button.

The screen for selecting switching line resources that are registered to the resource database is displayed.
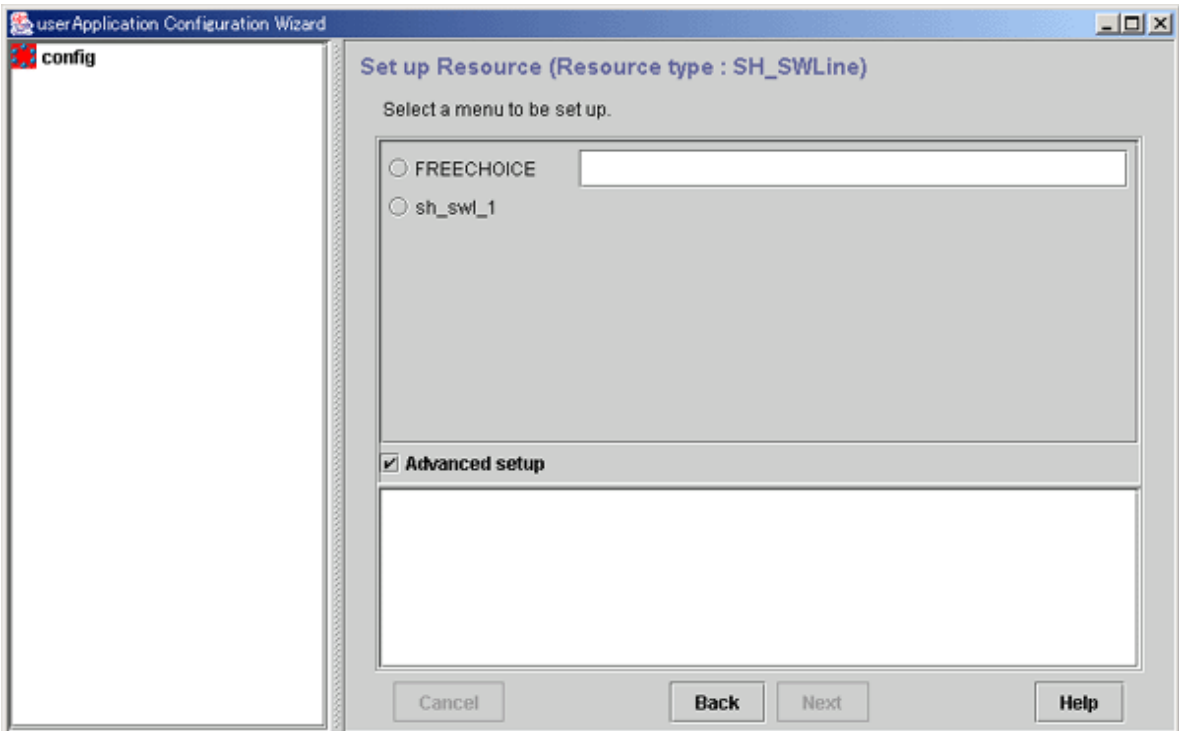
In the example shown in this screen, you can select the switching line resource called *sh_swl_1*. If no other data is displayed besides *FREECHOICE*, the following causes can be considered:

- There is an error in the nodes that were checked in step 5.

- No switching line resource is registered to the resource database.

For the first case, return to step 5 and check the nodes. To return to step 5, click the *Back* button.

For the second case, see "6.6.1.8.1 Prerequisites" and register a switching line resource to the resource database. To do this, click the *Back* button, and then click the *Cancel* button in the displayed screen. This stops the setup process, and you can start again from the beginning.

10. Select switching line resource from the candidate list, and click *Next*.

The top menu to which the switching line resource was added is displayed.

The information is displayed as follows:

```
cluster_resource[number]=rsc-name
        number    A number corresponding to the sequence in which the resource was added is
assigned.
        rsc-name   The resource name that was added is displayed.
```

Note

. . . .

The attribute must not be changed. If you add a switching line resource, the Attribute menu is added. If you use this menu to change any attribute of the switching line resource, RMS will not operate correctly.

. . . .

Go to step 1 of "Registering the information."

Registering the information

1. Select *SAVE+EXIT*.
    When you select *SAVE+EXIT*, the *Next* button changes to the *Registration* button.

    ![Note icon] **Note**

    ....

    If there is information that can be registered, the *SAVE+EXIT* menu will be added to the top screen.

    ....
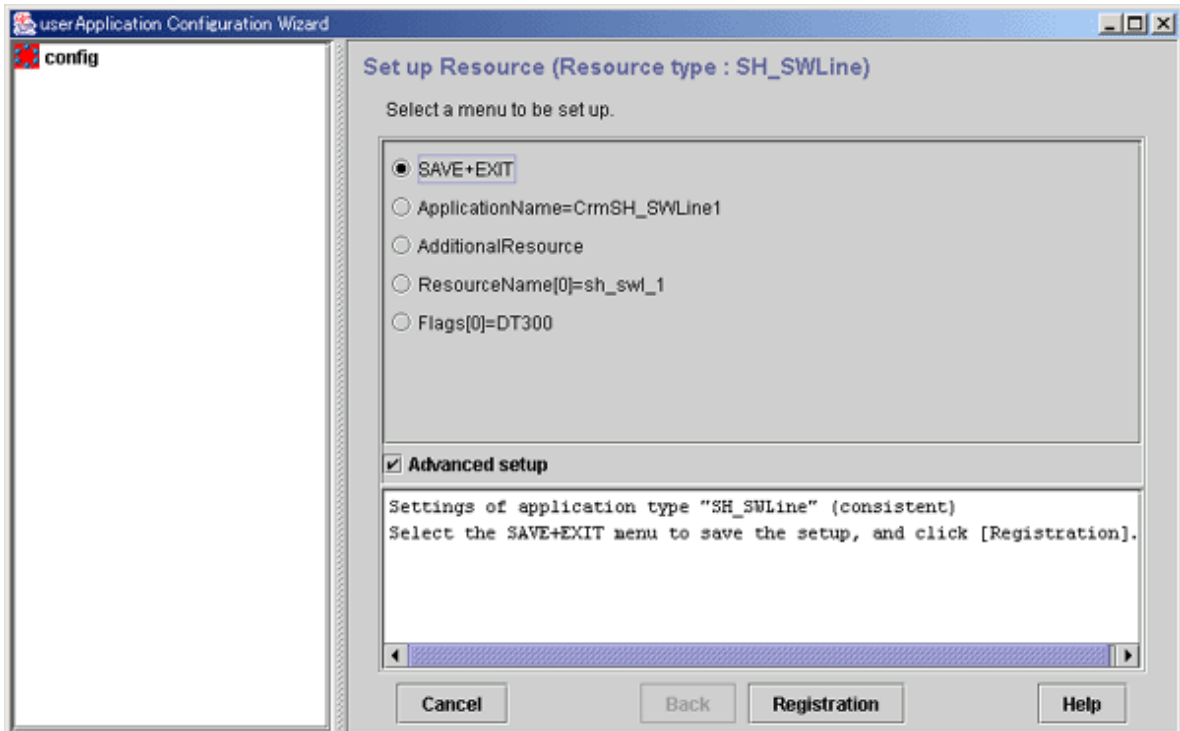
2. Click the *Registration* button.

    A line switching unit resource will be created according to the specified information, and the "userApplication Configuration Wizard Menu" screen will be displayed.
    This ends the procedure.

6.6.1.9 Creating ISV Resources

For information on ISV resource creation, refer to the manuals for the individual products.

6.6.2 Creating Cluster Applications

This section describes how to use the "userApplication Configuration Wizard" to register cluster applications (userApplication) to RMS.

Note that the operation procedure varies depending on the topology of the cluster application.

![Note icon] **Note**

....

userApplication name

When you create a cluster application, enter a name of the userApplication and then select a topology.
The userApplication name must be a string of **18 or fewer characters** starting with an alphabetic letter and consisting of alphanumerics and "_" (underscore) only.

Moreover, for this name, do not use the same name ("generic") as the type of userApplication.

Example: userApp_0

....

6.6.2.1 Creating Standby Cluster Applications

This section explains how to register standby cluster applications to RMS by using the "userApplication Configuration Wizard."

The flow of the process for registering the Standby cluster applications is shown below.

Figure 6.48 Flow of cluster application registration
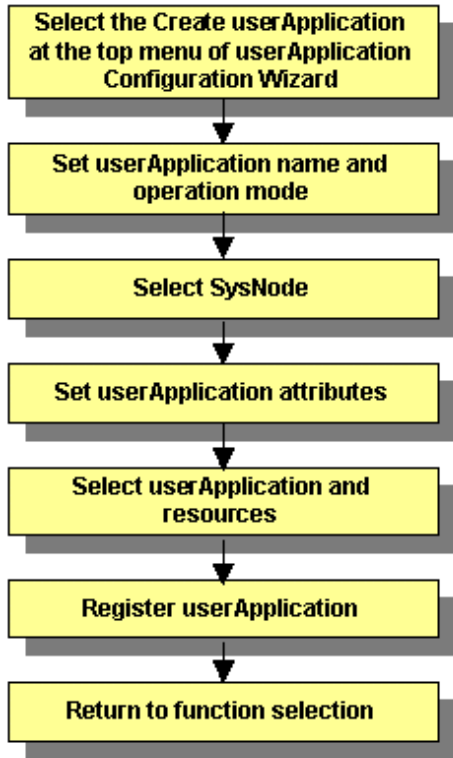
Starting cluster application creation

At the top menu of the "userApplication Configuration Wizard," select *Create userApplication*.

Figure 6.49 Top menu of userApplication Configuration Wizard
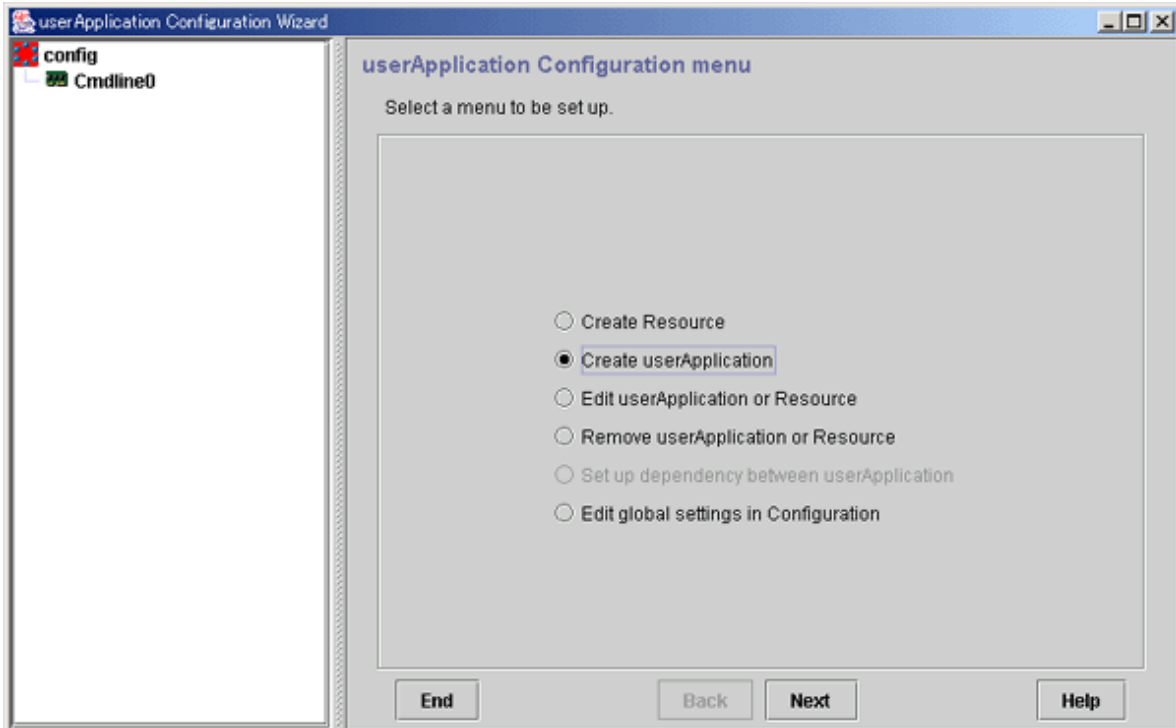
Click *Next* to go to the "Set up userApplication name and operation method" screen.

Setting the userApplication name and the operation method

Enter the userApplication name and the operation method.

Figure 6.50 Set up userApplication name and operation method
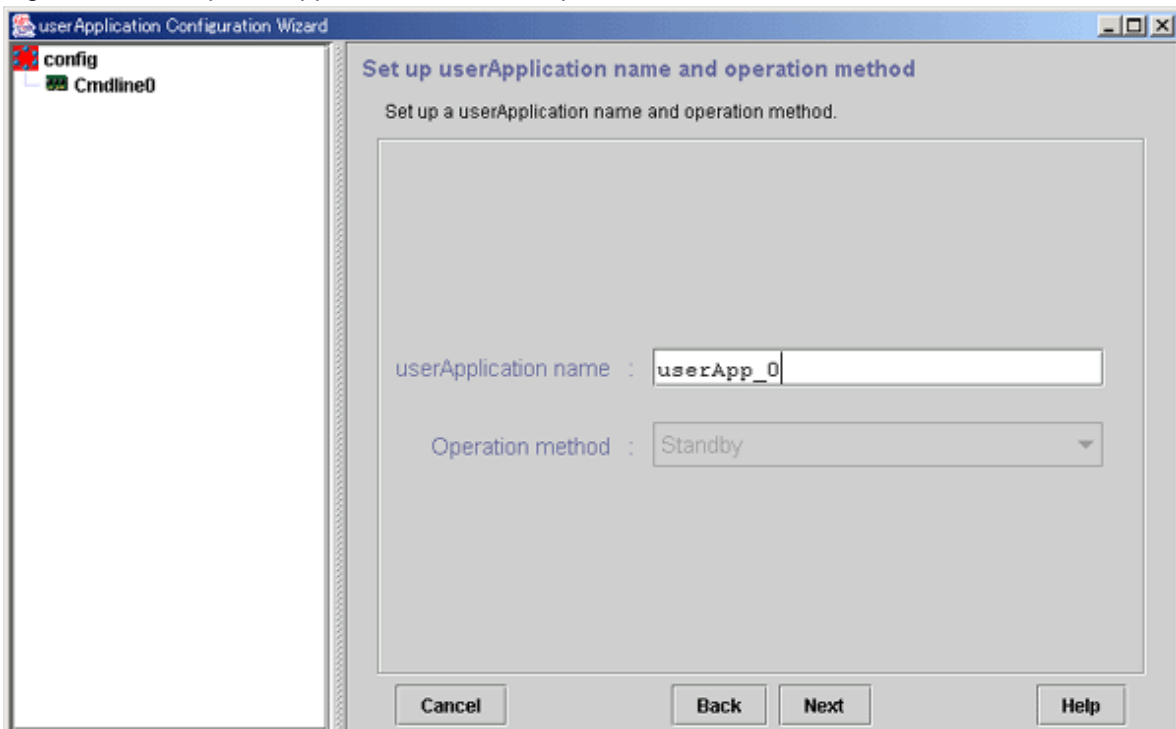
userApplication name

Specify the name of the cluster application.

Specify the input character string using **up to 18 characters** starting with an alphabetic letter and consisting of only alphanumeric characters and "_" (underscore).

Operation method

Specify Standby (standby operation) or Scalable (scalable operation) as the operation method of the cluster application.

- **Standby**

  This is the default operation method. When multiple SysNode are selected, "priority" values must be set.

- **Scalable**

  This is only for the cluster application resources displayed in *Resource Selection*. The priority does not need to be set even if multiple SysNode are selected.

After completing the setup, click *Next* to go to the "Select SysNode" screen.

Selecting a SysNode

Select the SysNode in which the cluster application is to be configured.

Figure 6.51 Select SysNode
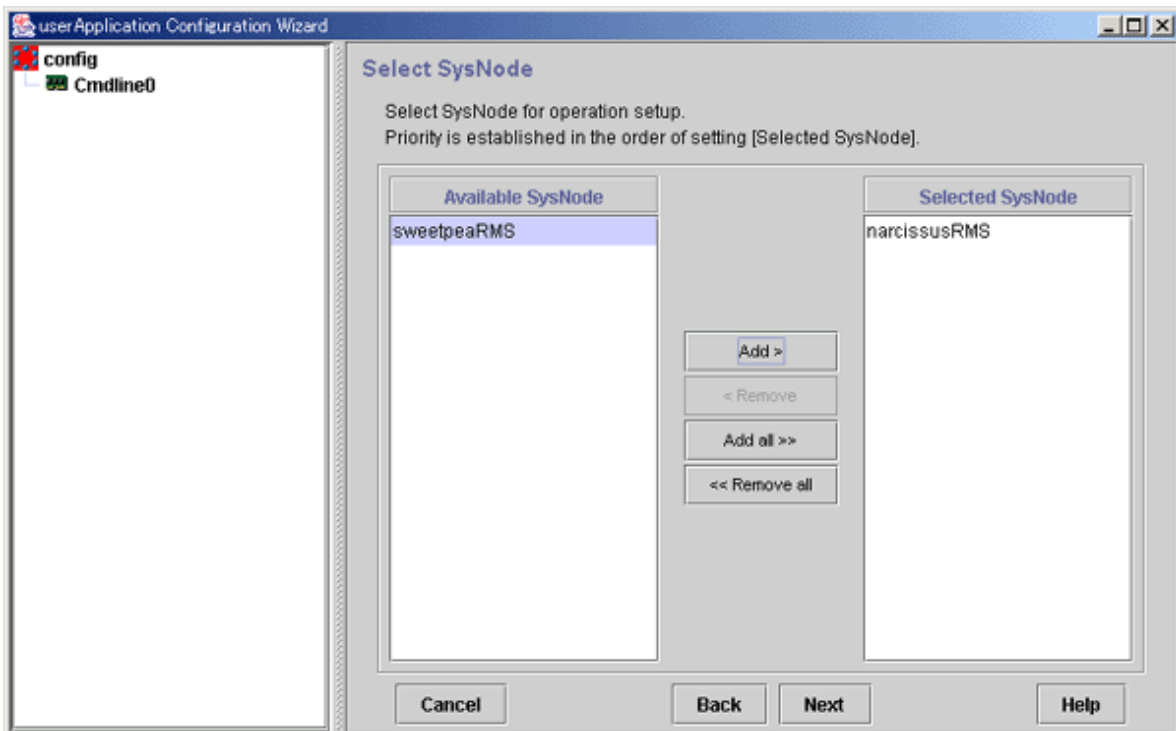
Available SysNode

Contains names of available SysNodes that belong to the specified RMS configuration.

Selected SysNode

The name of the SysNode in which the cluster application is to be operated. The priority of SysNode is determined by the sequence in which SysNode are listed in this box. SysNode displayed at the top of the list becomes the Online SysNode when userApplication is started.

From *Available SysNode*, select the SysNodes in which the cluster application is to be built, and then click *Add*. To add all listed SysNodes, click *Add all*. To remove a SysNode, select the SysNode to be removed from *Selected SysNode*, and then click *Remove*. To remove all listed SysNodes, click *Remove all*.

After completing the setup, click *Next* to go to the "Set up attribute" screen.

Setting cluster application attributes

Refer to the items below when setting up the cluster application attributes:

- To start the cluster application automatically when the cluster is started

  Set "AutoStartUp" to "Yes."

![Note icon] **Note**

If you use a standby cluster application for scalable operation, you must set "AutoStartUp" to "No." Scalable operation controls the startup of a standby cluster application.

- Action to take if another error occurs during failover (double failure)

  Set "HaltFlag" to "Yes."

![Note icon] **Note**

To ensure safe operation, be sure to set this item to "Yes."

- For Standby operation

  To monitor the resource state of the standby node, set "StandbyTransitions" to "ClearFaultRequest|StartUp|SwitchRequest."

![Note icon] **Note**

If GLS (Redundant Line Control Function) is used, please make sure to set "ClearFaultRequest|StartUp|SwitchRequest" for "StandbyTransitions" so that the Gls resources on a Standby node can be monitored.

For details on attribute, see "6.6.5 Attributes."

Figure 6.52 Set up attribute
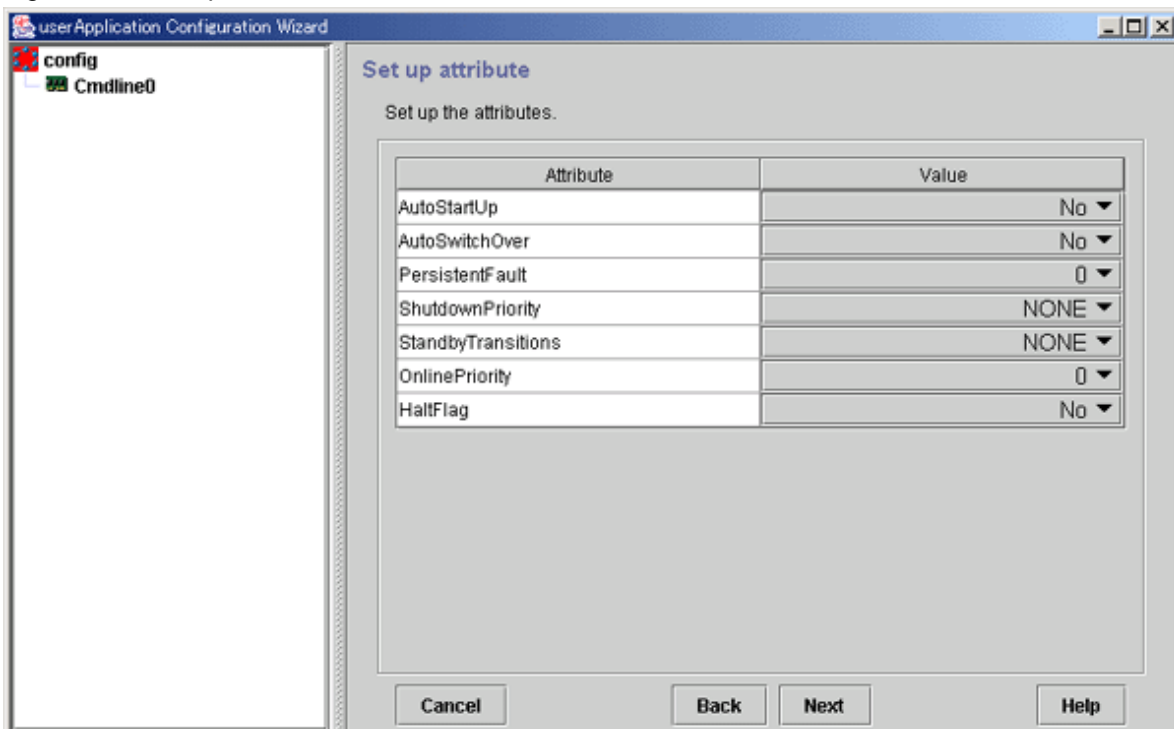
After completing the set up, click *Next* to go to the "Select Resource" screen.

Selecting resources

Enter the resources to be used by the cluster application.

Figure 6.53 Select Resource
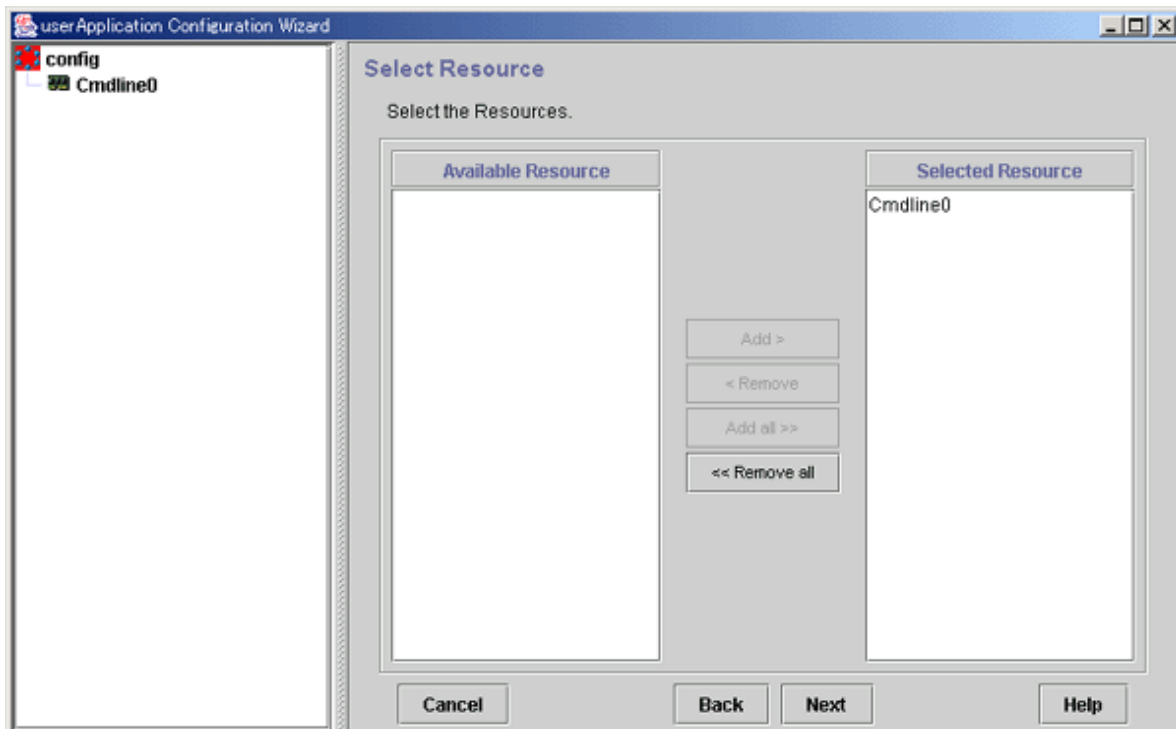
Available Resource

> The names of the resources that can be shared among the selected SysNodes are displayed. An available resource is one that was created with the same name in all the SysNodes that are specified by the cluster application.
> You cannot share one resource with multiple userApplication. The resource that is used for other userApplication is not displayed.

Selected Resource

> The resources to be registered to the cluster application are displayed.

From *Available Resource*, select the resources to be used by the cluster application, and then click *Add*. To add all listed resources, click *Add all*. To remove a resource, select the resource to be removed from *Selected Resource*, and then click *Remove*. To remove all listed resources, click *Remove all*.

After completing the setup, click *Next* to go to the "Confirm registration" screen.

Checking the registration information of a cluster application

Check the registration information of the cluster application.

Figure 6.54 Confirm registration
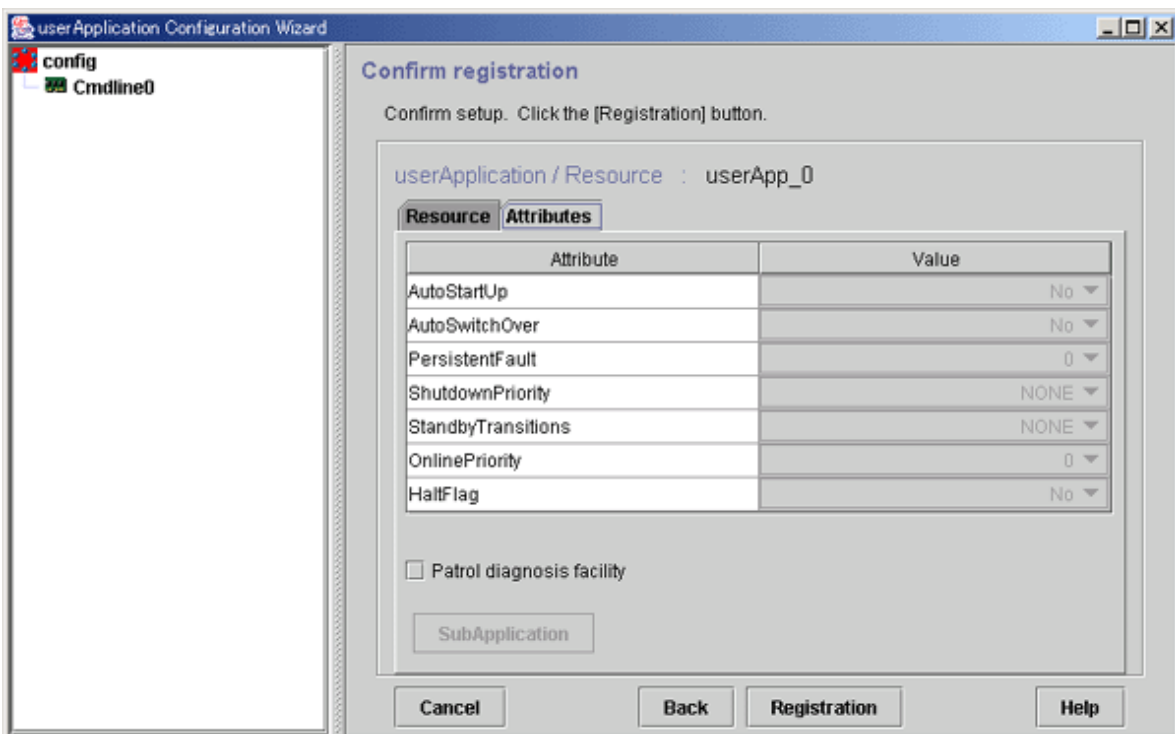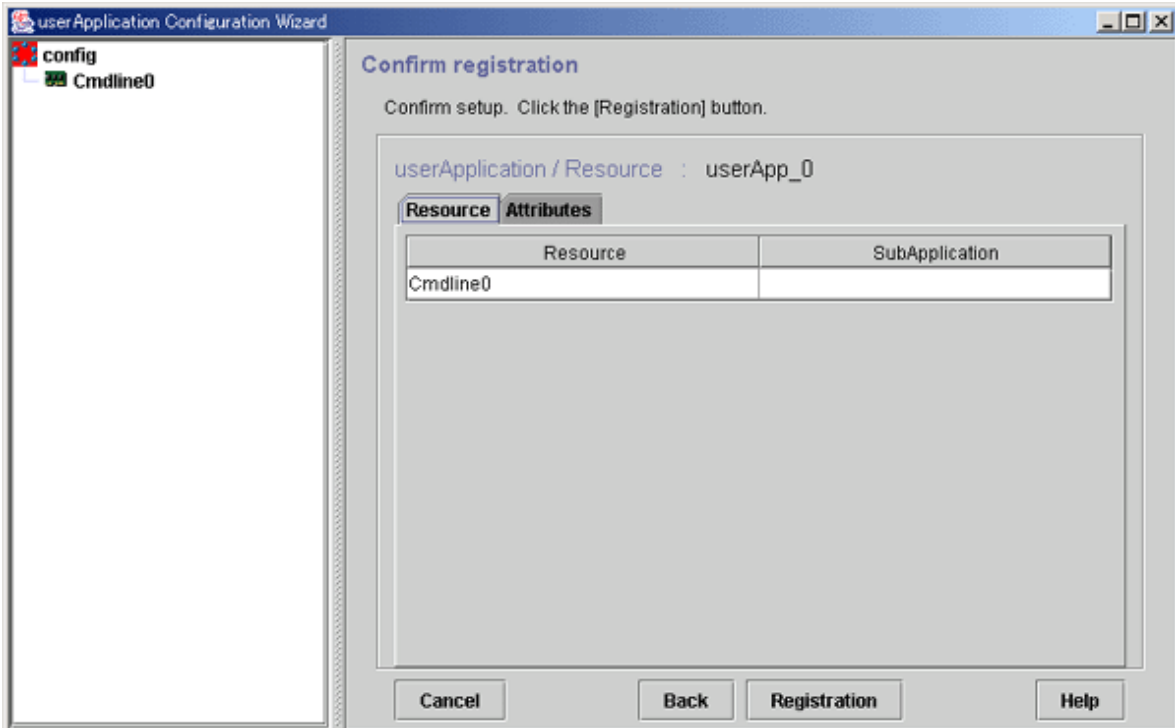
Patrol diagnosis facility checkbox

Select this checkbox to enable the patrol diagnosis facility in the cluster application. See "6.7 Setting Up Patrol Diagnosis."

SubApplication button

This button is always inactive.

The attributes displayed here cannot be changed. To change the attributes, you must return to "Setting cluster application attributes."

Check the registration information, and then click *Registration*.

6.6.2.2 Creating Scalable Cluster Applications

This section explains how to register scalable cluster applications to RMS by using the "userApplication Configuration Wizard."

The flow of the procedure for registering a scalable cluster application is shown below:

Figure 6.55 Flow of Scalable cluster application registration
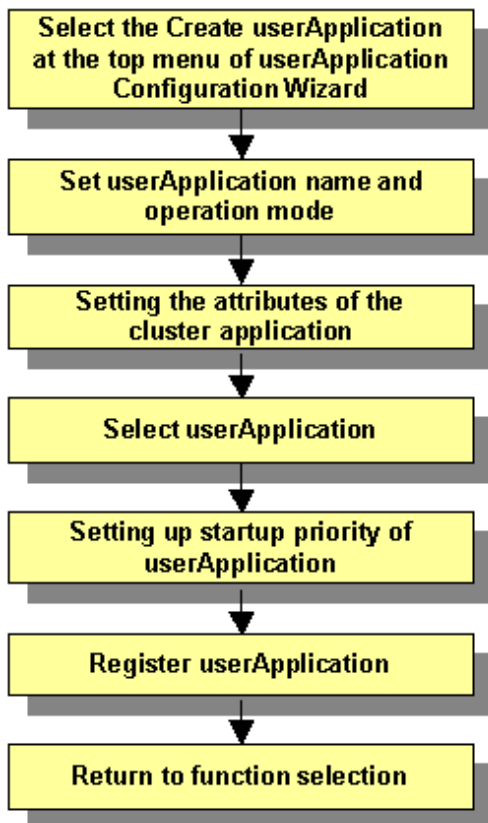
Preparing standby cluster applications

The scalable cluster application performs scalable operation by interconnecting multiple standby cluster applications. Before creating a scalable cluster application, you have to create the standby cluster applications that constitute the components of the scalable cluster application.

Figure 6.56 Preparing Standby cluster applications
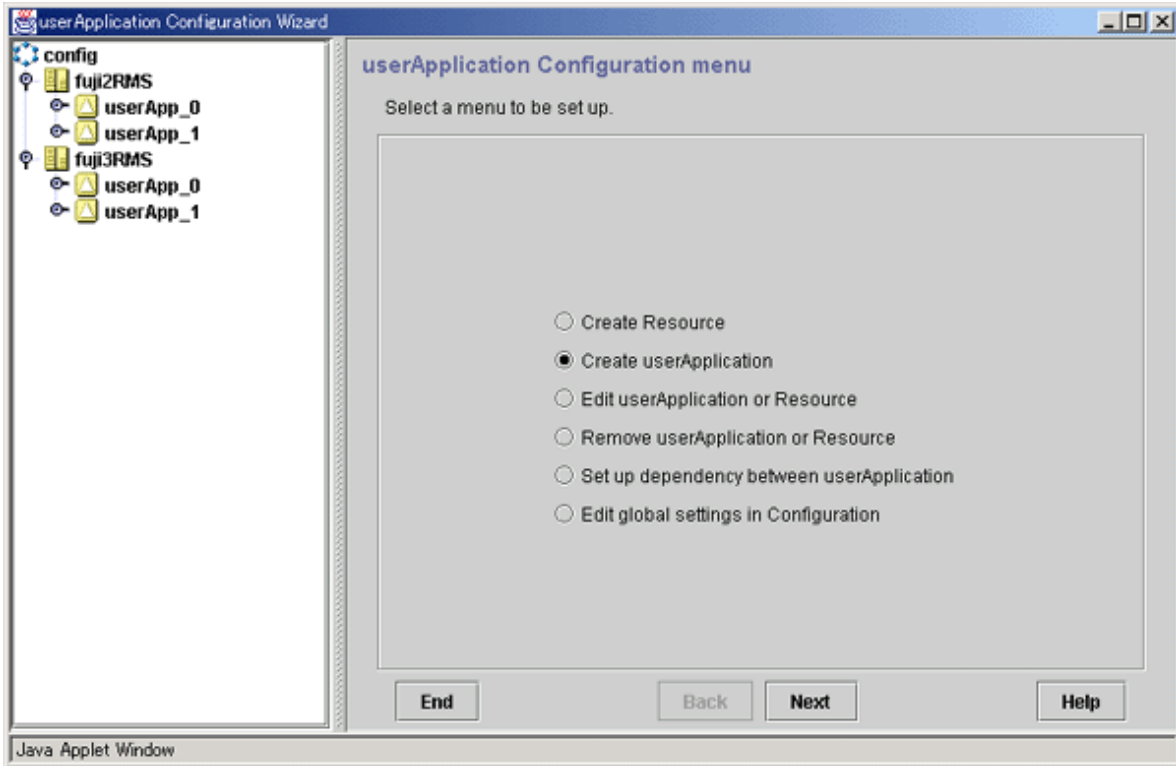
For details on how to create Standby cluster applications, see "6.6.2.1 Creating Standby Cluster Applications."

Starting up Create Cluster Applications

Specify *Create userApplication* from the top menu of "userApplication Configuration Wizard."

Figure 6.57 userApplication Configuration Wizard top menu
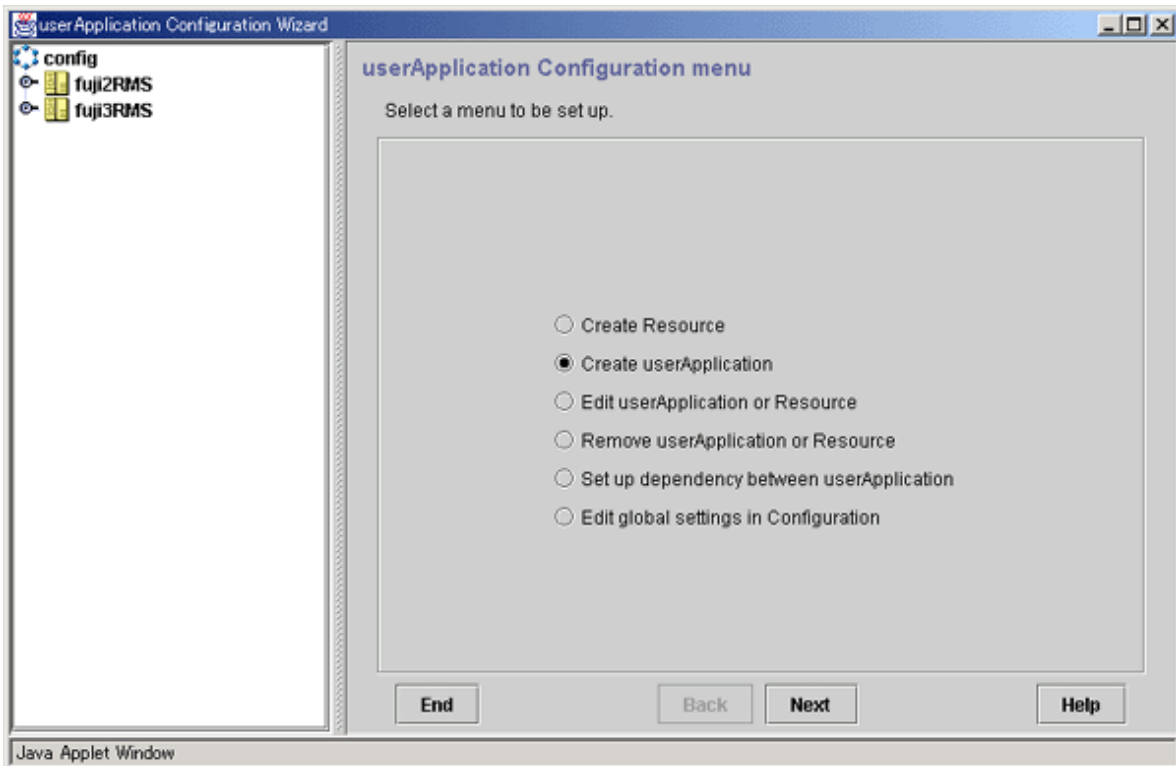
Click *Next* to go to the "Set up userApplication name and operation method" screen.

Setting userApplication name and the topology

Set up userApplication name and the topology.

Figure 6.58 Setting up userApplication name and the operation method
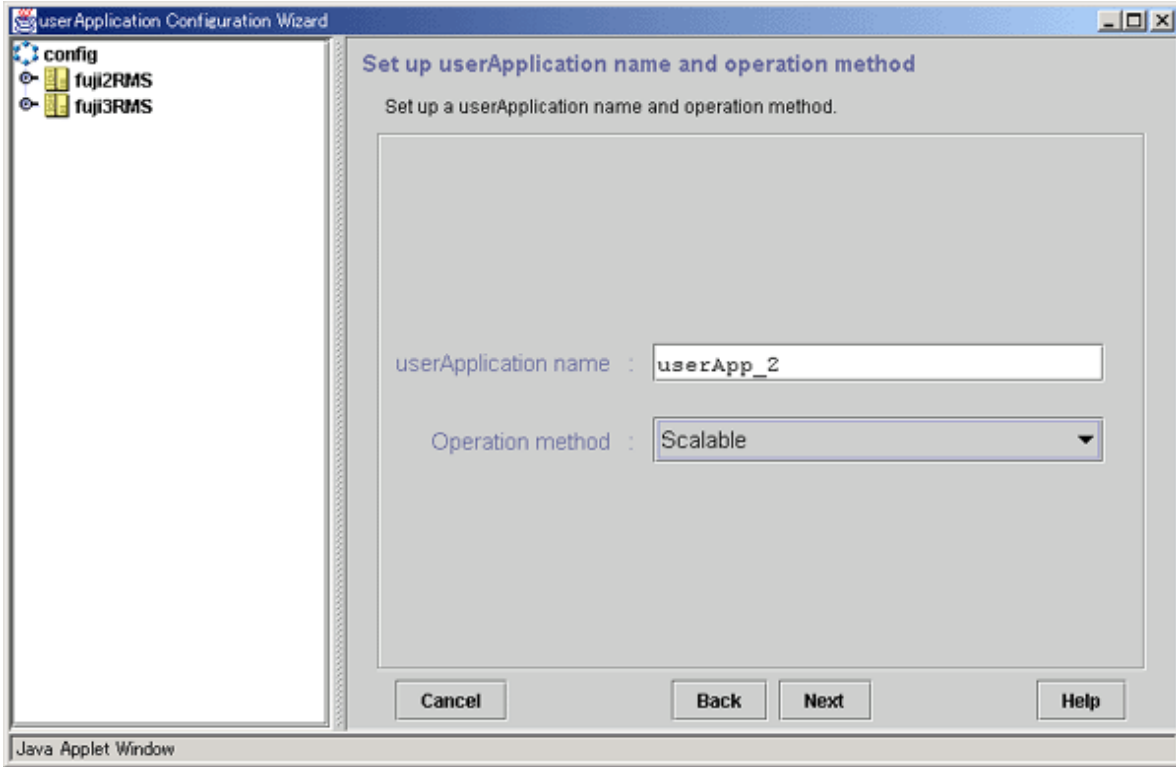
userApplication name

Specify a name for the cluster application that you want to set up.

Specify an input character string

Specify the input character string with **up to 18 characters** with an alphabetic letter and consisting of only alphanumeric characters and "_" (underscore).

Operation method

Select Scalable.

After you have made all the necessary settings, click *Next* to go to the "Set up attribute" screen.

Setting the attributes of the cluster application

Set up the attributes of the cluster application as described below:

- Set up automatic start of the scalable cluster application when the cluster starts up.

  Set up "*AutoStartUp*" to "*Yes*."

For details on this attribute, see "6.6.5 Attributes."

Figure 6.59 Set up attributes
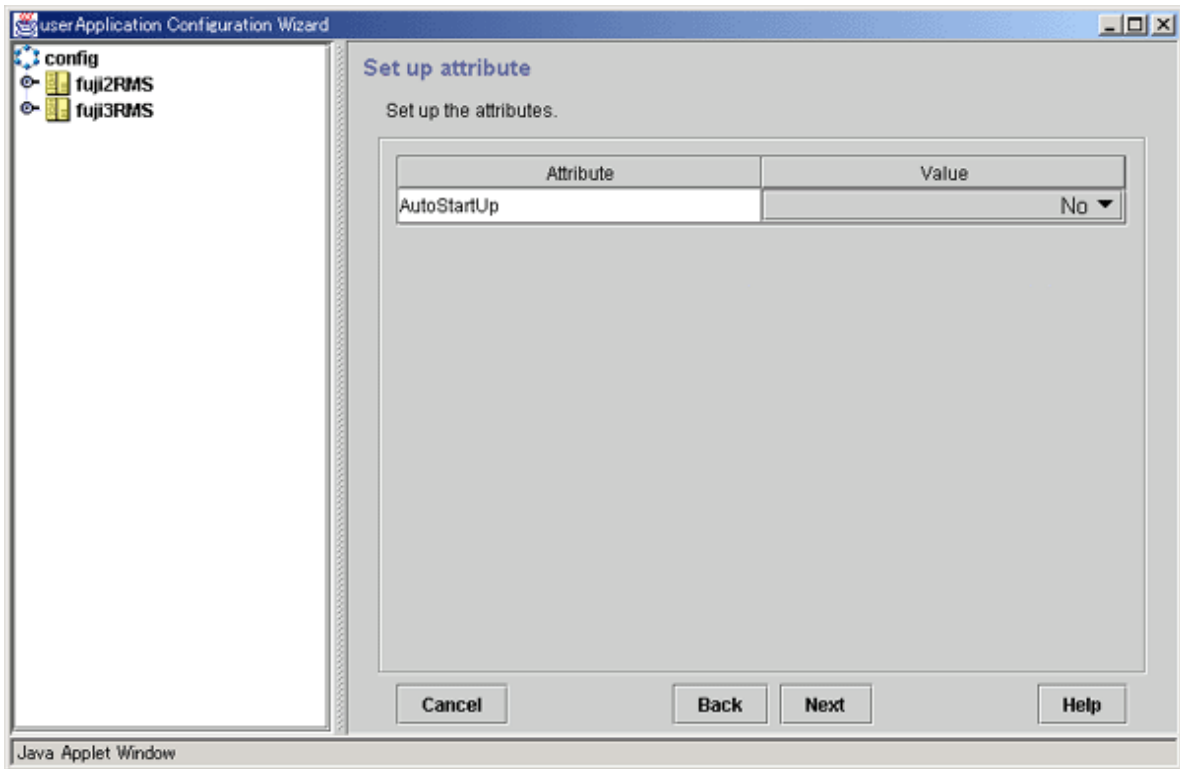
After you have made this setting, click *Next* to go to the "Select userApplication" screen.

Select userApplication

Select a cluster application for which scalable operation is to be performed.

Figure 6.60 Selecting userApplication
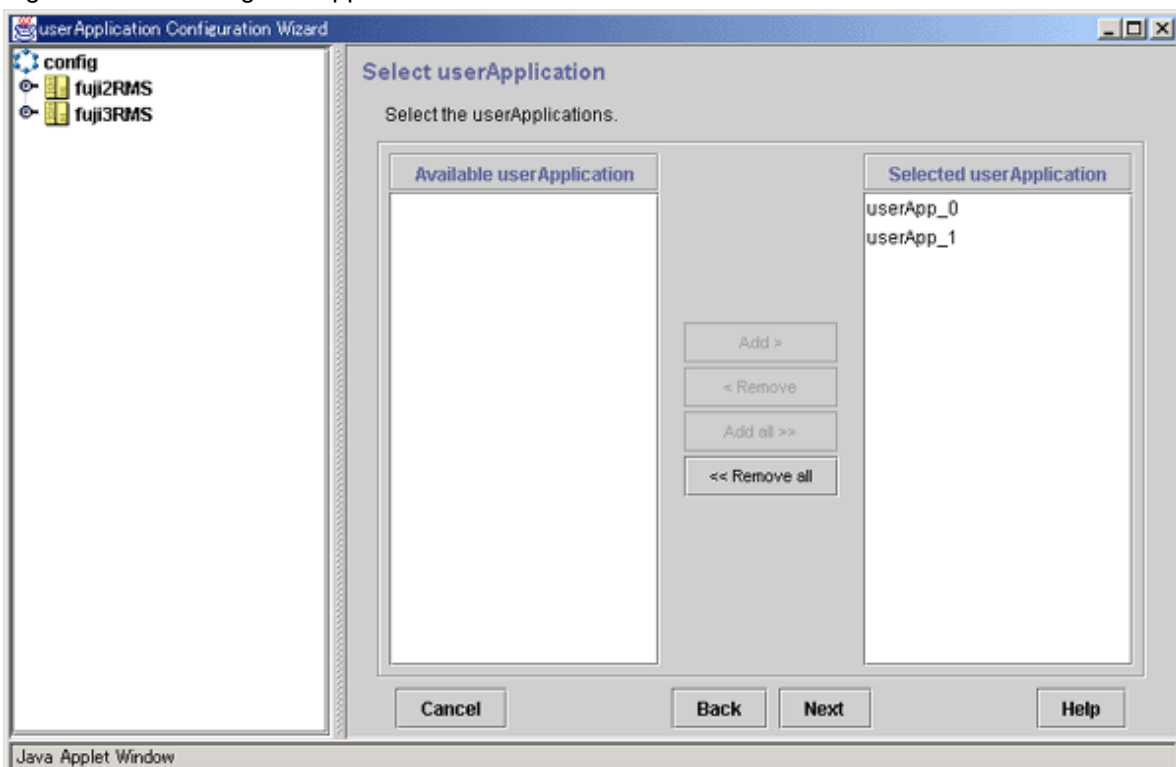
Available userApplication

Indicates a standby cluster application that can be used.

Selected userApplication

Indicates the cluster application that is being used for scalable operation.

Select a cluster application for scalable operation from *Available userApplication*, and then click *Add*. If you want to add all the cluster applications, click *Add all*. To delete a cluster application from scalable operation, select the cluster application to be deleted from [*Selected userApplication*] and then click *Remove*. To delete all the cluster applications, click *Remove all*.

After you have made these settings, click *Next* and then go to the "Set up startup priority of userApplication" screen.

Setting up startup priority of userApplication

Set up the startup order of the cluster applications.

Figure 6.61 Selecting userApplication
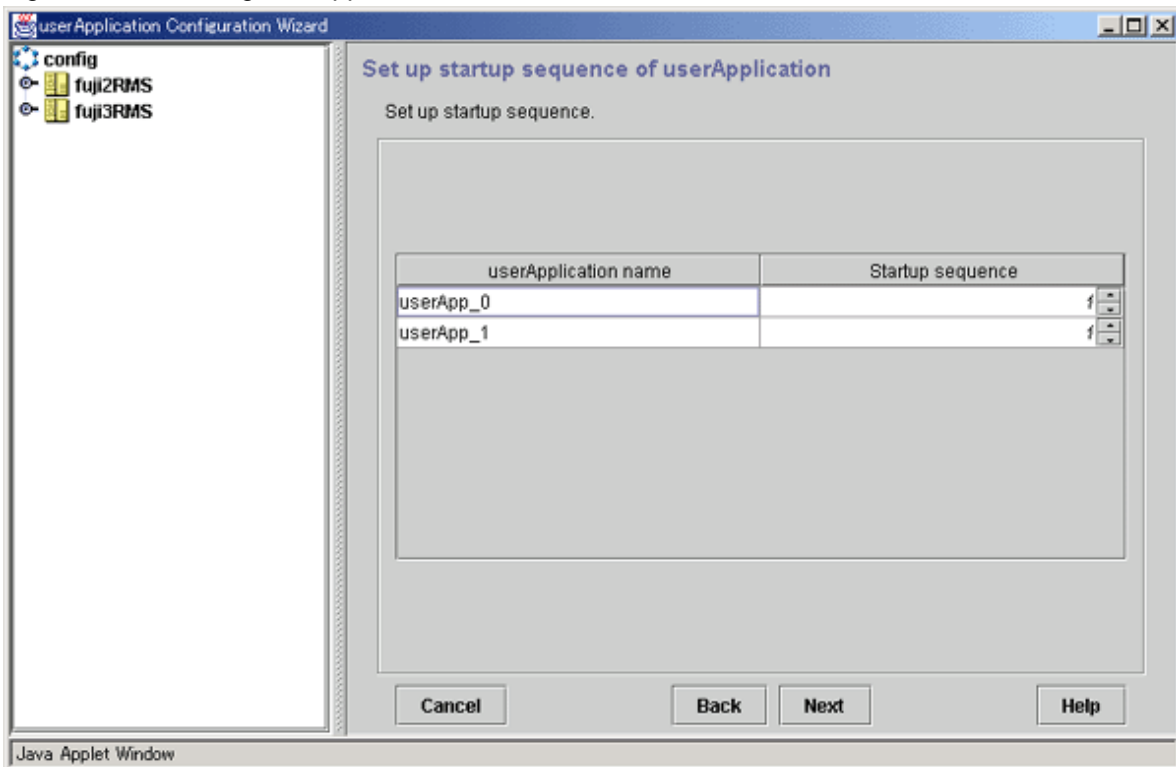
userApplication name

Cluster applications for scalable operation.

Startup priority

Set up a startup priority for the cluster applications. Cluster applications start up in order, starting from the smallest priority number. Conversely, cluster applications stop in order, starting from the greatest startup number. Cluster applications with the same startup priority start or stop in parallel.

After you have made these settings, click *Next* to go to the "Set up attribute" screen.

Confirming cluster application registration information

Confirm the registration information of the cluster application.

Figure 6.62 Confirm registration
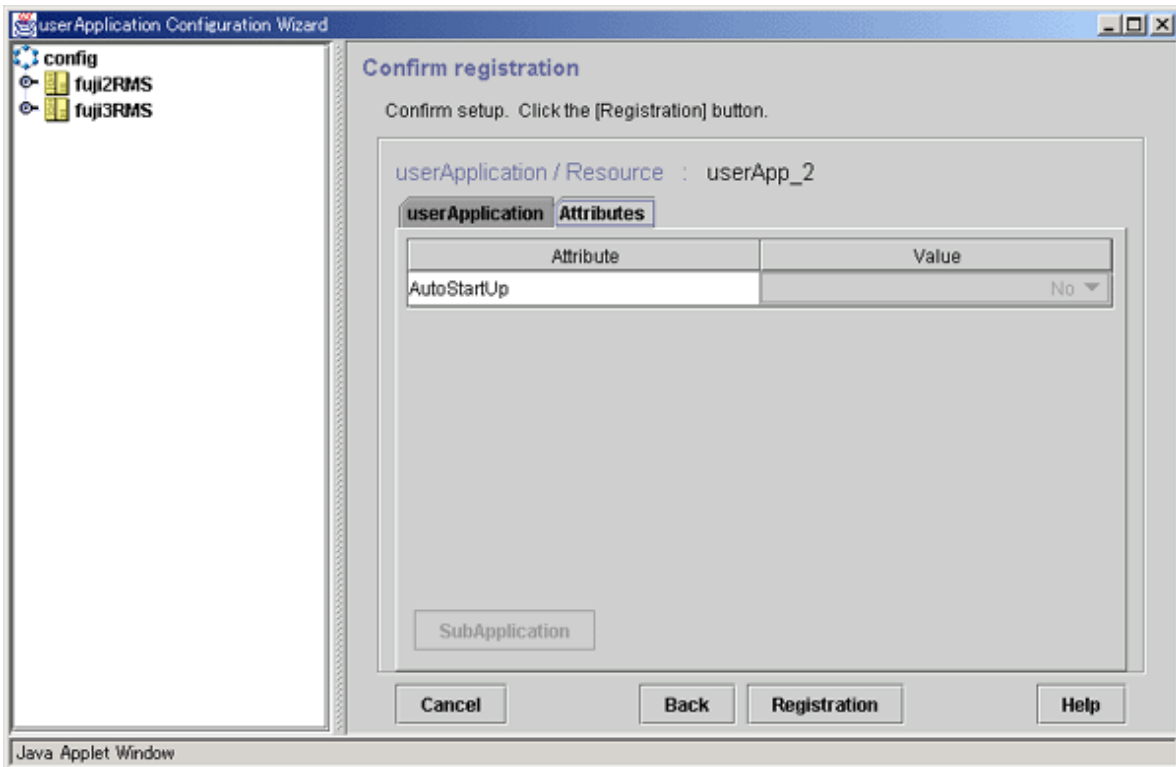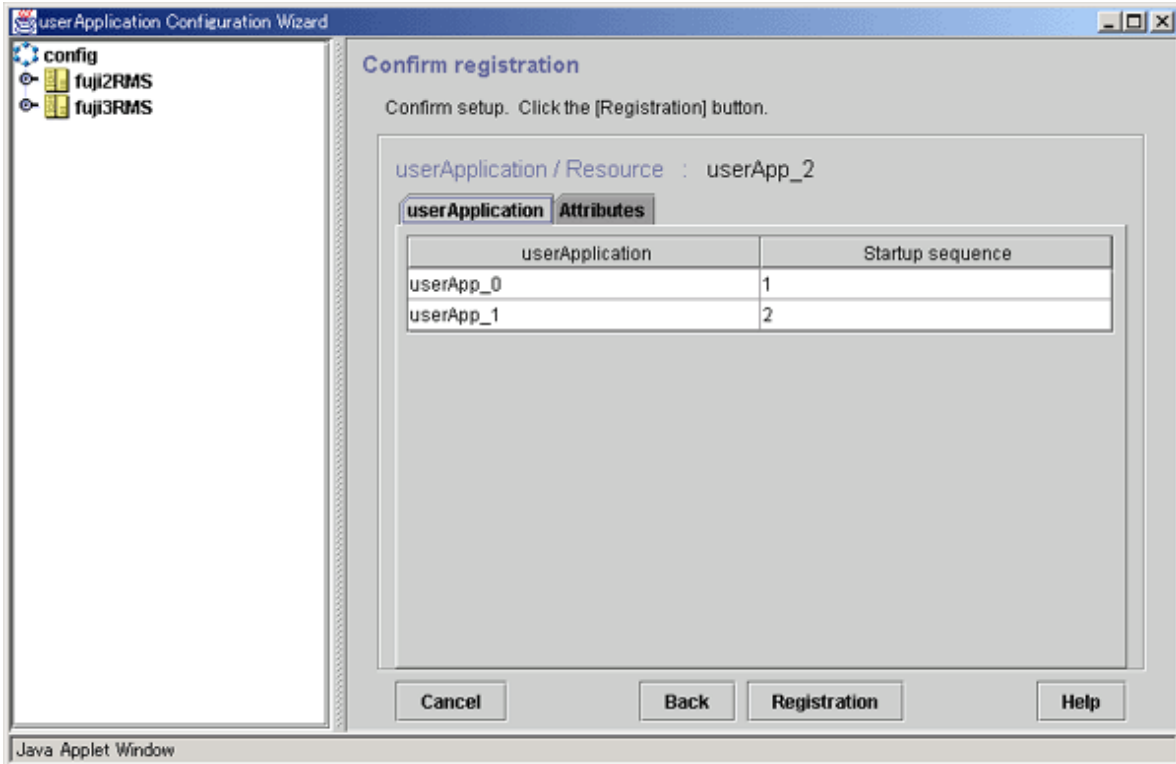
SubApplication button

This button is always disabled.

After you have confirmed the registration information, click *Registration*.

6.6.3 Setting Up Dependency Relationships Between Cluster Applications

You can set up exclusivity and startup priority as part of the RMS Configuration settings.

- **Setting up exclusive relationships between cluster applications**

  Set exclusive control so that multiple cluster applications will not run concurrently in the same cluster node.

- **Setting the startup priority of a cluster application**

  Determine the sequence in which the cluster applications are to be started.

  If you omit this setup, the startup sequence of the cluster applications will be undefined.

Note
. . . .

The startup priority setup function is not supported in this version.

. . . .

6.6.3.1 Setting Exclusive Relationships Between Cluster Applications

Set up exclusive relationships for the cluster applications to prevent cluster applications with exclusive relationships from operating simultaneously. An exclusive relationship can be set up only between standby cluster applications. If failover occurs, determine the cluster applications that should continue operating by using the job priority of the cluster applications between which an exclusive relationship is set.

The operation of cluster applications, between which an exclusive relationship is set up, during failover can be explained in the following two cases:

- When the job priorities are different

- When the job priorities are the same

When the job priorities of the cluster applications with an exclusive relationship are different

The "hvswitch(1M) -p" command is executed for a low-priority cluster application on the node where a high-priority cluster application is running, or a high-priority cluster application is switched. For the state transition of the cluster application, see the manual page of hvswitch(1M).

If a high-priority cluster application is switched to other node during a low-priority cluster application startup, the high-priority cluster application is given the highest priority for startup. The low-priority cluster application is eliminated with the "hvreset(1M)" command.

See
. . . .

For information on the hvreset(1M) command, see the manual page describing the hvreset(1M) command.

. . . .

Failover of the cluster application with a high job priority

If failover occurs in a cluster application with a high job priority, the cluster application with the high job priority will always be in the operating state.
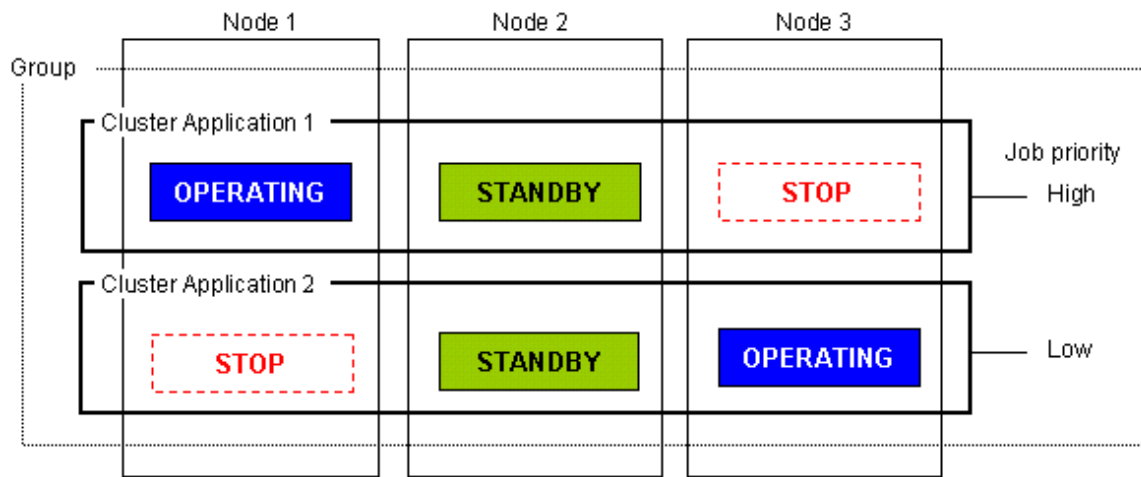
If OPERATING Node 1 using Cluster Application 1 stops, failover
is attempted to STANDBY Node 2. However, Cluster Application
2 with low job priority is STANDBY on Node 2.
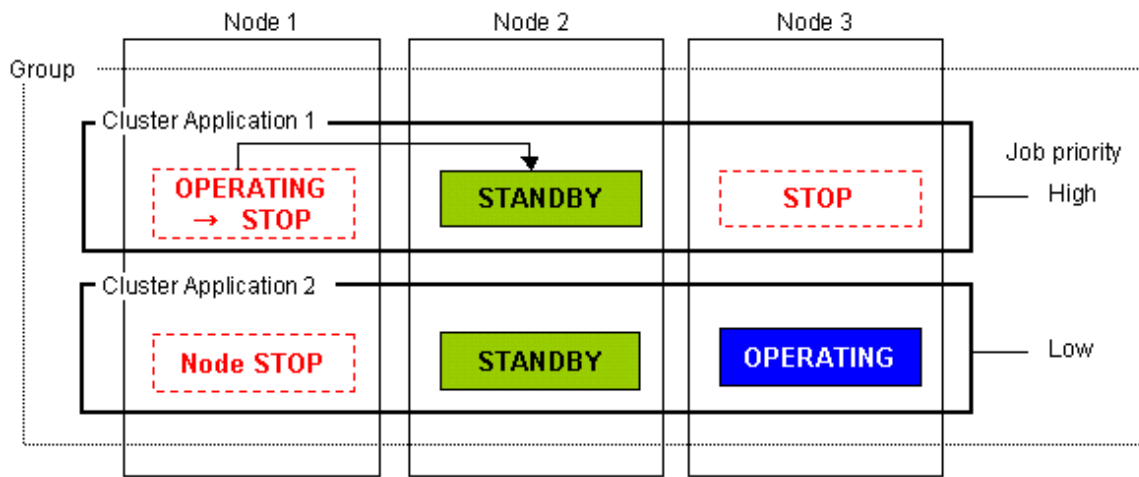
Before failover occurs for Cluster Application 1 whose job priority is high, Cluster Application 2, whose job priority is low, stops.
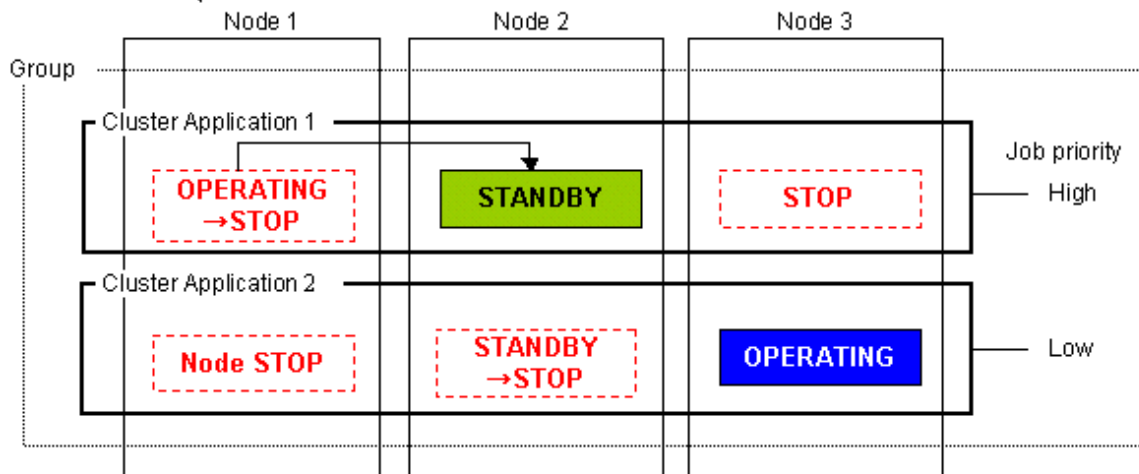
After Cluster Application 2 stops on Node 2, failover occurs for Cluster Application 1 whose job priority is high.
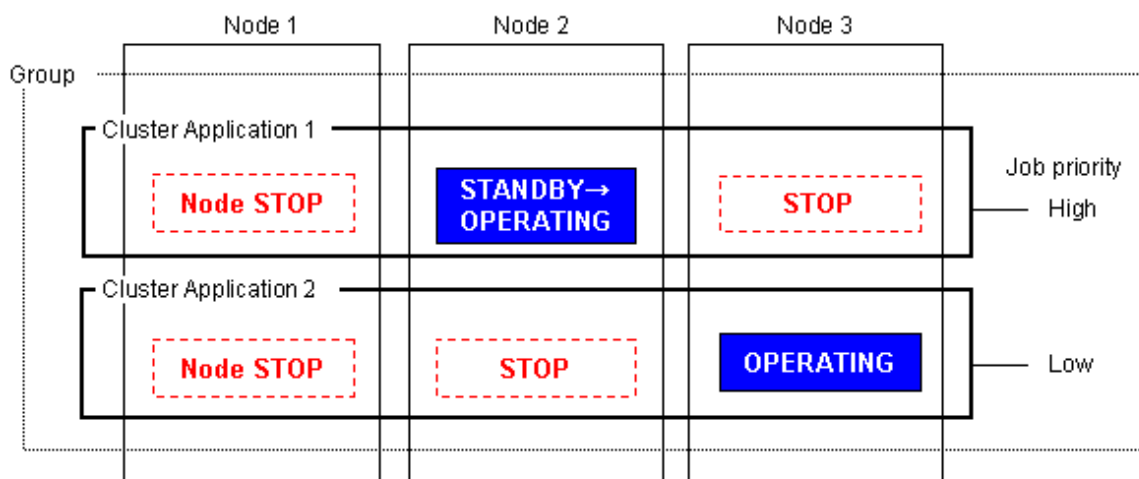
Moreover if Node 2 using Cluster Application 1 stops, failover is attempted to Node 3. However, Cluster Application 2 is OPERATING on Node 3.
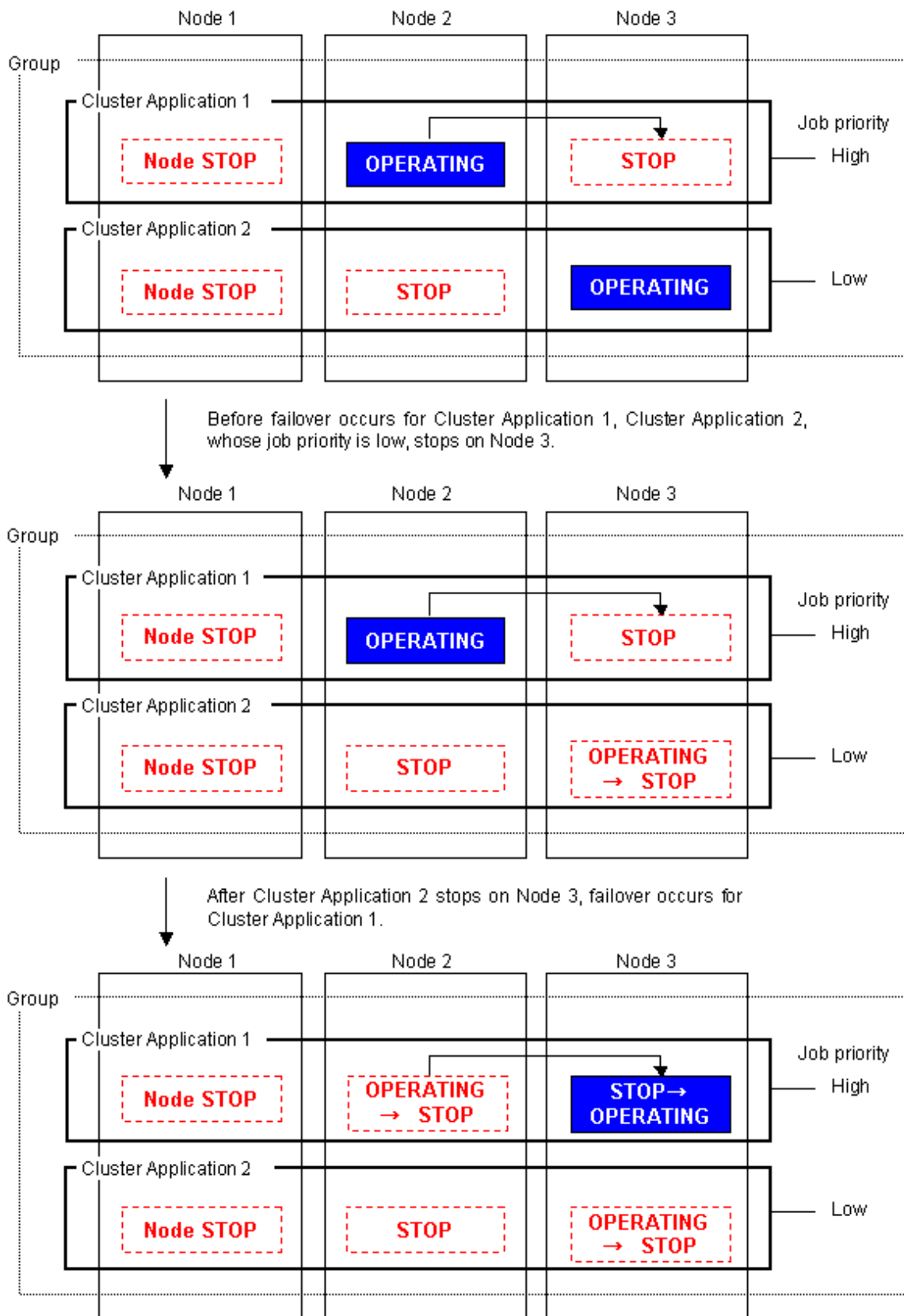
Failover of the cluster application with a low job priority

Failover occurs for a cluster application with a low job priority only when there is no cluster application with a high job priority included on the node to which the cluster application with a low job priority is to be failed over.
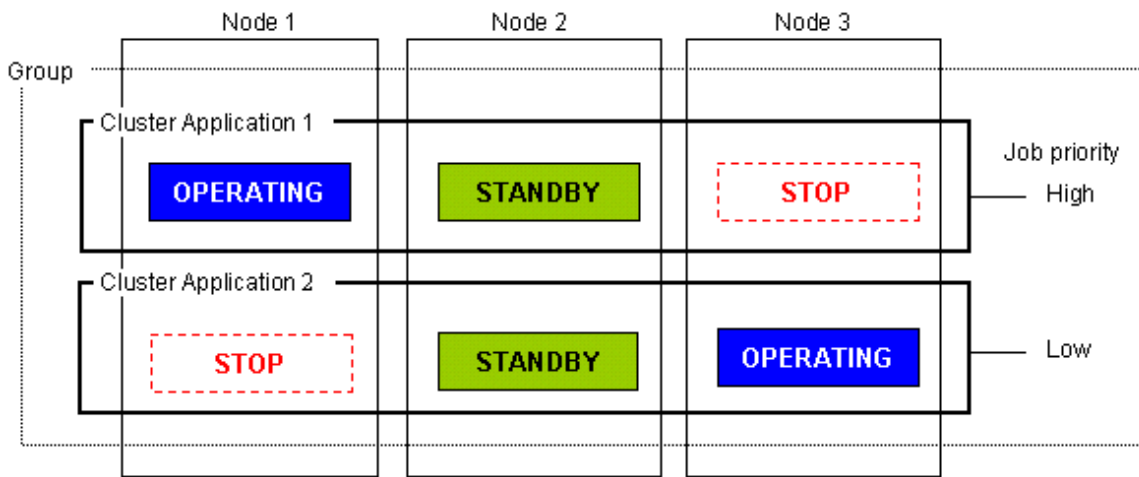
If OPERATING Node 3 using Cluster Application 2 stops, failover is attempted to STANDBY Node 2. However, Cluster Application 1 with high job priority is STANDBY on Node 2.
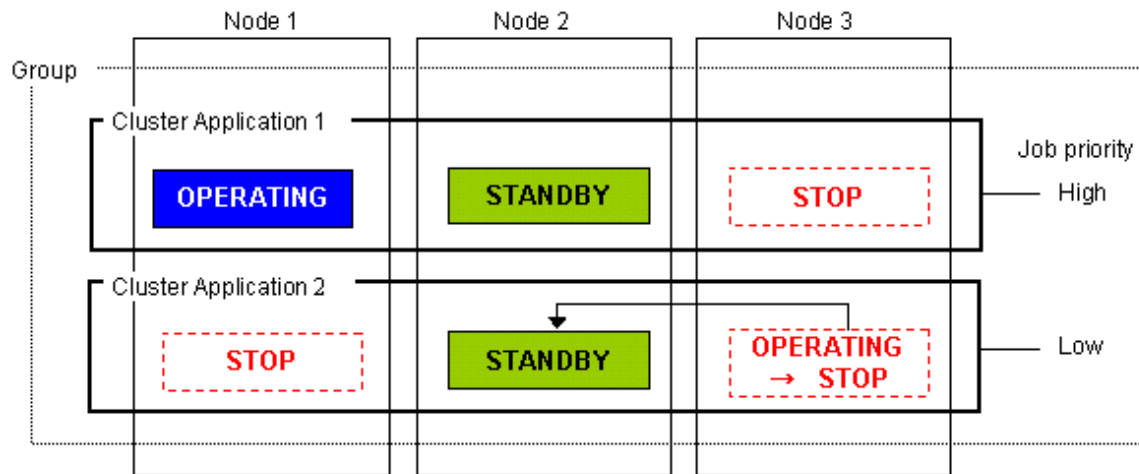
Before failover occurs for Cluster Application 2 whose job priority is low, Cluster Application 1, whose job priority is high, stops on Node 2.
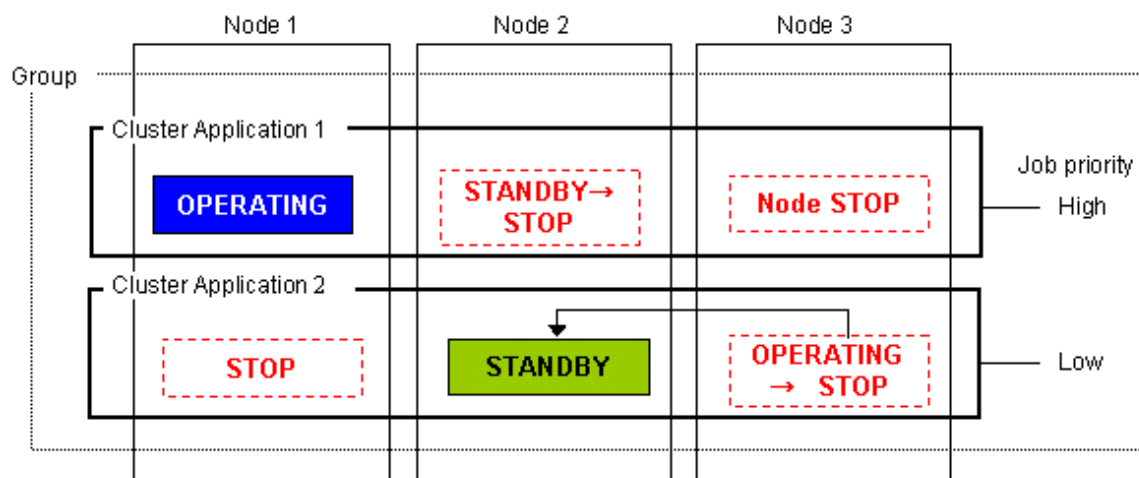
After Cluster Application 1 stops on Node 2, failover occurs for Cluster Application 2 whose job priority is low.
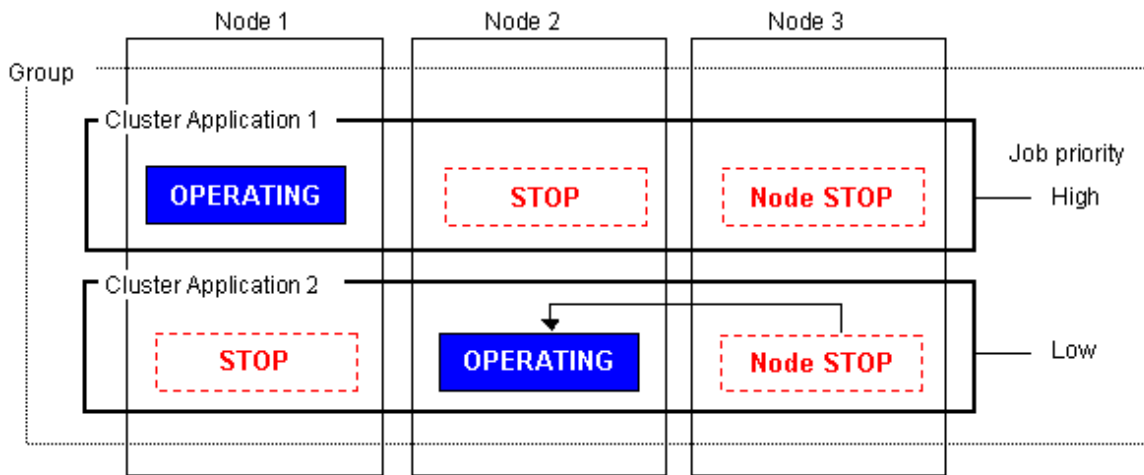
Moreover if Node 2 using Cluster Application 2 stops, failover is attempted to Node 1. However, Cluster Application 1, whose job priority is high, is OPERATING on Node 1.
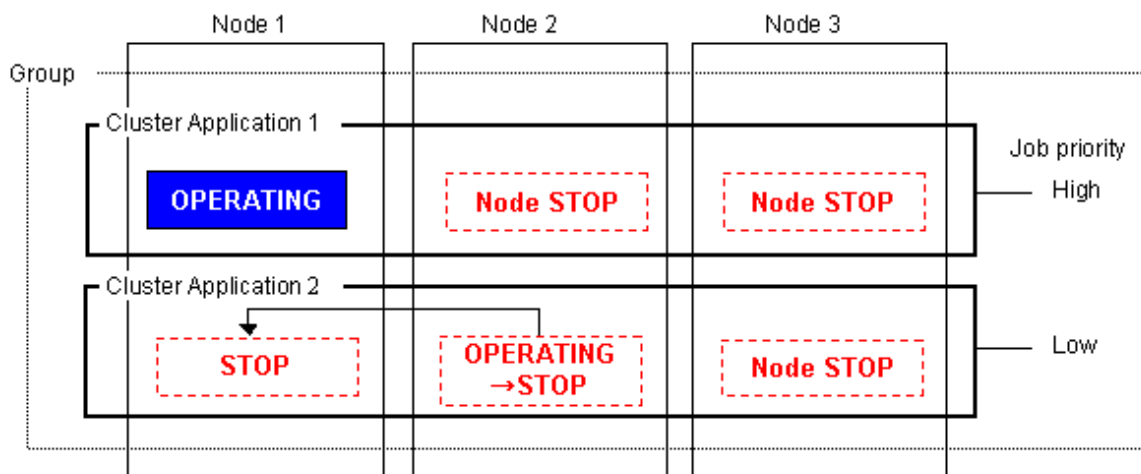
Since Cluster Application 1, whose job priority is high, is OPERATING on Node 1,Cluster Application 2, whose job priority is low, does not execute failover.
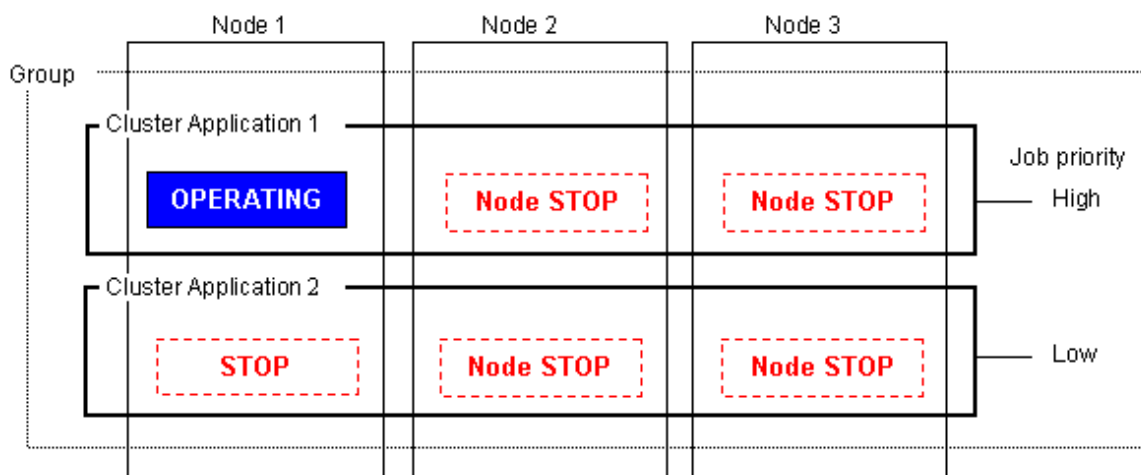
When the job priorities of cluster applications with an exclusive relationship are the same

Ongoing operations of the cluster application are never disrupted. If other cluster applications are started on the node where the cluster application has been running, or they are switched to that node, the "hvswitch(1M) -p" is executed. See the manual page of hvswitch.

After Cluster Application 2 stops on Node 2, failover occurs for Cluster Application 1.

	Node 1	Node 2	Node 3	
Group				
Cluster Application 1	Node STOP	STANDBY→OPERATING	STOP	
Cluster Application 2	Node STOP	STOP	OPERATING	Same job priority

Moreover if Node 2 using Cluster Application 1 stops, failover is attempted to Node 3. However, Cluster Application 2 is OPERATING on Node 3.

	Node 1	Node 2	Node 3	
Group				
Cluster Application 1	Node STOP	OPERATING →STOP	STOP	
Cluster Application 2	Node STOP	Node STOP	OPERATING	Same job priority
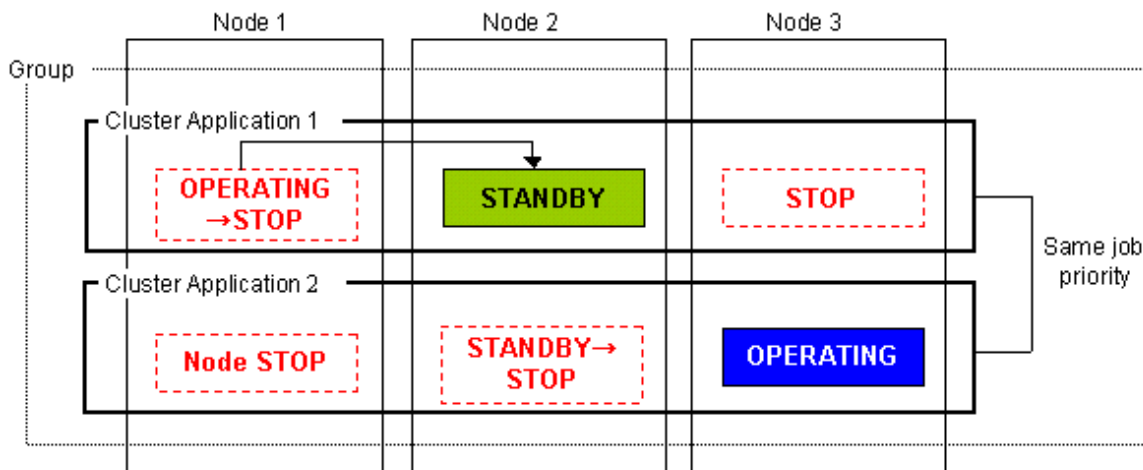
Cluster Application 2 continues to run, and Cluster Application 1 does not execute failover.

	Node 1	Node 2	Node 3	
Group				
Cluster Application 1	Node STOP	Node STOP	STOP	
Cluster Application 2	Node STOP	Node STOP	OPERATING	Same job priority
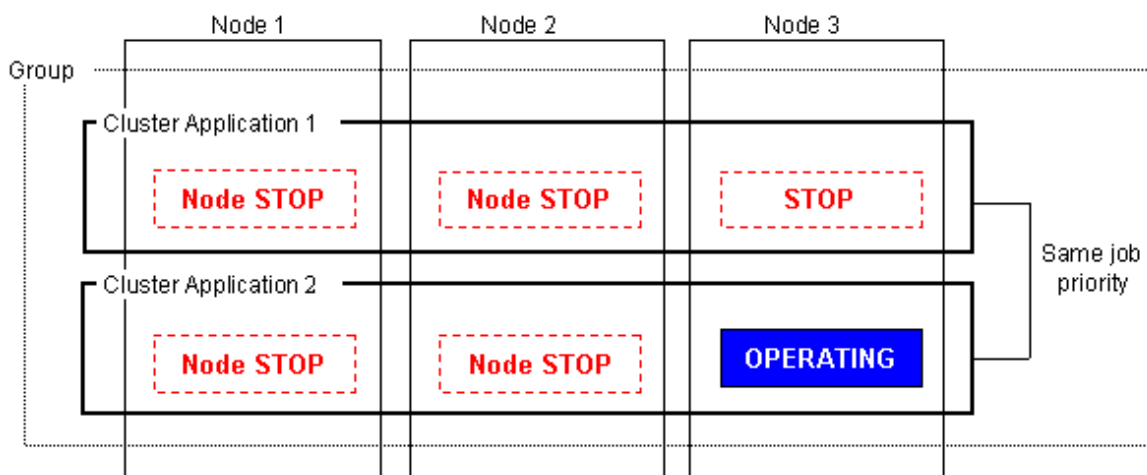
Information

To set up an exclusive relationship, create a group of cluster applications between which an exclusive relationship is to be set. Up to 52 groups can be created. Exclusive control is established between cluster applications within a single group.

Setting up an exclusive relationship

This section explains how to set up an exclusive relationship.

Figure 6.63 Flow for setting up an exclusive relationship
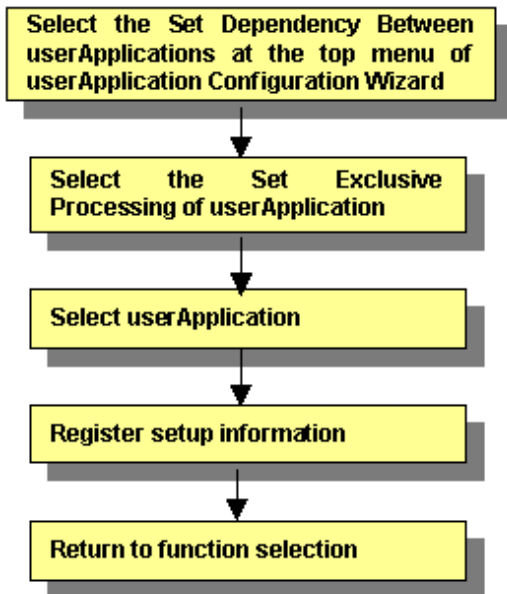
Setting up dependencies between userApplications

From the top menu of the "userApplication Configuration Wizard", select Set up dependency between userApplication.

Figure 6.64 userApplication Configuration Wizard top menu
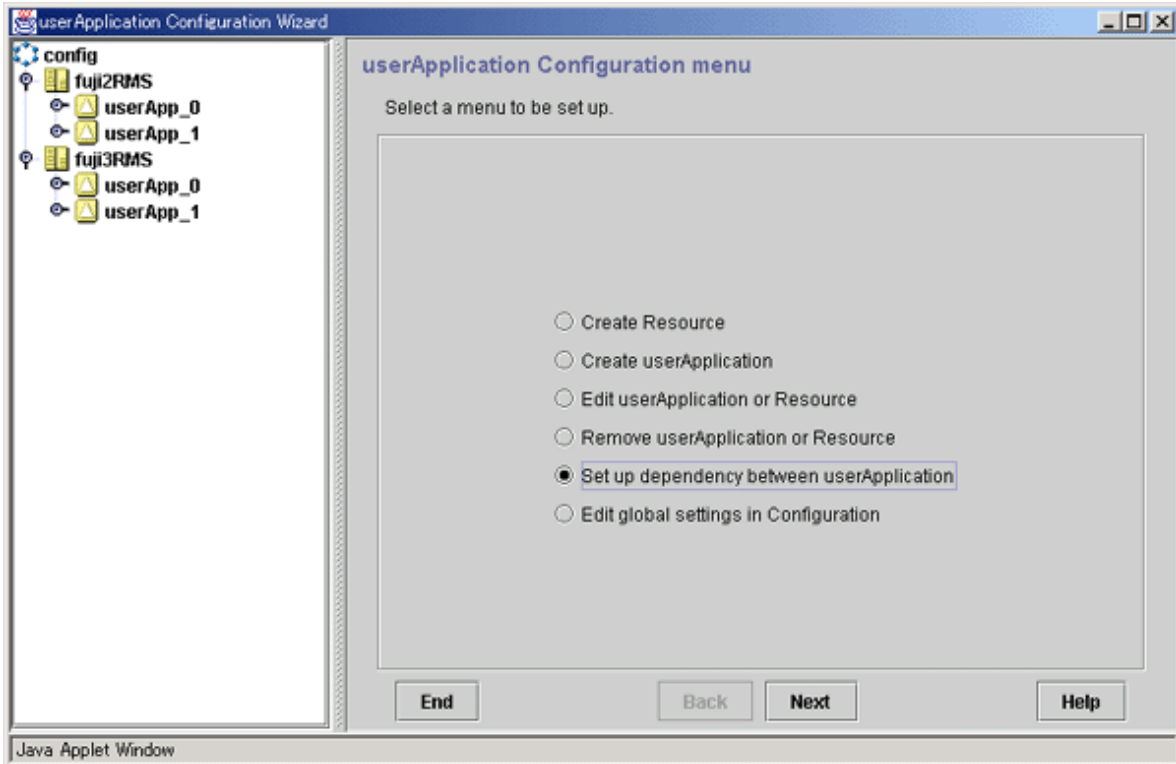
Click Next to go to the "Set up startup priority and exclusivity of userApplication" screen.

Note

The "Set up dependency between userApplication" menu can be selected only when there are two or more standby cluster applications in the cluster system.

Selecting Exclusivity of userApplication

Select Exclusivity of userApplication.

Figure 6.65 Setting up the startup priority and exclusivity of userApplication
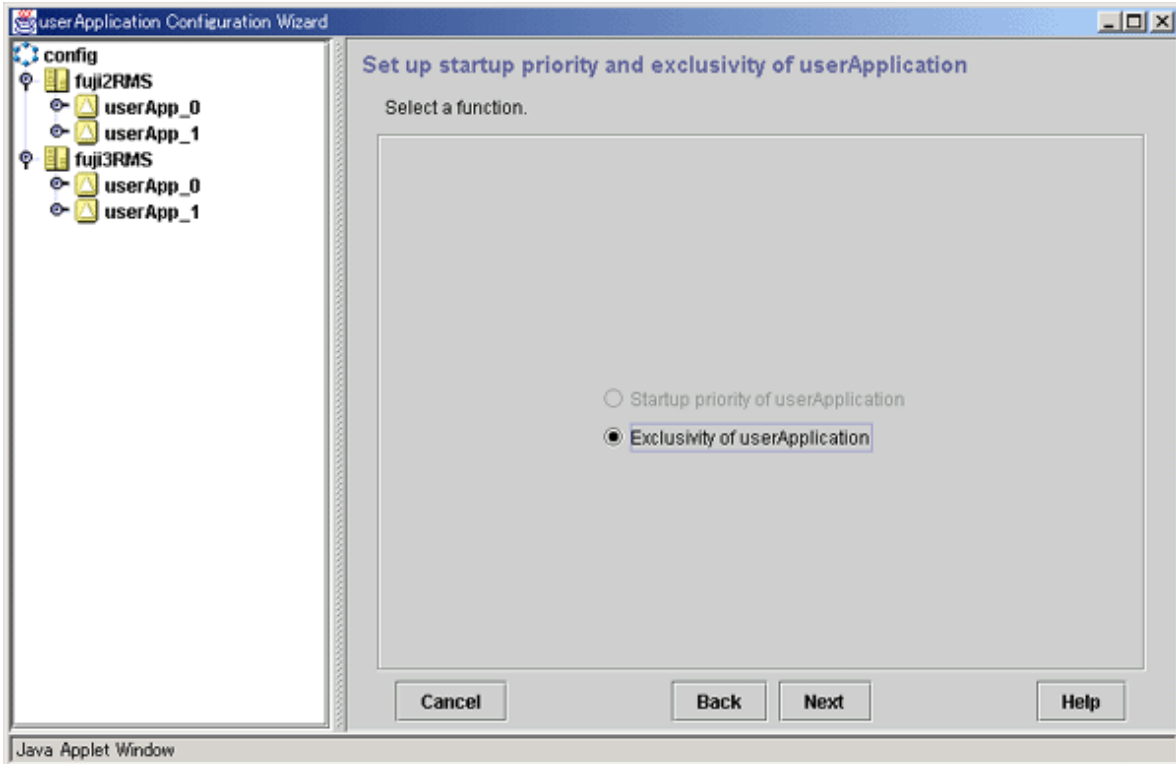
Click Next to go to the "Set up exclusivity of userApplication" screen.

Setting up exclusivity between userApplications

Create a group of cluster applications between which an exclusive relationship is to be set and then add cluster applications to the group.

Figure 6.66 Setting up exclusivity between userApplications
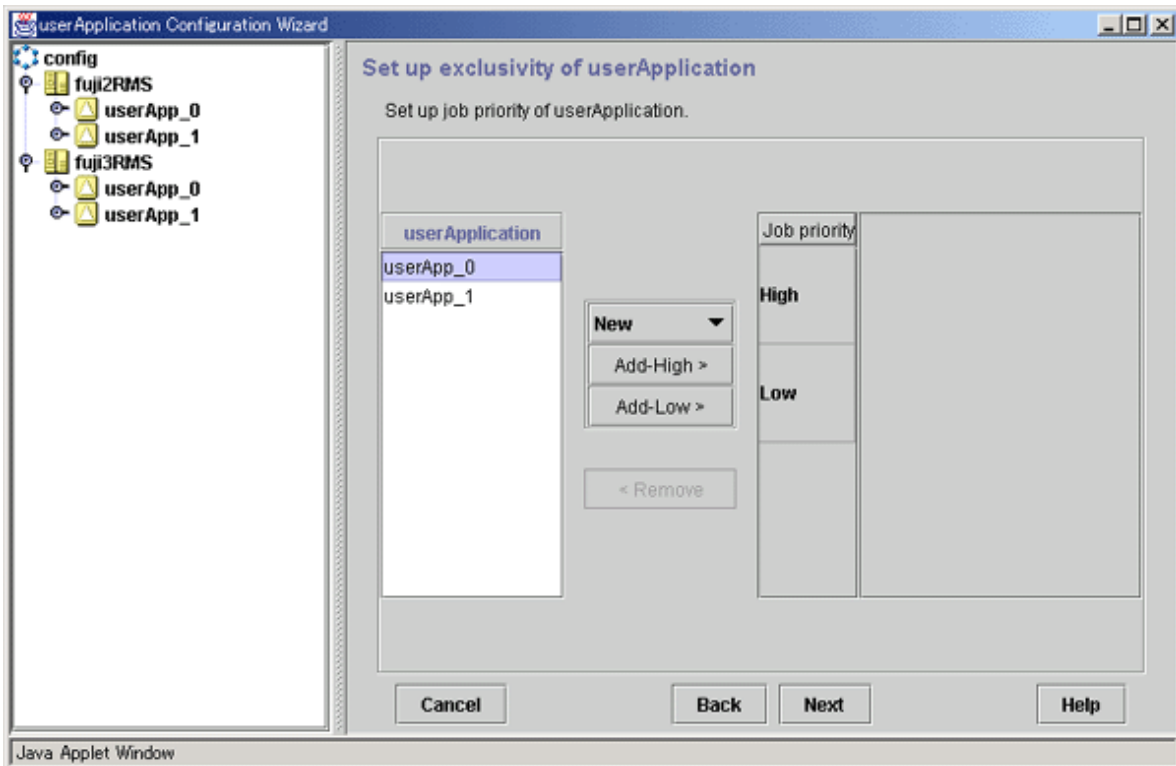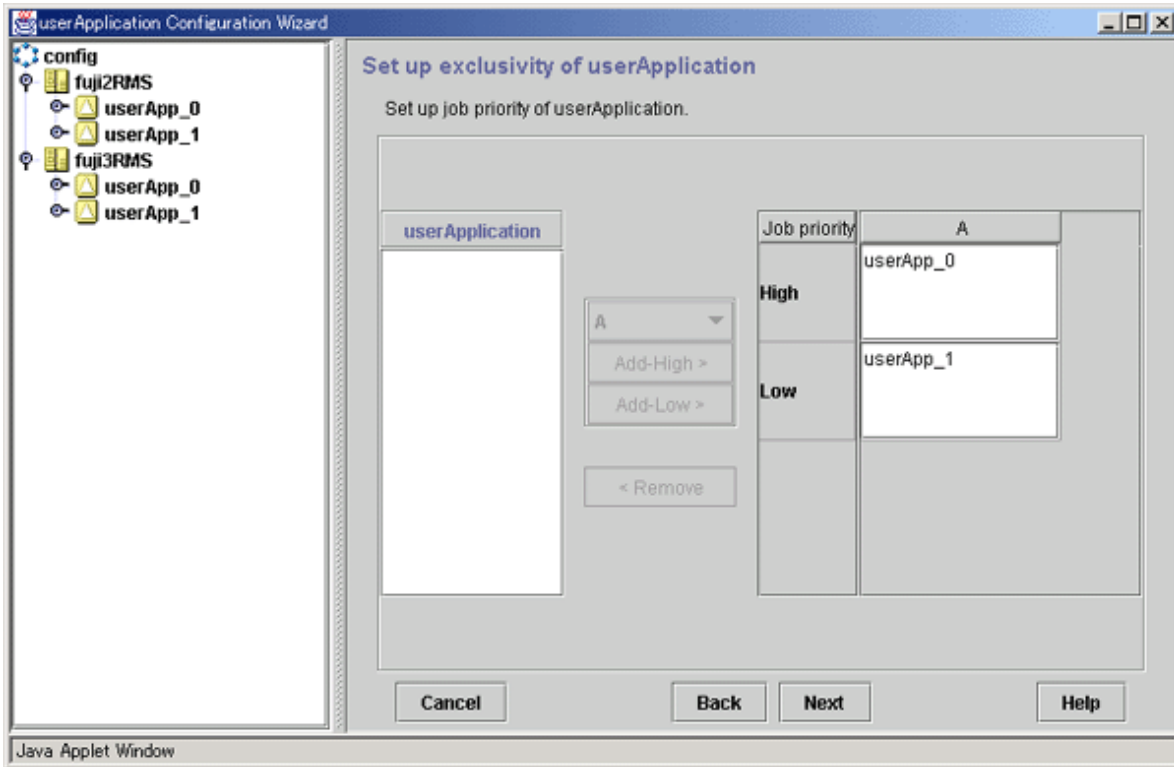
userApplication

Cluster applications are listed. A cluster application that has already been set up in either a group or a scalable cluster application will not be displayed.

Group selection box

From the list box in the center of the setup screen, select a group for which cluster applications are to be set up. If "New" is specified, a new group will be created. Up to 52 groups of A to Z and a to z can be created.
If a group has not been created, you can select "New" only.

Add-High/Add-Low buttons

Add the cluster application that you have selected from the userApplication list to the selected group. If you make this addition by using Add-High, the job priority will be set to "high" while, if you make the addition by using Add-Low, the job priority will be set to "low."
Select the addition destination group from the group selection box in advance.

Delete button

Used to cancel the exclusivity setting made for cluster applications. Select cluster applications within a group, and then click Delete. You can return the cluster applications to the [useApplication] list.

After you have made this setting, click Next to go to the "Registration of the exclusive information on userApplication" screen.

Information

If the job priorities of cluster applications are the same, the same operation will continue even when you add the cluster applications to "High" or "Low."

Confirming the registration information of exclusivity between cluster applications

Confirm the setup information for the exclusive relationship between cluster applications.

Figure 6.67 Registration of the exclusive information on userApplication
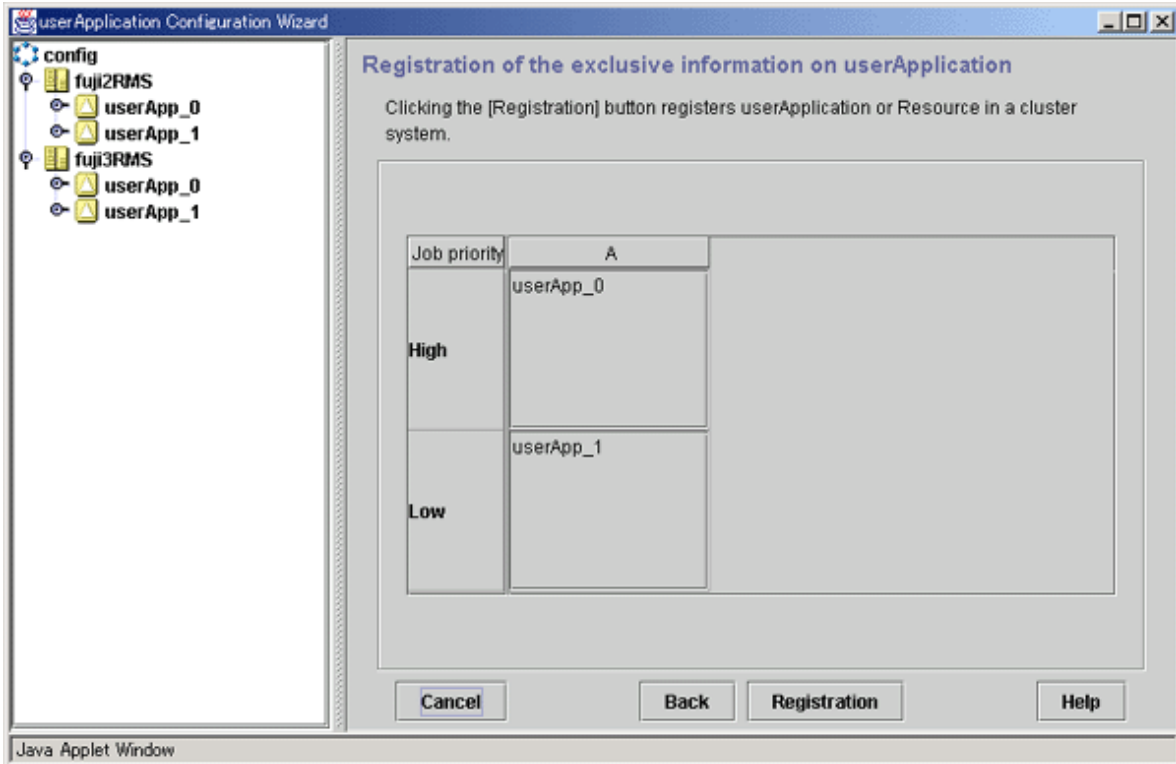
Setup information will be displayed.

After you have completed the registration information, click Registration.

6.6.4 Editing global settings in Configuration

Set global settings in the configuration for the RMS configuration.

This function is not used except to change the timeout value of PreCheck or to use NFS Lock Failover.

The procedure for changing NFS Lock Failover is described below.

Starting setup of global settings in the configuration

At the top menu of the "userApplication Configuration Wizard," select *Edit global settings in Configuration*.

Figure 6.68 Top menu of userApplication Configuration Wizard
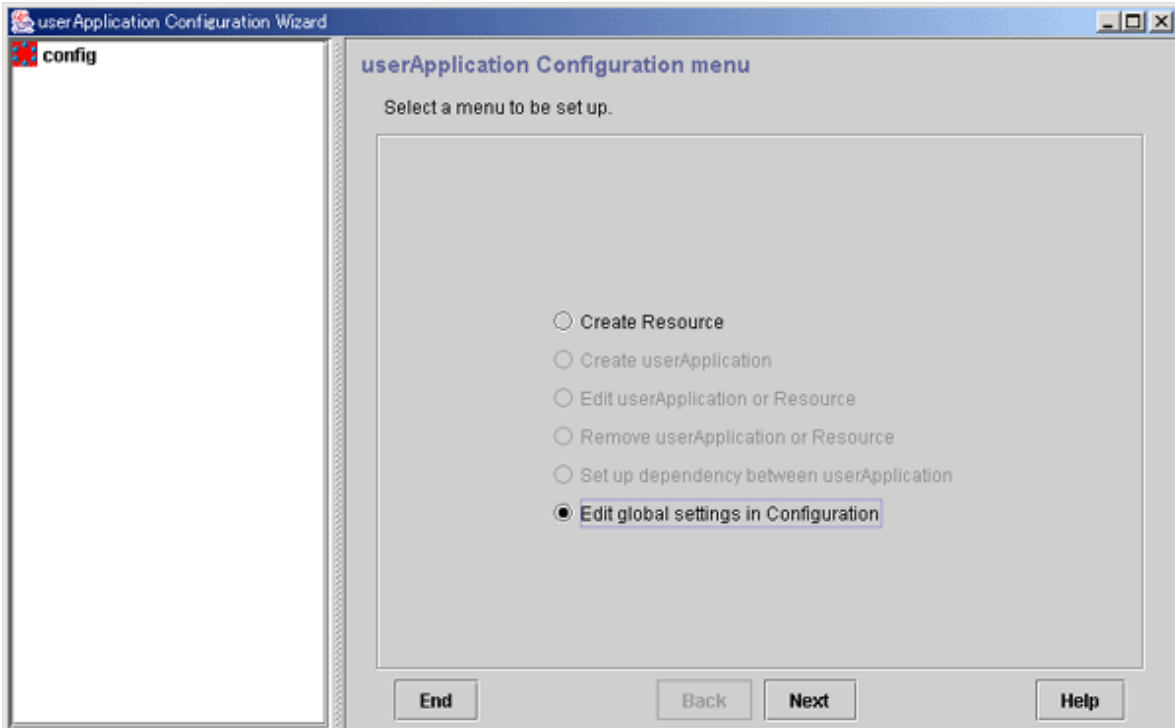
Click *Next* to go to the "Edit global settings in Configuration" screen.

Setting up global settings in the configuration

Set NFS Lock Failover to "yes."

Figure 6.69 Edit global settings in Configuration
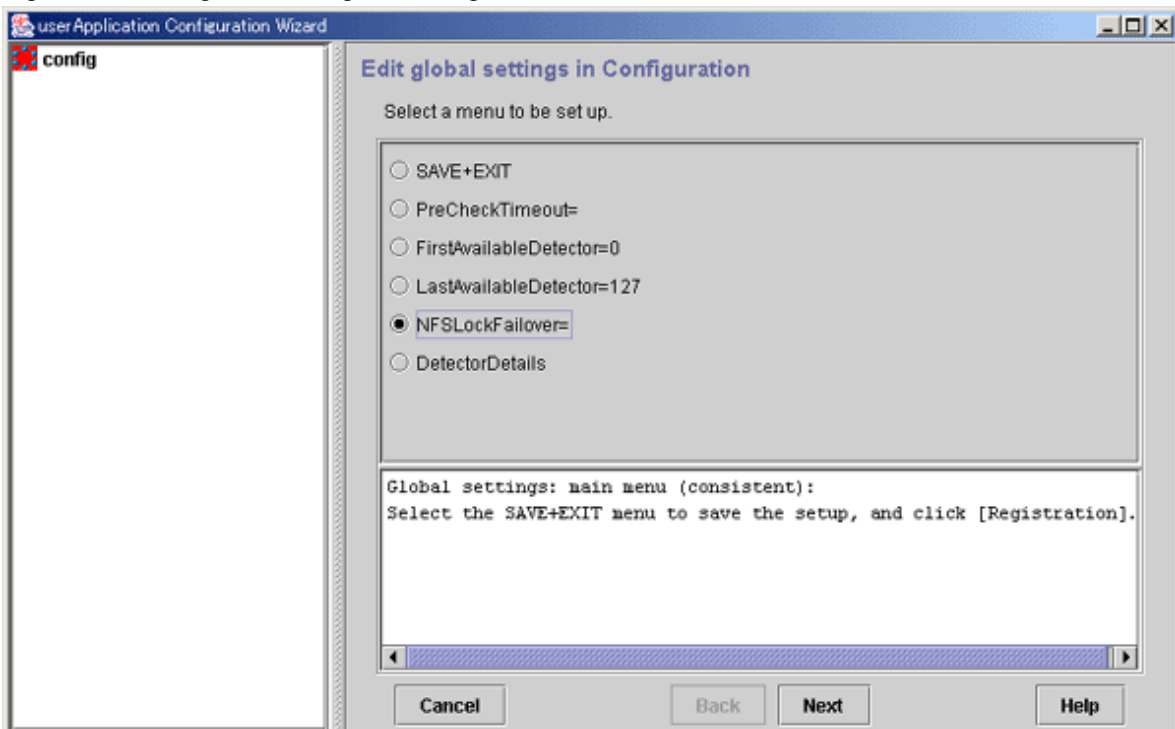
Select *NFS Lock Failover* and click *Next*.
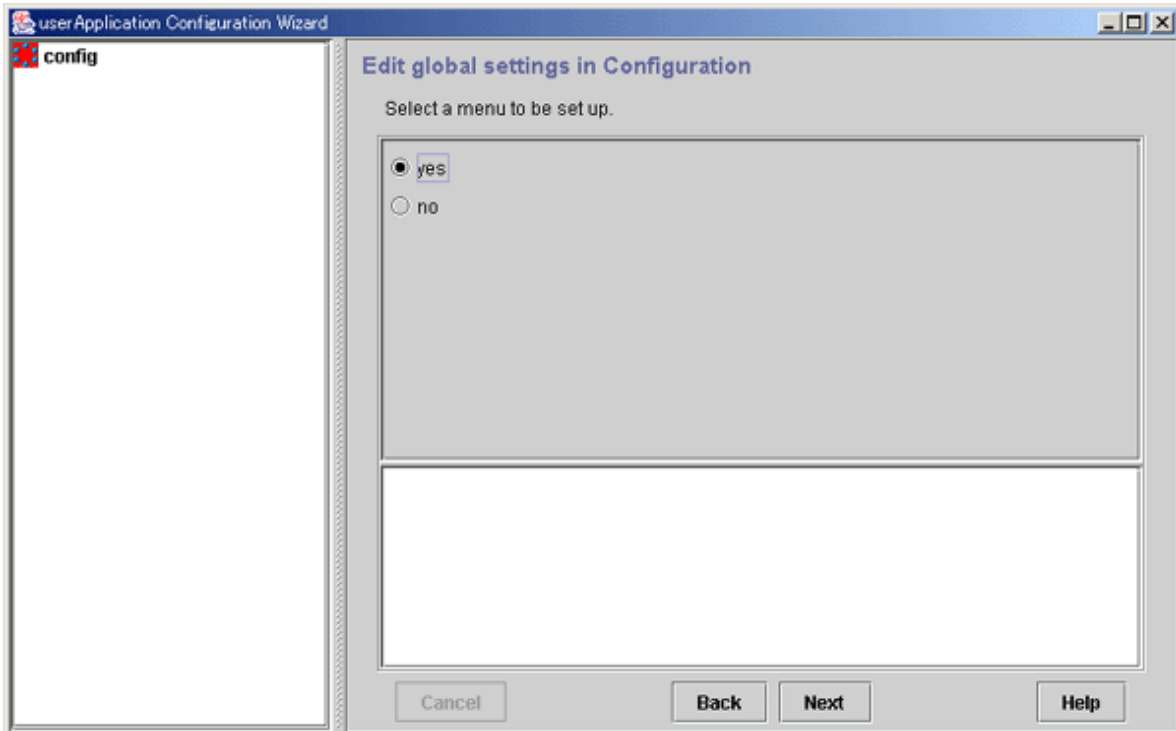
Select *yes* and click *Next* to go to the "Edit global settings in Configuration" screen.

Setting a common directory

Set a common directory.

Figure 6.70 Common directory setup
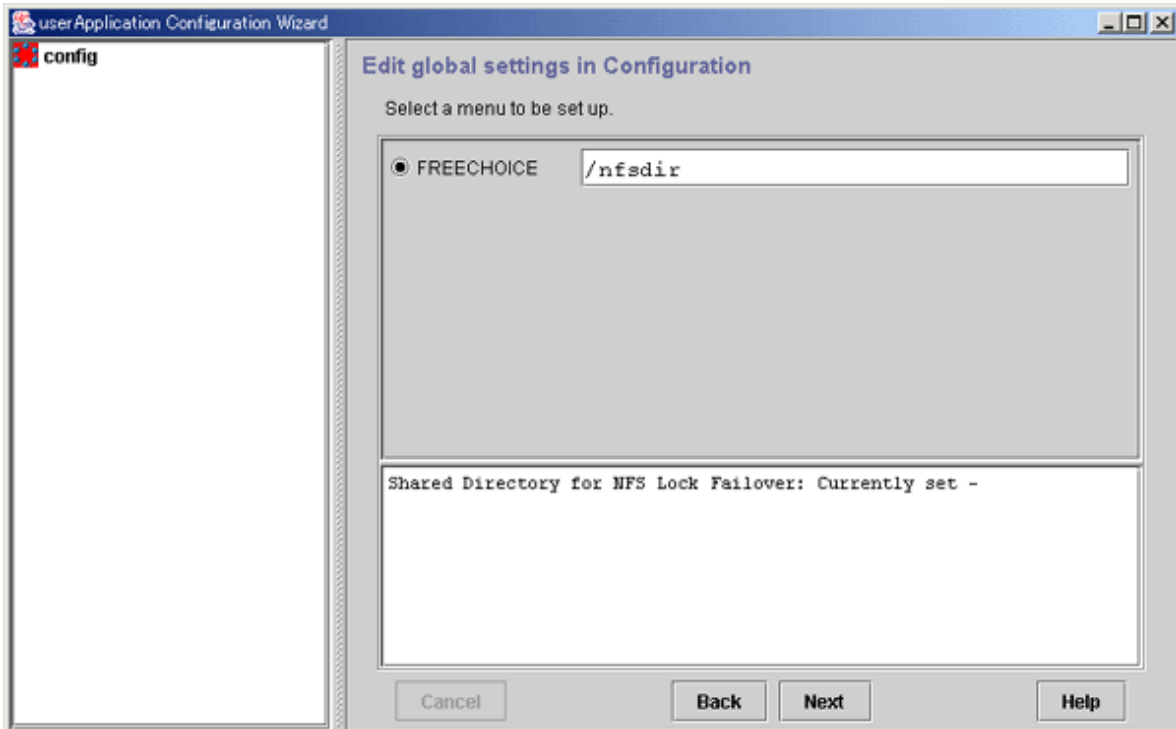
In *FREECHOICE*, enter the directory for storing NFS file lock information.

During Fsystem resource creation, the specified directory is created automatically in the file system in which the NFSLockFailover attribute was set (under the mount point).

After completing the setup, click *Next* to go to the Top of "Edit global settings in Configuration" screen.

Checking the setup information

Check that NFS Lock Failover is set to *yes*.

Figure 6.71 Top of global settings in Configuration
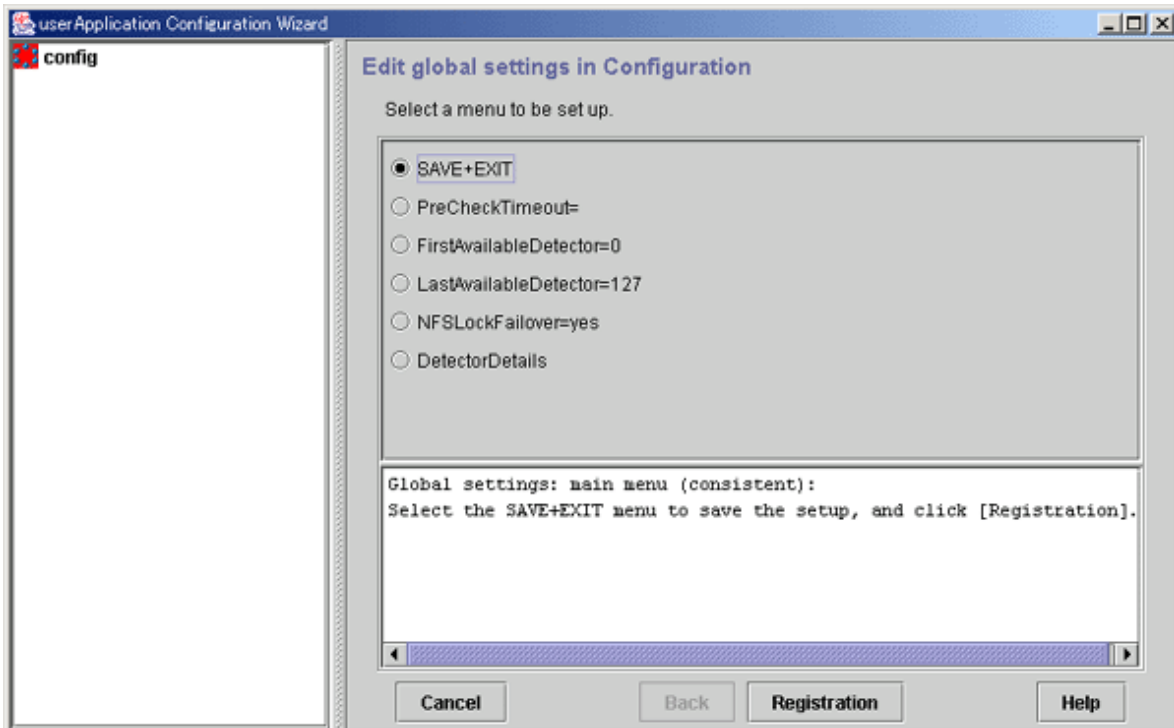
After completing the setup, select *SAVE+EXIT*, and then click the *Registration* button.

If there is one or more cluster applications and there is no inconsistency in RMS Configuration, the following message is displayed:
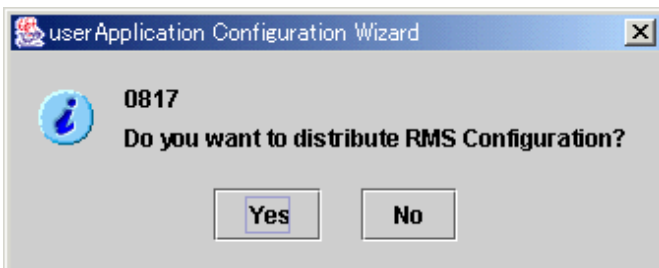
To distribute the Configuration information immediately, click *Yes* to this message.
If you still need to create or change a cluster application or resource, select *No*.

Make sure that you thoroughly verify and understand the operation of the Detector before changing *DetectorDetails*. If you are changing the settings for the first time, you need to follow the procedure described below because no default values have been set.

1. Select DetectorDetails and click the *Next* button.

2. Select "Return to Default Values" and click the *Next* button.

3. Click the *Back* button.

6.6.5 Attributes

Resource attributes

Attribute	Outline
InParallel	Normally, in Online or Offline processing, the programs for starting or stopping all the monitored programs run sequentially. If this attribute is set to yes, the programs run simultaneously. The default value is No (the programs are processed sequentially).
NeedAll	If this attribute is set to No, Online processing succeeds if there's at least one monitored Online program. If the attribute is set to "Yes", all defined monitoring programs must be Online to bring userApplication Online. The default value differs according to the resource. Cmdline: No Fsystem: Yes Takeover network: Yes Process monitoring: No
Timeout SCRIPTTIMEOUT	Sets the timeout duration (seconds) for resource start and stop processing. The default value differs according to the resource. Fsystem: 180 Gds: 1800 Gls: 60 Takeover network: 60 Procedure: 1800
AutoRecover	If this attribute is set to Yes, RMS attempts to recover the faulted resources to prevent userApplication from switching to other host. This attempt is executed only once. If this attribute is set to No, userApplication switches to other host in the event of a failure according to the AutoSwitchOver setting. The default value for Gds resources is "No."

userApplication attributes

Attribute	Outline
AutoStartUp	If this attribute is set to "Yes," the userApplication starts automatically when RMS is started. The default value is "No," and automatic userApplication startup is disabled.
AutoSwitchOver	If a failure occurs in a userApplication, this attribute allows userApplication to be switched over automatically. The switchover condition can be selected from the following: No Automatic switchover by RMS is not executed.

Attribute	Outline
	The cluster application must be switched over manually by the administrator.
	HostFailure If a failure occurs in the host (SysNode).
	ResourceFailure If a failure occurs in a resource.
	ShutDown If RMS shuts down.
	Combinations of the above values can be selected.
	Example)
	HostFailure|ResourceFailure
	HostFailure|ShutDown
	ResourceFailure|ShutDown
	HostFailure|ResourceFailure|ShutDown
	The default value is "No."
PersistentFault	If this attribute is set to 1, the Faulted state of userApplication remains the same even after RMS restart.
OnlinePriority	If this attribute is set to 1, when all clusters are stopped and restarted, RMS starts userApplication in the host that was Online last. The default value is "0," and after RMS is restarted, userApplication goes Online in the node with the highest priority. The priority is determined by the node setup sequence that was specified in "SysNode Setup" when userApplication was created.
ShutdownPriority	This attribute is a weight coefficient that is assigned to userApplication. If a failure occurs in a cluster interconnect, this attribute also becomes one of the factors in determining which node has the highest priority to survive. The higher value, the higher is the priority of that userApplication. Select a value from the range of 1 to 20. The default value is "NONE," which means no weight coefficient is assigned. The node in which this userApplication is Online may be stopped because it does not contain elements that have a high priority to survive. For details on ShutdownPriority, see "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide.*"
StandbyTransitions	Set up the condition whereby a standby resource will make the transition to the Standby state. You can select any of the following for the setup value: NONE Select this when you do not wish to set a condition for Standby state transition. ClearFaultRequest Select this when you want to clear the userApplication Faulted state and make a state transition to Offline. StartUp If AutoStartUp is set to "No," the operating userApplication will be also changed to Standby state when you start up RMS.

Attribute	Outline
	SwitchRequest The userApplication is switched to another node. You can select the following values in combination with the above values: Example) ClearFaultRequest|StartUp ClearFaultRequest|SwitchRequest StartUp|SwitchRequest ClearFaultRequest|StartUp|SwitchRequest The default is "NONE." Note) When you use the redundant line control function of GLS, set up the following value to enable monitoring of Gls resource state on the standby node: "ClearFaultRequest|StartUp|SwitchRequest"
HaltFlag	If this attribute is set to "Yes," the node is forcibly shut down when another failure occurs (double failure) during fault processing. In this case, RMS forcibly stops the node and completes failover of userApplication even if a resource failure occurs during failover processing. The default value is "No," and node elimination is not executed. Accordingly, userApplication does not execute failover. For details on HaltFlag (Halt attribute), see "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide*."

6.6.6 Changing the RMS Configuration Name

userApplication Configuration Wizard provides a command for changing the Configuration name ("config") to be used as the default.

You can change the RMS Configuration name to be used by userApplication Configuration Wizard by executing this command in the management server of Web-Based Admin View. For detailed usage instructions, see below.

wgcnfclient-RMS Configuration name setup/reference command

Function explanation

This command allows you to change the RMS Configuration name to be displayed and operated by userApplication Configuration Wizard. As much as possible, the modified information is distributed to all clusters. Normally, the RMS Configuration name does not need to be changed.

Prerequisites:

To enable this command, the following conditions must be satisfied:

- The cluster resource manager package must be installed.

  If this package is not installed, the 3-tier configuration of Web-Based Admin View is assumed, and only the information in the local node is changed. If the 3-tier configuration of Web-Based Admin View is being used, this command must also be executed in the secondary management server of Web-Based Admin View.

- Resource database installation must be complete.

  Open the Cluster Admin CRM main window, then check all the node icons are displayed and green on the CRM tree view.

Usage examples

1. Changing the RMS Configuration name

   As shown below, specify the new RMS Configuration name in the "-o" option and execute the command.

   ```
   # /etc/opt/FJSVwvucw/bin/wgcnfclient -s config-name -o config2
   #
   ```

2. Inquiring the RMS Configuration name

   ```
   # /etc/opt/FJSVwvucw/bin/wgcnfclient -g config-name
   config2
   #
   ```

   Since no value is set as the initial value of config-name, no data is displayed until you allocate one.

Supplement

If this command terminates abnormally, a system command error message may be displayed. If an error message is displayed, respond according to that message.

Note
. . . .

The RMS Configuration name must be **80 character or less**. If the RMS Configuration name has 81 or more characters, RMS Configuration generation and RMS activation will fail.

. . . .

6.7 Setting Up Patrol Diagnosis

The patrol diagnosis facility periodically diagnoses the following hardware units that are connected to the STANDBY node.

- **Shared disk units**

  The function diagnoses whether a shared disk unit has become unusable because the power is switched off, a cable is disconnected (adapter side or device side) or because of some other reason.

  If the diagnosis results indicate that an error was detected in a shared disk unit, a message is output to "syslogd(1M)" and the CRM main window.

- **Network interface cards**

  The function diagnoses whether any network interface card cannot communicate because a cable is disconnected or because of some other reason.

  If the patrol diagnosis results indicate that an error was detected in a network interface card, a message is output to "syslogd(1M)" and the CRM main window, and brings userApplication Faulted.

  If a cluster application uses a Global Link Services (GLS) resource, patrol diagnosis is disabled for the network interface card that is managed by the Gls resource. If a cluster application uses a Gls resource, use the GLS monitoring function. For details on the GLS monitoring function, see "Failure Monitoring Function" and "Standby Patrol Function" in the "*PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function.*"

When the Patrol Diagnosis Facility detects a hardware error, this error affects the Patrol Diagnosis Facility resource of a cluster application. This Patrol Diagnosis Facility resource is registered when the cluster application is created and a user enables the Patrol Diagnosis.

Set up patrol diagnosis according to the following flow:

1. Register the patrol diagnosis resource to userApplication.

2. Set up the patrol diagnosis configuration.

> 📝 **Note**
>
> - Set up patrol diagnosis after initial setup of the cluster resource manager is completed. If the patrol diagnosis setup command (clspconfig) is executed before initial setup of the cluster resource manager, the following message is displayed:
>
>   ```
>   6615: The cluster configuration management facility is not running.
>   ```
>
> - If you do not configure the patrol diagnosis configuration, the patrol diagnosis facility will not diagnose the hardware.
>
> - To view the manual pages of the individual commands, add "/etc/opt/FJSVcluster/man" to the MANPATH variable.

Registering a patrol diagnosis resource to a userApplication

To register a patrol diagnosis resource to userApplication, use the userApplication Configuration Wizard (GUI) or the RMS Wizard (CUI).

For instructions on resource registration with the userApplication Configuration Wizard (GUI), see "6.6.2 Creating Cluster Applications."

For instructions on resource registration with the RMS Wizard (CUI), see "/usr/opt/reliant/htdocs.solaris/wizards.en/wizards/CRM.htm."

Setting the patrol diagnosis configuration

Before starting RMS, set up the following configuration for patrol diagnosis by executing the "clspconfig(1M)" command on an arbitrary node.

For details on the "clspconfig(1M)" command, see the manual page for "clspconfig(1M)."

Specifying the target hardware for patrol diagnosis execution

Specify the hardware product that runs patrol diagnosis if either of the following conditions applies. This hardware is used by the specified userApplication.

- If the cluster application uses a Gds (Global Disk Services) resource

- If the cluster application uses an Ipaddress (takeover network) resource

Operation Procedure: (When the cluster application uses Gds resources)

If the cluster application is using Gds resources, specify the physical disks that were defined to the Gds resources as hardware for patrol diagnosis execution. The physical disks can be obtained by the procedure below.

1. Confirmation of the disk class

   Check the disk class to be used by the cluster application. The disk class is the class that was specified in "6.6.1.3 Creating Gds Resources."

   Example) **class0001**

2. Confirmation of the physical disks

   Check the physical disks that were defined in the disk class. The physical disks are those that were defined in "6.2.2.2 Setting Up Shared Disks." Use the class configuration setup screen or the "sdxinfo(1M)" command to check the physical disks. For information on the class configuration setup screen, see "6.2.2.2 Setting Up Shared Disks." For details on the "sdxinfo(1M)" command, see the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide.*"

   Example) **c1t4d0, mplb2048**

3. Setup of the target hardware for patrol diagnosis execution

   Set the physical disks that were defined in the disk class as hardware for patrol diagnosis execution.

   Example 1) For c1t4d0

   ```
   # /etc/opt/FJSVcluster/bin/clspconfig -u app1 TARGET_DISK=c1t4d0 TARGET_DISK=c1t4d1
   ```

   Example 2) For mplb2048

   ```
   # /etc/opt/FJSVcluster/bin/clspconfig -u app1 TARGET_DISK=mplb2048
   ```

Operation Procedure: (When the cluster application uses Ipaddress resources)

If the cluster application is using Ipaddress resources, specify the network interface card to be used in the takeover network as hardware for patrol diagnosis execution. Follow the procedure below to determine the network interface card.

1. Confirmation of the network interface card

   Check the network interface card to be used by the takeover network. The network interface card is the interface card that was selected in "Selecting an interface" of "6.6.1.5 Creating Takeover Network Resources."

   Example) hme0

2. Setup of target hardware for patrol diagnosis execution

   Configure the network interface card to be used by the takeover network as hardware for patrol diagnosis execution.

   Example)

   ```
   # /etc/opt/FJSVcluster/bin/clspconfig -u app1 TARGET_LAN=hme0
   ```

Time interval for patrol diagnosis

Specify the time interval for patrol diagnosis. If this setting is not specified, patrol diagnosis runs at 360-minute intervals.

Example) To run patrol diagnosis at 60-minute intervals

```
# /etc/opt/FJSVcluster/bin/clspconfig -u app1 INTERVAL=60
```

You can also check the current setting.

Example) Checking the current setting (INTERVAL=360)

```
    # /etc/opt/FJSVcluster/bin/clspconfig -u app1
ACTIVE=true
INTERVAL=360
TARGET_LAN=hme1
TARGET_DISK=c1t4d0
TARGET_DISK=c1t4d1
TARGET_DISK=c1t4d2
TARGET_DISK=c1t4d3
```

Stopping and restarting patrol diagnosis

Stop and restart patrol diagnosis by executing the "clspconfig(1M)" command. For details on the "clspconfig(1M)" command, see the manual page for "clspconfig(1M)."

To stop patrol diagnosis, specify "ACTIVE=false" as an operand in the "clspconfig(1M)" command.

Example)

```
# /etc/opt/FJSVcluster/bin/clspconfig -u app1 ACTIVE=false
```

To restart patrol diagnosis, specify "ACTIVE=true" as an operand in the "clspconfig(1M)" command.

Patrol diagnosis is restarted from the next time with the same configuration as when patrol diagnosis was stopped, including hardware for executing patrol diagnosis and the time interval for patrol diagnosis.

Example)

```
# /etc/opt/FJSVcluster/bin/clspconfig -u app1 ACTIVE=true
```

Part 3　Operations

Chapter 7 Operations

This chapter describes the functions that manage PRIMECLUSTER system operations.

The following user groups are allowed to do each specific operation:

Operation	Target
Referencing the operation management screens	All user groups
Operations	wvroot, clroot, cladmin
Monitoring	All user groups
Corrective actions for resource failures	wvroot, clroot, cladmin

7.1 Viewing the PRIMECLUSTER System Operation Management Screens

PRIMECLUSTER provides GUIs for viewing and performing cluster system operations.

- CF main window

  Use this screen to set up the configuration of the nodes that make up the cluster, manage the nodes, and display the node state.

- CRM main window

  Use this screen to identify faulted hardware and to execute hardware diagnosis.

- RMS main window

  Use this screen to monitor the state of the cluster system and to manage cluster applications and resources.

- MSG main window

  This screen displays messages for cluster control.

See
. . . .
For instructions on displaying each screen, see "4.5.3 Cluster Admin Functions."
. . . .

7.1.1 CF Main Window

The CF main window allows you to set up the configuration of the cluster nodes in the cluster, manage the nodes, and display the node state.
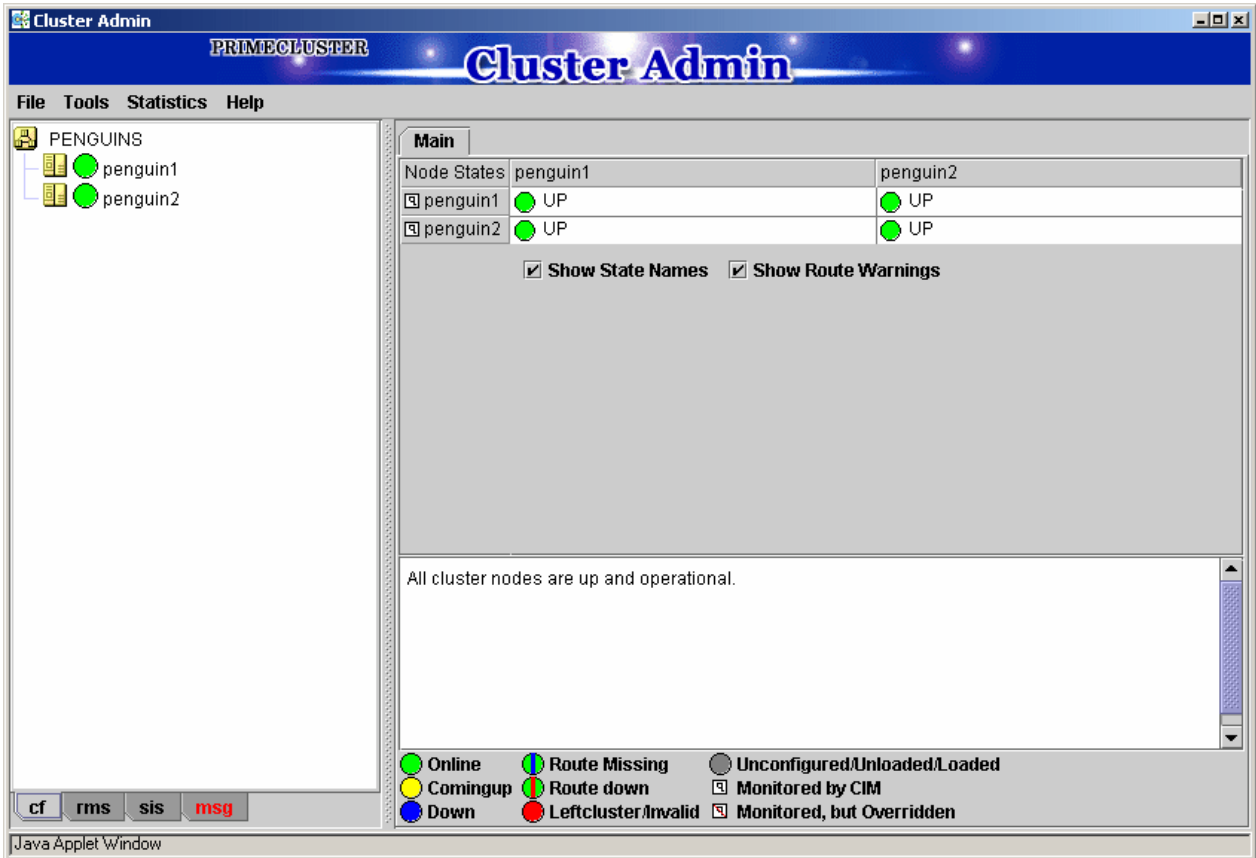
![See]

For details, see "5 GUI administration" in "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*."

7.1.2 CRM Main Window

When you monitor the state of the PRIMECLUSTER system, the CRM main window allows you to view the states of the resources managed by the cluster resource management facility and identify failed hardware.

This section describes resource states and operation methods related to hardware resources.
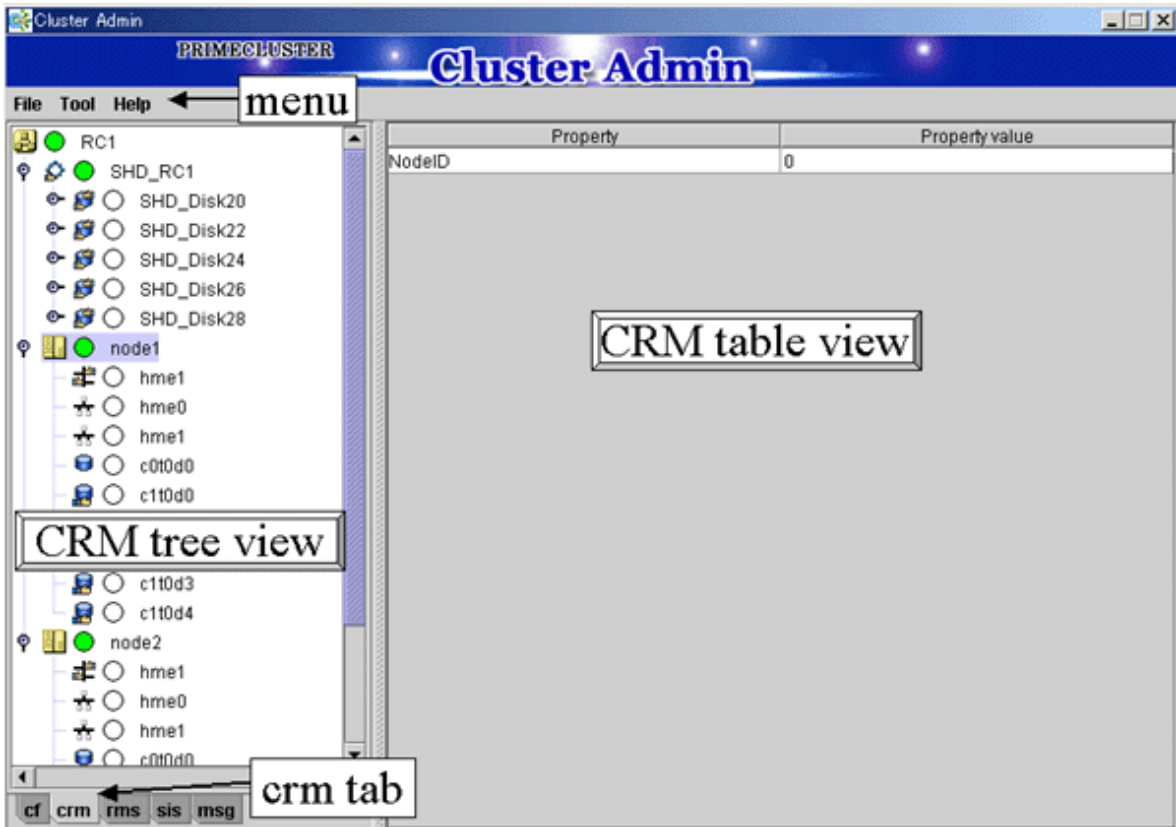
7.1.2.1 Displayed Resource Types

This section shows the resources associated with the hardware resources that are displayed in the CRM main window.

The detailed resource information lists the icons that are used in the CRM main window.

7.1.2.1.1 Resource Icons

This section describes the icons that are displayed in the CRM tree view.

Items that are related to resources under shared resources are displayed with overlapping .

Icon	Resource
	Shared resource
	DISK class managed by Global Disk Services
	Local disk Multipath disk
	Shared disk device Multipath disk
	IP address MAC address Network interface

Icon	Resource
	Takeover node name
	Node resource of a switching line used in a line switching unit
	Shared resource of a line switching unit
	Shared resource of a switching line used in a line switching unit
	Takeover network
	Resource that is not a multi-tiered resource under a shared resource
	Resource that is not related to a shared resource under a node and is not a multi-tiered resource
	Cluster
	Node

7.1.2.1.2 Resource States

In the CRM main window, the displayed icon types and states differ for each resource class.

For each resource, an icon that indicates a failure (OFF-FAIL or ON-FAILOVER) is displayed if there is a fault in a subordinate resource. In the CRM tree view, expand the icons sequentially from cluster icon to node icons to subordinate resource icons, and identify the faulted resource. If the resource state is OFF-FAIL or ON-FAILOVER, see "7.4 Corrective Actions for Resource Failures," and take corrective actions.

If the resource state is ON, OFF-STOP, or UNKNOWN, you don't need to take any corrective actions.

Cluster states

The following cluster states are displayed.

Icon	Icon color	Outline	Details
	⬤Green	ON	All nodes and shared resources are operating normally.
	⬤Red	OFF-FAIL	One of the nodes in the state other than the ON state, or a shared resource is in the OFF-FAIL state.

Node states

The following node states are displayed.

Icon	Icon color	Outline	Details
	⬤Green	ON	The node has been started normally.
	⬤Green with vertical red lines	ON-FAILOVER	One of the resources under the node is in the Faulted state.
	⬤Blue	OFF-STOP	The cluster resource management facility is stopped.
	⬤Red	OFF-FAIL	A failure has occurred in the node.
	◯White	UNKNOWN	The node has not been monitored or controlled.

If a failure occurs in the node, the following icons are displayed:

Icons	Outline	Details
✕	Error	Immediate replacement is necessary.

Icons	Outline	Details
⚠	Warning	Life replacement or preventive replacement is necessary.

If either icon is displayed, see the "*Machine Administration Guide*," and take corrective actions.

📝 **Note**

. . . .

- If a node is LEFTCLUSTER in CF, it becomes ON in the CRM tree view.

- After you execute the "clinitreset(1M)" command to initialize the resource database, the resource states displayed in the CRM main window differ according to the actual resource states. Execute "clinitreset(1M)", reboot the nodes, and then close the browser and redisplay the screen. Execute the "clinitreset(1M)" command after removing a cluster application.

. . . .

Shared resource states

The following states are displayed for shared resources.

Icon	Icon color	Outline	Details
🔧	🟢Green	ON	The sub-resource under the shared resource is ON, OFF-STOP, or UNKNOWN.
	🔴Red	OFF-FAIL	One of the sub-resources under the shared resource is OFF-FAIL.

Other resource states

The following states are displayed for other resources.

Icon color	Outline	Details
🟢Green	ON	The resource is operating normally.
🟢Green with vertical red lines	ON-FAILOVER	The resource is operating normally, but some devices or resources that are multiplexed and managed internally are in the Faulted state.
🔵Blue	OFF-STOP	The resource has been stopped normally.
🔴Red	OFF-FAIL	The resource is in the Faulted state.
⚪White	UNKNOWN	The resource has not been monitored or controlled.

7.1.2.1.3 Operations

You can perform the operations described below from the CRM main window.
In the table below, "Selection resource" is the resource class name of the selectable resource. For details on resource class names, see "7.1.2.2 Detailed Resource Information."

Table 7.1 Operations of the CRM main window

Feature	Operation method		Target group
	Menu	Selection resource	
Build CRM resource database	Tool - Initial setup	None (*1)	wvroot clroot
Configure automatically	Tool - Automatic configure	Cluster	wvroot clroot

Feature	Operation method		Target group
	Menu	Selection resource	
Run patrol diagnosis immediately	Tool - Diagnosis (*2)	DISK (*3) Ethernet (*3) MPDisk(*3)	wvroot clroot cladmin
Request Resource activation	Tool - Start	SDX_DC (*4)	wvroot clroot cladmin
Request Resource deactivation	Tool - Stop	SDX_DC (*4)	wvroot clroot cladmin
Display machine information	Tool - Machine Information	Node (*5)	wvroot cladmin clmon
Exit Cluster Admin screen	File - Exit	All No selection	All
View Help	Help - Content (*6)	All No selection	All
View version	Help - About	All No selection	All

*1 Set Initial Configuration menu can be selected only if the resource database has not been set up. This menu item is not displayed in the pop-up menu.

*2 This feature is enabled only for the resources for which the patrol diagnosis has been set.

*3 Only the resources of the DISK, Ethernet, and MPDisk classes, which have been registered in hardware where the patrol diagnosis facility runs, are enabled. For the hardware settings, see "6.7 Setting Up Patrol Diagnosis."

*4 Only the disk resources that are registered to Global Disk Services are enabled.

*5 This feature is enabled only if the Enhanced Support Facility is installed.

*6 Help for the CRM main window is displayed with a separate browser from the browser that displays Help for CF, RMS, and SIS.

Note

- Only available menus are displayed in the pop-up menu.

- If there are no available menu items for the resource selected from the CRM tree view, "None" is displayed in the pop-up menu. Selecting this item will not initiate any operation.

- For information about user groups, see "4.3.1 Assigning Users to Manage the Cluster."

Initial setup

Select this item to set up the resource database to be managed by the cluster resource management facility. Select *Tool* -> *Initial setup* to display the Initial Configuration Setup screen. The initial configuration setup cannot be operated simultaneously from multiple clients. See "5.1.3.1 Initial Configuration Setup."

Automatic configure

Select this item to register the devices connected to the system to the resource database. Automatic configuration cannot be operated simultaneously from multiple clients. See "5.1.3.1 Initial Configuration Setup."

Diagnosis

Select this item to diagnose the selected resource. Execute this diagnosis operation to check whether faulted hardware has been properly recovered after hardware recovery is performed. See "7.4.2 Corrective Action when Patrol Diagnosis Detects a Fault."

- *Yes* **button**

Executes resource diagnosis.

- *No* **button**

Does not execute resource diagnosis.

Start

This menu item activates the selected resource. The start operation is executed during maintenance work. If the selected resource is registered to a cluster application, the start operation can be executed only when that cluster application is in the Deact state. Use the RMS main window to check the cluster application state.

Note

- After completing the maintenance work, be sure to return the resource that you worked on to its state prior to the maintenance.

- If the resource that was maintained is registered to a cluster application, be sure to stop the resource before clearing the Deact state of the application.

- *Yes* **button**

Executes resource start processing.

- *No* **button**

Does not execute resource start processing.

Stop

This menu item deactivates the selected resource. The stop operation is executed during maintenance work. If the selected resource is registered to a cluster application, the startup operation can be executed only when that cluster application is in the Deact state. Use the RMS main window to check the cluster application state.

Note

- After completing the maintenance work, be sure to return the resource that you worked on to its state prior to the maintenance.

- If the resource that was maintained is registered to a cluster application, be sure to stop the resource before clearing the Deact state of the application.

- *Yes* **button**

Executes resource stop processing.

- *No* **button**

Does not execute resource stop processing.

Machine Information

This menu item displays the machine management screen of the selected node.
This menu item can be used only if Enhanced Support Facility (ESF) is installed.
For details on the machine management screen, see the "*Machine Administration Guide*."

Note

If a message is displayed during operating at the CRM main window and the frame title of the message dialog box is "Cluster resource management facility," then see the following references for information:

- PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide

- 12.10 Resource Database messages

・・・

7.1.2.2 Detailed Resource Information

This section describes the resource attributes that are defined in the CRM main window.

Table 7.2 Detailed resource information

Icon/ resource class name	Attributes	Meaning/attribute value (Top: Meaning, Bottom: Attribute value)
Node	NodeID	Node identifier number.
		Node identifier number (0 to 127).
SH_SWLine	lsu_mask	Mask information for the LSU used by the line switching unit
		Logical sum of the mask information for the LSU used by the line switching unit LSU machine information LSU00 0x0001 LSU01 0x0002 LSU02 0x0004 ... LSU14 0x4000 LSU15 0x8000
SH_SWU	lsu_num	Number of switchover units installed in the line switching unit
		4, 16
	addr	RCI device address of the line switching unit
		RCI device address
SWLine	Port	Port number of the LSU (group) used by the line switching unit
		0, 1
DISK	Disk_Attr	This item indicates the physical connection mode and usage mode of a disk that can be used from the cluster system.
		LOCAL Local disk that can be accessed only from one node **SHD_DISK** The disk is physically shared, but the usage mode (shared disk or switchover disk) is not specified. **SHD_SHARE** Shared disk that can be accessed from multiple nodes **SHD_SWITCH** Switching disk that is used exclusively between two nodes **Blank** The disk connection or usage mode is not set up.
MPDisk	Disk_Attr	This item indicates the physical connection mode and usage mode of a disk that can be used from the cluster system
		LOCAL Local disk that can be accessed only from one node **SHD_DISK** The disk is physically shared, but the usage mode (shared disk or switchover disk) is not specified. **SHD_SHARE** Shared disk that can be accessed from multiple nodes

Icon/ resource class name	Attributes	Meaning/attribute value (Top: Meaning, Bottom: Attribute value)
		SHD_SWITCH Switching disk that is used exclusively between two nodes
		Blank The disk connection or usage mode is not set up.
	path_stat	This item indicates the stage of each path of a multipath disk.
		The state value of each path in the multipath disk is indicated in sequence for up to eight paths starting from path 0. The values are shown with the format below. (The values in parenthesis are used when there are five or more paths.) **active** (A) Active **standby** (B) Standby **discon** (D) Low-order driver is disconnected. **fail** (F) Failed **stop** (S) Stopped because a path from another node to the same controller is faulted or is undergoing maintenance
SHD_DISK, SHD_MPDisk	Disk_Attr	This item indicates the physical connection mode and usage mode of a disk that can be used from the cluster system.
		SHD_DISK The disk is physically shared, but the usage mode (shared disk or switchover disk) is not specified.
		SHD_SHARE Shared disk that can be accessed from multiple nodes
		SHD_SWITCH Switching disk that is used exclusively between two nodes
SDX_DC, SDX_SHDDC	Disk_Attr	This class indicates the physical connection mode and usage mode of a GDS-managed disk class that can be used from the cluster system.
		SHD_DISK The disk is physically shared, but the usage mode (shared disk or switchover disk) is not specified.
		SHD_SHARE Shared disk class that allows access from multiple nodes
		SHD_SWITCH Switching disk class for exclusive use between two nodes
Ethernet	node_name	This item indicates the name of the node in which this LAN board is set.
		The node name is set.
	org_mac	This item indicates the MAC address that the operating system has assigned to this LAN board.
		The MAC address is in the format 00:11:22:33:44:55.
	WebView	This item indicates internal information used in the cluster.
		Undefined
SHD_Host	ip_addr	This item indicates the takeover IP address.
		The takeover IP address information is set in the format XXX.XXX.XXX.XXX. If IP address takeover has not been set, this item is blank.
		This item indicates the takeover MAC address.
	mac_addr	The takeover MAC address information is set in the format 00:11:22:33:44:55. If MAC address takeover has not been set, this item is blank.
	node_name	This item indicates the takeover node name.

Icon/ resource class name	Attributes	Meaning/attribute value (Top: Meaning, Bottom: Attribute value)
		The takeover node name is set. If node name takeover has not been set, this item is blank.
	LanDev_Rid	If a takeover MAC address is used, this item indicates the resource ID of LanDev (maintenance information for customer support representative).
		If a takeover MAC address is used, the resource ID of LanDev is set.

7.1.3 RMS Main Window

The RMS main window consists of the following elements:

- RMS tree

- Configuration information or object attributes

- Switchlogs and application logs

7.1.3.1 RMS Tree

The RMS tree displays the configuration information of the cluster in a hierarchical format. The tree has the following levels:

- Root of the tree - Represents the cluster.

- First level - Represents the system nodes in the cluster.

- Second level - Represents the userApplication objects running on each of the system nodes.

- Third level - Represents the sub-applications (an aggregation of objects if any exist).

- Fourth level - Represents the resources required for each of the sub-applications.

If an application has sub-applications, the fourth level represents resources used by that sub-application. If an application does not have sub-applications, then the third level represents all the resources used by userApplication.

Dependency relationships between cluster applications are indicated by controller objects in the RMS tree.

Figure 7.1 RMS main window
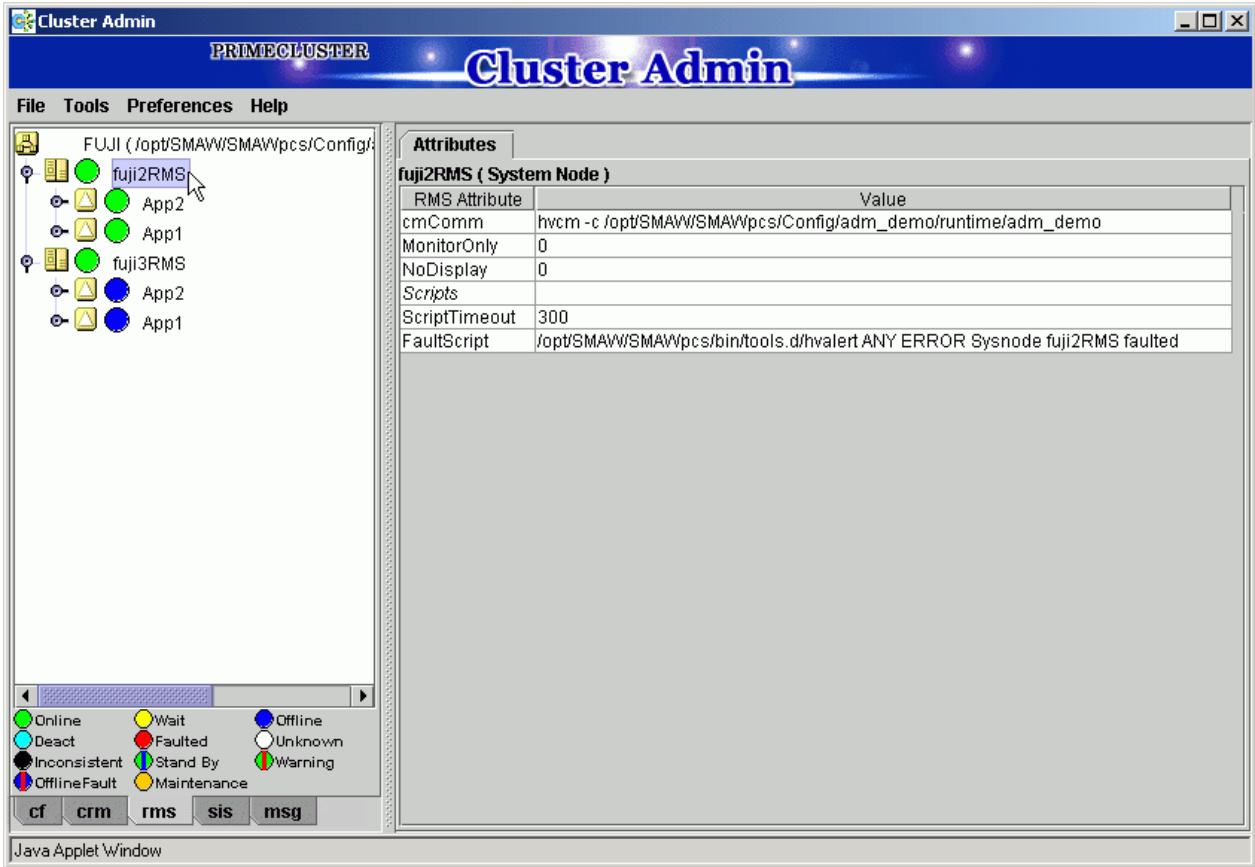
Meanings of object icons

Icon	Meaning
	Represents the cluster.
	Represents a node.
	Represents a parent object (cluster application) that has a child object.
	Represents a child object (cluster application or resource).
	Represents a leaf object (cluster application or resource). A leaf object is an object that cannot have a child object.
	Represents a controller object (cluster application). This object controls an object of another cluster application.

Meanings of state display icons

On the right side of the object icons shown above is displayed a color-coded circle that indicates the state of each object. This section describes the meaning of the colored circles (state display icons).

Information

State display icons are not displayed in cluster icons. Instead, the RMS cluster table can be displayed. For details, see "7.3.3 Concurrent Viewing of Node and Cluster Application States."

State indication of node

The state icons that are displayed in nodes are shown below.

Icon	Icon color	Outline	Details
Node	🟢Green	Online	Node is enabled and ready for use.
	🔵Blue	Offline	Node is enabled but RMS is disabled.
	🔴Red	Faulted	Node is disabled. If the node is shutdown normally, "Shutdown" appears in the SysNode state detailed information (as the value of the StateDetails attribute). If the node is shutdown abnormally, "Killed" appears in the SysNode state detailed information (as the value of the StateDetails attribute).
	🟡Yellow	Wait	Node is undergoing a state transition.

State indication of other objects

The following state icons are displayed with object icons that fall under parent objects.

Icon	Icon color	Outline	Details
Parent object / Child object / Controller object / Leaf object	🟢Green	Online	Object is enabled and ready for use.
	Green with vertical red lines	Warning	Object has exceeded some warning threshold.
	🔵Blue	Offline	Object is disabled and should not be used.
	🔴Red	Faulted	Object encountered an error condition.
	⚪White	Unknown	Monitoring and control are not being executed for the object.
	🟡Yellow	Wait	Node is undergoing a state transition.
	Sky blue	Deact	Node is deactivated because of maintenance, etc.
	⚫Black	Inconsistent	Node state is inconsistent.
	Green with vertical blue lines	Stand By	Object is in such a state that it can be quickly brought Online when needed
	Blue with vertical read lines	OfflineFault	Object is Offline Fault, but a fault occurred in the past and has not yet been cleared

Pop-up menu

If you right-click an object in the RMS tree, a pop-up menu lists the operations for that object. You can also use this menu for monitoring the state.
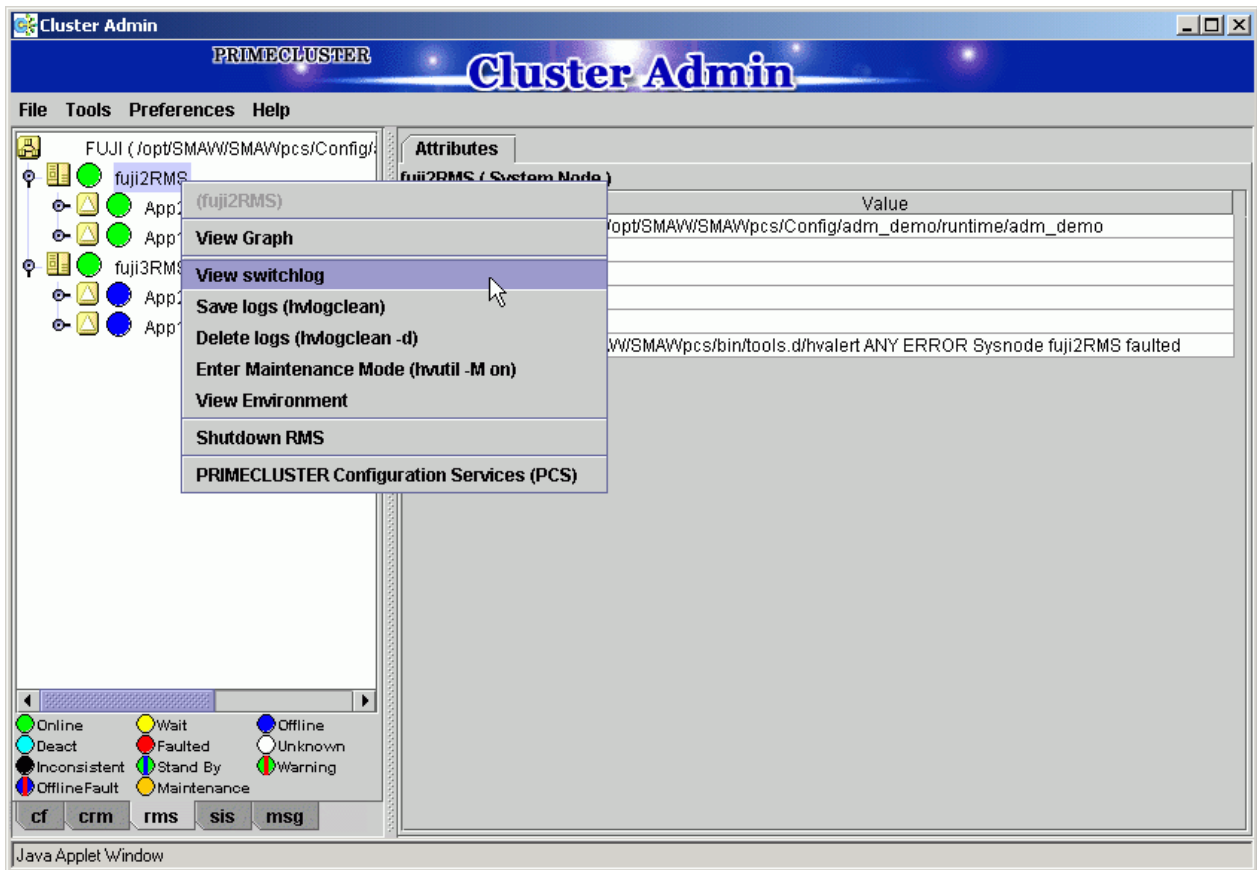
7.1.3.2 Configuration information or object attributes

View the configuration information for the individual objects by left-clicking the object in the RMS tree. The properties are displayed in a tabular format on the right-hand side panel of the RMS main window.

7.1.3.3 Switchlogs and application logs

Each node has a log file referred to as the switchlog. In this file, switchover requests and node failures are recorded. The switchlog is displayed in a tab on the right-side panel.

Display the application log by right-clicking on an application on the RMS tree and choosing *View Application Log*.
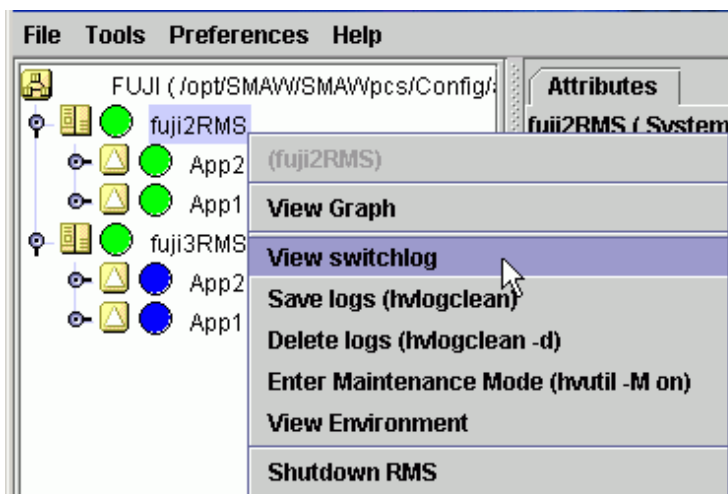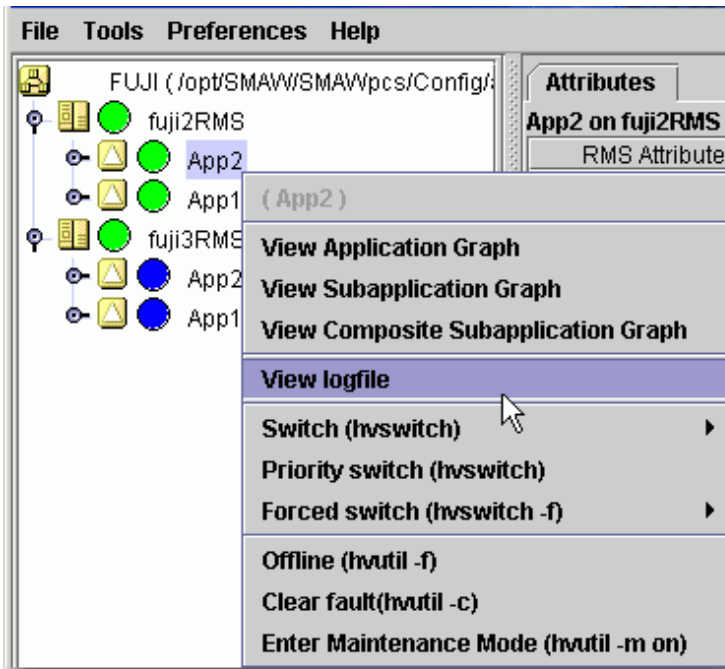
7.2 Operating the PRIMECLUSTER System

7.2.1 RMS Operation

To monitor RMS, RMS needs to be activated.

To stop multiple nodes at the same time, you must stop the user applications and RMS.

![Note icon] **Note**

· ·

To stop two or more nodes at the same time, it is necessary to first stop RMS.

Note that the user application is also stopped when you stop RMS. For instructions on stopping RMS, see "7.2.1.2 Stopping RMS."

· ·

The sections below explain how to start and stop RMS.

7.2.1.1 Starting RMS

This section explains how to start RMS.

Operation Procedure:

From the top screen of Web-Based Admin View, open Cluster Admin according to the following procedure:

1. Select Global Cluster Services.

2. Click *Cluster Admin* to switch to the cluster menu.

3. Select the *rms* tab.

4. Start RMS.

   1. Use the *Tools* pull-down menu of the RMS main window, and click *Start RMS -> all available nodes*. When the confirmation screen is displayed, click *Ok*.

   2. To specify an option for the RMS start command, click the *Option* box.

   You can also start RMS on one or all nodes using the command pop-up as follows:

1. Choose either one node or all the nodes from the pop-up window.

2. To specify an option for the RMS start command, click the *Option* box.

![See icon] **See**
....

See "4.13 Starting RMS" in "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide.*"
....

7.2.1.2 Stopping RMS

This section explains how to stop RMS.

Operation Procedure:

1. Use the *Tools* pull-down menu, or right-click on a system node, and select the mode of shutdown in the subsequent option screen.

    1. Choose either a specific node or all the nodes.

    2. Choose to stop all user applications, leave the applications running, or forcibly shutdown the node(s).

![See icon] **See**
....

See "4.1 Stopping RMS" in "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide.*"
....

7.2.2 Cluster Application Operations

This section explains how to change the operation state of the PRIMECLUSTER system. The operations include starting and stopping a cluster application.

7.2.2.1 Starting a Cluster Application

The procedure for starting a cluster application is described below.

Operation Procedure:

1. On the RMS tree in the RMS main window, right-click the cluster application to be started, and select *Online* from the pop-up menu that is displayed.

    The cluster application will start.

![Information icon] **Information**
....

You can also display the pop-up menu by right-clicking the target icon in an RMS graph or the RMS cluster table. For details on RMS graphs and the RMS cluster table, see "7.3.5 Viewing Detailed RMS Object Information" and "7.3.3 Concurrent Viewing of Node and Cluster Application States."
....

7.2.2.2 Stopping a Cluster Application

The procedure for stopping a cluster application is described below.

Operation Procedure:

1. On the RMS tree in the RMS main window, right-click the cluster application to be stopped, and select *Offline* from the displayed pop-up menu.

    The cluster application will stop.

7.2.2.3 Switching a Cluster Application

The procedure for switching a cluster application is described below.

Operation Procedure:

1. Right-click on the application object and select the *Switch* menu option.

   A pull-down menu appears listing the available nodes for switchover.

2. Select the target node from the pull-down menu to switch the application to that node.

7.2.2.4 Bringing Faulted Cluster Application to Online State

The procedure for bringing a Faulted cluster application to Online is described below.

Operation Procedure:

1. Right-click on the cluster application object in the RMS tree, and select *Clear Fault*.

   If you clear Fault while the cluster application is Online, RMS brings the faulted resource Online.

   If you clear Fault while the cluster application is Offline, RMS brings the faulted resource Offline.

7.2.2.5 Clearing the Wait State of a Node

The procedure for clearing the Wait state of a node is described below.

Operation Procedure:

1. Right-click on the system node and select the "*Online*" or "*Offline*" option from the pop-up menu.

 Note

· ·

If you clear the Wait state of a system node manually with the GUI, hvutil -u SysNode, or cftool -k, then RMS, SF, and CF assume that you have already checked that the target node is stopped.
If you clear the Wait state when the node has not been stopped, this may lead to the data corruption.

· ·

7.2.2.6 Entering maintenance mode for Cluster Application

The maintenance mode is a specific mode to temporarily restrict a specified cluster application switching.

If a cluster application enters maintenance mode, switching caused by a failure of a resource or a cluster application does not occur.

Note that cluster nodes and resources are monitored during maintenance mode, and a cluster application enters the Faulted state if a failure occurs.

Maintenance mode is only allowed when PRIMECLUSTER products specified.

For using maintenance mode, see "5.4.11 Using maintenance mode" in "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide.*"

7.3 Monitoring the PRIMECLUSTER System

This section describes how to monitor the state of a cluster application or node from the Cluster Admin screen.

7.3.1 Monitoring the State of a Node

Click a node on the CF tree. The node state will appear in the right panel.
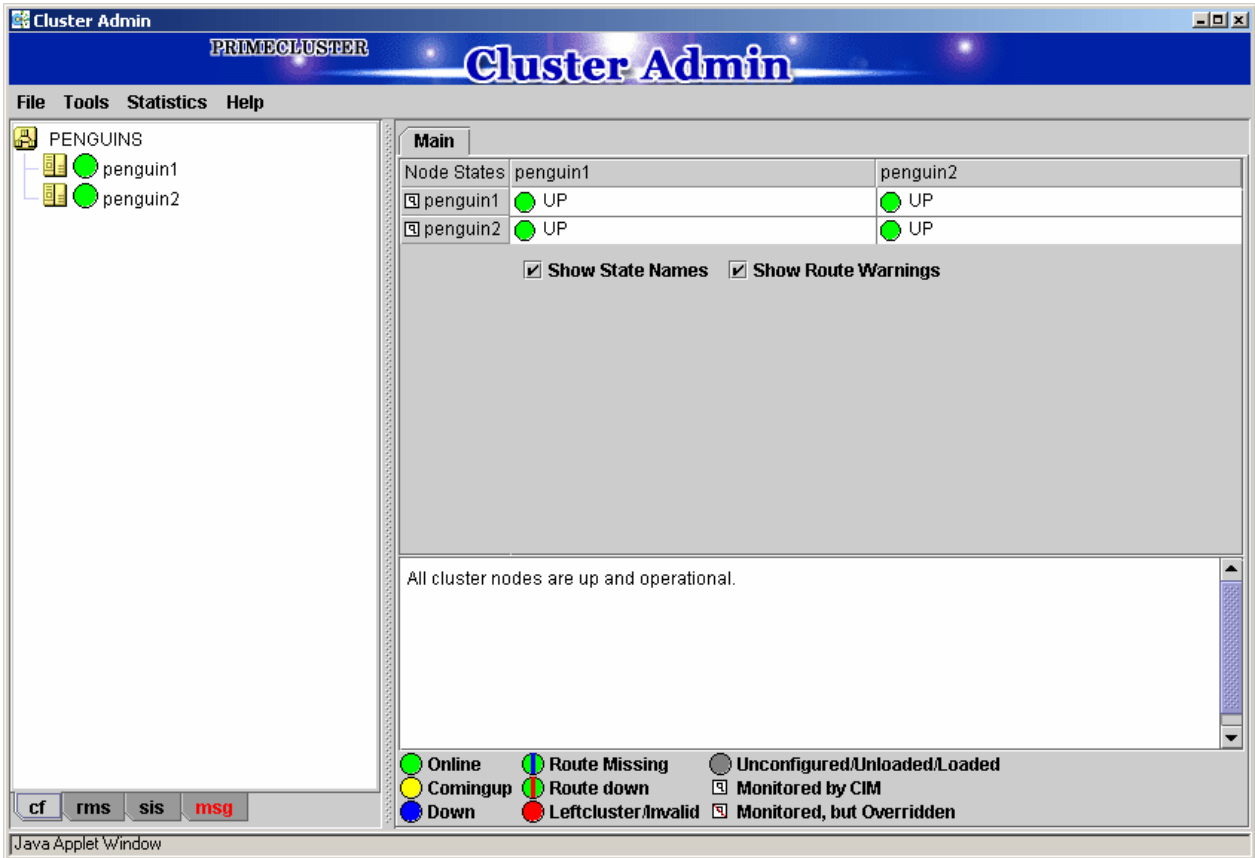
The node indicates one of the following states:

CF state		Description
Green	UP	The node is up and part of this cluster.
Blue	DOWN	The node is down and not in the cluster.
White	UNKNOWN	The reporting node has no opinion on the reported node.
Red	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.

7.3.2 Monitoring the State of a Cluster Application

When you display a cluster application in an RMS tree, the cluster application state appears to the right of the icon. The right panel displays the properties of the cluster application.
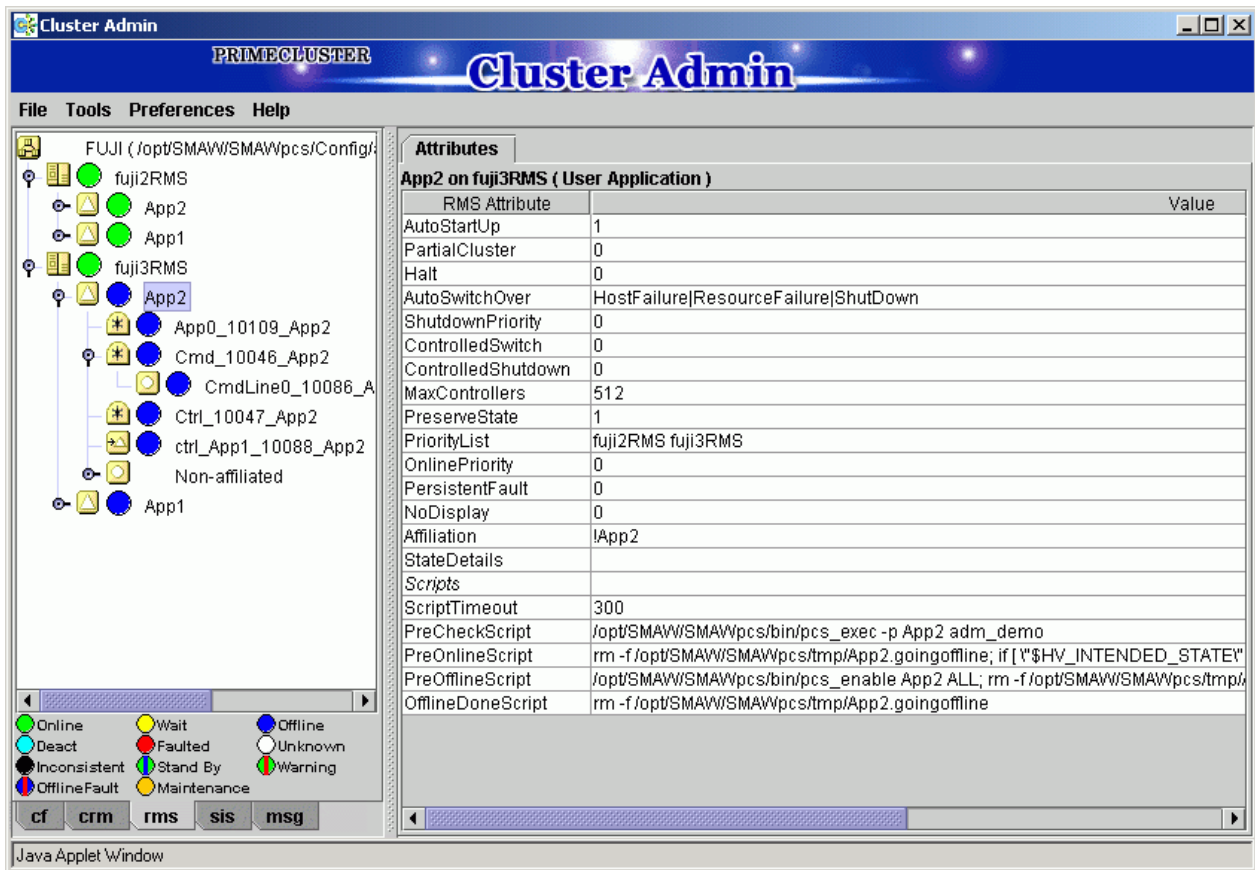
The cluster application shows one of the following states:

- Online

- Offline

- Unknown

- Wait

- Deact

- Inconsistent

- Faulted

- Stand By

- Warning

- OfflineFault

- Maintenance

🔖 See
· ·

See "State indication of other objects" in "7.1.3.1 RMS Tree."
· ·

7.3.3 Concurrent Viewing of Node and Cluster Application States

To view the states of the nodes and cluster applications concurrently, display the RMS cluster table.

Right-click the cluster icon on the RMS tree, and select *View Cluster Wide Table* from the displayed pop-up menu. The RMS cluster table is displayed as shown below.

RMS clusterwide table
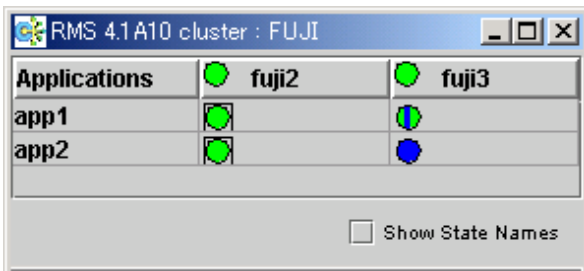
The first line shows the names of the nodes that RMS is managing (fuji2 and fuji3 in the example above). To the left of each node name is a state display icon that shows the state of that node.

The second and subsequent lines show the names of all cluster applications that RMS is managing and the states of those applications. The RMS cluster table enables you to display the states of nodes and cluster applications in one table.

Viewing the RMS Cluster Table

If the background color of the cluster application name is the same as that of the background of the window

It indicates that the cluster application is online.

If the background of the cluster application name is pink

This condition indicates that the cluster application is in the Faulted state and a failure has occurred in one or more SysNode.

If the background of the cluster application name is sky blue

This condition indicates that the cluster application is in the Offline state.

If the state display icon of a cluster application is enclosed in a rectangle

This condition indicates that the node has the highest priority among those nodes that configure the cluster application. If the cluster application is started after creating the cluster application, the node in a rectangle will be in the Online state.

Displaying/hiding state names

Select the *Show State Names* checkbox to display state names to the right of the state display icons.

 See
....

For details on the RMS cluster table, see "5.3.8 Using the RMS clusterwide table" in "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide.*"
....

7.3.4 Viewing Logs Created by the PRIMECLUSTER System

There are two types of logs that can be viewed in the PRIMECLUSTER system:

- Switchlog

  The switchover requests or failures that occur in nodes are displayed.

- Application log

  The operation log of the cluster application is displayed.

7.3.4.1 Viewing switchlogs

Right-click the system node and select the *View Switchlog* option from the pop-up menu. The switchlog is displayed on the right side of the screen.
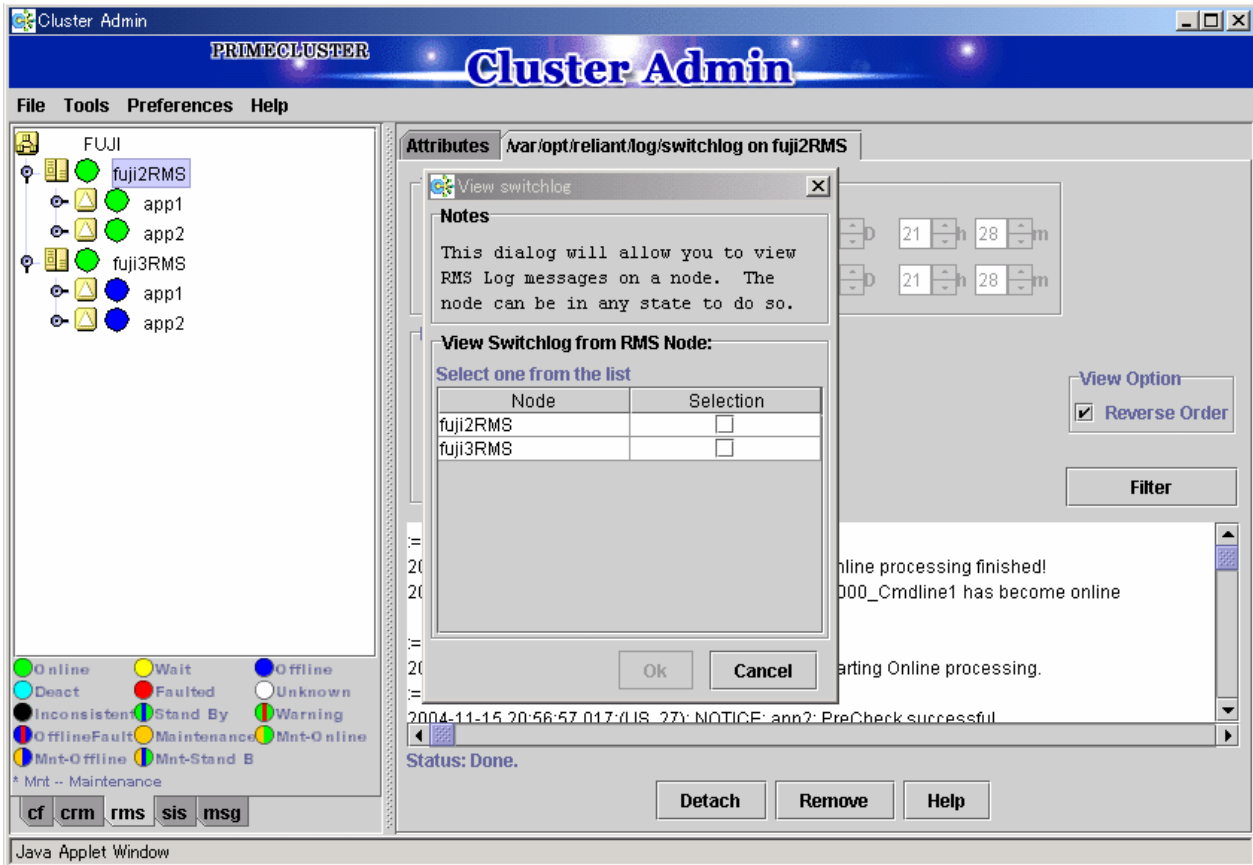
7.3.4.2 Viewing application logs

Right-click an application on the RMS tree and choose *View Log File*. The application log for that application will be displayed on the right side of the screen.
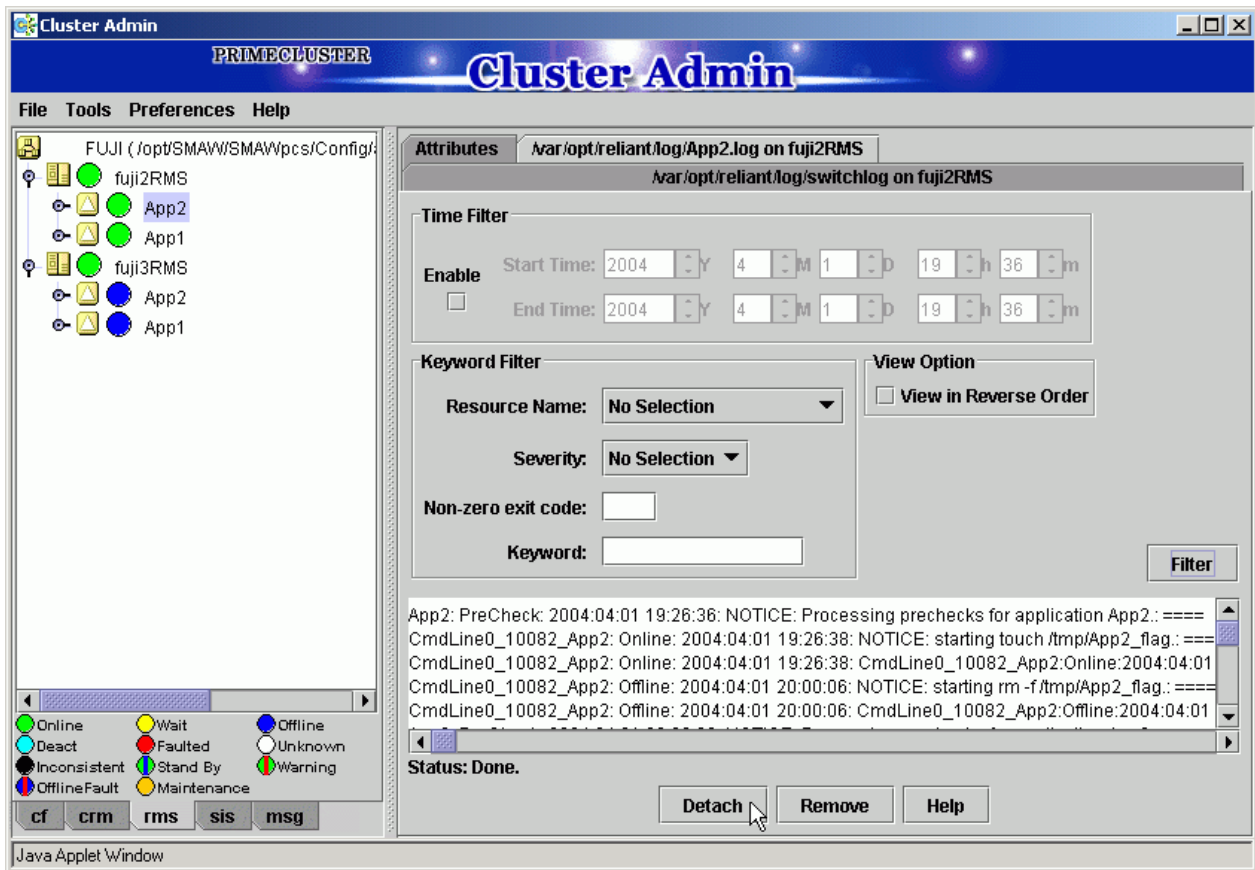
![Information icon] **Information**

. . . .

The following display formats are enabled for the log. For details, see "5 Administration" in "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide.*"

- Narrow the log by date or keyword.

- Scroll or jump to any entry.

- Search by keyword, date range, error message significance, or exit code other than 0 to exclude unrelated entries.

. . . .

7.3.5 Viewing Detailed RMS Object Information

Use RMS graphs to display detailed object information for each cluster application.

There are four types of RMS graphs. Each type can be displayed from the menu when you right-click an object on the RMS tree.

- Full graph

    Displays the configuration of the entire cluster system in which RMS is running.

- Application graph

    Shows all objects used by the specified application. You can check the details of the specific object using this graph.

- Sub-application graph

    Lists all sub-applications used by a given application and shows the connections between the sub-applications.

- Composite sub-applications graph

    Shows all sub-applications that the application depends on directly or indirectly.

RMS graphs
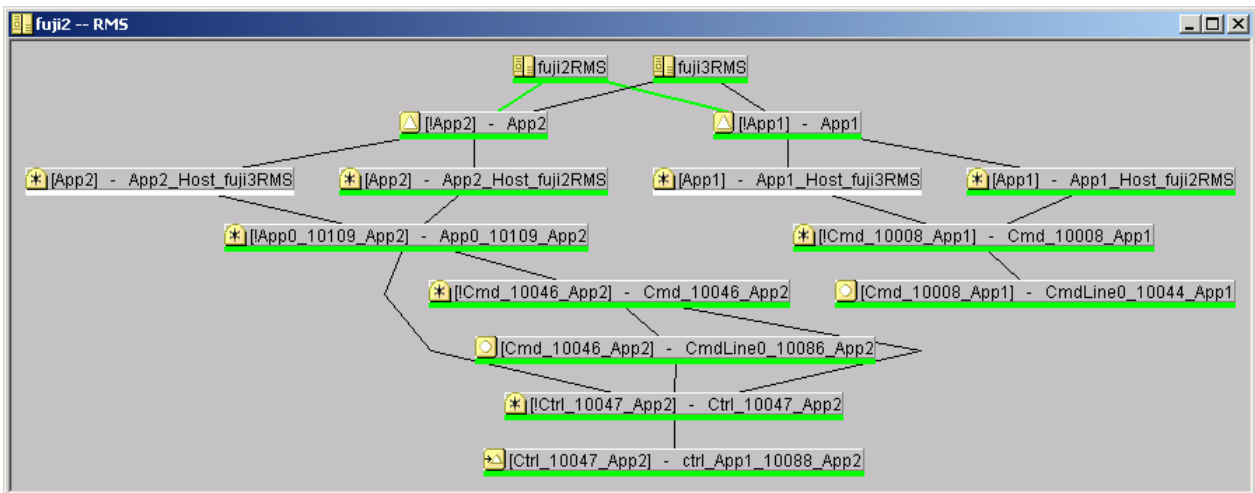
If you left-click the target object, the attributes of the object will be displayed on a pop-up screen.
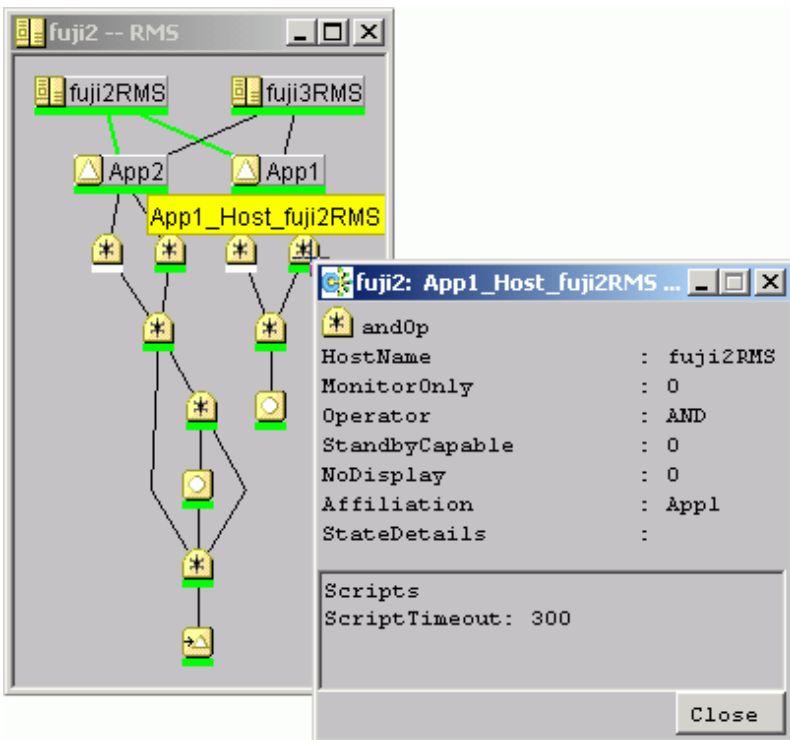
📚 **See**

. . . .

See "5.5 Using RMS graphs" in "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide.*"

. . . .

7.3.6 Displaying environment variables

Right-click a cluster in the RMS tree window and select *View Environment*. The local and global variables are displayed.
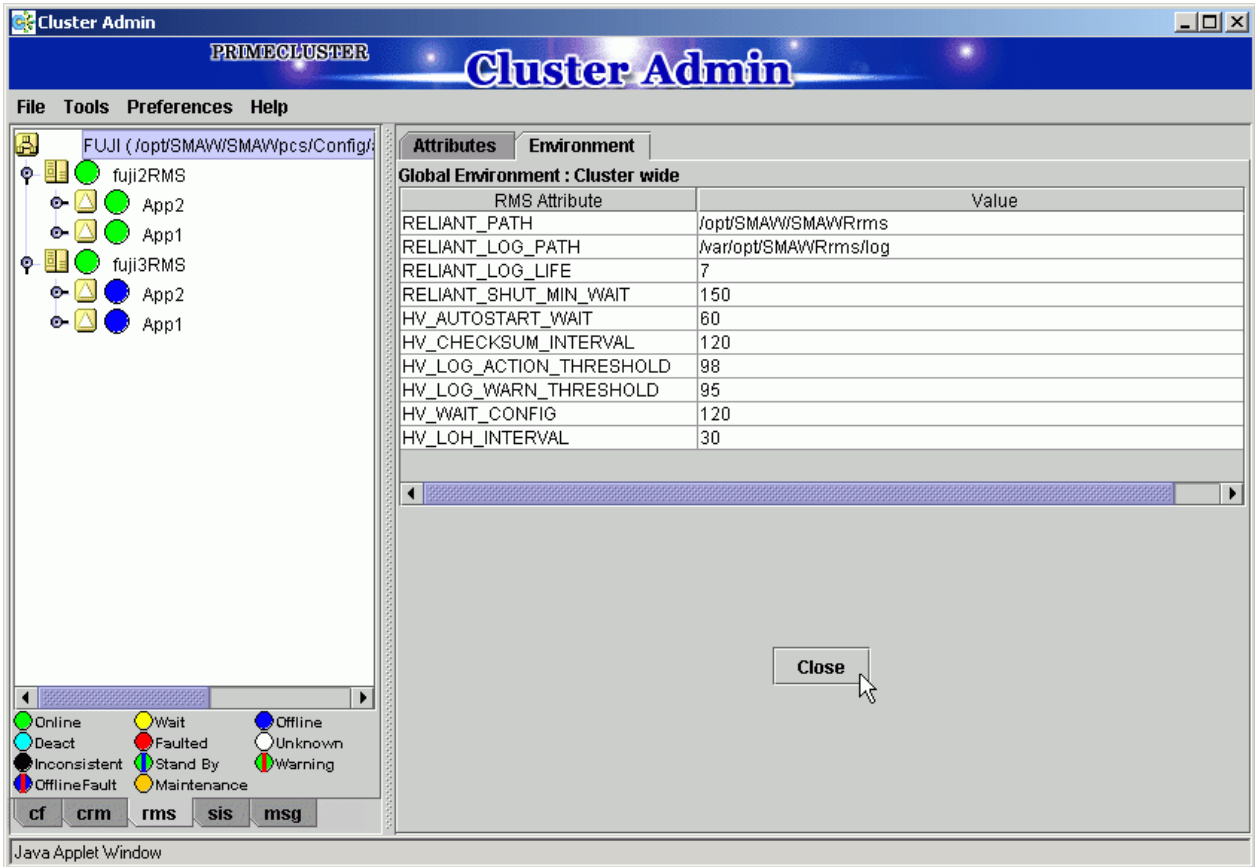
Right-click a node in the RMS tree, and select *View Environment*. The local variables are displayed.
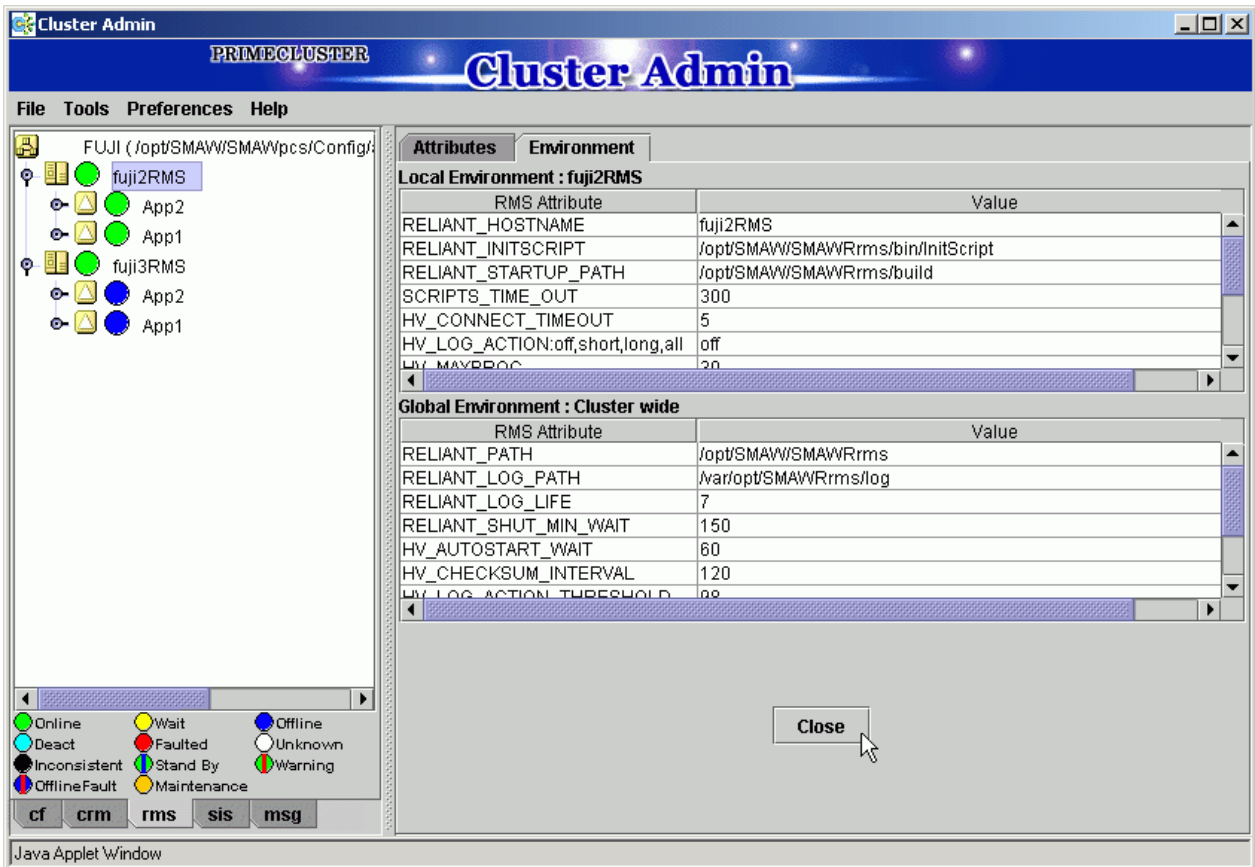
7.3.7 Monitoring Cluster Control Messages

Select the *msg* tab, which is found at the bottom of the tree panel. If a new message was added to the text area since the last time the area was displayed, this tab is displayed in red.

You can clear the message text area or isolate it from the main panel.

7.4 Corrective Actions for Resource Failures

A hardware or software failure might occur while the cluster system is running. If a failure occurs in a resource, a message indicating that a failure occurred is displayed in the Cluster admin pop-up screen. Based on this message, you need to identify the faulted resource using the CF, CRM, and RMS main window and take corrective actions to maintain high availability in the cluster system.

This section describes the actions to be taken for the following cases:

- If the resource state became Faulted.

- If the Patrol Diagnosis Facility detects a failure.

7.4.1 Corrective Action when the resource state is Faulted

This section describes the corrective actions to take when the resource state became Faulted.
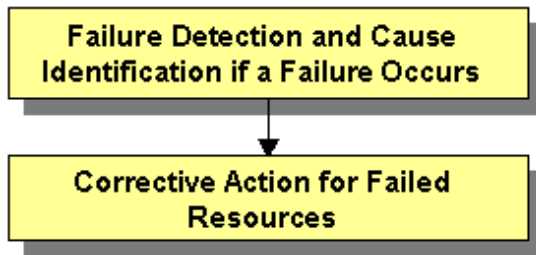
7.4.1.1 Failure Detection and Cause Identification if a Failure Occurs

If a failure occurs in a resource, you can use the functions of PRIMECLUSTER and the operating system to detect the failure and identify the faulted resource that caused the failure.

The descriptions given in (a) to (k) below are relevant to the "Failure confirmation features list" given below:

Failure detection

Normally, the RMS main window (b) is used to monitor the cluster applications.

- If a failure occurs in a resource or the system

    Failover of the userApplication or node panic will occur.

    In such a case, you can detect the failure by observing the following conditions:

    - A pop-up message screen (a) is displayed.

    - The color of the icons in the RMS main window (b) changes.

    - A message is output to the msg main window (g), Syslog(j), and the console (k).

- If a warning-level failure occurs in the system

    If a warning-level failure (for example, insufficient disk space or insufficient swap area) occurs in the system, you can detect the failure by observing the following conditions:

    - The node icon in the CRM main window (d) changes.

- A message is output to Syslog(j) and the console (k).

- If RMS fails to start on all the nodes, the userApplication will not start. You can start the userApplication by executing the "clreply" command.

    - By executing the "clreply" command, you can confirm an operator intervention request to which no response has been entered and start up the userApplication by responding to it. For information on the "clreply" command, see the manual pages.

    - The operator intervention request message will be output to Syslog(j) and the console (k). By responding to the operator intervention request message, you can start the userApplication.

    For further details, see "D.5 Operator Intervention Messages."

Note

If there are multiple operator intervention request messages for which no response has yet been entered, you need to respond to each of them.

In addition, you can use the features described in "Failure confirmation features" to detect the failure.

Cause identification

You can also use the function that detected the failure and the features listed in "Failure confirmation features" below to identify the faulted resource that caused the failure.

Failure confirmation features list

	Failure confirmation features	Manual reference
(a)	Message screen	C.3.1 Failed Resource Message
(b)	**RMS main window** The RMS tree and the RMS cluster table can be used from this screen.	7.1.3 RMS Main Window
(c)	**CF main window** The CF tree can be used from this screen.	7.1.1 CF Main Window
(d)	**CRM main window** The CRM tree can be used from this screen. This screen is useful in detecting hardware resource faults.	7.1.2 CRM Main Window
(e)	**"Resource Fault History" screen** This screen is useful in detecting hardware resource faults.	C.3.2 Resource Fault History
(f)	Current list of resources in which a failure has occurred	C.3.3 Fault Resource List
(g)	**MSG main window** The cluster control messages can be viewed in this screen. To display this screen, select the msg tab in the Cluster Admin screen.	-
(h)	Application log	7.3.4.2 Viewing application logs
(i)	switchlog	7.3.4.1 Viewing switchlogs
(j)	Syslog	-
(k)	**Console** Messages that are displayed on the console can be checked. Viewing the "console problem" information on the console can help you identify the fault cause.	Appendix D Messages
(l)	Machine management GUI	Machine Administration Guide

Failure confirmation features		Manual reference
(m)	MultiPathDisk view	Multipath Disk Control Load Balance option x.x Guide
(n)	GDS GUI	PRIMECLUSTER Global Disk Services Configuration and Administration Guide

7.4.1.2 Corrective Action for Failed Resources

Take the following steps for failed resources;

1. Correct the faulted resource

   Correct the problem in the failed resource. For details, see "2 Troubleshooting" in "*PRIMECLUSTER RMS Troubleshooting Guide.*"

   If an error message of patrol diagnosis is displayed, see "7.4.2 Corrective Action when Patrol Diagnosis Detects a Fault."

   "hvdet_sptl" is displayed in the name of the program that outputs the patrol diagnosis message.

   ## 🔰 Note
   ・・・

   If you are using an operation management product other than a PRIMECLUSTER product, you may need to take corrective actions prescribed for that product.

   For details, see the manual provided with each operation management product.

   [Examples] Machine Administration, MultiPathDisk view, GDS
   ・・・

2. Recover the cluster application

   At the RMS main window, check the state of the cluster application to which the corrected resource is registered. If the cluster application is in the Faulted state, execute the Fault clear operation.

   For details on the Fault clear operation, see "7.2.2.4 Bringing Faulted Cluster Application to Online State."

7.4.2 Corrective Action when Patrol Diagnosis Detects a Fault

This section explains the actions to take when the patrol diagnosis facility detects a failure
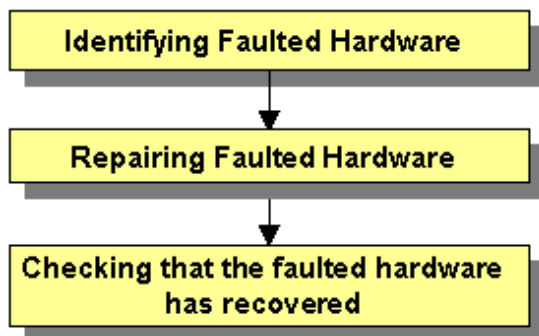
7.4.2.1 Identifying Faulted Hardware

Use one of the following methods to identify the faulted hardware:

- Message text output to the CRM main window or syslogd(1M)

  See "Display format 1" in "D.1 Searching for a Message."

- CRM main window

  The CRM main window displays the OFF-FAIL state for the faulted hardware. See "7.1.2 CRM Main Window."

- "clgettree(1)" command

  The "clgettree(1)" command displays the OFF-FAIL state for the faulted hardware. See the manual page for clgettree(1).

7.4.2.2 Corrective Action for Faulted Hardware

Correct the faulted hardware according to the operation procedure below.

Operation Procedure:

1. Stop the node to which the faulted hardware is connected.

2. Repair the faulted hardware.

3. Start the node.

> 📋 **Note**
>
> ....
>
> When a disk unit that is registered with GDS is to be exchanged, follow the steps described in the GDS disk replacement procedure. For information on GDS disk replacement, see "In Case of Disk Abnormality" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*."
>
> ....

4. Check that the faulted hardware has recovered using one of the following methods:

   1. Use the CRM main window.

   2. Execute the "clgettree(1)" command.

   If the above procedure shows that the fault was not corrected, you need to continue the following procedure:

   1. Execute the diagnosis operation for the faulted hardware from the CRM main window

      Then, use the CRM main window to check whether the fault was corrected. If the fault was corrected, the ON state is displayed.

   2. Execute the "clsptl(1M)" command to initiate the diagnosis operation.

      The "clsptl(1M)" command has two functions. One function allows you to specify a faulted hardware unit and diagnoses only the specified device. The other function runs batch diagnosis of all shared disk units or all network interface cards. If faults occur in multiple hardware units, it is convenient to use the batch diagnosis function.

      - Example in which a faulted shared disk unit is specified and diagnosis is executed:

      ```
      # /etc/opt/FJSVcluster/bin/clsptl -u generic -n c1t4d4
      ```

      - Example in which batch diagnosis is executed for all shared disk units:

      ```
      # /etc/opt/FJSVcluster/bin/clsptl -a DISK
      ```

      Execute the "clgettree(1)" command to check whether the fault was corrected. If the fault was corrected, the ON state will be displayed for the hardware.

5. Bring the Faulted cluster application Online.

   Confirm that the state of the cluster application to which the recovered hardware is registered, either in the RMS main window or with the "hvdisp(1M)" command.
   If the cluster application is Faulted, switch the cluster application from the failed to the active state, either in the RMS main window or with the "hvutil(1M)" command. For information on the procedures related to the CRM main window, see "7.2.2.4 Bringing Faulted Cluster Application to Online State."

   If operator intervention request is enabled, a message will be displayed with the "syslogd(1M)" command and Cluster Admin when RMS is started. By entering a response to this message, you can switch the state of the cluster application from the failed state to active. For information on the setup procedure for operator intervention requests, see "5.4 Setting Up Fault Resource Identification and Operator Intervention Request."
   An example of an operator intervention request is shown below. For details on the messages requesting operator intervention, see "D.7.2 Failed Resource and Operator Intervention Messages (GUI)" and "D.5 Operator Intervention Messages."

```
1422 On the SysNode "node1RMS", the userApplication "app0" is the Faulted state due to a fault
in the resource "apl1".
Do you want to clear fault? (yes/no)
Message number: 1001
```

Note

If "Yes" is set for the "AutoStartUp" attribute, an operator intervention request message will be displayed at node startup. You need to respond to the operator intervention message after executing step 4. of the procedure.

Part 4   System Configuration Modification

Chapter 8 Changing the Operation Configuration

This chapter explains how to add nodes and cluster applications to the PRIMECLUSTER system and change the configuration contents.

8.1 Changing the Cluster Configuration

This section explains the following cluster configuration modification methods:

1. Change a cluster application configuration.

2. Change an operation attribute of a cluster application.

3. Change a resource.

📌 Note

. . . .

- Be sure to stop RMS before changing the cluster configuration. If you start the "userApplication Configuration Wizard" screen while RMS is running, the message 0833 will be displayed. If this happens, stop RMS, and then restart the "userApplication Configuration Wizard" screen. For instructions on stopping RMS, see "7.2.1.2 Stopping RMS."

- If there are resources to be removed from the monitoring target by the standby patrol function after a cluster configuration is changed, delete resources from the target of the standby patrol function by using the following command after stopping RMS:

    - For a physical disk defined by a disk class

    ```
    # /etc/opt/FJSVcluster/bin/clspconfig -u <cluster application name> -d TARGET_DISK=<target
    resource>
    ```

    - For a network interface card

    ```
    # /etc/opt/FJSVcluster/bin/clspconfig -u <cluster application name> TARGET_LAN=<target
    resource>
    ```

. . . .

📖 Information

. . . .

After you finish changing the cluster application, execute "RMS startup" from the Cluster Admin screen. If necessary, also start the cluster application. For instructions on starting RMS, see "7.2.1.1 Starting RMS." For instructions on starting the cluster application, see "7.2.2.1 Starting a Cluster Application."

. . . .

8.1.1 Changing the Cluster Application Configuration

To change the configuration of a cluster application, you must first delete the cluster application. The procedure is shown below.

Operation flow
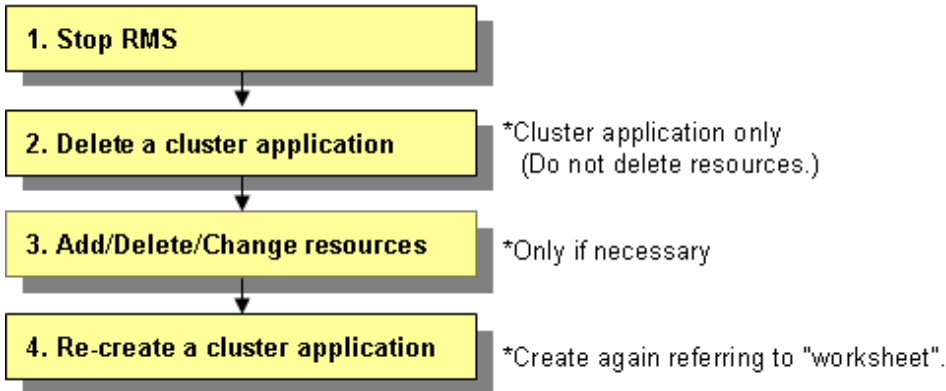
Operation Procedure:

1. Stop RMS.

   If RMS is running, see "7.2.1.2 Stopping RMS" and stop RMS.

2. Delete the cluster application.

   Delete only the target cluster application.

   For details on how to make this deletion, see "8.8.1 Deleting a userApplication."

   ![Note] **Note**

   ⋯⋯

   Do not delete a resource that is necessary to configure the cluster application.

   ⋯⋯

3. Add, delete, or change the resource.

   Add a new resource and delete any unnecessary resources.

   ![See] **See**

   ⋯⋯

   For details on how to perform the above operations, see the following:

   When a resource is to be added: 6.6.1 Setting Up Resources

   When a resource is to be deleted: 8.8.2 Deleting a Resource

   When a resource is to be changed: 8.1.3 Changing Resources

   ⋯⋯

   ![Note] **Note**

   ⋯⋯

   If you have changed a resource, check that all resources are registered to the cluster application. If there are resources that are not registered to the cluster application, delete them or add them to the cluster application.

   ⋯⋯

4. Create a cluster application.

   Create the same cluster application as that deleted in step 2.

   For details on how to create the cluster application, see "6.6.2 Creating Cluster Applications." If cluster application dependency relationships have been set up, they must be set up again.
   For details on how to make these settings, see "6.6.3 Setting Up Dependency Relationships Between Cluster Applications."

Note

. . . .

If it is not necessary to change the operation attributes of the cluster application, leave the attribute unchanged as explained in the "A.9 Cluster Application Worksheets."

. . . .

8.1.2 Changing the Operation Attributes of a Cluster Application

The procedure for changing the operation attributes of a cluster application is described below.

Operation Procedure:

1. Stop RMS.

   If RMS is running, see "7.2.1.2 Stopping RMS" and stop RMS.

2. Change the operation attributes of the target userApplication.

   At the "Global Cluster Services" screen, select *userApplication Configuration Wizard*.
   From the tree on the left of the "userApplication Configuration Wizard" screen, select the userApplication to be changed, right-click the mouse to display the pop-up menu, and select *Edit userApplication or Resource*.

Point

. . . .

To change the cluster application, you can also select *Edit userApplication or Resource* at the top menu of the userApplication Configuration Wizard, and click *Next*.

. . . .
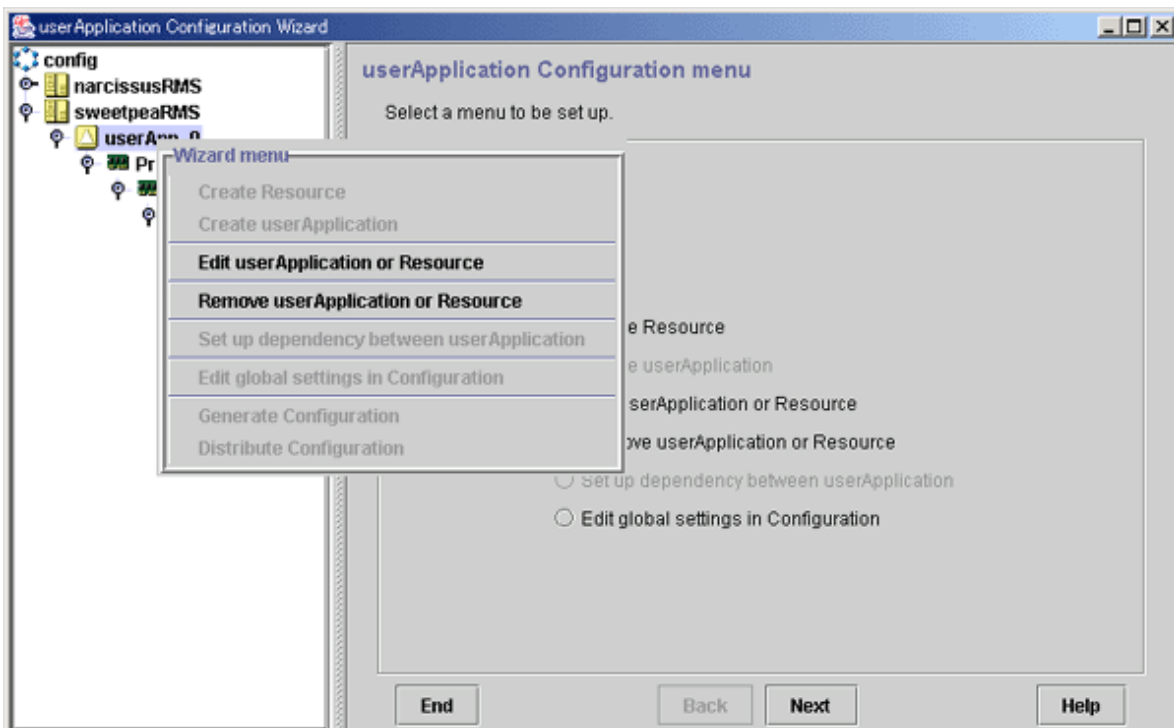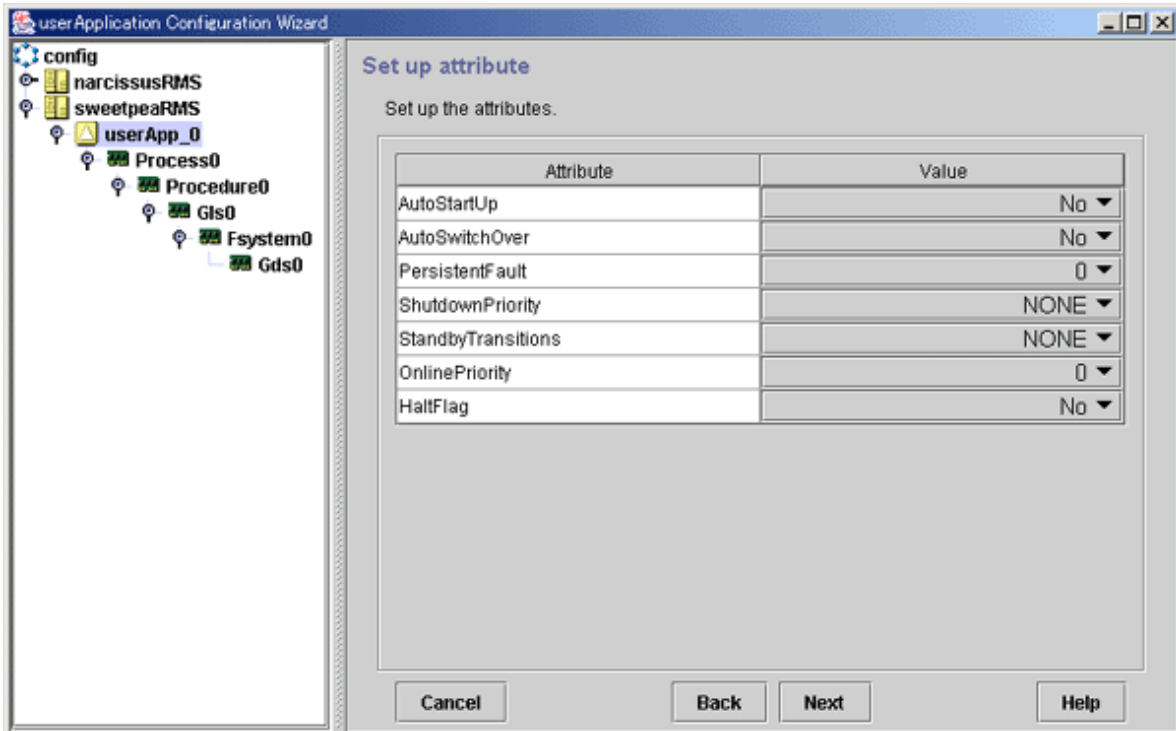
Change the operation attributes. For information on the individual operation attributes, see "6.6.5 Attributes."
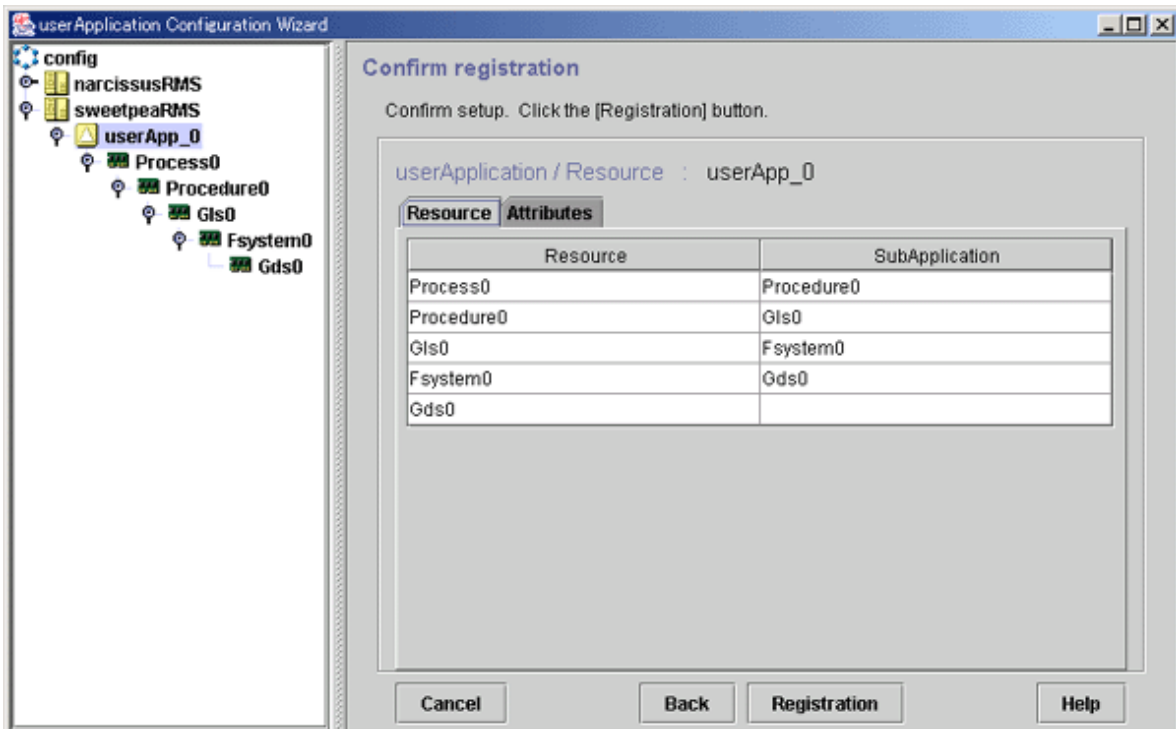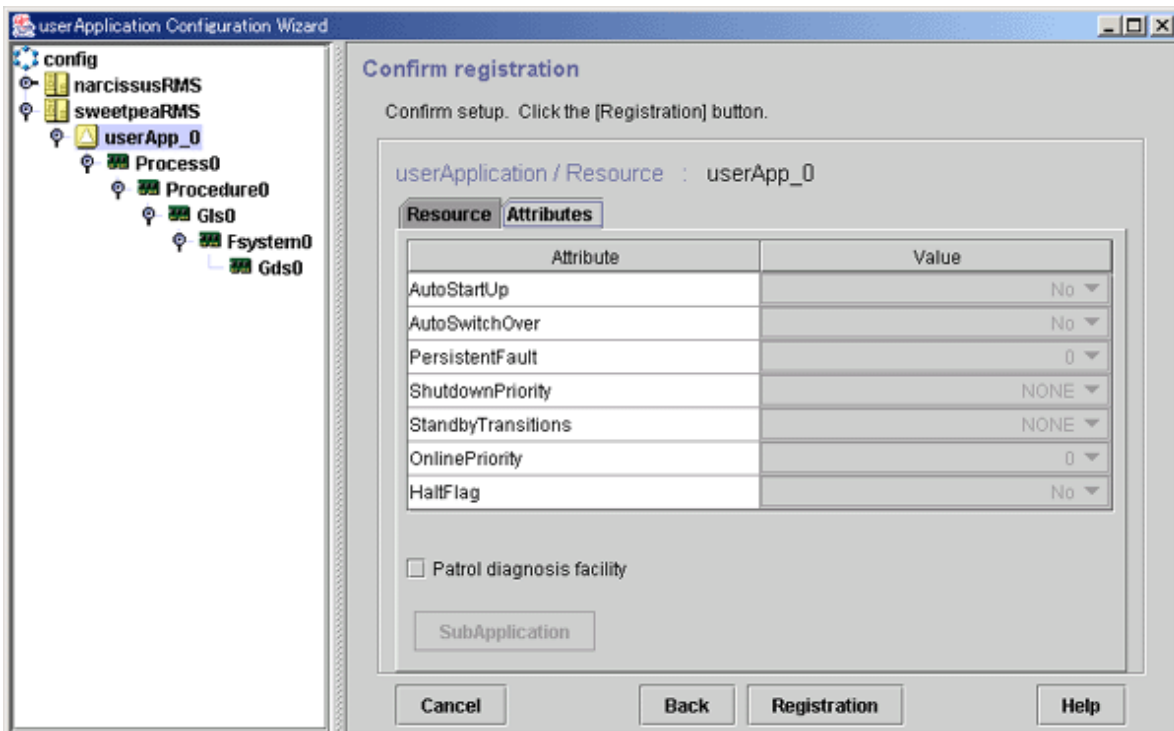
After you finish changing the operation attributes, register the attributes. Click the *Registration* button to register the attributes to RMS Configuration.

After registration to RMS Configuration is completed, the following message screen is displayed if the system judges that distribution of RMS Configuration is possible.
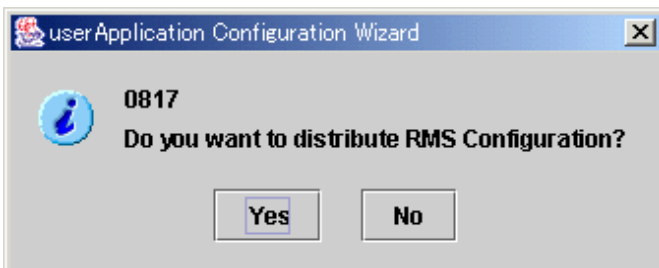
If you are not changing the operation attributes or resource attributes of another cluster application, click *Yes*.

Information
. . . .
PRIMECLUSTER allows operation attributes to be set by CUI operation, according to the needs of the user.

The table below shows the operation attributes and environment variables of userApplication objects and the values that can be set for those attributes and variables.

Table 8.1 Operation attributes of userApplication objects

Operation	userApplication attribute/environment variable	Setting value (Bold: Default)	Description
Failover operation	AutoSwitchOver	**No**	Automatic failover is disabled.
		HostFailure	Failover occurs if a node fails.
		ResourceFailure	Failover occurs if a resource fails.
		Shutdown	Failover occurs if RMS stops.

Operation	userApplication attribute/environment variable	Setting value (Bold: Default)	Description
Operation if resource inactivation fails	HaltFlag	**No**	The node is not forcibly stopped.
		Yes	The node is forcibly stopped.
Operation for cluster application startup	OnlinePriority	**0**	The highest node of the PriorityList attribute becomes the active system.
		1	The node that was the active system last becomes the active system.
Operation of failover function for node startup	HV_RCSTART environment variable in hvenv.local file	**0**	The failover function is not started automatically.
		1	The failover function is started automatically.
Handling of failed nodes	PersistentFault	**0**	The failed node is not eliminated after the node is restarted.
		1	The failed node is eliminated even after the node is restarted.
Hot-standby operation	StandbyTransitions	**No**	Hot-standby is not adopted.
		StartUp	Hot-standby is adopted when the failover function is started.
		SwitchRequest	Hot-standby is adopted after the switchover operation.
		ClearFaultRequest	Hot-standby is adopted after the fault clear operation.

8.1.3 Changing Resources

The resource modification includes the followings:

1. Change the interface used by a resource.

2. Change the attributes to the resources and resource interfaces.

P Point

Resources can be changed only when RMS is stopped.

8.1.3.1 Changing the Interface Used by a Resource

To change the interface used by a resource, you must first delete resource and then add it back again. Interface examples include a takeover network for an Ipaddress and a disk class for a Gds resource.

Operation flow
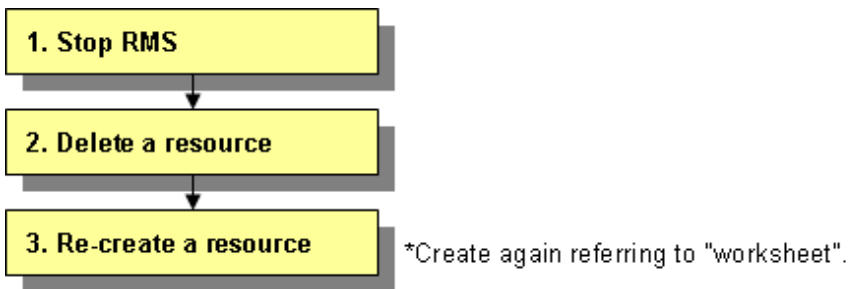
1. Stop RMS

2. Delete a resource

3. Re-create a resource

*Create again referring to "worksheet".

Operation Procedure:

Use the "userApplication Configuration Wizard" screen to change a resource used by a cluster application.

1. Stop RMS

   If RMS is running, see "7.2.1.2 Stopping RMS" and stop RMS.

2. Delete the resource.

   Return to the "Global Cluster Services" screen, and select *userApplication Configuration Wizard*.
   From the tree on the left of the "userApplication Configuration Wizard" screen, select the userApplication to be changed, right-click the mouse to display the pop-up menu, and select *Remove userApplication or Resource*.

   Note

   ....

   If you delete a resource included in the cluster application, the other resources found under the target resource are also deleted. In other words, the entire hierarchy of objects found under the resource on the tree in the left panel will be deleted regardless of those objects being related to other resources.

   ....
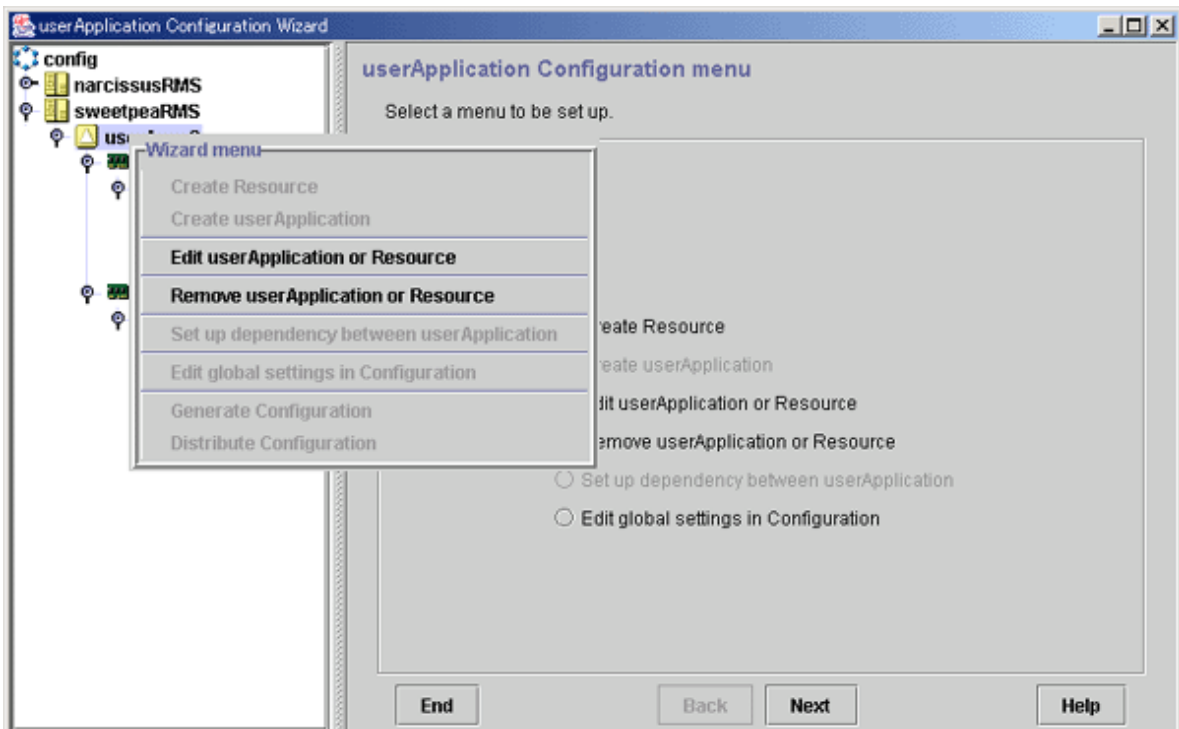
3. Re-create the resource.

   If new resources are required, create resources.

Select *Create Resource* from the *userApplication Configuration Wizard* menu.
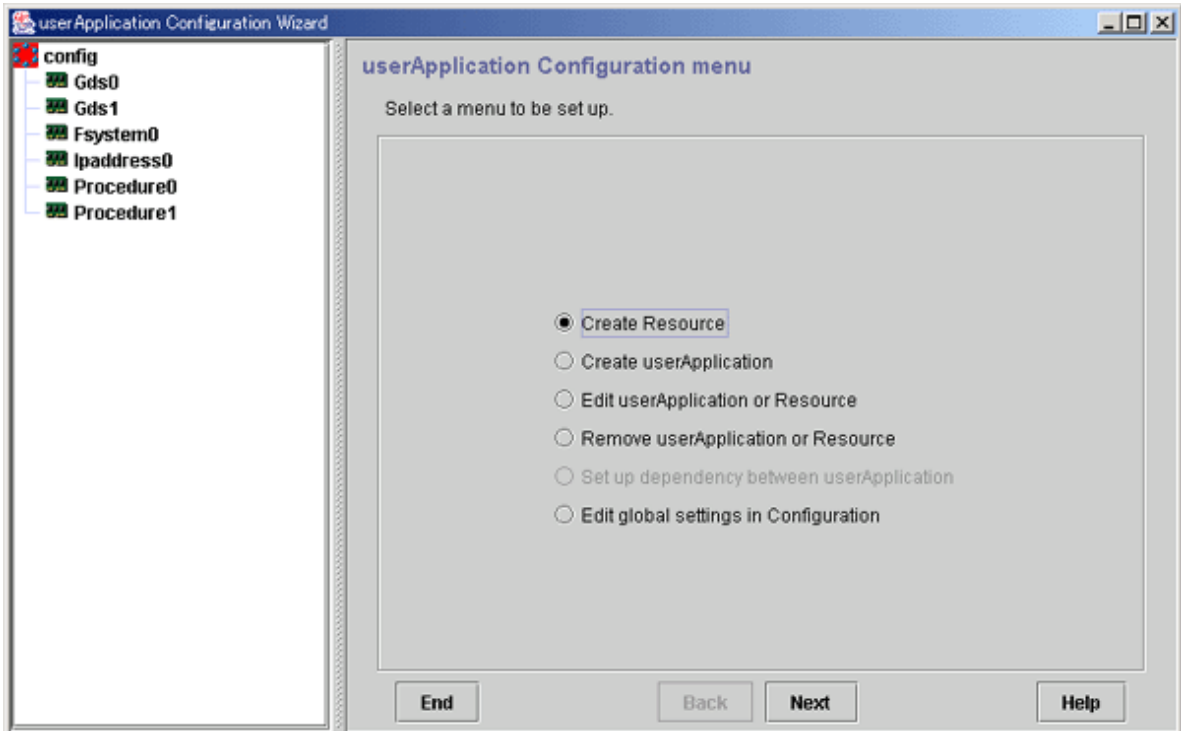
Note

If changes are not required, refer to the "A.9 Cluster Application Worksheets" and set up the resource so that the setup information is the same as before.

See

For setup instructions, see "6.6.1 Setting Up Resources."

8.1.3.2 Changing the Attributes Used by a Resource or a Resource Interface

Change the attributes used by a resource or a resource interface according to the following procedure.

Operation Procedure:

Use the "userApplication Configuration Wizard" screen to change a resource used by the cluster application.

1. Stop RMS.

   If RMS is running, see "7.2.1.2 Stopping RMS" and stop RMS.

2. Select the resource.

   Return to the "Global Cluster Services" screen, and select *userApplication Configuration Wizard*.
   From the tree on the left of the "userApplication Configuration Wizard" screen, select the resource to be changed, right-click the mouse to display the pop-up menu, and select *Edit userApplication or Resource*.

Point

To change the attributes used by the resource or an interface of the resource, you can also select *Edit userApplication or Resource* from the top menu of the userApplication Configuration Wizard, and then select *Next*.

Cmdline
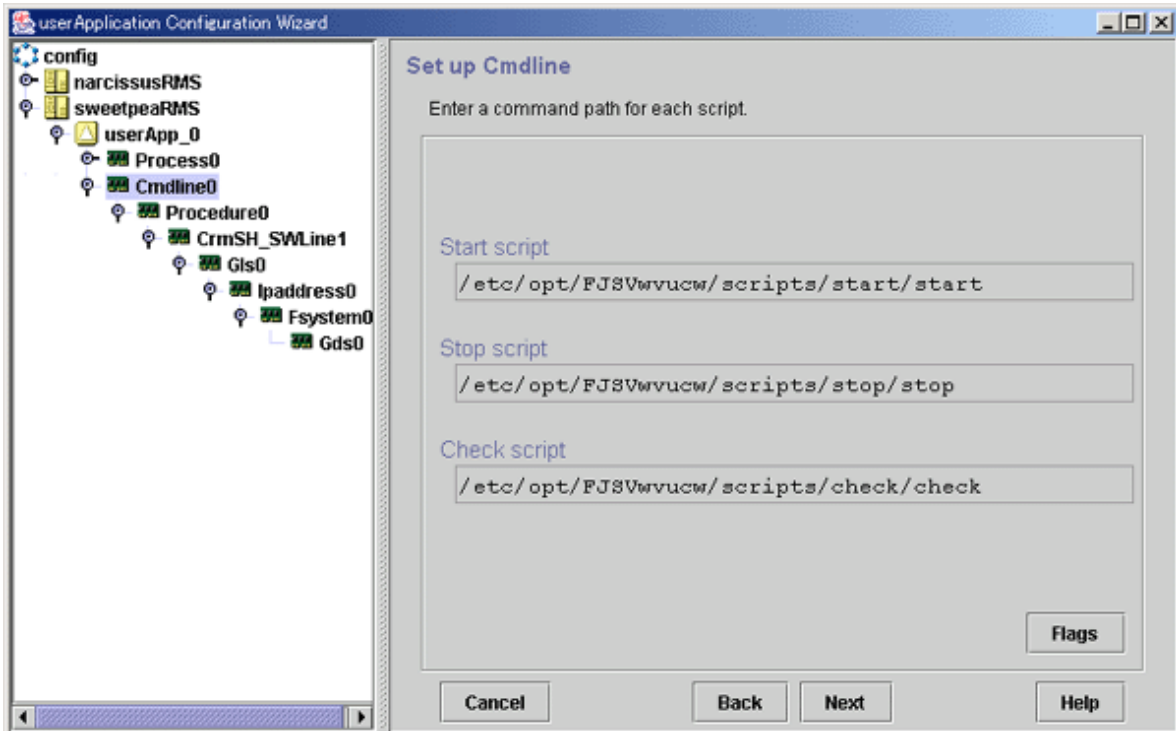
Flags button

Click this button to change the script attributes, if necessary.
A dialog box for setting the attributes of the created script will be displayed. For details, see "Setting up Cmdline flags" in "6.6.1.1 Creating Cmdline Resources."
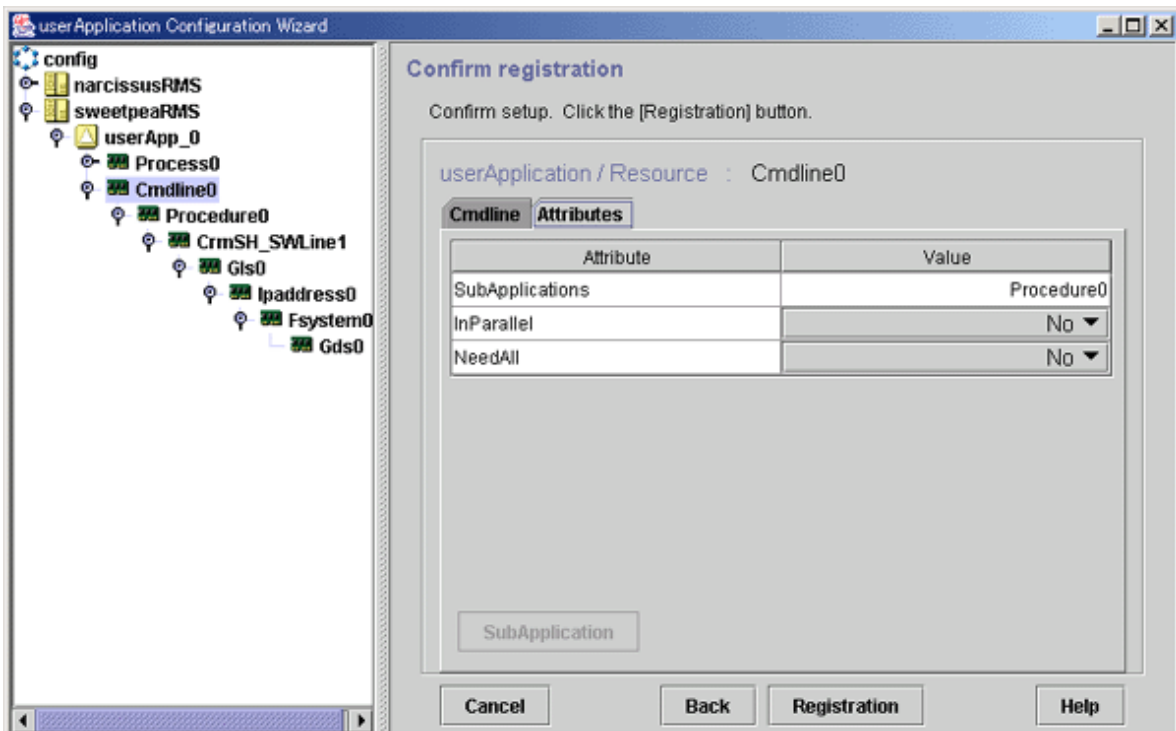
SubApplication button

Click this button to associate a previously created Cmdline resource or processing monitoring resource found under the current Cmdline resource. This button cannot be selected if the resource belongs to a userApplication.

Fsystem
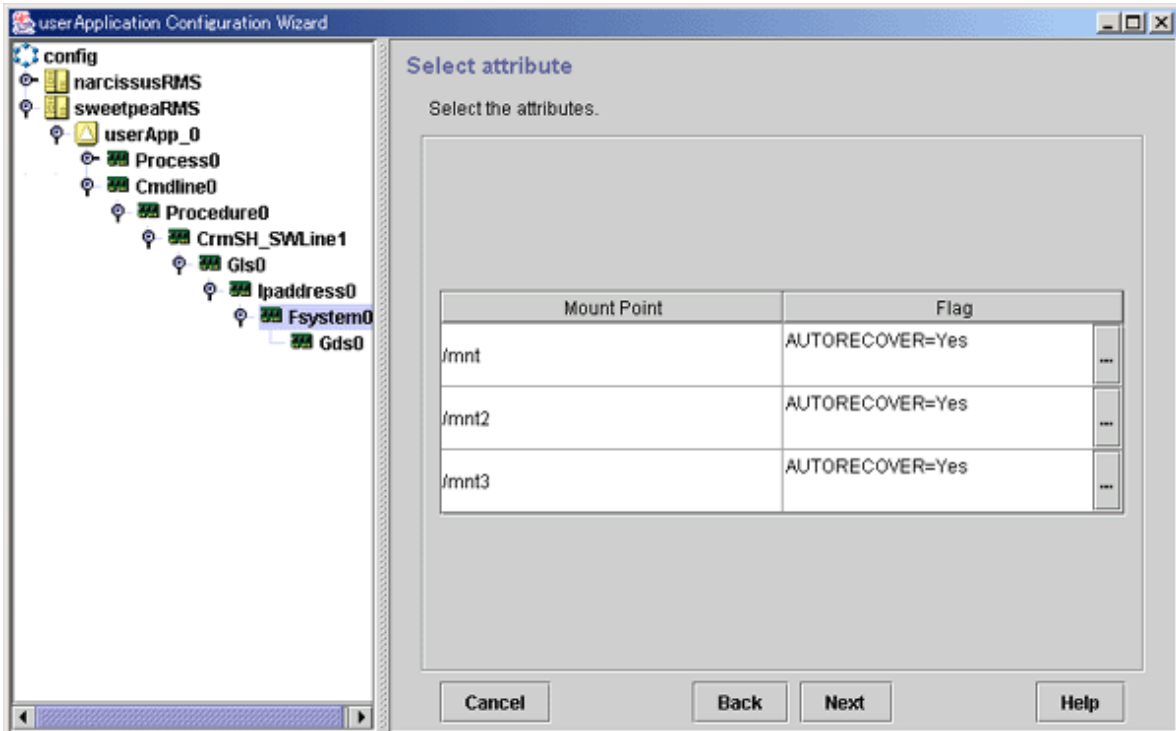
[...] button

Select this button to change an attribute that controls management of mount points. If you click this button, the "Change Mount Point Attributes" screen is displayed.
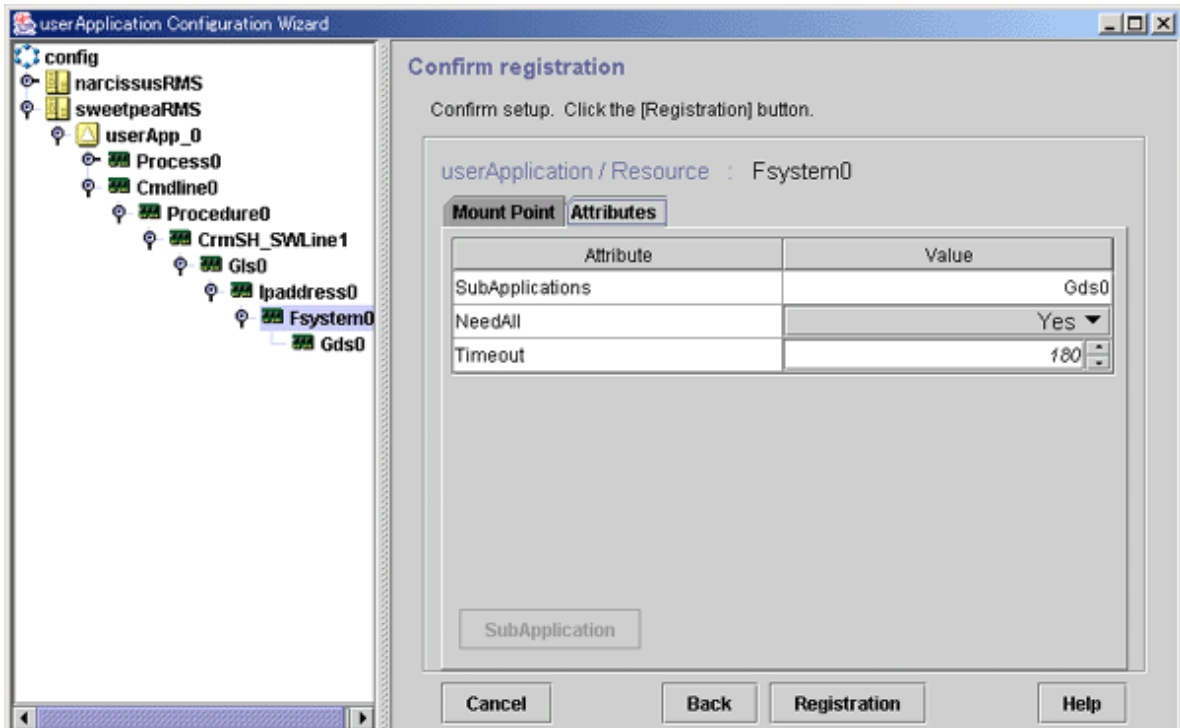
SubApplication button

This button is for specifying another resource that depends on the current resource. However, in Fsystem, this button is disabled.

Gds
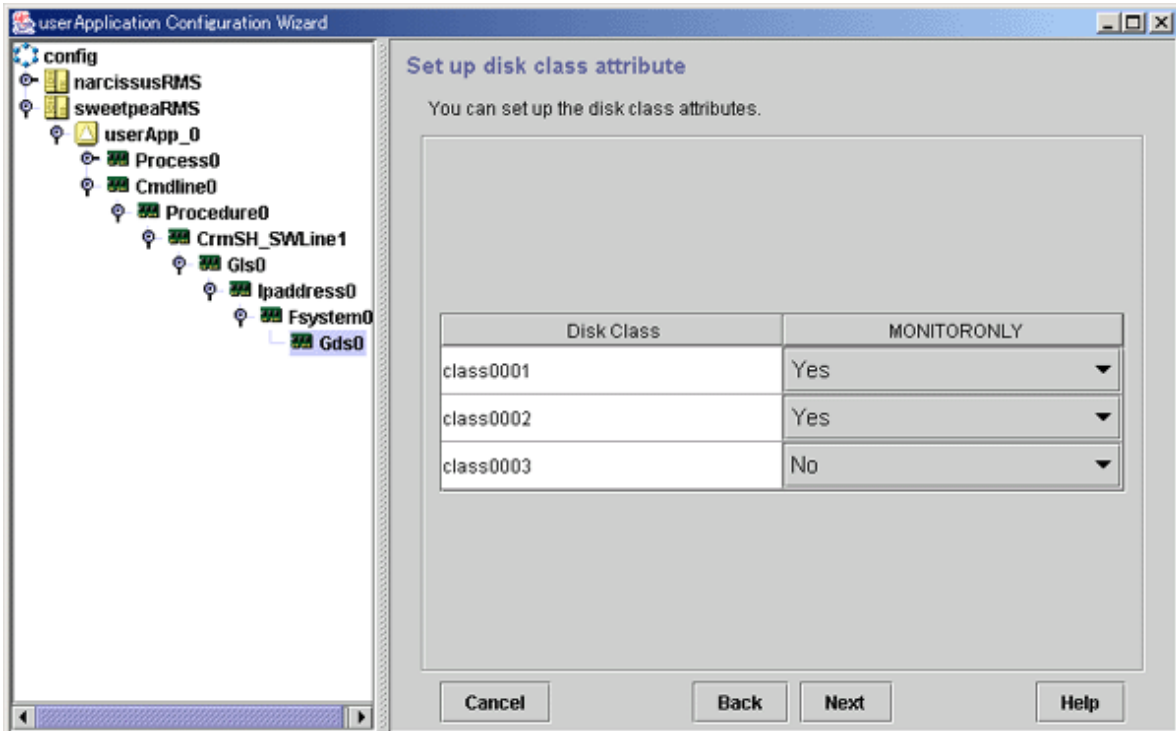
Flag	Outline
MONITORONLY	This flag sets whether the Faulted state of a disk class is to be reported to userApplication if the disk class fails. If this flag is set to "Yes" and a failure of the disk class occurs, the disk class is set to Faulted but the Gds resource remains Online and failover of userApplication does not occur.
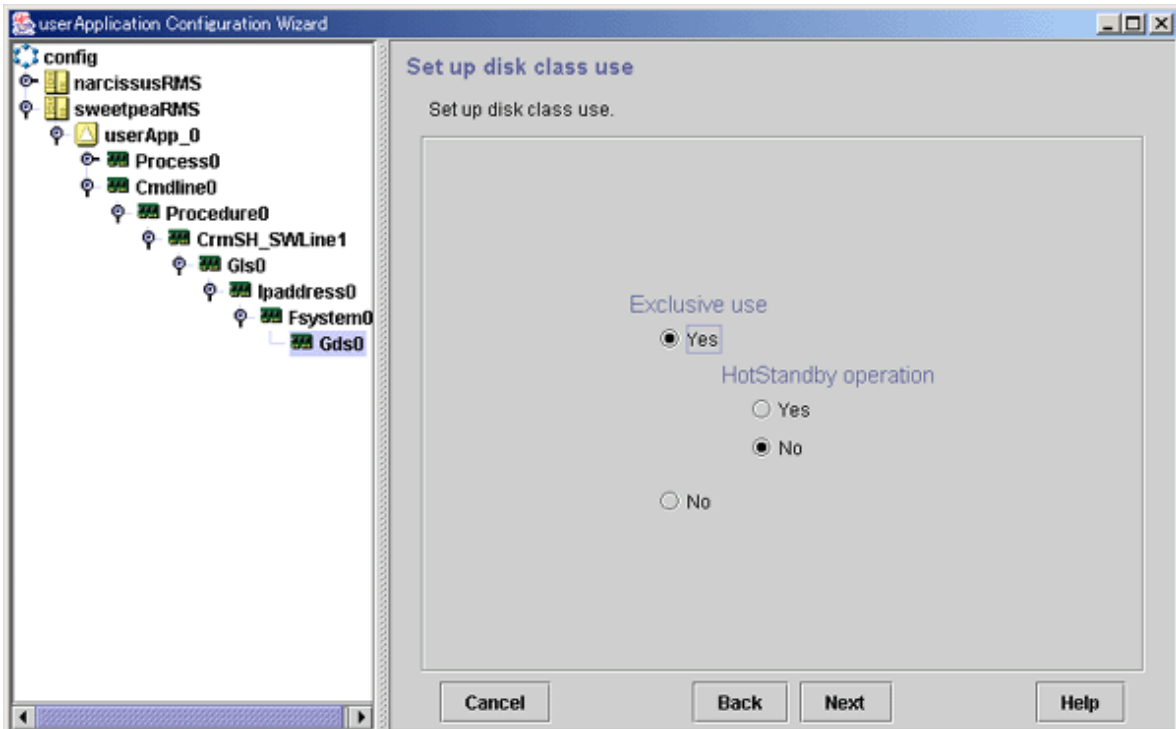
Exclusive use

Click "*Yes*" or "*No*" for *Exclusive use*. If you select "*Yes*", select "*Yes*" or "*No*" for *Hot-Standby operation*.
For information on the shared disk, see "Shared disk uses" in "6.6.1.3 Creating Gds Resources." Note that if you change *Exclusive use* from "*No*" to "*Yes*," there is the risk of disk damage if another Gds resource is using the same disk class.
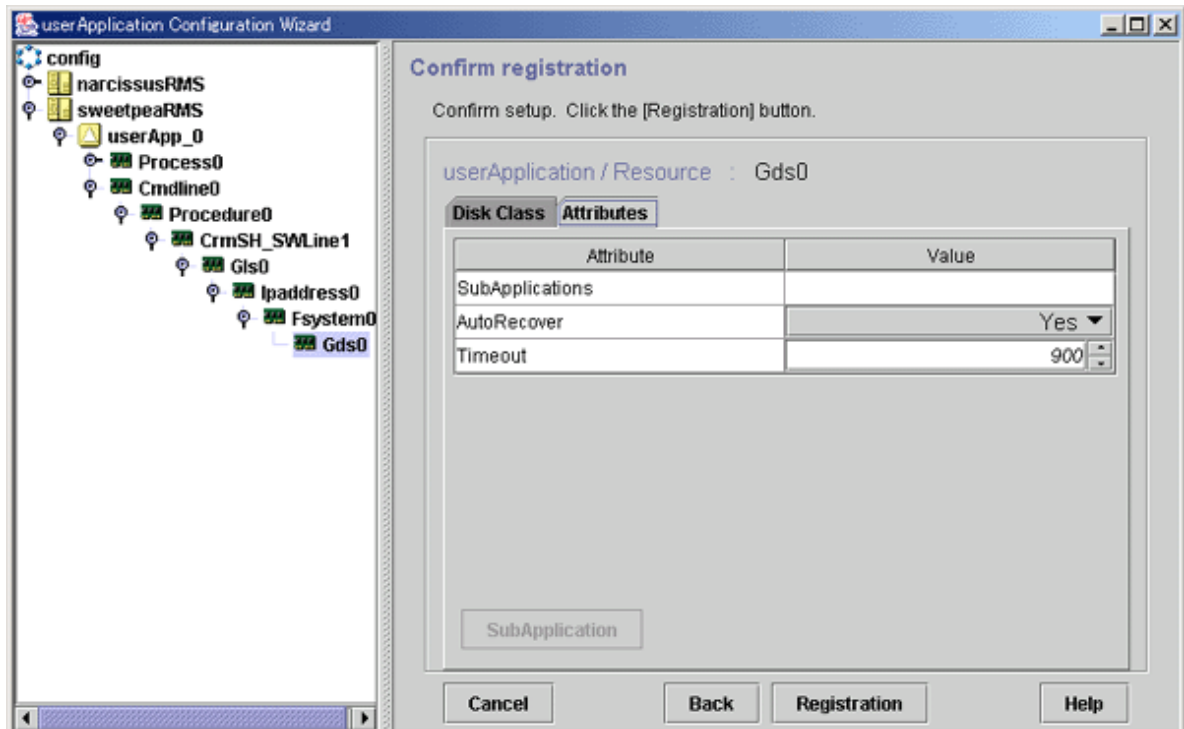
SubApplication button

Click this button to associate the other Gds resources with the Gds resource that has been created above. This button cannot be selected if the resource belongs to userApplication.

Gls
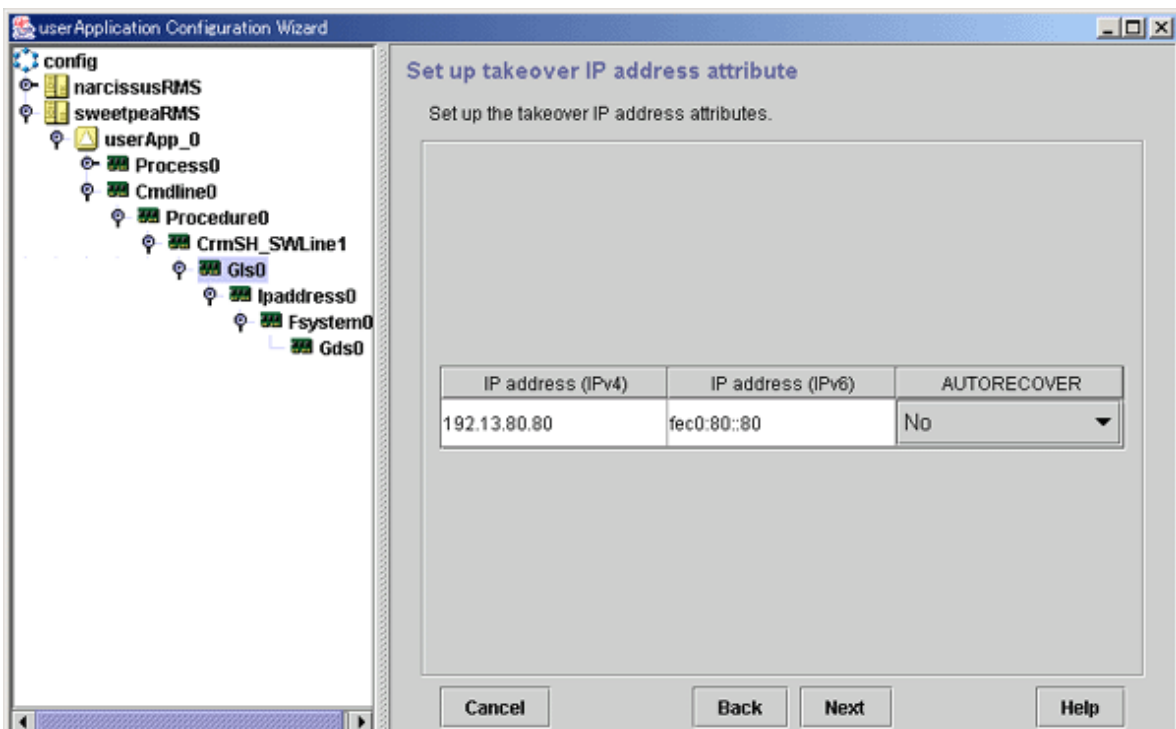
Flag	Outline
AUTORECOVER	If this is set to "Yes," RMS attempts to restore the faulted resource to prevent userApplication from being switched to other host. For GLS, this flag must be set to "No."
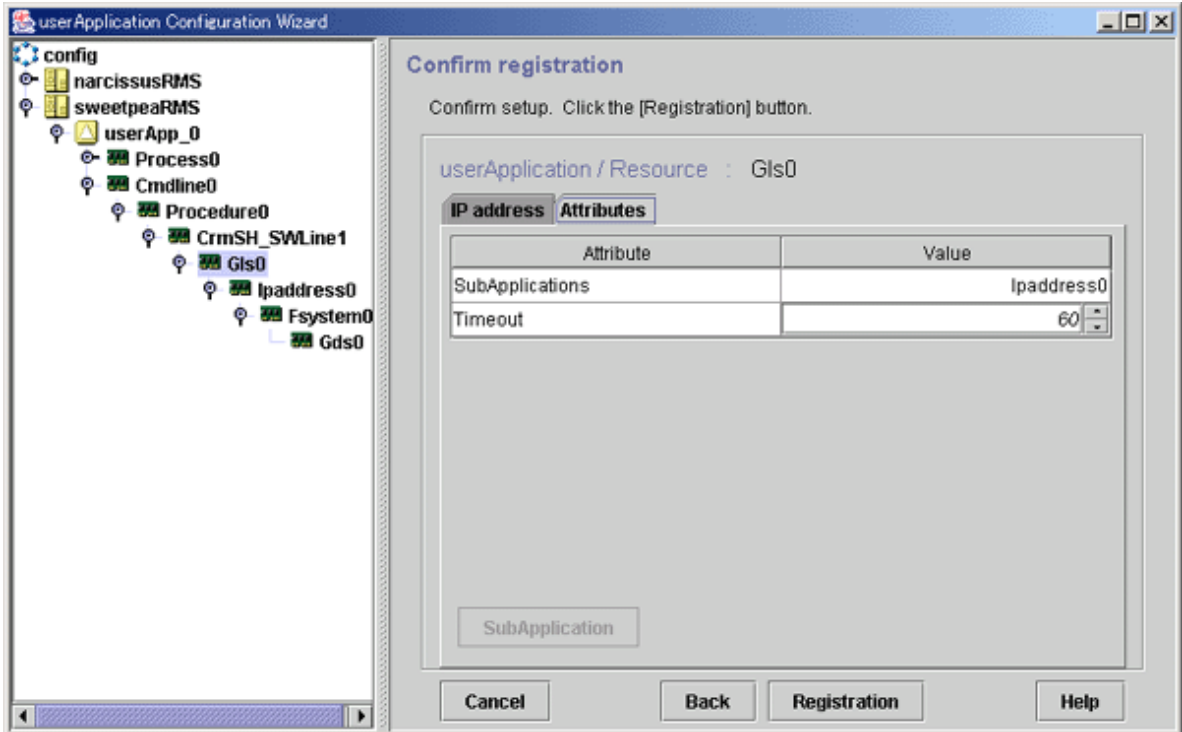
SubApplication button

Click this button to associate the other Gls resources with the Gls resource that has been created above. This button cannot be selected if the resource belongs to userApplication.

Ipaddress
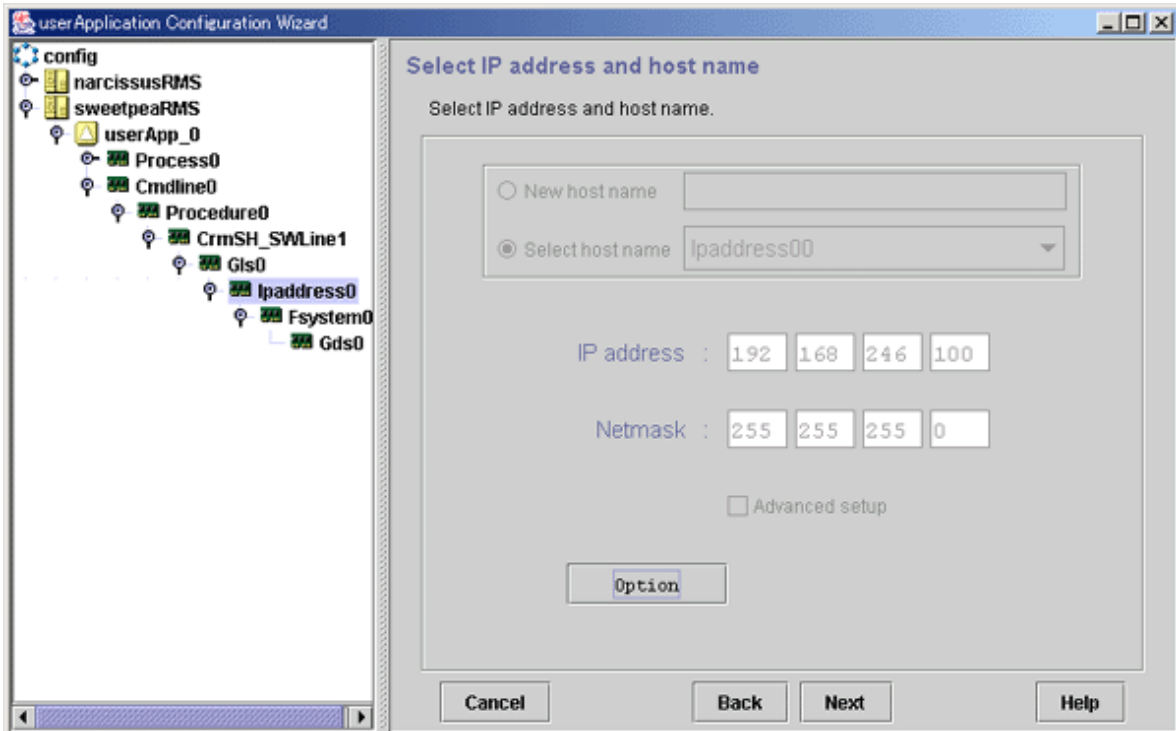
Options button

Use this button to change attributes of a takeover IP address. For operation instructions, see "Setting up Takeover IP Address Attributes."
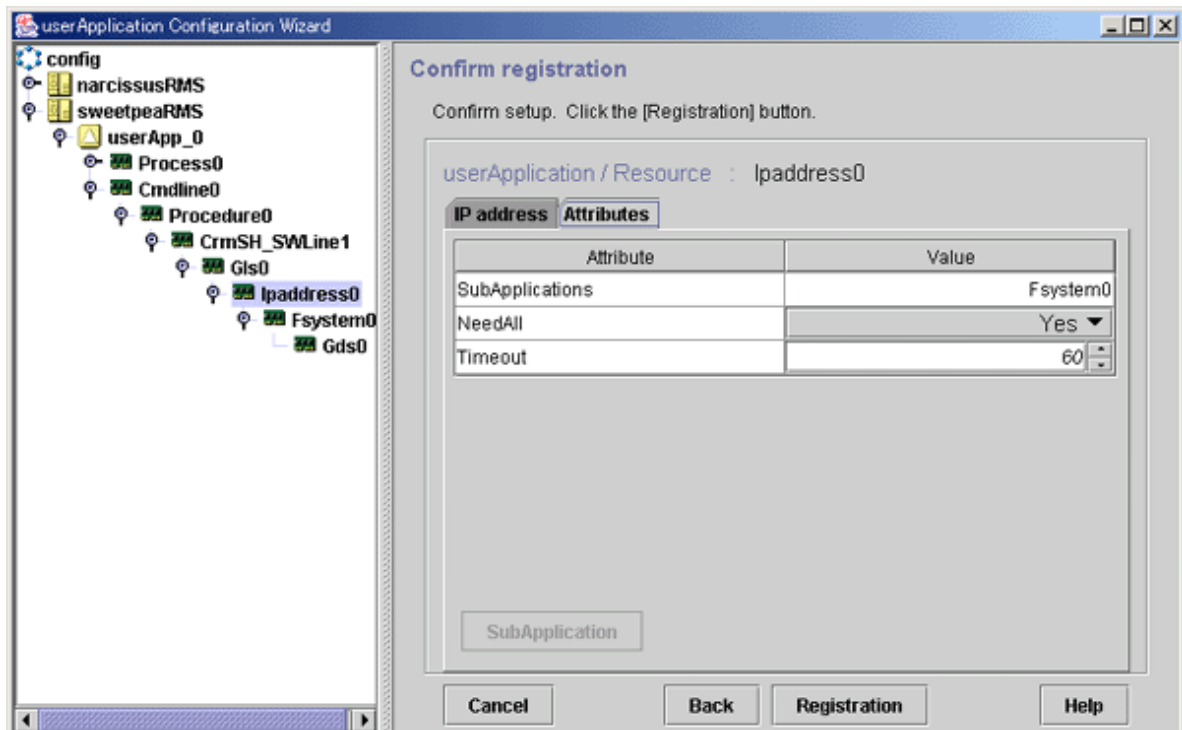
SubApplication button

Click this button to associate the other Ipaddress resources with the Ipaddress resource that has been created above. This button cannot be selected if the resource belongs to userApplication.
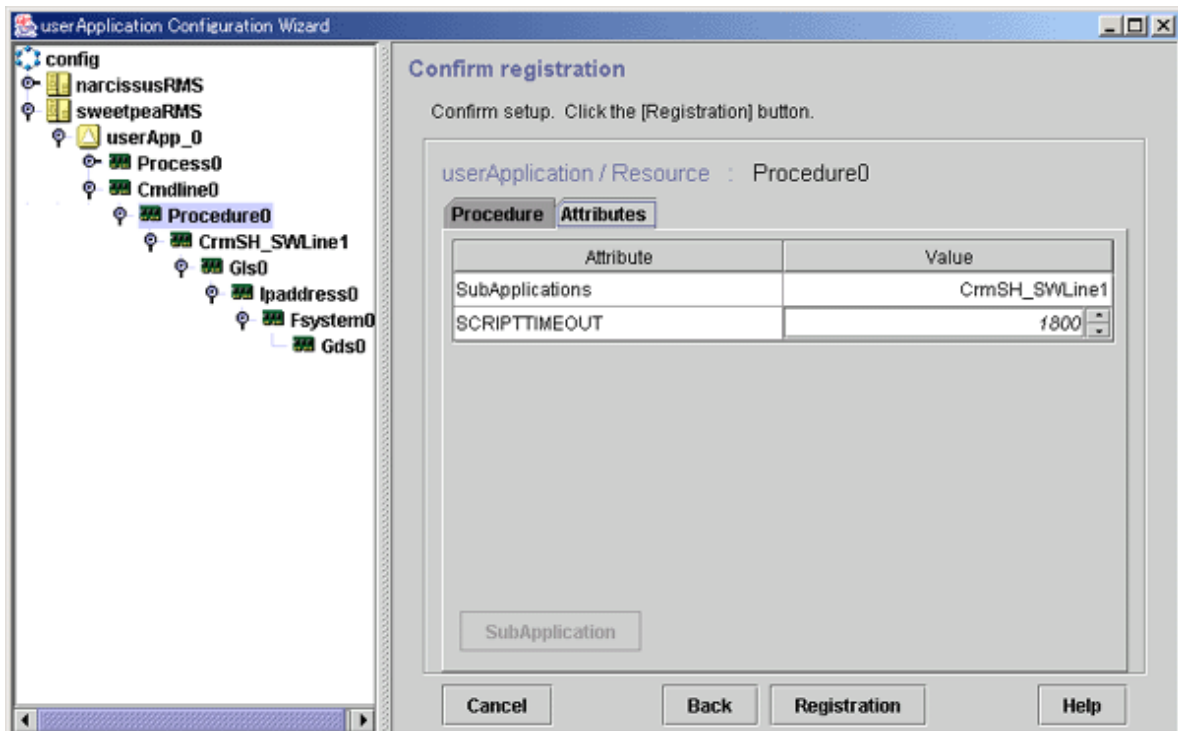
Procedure

SubApplication button

Click this button to associate the other Procedure resources with the Procedure resource that has been created above. This button cannot be selected if the resource belongs to userApplication.
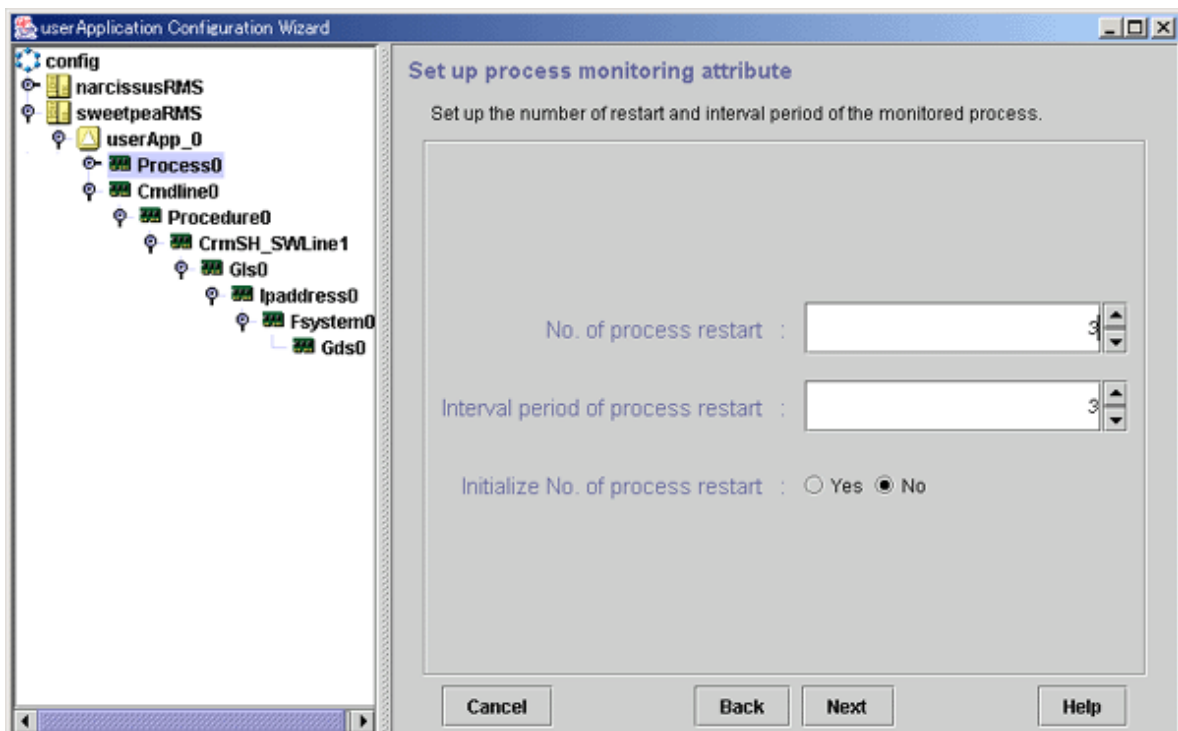
Process

No. of process restart

Specify the number of times the monitored process is to be restarted between 0 and 99 (default is 3). If 0 is specified and the monitored process stops, the process is not restarted and becomes Faulted.

Interval period of process restart

Specify the interval period when the process monitoring facility determines that the process has stopped until the facility executes restart. The specification range is between 0 and 3600 seconds (default is 3).

Initialize No. of process restart

Specify whether or not the counter that has the specified *No. of process restart* value in the process monitoring facility as its maximum value is to be initialized periodically. If you select "*Yes*," the counter is initialized every "*No. of process restart* value ´ 60 seconds." If you select "*No*," the counter is not initialized periodically.
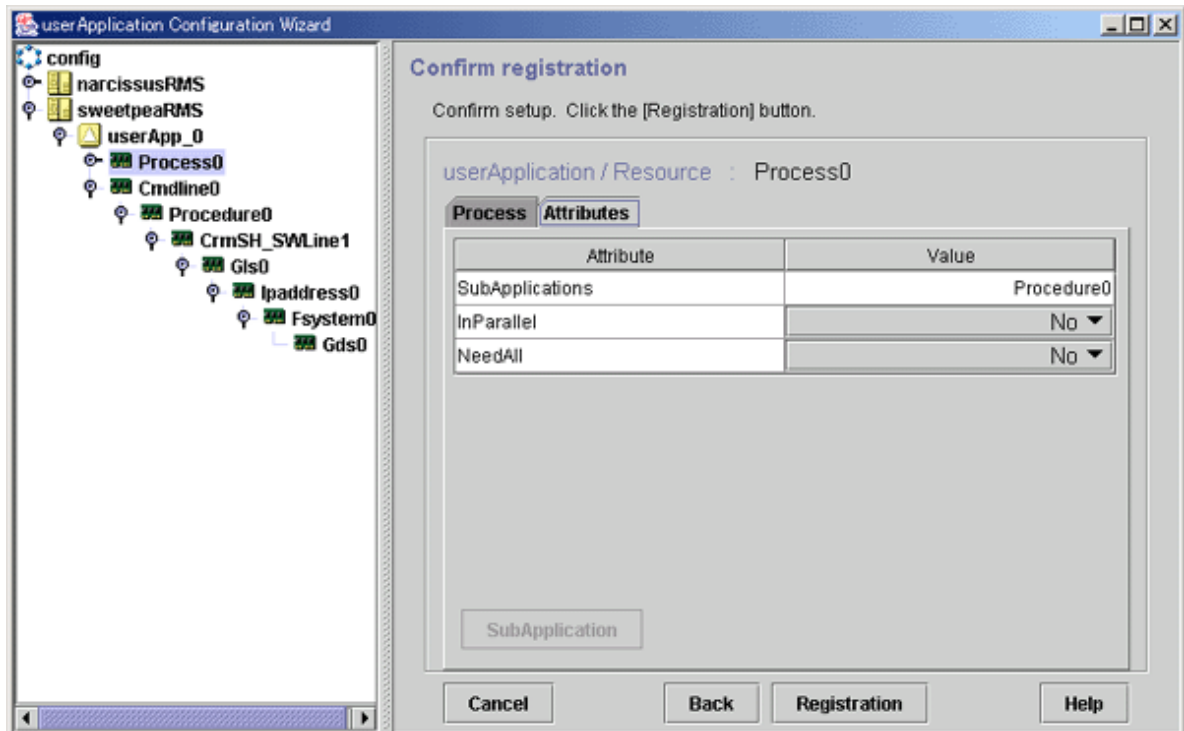
SubApplication button

Click this button to associate the Process resources or Cmdline resources with the Process resource that has been created above. This button cannot be selected if the resource belongs to userApplication.

SH_SWLine
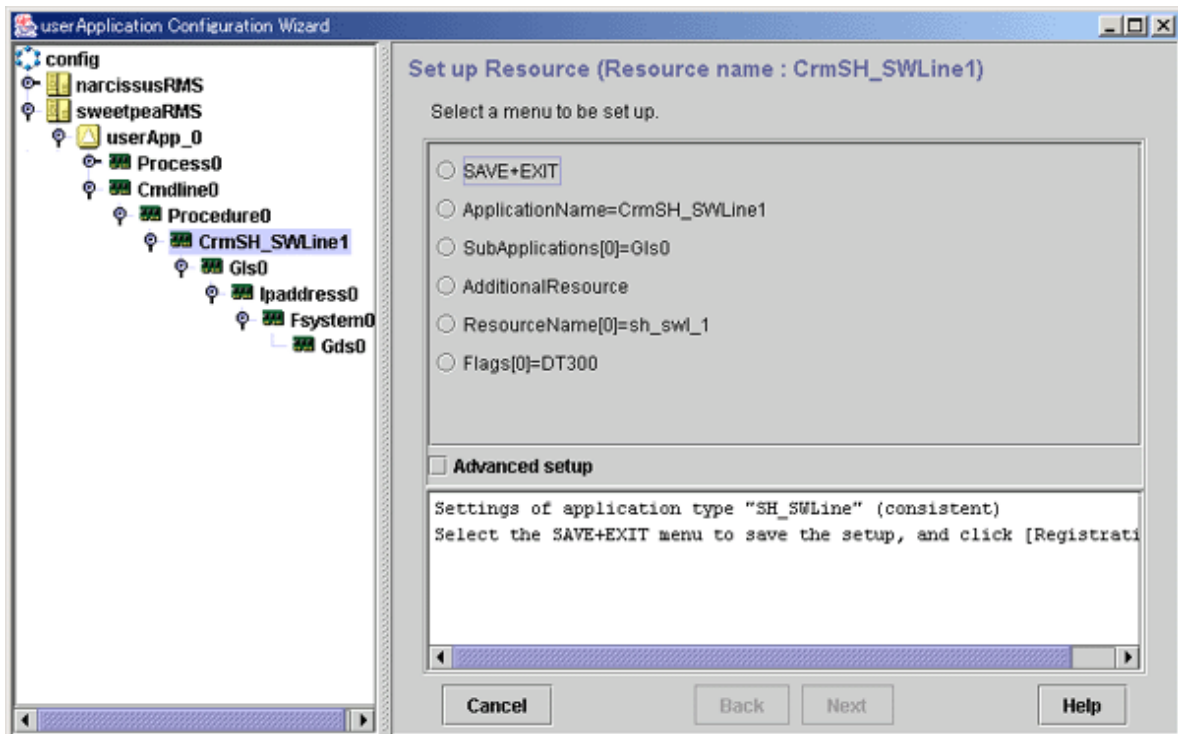
Advanced setup

Check this box to display attributes that do not need to be changed, during resource creation. Use this checkbox to display items such as the setup menu for the resource timeout value.

![Note icon] **Note**

. . . .

Specify "Application name" of the resource only if the resource is not associated with userApplication. If this resource is associated with userApplication, you must delete and then re-create the resource.

. . . .

ISV

For information on changing an ISV resource, see the manuals for the product.

![Note icon] **Note**

. . . .

Specify "Application name" of the resource only if the resource is not associated with userApplication. If this resource is associated with userApplication, you must delete and then re-create the resource.

. . . .

RMS Configuration distribution messages

After you finish changing the operation attributes, click the *Registration* button on the screen to register the new information to RMS Configuration.

If the system judges that registration to RMS Configuration is completed and RMS Configuration can be distributed, it displays the following message screen:
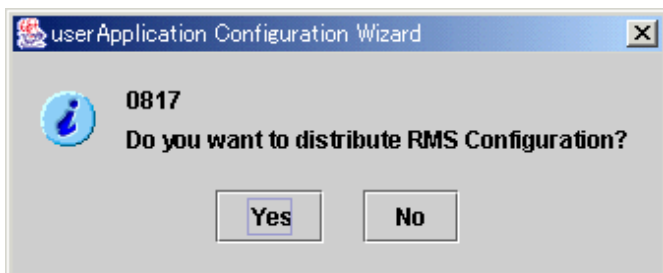
If you have completed making changes, click *Yes*.

8.2 Changing an IP Address on the Public LAN

This section explains how to change the IP address setting if the IP address of a public LAN or administrative LAN changes after the PRIMECLUSTER system is installed. Note that when you change an IP address, do not change the host name.

Operation Procedure:

1. Execute the following command on one of the cluster nodes to stop RMS operation:

```
# hvshut -a
```

2. Execute the following command on each cluster node to stop the shutdown facility:

```
# sdtool -e
```

3. If you use the shutdown facility to change an IP address when you change the public or administrative LAN, you must decide whether the /etc/opt/SMAW/SMAWsf/rcsd.cfg file must be changed on each cluster node. The file does not need to be changed if the host name registered to /etc/inet/hosts is defined in "admIP" of the rcsd.cfg file. The file must be changed if an IP address is defined in "admIP." In this case, change the "admIP" definition.
For details about the rcsd.cfg file, see the rcsd.cfg(4M) manual page.

4. To change the IP address of either the remote console connection unit or the XSCF when you change the public or administrative LAN, correct the /etc/inet/hosts file of each node.

5. To change the IP address of either the remote console connection unit or the XSCF when you change the public or administrative LAN, execute the following command in any cluster node and determine whether the IP address of either the remote console connection unit or the XSCF used by the shutdown facility must be changed. The IP address does not need to be changed if the host name registered to /etc/inet/hosts is displayed in the "IP-address" field of the command output
The IP address must be changed if an IP address is displayed in the "IP address" field.

```
# /etc/opt/FJSVcluster/bin/clrccusetup -l
```

Example: In the following example, the IP address must be changed.

```
# /etc/opt/FJSVcluster/bin/clrccusetup -l
device-name cluster-host-name  IP-address      host-name     user-name
-------------------------------------------------------------------
xscf       narcissus         10.10.10.11     -             user01
xscf       sweetpea          10.10.10.12     -             user02
```

6. If in Step 5 you determined that the IP address must be changed, execute the following command in each cluster node to change the IP address of the either remote console connection unit or the XSCF.
For *IP-address*, specify the IP address or host name of the remote console connection unit or XSCF.

```
# /etc/opt/FJSVcluster/bin/clrccusetup -m -i IP-address
```

Example: In the following example, the IP address of either the remote console connection unit or the XSCF is changed to 10.10.10.21.

```
# /etc/opt/FJSVcluster/bin/clrccusetup -m -i 10.10.10.21
```

7. Execute the following command in all nodes to start the system in single-user mode:

```
# /usr/sbin/shutdown -y -g0 -i0
      :
ok> boot -s
```

8. Execute the following command in all nodes to mount the file system:

```
# /usr/sbin/mountall -l
```

9. Edit the /etc/inet/hosts file, and change the IP address of each node.

10. On each node, change the IP address of the public LAN used by Web-Based Admin View.

![See icon] See
・・・
For details, see "7.1 Network address," "7.3 Management server," and "7.5 Multi-network between server and client by classified use" in the "*PRIMECLUSTER Web-Based Admin View Operation Guide*."
・・・

11. If a takeover IP address must be changed, correct the IP address being used as the takeover IP address in the "/etc/inet/hosts" file of each node.

12. If you used the userApplication Configuration Wizard to create an Ipaddress resource, edit the "/opt/SMAW/SMAWRrms/bin/hvenv.local" file on each cluster node as shown below to disable RMS automatic startup during system restart.

```
# vi /opt/SMAW/SMAWRrms/bin/hvenv.local <Return>
```

```
export HV_RCSTART=0
```

13. Restart the system on all nodes.

```
# /usr/sbin/shutdown -y -g0 -i6
```

14. If you used the userApplication Configuration Wizard to set up an Ipaddress resource, take the following steps to change the takeover IP address in the resource database.

    1. Identify the takeover IP address resource to be changed in the resource database.

       Execute the "clgettree(1)" command on one of the cluster nodes. The takeover IP address resource has the resource class name for IP addresses.

       Example) In the following example, the resources with resource IDs 56 and 57 are the takeover IP address resources.

```
# /etc/opt/FJSVcluster/bin/clgettree
Cluster 1 cluster
    Domain 2 RC2
            Shared 7 SHD_RC2
                    SHD_Host 58 config_Ipaddress0_ip1 UNKNOWN
                            IP_Address 56 hme0:1 UNKNOWN narcissus
                            IP_Address 57 hme0:1 UNKNOWN sweetpea
            Node 3 narcissus ON
                    Ethernet 21 hme0 UNKNOWN
                            IP_Address 56 hme0:1 UNKNOWN
                    Application 31 proc1 UNKNOWN
                    DISK 19 c0t0d0 UNKNOWN
            Node 5 sweetpea ON
                    Ethernet 22 hme0 UNKNOWN
                            IP_Address 57 hme0:1 UNKNOWN
                    Application 32 proc1 UNKNOWN
                    DISK 20 c0t0d0 UNKNOWN
```

2. Identify the shared resources of the takeover IP address.

To identify the shared resources, execute the "clgettree(1)" command on one cluster node.

The shared resources of the takeover IP address are resources that have the SHD_Host resource class name.

Example) For the example shown in Step 1), the resource that has resource ID 58 is the shared resource of the takeover IP address.

3. Change the takeover IP address.

To change the address, execute the "clsetrsc(1M)" command. The takeover IP address is defined in the attributes of the shared resource for the takeover IP address that was identified in Step 2.

On each cluster node, execute the "clsetrsc(1M)" command with the following format.

```
# /etc/opt/FJSVcluster/sys/clsetrsc -A ip_addr=new-IP-address resource-ID-of-shared-
resource-having-IP-address
```

Example) To change the takeover IP address of a shared resource (resource with resource ID=58) to 10.10.10.10

```
# /etc/opt/FJSVcluster/sys/clsetrsc -A ip_addr=10.10.10.10 58
```

4. Execute the "clgetrsc(1)" command to check that the takeover IP address was changed.

Using the command format below, execute the "clgetrsc(1)" command on one of the cluster nodes and check that the takeover IP address was changed.

```
# /etc/opt/FJSVcluster/sys/clgetrsc -a ip_addr resource-ID-of-shared-resource-having-IP-
address
```

Example) If the takeover IP address of a shared resource (resource with resource ID=58) was changed to 10.10.10.10

```
# /etc/opt/FJSVcluster/sys/clgetrsc -a ip_addr 58
ip_addr 10.10.10.10
```

15. If you edited the /opt/SMAW/SMAWRrms/bin/hvenv.local file in Step 12, return the file to its previous setting.

📄 Note

. . . .

- An interface or IP address to which MAC address takeover has already been configured cannot be used in Web-Based Admin View.

- If a takeover IP address has been defined by GLS and if the takeover IP address is specified directly with an IP address value, not a host name, in the GLS environment settings and the Gls resource settings, first delete the Gls resource. Then change /etc/inet/hosts and the GLS environment settings, and register the Gls resource.
For details, see "8.8.3 Supplement on Cluster Application and Resource Deletion," "6.1.1 GLS Setup," and "6.6.1.4 Creating Gls Resources."

- When the connection to XSCF is SSH connection, check that user inquiries of the first SSH connection (such as generation of the RSA key) have completed by connecting to XSCF from the cluster nodes via SSH using the log in user account for the shutdown facility after changing the IP address.

. . . .

8.3 Changing a CIP Address

This section describes how to change the IP address if the IP address of interconnect is changed after installation of the PRIMECLUSTER system.

Operation Procedure:

1. Start all the nodes that constitute the cluster system.

2. Stop CF on all nodes that constitute the cluster system.

   For information on how to stop CF, see "5.7 Starting and stopping CF" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide.*"

3. While referring to the cip.cf file, confirm the CIP name to change the IP address.

   For details on the cip.cf file, see "2.2 CIP Configuration File" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*" and the manual page describing cip.cf.

4. Change the IP address of the CIP name that is defined in the hosts(4) file.

   Perform this procedure on all the nodes constituting the cluster system.

5. Start CF on all the nodes constituting the cluster system.

   For instructions on how to start CF, see "5.7 Starting and stopping CF" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide.*"

6. Use the "ciptool(1M)" command to confirm that the IP address of CIP was changed.

```
# ciptool -a
```

 See
....

For details on the "ciptool(1M)" command, see the manual page describing ciptool(1M).
....

   1.

8.4 Changing Settings for the Shared Device Connection Confirmation Feature

For instructions on changing the operation environment for shared disk connection confirmation, see "5.3 Setting Up Shared Disk Connection Confirmation."

8.5 Changing the Operation Environment for Patrol Diagnosis

To change the operation environment for patrol diagnosis, execute the "clspconfig(1M)" command. For details on the "clspconfig(1M)" command, see the manual page for clspconfig(1M).

8.6 Changing a Node Name

The following explains how to change the node name after building a PRIMECLUSTER system.

 Note
....

Changing a node name may have a serious impact on the system. Therefore, make this change only when it is absolutely necessary.

Use the "sys-unconfig(1M)" command to change a node name. The "sys-unconfig(1M)" command is used to reset the system. All settings such as the name service, time zone, IP address, IP subnet mask, and root password are all erased and are then reset when the node next boots.
....

 See
....

For details on the "sys-unconfig(1M)" command, see the manual page describing sys-unconfig(1M).
....

Operation Procedure:

1. Stop CF on the node whose node name is to be changed.

   For details on how to stop CF, see "5.7 Starting and stopping CF" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide.*"

2. Define "export HV_RCSTART=0" in the hvenv.local file of the node whose node name is to be changed.

   This disables automatic startup RMS when the node is booted after changing the node name.

   For details on the hvenv.local file and the HV_RCSTART attribute, see the manual page describing hvenv.local.

3. Change the node name.

   1. Execute the "sys-unconfig(1M)" command.

      ```
      # sys-unconfig
      ```

   2. If you execute the "sys-unconfig(1M)" command, the system is shut down and the ok prompt will appear.
      Execute boot at the ok prompt, and then set up information such as node name according to the displayed message.

      ```
      ok boot
      ```

   📝 **Note**

   ....

   Do not attempt to change information other than node names with the "sys-unconfig(1M)" command.

   ....

4. Restore the hosts(4) file.

   Before the "sys-unconfig(1M)" command changes the system configuration, compare the contents of the backed-up /etc/inet/hosts.saved file with the hosts(4) file. Apply the necessary definitions included in the difference to the hosts(4) file.

5. Restart the system.

   ```
   # /usr/sbin/shutdown -y -i6 -g0
   ```

6. Set up the shutdown facility.

   For information on how to set up the shutdown facility, see "8.4 Configuring the Shutdown Facility" in the "*PRIMECLUSTER Cluster Foundation Configuration and Administration.*"

7. Start up RMS.

   For information on how to start up RMS, see "4.14 Starting RMS" in the "*PRIMECLUSTER RMS Configuration and Administration.*"

8. After confirming that RMS has started up, restore the definition of the HV_RCSTART environment variable corrected in step 2.

8.7 Changing PRIMECLUSTER Operation Attributes

The PRIMECLUSTER operation attributes are set with the RMS environment variables.

The environment variables are set in the RMS environment files shown below, and can be changed by editing the files directly to set or change the values. Also, you can reference the values that are currently set up.

File name	Contents
/opt/SMAW/SMAWRrms/bin/hvenv	**Global environment variable** The environment variable file that contains the variables used in the entire cluster
/opt/SMAW/SMAWRrms/bin/hvenv.local	**Local environment variable** The environment variable file that contains the variables defined only on the node

Note that the change procedure differs depending on whether the environment variable is global or local.

See

. . . .

- For information on the RMS environment variables and how to change the variable settings, see "2.9 Environment Variables" in "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide.*"

- For details on the RMS environment variables, see "10 Appendix - Environment Variables" in "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide.*"

- Create the hvenv.local file as required. When creating this file, set the permission of the file to 0644.
  For details on the hvenv.local file, see the manual page describing hvenv.local.

. . . .

8.8 Deleting a Cluster Application

This section explains how to delete a cluster application and its resources.

Note

. . . .

Be sure to stop RMS before deleting a cluster application and its resources. If the "userApplication Configuration Wizard" screen is started while RMS is operating, message 0833 will be displayed. If this happens, stop RMS, and then start the "userApplication Configuration Wizard" screen again. For instructions on stopping RMS, see "7.2.1.2 Stopping RMS."

. . . .

8.8.1 Deleting a userApplication

This section explains how to delete the userApplication.

Operation Procedure:

1. At the top menu of the *userApplication Configuration Wizard*, select userApplication to be deleted from the tree shown on the left of the screen.

2. Click the right mouse button to display the pop-up menu, and select *Remove userApplication or Resource*.
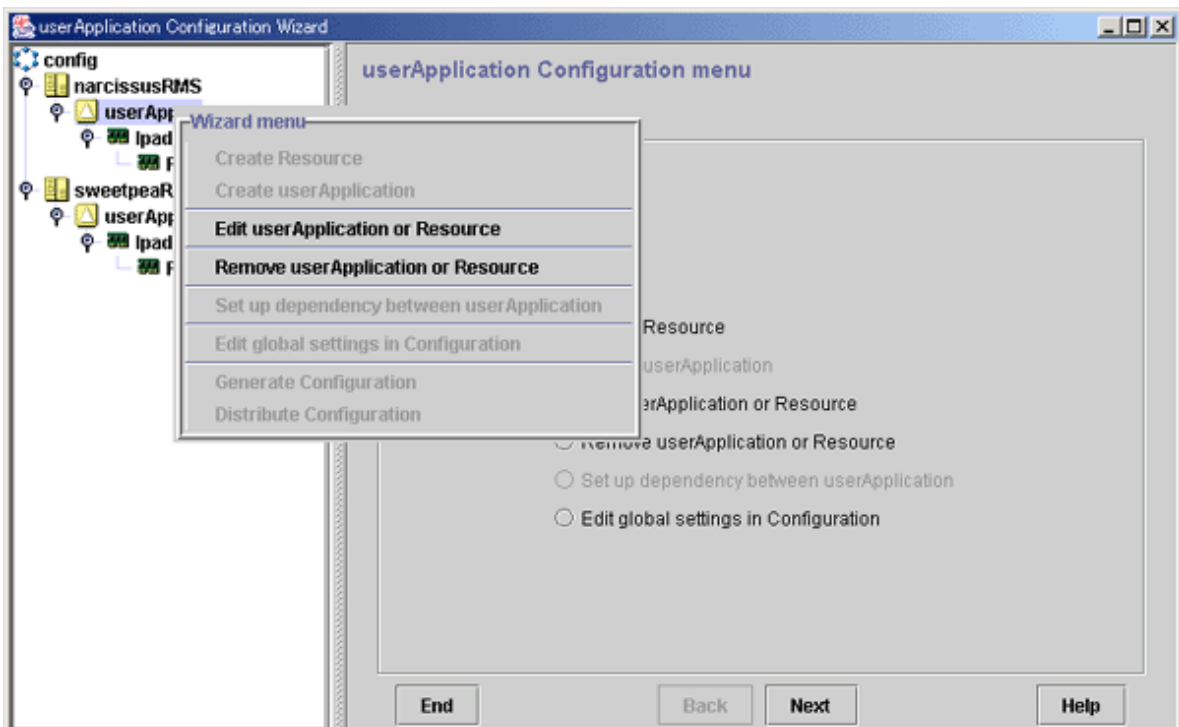
3. At the displayed dialog box, select *userApplication*. Only the userApplication will be deleted.
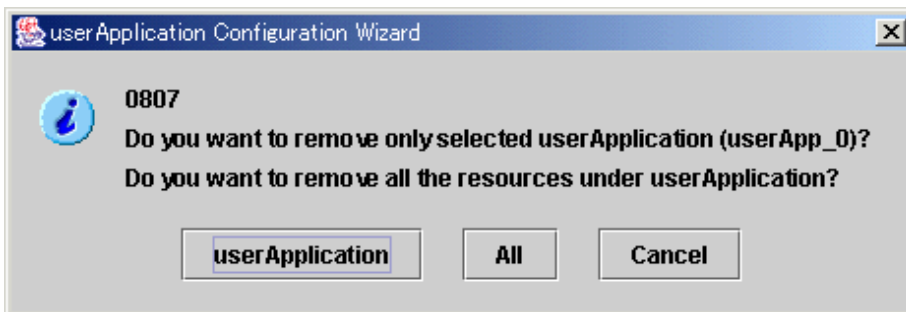
userApplication

Select this button to delete the specified userApplication.

![Note icon] Note

. . . .

The resources that are specified to userApplication will not be deleted.

The hierarchical relationships of the resources will be initialized. However, levels that were set with "Resource association" will not be initialized.

. . . .

All

Select this button to delete the specified userApplication and all related resources.
If resources are to be deleted, delete as described in the "8.8.3 Supplement on Cluster Application and Resource Deletion."

Cancel

Select this button to cancel the deletion.

You can confirm that userApplication has been deleted by observing that the userApplication no longer appears in the tree on the "userApplication Configuration Wizard" screen.

8.8.2 Deleting a Resource

This section explains how to delete a resource.

Operation Procedure:

1. At the top menu of the *userApplication Configuration Wizard*, select *Remove userApplication or Resource*.
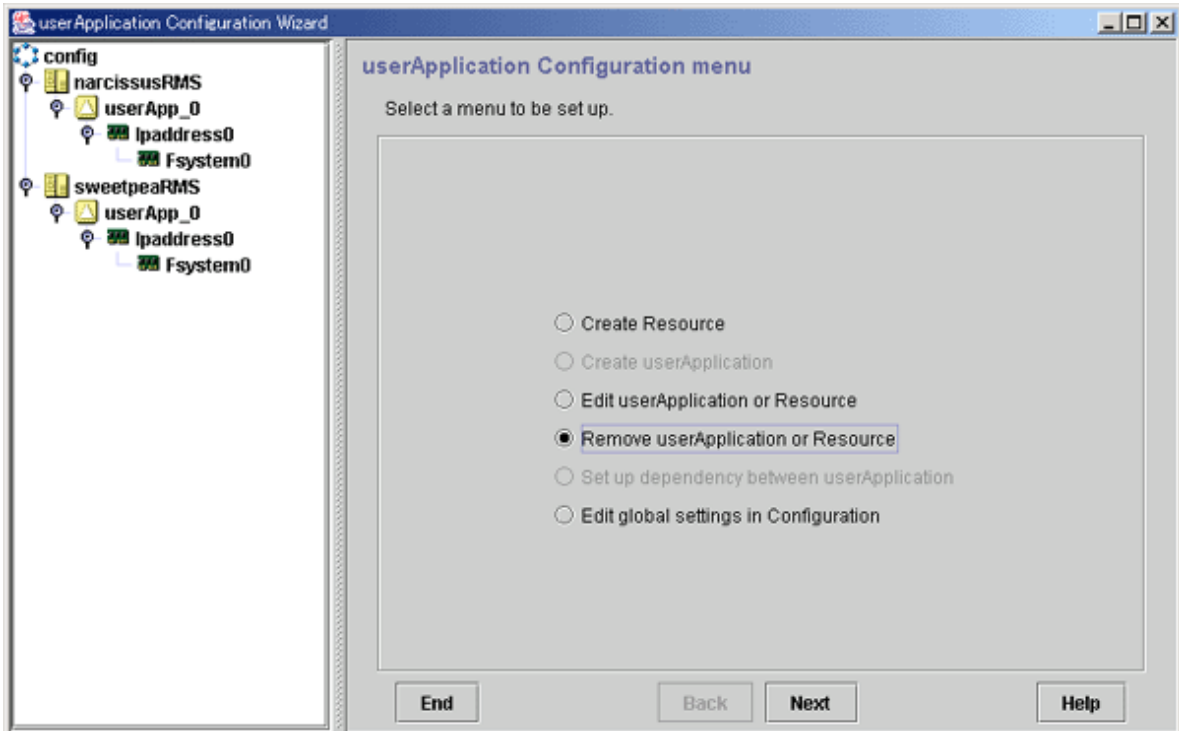
2. Select the resource to be deleted, and then select *Remove*.
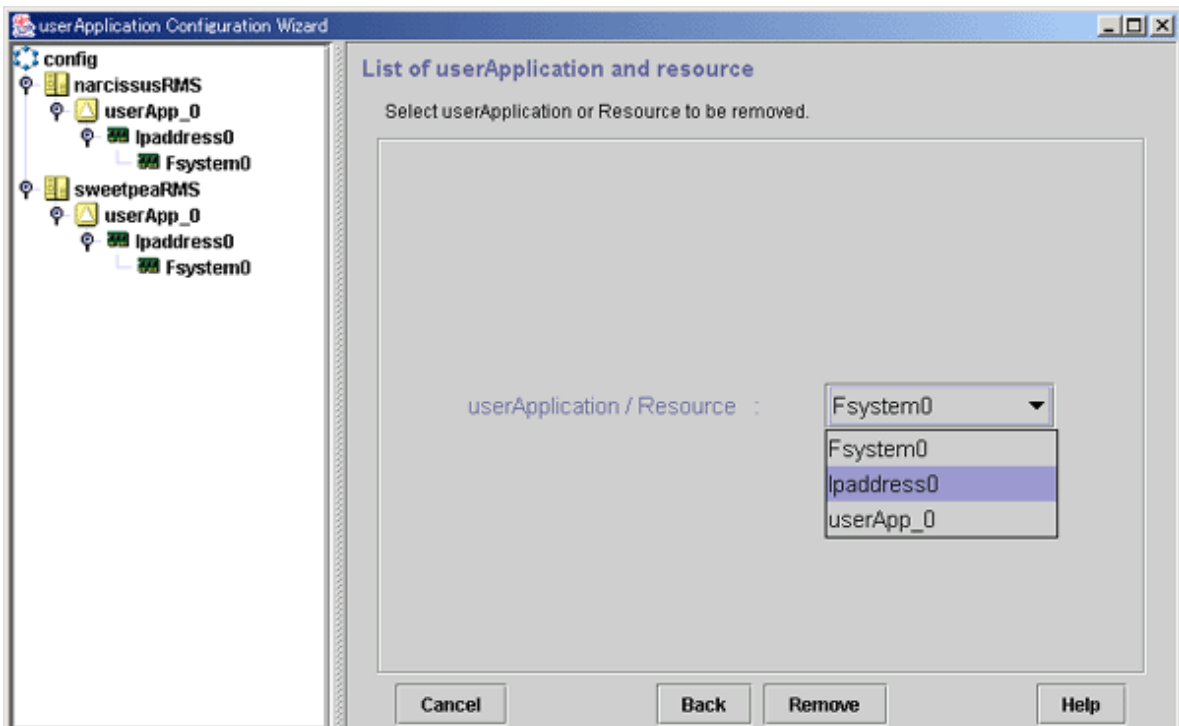
3. Respond to the displayed dialog box.

Yes

Select this button to delete the specified resource and all subordinate resources.
If deleting the selected resources deletes all the resources that construct userApplication, message 0835 will confirm with you whether you want to delete all the resources. To continue the deletion, select "*Yes*".

No

Select this button to cancel the deletion.

You can confirm the deletion by checking that the resource disappears in the tree on the "userApplication Configuration Wizard" screen.

Note
When deleting resources, refer to the "8.8.3 Supplement on Cluster Application and Resource Deletion."

8.8.3 Supplement on Cluster Application and Resource Deletion

This section explains the operations to be done after "8.8.1 Deleting a userApplication" and "8.8.2 Deleting a Resource."

8.8.3.1 Supplement on Cmdline resource deletion

Delete each script used in the Cmdline resources.

Note
Before deleting scripts, you need to delete the Cmdline resources.

Each script is stored in the following directories:

- /etc/opt/FJSVwvucw/scripts/start

- /etc/opt/FJSVwvucw/scripts/stop

- /etc/opt/FJSVwvucw/scripts/check

Execute the deletion command as follows:

```
# cd /etc/opt/FJSVwvucw/scripts/start

# rm <script name to be deleted>

# cd /etc/opt/FJSVwvucw/scripts/stop

# rm <script name to be deleted>

# cd /etc/opt/FJSVwvucw/scripts/check

# rm <script name to be deleted>
```

8.8.3.2 Supplement on Fsystem resource deletion

Restore the environment files related to the mount point to the original state.

Note

Before deleting the environment files, you need to delete the Fsystem resources.

Deleting mount point entries that were added at prerequisites

Delete the relevant entry from the following environment files that were edited as part of the "6.6.1.2.1 Prerequisites."

- /etc/vfstab

- /etc/dfs/dfstab

If a dummy entry is deleted from /etc/dfs/dfstab, delete the actual directory.

When an NFS Lock Failover was used

If the NFS Lock Failover functions are no longer used due to deletion of the Fsystem resources, take the following steps:

Procedure:

1. Change the settings for shared information with the configuration.

   Change the value of NFSLockFailover to "no."

See

See "6.6.4 Editing global settings in Configuration."

8.8.3.3 Supplement on Gds resource deletion

Delete the disk class if it is no longer necessary.

See the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide.*"

8.8.3.4 Supplement related to Gls resource deletion

Delete the takeover network definition if it is no longer necessary.

See "5.4 Deleting configuration for Cluster System" in the "*PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function.*"

8.8.3.5 Supplement on takeover network resource deletion

Restore the environment files related to the takeover network to the original state.

When MAC address takeover was used

Delete the entries that were added to the following environment files during the prerequisite procedure of the MAC address takeover.

- /etc/opt/FSUNnet/mactool/macaddr.conf

- /usr/opt/reliant/etc/hvipalias

- /etc/inet/hosts

For details on the added entries, see "6.6.1.5.1 Prerequisites."

When node name takeover was used

Change the node name in the cluster system used for node name takeover. To bring the system state back to the state before the change, take the following steps on all the nodes where node name takeover resources were created.

Operation

1. Execute the command on all the nodes, as follows:

```
# setuname -n fuji2
```

fuji2 is used as the node name in this example.

2. Reboot the system.

```
# /usr/sbin/shutdown -g0 -i6 -y
```

8.8.3.6 Supplement on procedure resource deletion

Delete the procedure resource from the cluster resource management facility when it is no longer necessary.

See "F.3 Deleting a Procedure Resource."

8.8.3.7 Supplement on process monitoring resource deletion

Delete scripts that were created automatically if the stop command of the process monitoring resource was used.

![Note icon] **Note**

. . . .

Before deleting the created scripts, you need to delete the process monitoring resources.

. . . .

Each script is stored in the /etc/opt/FJSVwvucw/scripts/process_stop directory.

Execute the deletion command as shown below.

```
# cd /etc/opt/FJSVwvucw/scripts/process_stop
# rm ./*
```

8.8.3.8 Supplement on line switching unit resource deletion

Delete the line switching unit resource from the cluster resource management facility if it is not necessary.

See the manual pages of the "cldelswursc(1M)" command.

8.8.3.9 ISV resource deletion procedure (supplement)

For information on the deletion procedure, see the manuals provided with each product.

8.9 Deleting the Hardware Resource

Delete the hardware resource that was registered with the automatic configuration facility using the following procedure.

Procedure

1. Stop RMS.

If RMS is activated, stop RMS as explained in "7.2.1.2 Stopping RMS."

2. Delete the hardware resource.

Use each command to delete the hardware resource that was registered with the automatic configuration facility.

See the relevant manual pages for details on each command.

Resource	Command
Shared disk device	cldeldevice(1M)
Network interface card	cldelrsc(1M)
Line switching unit	cldelswursc(1M)

Note

....

- If you have deleted an available network interface card by mistake, reregister the resources for the accidentally deleted network interface card by executing the "clautoconfig(1M)" command.

- If you have deleted other than the correct network takeover resource, reregister the deleted network takeover resource. See "G.2 Associating Resources to a Cluster Service Instance."

....

8.10 Deleting and Changing Hardware

This section explains how to delete and change hardware in the existing configuration.

Note

....

Do not restart the network when changing a system board or a network interface card.

....

8.10.1 Deleting Hardware

This section describes how to delete the following hardware from the existing configuration:

- Shared disk device

- Network interface card used for the public LAN and the administrative LAN

8.10.1.1 Deleting a shared disk device

To delete a shared disk device, you need to change a cluster application which includes resources of the shared disk device to be deleted beforehand.

Figure 8.1 Procedure to delete a shared disk device
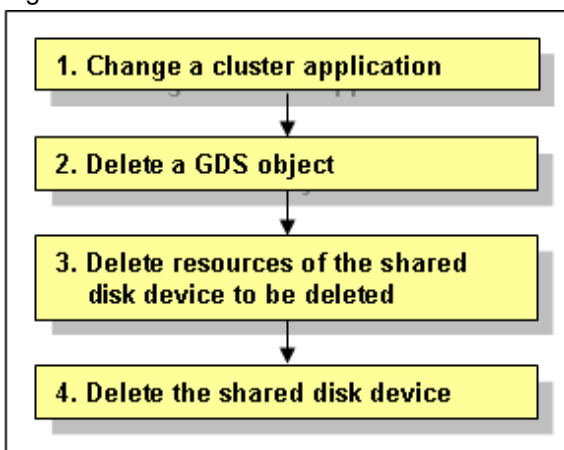
1. Change a cluster application.

   Delete the following resources using the shared disk device to be deleted from a cluster application:

   - Fsystem resource

- Gds resource

See

See
....
To change the configuration of a cluster application and delete a resource, see "8.1.1 Changing the Cluster Application Configuration."
....

2. Delete a GDS object.

   Delete a GDS object related to the shared disk device to be deleted.

See
....
To delete a GDS object, see the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*."
....

3. Delete resources of the shared disk device to be deleted.

   Delete resources of the registered shared disk device by using the "cldeldevice" command.

   For details on the "cldeldevice" command, see the manual page.

Note
....
When resources of the shared disk device to be deleted are registered to a GDS class, delete the shared disk device from the GDS class first, and then delete resources of the shared disk device. To delete the shared disk device from the GDS class, see the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*."
....

4. Delete the shared disk device.

   See "10.2 Maintenance Flow" and ask your Fujitsu customer engineer (CE) to delete the shared disk device.

8.10.1.2 Deleting a network interface card used for the public LAN and the administrative LAN

To delete a network interface card used for the public LAN and the administrative LAN, you need to change a cluster application which includes resources of the network interface card to be deleted beforehand.

Figure 8.2 Procedure to delete a network interface card
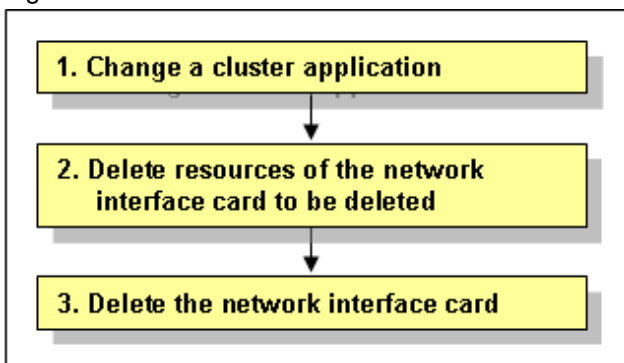
1. Change a cluster application.

   Delete the following resources using the network interface card to be deleted from a cluster application:

   - Takeover network resource

   - Gls resource

2. Delete resources of the network interface card.

   Delete resources of the registered network interface card by using the "cldelrsc" command.

   For details on the "cldelrsc" command, see the manual page.

3. Delete the network interface card.

   See "10.2 Maintenance Flow" and ask your Fujitsu customer engineer (CE) to delete the network interface card.

8.10.2 Changing Hardware

This section describes how to change the following hardware in the existing configuration:

- Shared disk device

- Network interface card used for the public LAN and the administrative LAN

8.10.2.1 Changing a shared disk device

To change a shared disk device, you need to delete a cluster application which includes resources of the shared disk device to be changed beforehand, and then create the cluster application again after changing the shared disk device.

Figure 8.3 Procedure to change a shared disk device
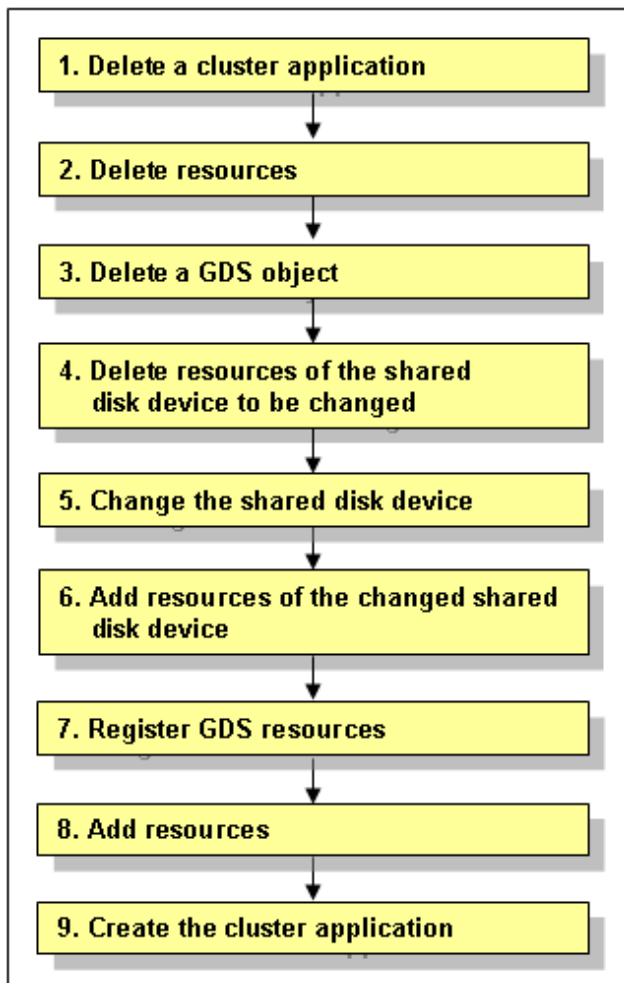
1. Delete a cluster application.

> 📖 **Note**
>
> To delete a cluster application, you must stop RMS beforehand. If RMS is running, see "7.2.1.2 Stopping RMS" and stop RMS.

Delete the cluster application which includes the following resources related to the shared disk device to be changed:

- Fsystem resource

- Gds resource

> 📖 **See**
>
> To delete a cluster application, see "8.8.1 Deleting a userApplication."

2. Delete resources.

Delete Fsystem resources and Gds resources related to the shared disk device to be changed.

> 📖 **See**
>
> To delete resources, see "8.8.2 Deleting a Resource."

3. Delete a GDS object.

Delete a GDS object related to the shared disk device to be changed.

> 📖 **See**
>
> To delete a GDS object, see the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide.*"

4. Delete resources of the shared disk device to be changed.

Delete resources of the registered shared disk device by using the "cldeldevice(1M)" command.

For details on the "cldeldevice(1M)" command, see the manual page.

> 📖 **Note**
>
> When resources of the shared disk device to be deleted are registered to a GDS class, delete the shared disk device from the GDS class first, and then delete resources of the shared disk device. To delete the shared disk device from the GDS class, see the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide.*"

5. Change the shared disk device.

See "10.2 Maintenance Flow" and ask your Fujitsu customer engineer (CE) to change the shared disk device.

6. Add resources of the changed shared disk device.

Register resources corresponding to the changed shared disk device to the resource database with the automatic configuration facility.

> 📖 **See**
>
> For information on automatic configuration of resources, see "5.1.3.2 Automatic Configure."

7. Register Gds resources.

To use Global Disk Services (GDS), set up GDS and create Gds resources.

See

For information on how to set up GDS and create Gds resources, see "6.2.2 GDS Configuration Setup" and "6.6.1.3 Creating Gds Resources."

8. Add resources.

If you have deleted Fsystem resources in Step 2, add Fsystem resources.

See

To add resources, see "6.6.1 Setting Up Resources."

9. Create the cluster application.

Create the cluster application deleted in Step 1 again.

If cluster application dependency relationships have been set up, they must be set up again.

After creating the cluster application, execute "RMS startup" from the Cluster Admin screen. If necessary, also start the cluster application.

See

To create a cluster application, see "6.6.2 Creating Cluster Applications."

For information on how to set up cluster application dependency relationships, see "6.6.3 Setting Up Dependency Relationships Between Cluster Applications."

For instructions on starting RMS, see "7.2.1.1 Starting RMS." For instructions on starting the cluster application, see "7.2.2.1 Starting a Cluster Application."

8.10.2.2 Changing a network interface card used for the public LAN and the administrative LAN

To change a network interface card used for the public LAN and the administrative LAN, you need to delete a cluster application which includes resources of the network interface card to be changed beforehand, and then create the cluster application again after changing the network interface card.

Figure 8.4 Procedure to change a network interface card
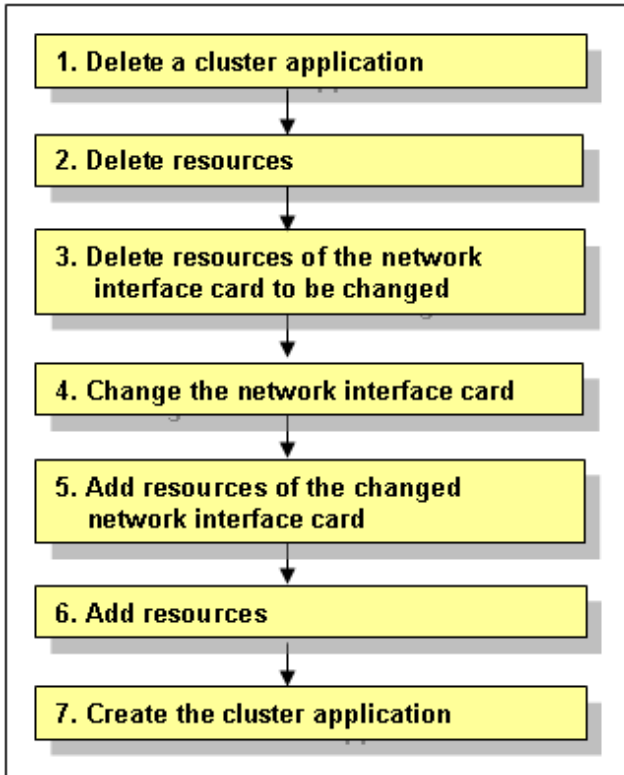
1. Delete a cluster application.

   > 📌 **Note**
   > ....
   > To delete a cluster application, you must stop RMS beforehand. If RMS is running, see "7.2.1.2 Stopping RMS" and stop RMS.
   > ....

   Delete the cluster application which includes the following resources related to the network interface card to be changed:

   - Takeover network resource

   - Gls resource

   > 🗂 **See**
   > ....
   > To delete a cluster application, see "8.8.1 Deleting a userApplication."
   > ....

2. Delete resources.

   Delete takeover network resources and Gls resources related to the network interface card to be changed.

   > 🗂 **See**
   > ....
   > To delete resources, see "8.8.2 Deleting a Resource."
   > ....

3. Delete resources of the network interface card to be changed.

   Delete resources of the registered network interface card by using the "cldelrsc" command.

   For details on the "cldelrsc" command, see the manual page.

4. Changing the network interface card.

   See "10.2 Maintenance Flow" and ask your Fujitsu customer engineer (CE) to change the network interface card.

5. Add resources of the changed network interface card.

   Register resources corresponding to the changed network interface card to the resource database with the automatic configuration facility.

   ## See

   For information on automatic configuration of resources, see "5.1.3.2 Automatic Configure."

6. Add resources.

   If you have deleted takeover network resources and Gls resources in Step 2, add takeover network resources and Gls resources.

   ## See

   To add resources, see "6.6.1 Setting Up Resources."

7. Create the cluster application.

   Create the cluster application deleted in Step 1 again.

   If cluster application dependency relationships have been set up, they must be set up again.

   After creating the cluster application, execute "RMS startup" from the Cluster Admin screen. If necessary, also start the cluster application.

   ## See

   To create a cluster application, see "6.6.2 Creating Cluster Applications."

   For information on how to set up cluster application dependency relationships, see "6.6.3 Setting Up Dependency Relationships Between Cluster Applications."

   For instructions on starting RMS, see "7.2.1.1 Starting RMS." For instructions on starting the cluster application, see "7.2.2.1 Starting a Cluster Application."

Chapter 9 Expanding the Operation Configuration

This chapter explains how to expand a cluster configuration.

9.1 Adding a Node

Node expansion means adding on an extra node to an operating cluster system. The purpose of node expansion is to provide high availability and extend available business operations.

The node expansion procedure is explained below:

(1) Setting up a new node

1) Installing OS

↓

2) Installing PRIMECLUSTER

↓

3) Registering the host information for the existing node

(2) Preparing existing nodes

1) Backing up the entire system

↓

2) Backing up the files that configure the cluster system and its components

↓

3) Registering host information for the new node

↓

4) Setting up to avoid automatic startup of the RMS upon a reboot

↓

5) Turning off the power to the system

(3) Connecting a node

1) Connecting a LAN, cluster interconnector, and RCI of the new node to the existing cluster environment

↓

2) Confirming the connection

(4) Setting up Web-based Admin View

↓

(5) Setting up CF and SF

↓

(6) Setting up CRM

To use the volume of an existing node by a new node:

······→ **(7) Setting up GDS**

To change the configuration of the user application is to be changed:

······→ **(8) Setting up RMS**

1) Deleting userApplication

↓

2) Recreating resources

↓

3) Recreating userApplication

To use the GFS shared file system for the existing node:

······→ **(9) Setting up GFS**

🔰 **Note**

· ·

- Use the same OS version, collective updates, and patches as those of the existing cluster nodes.

- Stop operation when nodes are to be added.

- Two or more nodes cannot be added at the same time. When multiple nodes are to be added, add them one by one.

- The nodes to be added must be of the same model as the existing cluster nodes.

In the following explanation, node1 and node2 are used as the node names of the existing cluster nodes while node3 is used as the node name of the new node.

9.1.1 Setting Up a New Node

Take the following steps to set up the new node.

The new node should be prepared such that the operating system, PTFs, FibreChannel, and packages such as the multipath software have already been installed.

Procedure

1. Install PRIMECLUSTER on the new node.

   This must be done in a single user mode.
   For details, see the "*PRIMECLUSTER Installation Guide*."

2. Configure NTP.

   Configure NTP for the new node to match the NTP of the existing nodes.

3. Define the following information in the "/etc/inet/hosts" file of the new node.

   - IP address of the existing cluster nodes and the host name that is associated with the IP address

   - IP address of the remote console that is connected to the existing cluster nodes and the host name that is associated with the IP address

   - The IP address of the CIP interface of the existing cluster nodes and the CIP name which uses that IP address

4. Turn off the power to the new node.

9.1.2 Preparing the Existing Nodes

Take the following steps to prepare the existing nodes.

Procedure

1. Preparing for unexpected failures, you need to back up the entire system of all existing cluster nodes, the PRIMECLUSTER system and the configuration files of each component.

   1. Back up the entire system.

      1. Stop RMS by executing the following command on any one of the existing cluster nodes.

         ```
         node1# hvshut -a
         ```

      2. Reboot all the existing cluster nodes from a single user mode.

         ```
         node1# /usr/sbin/shutdown -g0 -i0 -y
           ....
         ok boot -s
           ....

         Type control-d to proceed with normal startup,
          (or give root password for system maintenance):
            ....
         ```

      3. Mount the file system on all the existing cluster nodes.

         ```
         node1# mountall -l
         /dev/rdsk/c0t0d0s1: is clean.
         node1#
         ```

4. Back up the entire system or property in the shared disk by executing the "ufsdump(1M)" or "dd(1M)" command.

2. Back up the PRIMECLUSTER system and the configuration files of each component.

1. Back up the configuration files of the PRIMECLUSTER system on all existing cluster nodes. See "Chapter 11 Backing Up and Restoring a PRIMECLUSTER System".

2. Back up the configuration files that are used for GLS on all existing cluster nodes.
To back up the configuration files for GLS (redundant line control), use the following command (For details on the "hanetbackup" command, see the "*PRIMECLSUTER Global Link Services Configuration and Administration Guide*"):

```
node1# /opt/FJSVhanet/usr/sbin/hanetbackup -d /var/tmp/backup
```

To back up the configuration files for the GLS multipath function:

```
node1# cd /etc/opt/FJSVmpnet
node1# tar cvf - conf | compress > /var/tmp/backup/mpnetfile.tar.Z
```

2. Define the following information in the /etc/inet/hosts file of all the existing cluster nodes.

- The IP address of the node to be added and the name of the host that is associated with the IP address

- The IP address of the remote console that is connected to the node to be added and the host name that is associated with the IP address

3. Edit the "/opt/SMAW/SMAWRrms/bin/hvenv.local" file as shown below so that RMS does not start automatically on any of the existing cluster nodes, even when a cluster node is rebooted.

```
node1# vi /opt/SMAW/SMAWRrms/bin/hvenv.local
```

```
export HV_RCSTART=0
```

4. To add a new node, all the existing cluster nodes must be turned off.

9.1.3 Connecting a Node

Join a new node with the existing cluster nodes.

Procedure

1. Connect the LAN, Cluster Interconnect, and the RCI of the new node to the existing cluster environment.

   At this time, configure the RCI address for the new node.
   (This operation is done by your Fujitsu CE.)

2. After setting the RCI address, boot up the existing cluster nodes and check that no error message is output to the console or syslog.

3. Boot the new node and confirm that the new node and its remote console are correctly configured in the network by executing the "ping(1M)" command.

4. Confirm that the RMS is stopped on any one of the existing cluster nodes, and then stop the SF by executing the following commands on each existing cluster node.

   Confirm that RMS is stopped.

5. If the GFS shared file system is used in an existing node, take the following steps to stop the GFS operation.

   1. Execute the following command for the entire GFS shared file system on any one of the existing cluster nodes, and then unmount the file system.

   ```
   node1# sfcumntgl <mount point>
   ```

   2. Execute the following command, and then stop the GFS daemon on all cluster nodes.

   ```
   node1# sfcfrmstop
   ```

```
node1# hvdisp -a
hvdisp: RMS is not running
```

Stop SF. Execute the following command on all the existing cluster nodes.

```
node1# sdtool -e
node1# sdtool -s
(SMAWsf, 30, 13) : The RCSD is not running
```

9.1.4 Configure Web-Based Admin View

This section explains how to configure Web-Based Admin View.

The nodes on which you need to configure Web-Based Admin View vary depending on the following cases;

- When the existing management server is used

  The management server must be defined on the new node.

  Configure Web-Based Admin View on the new node.

- When the new node is used as the management server

  The definition of the new management server must be defined on all the nodes.

  Configure Web-Based Admin View on all the nodes.

Procedure

1. Set up Web-Based Admin View on the node.

   See "4.3.3.1 Initial setup of the operation management server."

2. Confirm that Web-Based Admin View is running correctly.

   For confirmation, use any one of the cluster nodes as explained in "4.3.3.2 Confirming Web-Based Admin View Startup."

9.1.5 Setting Up CF and SF

Make the CF and SF configuration by using Cluster Admin. This section explains how to configure CF and SF. See "2.1.4 Example of creating a cluster" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide.*"

Procedure

1. Start the Web browser on a client and connect to the primary management server to display Web-Based Admin View.

2. Log on to Web-Based Admin View and then select the "Global Cluster Services" menu (see Figure 3).

3. Select a node name to be added on the node selection screen (see Figure 6).

4. Click the "*Load driver*" button in the left-hand side panel on the screen, and then select the "*Configure*" button to start the CF wizard (see Figures 7 and 8).

5. Select an existing cluster system name from the "Create/Join Cluster" screen and then click the "*Add local node to an existing CF Cluster*" option button. Then, choose the "*Next*" button (see Figure 10).

6. When the Cluster Interconnect confirmation screen appears, confirm that the combination of network interface cards is correct on all nodes. Then, click the "*Next*" button (see Figure 14).

7. Check the "*For RMS*" checkbox for the CIP subnet settings (note that the RMS cannot use the CIP if this is not set.)

8. When the "*Complete Configuration*" dialog box appears, close that screen and click the "*Finish*" button. This completes the CF settings.

9. Configure SF.
   For details, see "8. Shutdown Facility" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide.*"

10. Confirm that the CF and SF have been configured correctly by executing the following commands on the GUI screen or on any one of the cluster nodes.

```
node1# cftool -n
Node    Number State      Os      Cpu
node1    1     UP         Solaris Sparc
node2    2     UP         Solaris Sparc
node3    3     UP         Solaris Sparc
```

```
node1# sdtool -s
Cluster Host   Agent          SA State   Shut State   Test State   Init State
------------   -----          --------   ----------   ----------   ----------
node1          SA_pprcip.so   Idle       Unknown      TestWorked   InitWorked
node1          SA_rccu.so     Idle       Unknown      TestWorked   InitWorked
node1          SA_pprcir.so   Idle       Unknown      TestWorked   InitWorked
node2          SA_pprcip.so   Idle       Unknown      TestWorked   InitWorked
node2          SA_rccu.so     Idle       Unknown      TestWorked   InitWorked
node2          SA_pprcir.so   Idle       Unknown      TestWorked   InitWorked
node3          SA_pprcip.so   Idle       Unknown      TestWorked   InitWorked
node3          SA_rccu.so     Idle       Unknown      TestWorked   InitWorked
node3          SA_pprcir.so   Idle       Unknown      TestWorked   InitWorked
```

9.1.6 Setting Up CRM (Resource Database)

This section explains how to set up the Customer Resource Management (CRM) resource database.

Procedure

1. Reconfigure the resource database on the existing nodes.

   Confirm the following:

   - All of the existing nodes have been started.

   - CIP entry of the new node is in /etc/cip.cf of the existing nodes.

     The resource database uses /etc/cip.cf to associate CF node names with CIP names.

   Reconfigure the resource database.

   The procedure is shown below:

   1. Log into any one of existing nodes using a system administrator access privilege.

   2. Specify the "-a" and the "-g" options in the "clsetp(1M)" command.

      ```
      # /etc/opt/FJSVcluster/bin/clsetup -a node -g file
      ```

      As a result of this operation, a new node will be added to the resource database of the existing nodes.

      The configuration information on the resource database, created after the execution of the "clsetup(1M)" command, is used when the resource database of the new node is configured. Therefore, do not specify a directory that will be automatically deleted when rebooting the node with the "-g" option (for example: /tmp).

      Specify the CF node name in *node*, and a full path name of the file name of the resource database configuration information. "tar.Z" extension will be appended to the resource database configuration information.

      For example, to add a new node which has a CF node name of fuji4, and a configuration information file name of /mydir/rdb, the command to be executed will be as shown below:

      ```
      # /etc/opt/FJSVcluster/bin/clsetup -a fuji4 -g /mydir/rdb
      ```

The configuration information file of the resource database will be created as /mydir/rdb.tar.Z.

3. Confirm that the new node has been added to the resource database.

    Execute the "clgettree(1)" command, and then confirm that the new node is displayed on the output result. At this time, the state of the new node is displayed as UNKNOWN.

2. Set up the resource database of the new node.

    Confirm the following:

    - The content of /etc/cip.cf of the new node must be the same as that of the existing nodes.

        Confirm that the CIP of the new node is in /etc/cip.cf of the new node and that the content of /etc/cip.cf is the same as that of the existing nodes.

    - Confirm that the new node is connected to all of the existing nodes by CIP using the "ping(1M)" command.

        If two or more CIPs are configured in the new nodes, use the first CIP for the resource database. Then, connection will be enabled. An example using "fuji4RMS" as the new node is shown below:

    ```
    # ping fuji4RMS
    ```

    After confirming the above, set up the resource database of the new node.
    The procedure is as follows:

    1. Log into the new node using a system administrator access privilege.

    2. Copy the "resource database configuration information" file created in Step 2) in "Recreate the settings for the resource database of the existing nodes" to the new node.

    3. Specify the -s option in the clsetup(1M) command, and execute it.

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

    Specify *file* with a full path name of the resource database configuration file.

    When the resource database configuration information file "rdb.tar.Z" is copied to /mydir, the command to be executed will be as shown below:

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

    4. Confirm that the resource database of the new node is configured.

        Execute the clgettree(1) command for the new node and confirm the following:

        - The new node is displayed.

        - The state of the new node is displayed as ON.

        - The output result is the same as that of the existing nodes.

3. Register the hardware, which is connected to the new node, to the resource database.

    Log into any one of nodes using a system administrator access privilege, and execute the command shown below:

    ```
    # /etc/opt/FJSVcluster/bin/clautoconfig -r
    ```

Setting up synchronization with the resource database

If the individual nodes are restarted at different times after node expansion, the tuning parameter must be set up to be synchronized with the resource database. For details, see "4.5.1 Start up synchronization and the new node" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*."

9.1.7 Setting Up GDS

If a new node was added to the GDS configuration, it is necessary to change the class scope by executing the sdxattr command.

For information on the change procedure, see the "D.7 sdxattr - Set objects attributes" in the "*Global Disk Services Configuration and Administration Guide*".

9.1.8 Setting Up RMS

This section explains how to register a new node (SysNode) in a userApplication that has been running on an existing node.

Procedure

1. Configuration for each resource

    Take the following steps depending upon the resources in the existing userApplication:

    - Cmdline

    Create the Start, Stop, and Check scripts in the new node or copy them from the existing node. If "*Path enter*" is selected from "*Creation method*" when creating the existing Cmdline resource, specify the paths to the scripts. If "New" is selected, the scripts under /opt/FJSVwvucw/scripts/start, /opt/FJSVwvucw/scripts/stop, and /opt/FJSVwvucw/scripts/check must be stored in the same directory of the new node. You also need to add the access privilege by executing the "chmod(1)" command.

    - Gds

    Take the following steps to expand the cluster scope:

    1. Expand the cluster scope.

        See "5.4 Changes" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*."

    2. Execute the following command on any one of the cluster nodes:

        ```
        # /opt/SMAW/SMAWRrms/bin/hvgdsetup -a class
        hvgdsetup with -a option performs the following tasks on
        nodes to which the specified disk class belongs.
          1) Make GDS disk class on resource database not activated
             automatically when the node boots. If this operation has
              been done before, nothing will be performed anymore.
          2) Next make volumes of the specified disk class
             enabled manual online on a node on which an application
             is offline or faulted while the application is online
             or standby on another node.
          3) Then make volumes of the specified disk class stopped
             immediately.
             This process is executed on the nodes to which the disk
             class belongs.
        Do you want to continue with these processes ? [yes/no]
        ```

        The following message might appear after executing the "hvgdsetup" command. This does not disrupt ongoing operation.

        ```
        FJSVcluster: error: clrmd: 7516: An error occurred in the resource deactivation
        processing. (resource:resource rid:rid detail:detail)
        WARNING !!
        Failed to control 'dc_class' in the following node(s).
          node(s) node_name:
        Check the state of the nodes. If any nodes have failed, you may ignore this message.
        ```

    - Fsystem

    If the file system is UFS, add the mount point entry to /etc/vfstab on the new node.
    If the file system is GFS, take the following steps and then create a GFS local file system on the GDS volume of the new node.

    Confirm that a GDS volume is available, and that the volume is ACTIVE. If the volume is STOP, start the volume. For information on how to start a volume, see "4.2.3 Operation" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*."

1. Acquire hostid for the new node.

```
node3 # hostid
80b1a621
```

2. Add the information on the new node to the GFS local file system by executing the following command on any one of the existing cluster nodes.

```
node1 # sfxnode -a -n node3 -i 80b1a621 /dev/sfdsk/class/rdsk/volume
```

3. Register the GFS local file system on the new node by executing the following command:

```
node3 # sfxadm -A /dev/sfdsk/class/rdsk/volume
```

4. Using the vi text editor, add the entry of the GFS local file system to /etc/vfstab in the new node.
   Each line must begin with #RMS#."

```
node3# vi /etc/vfstab
#RMS#/dev/sfdsk/class/dsk/volume /dev/sfdsk/class/rdsk/volume /disk1 sfxfs - no -
```

5. Create a mount point in the new node.

```
node3 # mkdir /disk1
```

- Gls

   Take the following steps for Gls:

   1. Set up the virtual interface for the takeover IP address on the new node and register it as a cluster resource. For details, see the "*PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function*."

   2. Restart Gls by executing the following command:

```
node3# /opt/FJSVhanet/usr/sbin/resethanet -s
```

- Takeover network

   Nothing needs be done at this time. In Procedure 3, however, it is necessary to recreate the resources.

- Procedure

   Create a state transition procedure on the new node and register the procedure resource with the cluster resource manager. For more details, see "F.1 Registering a Procedure Resource".

- Process monitoring

   Add the startup command to the new node. Also, you need to add the access privilege by using the "chmod(1)" command. Then, recreate the resources in Procedure 3.

2. Deleting userApplication

   Delete the existing userApplication by using the userApplication Configuration Wizard. At this time, select "*Delete only userApplication*."

   For more details, see "8.1.1 Changing the Cluster Application Configuration."

3. Recreating the takeover network and process monitoring resources

   If the takeover network resource and the process monitoring resource are registered in the cluster system, first delete and then recreate those resources.

   See "8.8.2 Deleting a Resource," "6.6.1.5 Creating Takeover Network Resources," and "6.6.1.7 Creating Process Monitoring Resources."

4. Recreating userApplication

   Recreate the userApplication that was deleted in Procedure 2, using the same procedure as that used to create it. Note that the new node must be registered when SysNode is to be selected. For details, see "6.6.2 Creating Cluster Applications."

5. Copy /opt/SMAW/SMAWRrms/bin/hvenv.local of the existing node to /opt/SMAW/SMAWRrms/bin/ of the new node.

6. Edit /opt/SMAW/SMAWRrms/bin/hvenv.local in each node with the "vi" editor, and delete the following entry:

```
export HV_RCSTART=0
```

9.1.9 GFS Shared Settings

If the GFS shared file system is used in an existing node, set up the GFS shared file system on the new node by using the following procedure:

Procedure

1. Confirm the GFS daemon (sfcfrmd) is not running by executing the "ps" command on all cluster nodes. If GFS daemon is running, see Step 5 of "9.1.3 Connecting a Node" when stop the GFS daemon.

2. Execute sfcsetup on the new node, and then register the node information in the management partition.

3. Execute sfcfrmstart and then start up the GFS daemon on all cluster nodes.

4. Execute sfcnode on any one of the cluster nodes, and then add the node configuration information of the new node.

5. Create a mount point and set up /etc/vfstab on the new node.

6. Execute sfcmntgl on any one of the cluster nodes and then mount the GFS shared file system.

See
. . . .
For information on how to use each command, see the "*PRIMECLUSTER Global File Services Configuration and Administration Guide.*"
. . . .

9.2 Recovering the Original Cluster Configuration at Node Expansion

When the original cluster configuration must be restored due to a fault or interruption during node expansion, take the following steps for recovery:

Procedure

1. Cancel the configuration changes in GFS.

   If a new node was added to the GFS configuration at node expansion, it is necessary to cancel the GFS configuration in multi-user mode before deleting the new node.

   - When the GFS local system is used:

     Execute sfxnode, and then delete the node configuration information on the node to be deleted.

   - When the GFS shared file system is used:

     Execute sfcnode on any one of the cluster nodes and delete the node configuration information. Then, execute sfcfrmstop and stop GFS on all the nodes. In this state, execute sfcsetup on the node to be deleted, and then delete the node information from the management partition.

   ### See
   . . . .
   For information on how to use each command, see the "*PRIMECLUSTER Global File Servers Configuration and Administration Guide.*"
   . . . .

2. Change the class scope.

   If a new node was added to the GDS configuration, it is necessary to change the class scope by executing the sdxattr command. For information on the change procedure, see the "*Global Disk Services Configuration and Administration Guide.*"

3. Recover the PRIMECLUSTER system configuration files in single user mode.

   See "Chapter 11 Backing Up and Restoring a PRIMECLUSTER System."

4. Recover the GLS configuration files.

   Restore the environment definition files that are used by the GLS redundant line control by executing the following command.

   ```
   node1# /opt/FJSVhanet/usr/sbin/hanetrestore -f /var/tmp/backup/<save file name>
   ```

   Use "hanetYYYYMMDD.bk" as the *save file name*.

   For details on the hanetrestore command, see the "*PRIMECLUSTER Global Link Services Configuration and Administration Guide.*"

   Restore the environment definition files used by the GLS multipath function by using the following procedure:

   ```
   node1# cd /etc/opt/FJSVmpnet
   node1# uncompress /var/tmp/backup/mpnetfile.tar.Z
   node1# tar xvf /var/tmp/backup/mpnetfile.tar
   ```

5. Restore the original cluster configuration.

   Restore the cluster interconnect, LAN cable, and RCI cable to the original cluster configuration. At this time, restore the original RCI settings. This should be done by your Fujitsu customer engineer (CE).

9.3 Adding Hardware

This section describes how to add the following hardware to the existing configuration:

- Shared disk device

- Network interface card used for the public LAN and the administrative LAN

9.3.1 Adding a Shared Disk Device

To add a shared disk device, you need to change a cluster application related to the added shared disk device.

Figure 9.1 Procedure to add a shared disk device
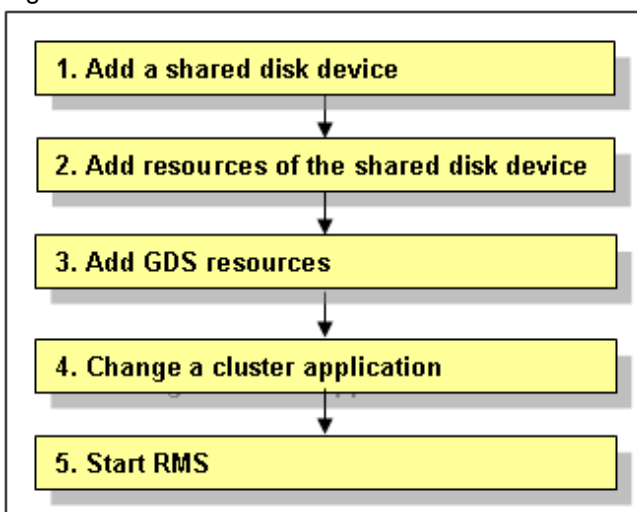
1. Add a shared disk device.

   See "10.2 Maintenance Flow" and ask your Fujitsu customer engineer (CE) to add a shared disk device.

2. Add resources of the shared disk device.

   Register resources corresponding to the added shared disk device.

   ![See icon] **See**

   ....

   To register resources, see "5.1.3.2 Automatic Configure."

   ....

3. Add Gds resources.

   To use Global Disk Services (GDS), set up GDS and create Gds resources.

   ![See icon] **See**

   ....

   For information on how to set up GDS and create Gds resources, see "6.2.2 GDS Configuration Setup" and "6.6.1.3 Creating Gds Resources."

   ....

4. Change a cluster application.

   Add the following resources related to the added shared disk device to a cluster application.

   - Fsystem resource

   - Gds resource

   ![See icon] **See**

   ....

   For more details, see "8.1.1 Changing the Cluster Application Configuration."

   ....

5. Start RMS.

   If RMS is not running, see "7.2.1.1 Starting RMS" and start RMS on all nodes.

9.3.2 Adding a Network Interface Card Used for the Public LAN and the Administrative LAN

To add a network interface card used for the public LAN and the administrative LAN, you need to change the cluster application related to the added network interface card according to the procedure below.

Figure 9.2 Procedure to add a network interface card
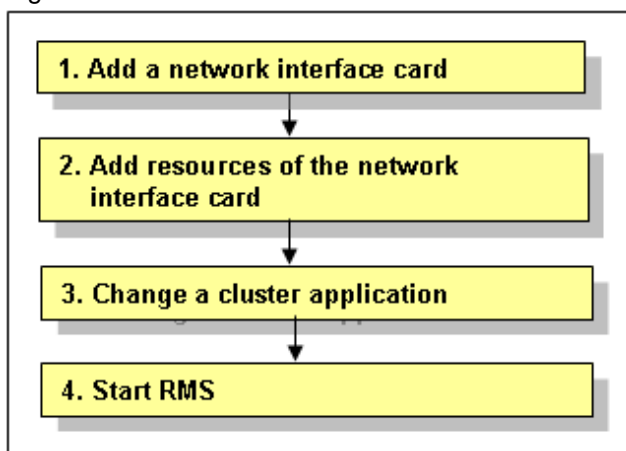
1. Add a network interface card.

   See "10.2 Maintenance Flow" and ask your Fujitsu customer engineer (CE) to add a network interface card.

2. Add resources of the network interface card.

   Register resources corresponding to the added network interface card.

   ### See

   To register resources, see "5.1.3.2 Automatic Configure."

3. Change a cluster application.

   Add the following resources related to the added network interface card to a cluster application.

   - Takeover network resource

   - Gls resource

   ### See

   For more details, see "8.1.1 Changing the Cluster Application Configuration."

4. Start RMS.

   If RMS is not running, see "7.2.1.1 Starting RMS" and start RMS on all nodes.

Part 5   Maintenance

This part explains the procedure for maintaining the PRIMECLUSTER system as well as the procedure for backing up and restoring the PRIMECLUSTER operation environment.

Chapter 10 Maintenance of the PRIMECLUSTER System

This chapter explains items and procedures related to maintenance of the PRIMECLUSTER system.

10.1 Maintenance Types

The maintenance of the PRIMECLUSTER system is divided as described below, depending on whether maintenance is performed while the job is stopped:

Stop maintenance

Maintenance that is performed while the entire cluster system is stopped.

Hot maintenance

Maintenance that is performed while the job continues to operate without state transition of the cluster application (failover or degeneration).
Hot maintenance is executed mainly after disk replacement of the disk array unit.
Also, the PRIMEPOWER PCI Hot Plug function can be used for hot maintenance.

Job hot maintenance

Maintenance that is performed while the maintenance target node is detached from the cluster by state transition of the cluster application (failover or degeneration), while the job is allowed to continue operating.

Of these, the type to be performed depends on the location and contents of the failure. Determine the maintenance that is to be performed after consulting with your Fujitsu systems engineer (SE).

10.2 Maintenance Flow

Your Fujitsu customer engineer (CE) shall perform parts replacement. The flow of maintenance is as follows:

Identifying the location of a failure

After an error is detected, identify the location of that error by using the CRM main window or RMS main window. For details on how to identify the location of the error, see "7.4.1.1 Failure Detection and Cause Identification if a Failure Occurs."

When stop maintenance is to be performed

1. The running PRIMECLUSTER system shall be stopped by administrator of the PRIMECLUSTER system.

2. Pass the operation over to your Fujitsu customer engineer (CE).

3. Your Fujitsu customer engineer (CE) shall then perform maintenance of the erroneous location (repair or replacement). Confirm that the system operates normally by running a test program, etc.

4. After the completion of maintenance by your Fujitsu customer engineer (CE), check the relevant equipment and then boot the PRIMECLUSTER system.

When job hot maintenance is to be performed

1. The administrator of the PRIMECLUSTER system shall shut down the node that contains the target equipment, so as to separate it from the operation, and then pass the operation over to your Fujitsu customer engineer (CE).

   For details on how to separate the node from the operation, see "10.2.1 Detaching Resources from Operation."

2. Your Fujitsu customer engineer (CE) shall confirm the target equipment and shall perform maintenance of the erroneous equipment (repair or replacement). Operation shall be confirmed by using a test program, etc.

3. After your Fujitsu customer engineer (CE) completes the maintenance and confirms the operation of the relevant equipment, boot the node and then execute standby restoration for the operation.

   For details on standby restoration for the operation, see "10.2.2 Executing Standby Restoration for an Operating Job."

10.2.1 Detaching Resources from Operation

Execute the following for the node that you are going to shut down.

Cluster application failover

If the relevant node is operating, you must first execute failover operation with the hvswitch(1M) command.

See
....

For details on how to determine whether the relevant node is operating, see "7.1.3.1 RMS Tree."
....

Stopping RMS

After confirming that the relevant node is in either the Offline or Standby state, stop RMS running on the relevant node by executing the hvshut(1M) command.

See
....

For details on how to stop RMS, see "4.1 Stopping RMS" in "*PRIMECLUSTER RMS Configuration and Administration.*"
....

Stopping a node

Execute the shutdown(1M) command to stop the relevant node.

10.2.2 Executing Standby Restoration for an Operating Job

Perform standby restoration for an operating job, as described below.

Procedure

1. Power on the relevant node.

2. Perform standby restoration for the relevant node (if necessary, subsequently execute failback).

See
....

For details on how to execute cluster application standby restoration, see "7.2.2.1 Starting a Cluster Application." For details on how to execute failback, see "7.2.2.3 Switching a Cluster Application."
....

10.3 Software Maintenance

This section provides notes on batch correction, emergency repair, and the application of patches to the PRIMECLUSTER system.

10.3.1 Notes on Applying Corrections to the PRIMECLUSTER System

Note the following when you apply batch corrections to the cluster system.

- Back up the system environment before you attempt to apply a correction.

- The software version to be installed on each node must be the same on all nodes in the cluster system. Also, the corrections must be the same on all the nodes constituting the system. Note, however, that this is not always true when rolling update, described below, is allowed.

- To apply a batch correction, you must stop the node temporarily. This means that the job must be stopped, albeit temporarily. You should consider a maintenance plan to ensure that the maintenance is completed within a specified period. You must also examine the time and duration of the maintenance to minimize the impact on a job.

- **Rolling update** is a method by which software is updated while the job continues to operate by executing job failover for a node in a cluster to separate the standby node from the operation in order to apply corrections to the node one by one.
  If you apply this method, the job stop time required for software update can be minimized. To perform update with this method, however, you must satisfy the prerequisites for rolling update (the items to be corrected must be correctible with rolling update).
  To apply this method, you must confirm the contents of the README file for the relevant patch and then contact your Fujitsu systems engineer (SE).

10.3.2 Overview of the Correction Application Procedure

This section provides an overview of the procedure for applying corrections such as a batch correction to a cluster system. The procedure explained here is a general procedure.

10.3.2.1 Procedure for Applying Corrections by Stopping an Entire System

This section explains the procedure for applying corrections by stopping the entire cluster system. An example of a two-node 1:1 standby configuration is used here.

Flow of operation
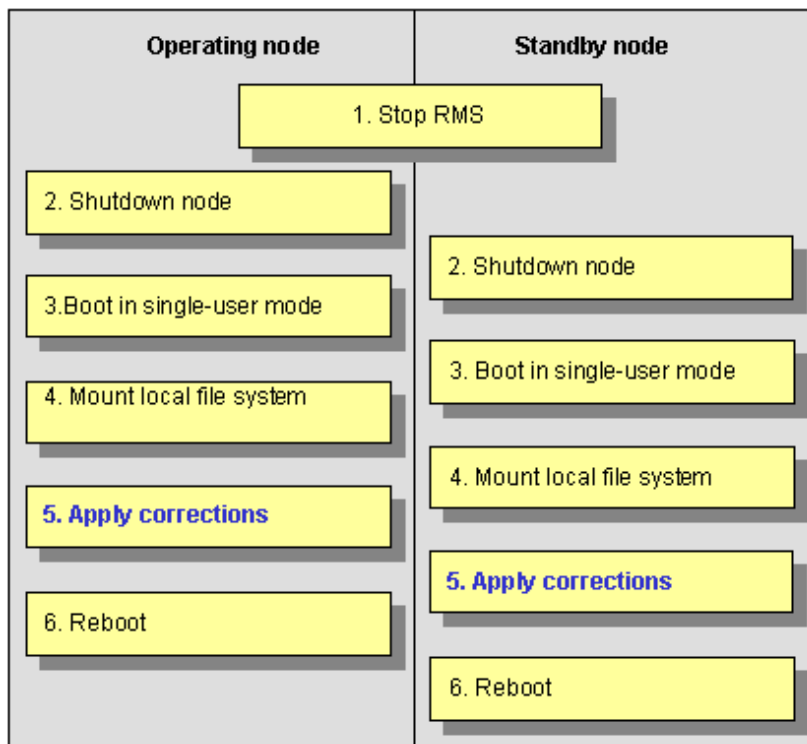
Procedure

Copy the correction to be applied to each node to the local file system in advance.

1. Stop RMS.

   Execute **hvshut -a** on either cluster node to stop the operation of RMS.

2. Shut down all nodes.

   Shut down all nodes.

3. Boot in single-user mode.

   Boot all the nodes that were shut down in single-user mode.

4. Mount the local file system.

   Execute **mountall -1** on all the nodes to mount the local file system.

5. Apply corrections.

    Apply the corrections that were copied to the local file system in advance.

6. Reboot.

    After applying the corrections, boot the nodes by using **shutdown -i6**.

![Note icon] Note
. . . .

- For details on the corrections, refer to the manuals provided with the corrections.

- For details on the standby restoration of cluster applications, see "7.2.2.1 Starting a Cluster Application." For details on failback, see "7.2.2.3 Switching a Cluster Application."
. . . .

10.3.2.2 Procedure for Applying Correction by Rolling Update

This section explains the procedure for applying corrections by rolling update. An example of two-node 1:1 standby configuration is used for this explanation.

Flow of operation
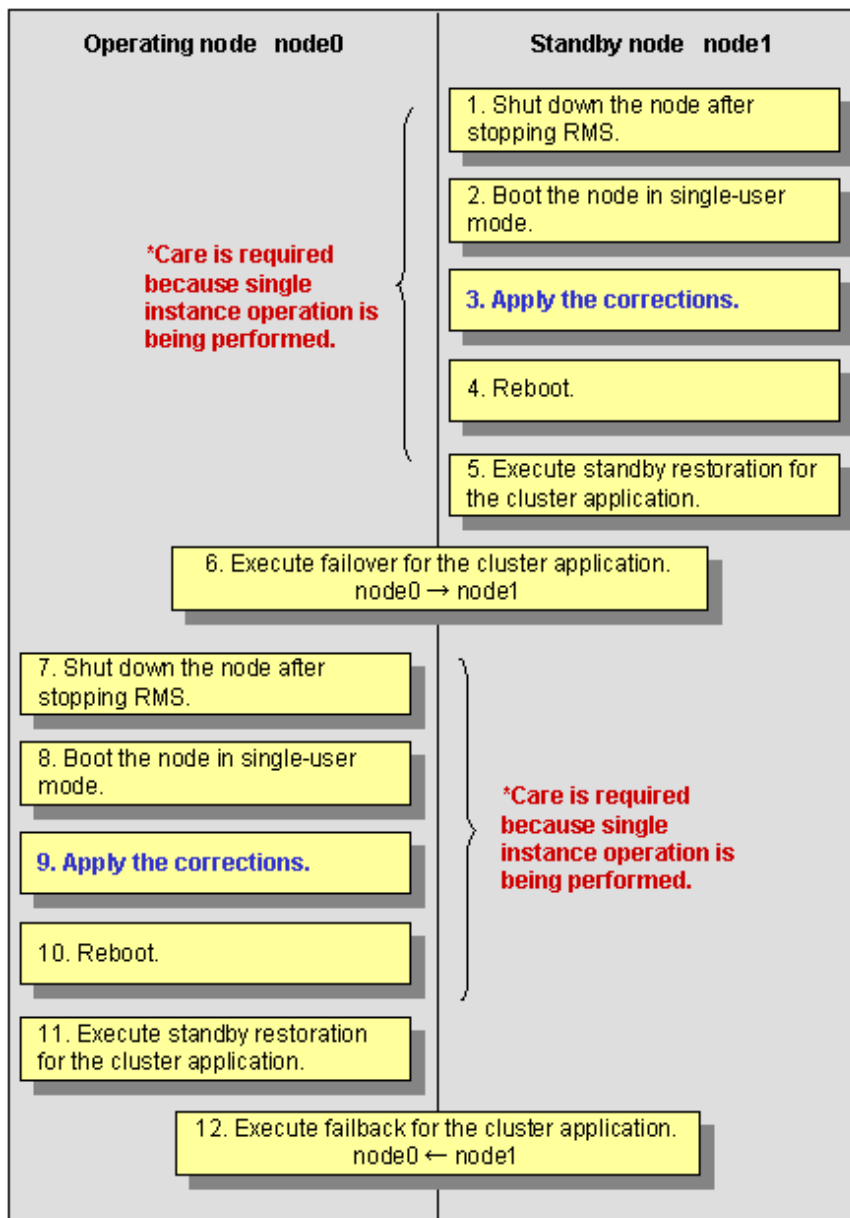
Procedure

1. Shut down the standby node (node1).

   To apply corrections to the standby node (node1), shut down the node after stopping RMS.

   Note that, as a result of this shutdown, a cutoff state transition occurs and dual instance operation is disabled until standby restoration is performed.

2. Boot the standby node (node1) in single-user mode.

   Boot in single-user mode.

3. Apply corrections.

   Apply the necessary corrections.

4. Reboot the standby node (node1).

   Reboot the node.

5. Execute standby restoration for the standby node (node1).

   Since the shutdown node (node1) to which corrections have been applied is cut off from the cluster system, execute standby restoration for the node.

6. Execute failover of the cluster application.

   To apply corrections to the operating node (node0), switch the cluster application with the **hvswitch(1M)** command.

7. Shut down the operating node (node0).

   After the completion of failover, stop RMS, and then shut down the node.

8. Boot the operating node (node0) in single-user mode.

   Boot the node in single-user mode.

9. Apply corrections.

   Apply the necessary corrections.

10. Reboot the operating node (node0).

    Reboot the node to apply the corrections to the system.

11. Execute standby restoration for the operating node.

    Since the shut-down node (node0) to which corrections were applied is cut off from the cluster system, execute standby restoration for the node.

12. Execute failback.

    Restore the state of the standby layout defined at installation by executing failback operation, as required.

Note

- For details on the corrections, refer to the manuals provided with the corrections.

- For details on standby restoration of cluster applications, see "7.2.2.1 Starting a Cluster Application." For details on failback, see "7.2.2.3 Switching a Cluster Application."

Chapter 11 Backing Up and Restoring a PRIMECLUSTER System

The PRIMECLUSTER system supports the collective backup and restoration of a variety of environmental files that constitute PRIMECLUSTER, for the purpose of recovering from problems caused by operating errors during upgrading of, or configuration changes to, the system.

The following properties of the PRIMECLUSTER system can be backed up and restored with this function.

- CF configuration

- SF (Shutdown Facility) configuration

- RMS configuration

- GUI (including Web-Based Admin View) configuration

- GLS (redundant line control function) configuration

Note

- The hardware configuration must not change between back up and the restoring .

- Among the RMS configuration files, special commands and utilities that can be invoked by the user from the configuration are backed up.

- The PRIMECLUSTER backup/restore function cannot backup or restore a GLS (multipath function), GFS, or GDS configuration. For backing up and restoring GLS (multipath function), GFS, and GDS configurations, please refer to the corresponding manuals.

- For details on the backup and restoration of PRIMECLUSTER-related products, refer to the corresponding manual.

Prerequisites for backup and restoration

The following conditions must be satisfied to use the backup and restoration functions of PRIMECLUSTER:

a. PRIMECLUSTER 4.1A10 or later.
   The backup/restore function including GLS (redundant line control function) will be available on PRIMECLUSTER 4.1A20 or later.

b. The version to be used for restoration must be of the same PRIMECLUSTER version as that was used for backup, or later.

c. Execute the backup procedure on all the nodes in the PRIMECLUSTER operation environment. It is preferable to execute this procedure simultaneously. You can back up in either multi-user or single user mode.

Note

If backup is not executed for all nodes simultaneously, the machines may contain backup data with a different configuration. If the configuration of the backup data differs between machines, the consistency of the cluster configuration cannot be maintained at restoration.

d. Execute the restoration procedure in single user mode.

Note

- If the OS is newly installed, and the PRIMECLUSTER environment is restored, the normal operations of a cluster system cannot be guaranteed.

- If restoration is not done on all nodes but only on some nodes, the consistency of the cluster configuration might not be maintained between nodes, and failures might occur.

e. When restoration is to be performed, the hardware configuration must be the same as that used for backup.

f. With the restoration function, Solaris standard system files cannot be restored. The backup function also cannot be used to back up all the system files in a PRIMECLUSTER system. You need to manually back up or restore the system files, if necessary.
For details on the method of restoring system files that have been backed up with the backup function, refer to Procedure 3) in the explanation of the restoration procedure.

Information

You should back up a cluster configuration each time it is modified.

11.1 Backing Up the PRIMECLUSTER Operation Environment

Execute the following command to back up the PRIMECLUSTER operation environment:

```
/opt/SMAW/bin/cfbackup [ -f ] [ n ]
```

See

For details on the command parameters, refer to the description of the "cfbackup(1M)" command.

Example

Back up the PRIMECLUSTER environment.

```
# cfbackup
08/07/02 11:17:55 cfbackup 1 started
...
08/07/02 11:21:23 cfbackup 1 ended
#
```

Example

Confirm the backup files.

```
# ls -l /var/spool/SMAW/SMAWccbr
-rw-rw-rw-   1 root     other       2860 Aug  7 11:21 fuji2_ccbr1.tar.Z
```

11.2 Restoring the PRIMECLUSTER Operation Environment

Take the following steps to restore the PRIMECLUSTER operation environment:

Procedure

1. Boot the machine to be restored in single-user mode.

Note

After booting the machine in single-user mode, mount the file system by executing the "mountall -l" command.

Example: Reboot in single-user mode.

```
# shutdown -g0 -i0 -y
...
```

```
ok boot -s
...
INIT: SINGLE USER MODE

Enter control-d to proceed with normal startup,
(or enter the root password for system maintenance):
...
# mountall -l
/dev/rdsk/c0t0d0s4: is clean.
/dev/rdsk/c0t0d0s3: is clean.
#
```

2. To restore the PRIMECLUSTER operation environment, execute the following command.

```
/opt/SMAW/bin/cfrestore [ -f ] [ -p ] [ -y ] [ -M ] [ n ]
```

See

For details on the command parameters, refer to the explanation of the cfrestore(1M) command.

Example: Restore the PRIMECLUSTER operation environment.

```
# cfrestore
08/07/02 11:17:55 cfbackup 1 started
...
08/07/02 11:21:23 cfbackup 1 ended
#
```

3. Restore a single configuration file.

Note

The restoration function does not restore backed-up files. The system files mentioned here in are provided by the operation system (environment). Restore the backed-up files manually if necessary.

Example: Pick up the "/etc/vfstab" file from the backup data "fuji2_ccbr1.tar.Z".

1. Confirm the backup data "fuji2_ccbr1.tar.Z" and then check for the backed-up system file.

```
# ls /var/spool/SMAW/SMAWccbr/*.tar.Z
fuji2_ccbr1.tar.Z
# zcat /var/spool/SMAW/SMAWccbr/fuji2_ccbr1.tar.Z | tar tvf - ./OS
drwxrwxrwx   0/1          0 Nov 20 13:58 2002 ./OS/
drwxrwxrwx   0/1          0 Nov 20 13:57 2002 ./OS/etc/
...
-rw-r--r--   0/1        909 Nov 20 13:57 2002 ./OS/etc/vfstab
...
#
```

2. Put the "/etc/vfstab" file in the "/tmp" directory.

```
# cd /tmp
# zcat /var/spool/SMAW/SMAWccbr/fuji2_ccbr1.tar.Z | tar xvf - ./OS
x ./OS, 0 bytes, 0 tape blocks
x ./OS/etc, 0 bytes, 0 tape blocks
x ./OS/etc/hosts, 401 bytes, 1 tape blocks
...
x ./OS/etc/vfstab, 909 bytes, 2 tape blocks
...
#
```

3. Restore part or all of the file contents by executing the "cp(1)" or "vi(1)" command.

> **Note**
>
> If an invalid correction is made to the system file, login may be disabled and the system may fail to boot.

4. If the AutoStartUp function of the RMS has been validated, invalidate it.

   For details, see the "10 Appendix - Environment variables" in the *"PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide."*

5. Reboot the machine.

   Example: The machine is rebooted.

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

6. Activate the RMS configuration.

   For details, see the "4.8 Activating the configuration" in the *"PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide."*

7. Enable the AutoStartup function if it is disenabled.

   For details, see the "10 Appendix - Environment variables" in the *"PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide."*

8. Start up the RMS.

   For details, see the "5.4.1 Starting RMS" in the *"PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide."*

11.3 Environment Setup File

The two environment configuration files for backup and restoration are "ccbr.conf" and "ccbr.gen". The storage destination of the data to be backed up or restored can be changed by editing these files using the text editor "vi(1)".

Environmental variables and environment configuration files used for backup and restoration are shown below:

- /opt/SMAW/ccbr/ccbr.conf

   This file defines variables for PRIMECLUSTER backup and restoration. The variables include the following:

   CCBRHOME :

      This variable indicates a directory. The backup data is stored into this directory.
   The following is a default value:

      /var/spool/SMAW/SMAWccbr

- /opt/SMAW/ccbr/ccbr.gen

   This file defines the number of records for PRIMECLUSTER backup. This number increases one by one whenever the backup process is done.

> **Information**
>
> The number of generations can be also changed with the arguments of cfbackup (1M) / cfrestore(1M). For details, refer to the explanation of each command.

> **Example**
>
> The backup configuration is as follows:

<Setup contents>

- Host name: fuji2

- CCBRHOME variable: /var/spool/SMAW/SMAWccbr

- The value of the "ccbr.gen" file: 1

In the case of above, the following file will be created as the backup data:

```
/var/spool/SMAW/SMAWccbr/fuji2_ccbr1.tar.Z
```

Configuration procedure example

Configure the backup data storage destination to "/var/spool/SMAW/SMAWccbr" and the number of generations to "1".

1. Edit the "/opt/SMAW/ccbr/ccbr.conf" file using the "vi(1)" text editor, and then change the value of the CCBRHOME file to "/var/spool/SMAW/SMAWccbr".
   If the value has already been entered, go to Procedure 2.

   Confirm the contents.

   ```
   # cat /opt/SMAW/ccbr/ccbr.conf
   ...
   CCBRHOME=/var/spool/SMAW/SMAWccbr
   export CCBRHOME
   ...
   #
   ```

2. Edit the "/opt/SMAW/ccbr/ccbr.gen" file using the "vi(1)" text editor, and then change the value to "1". If the value has already been configured, go to Procedure 3.

   Confirm the contents.

   ```
   # cat /opt/SMAW/ccbr/ccbr.gen
   1
   #
   ```

3. Back up the PRIMECLUSTER configuration.

   ```
   # cfbackup
   08/07/02 11:17:55 cfbackup 1 started
   ...
   08/07/02 11:21:23 cfbackup 1 ended
   #
   ```

   Confirm the contents of the backup file.

   ```
   # ls /var/spool/SMAW/SMAWccbr
   fuji2_ccbr1
   ```

🖼️ Example
. . . .

The restoration configuration is as follows:

<Setup contents>

- Host name: fuji2

- CCBRHOME variable: /var/spool/SMAW/SMAWccbr

- Value of the "ccbr.gen" file: 2

In the above case, the following file will be used as the data to be restored:

```
/var/spool/SMAW/SMAWccbr/fuji2_ccbr1
```

Configuration procedure example

Configure the destination of the restoration data to "/var/spool/SMAW/SMAWccbr" and the number of generations to "1".

1. Edit the "/opt/SMAW/ccbr/ccbr.conf" file using the "vi(1)" text editor, and then make the value of the CCBRHOME variable "/var/spool/SMAW/SMAWccbr".
   If the value has already been configured, proceed to Procedure 2.

   Confirm the contents.

   ```
   # cat /opt/SMAW/ccbr/ccbr.conf
   ...
   CCBRHOME=/var/spool/SMAW/SMAWccbr
   export CCBRHOME
   ...
   #
   ```

2. Edit the "/opt/SMAW/ccbr/ccbr.gen" file using the "vi(1)" text editor, and then change the value to "2" (specify a value that is 1 greater than the target number of generations).
   If the value has already been configured, proceed to Procedure 3.

   Confirm the contents.

   ```
   # cat /opt/SMAW/ccbr/ccbr.gen
   2
   #
   ```

3. Restore the PRIMECLUSTER environment.

   ```
   # cfrestore
   08/07/02 11:17:55 cfrestore 1 started
   ...
   08/07/02 11:21:23 cfrestore 1 ended
   #
   ```

Part 6 PRIMECLUSTER Products

This section explains the PRIMECLUSTER products.

- PRIMECLUSTER Product List

- PRIMECLUSTER Wizard for NetWorker

- PRIMECLUSTER Wizard for Oracle

- Interstage Application Server Enterprise Edition

- Softek AdvancedCopy Manager

- Symfoware Server

- Systemwalker Centric Manager

- Systemwalker Operation Manager

- Systemwalker WebMGR

Chapter 12 PRIMECLUSTER Product List

The following lists the version levels of PRIMECLUSTER products and the range of support:

Table 12.1 PRIMECLUSTER product list

Product name	Operation mode	Standby operation				Scalable operation	Notes
	System configuration	Standby		Mutual standby	Cascade		
		1:1	N:1				
	IP address takeover	Yes		Yes	Yes		
	Node name takeover	Yes	No	No	No	No		
Interstage Application Server Enterprise Edition 9.0		N	Y	N	Y	N	N	
Symfoware Server 8.0.0		Y	Y	Y	N	Y	N	
Symfoware Server Hot Standby Option 8.0.0		Y	Y	Y	N	Y	N	
Symfoware Server Parallel Cluster Option 8.0.0		N	Y	Y	N	Y	Y	
Softek AdvancedCopy Manager 13.1		Y	Y	N		N	N	
Systemwalker CentricMGR EE/GEE 13.2		N	Y	N	N	N	N	
Systemwalker OperationMGR EE/GEE 13.2		Y	Y	Y	Y	N	N	
Systemwalker WebMGR-M EE 10.1		N	Y	N	N	N	N	
PRIMECLUSTER Wizard for Oracle 4.2		Y	Y	Y	Y	Y	Y	For scalable operation, Oracle Real Application Clusters for PRIMECLUSTER is required.
PRIMECLUSTER Wizard for NetWorker 4.2		Y	Y	Y	Y	Y	N	
Oracle Real Application Clusters for PRIMECLUSTER 10g		N	N	N	N	N	Y	PRIMECLUSTER Wizard for Oracle is required.

12.1 PRIMECLUSTER Operation Modes

This section explains the cluster operation mode of PRIMECLUSTER, based on SynfinityCluster.
Note that the "cluster service" of SynfinityCluster is relevant to the "cluster application" of PRIMECLUSTER.

Also, "network takeover" of SynfinityCluster relates to the "public LAN takeover" of PRIMECLUSTER.

For the terminologies, see Appendix "E.1 Terminology".

Cluster operation mode

Cluster operation mode	Description
1:1 standby	A single standby cluster application runs on a system that consists of two nodes. One is an operating node, and the other is a standby node.

Cluster operation mode	Description
N:1 standby	"N" standby cluster applications run on a system consisting of "N+1" nodes. One application runs on each operating node. Two or more applications stand by on one node.
Mutual standby	One cluster application runs, and the other cluster application stands by on each node. Generally, a cluster system consists of 2 nodes. This is referred to as two-node mutual standby.
Cascade	One operating node has two or more standby nodes. Even when one node stops, the cluster application assures its redundancy by using other node. This operation mode is effective for maintaining the availability of a cluster application during maintenance.
Priority transfer	Topology applied from N:1 standby Effective in maintaining availability during maintenance because the restored node becomes a standby node through the transfer of the standby node, even after the standby node stops and the cluster application can adopt a redundant configuration.
Scalable	Topology that is best suited for parallel job execution because one or more cluster applications operate together.
High-availability scalable	Topology in which standby operation is performed for each cluster application that constitutes a scalable configuration.

Cluster operation mode

Cluster operation mode	Description
Standby operation mode	General names for the four operation modes are 1:1 standby, N:1 standby, mutual standby, and cascade.
Scalable operation mode	One scalable cluster application consists of two or more operating nodes. If one node stops in the event of a failure, ongoing operations are continued on other operating nodes. Each operating node can be configured to be in the standby operation mode.

Public LAN takeover function

Public LAN takeover function	Description
IP address takeover	If a failover occurs during standby operation, one or more IP addresses are taken over to other node. In such a case, two or more network interface cards can be used.
MAC takeover	If a failover occurs during standby operation, the MAC address of the network interface card is taken over to other node. This function enables ongoing operations on a client that uses MAC addresses.
Node name takeover	If a failover occurs during standby operation, the node name is taken over to other node. This function enables access to the logical IP address corresponding to the node name.

Combination of the takeover network

Status of support		Type of takeover		
SynfinityCluster	PRIMECLUSTER	IP	MAC	Node name
Y	Y	Y	N	N
Y	Y	Y	N	Y
Y	Y	Y	Y	Y
Y	N	N	Y	N

Status of support		Type of takeover		
SynfinityCluster	PRIMECLUSTER	IP	MAC	Node name
Y	Y	Y	Y	N

* MAC takeover **must be used at the same time** as IP address takeover with PRIMECLUSTER.

![Note icon] **Note**

. . . .

Physical IP address takeover and MAC address takeover deactivate a physical IP address at failover. To avoid disrupting ongoing operations of Web-Based Admin View or Shutdown Facility, set up physical IP address takeover or MAC address takeover for a network interface that is not being used by that Web-Based Admin View or Shutdown Facility.

. . . .

Chapter 13 PRIMECLUSTER Wizard for NetWorker

13.1 Functional Overview

You can back and restore the data on the shared disk managed by PRIMECLUSTER using NetWorker regardless of the operation mode. The standby configuration is enabled for NetWorker operating on a PRIMECLUSTER system.
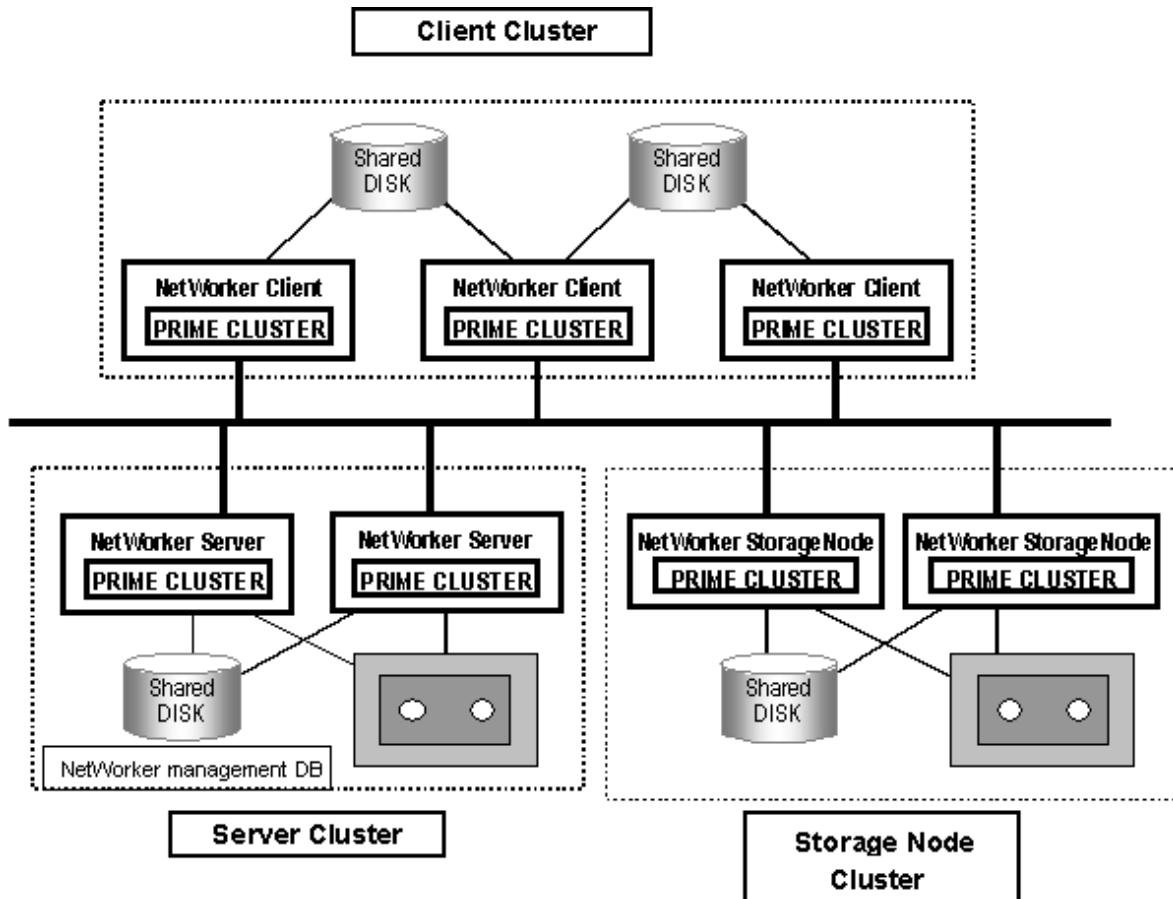
The configuration in which a NetWorker client operates on a cluster system is called a client cluster.

The configuration in which a NetWorker storage node operates on a cluster system is called a storage node cluster.

The configuration in which a NetWorker server operates on a cluster system is called a server cluster.

13.1.1 Client Cluster

Takeover of consistent settings for backup operation in the cluster environment (that does not require the NetWorker definition changes in the event of a failover) is enabled for the data on a shared disk or local disk of a PRIMECLUSTER system.

Also, standby operation of the NetWorker client is supported.
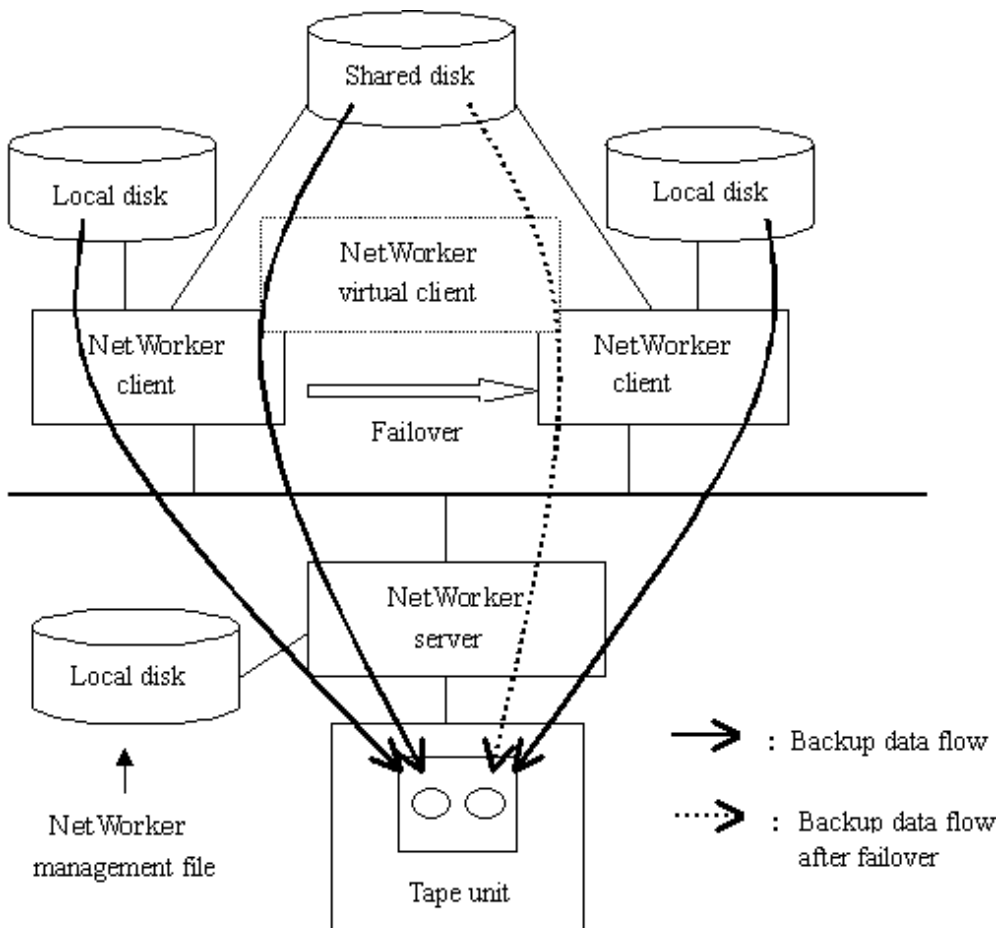
Virtual client backup

Data on the shared disk, which is managed by PRIMECLUSTER, can be backed up with a logical node name (virtual server name).

The logical node name is a virtual name associated with a virtual IP address that is defined by the PRIMECLUSTER takeover IP address. This name is controlled by the PRIMECLUSTER Wizard for NetWorker control to support virtual NetWorker client/server/storage nodes.

By using the logical node name, NetWorker can handle the cluster system easily in the same way as a single system. Thus, data on the shared disk can be backed up from the standby node (new operating node) without changing the definition of NetWorker and the backup data can be recovered from either of the nodes.

The logical node names can be specified for each user application in the cluster. Even in a cluster system with multiple user applications that are shown in an operation mode such as N:1 standby or mutual standby, each user application is handled as a single system. This enables the data on each shared disk to be backed up and recovered with the logical node name.

Local backup of the cluster configuration node

The data on the local disk of each cluster node can be backed up with the physical node name of the cluster node. Since the NetWorker client operates on all nodes that configure the cluster, the data on the local disk of the standby node can also be backed up. The settings for backup operation are taken over to the standby node without the NetWorker definition changes in the event of a failover.

Standby operation of NetWorker client

Standby operation of the NetWorker client can be provided on PRIMECLUSTER by:

- Starting up the NetWorker client for data backup on the local and shared disks on an operating node.

- Starting up the NetWorker client for data backup on the local disk on a standby node.

Also, the process monitoring facility (detector) that is exclusive to the NetWorker client monitors the NetWorker client process, and if a failure occurs, detects and notifies PRIMECLUSTER of the failure. This feature enhances the availability of the backup system.

13.1.2 Storage Node Cluster

By using NetWorker storage nodes and SAN storage nodes in a cluster system, a tape unit can be shared, and the process monitoring facility (detector) that is exclusive to NetWorker can monitor the process of NetWorker storage nodes and SAN storage nodes. This enhances the availability of the backup system.
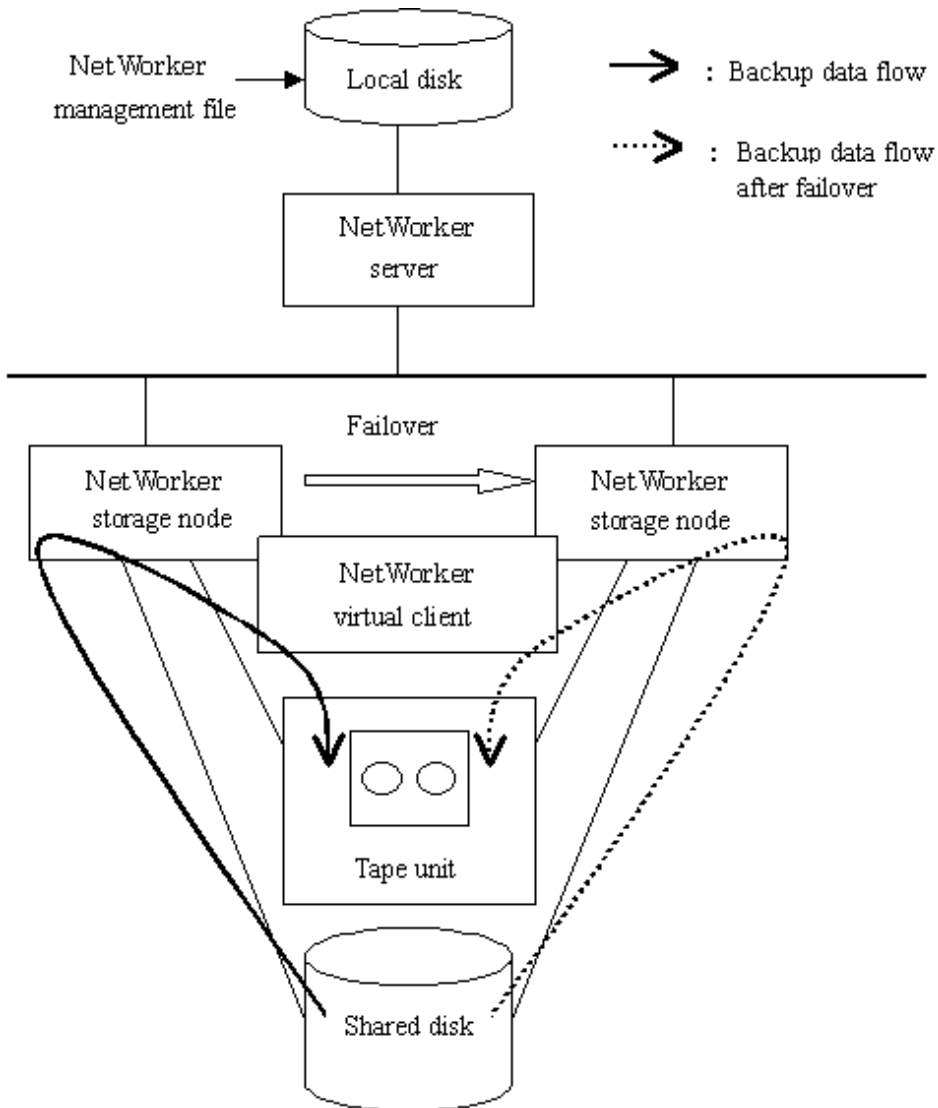
With the standby operation of the storage node and SAN storage node, the availability of LAN-free backup that is compatible with the SAN environment is enhanced.

By operating the NetWorker storage node on PRIMECLUSTER, the PRIMECLUSTER failover function enables the standby node to take over the functions provided by the NetWorker storage node without the NetWorker definition changes.

The client cluster configuration also allows backup operation using a logical node. The process monitoring facility (detector) that is exclusive to NetWorker can monitor the NetWorker client. This enhances the availability of the backup system.

Also, a tape unit can be shared between cluster nodes. Note that a tape unit can be shared between the NetWorker server and the NetWorker storage node using an optional NetWorker product.

Note

The NetWorker storage node serves as a NetWorker client.

13.1.3 Server Cluster

Standby operation of the NetWorker server is provided on PRIMECLUSTER. The process monitoring facility (detector), which is exclusive to NetWorker provided by PRIMECLUSTER Wizard for NetWorker, automatically detects and recovers from a failure. This enhances the availability of the backup system.
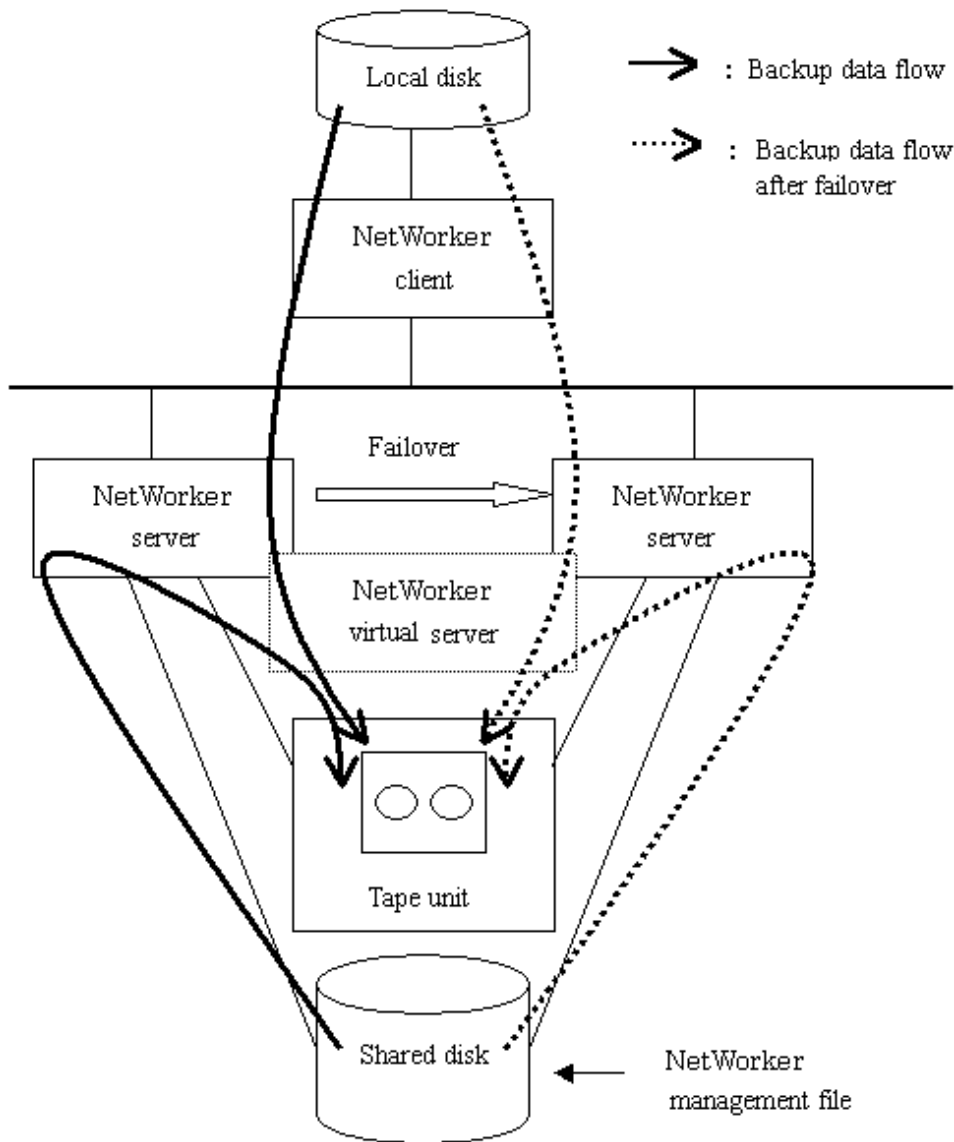
By running the NetWorker server on a cluster system, the tape unit and the NetWorker management database (index and NetWorker settings) can be shared. The backup operation is taken over to the standby node in the event of a failover.

Backup using the logical node is also enabled.

Note

The NetWorker storage node serves as a NetWorker client.

13.1.3.1 Backup takeover

The backup takeover is enabled by running RMS (Reliant Monitor Services) configuration script that is exclusive to NetWorker provided by PRIMECLUSTER Wizard for NetWorker and using the process monitoring facility that monitors the NetWorker server on PRIMECLUSTER.

When a failover occurs during backup operation, backup is automatically switched to the standby node. The NetWorker server automatically starts on the standby node (Next Online node), and backup is restarted without the NetWorker backup definitions.

Also, the backup data is recovered on the next online node using the data that is backed up with the previously online node.

13.1.3.2 Automatic detection and NetWorker failure recovery

NetWorker errors are automatically detected and corrected by running the RMS (Reliant Monitor Services) configuration script that is exclusive to NetWorker provided by PRIMECLUSTER Wizard for NetWorker and the process monitoring facility that monitors the NetWorker server on PRIMECLUSTER.

This automatic detection and recovery function reduces downtime of backup application.

RMS configuration script

RMS (Reliant Monitor Services) starts the RMS configuration script to start or stop NetWorker. Also, when an error occurs during NetWorker operation, the RMS configuration script is restarted with the AutoRecover function to recover the NetWorker operation. The RMS configuration script checks the NetWorker management database index during NetWorker startup. If an error occurs in the index, then it is automatically corrected.

NetWorker process monitoring facility

The PRIMECLUSTER NetWorker process monitoring facility monitors the process of NetWorker. When the process monitoring facility detects a NetWorker error, it notifies RMS of the error. First, RMS attempts to restart NetWorker once, but if it doesn't work, it brings the Online node Faulted. The process monitoring facility continues monitoring NetWorker even after notifying RMS of the error. When NetWorker is restarted with the RMS configuration script, the process monitoring facility starts monitoring NetWorker again.

Details on each error and its recovery are given below:

Failure recovery

The process monitoring facility monitors NetWorker, if an error is detected, it notifies RMS of the error. If the AutoRecover attribute is enabled, the RMS starts up NetWorker startup script after receiving the error notification.

📑 See
. . . .
To determine whether the AutoRecover function should be enabled, refer to the PRIMECLUSTER manuals.
. . . .

Recovery of the NetWorker management database index

The RMS configuration script checks the index by executing the NetWorker commands. If it determines that it is necessary to correct the index of the NetWorker management database, it corrects the index of the target client by executing the NetWorker commands. The RMS configuration script determines if the index needs to be recovered using the NetWorker command.

The index error might be caused by the following cases:

- A failure occurs in NetWorker during backup

- A running node is powered off during backup, and a failover occurs

You can select Yes or No in the recovery mode parameter provided by PRIMECLUSTER Wizard for NetWorker for index correction.

📑 See
. . . .
For details on how to set the recovery mode, see the "*PRIMECLUSTER Wizard for NetWorker Configuration and Administration Guide*."

For details on how to correct index errors manually, see the *"PRIMECLUSTER Wizard for NetWorker Configuration and Administration Guide."*
. . . .

13.1.4 Suspending NetWorker Monitoring

You can stop NetWorker monitoring by executing the "hvnwdisable" command of PRIMECLUSTER Wizard for NetWorker. By doing this, you can shut down NetWorker without disturbing the ongoing operations of other user applications and do maintenance work of NetWorker regardless of cluster operation. At this time, the state of user applications of NetWorker is Online. When restarting the NetWorker monitoring feature, you need to execute the "hvnwenable" command of PRIMECLUSTER Wizard for NetWorker. For details, see *"PRIMECLUSTER Wizard for NetWorker Installation/Administration Guide."*

13.1.5 Standby Operation

In the server cluster topology, the online node runs as the NetWorker server, and the standby node runs as a NetWorker client. The NetWorker process is monitored on the standby node. If NetWorker on the standby node is abnormally terminated, RMS attempts to restart NetWorker once. If it does not start normally, the NetWorker user application resource enters the faulted state. When a failure occurs, and the monitoring process is killed, the standby node enters Faulted. The Faulted node can restart the standby operation by clearing the faulted state if the database on the standby node is not damaged.

Also, if the NetWoker server is mistakenly booted on the standby node, the NetWorker process monitoring facility considers it as an error and notifies RMS of the error. Then, RMS brings that node Faulted.

13.2 Operation Environment

13.2.1 System Configuration

13.2.1.1 Client cluster
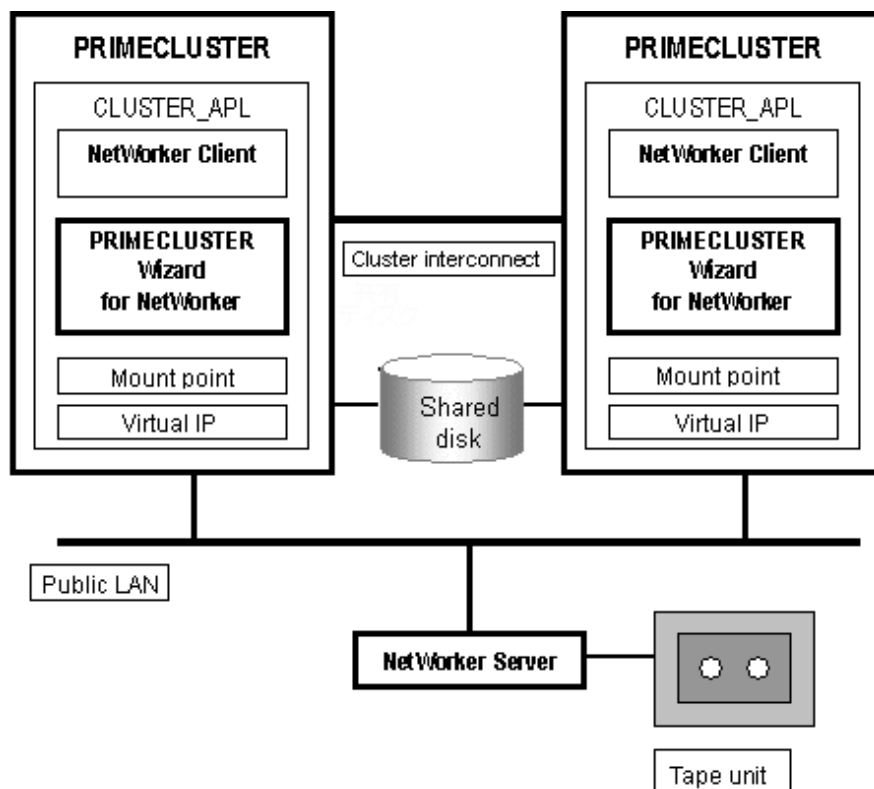
13.2.1.2
Storage node cluster
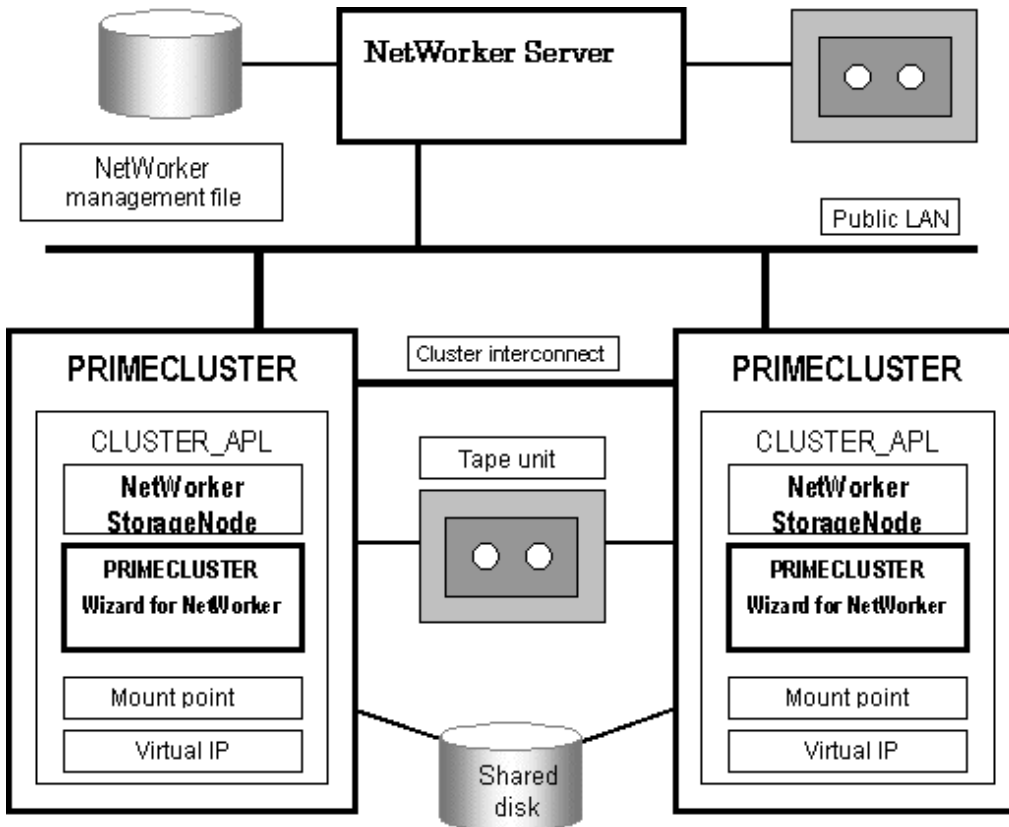
13.2.1.3
Server cluster
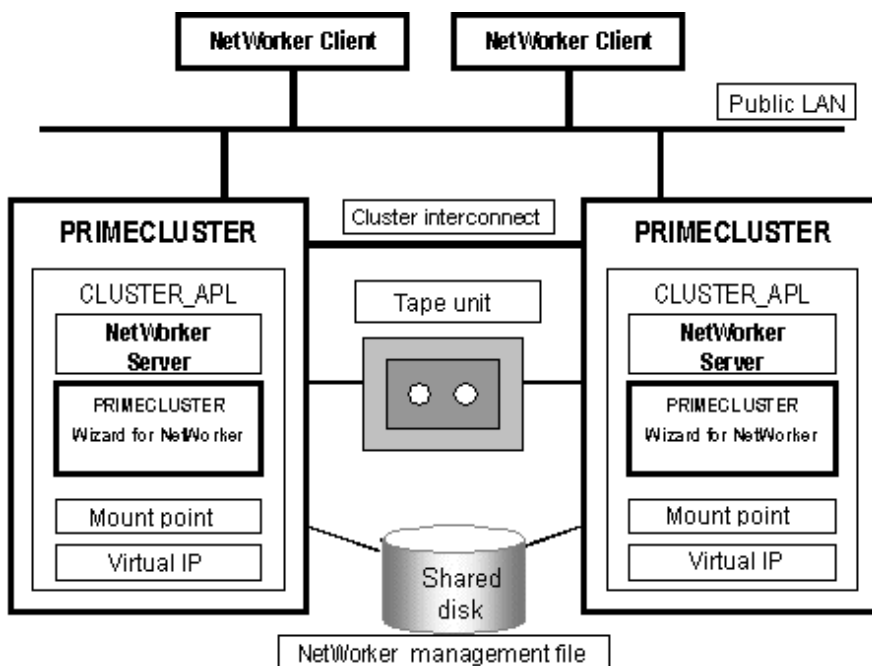
13.2.2 Supported Modes

For details, see the "Table 12.1 PRIMECLUSTER product list".

13.2.3 Overview of Cluster System Support

The PRIMECLUSTER Wizard for NetWorker provides the environment setup wizard that is used to build a backup system on the cluster system by using NetWorker, startup/stop scripts, and state monitoring detectors, and supports standby operation using NetWorker.

13.3 Notes

For details on the environment setup with the PRIMECLUSTER Wizard for NetWorker, see the "*PRIMECLUSTER Wizard for NetWorker Configuration and Administration Guide,*" which is provided with the PRIMECLUSTER Wizard for NetWorker.

Chapter 14 PRIMECLUSTER Wizard for Oracle

14.1 Functional Overview

PRIMECLUSTER Wizard for Oracle is a software product that facilitates cluster operation management for Oracle operating on a PRIMECLUSTER system.

This is required for scalable operation and standby operation for a cold standby Oracle instance. The scalable operation requires Oracle9i Real Application Clusters (Oracle9i RAC). The standby operation does not require Oracle9i RAC, but requires Oracle cold standby instead.

Startup and stop control

In line with the state transition of the cluster system, scripts automatically start or stop an Oracle instance or an Oracle listener. These scripts first check and then start up the Oracle instance. If damage has occurred, the scripts automatically recover the Oracle instance. For example, if the Oracle instance fails during online backup, the scripts automatically execute "end backup" to start up the Oracle instance. (When the AutoRecover is set to "Yes".) When stopping the Oracle instance, the scripts stop it in the immediate mode (default). If this does not occur properly, the scripts stop the instance in the abort mode. This means operational node switchover can be performed at high-speed.

Monitoring

A detector monitors an Oracle instance and an Oracle listener. The detector connects itself to Oracle using system user. Then, this enables the detector periodically to create, update and delete a dummy table, which should be generated in the system user's default tablespace, by executing an SQL command. Consequently, a logical failure can be detected as well as the process of the Oracle instance being active or inactive. When the Oracle instance hangs, and if the SQL command is not returned within a specified time, the detector times out and notifies a user of the failure. The two types of monitoring of the Oracle listener are process detection and the "tnsping" command.

Environment setup

The environment setup tool provides an "ORACLE" Wizard that generates userApplications in the RMS configuration by executing the "hvw" command of PRIMECLUSTER RMS Wizard.

14.2 Operation Environment

14.2.1 System Configuration

System configuration figures are shown below:

Scalable operation (Oracle9i RAC)

In scalable operation with Oracle9i RAC, Oracle is operated on all nodes. A client can use a database that is connected to either node.
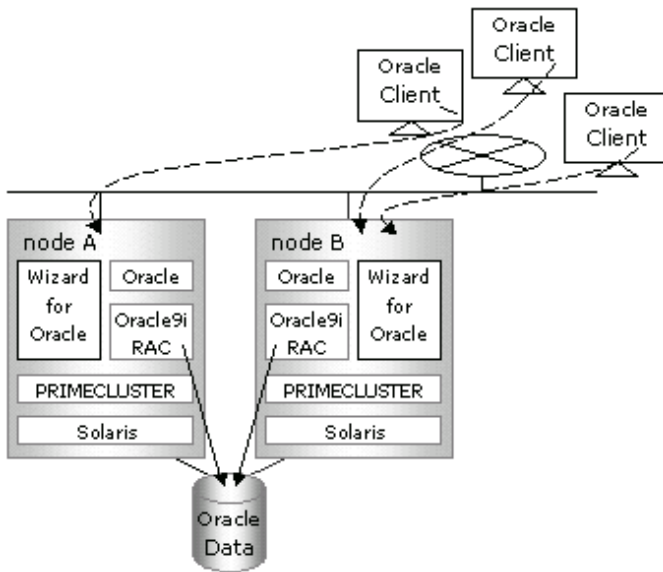
Standby operation (cold standby)

In standby operation, a cluster system consists of an operational node and standby nodes. On the operational node, Oracle applications, a logical IP address and shared disks are active. On the standby nodes, these resources are inactive. In the event of a failure, production operation is switched to one of the standby nodes, and the resources in this standby node are activated. The client can connect to the running node using the logical IP address without configuration changes.
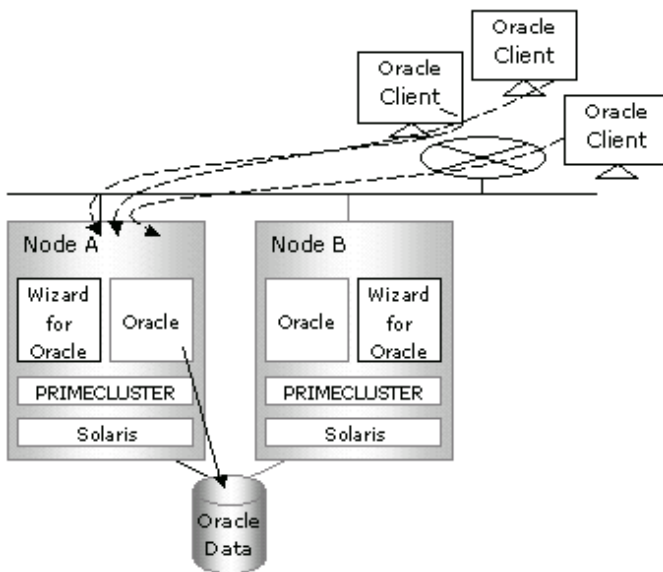
14.2.2 Supported Modes

For details, see the "Table 12.1 PRIMECLUSTER product list."

14.2.3 Overview of Cluster System Support

The PRIMECLUSTER Wizard for Oracle provides the environment setup wizard that is used to build Oracle on the cluster system, the start/stop script, and the state-monitoring detector. It also supports scalable operation using Oracle9i RAC and Oracle cold standby operation.

14.3 Notes

For details on the environment setup, etc., see the *"PRIMECLUSTER Wizard for Oracle Configuration and Administration Guide,"* which is attached to the PRIMECLUSTER Wizard for Oracle.

Chapter 15 Interstage Application Server Enterprise Edition

15.1 Functional Overview

The Interstage Application Server is an infrastructural support software product in which the latest international standard technologies including J2EE, SOAP, UDDI, ebXML, and CORBA are integrated with well-proven high-reliability and high-performance trunk system construction technologies.

The Interstage Application Server Enterprise Edition is the highest-ranking of the editions provided by the Interstage Application Server. Assuming the need to provide round-the-clock service, the Interstage Application Server Enterprise Edition guarantees a stable response even under high load.

15.1.1 Overview of Cluster System Support

The Interstage Application Server Enterprise Edition supports two modes: Namely, 1:1 standby, which is a so-called hot standby system in which the operation is automatically taken over by the standby node if a failure occurs in the OPERATING node; and mutual standby, in which two nodes perform each operation and, if a failure occurs on one node, the other node takes over the operation.

With the Interstage Application Server Enterprise Edition, operating on the cluster system allows high-speed switching to another server to take over the operation even in the event of a failure. The client can continue the transaction through reconnect processing and can minimize downtime.

15.1.2 Supported Modes

For details, see the "Table 12.1 PRIMECLUSTER product list."

15.2 Registering a SynfinityCluster-compatible State Transition Procedure

This section explains how to register a SynfinityCluster-compatible state transition procedure of Interstage Application Server 5.0.

- Topics other than the registering methods of a state transition procedure

   Refer to the description of how to use SynfinityCluster, which is included in the Interstage Application Server 5.0 manual.

- How to register a state transition procedure

   The registration process is explained below:

15.2.1 Setting Resource Information

Set:
"No. of process restart" to "0";
"Interval period of process restart" to "0"; and
"Initialize No. of process restart" to "No"
for the resource information of the state transition procedure.

15.2.2 Setting State Transition Instruction Timings

Set up the state transition procedures provided by Interstage so that they can be invoked at the timings shown below:

[START-RUN]

- AFTER

- CUTOFF/BUILDIN

[START-WAIT]

- AFTER

[STOP-RUN]

- BEFORE

[STOP-WAIT]

- BEFORE

[FAIL-RUN]

- BEFORE

[FAIL-WAIT]

- BEFORE

15.2.3 Registering a State Transition Procedure to a userApplication

Register the state transition procedures to userApplication. The "Resource startup procedure" must be set up so that resources can be invoked with the following priority:

Resource startup priority

- Startup priority to be registered to BasicApplication in the procedure
  (Set up to be invoked in the descending order.)

  service name.IS_INTERSTAGE

  service name.OTS_RMP_INTERSTAGE

  service name.ES_INTERSTAGE

  service name.SOAP_INTERSTAGE

- Startup priority of the resources to be registered to Application of the procedure class (optional)

  service name.TDWU_INTERSTAGE

  service name.EJBWU_INTERSTAGE

  service name.UTYWU_INTERSTAGE

  service name.ODWU_INTERSTAGE

  service name.EB_INTERSTAGE

Operation Procedure

Parameters to be specified for options:

*: Must be omitted if it is to be specified for TDWU_INTERSTAGE/ EJBWU_INTERSTAGE/ UTYWU_INTERSTAGE.

**: Must be changed according to the name of the resource to be registered.

Execute the following operation in each procedure on all nodes in which the state transition procedure is to be executed:

1. Register the state transition procedure.

```
clsetproc -c BasicApplication *
-m IS_INTERSTAGE **
-o /etc/opt/FJSVisas/HA/SynfinityCLUSTER/IS_INTERSTAGE **
```

2. Register the application resource that uses the state transition procedure.

```
claddprocrsc -k IS_INTERSTAGE **
-m IS_INTERSTAGE **
-c BasicApplication *
-K AFTER -w  -L AFTER -S BEFORE -T BEFORE -V BEFORE -W BEFORE -u 0 -t 0 -p 100
```

3. Register a resource used for the state transition procedure as a resource for PRIMECLUSTER.

   Register the resource created in "Procedure 2" as a resource for PRIMECLUSTER.

   Operation must be implemented from "*Create Resource*" of "*userApplication Configuration Wizard*."

   Use the format described in "Resource startup priority," above, for the resource name.

   ![See icon] **See**

   For details on the "userApplication Configuration Wizard", see "6.6.1.6 Creating Procedure Resources."

   ![Note icon] **Note**

   If two or more resources of the same kind are registered, the *SubApplication* button on the lower left will be enabled when the *Attributes* tab is selected on the "*Confirm registration*" screen.

   When the final resource is to be registered, click *SubApplication* and then **specify the resource startup priority** on the "*Associate resources*" screen.

   The startup priority can be set up by setting the resources so that resources are listed in the order described in "Resource startup priority," above, in the *Selected Resource* list of the "*Associate resources*" screen.

4. Registering state transition procedure resources to userApplication.

   Register all the resources created in Procedure 3. to userApplication.

   The operation must be implemented from "*Create userApplication*" of "*userApplication Configuration Wizard*."
   With this operation, set "*ClearFaultRequest|StartUp|SwitchRequest*" to "*StandbyTransitions*" of "userApplication attributes."

   ![See icon] **See**

   For more details on "userApplication Configuration Wizard," see "6.6.2 Creating Cluster Applications".

Chapter 16 Softek AdvancedCopy Manager

16.1 Functional Overview

This product realizes storage management for open systems and consists of the following three solutions:

High-speed backup (round-the-clock support)

Data can be backed up at high speed by using the hardware [advanced copy function (OPC, EC)] of the Fujitsu disk array unit (ETERNUS3000, GR720, GR730, GR740). Since the advanced copy function is a high-speed data copy function that is not dependent on the volume of data, a constant, high-speed backup time can be always maintained even when the amount of data to be backed up increases. Thus, the backup time does not increase even when the data quantity increases.

By linking this solution with the following DBMS, high-speed backup can be done without stopping the job:

- ORACLE

- SymfoWARE

For those jobs that are constructed using a database other than the above or a general file system, the job non-stop time required for the backup is significantly reduced compared with that in the past.

Database backup without stopping the job

Oracle and SymfoWARE database backup can both be done without stopping the job by using Softek AdvancedCopy Manager.

When combined with Global Disk Services, Softek AdvancedCopy Manager can be used to perform backup without stopping Global Disk Services, even in the mirror volume (SDX object) environment.

High-speed replication

By using the hardware [advanced copy function (OPC, EC)] of the Fujitsu disk array unit (ETERNUS3000, GR720, GR730, GR740), the data can be used for various purposes and a data replication (copy) can be collected to guard against accidental data destruction.

A remote disk array unit in the SAN environment is also available for storing the replication (copy) for the purpose of restoring the data after a catastrophic failure.

16.1.1 Overview of Cluster System Support

With Softek AdvancedCopy Manager, units constituting a cluster system that has been constructed by using cluster software are controlled as a single business server. By monitoring failover and the like, recovery from errors can be done quickly.

16.1.2 Supported Modes

Softek AdvancedCopy Manager supports 1:1 standby and mutual standby by means of IP address takeover. For details, see "Table 12.1 PRIMECLUSTER product list."

16.2 Standby Classes

16.2.1 System Configuration

The supported system configurations are shown below:

Standby type cluster system configuration
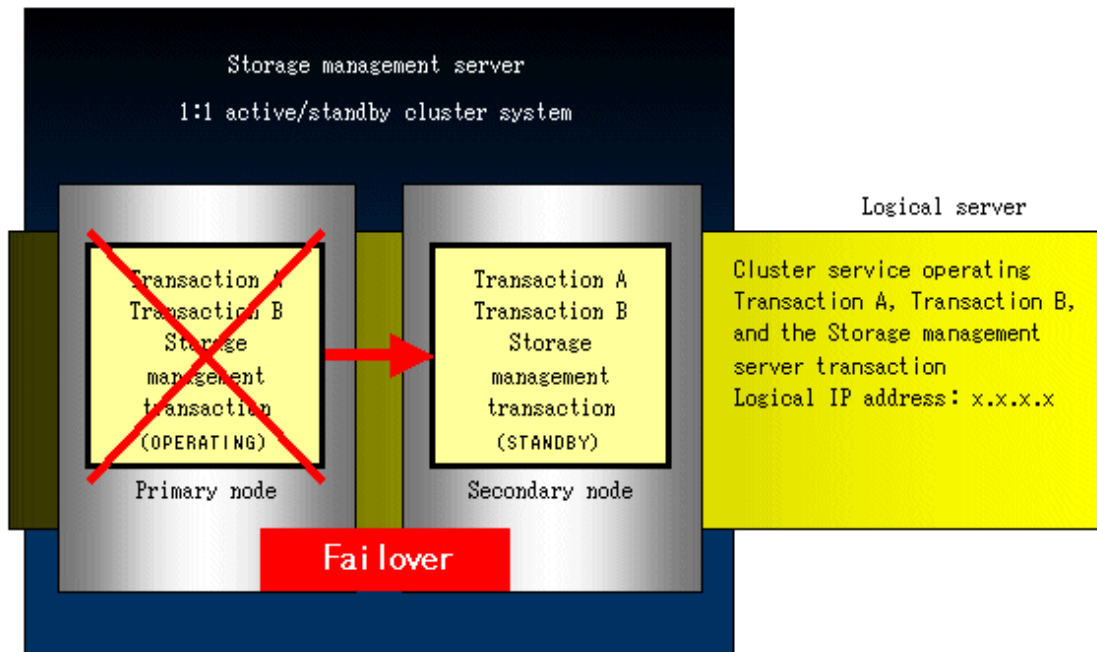
Mutual standby type cluster system configuration
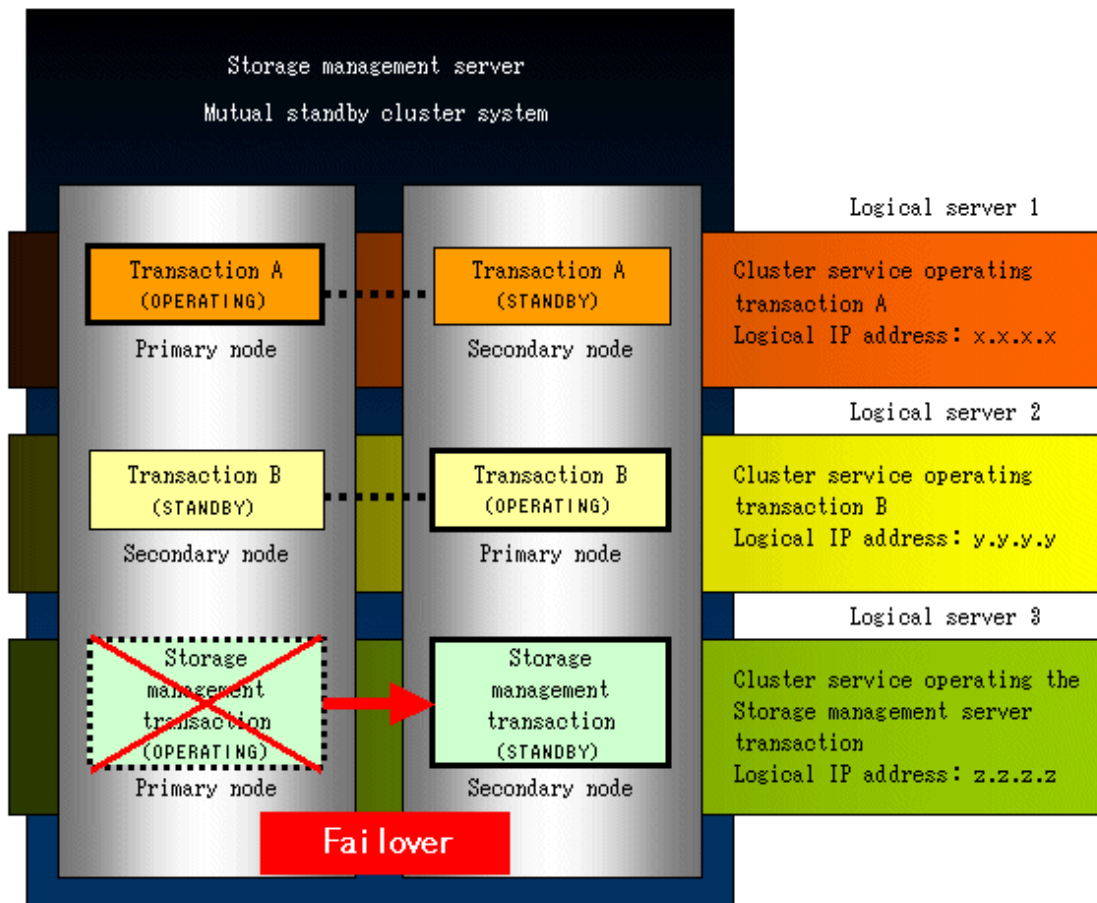
16.2.2 Setting Up the Environment

For details, see the "Softek AdvancedCopy Manager Cluster Application Guide (SynfinityCluster/PRIMECLUSTER)."

Chapter 17 Symfoware Server

This chapter explains the support provided for the Symfoware Server cluster system.

17.1 Functional Overview

Symfoware Server is a database system that can be applied to fields involving mission-critical transactions. It offers high-performance inquiry, a loaded, mass data maintenance, efficient mass data maintenance, and highly reliable operation. Levels of performance and throughput as great as those of a mainframe can be realized by making use of open system resources. Symfoware Server provides complete support for network computing technologies that enable information exchange with multiple servers and information access from any client, as well as future job extension.

17.1.1 Overview of Cluster System Support

Symfoware Server supports 1:1 standby.

Symfoware Server supports failover operation of 1:1 standby by using the Symfoware Server Hot Standby Option or Symfoware Server Parallel Cluster Option.

Symfoware Server Parallel Cluster Option enables 1:1 standby load-shared operation.

For details, see the "*SymfoWARE (R) Server Cluster Installation/Administration Guide for PRIMECLUSTER.*"

17.1.1.1 Failover Operation with Symfoware Server

Because Symfoware Server realizes failover, you can build a high-availability system. The operation involved in performing failover is referred to as the failover operation.

Failover refers to the facility that enables the taking over of a job that is being operated on the relevant node to another node (standby system) if an error occurs on one of the nodes in the cluster system. The following two functions of Symfoware Server are used to realize failover:

- Standby function

- Hot standby function

The standby function is such that, if an error occurs in a node, the standby RDB system to which the operation is to be switched is started up to take over the job operation.

To cope with a node error, the RDB system on the standby system is started up in advance and, if an error occurs on the node, failover of the operation can be completed in a short time. This is called the hot standby function.

17.1.1.2 Load-Shared Operation with Symfoware Server

Symfoware Server realizes load sharing to support high scalability. Load sharing refers to the function to process transactions in parallel on multiple nodes in the cluster system. The operation for performing load sharing is referred to as the load-shared operation.

Load sharing has the following two features:

- Performance on a par with a high-reliability system can be realized.

  One of the features of the load-shared operation is that you can realize a level of performance that is comparable with that of a high-reliability system.

  A load-shared operation is controlled by PRIMECLUSTER. The purpose of this is to ensure compatibility with other middleware products included in the cluster system by entrusting cluster system management and operations to cluster control. You can thus make maximum use of the high-scalability and high-reliability that are realized when cluster control is combined with the hardware

  Load-shared operation enables communication between the nodes that constitute the cluster system.

- An entire system can be realized as a high-reliability system.

  With load-shared operation, you can restart a job within a few seconds of a node failing.

With load-shared operation, database updated logs are transmitted sequentially, during operation, into memory of the standby node by using TCP/IP or BLASTBAND. If a switchover event occurs, the flash treatment recovery will be instigated and the first data will be closed by using the updated logs on the standby node. This enables the restart of the job transaction.

Failure recovery for lock-out data will be performed in the background in parallel with the job transaction.

17.1.2 Supported Products

For details, see "Table 12.1 PRIMECLUSTER product list."

17.2 Cluster Applications

This section explains about cluster applications when the following operation is being made:

- Failover operation

- Load-shared operation

17.2.1 Failover Operation

This section explains cluster applications for failover operation.

17.2.1.1 System Configuration

The system configuration for 1:1 standby failover operation is shown below:
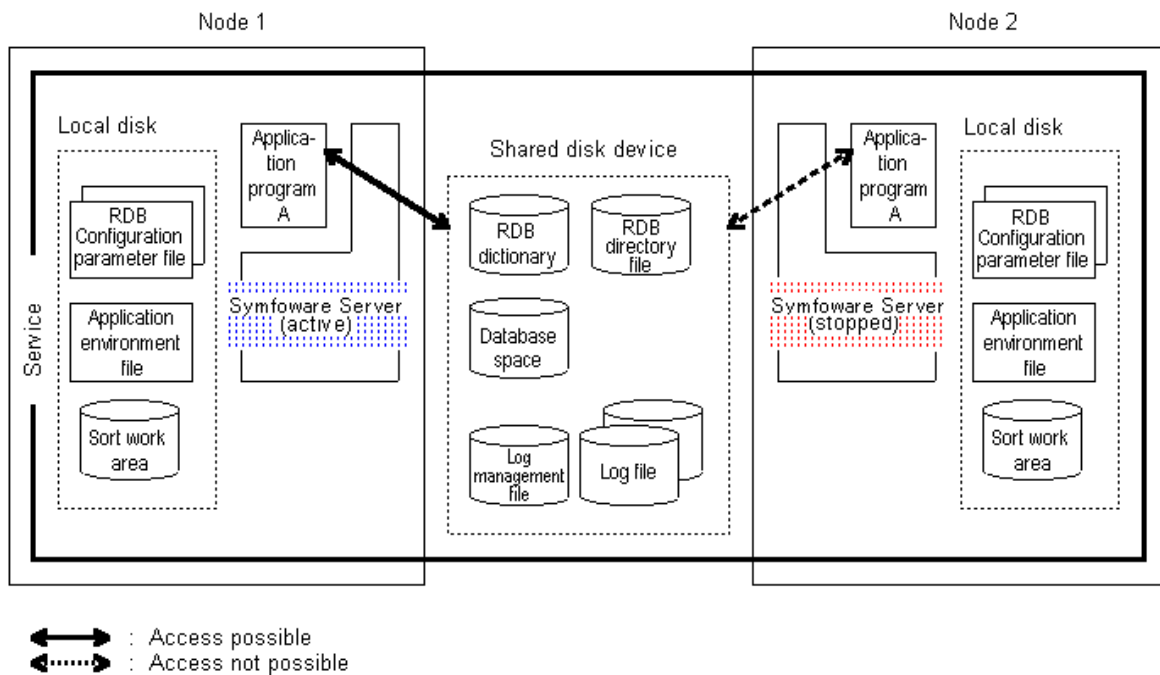
17.2.1.2 Environment Setup

You must register the RDB system with the cluster system as a resource and then set it up for the userApplication that is to be operated.

For details on the environment setup, see the "*SymfoWARE (R) Server Cluster Installation/Administration Guide for PRIMECLUSTER*."

17.2.2 Load-Shared Operation

Cluster applications in load-shared operation are explained here.

17.2.2.1 System Configuration

The system configuration of 1:1 standby in load-shared operation is shown below:
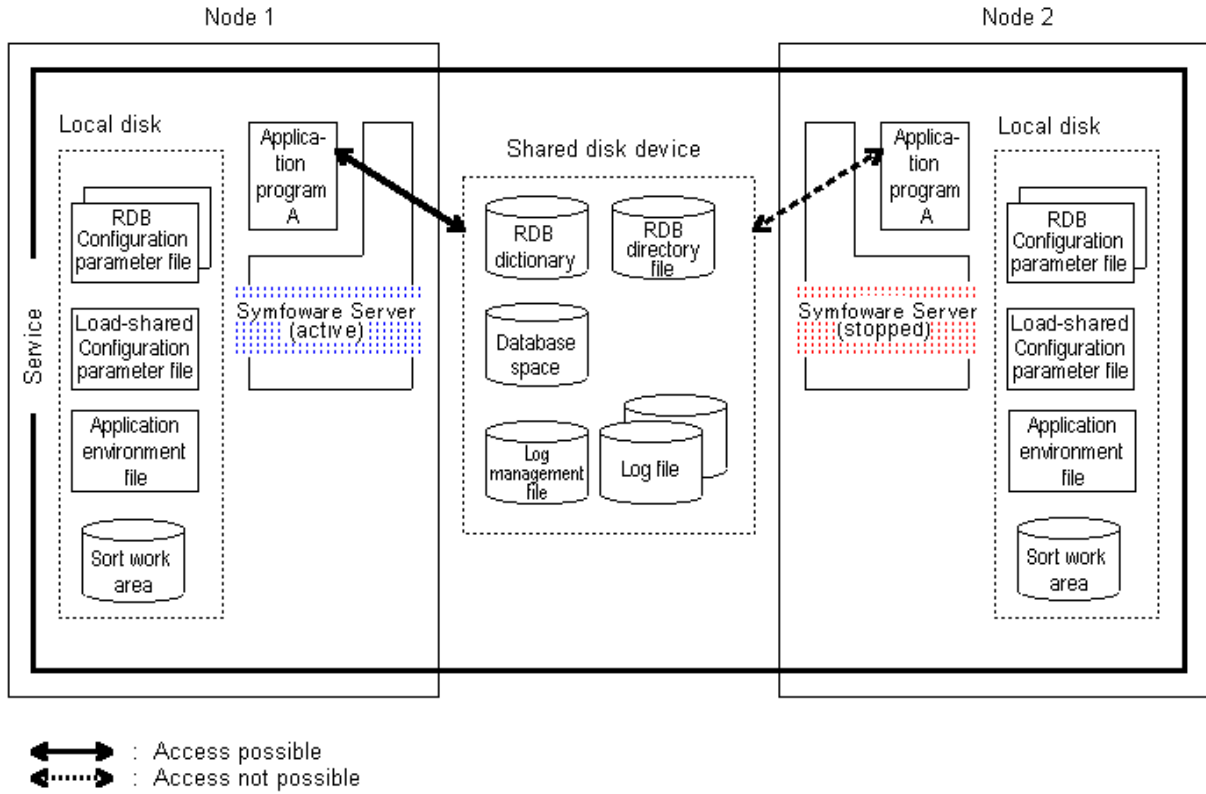
17.2.2.2 Environment Setup

You must register the RDB system with the cluster system as a resource and then set it up for the userApplication that is to be operated.

For details on the environment setup, see the "*SymfoWARE (R) Server Cluster Installation/Administration Guide for PRIMECLUSTER*."

17.3 Notes

None.

Chapter 18 Systemwalker Centric Manager

18.1 Functional Overview

Systemwalker Centric Manager is a software product that acts as the integrative infrastructure for the management of an information system. It provides life cycle management of IT resources such as networks, hardware, OS, applications, and users that consist of the system according to the policy based management.

Life cycle management

Systemwalker Centric Manager controls the life cycle of IT resources. It provides an effective operation management by categorizing the IT resources into the following 4 phases:
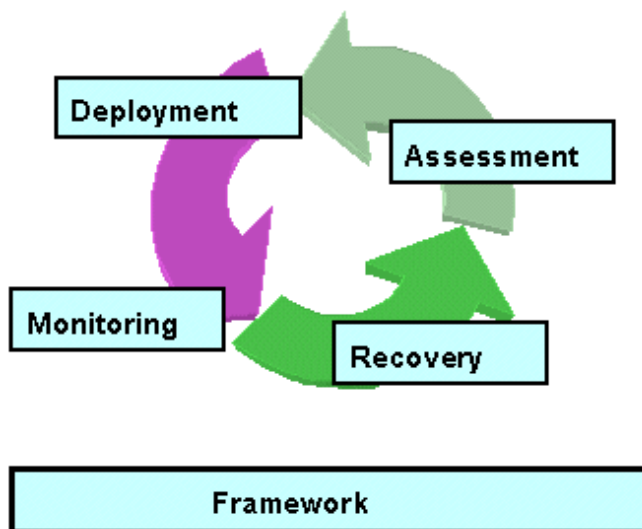
Deployment

Refers to the phase in which IT resources and transactions are installed on the system and set up. The resources to be used for the information system are deployed and applied.

Monitoring

Refers to the phase in which the performance of IT resources and failures are monitored.

Recovery

Refers to the phase in which corrective action is taken in the event of a failure. Troubleshooting and recovery are performed.

Assessment

Refers to the phase in which the stability of operation is evaluated. The service level is evaluated and analyzed from operation information of IT resources. Performance bottlenecks are detected, and capacity planning is done.

Framework

The "framework" refers to the operation management infrastructure that is used to share information between the functions of the life cycle and to enable seamless functional cooperation. The framework collects the configuration information for the entire information system, such as the network, hardware, OS, applications and others, and maintains the integrated management of the configuration information.

18.1.1 Overview of Cluster System Support

Systemwalker Centric Manager enables the integrated management of a cluster system. Systemwalker Centric Manager monitors that the entire cluster system is normally running. Also, Systemwalker Centric Manager eliminates downtime caused by node-down, and provides a highly reliable system by performing quick recovery at failover.

18.1.2 Supported Modes

Systemwalker Centric Manager supports a 1:1 standby operation mode using IP address takeover.

For details, see the "Table 12.1 PRIMECLUSTER product list."

18.2 Standby class

18.2.1 System Configuration

Systemwalker Centric Manager enables a user to use an operation management server, section management server, or job server in a cluster system. The system configuration of using each type of server is illustrated below:

Operation Management Server
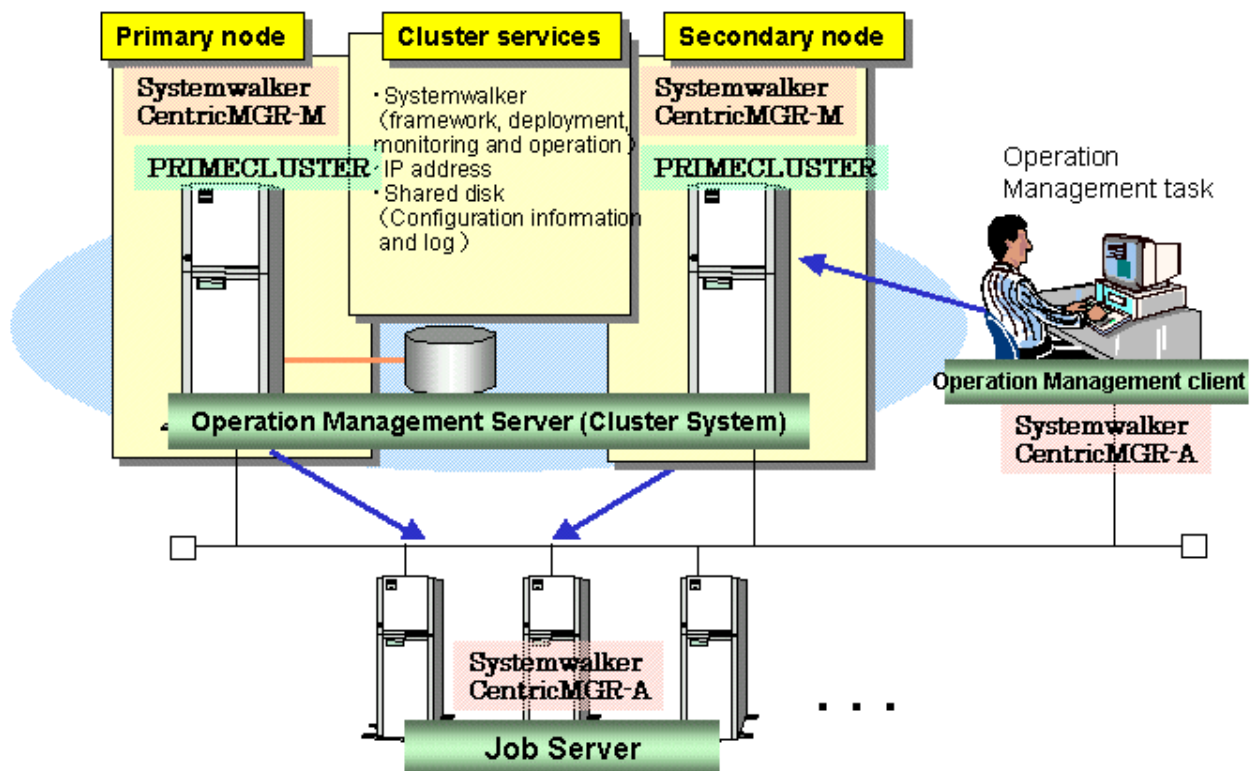
Section Management Server
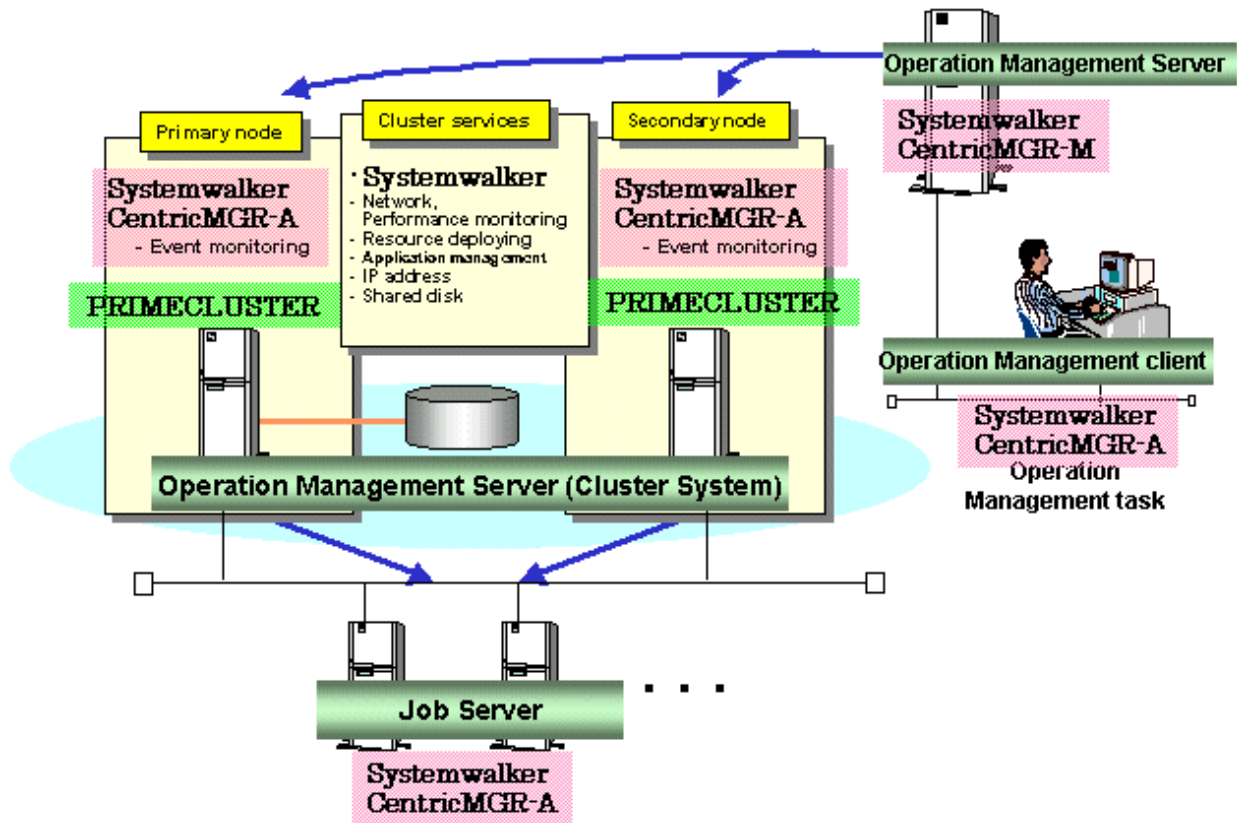
Job server
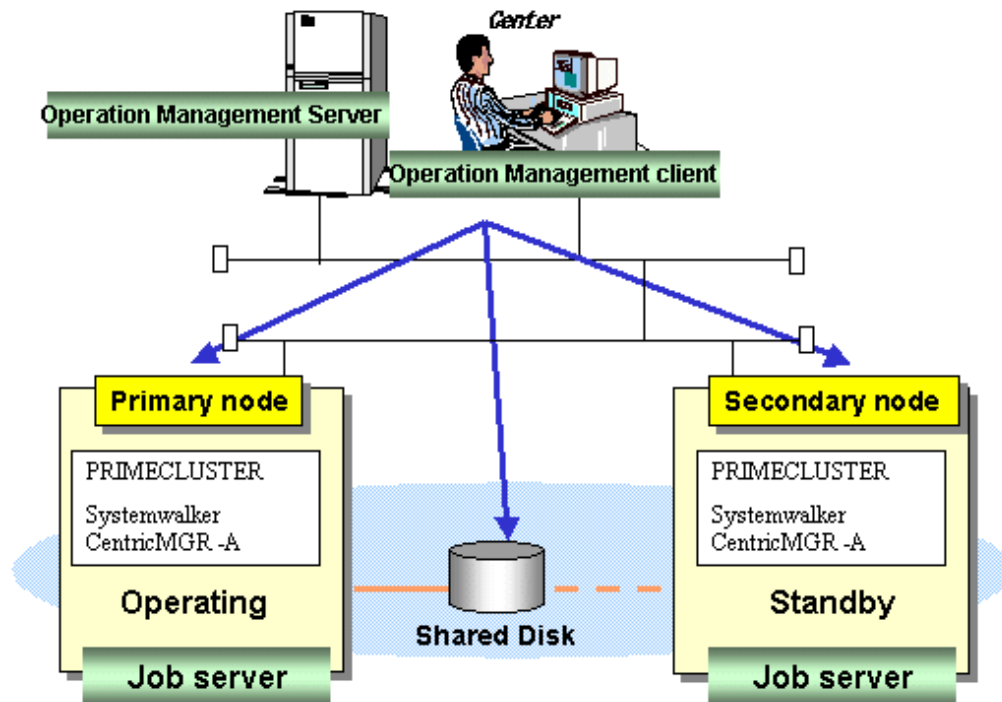
18.2.2 Environment Setup

For details, see the *"Systemwalker Centric Manager Customization Guide."*

Chapter 19 Systemwalker Operation Manager

19.1 Functional Overview

Systemwalker Operation Manager supports the following functions:

Automation of routine system operation and management

By installing Systemwalker Operation Manager, routine system operation and management can be automated as follows:

Automatic control of power on/off

Power to the server can be automatically turned on/off. Power on/off of the server is conducted according to a time schedule that is defined in advance. Also, power to the server can be automatically turned on when a client starts.

Note that this function is enabled only for Solaris servers that are running Systemwalker Operation Manager UNIX version.

Automatic activation of the services/applications

When the server is booted, services and applications are activated in the order in which they are registered. You can change the services and applications on any given day.

Note that this automatic activation of the services can be used only when the connected server is for the Windows version.

Automatic execution of a routine batch job

If you register the times and dates at which a routine batch job is run, the routine batch job is automatically executed according to that schedule. If an error occurs in a routine batch job, it can be corrected automatically without the intervention of the operator.

Scheduling file compression and transfer

The commands used for file compression and transfer are provided as job-associated commands. If you register these files with the job scheduler, you can execute these commands by scheduling.

Monitoring and operating a job on the screen

By installing Systemwalker Operation Manager, you can confirm the job execution status at a glance because the job status is displayed in different colors. You can monitor jobs running not only on a single system but also those running on multiple systems.

Moreover, you can operate a job while monitoring its status.

Control of job execution according to the system operation and status

Systemwalker Operation Manager enables you to control the job execution, as follows:

Control of job execution environment

Jobs can be executed efficiently by adjusting the job execution environments such as the execution order and the number of jobs that can be executed simultaneously.

Job execution with networked servers

Jobs can be executed on any server connected to the network.

Control of networked clients

You can provide clients that are connected over the network with controls such as power-on and file transfer.

19.1.1 Overview of Cluster System Support

Systemwalker Operation Manager provides high availability batch jobs and enables a high-reliability transaction system by running on a cluster system.

19.1.2 Supported Modes

Systemwalker Operation Manager supports the following operation modes:

- 1:1 standby

- N:1 standby

- Two-node mutual standby

For details, see the "Table 12.1 PRIMECLUSTER product list".

19.2 Standby Classes

19.2.1 System Configuration

Systemwalker Operation Manager enables the entire servers, schedule server, execution server, and schedule server + execution server to operate on a cluster system in 1:1 standby, N:1 standby, or two-node mutual standby operation mode.

Each cluster system configuration when the entire Systemwalker Operation Manager servers are set up in a cluster configuration is outlined below:

1:1 standby

The following figure shows an example in which node 1 operates as an active and node 2 operates as a standby:
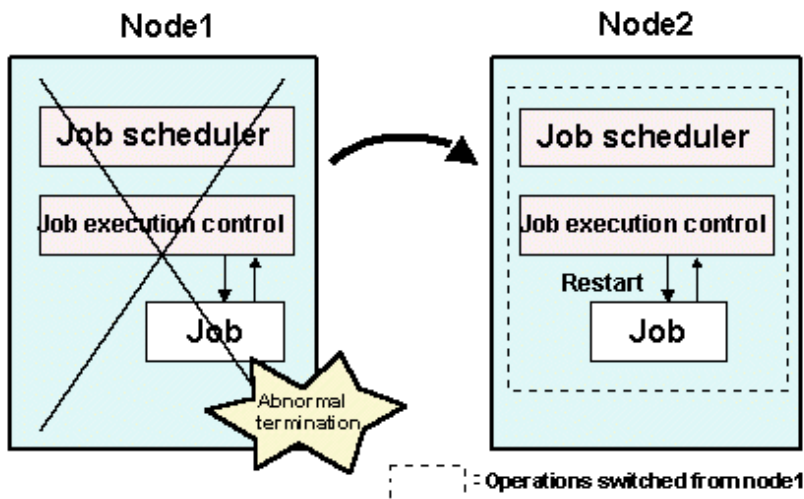
If an error occurs in node 1, ongoing operations are switched to node 2.

N:1 standby

The following figure shows an example in which node 1 and node 3 are active and node 2 operates as a standby. The following shows an example where an error occurs on node 1. Even if an error occurs on node 3, ongoing operations are switched to node 2.
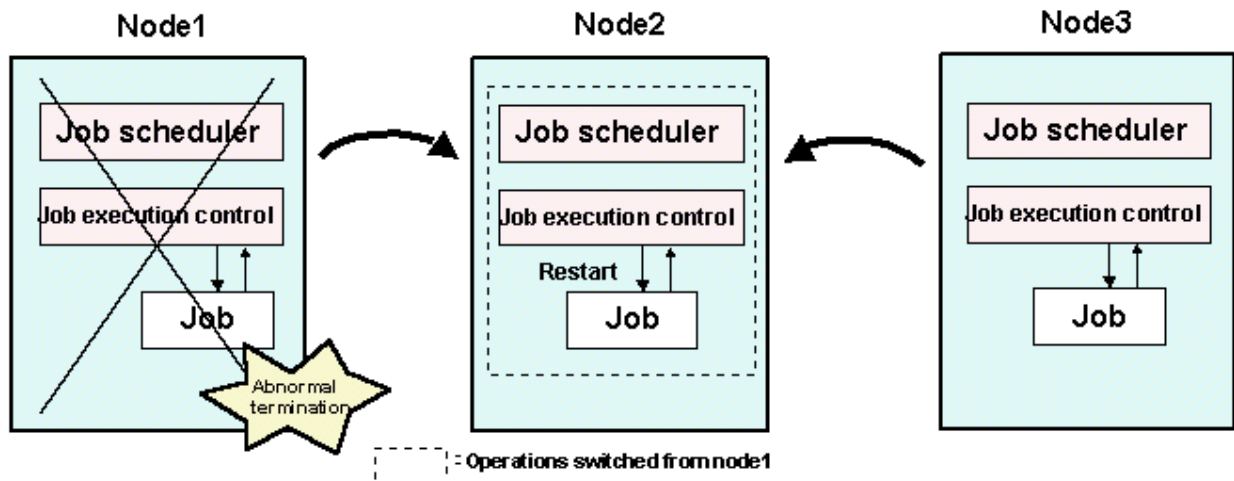
If an error occurs on node 1 or node 3, ongoing operations are switched to node 2. If errors occur on two or more nodes, the transaction running on the node on which the error occurred first is taken over by the standby node. However, the transactions of the nodes on which the errors occurred subsequently will not be taken over.

Two-node mutual standby

The following figure shows an example in which node 1 and node 2 execute their transactions while each operates as a standby for the other. The following shows an example where an error has occurred on node 1. If an error occurs on node 2, ongoing operations are switched to node 1.
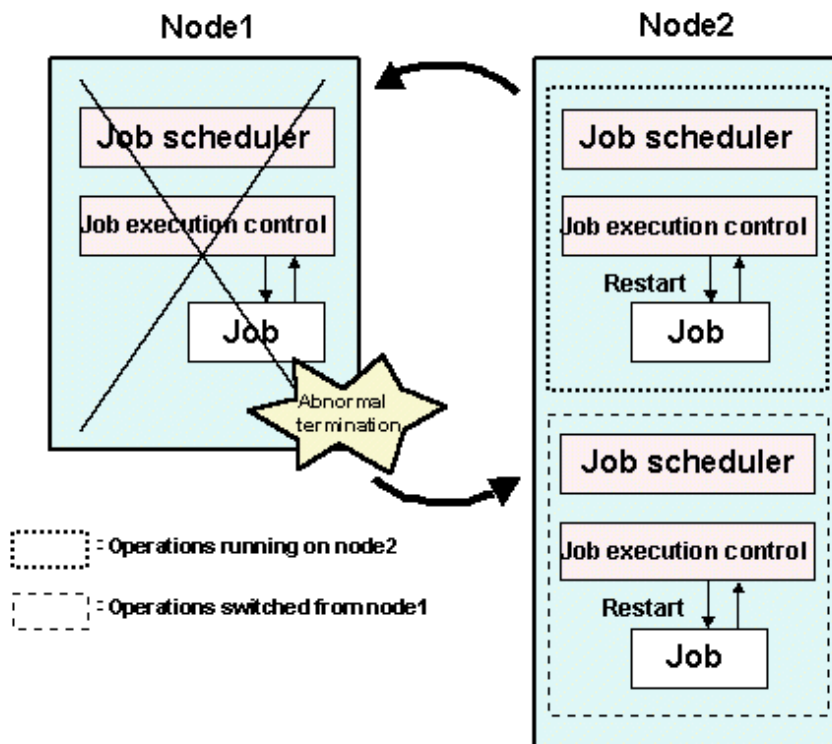
For details, see the *"Systemwalker Operation Manager Installation Guide."*

19.2.2 Environment Setup

For details, see the *"Systemwalker Operation Manager Installation Guide."*

Chapter 20 Systemwalker WebMGR

20.1 Functional Overview

Systemwalker WebMGR provides the functions that are required for supporting the integrated management of a Web system, ranging from SLA operation to the marketing strategies.

Monitoring, evaluation, and prediction of the Web system service level

By accessing services such as HTTP/HTTPS, CGI, ASP, Servlet, and JSP to collect information on the throughput and the operation status, you can monitor, evaluate, and predict the service levels.

From the user's point of view, you can collect and evaluate service level information such as response/page download performance and the operating ratio.

Analysis of the state of use

You can optimize business performance because customer needs can be comprehended and Web page configurations can be evaluated by analyzing the usage of the Web server.

Revision monitoring

You can check for invalid corrections or revisions of the contents on the Web server.

20.1.1 Overview of Cluster System Support

Systemwalker WebMGR supports 1:1 standby and provides commands and state transition procedures to be built on a cluster system.

20.1.2 Supported Modes

For details, see the "Table 12.1 PRIMECLUSTER product list."

20.1.3 Notes

None

20.2 Standby Classes

20.2.1 System Configuration

For details, see "15.7 Operation in a Cluster Environment" in the "Systemwalker WebMGR" manual.

20.2.2 Notes

None

Appendix A PRIMECLUSTER System Design Worksheets

A.1 Worksheet Purpose

The PRIMECLUSTER System Design Worksheets allow you to determine necessary parameters before you configure the PRIMECLUSTER system. Filling out all items in the worksheet will enable you to proceed smoothly with the PRIMECLUSTER installation work.

The worksheets consist of the following:

Overall Design Worksheet

This worksheet is used to design basic items for the PRIMECLUSTER system.

Cluster Installation Environment Setup Worksheet

Use this worksheet to configure the environment before you install Solaris, PRIMECLUSTER, and other required and related software. Also use the worksheet to configure the PRIMECLUSTER system using Web-Based Admin View.

For the installation of each software product, you can select either installing on each node using an installation script or using the One Shot Installer to automatically install software on all nodes.

🔖 See
. . . .
For details on each installation method, see the *"PRIMECLUSTER Installation Guide."*
. . . .

This section identifies the worksheet for preparing and organizing necessary information for each installation method. This worksheet will enable a smooth installation.

Kernel Parameter Worksheet

This worksheet allows you to determine the fundamental PRIMECLUSTER values for estimating the kernel parameters. After installing the Solaris software and PRIMECLUSTER packages, configure the new values for the kernel parameters, as necessary.

Cluster Configuration Worksheet

This worksheet organizes the necessary items for configuring the PRIMECLUSTER system with Web-Based Admin View. Using this worksheet and determining the parameter values beforehand will enable you to smoothly set up operations using Web-Based Admin View.

Cluster Application Configuration Worksheets

This worksheet is used to design how the application should operate when building a cluster application.

Cluster Application Setting Worksheet

This worksheet is used for configuring cluster applications and resources.

A.2 Notes on Worksheet Creation

Note the following points when creating PRIMECLUSTER design worksheets.

- Specify appropriate values in the blank sections of each design sheet. For instructions on specifying each item, see the point explanations.

- Items marked with "<>"are selection items. Select one of the items written in the box.

- Items marked with "[ ]" are selection items. You can select multiple items written in the box.

- Different values can be specified depending on the model in which PRIMECLUSTER is to be installed. Read the point explanations before specifying values.

A.3 Overall Design Worksheet

Select the operation mode of the cluster system.

📑 See

. . . .

See "A.10 System Design Examples" when determining values for the setup items in the Overall Design Worksheet.

. . . .

Cluster :	
System requirement	**Selective product/function/configuration**
Cluster system topology	
(1) Installation objective	<>Standby operation
( )Improvement of availability	( )1:1 standby
( )Improvement of performance	
(2) Operation after failover	( )Mutual standby
( )Degenerate operation enabled	
( )Degenerate operation disabled	( )N:1 standby
(3) Intersystem shared DB	
( )Provided	( )Cascade operation
( )Not provided	
(4) Node count	( )Priority transfer
( )2 nodes	
( )3 or more nodes	<>Scalable operation
(5) Web-Based Admin View operation modes	
( )2-layer model	<>1 PC required
( )Integrated console management	[]RC2000 ※Required on the PC
( )3-layer model	<>Operation management server and 1 PC required
( )Integrated console management	[]RC2000 ※Required on the management server or PC

```
(6) Select the shared disk units
 1) 2-node configuration

 <>RAID                                  [ ]Multipath disk control facility
    <>Interface                             <>MPHD  <>ETERNUS multipath driver
    <>SCSI                              Note: MPHD is supported only by GR710 and higher models
    <>Fibre Channel                     [ ]Multipath load balance (MPLB)
    <>Path                                  <>MPLB  <>ETERNUS multipath driver
     <>Single path                      [ ]Fujitsu PCI Fibre Channel (PFCA)
     <>Path duplex switchover               (hardware attachment)
     <>Path duplex load balancing       [ ]GDS (Global Disk Services)
     <>Path triplex or higher
 <>Use FDU/FU
  <>Manage disk cabinet
 2) 3 or more nodes (*1)                 [ ]Multipath disk control facility
 <>RAID                                     <>MPHD  <>ETERNUS multipath driver
    <>Interface                         Note: MPHD is supported only by GR710 and higher models
    <>Fibre Channel                     [ ]Multipath load balance
    <>Path                                  <>MPLB  <>ETERNUS multipath driver
   <>Single path                        [ ]Fujitsu PCI Fibre Channel (PFCA)
     <>Path duplex switchover               (hardware attachment)
     <>Path duplex load balancing       [ ]GDS (Global Disk Services)
     <>Path triplex or higher
 <>Manage disk partitions
 3) Switchover file system selection
 <>ufs
 <>High reliability above ufs           [ ]GFS (Global File Services)
 <>Simultaneous sharing
  <>File system

 <>RAW
```

```
(7)NTP
 1)Cluster topology                     <>NTP broadcast
  ( )No existing NTP network.           <>Place NTP server on the cluster node
  ( )Existing NTP network.              <>Synchronous with the NTP server outside the cluster
 2)Protocol (only when  broadcast is specified)
  ( )Node name takeover                 <>multicastclient
  ( )No node name takeover              <>broadcast
 3)NTP network setting                  Host name               IP address
  ( )When there is an NTP server outside
     the cluster system


  ( )NTP setting within the cluster system
    Primary NTP server
    Secondary NTP server
```

In this chart, ISV products (Veritas, Emulex, EMC PowerPath) should be mentioned.

A.4 Cluster Installation Environment Worksheet

The Cluster Installation Environment Worksheet organizes items related to installation and initial configuration of Solaris software, PRIMECLUSTER, required and related software. Create the Cluster Installation Environment Worksheet and the Solaris Installation Worksheet.

The Cluster Installation Environment Worksheet contains initial configuration items related to the entire PRIMECLUSTER system. If you are using the One Shot Installer, respond to queries from One Shot Installer. If you are not using One Shot Installer, configure a system on each node according to the information specified here.

In the Solaris Installation Worksheet, specify the configuration items for installing Solaris software on each node including the installation server when using One Shot Installer.

Cluster system name:		Cluster Installation Environment Worksheet
Node name:		

Settings		Notes			
(1) Software to be installed	[ ]Optional software	[ ]RAID Disk Driver (HDDV): Depends on hardware configuration [ ]Fibre Channel driver: Depends on hardware configuration [ ]Multi-path disk control (MPHD) [ ]Multipath load balance (MPLB) [ ]ETERNUS multipath driver			
(2) Version of Solaris	<>9 <>10	Use the same version on all nodes.			
(3) Partition information		Logical path	Size		Configuration should besynchronized on all nodes
	/				
	swap				
	/usr				
	/var				
	/opt				
	/export/home				
(4) System locale	<> C <> Ja <> Other ( )				
(5) Time zone	<> Japan <> Other ( )				

(6) Node information
 	
  Node name
  IP
  MAC					

(7)Web-Based Admin View	Node name	IP	Access I/F	Http port number	RMI port number
Primary management serve Secondary management server	_____ _____ _____ _____	_____ _____ _____ _____	_____ _____ _____ _____	_____ _____ _____ _____	_____ _____ _____ _____

Note

Use a single primary management server of Web-Based Admin View to monitor a single cluster system.

Solaris Installation Worksheet

Specify information needed for installing the Solaris software in nodes. Refer to the worksheet created here when installing various software programs to each node individually and when building the installation server.

In a cluster system, the items other than the host names and IP addresses of the individual nodes have the same values in all nodes.

See

For details of each item, see the "Solaris X Information Library."

Setting item	Setting value	Remarks
Select a Language	0) English	This item is displayed only for the English version of Solaris. If the environment is not a Japanese language environment, select 0 (English).
Select a Locale	0) USA - English (ASCII only)	This item is displayed only for the English version of Solaris. If the environment is not a Japanese language environment, select 0 (English).
What type of terminal are you using?		Set this item based on the terminal emulator. Select vt52 for using Windows telnet.
Host Name		Specify the node name of each node recorded on the initial setup worksheet.
Network Connectivity	Yes	Specify Yes.
Primary Network Interface	_____	Specify the network interface connected to the public LAN.
IP address		Specify the node name of each node recorded on the initial setup worksheet.
Name Service	None	Specify None.
Subnets	Yes/No	Specify Yes if the public LAN is part of the subnet.
Netmask	___.___.___.___	If Yes is selected for Subnets, set the net mask of the subnet.
Time Zone		Specify the default time zone.
Data and Time		The initial setup tool matches the date and time of other nodes to the date and time set for the installation server. Set the correct date and time.
Solaris Interactive Installation	Initial	Select Initial.
Allocation Client Services?	Continue	Select Continue.
Select Languages	ja,_____	Select the required language. For a Japanese language environment, you also need to select ja.
Select Software	Entire Distribution	Select Entire Distribution.
Select Disks		Select the disk on which Solaris is installed.
Preserve Data?	Continue	Select Continue.
Automatically Layout File Systems?	Manual Layout	Select Manual Layout. Set so that the disk slice configuration of all nodes is the same.
Customize Disk		Select Manual Layout, so the disk slice configuration is synchronized on all nodes.
Mount Remote File System?	Continue	Select Continue.
Reboot After Installation?	Auto Reboot	Select Auto Reboot to automatically restart after installation is done.

A.5 Kernel Parameter Worksheet

When you build the PRIMECLUSTER system, the default values for the kernel parameters of the Solaris may be inadequate. The Kernel Parameter Worksheet shows the kernel parameter values used for PRIMECLUSTER.

Set up the kernel parameter as follows according to the type of "Characteristics" included in the table.

- Addition

   Specify the total of the system default value and the recommended or specified values for each software product.

- Maximum value

Specify a maximum for the recommended or specified values for each software product.

You must, however, use the default value if the maximum value is smaller than the system default value.

The kernel parameter values differ depending upon:

- CF configuration

- RMS configuration

- Using GFS

- Using GLS

See
. . . .

- For details of the kernel parameters and instructions on changing parameter values, see "Solaris System Administration" in *"Solaris X System Administrator Collection Vol. 1."*

- For information on the default values of the Solaris, see the *"Solaris Tunable Parameters Reference Manual."*
. . . .

Note
. . . .

- The values used by products and user applications that operate in the PRIMECLUSTER system must also be included in the kernel parameter values.

- To enable values that have been changed, you must restart the system after the installation.

- If a kernel parameter value is already maximized, the change will not be added to the system.

- If a process is started without process monitoring from a state transition procedure, the hard- and software limits in the file descriptor (fd) for that process will both be "1024", even after changing the kernel parameter values.
If a process started from a state transition procedure requires a file descriptor above "1024", please add a "ulimit" definition at the beginning of processing for the state transition procedure, and specify the maximum value for the file descriptor.
. . . .

A.5.1   CF Configuration

When the resource database is used

The table below shows the kernel parameter values that are required in the CF configuration when the resource database is used.

The kernel parameter values in /etc/system are automatically changed by the installer. Be sure to return the settings in /etc/system to their original state when you uninstall the package.

Note
. . . .

For system expansion, if you increase the number of nodes and logical disks, you need to re-estimate the resources and restart each node in the cluster system. If you want to add nodes or logical disks to a cluster system after it is configured, it is necessary to set a kernel parameter in advance considering the number of the nodes and logical disks.
. . . .

Solaris 9

Kernel parameter	Characteristics	Recommended value/minimum value	Comments (parameter description)
shmsys:shminfo_shmmax	Maximum value	4194304 *	Value required for resource database

Kernel parameter	Characteristics	Recommended value/minimum value	Comments (parameter description)
shmsys:shminfo_shmmni	Addition	30	Value required for resource database
semsys:seminfo_semmni	Addition	20	Value required for resource database
semsys:seminfo_semmns	Addition	40	Value required for resource database
semsys:seminfo_semmnu	Addition	40 **	Value required for resource database

*

Change the shmsys:shminfo_shmmax value required by the cluster system to the value shown below.

Estimate the number of cluster system resources according to the following equation:

Number of resources = (a) + (b)

(a) Number of disks in shared system devices x (number of shared nodes + 1) x 2

(b) Total number of local disks (number of local disks in all cluster configuration nodes)

Value required for resource database = 1048576 + 2776 x number of resources

  - If the value calculated above is larger than the installation default value of Solaris 9 (8388608):

    shmsys:shminfo_shmmax = Value required for resource database

  - If the value calculated above is smaller than the installation default value of Solaris 9 (8388608):

    You do not need edit shmsys:shminfo_shmmax .

    (The installation default value of Solaris 9 is used.)

**

The installer automatically changes a kernel parameter value. It also sets semsys:seminfo_semmnu, which is the same as semsys:seminfo_semmns calculated by PRIMECLUSTER. Be sure to check and adjust the value according to the system environment.

Solaris 10

Kernel parameter	Characteristics	Recommended value/minimum value	Comments (parameter description)
shmsys:shminfo_shmmax	Maximum value	4194304 *	Value required for resource database
shmsys:shminfo_shmmni	Addition	30	Value required for resource database
semsys:seminfo_semmni	Addition	20	Value required for resource database

*

Change the shmsys:shminfo_shmmax value required by the cluster system to the value shown below.

Estimate the number of cluster system resources according to the following equation:

Number of resources = (a) + (b)

(a) Number of disks in shared system devices x (number of shared nodes + 1) x 2

(b) Total number of local disks (number of local disks in all cluster configuration nodes)

Value required for resource database = 1048576 + 2776 x number of resources

- If the value calculated above is larger than the installation default value of Solaris 10 (8388608):

  shmsys:shminfo_shmmax = Value required for resource database

- If the value calculated above is smaller than the installation default value of Solaris 10 (8388608):

  You do not need edit shmsys:shminfo_shmmax .

  (The installation default value of Solaris 10 is used.)

RCI monitoring agent setup

When you set up asynchronous RCI monitoring, you must specify the timeout interval (kernel parameter) in /etc/system for monitoring via SCF/RCI. Kernel parameters vary depending on the server type. Then check your server type so you can set the appropriate timeout interval.

📒 Note

. . . .

This setting is not required in the following cases:

- SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 provided by companies other than Fujitsu in Japan

- SPARC Enterprise M3000, M4000, M5000, M8000, and M9000 with logos of both Fujitsu and Oracle provided in other than Japan

. . . .

Below table shows the server types that require setting of the monitoring timeout interval.

Server type	Model	Kernel parameter (driver name: scf_rdctrl_sense_wait)
SPARC Enterprise	M3000	**scfd**:scf_rdctrl_sense_wait
	M4000	
	M5000	
	M8000	
	M9000	
PRIMEPOWER	800	**FJSVscf2**:scf_rdctrl_sense_wait
	1000	
	2000	
	900	**FJSVscf3**:scf_rdctrl_sense_wait
	1500	
	2500	

a. Method for Calculating Monitoring Timeout Intervals

Calculate monitoring timeout intervals as follows:

- Up to 2 domains: 2 seconds

- 3 or more domains: 1 second + (0.5 x number of domains)

```
Example:
 - 3 domains: 2.5 seconds
 - 4 domains: 3.0 seconds
```

📒 Note

. . . .

- In case of PRIMEPOWER servers, count each partition as 1 domain.

- Calculate timeout intervals based on the number of domains in the server that contains the largest number of domains in the RCI network.

・・

b. Method for Setting Timeout Intervals in /etc/system

Before setting up the initial cluster configuration, modify /etc/system for all nodes according to below procedure.

1. Make a backup of /etc/system.

   Example:
   Make a copy of /etc/system and save it under the filename /etc/system.org.

   ```
   # cp /etc/system /etc/system.org
   ```

2. Set the monitoring timeout interval in /etc/system.

   As the monitoring timeout interval is specified in microseconds, you have to multiply the seconds calculated in above item "a." by 1,000,000 for this setting.

   ```
   set driver name: scf_rdctrl_sense_wait = monitoring timeout
   interval [microseconds]
   ```

   Example:
   Setting a 2-second monitoring timeout interval for a SPARC Enterprise server with 2 domains.

   ```
   set scfd:scf_rdctrl_sense_wait = 2000000
   ```

3. Restart the node.

   Example:

   ```
   # /usr/sbin/shutdown -y -g0 -i6
   ```

A.5.2 RMS Configuration

The table below shows the values of the kernel parameters required for the RMS configuration.

The kernel parameter values in /etc/system are automatically changed by the installer. Be sure to return the settings in /etc/system to their original state when you uninstall the package.

Solaris 9

Kernel parameter	Characteristics	Recommended value/ minimum value	Comments (parameter description)
msgsys:msginfo_msg max	Maximum value	16384	Queue size that can be used for a single message transmission.
msgsys:msginfo_msg mnb	Maximum value	4194304	Maximum size of the message that can be stored in a single message.
msgsys:msginfo_msg mni	Addition	8142	Maximum number of message queue identifiers that can be used for the entire system.
msgsys:msginfo_msgt ql	Maximum value	65495	Maximum number of message headers

Solaris 10

Kernel parameter	Characteristics	Recommended value/ minimum value	Comments (parameter description)
msgsys:msginfo_msgm nb	Maximum value	4194304	Maximum size of the message that can be stored in a single message.
msgsys:msginfo_msgm ni	Addition	8142	Maximum number of message queue identifiers that can be used for the entire system.
msgsys:msginfo_msgtq l	Maximum value	65495	Maximum number of message headers

A.5.3 Using GFS

The kernel parameters required to enable the use of the GFS shared file system or GFS local file system are shown below:

Kernel parameter	Characteristics	Value	Comments (parameter description)
lwp_default_stksize	Maximum value	16384 (*1)	kernel stack size
svc_default_stksize (*2)	Maximum value	16384	Kernel stack size of the NFS service thread

(*1)

The size of the kernel stack area is set to this value during GFS installation if the size is not specified as a parameter in a Solaris 32-bit environment. Increase the value as necessary.

(*2)

This is needed only if you are accessing a GFS shared file system via NFS in a Solaris 32-bit environment.

Note

When you use a GFS local file system, you have to set up the kernel parameter of the GFS local file system. For details, see "9.2.3 Turning Parameters" in the "*PRIMECLUSTER Global File Services Configuration and Administration Guide.*"

A.5.4 Using GLS

The table below shows values for the kernel parameters required for when Global Link Services (GLS) is used.

Solaris 9

Kernel parameter	Characteristics	Recommended value/minimum value	Comments (parameter description)
shmsys:shminfo_shmmax	Addition	5120 or greater	Maximum size of a shared memory segment
shmsys:shminfo_shmmni	Addition	2 or greater	Maximum number of shared memory segments
semsys:seminfo_semmni	Addition	1 or greater	Maximum number of semaphore identifiers

Kernel parameter	Characteristics	Recommended value/minimum value	Comments (parameter description)
semsys:seminfo_semmns	Addition	1 or greater	Maximum number of semaphores in the system

Solaris 10

Kernel parameter	Characteristics	Recommended value/minimum value	Comments (parameter description)
project.max-shm-memory	Addition	5120 or greater	Total bytes of shared memory
project.max-shm-ids	Addition	2 or greater	Total number of shared memory identifiers
project.max-sem-ids	Addition	1 or greater	Total number of semaphore identifiers

The system default values are necessary for the other kernel parameters.

The values shown above are necessary regardless of whether the GLS environment is set up.

A.6 Cluster Configuration Worksheet

The Cluster Configuration Worksheet is required for building a cluster.

Use this worksheet to organize the items before using Web-Based Admin View to set up the PRIMECLUSTER system.

 See
. . . .

For information on the setup items for this worksheet, see "Chapter 5 Building a Cluster." Also see the "A.10.1 Cluster Configuration Worksheet" in the "A.10 System Design Examples."
. . . .

	Item		Setting	Remarks
Shared disk unit: Type 1	Machine model			
	Device name			
	IP address		- 369 -	
	Subnet mask			
	Interface			
	LUN number			
	Ports			
	Connection paths			
	FC-SW	Use of fibre channel switching hub	☐ yes ☐ no	
		Machine model		
		Device name		
		IP address		
		Subnet mask		
	File system selection (type of file system used in cluster)			
Shared disk unit: Type 2	Machine model			
	Device name			
	IP address			
	Subnet mask			
	Interface			
	LUN number			
	Ports			
	Connection paths			
	FC-SW	Use of fibre channel switching hub	☐ yes ☐ no	
		Machine model		
		Device name		
		IP address		
		Subnet mask		
	File system selection (type of file system used in cluster)			
NTP	Operation mode		☐ NTP server (integrated in cluster) ☐ NTP client	
	Protocol (only when broadcast is specified)			
	Server	Primary	Host name		
			IP address		
			Subnet mask		
		Secondary	Host name		
			IP address		
			Subnet mask		
Web-Based Admin View	Operation mode		☐ 1 ☐ 2 ☐ 3 layers	
	Server	Primary	Host name		
			IP address		
			Subnet mask		
			URL		
		Secondary	Host name		
			IP address		
			Subnet mask		
			URL		
	User ID	wvroot group		
		clroot group		
		cladmin group		
		clmon group		
		sdxroot group		

	Cluster name			
	Cluster nodes			
	Cluster node names			
Cluster Foundation (CF)	Subnets			
	Subnet IP (network address)			
	Subnet mask			
	Use of CF remote services		☐ cfcp (CF file copy) ☐ cfsh (CF remote command execution)	
	Use by RMS		◇ [yes] ◇ no	[] default
Shutdown facility (SF)	Shutdown agent		☐ RCI Panic (SA_pprcip) ☐ RCI Reset (SA_pprcir) ☐ XSCF Panic (SA_xscfp) ☐ XSCF Reset (SA_xscfr) ☐ XSCF Console Break (SA_rccu) ☐ RCCU Console Break (SA_rccu) ☐ ALOM Console Break (SA_sunF) ☐ ILOM Panic (SA_ilomp) ☐ ILOM Reset (SA_ilomr)	
Cluster resource management facility (CRM)	Hardware devices stored in resource database		☐ Network unit	Required for using takeover of public LAN (IP) and node names
			☐ Line switching unit	Required for using line takeover function
			☐ Disk unit	Required when using GDS and GFS
Node 1	Machine model			
	Device name			
	Domain ID / domain name			
	Node name (#uname -n)			
	CF node name			
	Interconnect	IP address		
		Subnet mask		
		Device name		
	Node in CF quorum set		◆ [yes] ◇ no	Do not change the default setting.
	Use by RMS		◇ [yes] ◇ no Suffix:	
	weight			Node weight (priority if cluster partition occurs) used by SF
	RCI	Address		Only when using RCI asynchronous monitoring
	System console	IP address		RCCU, XSCF, ALOM, and ILOM data
		Subnet mask		
		User ID		
		User authority		
		Connection type	■ SSH ☐ telnet	
	Use of shared disk		■ yes ☐ no	
	scsi-initiator-id (OBP variable #eeprom command)			
	SCSI / fibre channel	Circuits (paths)		
		Device name 1		
		sd device name		
		Device name 2		
		sd device name		
	Management LAN	IP address		
		Subnet mask		
		Device name		
	Public LAN 1	IP address		
		Subnet mask		
		Device name		
	Public LAN 2	IP address		
		Subnet mask		
		Device name		

Node 2	Machine model			
	Device name			
	Domain ID / domain name			
	Node name (#uname -n)			
	CF node name			
	Interconnect	IP address		
		Subnet mask		
		Device name		
	Node in CF quorum set		◆ [yes] ◇ no	Do not change the default setting.
	Use by RMS		◇ [yes] ◇ no Suffix:	
	weight			Node weight (priority if cluster partition occurs) used by SF
	RCI	Address		Only when using RCI asynchronous monitoring
	System console	IP address		RCCU, XSCF, ALOM, and ILOM data
		Subnet mask		
		User ID		
		User authority		
		Connection type	■ SSH □ telnet	
	Use of shared disk		■ yes □ no	
	scsi-initiator-id (OBP variable #eeprom command)			
	SCSI / fibre channel	Circuits (paths)		
		Device name 1		
		sd device name		
		Device name 2		
		sd device name		
	Management LAN	IP address		
		Subnet mask		
		Device name		
	Public LAN 1	IP address		
		Subnet mask		
		Device name		
	Public LAN 2	IP address		
		Subnet mask		
		Device name		

 See
. . . .

For information on the "setup policy for survival priority if cluster partition occurs" in the Remarks column for weight, see "Survival scenarios" in "5.1.2.1.2 Using the Shutdown Configuration Wizard."
. . . .

A.7 Cluster Application Configuration Worksheets

Use the Cluster Application Configuration Worksheet to organize the items before building a cluster application.
The following worksheets are required for building a cluster application.

- A.7.1 GFS Local File System Setup Worksheet

- A.7.2 GLS Setup Worksheet

- A.8 GDS Setup Worksheets

- A.9 Cluster Application Worksheets

A.7.1 GFS Local File System Setup Worksheet

If you plan to use a GFS (Global File Services) Local File System, use the GFS Local File System Setup Worksheet to organize the input items beforehand.

Item			Setting
GFS Local File System settings		Device name	
		Mount point	
		Device name	
		Mount point	

A.7.2 GLS Setup Worksheet

If you will be using Global Link Services (GLS), use the GLS Setup Worksheet to organize the input items beforehand.

This section describes the worksheet settings for the following typical duplexing modes:

- **NIC switching mode** (logical IP takeover)

- **Fast switching mode**

- **GS/SURE linkage mode**

NIC Switching Mode (Logical IP Takeover)

	Item			Setting
GLS settings	Switching mode			◆ NIC switching mode (logical IP takeover) ◇ Fast switching mode
	Takeover virtual interface name			
		Takeover virtual IP address (or host name)		
		Subnet mask		
	Node name (1)			
		Configuration information	Virtual interface name	
			Primary physical interface name	
			Secondary physical interface name	
			Physical IP address (or host name)	
			Logical IP address (or host name)	
		Monitoring destination information	IP address (or host name) of primary monitoring destination	
			IP address (or host name) of secondary monitoring destination	

		Optional function	Standby NIC patrol	◇ Disable
				◇ Enable
				  Set to standby NIC
				  Local MAC address ( )

	Node name (2)			
		Configuration information	Virtual interface name	
			Primary physical interface name	
			Secondary physical interface name	
			Physical IP address (or host name)	
			Logical IP address (or host name)	
		Monitoring destination information	IP address (or host name) of primary monitoring destination	
			IP address (or host name) of secondary monitoring destination	

		Optional function	Standby NIC patrol	◇ Disable
				◇ Enable
				  Set to standby NIC
				  Local MAC address ( )

[Setup items]

- Switching mode

  Specify a redundant operation type for GLS; Fast switching mode, NIC (Network Interface Card) switching mode, or RIP mode.
  Add d (logical IP takeover of the NIC switching mode) to the "-m" option of the "hanetconfig create" command.

- Takeover virtual interface name

  Specify the name of the virtual interface for cluster switchover.
  The virtual interface is automatically created under the name of shaX:NN (X and NN are numbers) as a result of executing the "hanethvrsc create" command.

- Takeover virtual IP address (or host name)

  Specify the IP address (or host name) for cluster switchover.
  The same logical IP address (or host name) that was set in nodes (1) and (2) is set as a result of executing the "hanethvrsc create" command.

- Subnet mask

  Specify the subnet mask value corresponding to the takeover virtual IP address in the "/etc/inet/netmasks" file.
  This value is used for a physical IP and logical IP.
  If subnet is not used, you don't need to specify the value.

- Virtual interface name

  Specify the virtual interface name to be assigned to the physical interface to be duplexed.

Specify shaX (X is a number) to the "-n" option of the "hanetconfig create" command.
Set the same name for node (1) and node (2).

- Primary physical interface name

  Specify the name of the interface to be first activated among the duplexed physical interfaces (e.g. hme1, qfe1, etc.)
  Specify the name to the "-t" option of the "hanetconfig create" command.

- Secondary physical interface name

  Among the duplexed physical interfaces, specify the name of the interface to be activated after switchover from the primary physical interface. Specify the name to the "-t" option of the "hanetconfig create" command.

- Physical IP address (or host name)

  Specify the physical IP address (or host name) to be assigned to the physical interface to be duplexed.
  Specify the value to the "-e" option of the "hanetconfig create" command.
  Specify different values for node (1) and node (2).

- Logical IP address (or host name)

  Specify the logical IP address (or host name) to be assigned to the physical interface to be duplexed.
  Specify the value to the "-i" option of the "hanetconfig create" command. Specify the same value for node (1) and node (2).

- Primary monitoring-destination IP address (or host name)

  Specify the IP address (or host name) of the switching hub to be monitored when the primary physical interface is used.
  Specify the value to the "-p" option of the "hanetpoll create" command.

- Secondary monitoring-destination IP address (or host name)

  Specify the IP address (or host name) of the switching hub to be monitored when the secondary physical interface is used.
  Specify the value to the "-p" option of the "hanetpoll create" command.

- Standby NIC patrol

  When monitoring the standby NIC, specify p for standby patrol (automatic switchback in the event of a failure) or q for standby patrol (immediate automatic switchback) to the -m option of the "hanetconfig create" command.
  You need to specify the local MAC address to the "-a" option.
  Specify the different local MAC address values for node (1) and node (2).

Fast Switching Mode

Item				Setting
GLS settings	Switching mode			◇NIC switching mode (logical IP takeover) ◆Fast switching mode
	Takeover virtual interface name			
		Takeover virtual IP address (or host name)		
		Subnet mask		
	Node name (1)			
		Configuration information	Virtual interface name	
			Physical interface name (1)	
			Physical interface name (2)	
			Virtual IP address (or host name)	
			Subnet mask	
			Physical IP address (or host name) (1)	
			Subnet mask	
			Physical IP address (or host name) (2)	
			Subnet mask	
	Node name (2)			
		Configuration information	Virtual interface name	
			Physical interface name (1)	
			Physical interface name (2)	
			Virtual IP address (or host name)	
			Subnet mask	
			Physical IP address (or host name) (1)	
			Subnet mask	
			Physical IP address (or host name) (2)	
			Subnet mask	

[Setup items]

- Switching mode

  Specify a redundant operation type for GLS; Fast switching mode, NIC (Network Interface Card) switching mode, or RIP mode.
  Add t (fast switching mode) to the "-m" option of the "hanetconfig create" command.

- Takeover virtual interface name

  Specify the name of the virtual interface for cluster switchover.
  The virtual interface is automatically created under the name of shaX:NN (X and NN are numbers) as a result of executing the "hanethvrsc create" command.

- Takeover virtual IP address (or host name)

  Specify the IP address (or host name) for cluster switchover.
  Specify the value to the "-i" option of the "hanethvrsc create" command.

- Subnet mask

  Specify the subnet mask value corresponding to the takeover virtual IP address in the "/etc/inet/netmasks" file.
  This value should be the same as the virtual IP for the subnet mask.
  If subnet is not used, you don't need to specify the value.

- Virtual interface name

  Specify the virtual interface name to be assigned to the physical interface to be duplexed.
  Specify shaX (X is a number) to the "-n" option of the "hanetconfig create" command.
  Set the same name for node (1) and node (2).

- Physical interface name (1) and (2)

  Specify the name of the interface to be first activated among the duplexed physical interfaces (e.g. hme1, qfe1, etc.)
  Specify the name to the "-t" option of the "hanetconfig create" command.

- Virtual IP address (or host name)

  Specify the virtual IP address (or host name) to the "-i" option of the "hanetconfig create" command.
  Specify different values for node (1) and node (2).

- Subnet mask

   Specify the subnet mask value corresponding to the takeover virtual IP address in the "/etc/inet/netmasks" file.
   This value should be the same as the virtual IP for the subnet mask.
   If subnet is not used, you don't need to specify the value.

- Physical IP address (or host name) (1) and (2)

   Specify the physical interface name to be assigned to the physical interface to be duplexed.
   Create the "/etc/hostname. physical interface name" file and specify the IP address (or host name) in the file.
   This value should be different than other IP.

- Subnet mask

   Specify the subnet mask value corresponding to the physical IP address in the "/etc/inet/netmasks" file.
   If subnet is not used, you don't need to specify the value.

GS/SURE linkage mode

Item				Setting
GLS settings	Switching mode			<Y>GS/SURE linkage mode
	Takeover virtual interface name			
		Takeover virtual IP address (or host name)		
		Subnet mask		
	Node name (1)			
		Configuration information	Virtual interface name	
			Physical interface name (1)	
			Physical interface name (2)	
			IP address (or host name)(1)	
			Subnet mask	
			IP address (or host name)(2)	
			Subnet mask	
	Node name (2)			
		Configuration information	Virtual interface name	
			Physical interface name (1)	
			Physical interface name (2)	
			IP address (or host name)(1)	
			Subnet mask	
			IP address (or host name)(2)	
			Subnet mask	
	Remote host name			
		Configuration information	Virtual IP address	
			IP address (1)	
			IP address (2)	
		Monitoring destination information	Remote virtual IP address monitoring	<>Enable <>Disable
		Optional function	RIP packet is sent from remote host	<>Wait <>Not Wait
			TCP relay	Network information of relay destinationNetwork IP address or host name ( )

[Setup items]

- Switching mode

   Specify a redundant operation type for GLS; GS/SURE linkage mode.
   Add n (Physical interface setting) and c (Virtual interface setting) to the "-m" option of the "hanetconfig create" command.

- Takeover virtual interface name

  Specify the name of the virtual interface for cluster switchover.
  The virtual interface is automatically created under the name of shaX:NN (X and NN are numbers) as a result of executing the "hanethvrsc create" command.

- Takeover virtual IP address (or host name)

  Specify the IP address (or host name) for cluster switchover.
  The same logical IP address (or host name) that was set in nodes (1) and (2) is set as a result of executing the "hanethvrsc create" command.

- Subnet mask

  Specify the subnet mask value corresponding to the takeover virtual IP address in the "/etc/inet/netmasks" file.
  This value is used for a physical IP and logical IP.
  If subnet is not used, you don't need to specify the value.

- Virtual interface name

  Specify the virtual interface name to be assigned to the physical interface to be duplexed.
  Specify shaX (X is a number) to the "-n" option of the "hanetconfig create" command.
  Set the same name for node (1) and node (2).

- Primary physical interface name (1) and (2)

  Specify the name of the interface to be first activated among the duplexed physical interfaces (e.g. hme1, qfe1, etc.)
  Specify the interface name followed by "-t" option along with executing "hanetconfig create" command with "-m" option and parameter "n" to set the physical interface.

- IP address (or host name) (1) and (2)

  Specify the interface name followed by "-i" option along with executing "hanetconfig create" command with "-m" option and parameter "n" to set the physical interface.

- Subnet mask (1) and (2)

  Specify the subnet mask value corresponding to the physical IP address in the "/etc/inet/netmasks" file.
  If subnet is not used, you don't need to specify the value.

- Remote host name

  Specify the remote host name.
  Specify the remote host name to the "-n" option of the "hanetobserv create" command. (e.g. GS-1, SURE-1, etc.)

- Virtual IP address

  Specify the virtual IP address of the remote host.
  Specify the virtual IP address to the "-i" option of the "hanetobserv create" command.

- IP address (1) and (2)

  Specify the physical IP address of the remote host.
  Specify the physical IP address to the "-t" option of the "hanetobserv create" command.

- Remote virtual IP address monitoring

  When monitoring the remote virtual IP address, specify "on" to the "-m" option of the "hanetobserv create" command.
  When not monitoring the remote virtual IP address, specify "off" to the "-m" option of the "hanetobserv create" command.

- RIP packet is sent from remote host

  When sending out a report regarding node switch against the remote host, specify whether or not to wait until receiving RIP from the remote host. Specify the parameter "on" using hanetobserv create command with "-r" option to hold off for the report, or "off" for not holding off.

- TCP relay

  Specify the network information of relaying system for TCP relay. Specify client IP address, network address, and subnet mask using hanetobserv create command with "-c" option. (e.g. -c 192.168.72.1,192.168.73.0:255.255.255.0)

A.8 GDS Setup Worksheets

These worksheets serve to organize necessary input items beforehand for setting the GDS (Global Disk Services) configuration.

There are two types of GDS Setup Worksheet, "System Disk Mirror Setup" for setting up system disk mirroring and "GDS Configuration" for setting up shared disks.

- System Disk Mirror Setup Worksheet

- GDS Configuration Worksheet

See
. . . .

- Before setting up the GDS configuration, you need to know basic information from reading *"PRIMECLUSTER Global Disk Services Configuration and Administration Guide."*

- For information on the items in this worksheet, see "6.2.2 GDS Configuration Setup."
. . . .

Note
. . . .

The class name should be unique in the cluster system.
. . . .

A.8.1 System Disk Mirror Setup Worksheet

Use the System Disk Mirror Setup Worksheet to organize items before mirroring the system disk.

See
. . . .

- Before setting up the GDS configuration, you need to know basic information from reading *"PRIMECLUSTER Global Disk Services Configuration and Administration Guide."*

- For information on the items in this worksheet, see "6.2.2.1 Setting Up System Disk Mirroring." Also see "A.10.4 System Disk Mirror Setup Worksheet" in "A.10 System Design Examples."
. . . .

System Disk Mirror Setup Worksheet

Item				Setting
System disk mirror settings	Node name			
		Class name		
			Group name	
			Mount point	
			Physical disk name	
			Mirror disk name	
			Spare disk name	
	Node name			
		Class name		
			Group name	
			Mount point	
			Physical disk name	
			Mirror disk name	
			Spare disk name	

A.8.2 Shared Disk Setup Worksheet

If you are setting up shared disks, you must determine the GDS configuration beforehand.
Use the Shared Disk Setup Worksheet to organize the GDS configuration beforehand.

See

- Before setting up the GDS configuration, you need to know basic information from reading *"PRIMECLUSTER Global Disk Services Configuration and Administration Guide."*

- For information on the items in this worksheet, see "6.2.2.2 Setting Up Shared Disks." Also see "A.10.5 GDS Configuration Worksheet" in "A.10 System Design Examples."

GDS Configuration Worksheet

Item				Setting
GDS configuration	Class 1	Class name		
		Class scope (node name)	Node 1	
			Node 2	
		Spare disk 1 (*1)	SDX disk name	
			Physical disk name in node 1	
			Physical disk name in node 2	
		Single disk 1 (*2)	SDX disk name	
			Physical disk name in node 1	
			Physical disk name in node 2	
			Single volume 1	Volume name	
				Size	
			Single volume 2	Volume name	
				Size	
		Single disk 2 (*2)	SDX disk name	
			Physical disk name in node 1	
			Physical disk name in node 2	
			Single volume 1	Volume name	
				Size	
			Single volume 2	Volume name	
				Size	

		Disk 1 to be connected to group (*3)	SDX disk name		
			Physical disk name in node 1		
			Physical disk name in node 2		
		Disk 2 to be connected to group (*3)	SDX disk name		
			Physical disk name in node 1		
			Physical disk name in node 2		
		Low-order group 1 (*4)	Group name		
			Group type		
			Stripe width (*6)		
			Disk /low-order group name	Disk /low-order group 1	
				Disk /low-order group 2	
		Low-order group 2 (*4)	Group name		
			Group type		
			Stripe width (*6)		
			Disk /low-order group name	Disk /low-order group 1	
				Disk /low-order group 2	

		Highest-order group 1 (*5)	Group name		
			Group type		
			Stripe width (*6)		
			Disk /low-order group name	Disk /low-order group 1	
				Disk /low-order group 2	
			Volume 1	Volume name	
				Size	
			Volume 2	Volume name	
				Size	
		Highest-order group 2 (*5)	Group name		
			Group type		
			Stripe width (*6)		
			Disk /low-order group name	Disk /low-order group 1	
				Disk /low-order group 2	
			Volume 1	Volume name	
				Size	
			Volume 2	Volume name	
				Size	

Class 2	Class name			
	Class scope (node name)	Node 1		
		Node 2		
	Spare disk 1 (*1)	SDX disk name		
		Physical disk name in node 1		
		Physical disk name in node 2		
	Single disk 1 (*2)	SDX disk name		
		Physical disk name in node 1		
		Physical disk name in node 2		
		Single volume 1	Volume name	
			Size	
		Single volume 2	Volume name	
			Size	
	Single disk 2 (*2)	SDX disk name		
		Physical disk name in node 1		
		Physical disk name in node 2		
		Single volume 1	Volume name	
			Size	
		Single volume 2	Volume name	
			Size	
	Disk 1 to be connected to group (*3)	SDX disk name		
		Physical disk name in node 1		
		Physical disk name in node 2		
	Disk 2 to be connected to group (*3)	SDX disk name		
		Physical disk name in node 1		
		Physical disk name in node 2		

		Low-order group 1 (*4)	Group name	
			Group type		
			Stripe width (*6)	
			Disk /low-order group name	Disk /low-order group 1	
				Disk /low-order group 2	
		Low-order group 2 (*4)	Group name	
			Group type		
			Stripe width (*6)	
			Disk /low-order group name	Disk /low-order group 1	
				Disk /low-order group 2	
		Highest-order group 1 (*5)	Group name	
			Group type	
			Stripe width (*6)	
			Disk /low-order group name	Disk /low-order group 1	
				Disk /low-order group 2	
			Volume 1	Volume name	
				Size	
			Volume 2	Volume name	
				Size	

		Highest-order group 2 (*5)	Group name	
			Group type	
			Stripe width (*6)	
			Disk /low-order group name	Disk /low-order group 1	
				Disk /low-order group 2	
			Volume 1	Volume name	
				Size	
			Volume 2	Volume name	
				Size	

(*1) This item can be defined only if mirroring is performed.
(*2) This item must be defined only if a single volume is to be created.
(*3) This item must be defined only if mirroring, striping, or concatenation is performed.
(*4) This item must be defined only if the groups are to be arranged into a hierarchy.
(*5) This item must be defined only if mirroring, striping, or concatenation is performed.
(*6) This item must be defined only if the group type is stripe.

A.9 Cluster Application Worksheets

The Cluster Application Worksheets are necessary for setting up cluster applications and resources.

The following worksheets are available:

Resource configuration:

- Cmdline Resource Worksheet

- Fsystem Resource Worksheet

- Gds (Global Disk Services) Resource Worksheet

- Gls (Global Link Services) Resource Worksheet

- Takeover Network Resource Worksheet

- Procedure Resource Worksheet

- Process Monitoring Resource Worksheet

- Line Switching Unit Resource Worksheet

Cluster application configuration:

- Cluster Application Setup Worksheet

Exclusive relationship configuration for cluster applications:

- Cluster Application Exclusion Setup Worksheet

Global Settings in Configuration Setup:

- Configuration Global Settings Setup Worksheet

See

For information on the items in this worksheet, see "6.6 Setting Up Cluster Applications."

A.9.1 Cmdline Resource Worksheet

The Cmdline Resource Setup Worksheet is used for configuring a Cmdline resource.

See

For information on the items in this worksheet, see "6.6.1.1 Creating Cmdline Resources."

Item			Setting value
Cmdline resource	Active node candidate	SysNode	
	Resource name		
	Resource operation	Attribute	InParallel	◇Yes ◇[No]
			NeedAll	◇Yes ◇[No]
		Association	SubApplication	
	Script to be controlled	Object	Start script	
			Stop script	
			Check script	
		Flag	NULLDETECTOR	◇Yes ◇[No]
			ALLEXITCODES	◇Yes ◇[No]
			LIEOFFLINE	◇Yes ◇[No]
			CLUSTEREXCLUSIVE	◇Yes ◇[No]
			AUTORECOVER	◇Yes ◇[No]
			MONITORONLY	◇Yes ◇[No]
			STANDBYCAPABLE	◇Yes ◇[No]
			REALTIME	◇Yes ◇[No]
			TIMEOUT	(Range: 45 to 31535999) [300]

Note: Values enclosed in brackets ([ ]) are default values.

A.9.2 Fsystem Resource Worksheet

The Fsystem Resource Worksheet is used for configuring an Fsystem resource.

 See
. . . .

For information on the setup items in this worksheet, see "6.6.1.2 Creating Fsystem Resources."
. . . .

Item			Setting value
Fsystem resource	Active node candidate	SysNode	
	Resource name		
	Resource operation	Attribute	NeedAll	◇[Yes]　◇No
			Timeout	(Range: 45 to 3600) [180]
	Mount point to be controlled 1	Object	Mount point	
		Flag	AUTORECOVER	◇Yes　◇[No]
			SHARE	◇Yes　◇[No]
			NFSLOCKFAILOVER	◇Yes　◇[No]
	Mount point to be controlled 2	Object	Mount point	
		Flag	AUTORECOVER	◇Yes　◇[No]
			SHARE	◇Yes　◇[No]
			NFSLOCKFAILOVER	◇Yes　◇[No]
	Mount point to be controlled 3	Object	Mount point	
		Flag	AUTORECOVER	◇Yes　◇[No]
			SHARE	◇Yes　◇[No]
			NFSLOCKFAILOVER	◇Yes　◇[No]

Note: Values enclosed in brackets ([ ]) are default values.

📒 Note

Before you can set the NFSLOCKFAILOVER to Yes, you need to select Yes for NFS Lock Failover and then specify the Lock information directory name for the Global setting in the configuration.

For information on the Global Settings in the Configuration, see "A.9.11 Configuration Global Settings Worksheet."

A.9.3　Gds Resource Worksheet

The Gds Resource Worksheet is used for configuring a Gds (Global Disk Services) resource.

📑 See

For information on the setup items in this worksheet, see "6.6.1.3 Creating Gds Resources."

Item			Setting value
Gds resource	Active node candidate	SysNode	
	Resource name		

	Resource operation	Attribute	Use	Exclusive use
				◇[Enable]
				  Hot Standby operation
				  ◇Enable
				  ◇[Disable]
				◇Disable

			AutoRecover	◇Yes ◇[No]
			Timeout	(Range: 5 to 3600) [1800]
	Disk class to be controlled 1	Object	Disk Class	
		Flag	MONITORONLY	◇Yes ◇[No]
	Disk class to be controlled 2	Object	Disk Class	
		Flag	MONITORONLY	◇Yes ◇[No]
	Disk class to be controlled 3	Object	Disk Class	
		Flag	MONITORONLY	◇Yes ◇[No]

Note: Values enclosed in brackets ([ ]) are default values.

A.9.4 Gls Resource Worksheet

The Gls Resource Worksheet is used for configuring a Gls (Global Link Services) resource.

See
....

For information on the setup items in this worksheet, see "6.6.1.4 Creating Gls Resources."
....

Item			Setting value
Gls resource	Active node candidate	SysNode	
	Resource name		
	Resource operation	Attribute	Timeout	(Range: 5 to 300) [60]
		Association	SubApplication	
	Takeover IP address to be controlled 1	Object	Takeover IP address(IPv4)	
			Takeover IP address(IPv6)	
		Flag	AUTORECOVER	◇Yes ◇[No]
	Takeover IP address to be controlled 2	Object	Takeover IP address(IPv4)	
			Takeover IP address(IPv6)	
		Flag	AUTORECOVER	◇Yes ◇[No]
	Takeover IP address to be controlled 3	Object	Takeover IP address(IPv4)	
			Takeover IP address(IPv6)	
		Flag	AUTORECOVER	◇Yes ◇[No]

Note: Values enclosed in brackets ([ ]) are defaul values.

A.9.5 Takeover Network Resource Worksheet

The Takeover Network Resource Worksheet is used for configuring a takeover network resource.

 See
. . . .

For information on the items in this worksheet, see "6.6.1.5 Creating Takeover Network Resources."
. . . .

	Item		Setting value
Takeover IP address resource	Active node candidate	SysNode	
	Resource name		
	Takeover network type		◇[IP address takeover] ◇Node name takeover + IP address takeover
	Network interfaces	Node 1 ( )	
		Node 2 ( )	
		Node 3 ( )	
		Node 4 ( )	
	Resource operation	Attribute	NeedAll	◇[Yes]　◇No
			Timeout	(Range: 45 to 3600) [60]
			PingHost	
		Association	SubApplication	
	Takeover IP address to be controlled	Object	Host name	
			IP address	
			Net mask	
		Flag	BASE	◇[virtual]　◇base
			AUTORECOVER	◇Yes　◇[No]
			PingHost	

Note: Values enclosed in brackets ([ ]) are default values.

A.9.6　Procedure Resource Worksheet

The Procedure Resource Worksheet is used for configuring a procedure resource.

 See

For information on the items in this worksheet, see "6.6.1.6 Creating Procedure Resources."

Item			Setting value
Procedure resource	Active node candicate	SysNode	
	Resource name		
	Resource operation	Attribute	SCRIPTTIMEOUT	(Range: 45 to 31535999) [1800]
		Association	SubApplication	
	Procedure resource to be controlled	Procedure class	◇Application ◇BasicApplication ◇SystemState2 ◇SystemState3 ◇( )
		Procedure resource	

Note: Values enclosed in brackets ([ ]) are default values.

A.9.7   Process Monitoring Resource Worksheet

The Process Monitoring Resource Worksheet is used for configuring a process monitoring resource.

 See
••
For information on the items in this worksheet, see "6.6.1.7 Creating Process Monitoring Resources."
••

Item			Setting value
Process resource	Active node candidate SysNode		
	Resource name		
	Resource operation	Attribute	InParallel	<>Yes <>[No]
			NeedAll	<>Yes <>[No]
		Association	SubApplication	
	Command pathto be controlled	Object	Start command path	
			Stop command path	
			Monitored process is in different process group from parent process	If "Yes," check "Process is a daemon."
		Attribute	Restart count for monitored process	(Range: 0 to 99) [3]
			Restart interval for monitored process	(Range: 0 to 3600) [3]

Note: Values enclosed in brackets ([ ]) are default values.

A.9.8 Line Switching Unit Resource Worksheet

The Line Switching Unit Resource Worksheet is used for configuring a line switching unit resource.

🔖 See
....
For information on the items in this worksheet, see "6.6.1.8 Creating Line Switching Unit Resources."
....

Item			Setting value
SH_SWLine resource	Active node candidate	SysNode	
	Resource name		
	Control target	Line switching unit resource	

Note: Values enclosed in brackets ([ ]) are default values.

A.9.9 Cluster Application Worksheet

The Cluster Application Worksheet is used for configuring a cluster application.

	Item			Setting value	Remarks (value specified with GUI/CUI)
Cluster application settings	Configuration name			config	Fixed
	Cluster application name				
	Operating node	SysNode (Priority 1: High)			
		SysNode (Priority 2)			
		SysNode (Priority 3)			
		SysNode (Priority 4: Low)			
	Attribute	Start	Operating node at startup	OnlinePriority	<>[According to cluster application	[0]
					<>Operating node when RMS stopped	1
			Automatic startup	AutoStartUp	<>[Manual]	[No]
					<>Automatic	Yes
		Switchover	Switchover operation	AutoSwitchOver	<>[Manual]	[No]
					<>Automatic (more than one of the following can be selected)	
					[ ]At node failure	HostFailure
					[ ]At resource failure	ResourceFailure
					[ ]When RMS stops	ShutDown
		Standby	Trigger to be Standby	StandbyTransitions	<>[Manual]	[NONE]
					<>Automatic (more than one of the following can be selected)	
					[ ]At startup	StartUp
					[ ]At switchover	SwitchRequest
					[ ]When fault is cleared	ClearFaultRequest
		Operation in the event of a failure	Operation when RMS fails to control 2	HaltFlag	<>[Do not forcibly stop node]	[No]
					<>Forcibly stop node	Yes
			Handling of failed node	PersistentFault	<>[Use if fault is not detected when RMS is restarted]	[0]
					<>Use after operator clears fault	1
			Operation if cluster interconnect fails	ShutdownPriority	<>[NONE]	[NONE]
					<>Priority ( )	Priority (0 to 20) * Higher number indicates higher priority. * See setup policy for survival priority if cluster partition occurs.
	Configuration	Resource				Specified resource name
		Patrol diagnosis			<>Required	Check
					<>[Not required]	[Do not check]

Note: Values enclosed in brackets ([ ]) are default values.

A.9.10 Cluster Application Exclusion Worksheet

The Cluster Application Exclusion Worksheet is used for configuring exclusion relationships between cluster applications.

Item		Setting value	
Dependency relationship settings for cluster applications	Exclusion group 1	Group name	Exclusive
		Cluster application name 1	(Priority:)
		Cluster application name 2	(Priority:)
		Cluster application name 3	(Priority:)
		Cluster application name 4	(Priority:)
		Cluster application name 5	(Priority:)
	Exclusion group 2	Group name	Exclusive
		Cluster application name 1	(Priority:)
		Cluster application name 2	(Priority:)
		Cluster application name 3	(Priority:)
		Cluster application name 4	(Priority:)
		Cluster application name 5	(Priority:)
	Exclusion group 3	Group name	Exclusive
		Cluster application name 1	(Priority:)
		Cluster application name 2	(Priority:)
		Cluster application name 3	(Priority:)
		Cluster application name 4	(Priority:)
		Cluster application name 5	(Priority:)
	Exclusion group 4	Group name	Exclusive
		Cluster application name 1	(Priority:)
		Cluster application name 2	(Priority:)
		Cluster application name 3	(Priority:)
		Cluster application name 4	(Priority:)
		Cluster application name 5	(Priority:)

A.9.11 Configuration Global Settings Worksheet

This worksheet is used for configuring the Global Settings in the Configuration.

 See
....

For information on the setup items in this worksheet, see "6.6.4 Editing global settings in Configuration."
....

Item				Setting value
Global Settings in the Configuration	PreCheck Timeout			〈Range: 5 to 2147483647〉 [300]
	First Available Detector			〈Range: 0 to127〉 [0]
	Last Available Detector			〈Range: 0 to127〉 [127]
	NFS Lock Failover			◇Yes ◇[No]
		Lock information directory		*1)Only if Yes is selected above [ ]
	Detector Details	Monitoring Intervals	hvdet_execbin Used for ops, rtp	〈Range: 5 to 2147483647〉 [10]
			hvdet_ckhost Used for ops, rtp.	〈Range: 5 to 2147483647〉 [10]
			hvdet_ddm Used for Ddm (invalid under Solaris OE).	〈Range: 5 to 2147483647〉 [19]
			hvdet_glbassrt Used for Used for ops.	〈Range: 5 to 2147483647〉 [10]
			hvdet_gmount Used for Fsystem.	〈Range: 5 to 2147483647〉 [10]
			hvdet_icmp Used for Ipaddress.	〈Range: 5 to 2147483647〉 [10]
			hvdet_locassrt Used for userApplication.	〈Range: 5 to 2147483647〉 [10]
			hvdet_lvm Used for Lvm (invalid under Solaris OE).	〈Range: 5 to 2147483647〉 [18]
			hvdet_nfs Used for Fsystem.	〈Range: 5 to 2147483647〉 [10]
			hvdet_rcfs Rcfs.	〈Range: 5 to 2147483647〉 [9]
			hvdet_rcvm Rcvm.	〈Range: 5 to 2147483647〉 [33]
			hvdet_read Used for Rawdisk.	〈Range: 5 to 2147483647〉 [10]
			hvdet_srdf Used for Srdf.	〈Range: 5 to 2147483647〉 [60]
			hvdet_stopclnt Used for Fsystem.	〈Range: 5 to 2147483647〉 [10]
			hvdet_system Used for Cmdline, Fsystem, userApplication.	〈Range: 5 to 2147483647〉 [10]
			hvdet_vxvm Used for Vxvm.	〈Range: 5 to 2147483647〉 [30]
			ForeignDetectors Used for foreign-code.	〈Range: 5 to 2147483647〉 [30]
		Log level (MemoryLevel)		〈Range: 0 to 2147483647〉 [1]
		Detector logging		〈Range: 0 to 2147483647〉 [0]

*1) Values enclosed in brackets ([ ]) are default values.
Note: If the settings for items other than NFS Lock Failover need to be changed, please contact Fujitsu systems engineer (SE). However, no changes may be necessary for the normal operation.

A.10 System Design Examples

This section shows design examples of the PRIMECLUSTER system. Refer to these examples when designing the PRIMECLUSTER system.

System configuration

Figure A.1 System configuration example
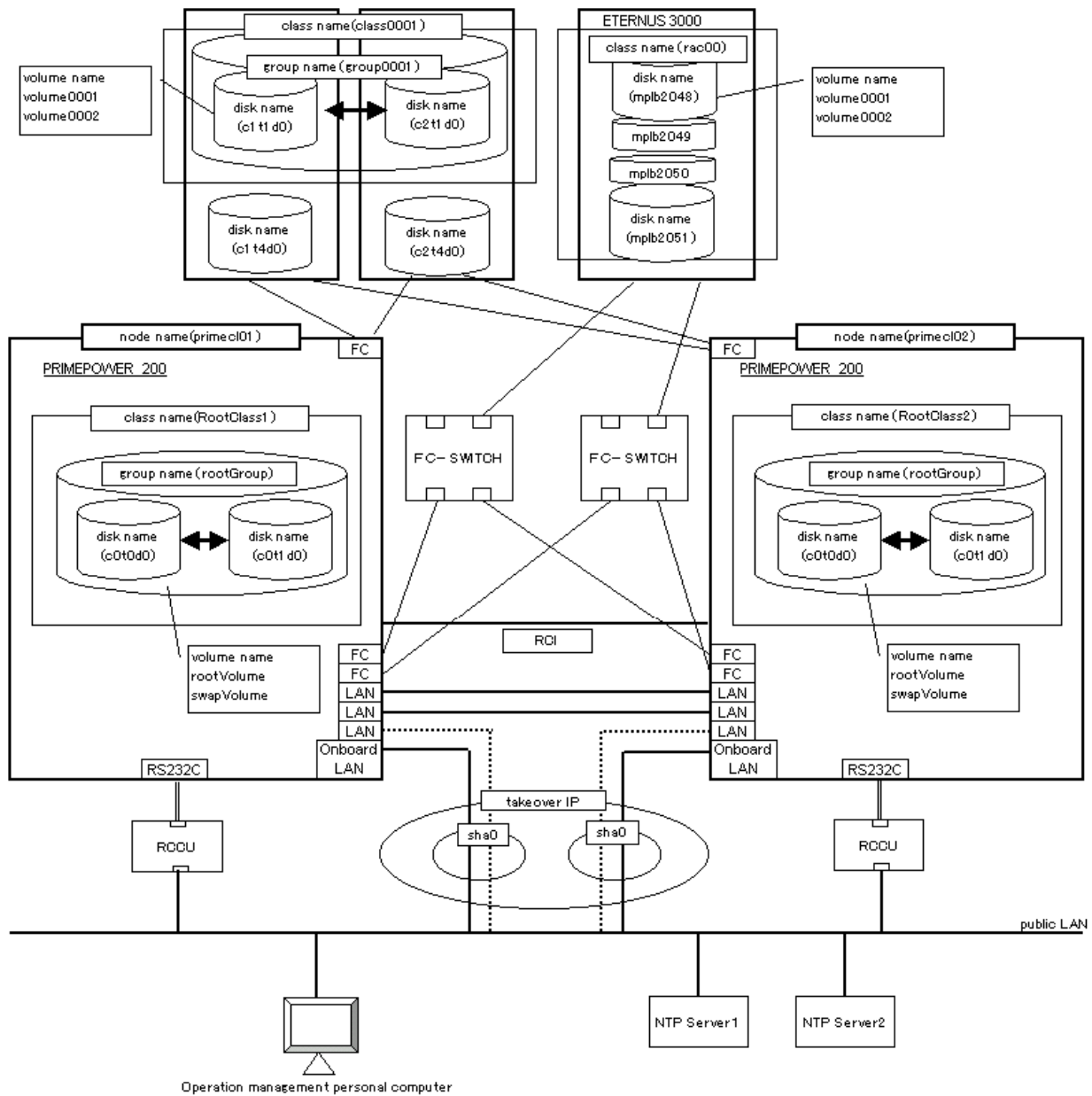
Note: The NTP server should be set to one of the following operation mode;

- Maintains time that is reliable for special hardware devices such as time units and GPS units.

- Uses another server that maintains time that is reliable for special hardware devices such as time units and GPS units, and uses NTP for time synchronization.

Used hardware

- PRIMEPOWER 200 x 2

- Expansion file unit x 2

- Remote console unit (RCCU) x 2

- Fibre channel switch (SN200) x 2

- Fast Ethernet card x 6

- Fibre channel card x 6

- Disk array unit (ETERNUS3000 model 300 (FC)) x 1

- Operation management personal computer x 1

Used software

- Solaris 10

- Enhanced Support Facility 2.0

- PRIMECLUSTER 4.2A00

- Multipath Disk Control 2.0

- Multipath Control Load Balance Option 2.1

- Fujitsu PCI Fibre Channel 2.2

- Netscape Communicator 4.7x (installed in the operation management personal computer)

- Java(TM) 2 Runtime Environment Standard Edition v1.4.2_08 (installed in the operation management personal computer)

System disk settings

Node name	Physical disk name	Slice number	Slice name	Size (MB)
node:primecl01	c0t0d0	0	/	15374
		1	swap	2048
		2	backup	17423
node:primecl02	c0t0d0	0	/	15374
		1	swap	2048
		2	backup	17423

Note

If GDS is used for system disk mirroring, one slice number in the system disk and at least 32 megabytes of free area (or swap area) are used. You need to consider the resources used by GDS when designing the system.

Shared disk array settings

Device name	Number of LU	Number of RAID groups	RAID level	Size
ETERNUS3000 model 300(FC)	4	1	5	248.2 GB
Expansion file unit (FC) x 2	-	-	-	18.2 GB

Note

To use the disk array units on UNIX, you need to format them by executing the "format (1M)" command.
When formatting the disk array units as part of the cluster system configuration work, you need to estimate the formatting time.

A.10.1 Cluster Configuration Worksheet

Shown below is an example of creating the Cluster Configuration Worksheet.

	Item			Setting	Remarks
Shared disk unit: Type 1	Machine model			ETERNUS 3000 Model 300	
	Device name			ETERNUS3000#003	
	IP address			10.33.66.215	
	Subnet mask			255.255.255.0	
	Interface			Fibre channel	
	LUN number			0 - 0F	
	Ports			CM0CA2P0, CM1CA2P0	
	Connection paths			2	
	FC-SW	Use of fibre channel switching hub		■ yes □ no	
		Machine model		ETERNUS SN200 Model 210	
		Device name		ETERNUSSN200#003	
		IP address		10.33.66.203	
		Subnet mask		255.255.255.0	
	File system selection (type of file system used in cluster)			UFS	
Shared disk unit: Type 2	Machine model			ETERNUS 4000 Model 300	
	Device name			ETERNUS4000#004	
	IP address			10.33.66.216	
	Subnet mask			255.255.255.0	
	Interface			Fibre channel	
	LUN number			0 - 0F	
	Ports			CM0CA0P0, CM1CA0P0	
	Connection paths			2	
	FC-SW	Use of fibre channel switching hub		■ yes □ no	
		Machine model		ETERNUS SN200 Model 210	
		Device name		ETERNUSSN200#004	
		IP address		10.33.66.204	
		Subnet mask		255.255.255.0	
	File system selection (type of file system used in cluster)			GFS shared file system	
NTP	Operation mode			□ NTP server (integrated in cluster) ■ NTP client	
	Protocol (only when broadcast is specified)				
	Server	Primary	Host name	pclntp1	
			IP address	10.124.95.11	
			Subnet mask	255.255.255.0	
		Secondary	Host name	pclntp2	
			IP address	10.21.8.3	
			Subnet mask	255.255.255.0	
Web-Based Admin View	Operation mode			□ 1 ■ 2 □ 3 layers	
	Server	Primary	Host name	primecl01	
			IP address	10.34.214.181	
			Subnet mask	255.255.255.0	
			URL	http://10.34.214.181:8081/Plugin.html	
		Secondary	Host name	primecl02	
			IP address	10.34.214.182	
			Subnet mask	255.255.255.0	
			URL	http://10.34.214.182:8081/Plugin.html	
	User ID	wvroot group		pclwv	
		clroot group		pclcl	
		cladmin group		pcladm	
		clmon group		pclmon	
		sdxroot group		pclsdx	

Cluster Foundation (CF)	Cluster name		SLV002	
	Cluster nodes		2	
	Cluster node names		primecl01,primecl02	
	Subnets		1	
	Subnet IP (network address)		192.168.1.0	
	Subnet mask		255.255.255.0	
	Use of CF remote services		■ cfcp (CF file copy) ■ cfsh (CF remote command execution)	
	Use by RMS		◆ [yes]　◇ no	
Shutdown facility (SF)	Shutdown agent		■ RCI Panic (SA_pprcip) ■ RCI Reset (SA_pprcir) ■ XSCF Panic (SA_xscfp) ■ XSCF Reset (SA_xscfr) ■ XSCF Console Break (SA_rccu) □ RCCU Console Break (SA_rccu) □ ALOM Console Break (SA_sunF) □ ILOM Panic (SA_ilomp) □ ILOM Reset (SA_ilomr)	
Cluster resource management facility (CRM)	Hardware devices stored in resource database		■ Network unit	Required for using takeover of public LAN (IP) and node names
			■ Line switching unit	Required for using line takeover function
			■ Disk unit	Required when using GDS and GFS
Node 1	Machine model		SPARC Enterprise M8000	
	Device name		DBSV#001	
	Domain ID / domain name		domain0	
	Node name (#uname -n)		primecl01	
	CF node name		primecl01	
	Interconnect	IP address	192.168.1.1	
		Subnet mask	255.255.255.0	
		Device name	bge1 , bge2	
	Node in CF quorum set		◆ [yes]　◇ no	Do not change the default setting.
	Use by RMS		◆ [yes] ◇ no Suffix: primecl01RMS	
	weight		2	Node weight (priority if cluster partition occurs) used by SF
	RCI	Address	000101ff	Only when using RCI asynchronous monitoring
	System console	IP address	10.33.66.71	RCCU, XSCF, ALOM, and ILOM data
		Subnet mask	255.255.255.0	
		User ID	cl01sa	
		User authority	platadm	
		Connection type	■ SSH　□ telnet	
	Use of shared disk		■ yes　□ no	
	scsi-initiator-id (OBP variable #eeprom command)		7	
	SCSI / fibre channel	Circuits (paths)	2	
		Device name 1	fjpfca#0	
		sd device name	sd11 - sd42	
		Device name 2	fjpfca#1	
		sd device name	sd43 - sd74	
	Management LAN	IP address	10.33.66.91	
		Subnet mask	255.255.255.0	
		Device name	bge0	
	Public LAN 1	IP address	10.34.214.181	
		Subnet mask	255.255.255.0	
		Device name	fjgi0	
	Public LAN 2	IP address	10.34.215.21	
		Subnet mask	255.255.255.0	
		Device name	fjgi1	

		Machine model		SPARC Enterprise M8000	
Node 2		Device name		DBSV#002	
		Domain ID / domain name		domain0	
		Node name (#uname-n)		primecl02	
		CF node name		primecl02	
	Interconnect	IP address		192.168.1.2	
		Subnet mask		255.255.255.0	
		Device name		bge1,bge2	
	Node in CF quorum set			◆ [yes]　◇ no	Do not change the default setting.
	Use by RMS			◆ [yes] ◇ no Suffix: primecl02RMS	
	weight			2	Node weight (priority if cluster partition occurs) used by SF
	RCI	Address		000102ff	Only when using RCI asynchronous monitoring
	System console	IP address		10.33.66.72	RCCU, XSCF, ALOM, and ILOM data
		Subnet mask		255.255.255.0	
		User ID		cl02sa	
		User authority		platadm	
		Connection type		■ SSH　□ telnet	
	Use of shared disk			■ yes　□ no	
	scsi-initiator-id (OBP variable #eeprom command)			8	
	SCSI / fibre channel	Circuits (paths)		2	
		Device name 1		fjpfca#0	
		sd device name		sd11 - sd42	
		Device name 2		fjpfca#1	
		sd device name		sd43 - sd74	
	Management LAN	IP address		10.33.66.92	
		Subnet mask		255.255.255.0	
		Device name		bge0	
	Public LAN 1	IP address		10.34.214.182	
		Subnet mask		255.255.255.0	
		Device name		fjgi0	
	Public LAN 2	IP address		10.34.215.22	
		Subnet mask		255.255.255.0	
		Device name		fjgi1	

In each node, specify /etc/inet/ntp.conf as shown below.

```
server 10.34.214.100
server 10.34.214.101
server 127.127.1.0
fudge 127.127.1.0 stratum 9

enable auth monitor

disable pll
driftfile /etc/ntp/ntp.drift
statsdir /var/ntp/ntpstats/
filegen peerstats file peerstats type day enable
filegen loopstats file loopstats type day enable
filegen clockstats file clockstats type day enable
```

For detailed instructions on specifying /etc/inet/ntp/conf, see the online manual by executing the "xntpd(1M)" command.

A.10.2 GFS Local File System Setup Worksheet

Shown below is an example of creating the GFS Local File System Setup Worksheet.

	Item		Setting
GFS Local File System setup		Device name	class0001:volume0001
		Mount point	/disk1
		Device name	class0001:volume0002
		Mount point	/disk2

A.10.3 GLS Setup Worksheet

Shown below are examples of creating the GLS Setup Worksheet.

NIC Switching Mode (Logical IP Takeover)

	Item			Setting
GLS setup	Switching mode			◆NIC switching mode (logical IP takeover) ◇Fast switching mode
	Takeover virtual interface name			sha0:65
		Takeover virtual IP address (or host name)		10.34.214.185
		Subnet mask		255.255.255.0
	Node name (1)			primecl01
		Configuration information	Virtual interface name	sha0
			Primary physical interface name	hme0
			Secondary physical interface name	hme3
			Physical IP address (or host name)	10.34.214.181
			Logical IP address (or host name)	10.34.214.185
		Monitoring destination information	Primary monitoring destination IP address (or host name)	10.34.214.188
			Secondary monitoring destination IP address (or host name)	10.34.214.189
		Optional function	Backup NIC patrol	◇Disable ◆Enable Set to backup NIC. Local MAC address (02:00:00:00:00:01)
	Node name (2)			primecl02
		Configuration information	Virtual interface name	sha0
			Primary physical interface name	hme0
			Secondary physical interface name	hme3
			Physical IP address (or host name)	10.34.214.182
			Logical IP address (or host name)	10.34.214.185
		Monitoring destination information	Primary monitoring-destination IP address (or host name)	10.34.214.188
			Secondary monitoring-destination IP address (or host name)	10.34.214.189
		Optional function	Backup NIC patrol	◇Disable ◆Enable Set to backup NIC Local MAC address (02:00:00:00:00:02)

Fast Switching Mode

Item				Setting
GLS setup	Switching mode			◇NIC switching mode (logical IP takeover) ◆Fast switching mode
	Takeover virtual interface name			sha0:65
		Takeover virtual IP address (or host name)		10.34.214.185
		Subnet mask		255.255.255.0
	Node name (1)			primecl01
		Configuration information	Virtual interface name	sha0
			Physical interface name (1)	hme0
			Physical interface name (2)	hme3
			Virtual IP address (or host name)	10.34.214.181
			Subnet mask	255.255.255.0
			Physical IP address (or host name) (1)	10.34.215.181
			Subnet mask	255.255.255.0
			Physical IP address (or host name) (2)	10.34.216.181
			Subnet mask	255.255.255.0
	Node name (2)			primecl02
		Configuration information	Virtual interface name	sha0
			Physical interface name (1)	hme0
			Physical interface name (2)	hme3
			Virtual IP address (or host name)	10.34.214.182
			Subnet mask	255.255.255.0
			Physical IP address (or host name) (1)	10.34.215.182
			Subnet mask	255.255.255.0
			Physical IP address (or host name) (2)	10.34.216.182
			Subnet mask	255.255.255.0

A.10.4 System Disk Mirror Setup Worksheet

Shown below is an example of creating the System Disk Mirror Setup Worksheet.

Item				Setting
System disk mirroring setup	Node name			primecl01
		Class name		Rootclass1
			Group name	rootGroup
				Mount point	/, /swap, /usr, /opt
				Physical disk name	c0t0d0
				Mirror disk name	c0t1d0
				Spare disk name	-
	Node name			primecl02
		Class name		Rootclass2
			Group name	rootGroup
				Mount point	/, /swap, /usr, /opt
				Physical disk name	c0t0d0
				Mirror disk name	c0t1d0
				Spare disk name	-

A.10.5 GDS Configuration Worksheet

Shown below are examples of creating the GDS Configuration Worksheet.

System configuration example

For a system configuration example, set up the worksheet as shown below.

Item				Setting
GDS configuration	Class 1	Class name		rac00
		Class scope (node name)	Node 1	primecl01
			Node 2	primecl02
		Single disk 1	SDX disk name	diskmplb0001
			Physical disk name in node 1	mplb2048
			Physical disk name in node 2	mplb2048
			Single volume 1	Volume name	volume0001
				Size	1024 MB
			Single volume 2	Volume name	volume0002
				Size	1024 MB
		Single disk 2	SDX disk name	diskmplb0002
			Physical disk name in node 1	mplb2049
			Physical disk name in node 2	mplb2049
		Single disk 3	SDX disk name	diskmplb0003
			Physical disk name in node 1	mplb2050
			Physical disk name in node 2	mplb2050
		Single disk 4	SDX disk name	diskmplb0004
			Physical disk name in node 1	mplb2051
			Physical disk name in node 2	mplb2051
	Class 2	Class name		class0001
		Class scope (node name)	Node 1	primecl01
			Node 2	primecl02
		Spare disk 2	SDX disk name	-
			Physical disk name in node 1	-
			Physical disk name in node 2	-
		Disk 1 connected to group	SDX disk name	disk010001
			Physical disk name in node 1	c1t1d0
			Physical disk name in node 2	c1t1d0
		Disk 1 connected to group	SDX disk name	disk020001
			Physical disk name in node 1	c2t1d0
			Physical disk name in node 2	c2t1d0
		Highest-order group 1	Group name	group0001
			Group type	mirror
			Stripe width (*6)	-
			Disk/low-order group name	Disk/low-order group 1	disk010001
				Disk/low-order group 2	disk010002
			Volume 1	Volume name	volume0001
				Size	1024 MB
			Volume 2	Volume name	volume0002
				Size	1024 MB

Examples of creating the GDS Configuration Worksheet are provided below.

Example 1. Managing a disk as a single disk:

In this example, single volumes Volume1 and Volume2 are created in physical disk c3t4d5.

Item			Setting
Class 1	Class name		Class1
	Class scope (node name)	Node 1	Node1
		Node 2	Node2
	Single disk 1	SDX disk name	Disk1
		Physical disk name in node 1	c3t4d5
		Physical disk name in node 2	c2t3d4
		Single volume 1 — Volume name	Volume1
		Single volume 1 — Size	1024 blocks
		Single volume 2 — Volume name	Volume2
		Single volume 2 — Size	1024 blocks

Example 2. If a disk is not mirrored but may be changed to a mirror configuration later:

In this example, mirror group Group1, which consists only of physical disk c3t4d5, is created, and single volumes Volume1 and Volume2 are created in Group1.

Since Group1 consists of only one disk, Volume1 and Volume2 are not mirrored. If another disk is added to Group1, then Volume1 and Volume2 can be changed to a mirror configuration.

Item			Setting
Class 1	Class name		Class1
	Class scope (node name)	Node 1	Node1
		Node 2	Node2
	Disk 1 connected to group	SDX disk name	Disk1
		Physical disk name in node 1	c3t4d5
		Physical disk name in node 2	c2t3d4
	Highest-order group 1	Group name	Group1
		Group type	mirror
		Stripe width	–
		Disk/low-order group name — Disk/low-order group 1	Disk1
		Volume 1 — Volume name	Volume1
		Volume 1 — Size	1024 blocks
		Volume 2 — Volume name	Volume2
		Volume 2 — Size	1024 blocks

Example 3. If disk mirroring is to be executed:

In this example, mirror group Group1, which consists of physical disks c3t4d5 and c4t5d6, is created, and mirrored volumes Volume1 and Volume2 are created in Group1.

Item			Setting
Class 1	Class name		Class1
	Class scope (node name)	Node 1	Node1
		Node 2	Node2
	Spare disk 1	SDX disk name	Disk3
		Physical disk name in node 1	c1t2d3
		Physical disk name in node 2	c2t2d2
	Disk 1 connected to group	SDX disk name	Disk1
		Physical disk name in node 1	c3t4d5
		Physical disk name in node 2	c2t3d4
	Disk 2 connected to group	SDX disk name	Disk2
		Physical disk name in node 1	c4t5d6
		Physical disk name in node 2	c3t3d3
	Highest-order group 1	Group name	Group1
		Group type	mirror
		Stripe width	–
		Disk/low-order group name	Disk/low-order group 1	Disk1
			Disk/low-order group 2	Disk2
		Volume 1	Volume name	Volume1
			Size	1024 blocks
		Volume 2	Volume name	Volume2
			Size	1024 blocks

Example 4. If disk striping is to be executed:

In this example, striping group Group1, which consists of physical disks c3t4d5 and c4t5d6, is created, and striped volumes Volume1 and Volume2 are created in Group1.

Item				Setting
Class 1	Class name			Class1
	Class scope (node name)	Node 1		Node1
		Node 2		Node2
	Disk 1 connected to group	SDX disk name		Disk1
		Physical disk name in node 1		c3t4d5
		Physical disk name in node 2		c2t3d4
	Disk 2 connected to group	SDX disk name		Disk2
		Physical disk name in node 1		c4t5d6
		Physical disk name in node 2		c3t3d3
	Highest-rder group 1	Group name		Group1
		Group type		stripe
		Stripe width		32
		Disk/low-order group name	Disk/low-order group 1	Disk1
			Disk/low-order group 2	Disk2
		Volume 1	Volume name	Volume1
			Size	1024 blocks
		Volume 2	Volume name	Volume2
			Size	1024 blocks

Example 5. If disk concatenation is to be executed:

In this example, concatenation group Group1, which consists of physical disks c3t4d5 and c4t5d6, is connected, and volumes Volume1 and Volume2 are created in Group1.

Item				Setting
Class 1	Class name			Class1
	Class scope (node name)	Node 1		Node1
		Node 2		Node2
	Disk 1 connected to group	SDX disk name		Disk1
		Physical disk name in node 1		c3t4d5
		Physical disk name in node 2		c2t3d4
	Disk 2 connected to group	SDX disk name		Disk2
		Physical disk name in node 1		c4t5d6
		Physical disk name in node 2		c3t3d3
	Highest-order group 1	Group name		Group1
		Group type		concat
		Stripe width		-
		Disk/low-order group name	Disk/low-order group 1	Disk1
			Disk/low-order group 2	Disk2
		Volume 1	Volume name	Volume1
			Size	1024 blocks
		Volume 2	Volume name	Volume2
			Size	1024 blocks

Example 6. If disk striping and mirroring are to be executed:

In this example, mirror group Group1, which consists of striping group Group2 and striping group Group3, is created, and volumes Volume1 and Volume2 are created in Group1. Striping group Group2 consists of physical disks c3t4d5 and c4t5d6, and striping group Group3 consists of physical disks c3t4d6 and c4t5d7.

Item			Setting
Class 1	Class name		Class1
	Class scope (node name)	Node 1	Node1
		Node 2	Node2
	Spare disk 1	SDX disk name	Disk5
		Physical disk name in node 1	c1t2d3
		Physical disk name in node 2	c2t2d2
	Disk 1 connected to group	SDX disk name	Disk1
		Physical disk name in node 1	c3t4d5
		Physical disk name in node 2	c2t3d4
	Disk 2 connected to group	SDX disk name	Disk2
		Physical disk name in node 1	c4t5d6
		Physical disk name in node 2	c3t3d3
	Disk 3 connected to group	SDX disk name	Disk3
		Physical disk name in node 1	c3t4d6
		Physical disk name in node 2	c2t3d5
	Disk 4 connected to group	SDX disk name	Disk4
		Physical disk name in node 1	c4t5d7
		Physical disk name in node 2	c3t3d4
	Low-order group 1	Group name	Group2
		Group type	stripe
		Stripe width	32
		Disk/low-order group name / Disk/low-order group 1	Disk1
		Disk/low-order group name / Disk/low-order group 2	Disk2
	Low-order group 2	Group name	Group3
		Group type	stripe
		Stripe width	32
		Disk/low-order group name / Disk/low-order group 1	Disk3
		Disk/low-order group name / Disk/low-order group 2	Disk4
	Highest-order group 1	Group name	Group1
		Group type	mirror
		Stripe width	–
		Disk/low-order group name / Disk/low-order group 1	Group2
		Disk/low-order group name / Disk/low-order group 2	Group3
		Volume 1 / Volume name	Volume1
		Volume 1 / Size	1024 blocks
		Volume 2 / Volume name	Volume2
		Volume 2 / Size	1024 blocks

Appendix B Manual Pages

This appendix provides online manual page lists for CCBR, CF, CIP, operator intervention, PAS, the cluster resource management facility, RMS, shutdown facility (SF), SIS, tracing failed resource, Web-Based Admin View, procedure resource, process monitoring feature, and the RMS wizards.

To view a manual page, enter the following command:

$ **man** *man_page_name*

Note:

To view these manual pages, you must set the MANPATH environment variable so that /etc/opt/FJSVcluster/man is included.

To print a hard copy of a manual page, enter the following command:

% **man** *man_page_name* |**col-b** |**lpr**

B.1 CCBR

System administrator

Command	Function
cfbackup(1M)	Creates a backup copy of all the cluster configuration information related to the PRIMECLUSTER nodes.
cfrestore(1M)	Restores the cluster configuration information stored on the PRIMECLUSTER node.

B.2 CF

System administrator

Command	Function
cfconfig(1M)	Configures or unconfigures a node for a PRIMECLUSTER cluster.
cfset(1M)	Applies or modifies /etc/default/cluster.config entries into the CF module.
cftool(1M)	Prints the node communications state of a node or the cluster.

B.3 CIP

System administrator

Command	Function
cipconfig(1M)	Starts or stops CIP 2.0.
ciptool(1M)	Retrieves CIP information about local and remote nodes in the cluster.

File format

File	Format
cip.cf(4)	CIP configuration file format

B.4 Operator Intervention

System administrator

Command	Function
clreply(1M)	Responds to an operator intervention request message.

B.5 PAS

System administrator

Command	Function
mipcstat(1M)	MIPC statistics
clmstat(1M)	CLM statistics

B.6 Cluster Resource Management Facility

System administrator

Command	Function
clautoconfig(1M)	Executes automatic resource registration.
clbackuprdb(1M)	Saves the resource database.
clexec(1M)	Executes a remote command.
cldeldevice(1M)	Deletes a resource that was registered during automatic resource registration.
cldelrsc(1M)	Deletes a network interface card resource that was created with the automatic resource registration facility.
clinitscript(1M)	Reports the connection confirmation results for shared disk units.
clrestorerdb(1M)	Restores the resource database.
clsetacfparam(1M)	Checks the connections of shared disk units and sets up the operation for automatic resource registration.
clsetparam(1M)	Displays and changes the operation environment for a resource database, failed resource identification, or operation intervention request.
clsetrsc(1M)	Changes the takeover IP address.
clsetup(1M)	Sets up the resource database.
clspconfig(1M)	Sets up the operation of patrol diagnosis.
clsptl(1M)	Executes patrol diagnosis.
clstartrsc(1M)	Activates a resource (GDS only).
clstoprsc(1M)	Deactivates a resource (GDS only).
clsyncfile(1M)	Distributes a file between cluster nodes.

User command

🅿 Point
・・・

There is also a clgettree command in the Web-Based System Administration tool WSA.

・・

Command	Function
clgetrsc(1)	Acquires takeover IP address.
clgettree(1)	Outputs tree information for the resource database.

B.7 RMS

System administrator

Command	Function
hvassert(1M)	Asserts (tests for) an RMS resource state.
hvcm(1M)	Starts the RMS configuration monitor.
hvconfig(1M)	Displays or saves the RMS configuration file.
hvdisp(1M)	Displays RMS resource information.
hvdist(1M)	Distributes RMS configuration files.
hvdump(1M)	Collects debugging information about RMS.
hvgdmake(1M)	Compiles an RMS custom detector.
hvlogclean(1M)	Cleans the RMS log files.
hvlogcontrol(1M)	Controls the size of the log disk.
hvrclev(1M)	Changes the default startup execution level of RMS.
hvreset(1M)	Initializes the graph of the RMS user jobs on the nodes in the RMS configuration.
hvshut(1M)	Shuts down RMS.
hvswitch(1M)	Switches control of an RMS user application resource to another host.
hvthrottle(1M)	Prevents multiple RMS scripts from running simultaneously.
hvutil(1M)	Manipulates the availability of an RMS resource.

File format

File	Format
hvenv.local(4)	RMS local environment configuration file

B.8 Shutdown Facility (SF)

System administrator

Command	Function
sdtool(1M)	Interface tool for shutdown daemon
rcsd(1M)	Shutdown daemon for shutdown manager
clrcimonctl(1M)	Displays, starts, stops, and restarts the operation of the RCI asynchronous monitoring daemon.
clrccumonctl(1M)	Displays, starts, stops, and restarts the operation of the console asynchronous monitoring daemon.
clrccusetup(1M)	Registers, changes, deletes, or displays console information

File format

File	Format
rcsd.cfg(4M)	Configuration file for shutdown daemon
SA_wtinps.cfg(4M)	Configuration file for WTI NPS shutdown agent
SA_sunF.cfg(4M)	Configuration file for sunF system controller shutdown agent
SA_sspint.cfg(4M)	Configuration file for Sun E10000 shutdown agent

File	Format
SA_scon.cfg(4M)	Configuration file for SCON shutdown agent
SA_rps.cfg(4M)	Configuration file for shutdown agent that uses remote power switch (RPS)
SA_rccu.cfg(4M)	Configuration file for RCCU shutdown agent

B.9 SIS

System administrator

Command	Function
dtcpadmin(1)	Starts the SIS administration utility.
dtcpd(1)	Starts the SIS daemon for configuring VIPs.
dtcpdbg(1)	Displays SIS debugging information.
dtcpstat(1)	Displays state information on SIS.

B.10 Tracing Failed Resource

System administrator

Command	Function
cldispfaultrsc(1M)	Outputs a list of the current failed resources.

B.11 Web-Based Admin View

System administrator

Command	Function
fjsvwvbs(1M)	Stops Web-Based Admin View.
fjsvwvcnf(1M)	Starts, stops, or restarts the web server for Web-Based Admin View.
wgcnfclient(1M)	Sets and displays the RMS configuration name.
wvCntl(1M)	Starts, stops, or gets debugging information for Web-Based Admin View.
wvGetparam(1M)	Displays the Web-Based Admin View environment variables.
wvSetparam(1M)	Sets the Web-Based Admin View environment variables.
wvstat(1M)	Displays the operating state of Web-Based Admin View.

B.12 Procedure Resource

System administrator

Command	Function
claddprocrsc(1M)	Registers an application resource that uses a state transition procedure.
cldelproc(1M)	Deletes a state transition procedure.
cldelprocrsc(1M)	Deletes an application resource that uses state transition procedure.
clgetproc(1M)	Gets a state transition procedure.
clsetproc(1M)	Registers a state transition procedure.

Command	Function
clsetprocrsc(1M)	Changes the registered information of an application resource that uses a state transition procedure.

User command

Command	Function
cldspproc(1)	Outputs information on the resource that uses the state transition procedure.

B.13 Process Monitoring Feature

System administrator

Command	Function
clmonproc(1M)	Requests the process monitoring facility to monitor a process.

B.14 RMS Wizards

RMS Wizards and RMS Application Wizard

The RMS Wizard manual will be saved in the following directory when the SMAWRhvdo package is installed.

/usr/opt/reliant/htdocs.solaris/wizards.en

Appendix C Troubleshooting

This appendix explains how to collect troubleshooting information if an error occurs in the PRIMECLUSTER system.

C.1 Collecting Troubleshooting Information

If an error occurs in the PRIMECLUSTER system, collect the information required for the error investigation from all nodes that construct the cluster and the cluster management servers. Then, contact your customer support representative.

1. PRIMECLUSTER investigation information

    - Use fjsnap to collect information.

    - Retrieve the system dump.

       If a hardware error, an operating system error, a panic, or a login failure occurs, collect the followings:

         - System dump

         - SCF* dump

            * SCF: System Control Facility

    - Collect the Java Console on the clients.

       See "Appendix B.2.2 Java console" in the "*PRIMECLUSTER Web-Based Admin View Operation Guide*."

    - Collect a hard copy of the client screens.

       See "Appendix B.2.3 Screen hard copy" in the "*PRIMECLUSTER Web-Based Admin View Operation Guide*."

    - If a client is using RC2000, collect RC2000 information for the client.

       See "D.2 Collecting Troubleshooting Information" in the "*RC2000 User's Guide*."

2. Investigation information for the failed application

3. Error reproduction procedure description if the error can be reproduced

Information

. . . .

 - When reporting a problem, collect the information required for an error investigation. If you do not provide information for problem checking and error reproduction execution, it may take a long time to reproduce and diagnose the problem or it may become impossible to do so.

 - Collect investigation material promptly from all nodes of the PRIMECLUSTER system. Necessary information may become lost if a long time elapses after the error occurs. This applies especially to information collected by fjsnap.

 - If a node was forcibly terminated, execute sync in OBP mode, and collect a system dump.

. . . .

C.1.1 Executing the fjsnap Command

The "fjsnap" command is a system information tool of the Solaris, and it comes with the FJSVsnap Expanded Support Function Package. This tool enables you to collect information for troubleshooting if a failure occurs in the PRIMECLUSTER system.

The procedure for executing the fjsnap command is described below.

1. Log in with system administrator authority.

2. Execute the "fjsnap" command.

    /opt/FJSVsnap/bin/fjsnap -h *output*

    or

    /opt/FJSVsnap/bin/fjsnap -a *output*

- If -a is specified, the amount of data becomes large because all detailed information is collected. If -h is specified, only cluster control information is collected.

- Specify a special file name or an output file name (ex: /dev/rmt/0) to output information. The specified name is the file to which collected information is output when the "fjsnap" command is executed.

See
. . . .
For details on the "fjsnap" command, see the "README" file, which is the README file included in the "FJSVsnap" package.
. . . .

Information
. . . .
Execution timings for the fjsnap command

- For problems that occur during operation, for example, if an error message is output, execute the "fjsnap" command immediately after the problem occurs.

- If the "fjsnap" command cannot be executed because the system hangs, collect a system dump. Then start the system in single user mode, and execute the "fjsnap" command.
To collect the system dump, press the REQUEST switch or input the abort key sequence (for example, Break signal) to forcibly stop the node to OBP mode, and then execute "sync." For detailed instructions on forcibly stopping the node to OBP mode, see the PRIMEPOWER user's guides and the "Solaris System Management" manual of the Solaris.

- After an error occurs, if a node restarts automatically (the node could not be started in single-user mode) or if the node is mistakenly started in multi-user mode, execute the "fjsnap" command.

- If investigation information cannot be collected because the "fjsnap" command results in an error or the "fjsnap" command does not return, then collect a system dump.
. . . .

C.1.2 System dump

If a system dump is collected when a node is in a panic, obtain the system dump as investigation information. The system dump is written into the dump file during node startup. The default storage directory is /var/crash/node_name.

C.1.3 SCF Dump

Collect a SCF dump if conditions like those listed below occur:

- Message 7003 or 7004 is output.

- The SCF driver outputs a message.

See
. . . .
For information on the messages that the SCF driver outputs, see the "*Enhanced Support Facility User's Guide*."
. . . .

Collecting the SCF dump

The SCF RAS monitoring daemon that received the error notification from SCF stores the SCF dump to the "/var/opt/FJSVhwr/scf.dump" file. Collect the entire directory of this file according to the following operation.

```
# cd /var/opt
# tar cf /tmp/scf.dump.tar ./FJSVhwr
```

Regarding the machine type to which SCS is connected, you can collect troubleshooting information by using the getscfdump command. For details on the getscfdump command, see the "System Console Software User's Guide."

C.2 Troubleshooting

C.2.1 GUI in General

This section explains how to take corrective actions for problems that may occur while you are using the GUI.

If you find no relevant descriptions in this chapter, see also "B.1 Corrective action" in the "*PRIMECLUSTER Web-Based Admin Operation Guide.*"

Symptom 1: When disk class creation, deletion, or name change is executed on the GDS screen, the message, "An error occurred in the log monitoring facility in node XXXX. Monitoring stops. After confirming that the node is active, update to the latest information." or "0007 The connection to the management server is disconnected. Attempt to connect to the active management server?" is displayed. Web-Based Admin View is disconnected.

Corrective action

This error might be due to a JavaVM related problem. Close the browser, and then display the screen again.

If you are using the Cluster Admin screen, close the screen first, and create or delete a disk class or change the disk class name to avoid this error. If the same error occurs frequently even after closing the Cluster Admin screen, contact your Fujitsu engineer.

Symptom 2: If you operate the userApplication Configuration Wizard while the registration of cluster application and resource configuration is in progress (while 0805 or 0813 dialog box is displayed), the message, "0880 A non-classified error occurred." appears.

Corrective action

If this phenomenon occurs, respond to the message, exit the userApplication Configuration Wizard and then restart it. After restarting, re-execute the operation that was being made before the 0880 error message was displayed.

C.3 Detecting a Failed Resource

If a failure occurs in a resource, you can specify the resource by referring to the following:

- The message displayed if a failure occurs in the resource

- Resource Fault History

- Fault Resource List

Note

To use the history function of the failed resource, the resource database must be set up correctly. Also, the "AutoStartUp" and "PersistentFault" attributes of userApplication must be set to yes(1).

For information on the resource database settings, see the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide.*"

To use the detection function of the failed resources, you must enable an operator intervention request. For information on the use of the operator intervention request, see "5.4 Setting Up Fault Resource Identification and Operator Intervention Request."

The operator intervention function and the failed resource history function are both dependent on the "clwatchlogd" daemon. This daemon can be started automatically with the "rc" script in multi-user mode. The "clwatchlogd" daemon uses the "RELIANT_LOG_PATH" environment variable of RMS. The value of this variable is set when the "rc" script starts up for the first time.

When this value is changed, you need to send the "SIGHUP" signal to clwatchlogd. When clwatchlogd receives this signal, clwatchlogd acquires the latest value of RELIANT_LOG_PATH. After completing the above processing, start RMS.

This manual is installed in the /etc/opt/FJSVcluster/man directory.

Before executing the "man (1)" command, add this directory to the beginning of MANPATH. Usually, a directory name is added to the line beginning with "setenv MANPATH" within the ".cshrc" file or the line beginning with "export MANPATH" within the ".profile" file.

C.3.1 Failed Resource Message

If a resource failure or recovery is detected, a message screen pops up as the Cluster Admin popup screen. An example of failed resource message screen is shown below:
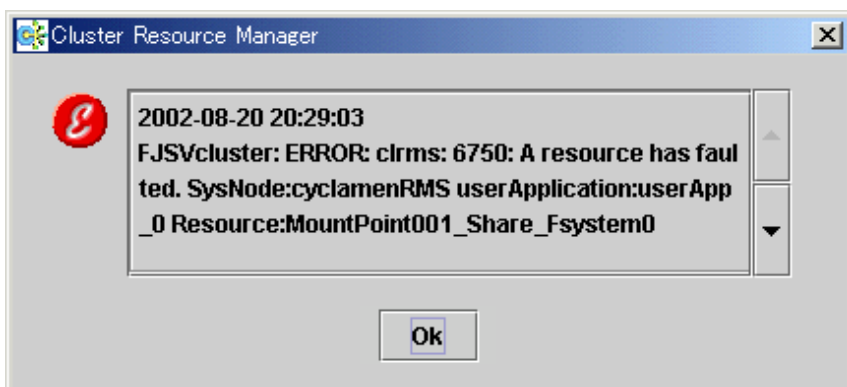
📒 **Note**

. . . .

If a message frame title says "Cluster resource management facility," refer to the following:

- PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide

    - 12.10 Resource Database messages

- Appendix D in this manual

    - D.7.2 Failed Resource and Operator Intervention Messages (GUI)

    - Display format 1 in D.1 Searching for a Message

. . . .

The severity icon is defined as follows:

Icon	Meaning
	Notice
	Warning
	Error
	Other

Respond to the failed resource message screen as follows:

Procedure

1. Click on the *OK* button to respond to the message.

2. Click the up arrow mark or down arrow mark to go to the previous or next message. Then, a message appears to remind you that you have not yet entered a response or confirmed the displayed message.

If you subsequently enter a response, the message is cleared and the next message appears. If the next message does not appear and the message prior to that for which a response was entered is still available, the previous message will appear. If there is any message for which confirmation or a response has not yet been entered, the message screen closes. For information on the message contents, refer to "Failed Resource and Operator Intervention Messages" and for information on how to display previous messages, refer to "Resource Fault History."

Note

. . . .

If you close Web-Based Admin View or Cluster Admin after this message is displayed, a fault resource message with the same contents will not be displayed. Therefore, you are recommended to confirm the message contents if a fault resource message is displayed for the first time. After you have closed the message, refer to the fault history on the "Resource Fault History" screen. For information on the message display language, refer to "4.3.3.3 Setting the Web-Based Admin View Language."

If the Cluster Admin screen is not displayed on the client PC when the fault resource message is displayed, the message is transmitted only to the client to which the management server was first connected.

Each management server administers its fault resource messages. If you change the management server after confirming the message, the same message will be displayed again. To delete these messages, select *Cluster Admin* by using the GUI of *Web-Based Admin View* after closing *Cluster Admin*, and then open *Cluster Admin* again.

. . . .
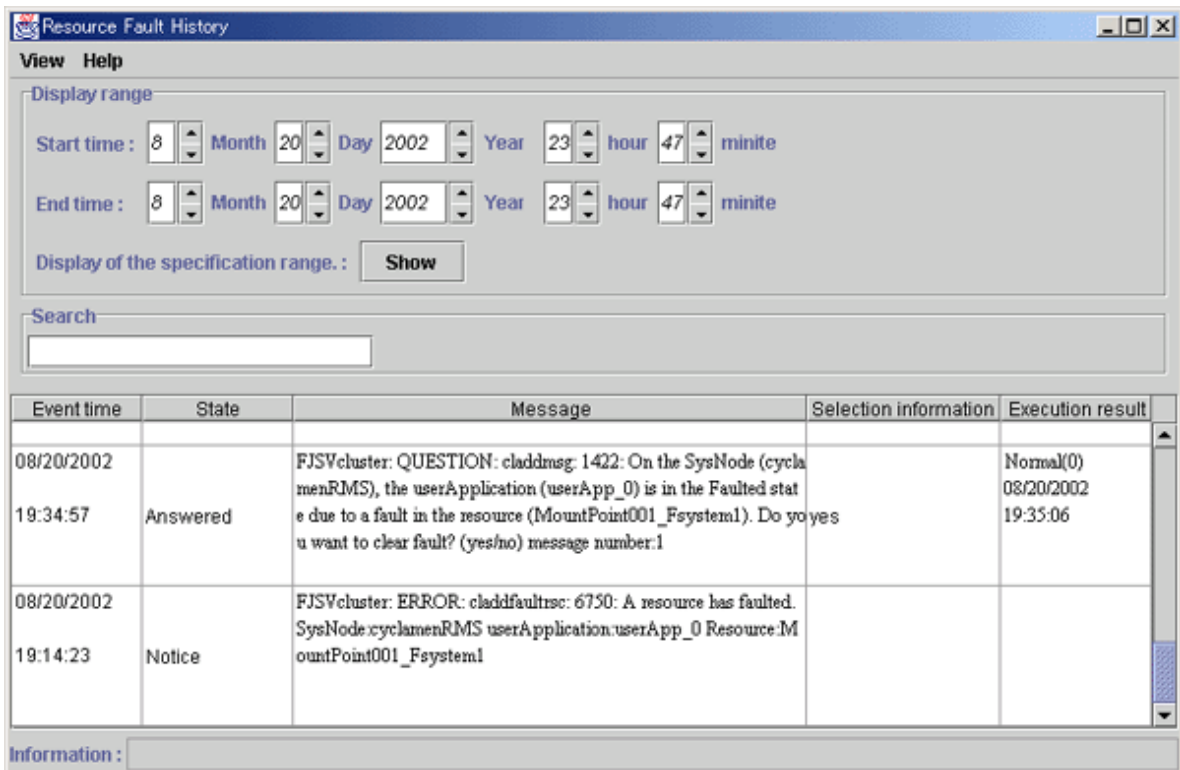
C.3.2 Resource Fault History

Display the "Resource Fault History" screen, in which the resource fault history is displayed, in the following procedure.

Procedure

1. Open the "Web-Based Admin View" screen and then select *Global Cluster Services*.

2. Choose *Resource Fault History*.
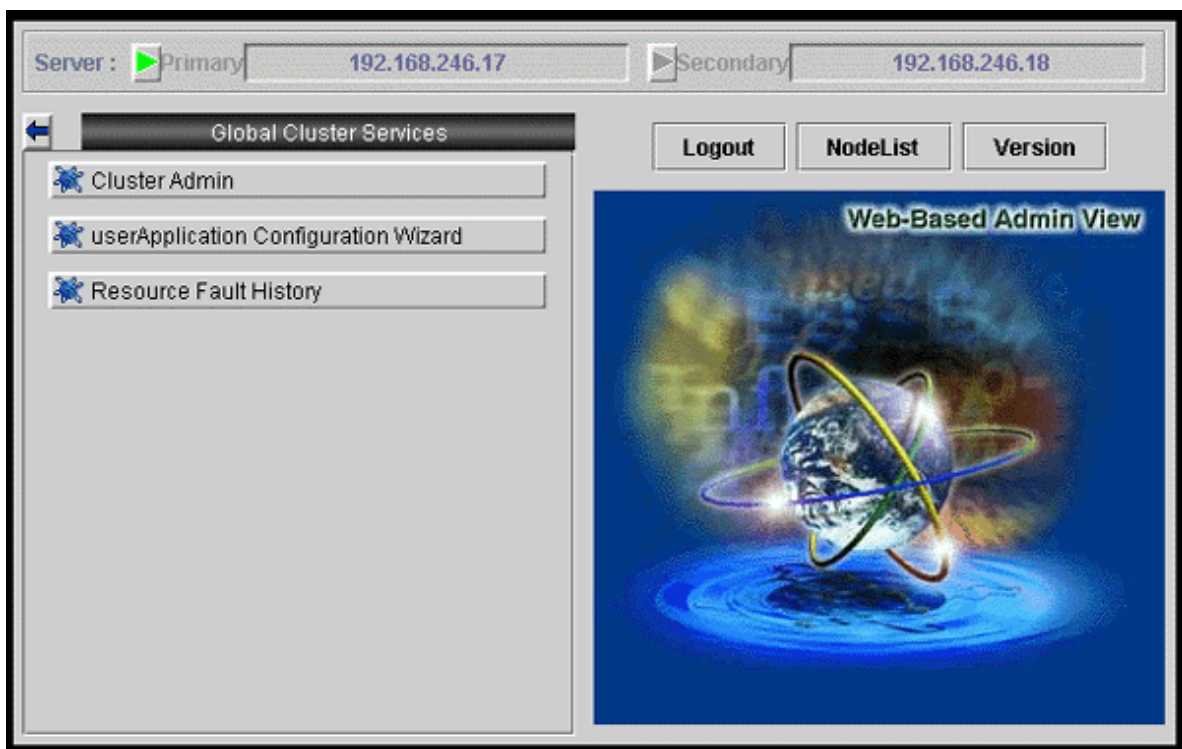
The "Resource Fault History" will be displayed.

📘 **Note**

The "Resource Fault History" cannot be displayed automatically. To display the latest history information, select *View -> Update* menu.

Menu of the fault resource list screen

The "Resource Fault History" screen contains the following menu items:

Menu	Function
View -> Update latest information	The duration is initialized to the present time and date. A maximum of 100 of the latest history resources are displayed.
View -> Fault Resource List	A list of resources in which failures are present is displayed (see "C.3.3 Fault Resource List").
View -> Exit	The "Resource Fault History" screen is cleared.
Help -> Help	The GUI help screen is displayed.

Setting the range of time

A fault resource history listing can be displayed by specifying a date and time.

- *Start time* - A start time is set up.

- *End time* - An end time is set up.

If you click the *View* button after setting up the required values, a maximum of 100 of the most recently failed resources within the specifiable range can be displayed.

Search with a keyword

The fault resource history list can be narrowed by specifying "*Keyword*".

If a duration is set, the history of up to the 100 latest failed resources that satisfy both conditions can be displayed.

How to read the list

The following information is displayed on the "Resource Fault History" screen.

- Event time - The time at which the RMS detected a resource failure is displayed.

- State - One of the following statuses is indicated.

    - Responded - The operator has already responded the message.

    - Not responded - The operator has not responded to the message for which a response is required.

    - Responding - The operator is currently responding to the message.

    - Confirm - Notification message for which no response is required.

- Message - The message is displayed.

- Selection information - Operator intervention message information from the client that is connected to the management server is displayed. If the message is canceled or if a response to the message is entered by executing the "clreply(1M)" command, nothing will be displayed.

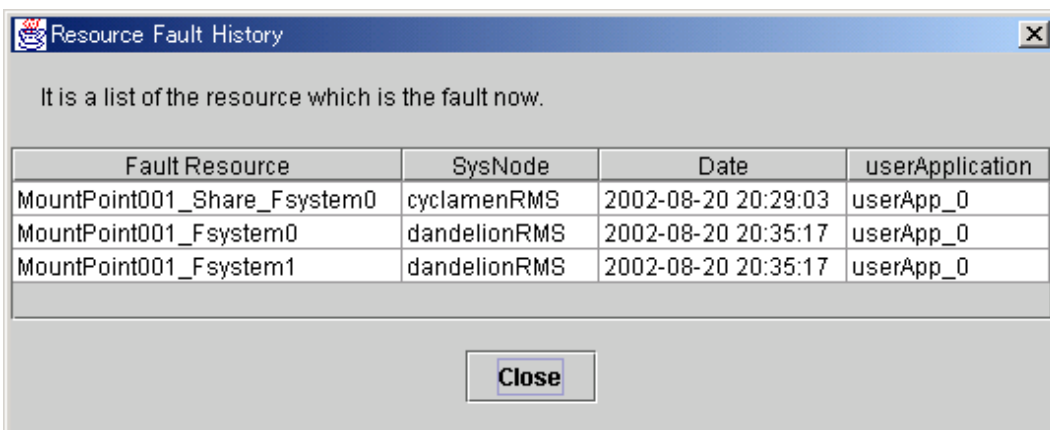- Execution result - The result and time of the response processing are displayed.

Information field

The information related to error detection during the acquisition or read-in of the history files is displayed. The following items will be displayed:

- Processing - History data is being collected from the management server.

- An error is included in the specified duration. - The specified duration is incorrect. Correct it and then click the *View* button.

- Part of the data acquisition failed. - Parts of the history files could be damaged. This will not disrupt ongoing operation, but the corrupted data will not be displayed.

C.3.3 Fault Resource List

If you select *View -> Fault Resource List* on the "Resource Fault History" screen, the fault resource list is displayed as follows:

The following information is displayed on the fault resource list screen:

- Fault Resource - The name of the fault resource is displayed.

- SysNode - The SysNode to which the fault resource belongs is displayed.

- Date - The date and time at which the RMS detected a resource failure are displayed.

- userApplication - The name of userApplication to which the failed resource belongs is displayed.

If you click *Close*, the failed resource list screen is closed.

The list of resources that are currently in the failed state can be displayed by executing the "cldispfaultrsc(1M)" command.

Appendix D Messages

This appendix explains how to search for the reference relating to the following messages, as well as the corresponding corrective actions.

For other than GUI messages, first refer to "D.1 Searching for a Message."

- Information messages

- Warning messages

- Error messages

- Operator intervention messages

- Wgcnfclient command messages

- GUI messages

    - Messages related to the CRM view

    - Failed resource and operator intervention messages (GUI)

    - userApplication Configuration Wizard GUI messages

D.1 Searching for a Message

This section explains the procedure for searching for a message reference.

The PRIMECLUSTER display formats are classified into six types. Determine the reference according to the format of the displayed message.

For details on those messages displayed by the GUI, see "D.7 GUI Messages."

Display format 1

```
FJSVcluster:severity:program:message-number:message
```

severity:

Indicates the message severity level.

There are five message severity levels. Stop (HALT), Information (INFO), Warning (WARNING), Error (ERROR), and Question (QUESTION).

program:

Indicates the name of the program that outputs this message.

message-number:

Indicates the message number.

message:

Indicates the message text.

Search procedure

After confirming the *severity* and *message-number*, determine the reference by referring to the table below.

severity	message-number	Reference
Stop (HALT)	0100, 0101, 0102	CF Configuration and Administration Guide 12.10.1 HALT messages (Resource Database messages)
Information (INFO)	2206, 2207, 2620, 2621, 2622, 2700, 2701, 2914, 2927, 3050, 3051, 3052, 3053	D.2 Information Messages

severity	message-number	Reference
	2100, 2200, 2201, 2202, 2203, 2204, 2205, 3200, 3201, 3202, 3203, 3204, 3205, 3206	CF Configuration and Administration Guide 12.10.2 INFORMATION messages (Resource Database messages)
	3040, 3041, 3042, 3043, 3044, 3045, 3046, 3070, 3071	CF Configuration and Administration Guide 12.12.1 INFORMATION message (Monitoring Agent messages)
Warning (WARNING)	5100	D.3 Warning Message
	4250, 5200	CF Configuration and Administration Guide 12.10.3 WARNING messages (Resource Database messages)
	5001	CF Configuration and Administration Guide 12.12.2 WARNING message (Monitoring Agent messages)
Error (ERROR)	????, 6000, 6001, 6002, 6003, 6004, 6006, 6007, 6008, 6009, 6010, 6208, 6602, 6615, 6616, 6617, 6618, 6619, 6621, 6623, 6624, 6651, 6655, 6657, 6658, 6659, 6662, 6680, 6690, 6691, 6692, 6750, 6751, 6752, 6753, 6754, 6755, 6780, 6781, 6782, 6807, 6817, 6836, 7050, 7051, 7052, 7053, 7054, 7055, 7056, 7101, 7102, 7103, 7104, 7105, 7106, 7108, 7109, 7110, 7111, 7112, 7113, 7116, 7117, 7119, 7121, 7122, 7125, 7126, 7130, 7131, 7132	D.4 Error Messages
	6005, 6021, 6025, 6200, 6201, 6202, 6203, 6204, 6206, 6207, 6209, 6210, 6211, 6212, 6213, 6214, 6215, 6216, 6217, 6218, 6219, 6220, 6221, 6222, 6223, 6226, 6250, 6300, 6301, 6302, 6303, 6600, 6601, 6602, 6603, 6604, 6606, 6607, 6608, 6611, 6614, 6653, 6661, 6665, 6668, 6675, 6900, 6901, 6902, 6903, 6904, 6905, 6906, 6907, 6910, 6911, 7500, 7501, 7502, 7503, 7504, 7505, 7506, 7507, 7508, 7509, 7510, 7511, 7512, 7513, 7514, 7515, 7516, 7517, 7518, 7519, 7520, 7521, 7522, 7535, 7536, 7537, 7538, 7539, 7540, 7541, 7542, 7543, 7545, 7546	CF Configuration and Administration Guide 12.10.4 ERROR messages (Resource Database messages)
	6007, 7003, 7004, 7018, 7019, 7026, 7027, 7030, 7031, 7033, 7034, 7035, 7040, 7042, 7200, 7201, 7202, 7203, 7204	CF Configuration and Administration Guide 12.12.3 ERROR message (Monitoring Agent messages)
Question (QUESTION)	1421, 1422, 1423	D.5 Operator Intervention Messages

Display format 2

```
wgcnfclient: ERROR: message-number: message
```

message-number:

    Indicates the message number.

message:

    Indicates the message text.

Search procedure

    Confirm the *message-number*, and then determine the reference destination by referring to the table below:

message-number	Reference
0000, 0001, 0002, 0003, 0004, 0005, 0006, 0007, 0009	D.6 Wgcnfclient Command Messages

Display format 3

```
command-name:message
```

command-name:

    Indicates the command name.

message:

    Indicates the message (describing the phenomenon, message number, reason, etc.).

Search procedure

    Messages of display format 3 are described in the *CF Configuration and Administration Guide*. Confirm *command-name*, and then determine the reference by referring to the following table:

Function	*command-name*	Reference in the CF Configuration and Administration Guide
Cluster node configuration	cfconfig	12.1 cfconfig message
CIP2.0 start/stop	cipconfig	12.2 cipconfig message
Outputting the node communication state	cftool	12.3 cftool message
Outputting and manipulating quorum information	rcqconfig	12.4 rcqconfig message
	rcqquery	12.5 rcqquery message
CF execution	CF	12.6 CF runtime message

Display format 4

```
(SMAWsf,number1,number2):message
```

number1,number2:

    Indicates the message number.

message:

    Indicates the message.

Search procedure

    Messages of display format 4 are described in the *CF Configuration and Administration Guide*.

    Confirm *number1* and *number2*, and then determine the reference by referring to the table below.

Function	(*number1-number2*)	Reference in the CF Configuration and Administration Guide
Shutdown facility	(10,2)-(10,12) (10,15)-(10,17), (10,19), (10,20) (10,30), (10,31), (10,34), (10,36), (10,38), (10,101) (30,2)-(30,10), (30,12)-(30,17) (50,3)-(50,4), (50,6), (50,9), (50,11)-(50,14), (50,20)-(50,22)	12.11 Shutdown Facility

Display format 5

`(err-code,message-number):message`

err-code:

Indicates the error code.

message-number:

Indicates the message number.

message:

Indicates the message.

Search procedure

Messages of display format 5 are described in the "*PRIMECLUSTER Reliant Monitor Services (RMS) Troubleshooting Guide.*" Confirm *err-code* and *message-number*, and then determine the reference by referring to the table below.

Function	*err-code*	*message-number*	Reference in the RMS Troubleshooting Guide
switchlog	ADC	1-5, 15, 17, 18, 20, 25, 27, 30, 31, 32, 33, 34, 35, 37, 38, 39, 41, 43- 49, 57, 58, 59, 60, 61, 62, 63, 68, 70	3 Non-fatal error messages
	ADM	3- 9, 11-60, 62, 63, 67, 70, 75-79, 81-91, 95, 96, 98, 99, 100, 106, 107, 118, 125	
	BAS	2, 3, 6, 9, 14, 15, 16, 17, 18, 19, 23-27, 29, 30, 31, 32, 36, 37, 38, 39, 40,41, 42	
	BM	8, 13-21, 23-26, 29, 42, 46, 54, 59, 68, 71-75, 80, 81, 90, 92, 94, 96-99, 101, 103, 105-111, 114	
	CML	11, 12, 17, 18, 19, 20, 21	
	CRT	1, 2, 3, 4, 5	
	CTL	1, 2	
	CUP	2, 3, 5, 7, 8	
	DET	1- 7, 11, 12, 13, 24, 26, 28, 33, 34	
	GEM	1-7	
	INI	1, 9	
	MAS	2, 4, 5	

Function	*err-code*	*message-number*	Reference in the RMS Troubleshooting Guide
	MIS	1, 7, 8	
	QUE	13, 14	
	SCR	8, 9, 20, 21, 26	
	SWT	4, 20, 25, 26, 45, 46	
	SYS	1, 8, 13, 14, 15, 17, 48-50, 84, 90, 93, 94, 97, 98, 100	
	UAP	1, 5, 6, 7, 8, 9, 15, 16, 19, 21, 22, 23, 24, 27, 28, 29, 32, 33, 35, 41-45	
	US	5, 6, 31	
	WLT	1, 3, 5	
	WRP	1-12, 14-18, 23-25, 28-35, 52	
RMS error	ADC	16, 21, 69	4 Fatal error messages
	ADM	1, 2	
	BM	3, 49, 51, 58, 67, 69, 82, 89	
	CML	14	
	CMM	1, 2	
	CRT	6	
	DET	8, 9, 16, 18	
	INI	4, 7, 10, 11-14	
	MAS	1, 2, 6, 7	
	MIS	4	
	QUE	1, 2, 5, 6, 11, 12	
	SCR	4, 5, 10, 12, 13, 28, 29	
	SYS	33, 52	
	UAP	36	
	US	1, 42	
	WLT	9	
	WRP	40, 44-48	

Display format 6 (other)

If the display format is other than display format 1 to 5, see the following:

- See "12.7 CF Reason Code Table" in the "*PRIMECLUSTER CF Configuration and Administration Guide*

- See "12.8 Error messages for different systems" in the "*PRIMECLUSTER CF Configuration and Administration Guide*

- See "5 Console error messages" in the "*PRIMECLUSTER RMS Troubleshooting Guide.*"

- See "Appendix E GDS Messages" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide.*"

- See "Appendix A Kernel Messages" or "Appendix E List of Messages" in the "*PRIMECLUSTER Global File Services Configuration and Administration Guide.*"

- See "Appendix A List of Error Messages" in the "*PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function.*"

D.2 Information Messages

2206 The process (*count*: *appli*) was restarted.

Content:

The process monitoring function detected that a process being monitored had stopped and restarted the process.
count indicates the restart count of the process to be monitored. *appli* indicates the absolute path name of the restarted process.

Corrective action:

Message 2207 is output before this message.

See the explanation of message 2207.

2207 Process (*appli*) has stopped.

Content:

This message warns that the monitoring target with the process monitoring feature was stopped by other than the correct processing for stopping by the execution of the Offline script.

appli indicates the absolute path name of the monitored process that has been stopped.

Corrective action:

If the process was stopped because of an error, search for the cause by checking the process specification or core file error. The core file is written to the current directory of the process. If the file is written-enabled, the normal access control is applied. The core file is not generated for the process that has an effective user ID different than the actual user ID. For detailed investigations, contact the creator of this process.

2620 On the SysNode "*SysNode*", the userApplication "*userApplication*" transitioned to state *state* . Therefore, message " *number* " has been canceled.

Content:

Since the userApplication state has changed, the message has been cancelled. You do not need to respond to the operator intervention message.
SysNode indicates the name of SysNode in which the userApplication state has been changed. *userApplication* indicates the name of userApplication whose status has been changed. *state* indicates the userApplication state. *number* indicates the message number.

2621 The response to the operator intervention message " *number* " was action.

Content:

The response to the operator intervention message has been made.
number indicates the number of the operator intervention message to which you responded. action indicates with or without response by yes or no.

2622 There are no outstanding operator intervention messages.

Content:

There is no operator intervention messages that have yet to be processed.

2700 The resource failure has been recovered. SysNode:*SysNode* userApplication:*userApplication* Resorce:*resource*

Content:

The resource recovered from the failure.
SysNode indicates the name of SysNode whose resource has been recovered. *userApplication* indicates the name of the userApplication to which the recovered resource belongs. *resource* indicates the name of the resource that has recovered from the error state.

2701 A failed resource has recovered. SysNode:*SysNode*

Content:

SysNode that was in the Faulted state due to a node panic entered the Online state.
SysNode indicates a recovered SysNode.

2914 A new disk device(*disk*) was found.

Corrective action:

Register a new detected shared disk device (*disk*) in the resource database by executing the "clautoconfig(1M)" or using the CRM main window. For details on the "clautoconfig(1M)" command, see the manual pages of "clautoconfig(1M)."

disk indicates a newly detected shared disk device.

Supplement

This message is also displayed when a node is started on a machine where a DVD or CD-ROM is in a DVD-ROM device. If the device name indicating the DVD-ROM is in "(*disk*)", it is not necessary to register it in the resource database.
Start up the node after checking a DVD or CD-ROM is not in the DVD-ROM.

2927 A node (*node*) detected an additional disk. (*disk*)

Corrective action:

Register a newly detected disk unit (disk) in the resource database. A disk unit can be registered in the resource database by using the clautoconfig(1M) command or the CRM main window. For details on the clautoconfig(1M) command, see the manual page describing clautoconfig(1M).
node indicates the node identifier of the node on which a new disk unit was detected while disk indicates the newly detected disk unit. If there are many newly detected disk units, "..." will be suffixed to disk. In this case, you can reference all of newly detected disk units by searching for the message having number 2914 in the /var/adm/messages file on the node where the message has been output.

Supplement

This message is also displayed when a node is started on a machine where a DVD or CD-ROM is in a DVD-ROM device. If the device name indicating the DVD-ROM in "(*disk*)", it is not necessary to register it in the resource database.
Start up the node after checking a DVD or CD-ROM is not in the DVD-ROM.

3050 Patrol monitoring started.

3051 Patrol monitoring stopped.

3052 A failed LAN device is found to be properly running as a result of hardware diagnostics. (device:*altname* rid:*rid*)

Supplement

altname indicates the interface name of the specified LAN device while *rid* indicates its resource ID.

3053 A failed shared disk unit is found to be properly running as a result of hardware diagnostics. (device:*altname* rid:*rid*)

Supplement

altname indicates the device name of the shared device that is operating normally while *rid* indicates its resource ID.

D.3 Warning Message

5100 An error was detected in the failover unit of the line switching unit. (RCI:*addr* LSU:*mask* status:*status* type:*type*)

Corrective action:

No corrective action is required because processing is retried. If, however, retries fail repeatedly and if this warning occurs frequently, contact your Fujitsu customer engineer (CE).

Supplement

An error was detected in the switching control board of the line switching unit.

addr: Indicates the RCI address of the line switching unit in which an error was detected.

mask: Indicates the LSU mask to be controlled.

status: Indicates the an internal status of the line switching unit by error type.

type: Indicates an error type.

- 3: Indicates that the reserved status was not canceled.

   status: The reserved status of each LSU included in the switching unit is indicated by the value of the LSU mask.

   0: Indicates that the LSU is in the released status.

   1: Indicates that the LSU is in the reserved status.

- 4: Indicates that the connection has not changed.

   status: The status of the connection of each LSU included in the switching unit is indicated by the LSU mask value.

   0: Indicates that the LSU is connected to port 0.

   1: Indicates that the LSU is connected to port 1.

- 5: Indicates that the reserved status could not be created.

   status: The reserved status of each LSU included in the switching unit is indicated by the value of the LSU mask.

   0: Indicates that the LSU is in the released status.

   1: Indicates that the LSU is in the reserved status.

LSU mask value

LSU15 LSU14 LSU13 LSU12 ... LSU03 LSU02 LSU01 LSU00

0x8000 0x4000 0x2000 0x1000 ... 0x0008 0x0004 0x0002 0x0001

D.4 Error Messages

If an error message is output, analyze the log of the time zone in which the message was output from the "/var/adm/messages" message to check if any other error message had been output before that. If you found another error message, you need to take corrective action for this error message first.

???? Message not found!!

Corrective action:

The text of the message that is correspondent to the message number is not available. Contact your Fujitsu systems engineer (SE).

6000 An internal error occurred. (function:*function* detail :*code1-code2-code3-code4*)

Content:

An internal error occurred in the program. *function*, *code1*, *code2*, *code3*, and *code4* indicate the debugging information for error diagnosis.

Corrective action:

Collect the debugging information. Then, contact your Fujitsu systems engineer (SE). See "Appendix C Troubleshooting."
code1, *code2*, *code3*, and *code4* indicate the information that is required by your Fujitsu systems engineer (SE).

6001 Insufficient memory. (detail:*code1-code2*)

Corrective action:

One of the following could be the case:

- The memory resources are insufficient.

- The kernel parameter has incorrect settings.

Collect information required for troubleshooting. For details on how to collect information, see "Appendix C Troubleshooting."
Reexamine the estimation of the memory resources that are required for the entire system. For information on the amount of memory required for cluster control, see the "*PRIMECLUSTER Installation Guide*," which is provided with each product.
If you still have the problem, confirm that the kernel parameter setting is correct by referring to the "A.5.1 CF Configuration."
Change the settings if necessary, and then reboot the system. If above actions do not help you solve the problem, contact your Fujitsu systems engineer (SE).
code1 and *code2* indicate the information required for troubleshooting.

6002 Insufficient disk or system resources. (detail:*code1-code2*)

Corrective action:

One of the following could be the case:

- The disk resources are insufficient.

- The kernel parameter has incorrect settings.

Collect the information required for troubleshooting. For details on how to collect information, see "Appendix C Troubleshooting."

Confirm that there is sufficient disk space to enable the operation of PRIMECLUSTER. If necessary, delete any unnecessary files to create sufficient free space and then restart the system. For information on the amount of required disk space, see the "*PRIMECLUSTER Installation Guide*," which is provided with each product.

If you still have the problem, confirm that the kernel parameter settings are correct by referring to the "A.5.1 CF Configuration". Change the settings if necessary, and then reboot the system.

If above actions do not help you solve the problem, contact your Fujitsu systems engineer (SE).

code1 and *code2* indicate the information required for troubleshooting.

6003 Error in option specification. (option:*option*)

Corrective action:

Specify the option correctly, and then re-execute the operation.

option indicates an option.

6004 No system administrator authority.

Corrective action:

Re-execute the process with system administrator's authority.

6006 Required option *option* must be specified.

Corrective action:

Specify the option, and then re-execute the operation.
option indicates an option.

6007 One of the required options (*option*) must be specified.

Corrective action:

Specify the option, and then re-execute the operation.
option indicates an option.

6008 If option *option1* is specified, option *option2* is required.

Content:

When option *option1* is specified, option *option2* is required.

Corrective action:

Specify the options correctly, and re-execute the operation.

6009 If option *option1* is specified, option *option2* cannot be specified.

Content:

Two options specified with the command conflict.

Corrective action:

Re-execute the command by specifying either one of the options.

6010 If any one of the options *option1* is specified, option *option2* cannot be specified.

Content:

If one of the options *option1* is specified, option *option2* cannot be specified.

Corrective action:

Specify the options correctly, and re-execute the operation.

6208 Access denied (*target*).

Corrective action:

Collect debugging information and then contact your Fujitsu systems engineer (SE). For details on how to collect information, see "Appendix C Troubleshooting."
target indicates a command name.

6602 The specified resource does not exist. (detail:*code1-code2*)

Content:

The specified resource does not exist. code1 and code2 indicate the information required for troubleshooting.

Corrective action:

Specify the resource correctly, and re-execute the operation.

6615 The cluster configuration management facility is not running.(detail:*code1-code2*)

Corrective action:

Reboot the node on which this message is displayed and then reactivate the resource database. If this message appears again, collect debugging information and then contact your Fujitsu systems engineer (SE). For details on how to collect information, see "Appendix C Troubleshooting."

code1 and *code2* indicate the debugging information required for your Fujitsu systems engineer (SE).

6616 Cluster configuration management facility: error in the communication routine.(detail:*code1-code2*)

Corrective action:

Collect debugging information, and then contact your Fujitsu systems engineer (SE). For details on how to collect information, see "Appendix C Troubleshooting."
code1 and *code2* indicate the debugging information required for your Fujitsu systems engineer (SE).

6617 The specified state transition procedure file does not exist.

Corrective action:

Specify the state transition procedure file correctly, and then re-execute.
If the problem cannot be solved with this corrective action, collect debugging information and then contact your Fujitsu systems engineer (SE). For details on collecting debugging information, see "Appendix C Troubleshooting."

6618 The state transition procedure file could not be written. A state transition procedure file with the same name already exists.

Corrective action:

To overwrite the state transition procedure file, specify the "-o" option and then re-execute.

6619 The state transition procedure file could not be written. There was an error in the resource class specification.

Corrective action:

Specify the resource class correctly, and then re-execute. The available resource classes are file names under /etc/opt/FJSVcluster/classes.

Confirm that the character string specified as the resource class contains no errors.

6621 Could not perform file operation on state transition procedure file. (detail:*code1-code2*)

Corrective action:

- Confirm the disk resource from which the state transition procedure is retrieved or the node resource of the file system is sufficient by executing the "df(1M)" command. Correct the insufficient resource problem or change the location from which the state transition procedure is retried, and do the operation again.

- Confirm the disk resource on which the cluster system is installed or the node resource of the file system is sufficient by executing the "df(1M)" command. Correct the insufficient resource problem, and do the operation again.

If the problem cannot be solved even by applying with the corrective actions, collect debugging information and then contact your Fujitsu systems engineer (SE). For details on debugging information, see "Appendix C Troubleshooting."

code1 indicates debugging information while *code2* indicates error number.

6623 Cannot delete the specified state transition procedure file.

Corrective action:

You can delete only the state transition procedure file that was defined by a user. For information on how to delete it, refer to the manual of the product that provides the state transition procedure file.

6624 The specified resource does not exist in cluster service. (resource:*resource* rid:*rid*)

Corrective action:

A procedure resource that is registered to the cluster application is not in the resource database. This message is not output if the cluster application was registered by using the userApplication Configuration Wizard(GUI) and executing the "hvw(1M)" command. However, if this occurs, collect debugging information and contact your Fujitsu system engineer. For details on collecting debugging information, see "Appendix C Troubleshooting."

If hvw(1M) was used to register the procedure resource to the cluster application by FREECHOICE, check that the procedure resource displayed in the message text is not registered to the resource database by executing the "clgettree(1)" command. For details on the "clgettree(1)" command, see the manual page for clgettree(1M).

If the procedure resource is not registered to the resource database, register the procedure resource displayed in the message to the resource database. See "Part 6 PRIMECLUSTER Products."

If the procedure resource is registered to the resource database, collect debugging information and contact your Fujitsu system engineer. For details on collecting debugging information, see "Appendix C Troubleshooting."

Resource indicates the resource name that is not registered to the resource database, and *rid* indicates the resource ID of the resource that is not registered to the resource database.

6651 The specified instruction contains an error.

Content:

This message will not be displayed on a client system on which a browser is active. This message is displayed on the standard output when you respond to the operator intervention message with the CLI command (clreply). If a string other than yes or no is to be entered, this message will be displayed.

Corrective action:

Specify the option correctly, and then re-execute the processing.

6655 Use the absolute path to specify the option (*option*).

Content:

Specify the option *option* with the absolute path name.

Corrective action:

Specify the option correctly, and re-execute the operation.

6657 The specified resource is not being monitored. (detail:*code*)

Content:

The specified resource is not being monitored. *code* represents debugging information.

Corrective action:

No action is necessary.

6658 The specified process does not exist. (pid:*pid*)

Content:

The specified process does not exist. pid represents the process ID of the specified process.

Corrective action:

Specify the process ID correctly, and re-execute the operation.

6659 The specified command does not exist. (command:*command*)

Content:

The specified command does not exist. *command* represents the specified command.

Corrective action:

Specify the command with a full path name, and then re-execute the operation.

6662 A timeout occurred in process termination. (detail:*code1-code2*)

Content:

The process did not terminate within the termination wait time of the process.

code1 and *code2* indicate the information required for troubleshooting.

Corrective action:

Search for a cause. For information on the investigation, contact the creator of the process.

6680 The specified directory does not exist.

Corrective action:

Specify a correct directory, and then re-execute.

6690 The specified userApplication or resource is not monitored. (*resource*)

Corrective action:

userApplication or the resource is not registered as it is monitored by the patrol diagnosis.

Reexamine the configuration of the patrol diagnosis by referring to "6.7 Setting Up Patrol Diagnosis."
The specified userApplication or resource will be output to *resource*.

6691 The userApplication cannot do the patrol monitoring because of *status* .

Corrective action:

Specify userApplication that is in the following state, and then re-execute:

```
        Standby, Offline, Faulted, Deact
```

The state of userApplication will be output to *status*.

6692 Patrol monitoring timed out.

Corrective action:

Confirm that power is being supplied to the node on which patrol is enabled. Power on the node if necessary, and then re-execute.
Confirm that RMS is activated.
If RMS is not activated, start the RMS with Cluster Admin or the "hvcm(1M)" command and then re-execute operation.

6750 A resource has failed. SysNode:*SysNode* userApplication:*userApplication* Resorce:*resource*

Corrective action:

Confirm the state of userApplication in which the corrected resource is registered by using the RMS main window or executing the "hvdisp(1M)" command. You can also confirm the userApplication state by responding to the operator intervention message that is displayed after RMS is started. If userApplication is Faulted, you need to clear the Fault by using the RMS main window or specifying the "-c" option of the "hvutil(1M)" command.
SysNode indicates the name of SysNode in which a resource has failed. *userApplication* indicates the name of userApplication to which the failed resource belongs. *resource* indicates the name of the resource that has failed.

6751 A SysNode has failed. SysNode:*SysNode*

Corrective action:

Respond to the operator intervention message that is displayed after the node is started or confirm the state of userApplication running on the node in which the failure occurred, by using the RMS main window or executing the "hvdisp(1M)" command. If userApplication is Faulted, you need to clear the Fault by using the RMS main window or specifying the "-c" option of the "hvutil(1M)" command.
SysNode indicates the SysNode in which a failure occurred.

6752 The processing was canceled due to the following error. Error message from RMS command

Content:

If you execute the RMS command (hvdisp) to enable the history function for the failed resource or the operation intervention function, and this fails, an error message will be displayed. Examine this error message and, after taking an appropriate correction, execute the required processing again.

Corrective action:

Confirm the contents of this message and, after taking the required corrective action, re-execute the processing.

6753 Failed to process the operator intervention message due to the following error. (message number:*number* response:*action* command :*command*) Error message from RMS command

Content:

Since the RMS command ended abnormally, the operator intervention function for the message failed.

Corrective action:

Note the contents of this message and, after applying an appropriate correction, re-execute the required processing.
number indicates the operator intervention message for which processing failed. *action* indicates the first response to the message from the operator with either yes or no. *command* indicates the RMS command which ended abnormally. "Error message from RMS command" indicates an error message that the RMS command transmits when a standard error occurs.

6754 The specified message number (*number*) does not exist.

Corrective action:

- This message is output in the following situations:
  The operator specified a nonexistent message number when executing the "clreply" command. This problem can be solved by specifying a message number that is included in the unfinished message list.

- The message for which a response was entered has been canceled. This occurs when the message gets old since the application state has been changed. You do not have to take any action.

- The operator intervention function has automatically replaced the message for which a response has been entered with other message. For example, when the operator intervention function's message number 1422 is replaced by the message number 1423 and vice versa. In such a case, the operator intervention function responds to the first message, displays another message, and advises a user that the message number has been changed.

number indicates the number of operator intervention messages.

6755 Failed to respond to the operator intervention message due to the SysNode (*SysNode*) stop. (message number:*number* response : *action*)

Content:

Since the node indicated by SysNode has stopped, the response to the operator intervention message failed.

Corrective action:

Reboot the node, and then confirm that RMS is running.

SysNode indicates the SysNode name of the node that is stopped. *number* indicates the number of the operator intervention messages that failed to respond. *action* indicates the operator's response by yes or no.

6780 Cannot request to the process monitoring daemon.

Content:

The "clmonproc" command failed to request start or stop monitoring of the process monitoring function.

Corrective action:

The daemon process for the process monitoring function might not be running. Check the "prmd" process by executing the "ps" command. If the "prmd" process does not exist, execute /etc/init.d/clprmd start to recover the process. If the "prmd" process exists, execute /etc/init.d/clprmd stop and then /etc/init.d/clprmd start, to recover the process.

6781 The process (*appli*) cannot be monitored because the process hasn't made a process group at starting.

Content:

A process to be monitored by specifying the "-g" option in the "clmonproc" command must satisfy the condition "process ends immediately after descendent processes are generated." However, the parent process did not end even after the specified time (default 10 seconds) elapsed after the process was started. If the system continues to monitor the process in this condition, the load on the system is getting high, so the process was removed from the processes to be monitored by the process monitoring function.
appli indicates the absolute path name of the monitored process.

Corrective action:

Take one of the following actions:

1. Do not monitor descendent processes. Do not specify the "-g" option in the "clmonproc" command.

2. If the monitored process can be changed, do not change the process group in the descendent processes, and do not specify the "-g" option in the "clmonproc" command.

3. If processes were generated, end the parent process immediately. Set up the process so that the first child process that is generated becomes the process leader immediately after it operates.

6782 The process(*appli*) was not able to be executed. (errno:*error*)

Content:

The command that was specified with the "-a" option of the "clmonproc" command could not be executed.
appli indicates the absolute path name of the process that has failed startup. *error* indicates the detailed code.

Corrective action:

Check whether the command that was specified with the "-a" option of the "clmonproc" command can be executed. Execute the command, and check whether an error occurs. If this message is output again even if the command could be executed, contact your Fujitsu system engineer (SE).

Since errno is output as a detail code, the Fujitsu SE should investigate the cause from this message information. For example, if errno is 13 (EACCES), the most likely cause is that the command path that was specified in the "-a" option of the "clmonproc" command does not have execute permission.

6807 Disk device (NodeID *NodeID* , *disk*) cannot be detected.

Content:

The power to the disk unit may not be turned on or the disk may be disconnected. If an attempt is made to start userApplication, userApplication may not start normally.

Two messages, namely, this message and that of message number 6836 may be displayed depending on connection.

Corrective action:

Confirm that the power to the shared disk device is turned on and that the shared disk device is connected correctly. If power is not being supplied to the shared disk device, stop the node, turn on the power to the shared disk device, and then boot the node. If the connection of the shared disk device is incorrect, stop the node, rectify the connection, and then boot the node.
NodeID indicates the identification number of the node to which the disk unit was connected while *disk* indicates the shared disk that could not be detected.

6817 An error occurred during state transition procedure execution. (*error* procedure:*procedure* detail:*code1-code2-code3-code4-code5-code6-code7*)

Corrective action:

Collect debugging information and contact your Fujitsu system engineer. For details on collecting debugging information, see "Appendix C Troubleshooting."

procedure indicates the state transition procedure in which the error occurred. If the state transition procedure indicated by *procedure* was created, see the subsequent information, and correct the state transition procedure.

code1, *code2*, *code3*, *code4*, *code5*, *and code6* indicate the state transition instruction that resulted in an error. *code7* represents debugging information.

- code1: the first argument (state transition instruction type)

- code2: the second argument (cluster service instance type)

- code3: the third argument (state transition instruction timing)

- code4: the fourth argument (resource ID)

- code5: the fifth argument (state transition event type)

- code6: the sixth argument (state transition event detail)

error indicates the error cause. The error causes include the following:

- procedure file exit error

  [Error cause details]

  The possible causes are as follows:

  - The state transition procedure (*procedure*) is not a Bourne shell script.

  - The state transition processes for the individual arguments (*code1* to *code6*) of the state transition procedure (*procedure*) returned an error. The control program of the procedure resource determines the success or failure of the state transition processes based on the exit code (exit function) of the state transition procedure.

    - Exit code 0: The control program determines that the state transition process is normal.

    - Exit code other than 0: The control program determines that the state transition process failed.

  [Corrective action]

Take the following actions:

- Obtain the state transition procedure, and check that the state transition procedure (*procedure*) is a Bourne shell script. If the procedure is not a Bourne shell script, specify the following in the first line of the state transition procedure, and then register the state transition procedure again:

```
#!/bin/sh
```

To obtain the state transition procedure, execute the "clgetproc(1M)" command. To register the state transition procedure, execute the "clsetproc(1M)" command. For details on each command, see the corresponding manual page.

- Check the return value of the state transition process for each argument (*code1* to *code6*). If there is no process that sets up an exit code (exit function), you need to add this exit code process. If there is no process, the execution result of the last command in the state transition procedure becomes the return value.

If the problem is not resolved by the above action, contact your Fujitsu system engineer.

6836 The disk device (NodeID *NodeID* , *disk*) has changed.

Content:

The wiring of the shared disk device could be incorrect. If an attempt is made to launch the userApplication in the current state, userApplication may not start up normally.
Two messages, namely, this message and that of message number 6807 may be displayed depending on connection.

Corrective action:

Confirm that the shared disk device is connected correctly. If the connection is incorrect, stop the node, correct the connection, and then boot the node.
NodeID indicates the node identification number to which the shared disk device is connected while *disk* indicates the shared disk in which an error has been detected.

7050 A failure is detected in a LAN device as a result of hardware diagnostics. (node:*nodename* device:*altname* rid:*rid* detail:*code*)

Corrective action:

Take corrective action by referring to "7.4.2 Corrective Action when Patrol Diagnosis Detects a Fault."
The interface name of the LAN device for which a failure is diagnosed is output with *altname*. *nodename*, *rid*, and *code* indicate the information required for debugging.

7051 A network device monitoring command is abnormally terminated as a result of diagnosing a LAN device. (node:*nodename* device:*altname* rid:*rid* detail:*code*)

Corrective action:

Collect debugging information, and then contact your Fujitsu systems engineer (SE). For details on how to collect information, see "Appendix C Troubleshooting."
An interface name of the LAN device for which a failure is diagnosed is output with *altname*. *nodename*, *rid*, and *code* indicate debugging information.

7052 A failure of the shared disk device is detected as a result of the hardware diagnostics. (node:*nodename* device:*altname* rid:*rid* detail:*code*)

Corrective action:

Take corrective action by referring to "7.4.2 Corrective Action when Patrol Diagnosis Detects a Fault."
A device name for which a failure is diagnosed is output with *altname*. *nodename*, *rid*, and *code* indicate debugging information.

7053 A disk monitoring command is abnormally terminated as a result of the hardware diagnostics. (node:*nodename* device:*altname* rid:*rid* detail:*code*)

Corrective action:

Collect debugging information, and then contact your Fujitsu systems engineer (SE). For details on how to collect information, see "Appendix C Troubleshooting."

A shared device name for which a failure is diagnosed is output with *altname*. *nodename*, *rid*, and *code* indicate debugging information.

7054 A designated device cannot be opened as a result of diagnosing the shared disk device. (node:*nodename* device:*altname* rid:*rid* detail:*code*)

Corrective action:

Specify the available hardware in which patrol diagnosis is enabled. See "6.7 Setting Up Patrol Diagnosis."

The device name that could not found can be output with *altname*. *nodename*, *rid*, and *code* indicate debugging information.

7055 The designated LAN device cannot be found as a result of the hardware diagnostics. (node:*nodename* device:*altname* rid:*rid* detail:*code*)

Corrective action:

Specify the available hardware in which patrol diagnosis is enabled. See "6.7 Setting Up Patrol Diagnosis."

The interface name of the LAN device that could not be found can be output with *altname*. *nodename*, *rid*, and *code* indicate debugging information.

7056 The flag settings of the activated LAN device is found improper as a result of the hardware diagnostics. (node:*nodename* device:*altname* rid:*rid* detail:*code*)

Corrective action:

The flag used when the LAN device is activated is in the inadequate state such as not UP and not BROADCAST. After confirming the system configuration, activate the device correctly. Confirm the flag by executing the "ifconfig(1M)" command.

- UP must be set.

- BROADCAST must be set.

- LOOPBACK must not be set.

- POINTOPOINT must not be set.

- NOARP must not be set.

The network device name, which is in the inadequate state, can be output with *altname*. *nodename*, *rid*, and *code* indicates debugging information.

7101 SCF cannot be accessed because it is in the busy state. (type:*type*)

Corrective action:

Leave the operation undone for several minutes, and then re-execute accessing.
type indicates information for research.

7102 SCF open failed. (errno:*errno*)

Corrective action:

Contact your Fujitsu customer engineer (CE) to check whether the SCF is operating normally.
errno indicates an error number.

7103 SCF access failed. (errno:*errno*)

Corrective action:

Contact your Fujitsu customer engineer (CE) to check whether the SCF is operating normally.

errno indicates an error number.

7104 The subclass of the line switching unit cannot be identified. (RCI:*addr* Subclass:*no*)

Corrective action:

A line switching unit whose subclass cannot be identified is connected.

- *addr*: Indicates the RCI address of the line switching unit.

- *no*: Indicates the subclass of the line switching unit.

- 0x01: Indicates a 4-line switching unit.

- 0x02: Indicates a 16-line switching unit.

Confirm if the line switching unit is a supported product.

7105 The specified line switching unit does not exist. (RCI:*addr*)

Corrective action:

No line switching unit that is specified with the RCI address exists.
Confirm with your Fujitsu customer engineer (CE) whether the RCI address of the specified line switching unit is correct or it is connected.

addr indicates an RCI address.

7106 The power to the line switching unit is not on, or the RCI cable has been disconnected. (RCI:*addr*)

Corrective action:

Confirm if the power of the specified line switching unit is turned on and that the RCI cable is not disconnected.
addr indicates the RCI address of the line switching unit.

7108 Reservation of the line switching device failed. (RCI:*addr* LSU:*mask* retry:*no*)

Corrective action:

Although *no* times of attempts were made to reserve the switching unit, which was specified with *mask* of the line switching unit that was specified with RCI address (*addr*), reservation failed.
Check whether the line switching unit has failed, whether the RCI connection has an error, and whether the power-supply voltage is abnormal. If errors occur frequently, contact your Fujitsu customer engineer (CE).
addr indicates the RCI address of the line switching unit, while *mask* indicates a mask that represents a switching unit of the line switching unit and *no* indicates the retry count until an error is displayed.

Maintenance information (for system administrators)

Determine whether the specified line switching unit has an error.

7109 An error was detected in the switching control board of the line switching unit. (RCI:*addr* status:*status* type:*type*)

Corrective action:

An error was detected in the switching control board of the line switching unit.

addr: Indicates the RCI address of the line switching unit in which an error was detected.

status: Indicates the internal status of the line switching unit by error type.

type: Indicates the error type.

- 1: Error in the switching control system (*status*: status 0)

  status: Indicates the value of status 0 (see below) of the line switching unit.

  - 0x80: QANS (0: normal, 1: abnormal)

    Indicates that the switching line of QSC remains in the asserted state (abnormal) although other than switchover processing is being performed.

  - 0x40: QAST (0: normal, 1: abnormal)

    Indicates that the switching line of QSC cannot be asserted during switchover processing. Once an abnormal state arises, that state remains until the power supply is turned off.

- 2: Power/circuit error (*status*: status 0)

  status: Indicates the value of status 0 (see below) of the line switching unit.

  - 0x10: QENA (0: normal, 1: abnormal)

    Indicates the QSC switchover function enabled state. If this Bit is ON, it indicates that every connected QSC is abnormal.

  - 0x08: DCNV (0: normal, 1: abnormal)

    Indicates that the output voltage of the DC-DC converter, which is mounted in the QSC, is normal. If this Bit is ON, it indicates that the output voltage of the DC-DC converter is abnormal.

  - Remark: This bit is enabled only for a 4-line switching unit.0x04: PW12

    0 is indicated for a single power unit while 1 indicates two power units.

  - Remark: This bit is enabled only for a 16-line switching unit.0x02: PRY1 (0: normal, 1: abnormal)

  - 0x01: PRY0 (0: normal, 1: abnormal)

    Indicates whether the power supply is normal/abnormal. (Normal means that power is supplied without any failure.)

    Remark: This Bit is enabled only for a 16-line switching unit.

- 3: QSC connection error (*status*: status 1)

  status: Indicates the value of status 1 (see below) of the line switching unit.

  - 0x80: HSC (0: 4-line switching unit, 1: 16-line switching unit)

    Indicates if a 16-line switching unit is used.

  - 0x20: QSC1 (slot 1 mounted)

  - 0x10: QSC0 (slot 0 mounted)

    Indicates the slot in which the QSC (switching control board) is mounted. Values other than 0x10 and 0x20 indicate errors.

  - 0x02: OBSY

    Indicates that the other side of the duplicated QSC is presently executing switchover processing (canceling of reserve, switchover command 0, and switchover command 1).

Check whether the line switching unit has a failure, if the RCI connection has an error, or whether the power-supply voltage is abnormal. If these errors occur frequently, contact your Fujitsu customer engineer (CE).

Maintenance information (for system administrators)

Check whether the specified line switching unit has an error.

7110 An error was detected in the switching unit of the line switching unit. (RCI:*addr* LSU:*mask* status:*status* type:*type*)

Corrective action:

An error was detected in the switching unit of the line switching unit.

addr: Indicates an RCI address of the line switching unit in to be controlled.

mask: Indicates an LSU mask which is the target of control

status: Indicates the internal status of the line switching unit by error type (for research).

type: Indicates an error type.

- 1: Indicates an erroneous state.

  status: Indicates the abnormal state of each LSU, which is included in the switching unit, as the value of LSU mask.

  0: Indicates that the relevant LSU is normal.

  1: Indicates that the relevant LSU is abnormal.

- 2: Indicates that a request of canceling switchover/reserve has been sent to the unconnected LSU.

status: Indicates whether each LSU that is included in the switching unit is connected using the value of the LSU mask.

0: Indicates that the LSU is unconnected.

1: Indicates that the LSU is connected.

LSU mask value

LSU15 LSU14 LSU13 LSU12 ... LSU03 LSU02 LSU01 LSU00

0x8000 0x4000 0x2000 0x1000 ... 0x0008 0x0004 0x0002 0x0001

Check whether the line switching unit has a failure, if the RCI connection has an error, or if the power-supply voltage is abnormal. If these errors occur frequently, contact your Fujitsu customer engineer (CE).

Maintenance information (for system administrators)

Check whether the specified line switching unit has an error.

7111 The cluster event control facility is not running. (detail:*code1-code2*)

Corrective action:

Collect information for troubleshooting for all nodes, and then contact your Fujitsu systems engineer (SE). For details on how to collect information, see "Appendix C Troubleshooting."

Reboot the node in which the error occurred. If you need to stop the node, execute the "shutdown(1M)" command.

code1 and *code2* indicate information items for diagnosis.

7112 Communication failed in the cluster event control facility (detail:*code1-code2*)

Corrective action:

Collect information for troubleshooting for all nodes, and then contact your Fujitsu systems engineer (SE). For details on how to collect information, see "Appendix C Troubleshooting."

Reboot the node in which an error occurred. If you need to stop the node, execute the "shutdown(1M)" command.

code1 and *code2* indicate the information required for troubleshooting.

7113 Cluster event control facility: internal error. (detail:*code1-code2*)

Corrective action:

Copy this message, collect debugging information, and then contact your Fujitsu systems engineer (SE). For details on how to collect debugging information, see "Appendix C Troubleshooting."

After collecting debugging information from all nodes, reboot the node in which the error occurred. To reboot the node, use the shutdown(1M) command.

code1 and *code2* indicate the information required for troubleshooting.

7116 Port number information is not set for resource SWLine. (rid:*rid*)

Corrective action:

Set the port number attribute (port) of the line switching unit used for the resource SWLine.

rid indicates the resource ID of SWLine in the line switching unit.

7117 The port number specified for resource SWLine is incorrect. (rid:*rid* port:*port*)

Corrective action:

Set up the correct port number.
rid indicates the resource ID of the SWLine in the line switching unit while *port* indicates a port number.

7119 The LSU mask information has not been set for the shared resource SH_SWLine. (rid:*rid*)

Corrective action:

Set up the mask (Isu_mask) attribute of the switching unit used for the shared resource SH_SWLine.
rid indicates the resource ID of the shared resource SH_SWLine in the line switching unit.

7121 The parent resource of the shared resource SH_SWLine is other than shared resource SH_SWU. (rid:*rid*)

Corrective action:

Re-create the shared resource SH_SWLine as a child resource of shared resource SH_SWU.

rid indicates the resource ID of the shared resource SH_SWLine in the line switching unit.

7122 The RCI address information has not been set for shared resource SH_SWU. (rid:*rid*)

Corrective action:

Set up the RCI address attribute (addr) of the line switching unit used for the shared resource SH_SWU.
rid indicates the resource ID of the shared resource SH_SWLine in the line switching unit.

7125 The resource ID of the node connected to the specified port *no* (rid: *rid*) is incorrect.

Corrective action:

For the resource ID of the node that is connected to the specified port of the line switching unit, an incorrect resource ID of other than a node or nonexistent resource ID is specified. Set up the correct resource ID.
no indicates a port number in the line switching unit while *rid* indicates the resource ID.

7126 The resource ID (*rid*) of the same node is specified for ports 0 and 1.

Corrective action:

The same resource ID cannot be specified for port 0 and port 1 in the line switching unit. The same node cannot be specified as the operating and standby node.
Set up the correct resource ID of the node.
rid indicates the resource ID of the specified node.

7130 The specified resource ID (*rid*) cannot be deleted because it is being used.

Corrective action:

The specified resource ID cannot be deleted because it has been used for GDS
Confirm the resource ID to be deleted, and then set up the correct resource ID.
rid indicates the resource ID.

7131 The specified resource ID (*rid*) is not present in the shared resource class (*class*).

Corrective action:

Set up the correct resource ID.
rid indicates a resource ID while *class* indicates the class name of the shared resource class.

7132 The specified resource name (*name*) is not present in the shared resource class (*class*).

Corrective action:

Set up the correct resource ID.
name indicates a resource name while *class* indicates the class name of the shared resource class.

D.5 Operator Intervention Messages

1421 The userApplication " *userApplication* " did not start automatically because not all of the nodes where it can run are online. Do you want to force the userApplication online on the SysNode " *SysNode*

"? Message No.:*number*
Do you want to do something? (no/yes) Warning:Forcing a userApplication online ignores potential error conditions. Used improperly, it can result in data corruption. You should not use it unless you are certain that the userApplication is not running anywhere in the cluster.

 Note

....

If this message is output to the console and /var/adm/messages, "month day time node name" might be output to the beginning of the second and subsequent lines of the message independently of the text of the message. You can ignore it.

....

Corrective action:

This operator intervention message is displayed when userApplication cannot be online because RMS is inactive on some of the nodes that are designed to execute applications. Use the clreply command to respond from the command prompt (see "Appendix B Manual Pages").

If you select yes, the operator intervention function will execute the "hvswitch" command and userApplication will go online. Confirm whether userApplication is in the online state by using Cluster Admin or executing the "hvdisp." (For details on the "hvswitch" command and the "hvdisp" command, see "Appendix B Manual Pages").

If you select no, the RMS command that is used to change the userApplication state cannot be executed.

userApplication indicates userApplication that is not online in any node.

SysNode indicates SysNode on which userApplication will be online if you select yes.

number indicates a message number.

1422 On the SysNode " *SysNode* **", the userApplication "** *userApplication* **" is in the Faulted state due to a fault in the resource "** *resource* **". Do you want to clear fault? Message No.:***number*
Do you want to do something? (yes/no)

Corrective action:

Confirm that the resource indicated by *resource* has recovered. If you need to reply from CLI, execute the "clreply" command.

If you select yes, the operator intervention function will execute the "hvutil" command by specifying the clear option in the "hvutil" command and also brings userApplication Offline. Confirm that userApplication has gone Offline by executing the "hvdisp" command.

If you select no, the operator intervention function does not execute the RMS command to bring userApplication Offline.

SysNode indicates the name of SysNode on which the failure occurred. *userApplication* indicates the name of userApplication in which the failure occurred. *resource* indicates the name of the resource that caused a failure in userApplication. *number* indicates a message number.

1423 On the SysNode " *SysNode* **", the userApplication "** *userApplication* **" has the faulted resource "** *resource* **". The userApplication "** *userApplication* **" did not start automatically because not all of the nodes where it can run are online.**
Do you want to force the userApplication online on the SysNode " *SysNode* **"?**
Message No.:*number*
Do you want to do something? (no/yes) Warning:Forcing a userApplication online ignores potential error conditions. Used improperly, it can result in data corruption. You should not use it unless you are certain that the userApplication is not running anywhere in the cluster.

 Note

....

If this message is output to the console and /var/adm/messages, "month day time node name" might be output to the beginning of the second and subsequent lines of the message independently of the text of the message. You can ignore it.

....

Corrective action:

Confirm that the resource indicated by *resource* has recovered. If you need to reply from CLI, execute the "clreply" command.

If you select yes, the operator intervention function will execute the "hvswitch" command and userApplication will go Online. Confirm that userApplication has gone Online by executing the "hvdisp" command.

If you select no, the operator intervention function will not execute the RMS command to bring userApplication Online.

SysNode indicates the name of SysNode in which the failure occurred. *userApplication* indicates the name of userApplication in which

the failure occurred. *resource* indicates the name of the resource that caused the failure in userApplication. *number* indicates the message number.

D.6 Wgcnfclient Command Messages

0000: Message not found!!

Corrective action:

An unexpected error occurred.

If the same error is output again when you re-execute, collect debugging data and then contact your Fujitsu system engineer.

0001: Illegal option.

Corrective action:

Confirm how to specify the option and option argument, and then execute the command correctly.

0002: No system administrator authority.

Corrective action:

Execute the command with system administrator authority.

0003: File not found. (file:*file-name*)

Corrective action:

The operation environment file does not exist.
Re-install the "FJSVwvucw" package.

0004: Editing of the file failed.

Corrective action:

This message could be output when the state of the system is unstable such as when there is insufficient memory.
If this message is displayed for any other reason, follow the instructions given as the corrective action.
If the same error is output even when you re-execute the command, collect the debugging information and then contact your Fujitsu system engineer.

0005: Unknown keyword. (keyword:*keyword*)

Corrective action:

Confirm the specified keyword, and then execute the command with the correct keyword.

0006: The distribution of the file failed.

Corrective action:

Confirm that CRM is running on all cluster nodes including the node on which the command was executed. Activate CRM on all nodes if necessary, and then execute the command.

Check whether the state of any cluster node is unstable because of insufficient memory, and then re-execute the command. If you specify the "-v" option at this time, you can determine the node on which the failure occurred.

If the same error is output again, collect debugging information and then contact your Fujitsu systems engineer.

0007: The cluster configuration management facility is not running.

Corrective action:

Confirm that CRM is running on all cluster nodes including the node on which the command was executed. Activate CRM on all nodes if necessary, and then execute the command. Check whether the state of any cluster node is unstable because of insufficient memory. Correct the problem if necessary, and then execute the command again.

If you specify the "-v" option at this time, you can determine the node on which a failure occurred. If the same error is output repeatedly, collect debugging data and then contact your Fujitsu system engineer.

0009: The command received a signal.

Corrective action:

A signal is received during the execution of the command.

If the same error is output again when you re-execute, collect debugging data and then contact your Fujitsu system engineer.

D.7 GUI Messages

The section explains GUI messages.

- Messages related to the CRM view.

- Failed resource and operator intervention messages

- userApplication Configuration Wizard GUI messages

D.7.1 Messages Related to CRM View

This section explains the messages related to the CRM view.

If a message of the following display format is displayed, take corrective action by referring to the following manual.

Display format

If a message is displayed during operation of the CRM view or if a message whose frame title in the message dialog box is "Cluster resource management facility" is displayed:

Reference manual

For the above display format, see the following:

- Determine the reference destination according to the search procedure described in Display format 1 in D.1 Searching for a Message

- D.7.2 Failed Resource and Operator Intervention Messages (GUI)

D.7.2 Failed Resource and Operator Intervention Messages (GUI)

The messages displayed on the GUI in the event of a failed resource, as well as the operator intervention messages, are described below:

Information Messages

0701 There is no failed resource.

Content:

The results output by the cldispfaultrsc command do not contain any failed resources.

Error Messages

0765 Communication with the management server failed.

Content:

A communication error occurred between the browser and the Web-Based Admin View management server. This error occurs when the client's browser attempts to respond to an operator intervention message.

Corrective action:

Click *OK* to close the error message screen.
If an error occurs even when you re-execute, use the "clreply" command to reply. If a message that is output by Web-Based Admin View is displayed, follow the instructions for that message.

If the problem cannot be solved through the above method, collect the maintenance information and then contact your Fujitsu systems engineer (SE). For information on collecting maintenance information, see "B.2 Troubleshooting information collection" in the "PRIMECLUSTER Web-Based Admin View Operation Guide."

For details on messages 0001 to 0099, see "Appendix A Message" in the "PRIMECLUSTER Web-Based Admin View Operation Guide."

0766 The command terminated abnormally.

Content:

The "cldispfaultrsc" command or the "clreply" command ended abnormally.

Corrective action:

Follow the corrective action shown in the message.

0767 Command execution failed.

Content:

The "clreply" command cannot be executed while accessing the node.

Corrective action:

If SysNode that cannot be referenced in the detailed information appears, you can respond to the operator intervention message by executing the "clreply" command on SysNode. To respond through with the GUI, restart Web-Based Admin View on SysNode.
In all other cases, restart Web-Based Admin View on the node that serves as the management server of Web-Based Admin View. For information on how to restart Web-Based Admin View, see "6.2 Restart" in the "Web-Based Admin View Operation Guide."
If the failure cannot be solved, collect maintenance information and then contact your Fujitsu systems engineer (SE).
For details on the maintenance information, see "B.2 Troubleshooting information collection" in the "PRIMECLUSTER Web-Based Admin View Operation Guide."

0790 An error occurred while collecting the fault resources.

Content:

Node connection failed while the resource data was being collected.

Corrective action:

Reboot the management server of Web-Based Admin View for the affected client. For details, see the "PRIMECLUSTER Web-Based Admin View Operation Guide."
If the problem still cannot be solved, collect maintenance information and then contact your Fujitsu systems engineer (SE). For details on maintenance information, see "B.2 Troubleshooting information collection" in the "PRIMECLUSTER Web-Based Admin View Operation Guide."

0791 You do not have the access authority to respond to this message.

Content:

The user account that logged on to Web-Based Admin View does not have the execute permission to respond to this operator intervention message.

Corrective action:

Log out from Web-Based Admin View. Retry to log in as root or as a user belonging to wvroot, clroot, or cladmin, and then respond to the operator intervention message.

0792 The error occurred while accessing the management server. Select [Continue], and then end the Resource Fault History.

Content:

An error occurred while accessing the management server of Web-Based Admin View.

Corrective action:

Click the *Continue* button to close the "Fault Resource List" screen. If the network between the browser and the management server is merely temporarily disconnected, select the *Continue* button to close the "Fault Resource List" screen. (For example, you need to

select *Connect* when rebooting or when a LAN cable is temporarily disconnected). If a message output by Web-Based Admin View is displayed, follow the instructions for that message. If no message is output by Web-Based Admin View, choose *Fault Resource List* from the top menu to restart.

If the problem cannot be corrected by clicking the *Continue* button to reconnect, collect the maintenance information and then contact your Fujitsu systems engineer (SE). For information on collecting maintenance information, see "B.2 Troubleshooting information collection" in the "PRIMECLUSTER Web-Based Admin View Operation Guide."

D.7.3   userApplication Configuration Wizard GUI messages

The messages that userApplication Configuration Wizard GUI displays are listed below:

Information Message

0801   Do you want to exit userApplication Configuration Wizard GUI?

Corrective action:

To exit the userApplication Configuration Wizard GUI, select *Yes*. If you do not want to exit the GUI, select *No*.

0802   Do you want to cancel the setup process?

Corrective action:

To interrupt the process being operated by the userApplication Configuration Wizard GUI, select *Yes*. To cancel the interruption, select *No*.

0803   Do you want to register setup in a cluster system?

Corrective action:

To register the setup information to the system, select *Yes*. If you do not want to register the setup information to the system, select *No*.

0805   GUI is generating RMS Configuration.

Corrective action:

This message is automatically cleared when RMS Configuration generation is completed. Wait until this message is cleared.

0807   Do you want to remove only selected userApplication (*userApplication name*)? Do you want to remove all the resources under userApplication?

Corrective action:

To delete only userApplication, select *userApplication only*. To delete all resources under userApplication as well, select *All*. To cancel the deletion, select *Cancel*. "*userApplication name*" represents the name of userApplication to be deleted.

0808   Do you want to remove only selected Resource (*resource name*) and all the resources under Resource?

Corrective action:

To delete the resource indicated by "*resource name*" in the message and the resources under that resource, select *Yes*. If you do not want to delete the resources, select *No*.

0810 Node name takeover is registered or removed in userApplication. You need to restart SysNode to enable or disable takeover network. Restart SysNode after completing setup.

Corrective action:

As soon as the setup process is completed, restart all SysNode that contain userApplication for which node name takeover was set.

0813 GUI is reading RMS Configuration.

Corrective action:

The userApplication Configuration Wizard GUI collects RMS Configuration information.

The GUI cannot be closed until the process in the GUI ends.

0814 GUI is saving RMS Configuration in a system.

Corrective action:

The userApplication Configuration Wizard GUI collects RMS Configuration information.

The GUI cannot be closed until the process in the GUI ends.

0815 GUI is generating RMS Configuration.

Corrective action:

The userApplication Configuration Wizard GUI collects RMS Configuration information.

The GUI cannot be closed until the process in the GUI ends.

0816 Do you want to generate RMS Configuration?

Corrective action:

If you want the userApplication Configuration Wizard GUI to generate RMS Configuration information, select *Yes*. If you do not want the GUI to generate the information, select *No*.

0817 Do you want to distribute RMS Configuration?

Corrective action:

If you want the userApplication Configuration Wizard GUI to distribute RMS Configuration information, select *Yes*. If you do not want the GUI to distribute the information, select *No*.

0818 Do you want to cancel the scalable settings for userApplication?

Corrective action:

If you want the userApplication Configuration Wizard GUI to delete the scalable settings for the userApplication, select *Yes*. Otherwise, select *No*.

Warning Messages

0830 Since other client is using userApplication Configuration Wizard GUI or the hvw(1M), GUI cannot be started.

Corrective action:

Check whether the userApplication Configuration Wizard GUI has already been started in another client or whether another client is using the "hvw" command. After the other operation is completed, start the GUI or command again.

Even if the Web browser or the cluster management server of Web-Based Admin View is restarted while the GUI is being displayed, this message may be displayed at the next startup. If this is the case, wait about five minutes, and then restart the Web browser and the GUI.

0832 Cluster resource management facility is not running. Since a list of candidate interfaces cannot be obtained, GUI is terminated.

Corrective action:

Reasons for this occurring include the following:

- **Cause 1**

  A node for which initial setup has not yet been completed is included among the nodes to be monitored with Web-Based Admin View.

- **Cause 2**

   A node on which the cluster resource management facility is not running is included among the nodes to be monitored with Web-Based Admin View.

How to confirm cause 1

Confirm whether a node for which initial setup of the cluster resource management facility has not yet been completed is included among the nodes to be monitored with Web-Based Admin View.

Make the confirmation as follows:

1. Confirm the nodes to be monitored with Web-Based Admin View.

   Execute the wvstat (1M) command to confirm the monitoring target nodes for each operation management server.
   For details on wvstat(1M), see the relevant manual page.

2. Confirm the settings of the resource database.

   Use the CRM main window of Cluster Admin or the clgettree(1) command to confirm whether all the nodes that were confirmed in 1) can be displayed.
   For details on clgettree(1), see the relevant manual page.

If Web-Based Admin View uses a network other than a public LAN for operation management, the node names displayed in 1) and 2) will be different.

Corrective action for cause 1

If the initial setup of the cluster resource management facility has not yet been completed, use the userApplication Configuration Wizard GUI after performing initial setup, and then automatic configuration on the node on which initial setup of the cluster resource management facility has not been completed. For details on the initial setup of the cluster resource management facility, see "5.1.3 Initial Setup of the Cluster Resource Management Facility."

Once installation of the product package has been completed, use the userApplication Configuration Wizard GUI after node expansion, as explained in "9.1 Adding a Node."

For all other cases, change the settings of the operation management server of Web-Based Admin View to a host that is not the current operation management server.

For details on how to change the operation management server of Web-Based Admin View, see "4.3.3.1 Initial setup of the operation management server."

How to confirm cause 2

Confirm whether the node for which the cluster resource management facility is not operating is contained among the Web-Based Admin View monitoring target nodes.

Make the confirmation as follows:

1. Confirm the nodes to be monitored with Web-Based Admin View.

   Execute the wvstat (1M) command to confirm the monitoring target nodes of each operation management server.
   For details on wvstat(1M), see the relevant manual page.

2. Confirm the operation status of the cluster resource management facility.

   Use the CRM main window of Cluster Admin or the clgettree(1) command to confirm which node is in other than the "NO" state.
   For details on clgettree(1), see the relevant manual page.

Corrective action for cause 2

Boot the node on which the cluster resource management facility is not operating.

0833 RMS is running. Since Configuration might not be saved, GUI is terminated.

Corrective action:

RMS is running. You might fail to generate or distribute the configuration information that you were trying to create or modify. Stop RMS, and then re-execute the operation.

0834 An invalid character is included.

Corrective action:

A userApplication name, a resource name, a script name, or a script path includes an disabled character. Reenter the correct value.

0835 Removing resource (*resource name*) will concurrently remove userApplication (*userApplication name*). Do you want to continue?

Corrective action:

The system is about to delete all resources under userApplication.
To delete userApplication indicated by "*userApplication name*" in the message and all resources under userApplication, select *Yes*. If you do not want to execute deletion processing, select *No*.

0836 A name is not entered.

Corrective action:

Enter the userApplication name, the resource name, or the script name.

0837 A value is invalid.

Corrective action:

The number that was entered for TIMEOUT or net mask is a disabled value. Enter an appropriate value.

0838 The specified takeover IP address is not available.

Corrective action:

Enter another takeover IP address.

0839 There is an incorrect setup.

Corrective action:

Change the attribute (that has the incorrect data), and continue the operation.

0840 The takeover network name has been defined. Do you want to use the following definitions?

Corrective action:

Information on the takeover network is already specified in the /usr/opt/reliant/etc/hvipalias file. To use the values that were set, select *Yes*.

If you select *No*, the userApplication Configuration Wizard GUI re-creates the takeover network information in the previous file and in /etc/inet/hosts. In this case, the takeover network information can be deleted automatically from /usr/opt/reliant/etc/hvipalias and /etc/inet/hosts when the Ipaddress resource is deleted.

0841 There is an attribute different than the ones of other resources. Do you want to continue?

Corrective action:

Review if the displayed attribute is correct. If it is correct, select *Yes*.

0848 The file name is not specified.

Corrective action:

Enter the path of the reference file.

0849 A required setup is missing.

Corrective action:

Set the required items in the screen.

0852 It is not a proper combination.

Corrective action:

You need to set two or more userApplication for one exclusion group.

0856 The selected userApplication or Resource cannot be edited.

Corrective action:

Change the userApplication name or resource name, and execute the setup process again.

0857 The specified takeover Ipaddress or host name has been used.

Corrective action:

Enter another IP address or host name.

0859 Invalid file name or path.

Corrective action:

Enter another reference file path.

0860 The specified file exists. Do you want to replace it?

Corrective action:

The file to be created already exists. If you select *Yes*, the existing file will be replaced. If you select *No*, the existing file will not be replaced.

0861 The specified interface is different. Do you want to set up IP address?

Corrective action:

The IP address segment that is assigned to the network interface card (NIC) for the takeover network is different. To continue the setup with the assigned IP address segment, select *Yes*. To select the NIC again, select *No*.

0866 The file system has been used.

Corrective action:

The file system that you tried to create is already being used. Set another device path and mount point.

0867 Since a list of candidate interfaces cannot be obtained. The process is exited.

Corrective action:

Candidate interfaces to be set to the resource could not be obtained. Check whether interfaces were previously registered to the cluster resource management facility, and then re-execute the operation.

0868 It is not an executable file.

Corrective action:

The specified file is not executable. Use another file, or add an execute permission to the file.

Error messages

0880 A non-classified error occurred.

Corrective action:

An error occurred in the userApplication Configuration Wizard GUI.

Collect Java console information, detailed error information, and maintenance information, and contact a Fujitsu technician (SE).

For maintenance information, see "Appendix C Troubleshooting."

0881 Connection to the server failed.

Corrective action:

An internal server contradiction or an event that caused a network disconnection between the Web browser and the cluster management server (restart, shutdown, emergency stop, LAN cable disconnection) may have occurred.

If a Web-Based Admin View message is displayed, take corrective action for that message.

If a Web-Based Admin View message is not displayed, respond to this message and restart the userApplication Configuration Wizard.

If the problem cannot be resolved by the action described above, collect Java console information, detailed error information, and maintenance information, and contact your Fujitsu system engineer.

For maintenance information, see "Appendix C Troubleshooting."

0882 A non-supported package is installed. Check the version.

Corrective action:

A package having a different version from the package version required by the userApplication Configuration Wizard GUI is installed in the cluster node. See the information that was output to (detailed information), and reinstall the correct package of the cluster node.

0883 Since the specified file is in the non-supported format, it cannot be edited.

Corrective action:

The file cannot be edited because it is not a Bourne, C, or Korn shell file. Specify another file.

0886 Since a list of candidate interfaces that can set in Resource is not acquired. The process is exited.

Corrective action:

A candidate list of the information to be set to Resource could not be obtained. Check whether the cluster resource management facility is operating, whether automatic resource registration was executed, and whether Gds and Gls have been set up. If any of these setup operations has not been completed, correct them, and execute the process again.

0888 The command is abnormally terminated.

Corrective action:

Execute the command again, and if you still have the same problem, collect Java console information, detailed error information, and maintenance information, and then contact your Fujitsu system engineer.
The message that is output from the command may contain the "FJSVcluster" keyword and a message number. In this case, take corrective action described in "12.10 Resource Database messages" in the "*CF Configuration and Administration Guide.*"

For maintenance information, see "Appendix C Troubleshooting."

0889 The command execution failed.

Corrective action:

Execute the command again, and if you still have the same problem, collect Java console information, detailed error information, and maintenance information, and then contact a Fujitsu technician (SE).

For maintenance information, see "Appendix C Troubleshooting."

0890 The SysNode for executing a command cannot be found.

Corrective action:

When the three-layer configuration of Web-Based Admin View is used, no cluster node, which can be accessed from the management server, can be found. The operating status of Web-Based Admin View can be confirmed by executing /etc/opt/FJSVwvbs/etc/bin/wvstat.
If the problem cannot be improved when the environment of Web-Based Admin View has no problem, collect Java console information, detailed error information, and maintenance information, and then contact your Fujitsu systems engineer (SE).

For maintenance information, see "Appendix C Troubleshooting."

0891 Reading of the RMS Configuration failed.

Corrective action:

The RMS configuration information cannot be analyzed.

If userApplication or Resource is not created, change to other Configuration information.
If the problem cannot be improved, collect Java console information, detailed error information, and maintenance information, and then contact your Fujitsu systems engineer (SE).

For maintenance information, see "Appendix C Troubleshooting."

0893 RMS Configuration generation failed.

Corrective action:

Start the operation again from the beginning.

If the problem is not resolved by the above action, collect Java console information, detailed error information, and maintenance information, and then contact your Fujitsu system engineer.

For maintenance information, see "Appendix C Troubleshooting."

0895 RMS Configuration distribution failed.

Corrective action:

Check the following:

1. Whether there are resources that are not related to any userApplication

2. Whether there is userApplication that does not have any resources

If one or both of the above conditions are found, delete any unnecessary userApplication and resources. If none of the above conditions is found, do the operation again from the beginning. If an error still occurs, collect Java console information, detailed error information, and maintenance information, and then contact your Fujitsu system engineer. For maintenance information, see "Appendix C Troubleshooting."

0896 Since a list of candidate procedure resources that can set in Resource has not been acquired, the process is exited.

Corrective action:

A list of candidate information to be set in Resource could not be acquired. Check if the cluster resource management facility is operating and whether the procedure resource has already been set up. If it has not yet been set up, first set up the procedure resource, and then re-execute the processing.

Appendix E Using SynfinityCluster Products in PRIMECLUSTER

Target users:

Users who use SynfinityCluster products

This appendix describes the required information for applying a SynfinityCluster-related topic in a reference manual to PRIMECLUSTER. The information focuses on the following points:

- Terminology

- Manual series

- Building procedure

E.1 Terminology

The terms of SynfinityCluster and PRIMECLUSTER are as follows:

SynfinityCluster	PRIMECLUSTER
Cluster services	Cluster applications, userApplication
SynfinityCluster agent products	PRIMECLUSTER products
State Transition Procedure	Online/Offline scripts
Cluster Service Instance	None (Expressions such as operating node and standby node are used instead.)
Cluster Domain	None
Node	Node or Host
Failover	Switch Over (Switchover or Failover)
Private LAN	Interconnect
Standby Patrol	Patrol diagnosis
Public LAN Takeover	Network Takeover
Console Function	RC2000
Monitoring Function	RMS Monitoring Function
SynfinityLink	GLS (Global Link Services)
SynfinityDisk, SynfinityDisk/Global	GDS (Global Disk Services)
SynfinityFile, SynfinityFile/Global	GFS (Global File Services)

E.2 Manual Series

The manual series for SynfinityCluster and PRIMECLUSTER are shown below.
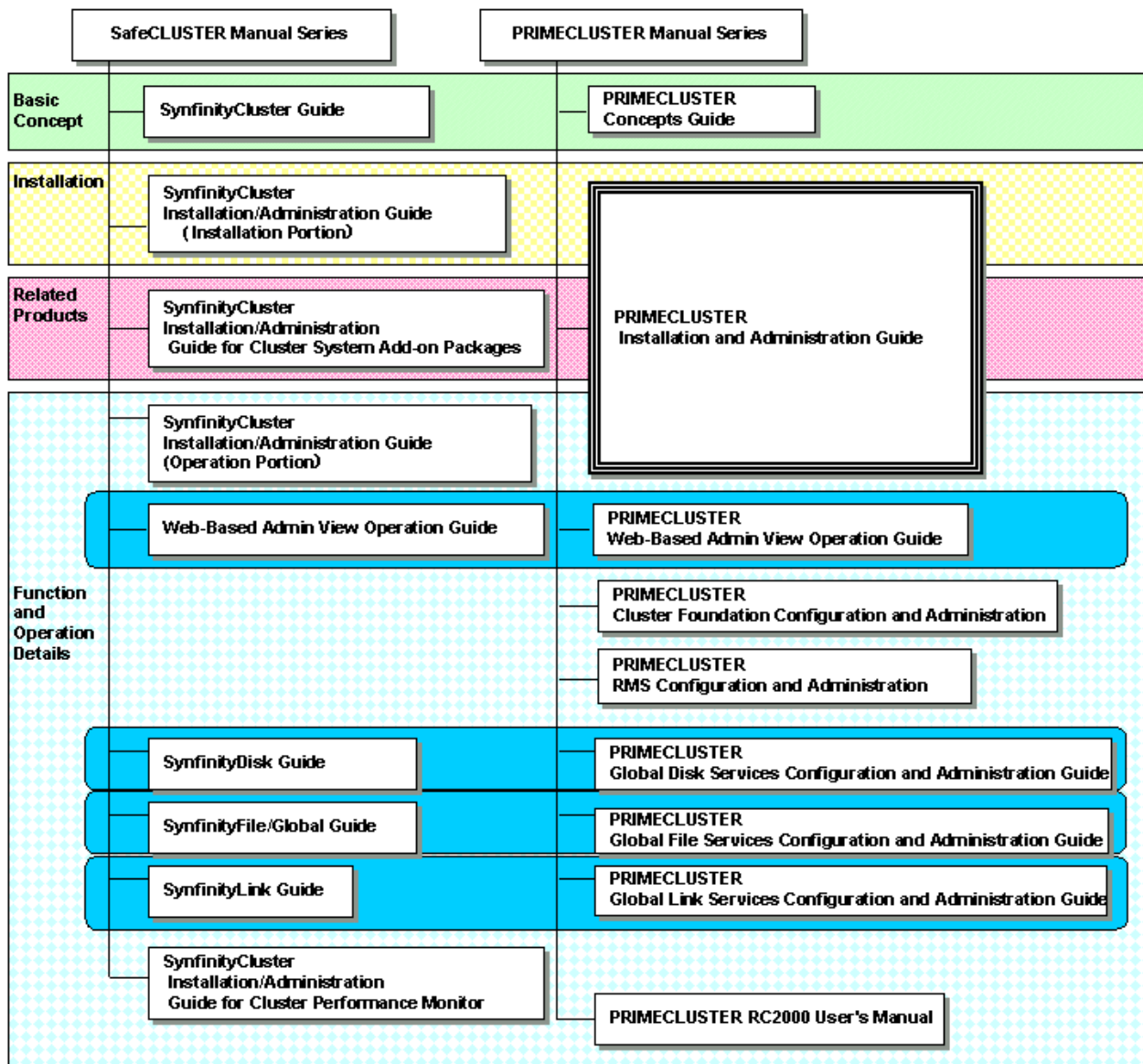
Table E.1 Manual list and manual uses

PRIMECLUSTER manual	Use	Reference
PRIMECLUSTER Installation/ Administration Guide **(This Manual)**	This manual is a general guide for all PRIMECLUSTER operations (installation, design, and maintenance). The manual also explains differences between PRIMECLUSTER and SynfinityCluster, the predecessor product, and information on PRIMECLUSTER products.	Installation - Design to maintenance
PRIMECLUSTER Concepts Guide	This manual allows readers to learn basic knowledge (technology, architecture, functions, etc.) related to each cluster system.	Design
PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide	This manual describes theory to setup information related to the Cluster Foundation (CF) functions.	Installation to operations
PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide	This manual describes theory to setup information related to the RMS functions.	Installation to operations

- 453 -

PRIMECLUSTER manual	Use	Reference
PRIMECLUSTER Web-Based Admin View Operation Guide	This manual describes theory to setup information related to Web-Based Admin View.	Installation to operations
PRIMECLUSTER Global Disk Services Configuration and Administration Guide	This manual describes information on Global Disk Services (GDS).	Installation to operations
PRIMECLUSTER Global File Services Configuration and Administration Guide	This manual describes information on Global File Services (GFS).	Installation to operations
PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function	This manual describes information related to the transmission path duplexing function of Global Link Services (GLS).	Installation to operations
PRIMECLUSTER Global Link Services Configuration and Administration Guide: Multipath Function	This manual describes information related to the multipath function of Global Link Services (GLS).	Installation to operations
RC2000 User's Guide	This manual presents an overview of RC2000 and describes the RC2000 functions.	Installation to maintenance

E.3   Building Procedure

This section compares the PRIMECLUSTER operation procedures and the SynfinityCluster operation procedures.

E.3.1   Software Installation

The installation methods for the cluster products and related software are shown below.

Table E.2 Installation method

Operation	SynfinityCluster	PRIMECLUSTER
Installation using a custom jumpstart	Installation using custom Jumpstart	One Shot Installer
Installation in each node	Installation in node units	CLI Installer (installation using an installation script)

E.3.2   Initial Cluster Setup

The initial setup procedures for the cluster products are shown below.

Table E.3 Initial setup procedure of the cluster

Operation	SynfinityCluster	PRIMECLUSTER
Initial setup of the cluster system	Setting up cluster configuration information and the resource database Cluster operation management view (GUI) *SynfinityCluster initial setup*	5.1 Initial Cluster Setup 5.1.3 Initial Setup of the Cluster Resource Management Facility
	Setting up cluster configuration information Setting up a private LAN	5.1.1 Setting Up CF and CIP (CF Wizard used)
	Setting up console asynchronous monitoring	5.1.2 Configuring the Shutdown Facility

Operation	SynfinityCluster	PRIMECLUSTER
	Setting up optional hardware resources	5.2 Setting Up Power Supply Linkage
	Setting up shared disk units	5.3 Setting Up Shared Disk Connection Confirmation
	Changing the operation environment for standby patrol	6.7 Setting Up Patrol Diagnosis

E.3.3 Application Setup

The setup procedures for operating applications with each product are shown below.
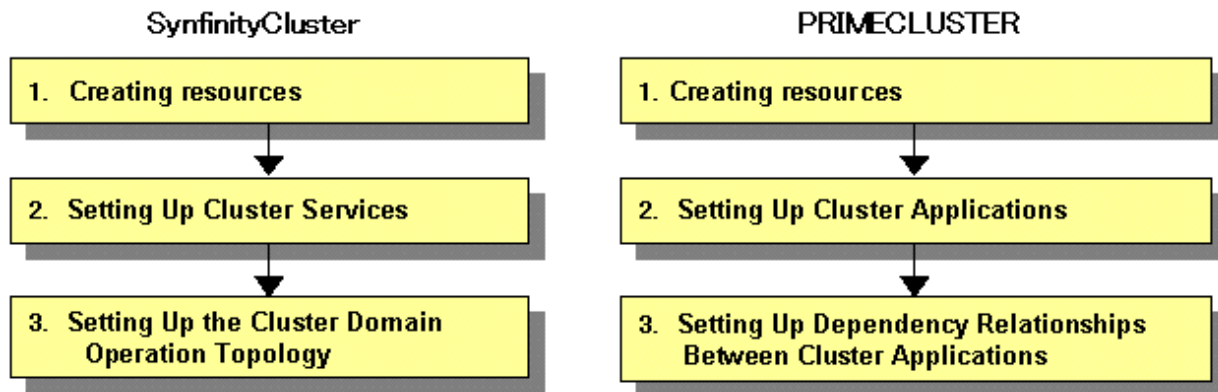
Setting up resources

Create the elements used by cluster services (cluster application in PRIMECLUSTER) as resources.

Operation	SynfinityCluster	PRIMECLUSTER
Setting up the takeover network	Set Cluster Operation GUI Setting up a takeover network	6.6.1.5 Creating Takeover Network Resources
Setting up state transition processing (script) of an application	Set Cluster Operation GUI Set State Transition Procedure of Set Application Resource	F.1 Registering a Procedure Resource 6.6.1.6 Creating Procedure Resources Note: This function is for making a SynfinityCluster product compatible for use in PRIMECLUSTER.
Using scripts and commands as resources	-	6.6.1.1 Creating Cmdline Resources
Mount information for shared disks	Manual	6.6.1.2 Creating Fsystem Resources
Setting up shared disks	Set Cluster Operation GUI Shared resource units	6.6.1.3 Creating Gds Resources
Setting up high availability resources	Manual operation of SynfinityLink (CLI)	6.6.1.4 Creating Gls Resources

Setting up services

Services that operate in the cluster are provided for switchover units. These services are called "cluster services" in SynfinityCluster and "cluster applications" in PRIMECLUSTER.

Operation	SynfinityCluster	PRIMECLUSTER
Setting the application startup priority	Set Cluster Operation GUI Set Cluster Service - Set Application Start/Stop Priority menu items	Resource association
Setting a disk group	Set Cluster Operation GUI Set Cluster Service-Set Group menu items	GDS GUI
Setting the type of shared disk device	Set Cluster Operation GUI Set Cluster Service-Set Shared Disk Unit by Use menu items	6.6.1.3 Creating Gds Resources

Setting up Inter-service relationships

Operation	SynfinityCluster	PRIMECLUSTER
Setting a launch priority	Set Cluster Domain Operation Mode-Set Cluster Service Startup Priority menu items	-
Setting a low job priority	Set Cluster Domain Operation Mode-Set Cluster Service Application Priority/Exclusive Relationship menu items	-
Setting an exclusive relationship	Set Cluster Domain Operation Mode-Set Cluster Service Application Priority/Exclusive Relationship menu items	"userApplication Configuration Wizard" menu-Set Dependency Relationship Between userApplications-Set userApplication exclusion 6.6.3.1 Setting Exclusive Relationships Between Cluster Applications

E.3.4 Operations

The operations of each cluster product are shown below.

System state monitoring

Operation	SynfinityCluster	PRIMECLUSTER
Start the operation management screen	From the top menu of the cluster operation management view, select *SynfinityCluster* and then select *Cluster Operation Management.*	From the top menu of Web-Based Admin View, select *Global Cluster Services* and then select *Cluster Admin.*

Configuring the operation management screen	- Node view
	  Display a hierarchical list of the resources (system resources) as seen from a node.
	
	- Service view
	  Display a hierarchical configuration of the cluster resources.
		- CF GUI
		
		- CRM main window (displays only resources related to hardware)
		
		- RMS main window
		
		- MSG main window

When the resource state changes

Operation	SynfinityCluster	PRIMECLUSTER
Checking resource errors	- Message - Node View - Service View	- Message - Resource Fault History screen - Failed resource list
Patrol diagnosis	1. Detection by standby patrol 2. Confirmation with messages	1. Detection by patrol diagnosis 2. Confirmation with messages and CRM main window

Appendix F Registering, Changing, and Deleting State Transition Procedure Resources for SynfinityCluster Compatibility

Operation target:

The procedures described here are necessary only if you will be using products that have been migrated from SynfinityCluster.

To use a procedure resource in a cluster application, you must register the procedure resource before setting up the cluster application.

This appendix explains how to register, change, and delete procedure resources.

F.1 Registering a Procedure Resource

This section explains how to register a procedure resource.

Take the following steps to register a procedure resource on all nodes where the procedure resource needs to be registered.

Note

If a process is started without process monitoring from a state transition procedure, the hard- and software limits in the file descriptor (fd) for that process will both be "1024".
If a process started from a state transition procedure requires a file descriptor above "1024", please add a "ulimit" definition at the beginning of processing for the state transition procedure, and specify the maximum value for the file descriptor.

Example

For setting the maximum value for the file descriptor to "4096":

```
#!/bin/sh
ulimit -n 4096
...
```

Operation Procedure:

1. Log in with the system administrator authority to the node in which the procedure resource is to be registered.

2. Execute the "clsetproc(1M)" command to register the state transition procedure.

See

For details on the "clsetproc(1M)" command, see the manual page.

Example

To register the "/tmp/program" state transition procedure as program (file name) to the BasicApplication class

```
# /etc/opt/FJSVcluster/bin/clsetproc -c BasicApplication -m program /tmp/program
```

Point

To overwrite a state transition procedure that is already registered, specify the -o option.

3. Registering the procedure resource

   Execute the "claddprocrsc(1M)" command to register the procedure resource.

   ![See icon] **See**

   ....

   For details on the claddprocrsc(1M) command, see the manual page for claddprocrsc (1M).

   ....

   ![Example icon] **Example**

   ....

   When registering a procedure resource, this procedure resource has to meet the following conditions;

   - The resource key of the procedure resource is SDISK,

   - The procedure name is program,

   - This procedure resource uses the state transition procedure registered in the BasicApplication class,

   - The procedure resource is registered in the node (NODE1), and

   - The state transition procedure receives the requests of START RUN AFTER and STOP RUN BEFORE

   ```
   # /etc/opt/FJSVcluster/bin/claddprocrsc -k SDISK -m program -c BasicApplication -s NODE1 -K
   AFTER -S BEFORE
   ```

   ....

F.2   Changing a Procedure Resource

This section explains how to change a procedure resource.

F.2.1   Changing a State Transition Procedure

Take the following steps to change the state transition procedure on all nodes where it needs to be changed.

Operation Procedure:

1. Log in with the system administrator authority to the node in which the state transition procedure is to be changed.

2. Execute the "clgetproc(1M)" command to retrieve the state transition procedure.

   ![See icon] **See**

   ....

   For details on the "clgetproc(1M)" command, see the manual page.

   ....

   ![Example icon] **Example**

   ....

   When retrieving a state transition procedure, this procedure resource has to meet the following conditions;

   - The state transition procedure is retrieved to the "/tmp" directory.

   - The file name registered in the BasicApplication class is program

   ```
   # /etc/opt/FJSVcluster/bin/clgetproc -c BasicApplication -f /tmp/program program
   ```

   ....

3. Modifying the state transition procedure

   Using a text editor such as vi(1), modify the state transition procedure that was retrieved in Step 2.

4. Registering the state transition procedure

   For instruction on registering the procedure, see Step 2 of "F.1 Registering a Procedure Resource."

F.2.2 Changing the Startup Priority of a State Transition Procedure

Take the following steps to change the startup priority of a state transition procedure on all nodes where it needs to be changed.

📗 Note

To change the startup priority of a state transition procedure, you need to delete a procedure resource with the procedure for changing a cluster application configuration and create the procedure resource again.

For more details, see "8.1.1 Changing the Cluster Application Configuration."

Operation Procedure:

1. Log in with the system administrator authority to the node in which the state transition procedure of the procedure resource is to be changed.

2. Execute the "clsetprocrsc(1M)" command to change the startup priority of the procedure resource.

🗔 See

For details on the "clsetprocrsc(1M)" command, see the manual page.

📄 Example

When changing the startup priority of the procedure resource to 10000, this procedure resource has to meet the following conditions;

- The resource class registered in the node (NODE1) is the BasicApplication class.

- The resource name is SDISK.

```
# /etc/opt/FJSVcluster/bin/clsetprocrsc -n SDISK -c BasicApplication -s NODE1 -p 10000
```

F.2.3 Changing Registration Information of a Procedure Resource

Take the following steps to change the registration information of the procedure resource on all nodes where the state transition procedure needs to be changed.

Operation Procedure:

1. Log in with the system administrator authority to the node in which the state transition procedure of the procedure resource is to be changed.

2. Execute the "clsetprocrsc(1M)" command to change the registration information of the procedure resource.

🗔 See

For details on the "clsetprocrsc(1M)" command, see the manual page.

📄 Example

When adding the requests of START RUN BEFORE state transition to the procedure resource that meet the following conditions;

- The resource key of the procedure resource is SDISK,

- This procedure resource uses the state transition procedure registered in the BasicApplication class,

- The procedure resource is registered in the node (NODE1), and

- The state transition procedure receives the requests of START RUN AFTER and STOP RUN BEFORE

```
# /etc/opt/FJSVcluster/bin/clsetprocrsc -n SDISK -c BasicApplication -s NODE1 -K BEFORE,AFTER
-S BEFORE
```

F.3   Deleting a Procedure Resource

Take the following steps to delete a procedure resource on all nodes where the procedure resource needs to be changed.

Operation Procedure:

1. Log in with the system administrator authority to the node from which the procedure resource is to be deleted.

2. Execute the "cldelprocrsc(1M)" command to delete the procedure resource.

🗂 See
For details on the "cldelprocrsc(1M)" command, see the manual page.

📋 Example
When deleting a procedure resource, the procedure resource needs to meet the following conditions;

- The resource key of the procedure resource is SDISK,

- This procedure resource uses the state transition procedure registered in the BasicApplication class, and

- The node identification name is NODE1

```
# /etc/opt/FJSVcluster/bin/cldelprocrsc -n SDISK -c BasicApplication -s NODE1
```

3. Deleting the state transition procedure

If a state transition procedure becomes unnecessary after all procedure resources that use that state transition procedure have been deleted, execute the cldelproc(1M) command to delete the state transition procedure.

🗂 See
For details on the "cldelproc(1M)" command, see the manual page.

📋 Example
When deleting a procedure resource, the procedure resource needs to meet the following conditions;

- The procedure name is program, and

- This procedure resource uses the state transition procedure registered in the BasicApplication class

```
# /etc/opt/FJSVcluster/bin/cldelproc -c BasicApplication program
```

Appendix G Registering SynfinityCluster-compatible Resources

This appendix explains how to register SynfinityCluster-compatible resources.

The procedures described here are necessary only when the following product is used:

- Interstage Traffic Director

See
....
See "G.3 Deleting a Cluster Service" only if the userApplication configuration was changed.
....

G.1 Registering a Cluster Service

If you used the CUI to set up userApplication, you need to register a SynfinityCluster-compatible cluster service and a cluster service instance resource to the Resource Database for the following product:

- Interstage Traffic Director

Note
....
- If you used the GUI to set up userApplication, the procedure below is unnecessary.

- If you are using the CUI to set up SynfinityCluster-compatible userApplication, see the following:

  "/opt/SMAW/SMAWRrms/htdocs.solaris/wizards.en/wizards/CRM.htm"
....

Describe below is the procedure for registering a SynfinityCluster-compatible cluster service and a cluster service instance resource.

Operation Procedure:

Note
....
Carry out Steps 1 and 2 in any one of the nodes that make up userApplication. The nodes refer to all nodes that were set to "Machines" in "(5)" of "CRM.htm."
....

1. Log in with the system administrator authority to any node of the cluster system.

2. Execute the "claddrsc(1)" command in the format shown below. If the registration is successful, the resource ID of the cluster service is output to standard output.

```
# /etc/opt/FJSVcluster/sys/claddrsc -c ClusterService -k ApplicationName-of-userApplication * -
p 0 -t nocontroler -s -1 -d self -e service
```

* For "*ApplicationName-of-userApplication*," specify the application name of the user application that was set to "(4)" of "CRM.htm."

Example) If the application name of userApplication is "generic."

```
# /etc/opt/FJSVcluster/sys/claddrsc -c ClusterService -k generic -p 0 -t nocontroler -s -1 -d
self -e service
100
```

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

4. Obtain the node identification number of the local node.

   Execute the clgetnode(1) command in the format shown below, and check the node identification number of the local node. If clgetnode(1) execution is successful, the node identification number of the local node is output to standard output.

   ```
   # /etc/opt/FJSVcluster/bin/clgetnode -i
   ```

   In the following example, the node identification number of the local node is "0."

   ```
   # /etc/opt/FJSVcluster/bin/clgetnode -i
   0
   ```

5. Register the cluster service instance resource.

   Execute the claddrsc(1) command in the format shown below. If the registration is successful, the resource ID of the cluster service instance is output to standard output.

   ```
   # /etc/opt/FJSVcluster/sys/claddrsc -c AndBinding -k ApplicationName-of-userApplication (*)  -p
   resource-ID-of-cluster-service (**)  -t nocontroler -s node-identification-number-of-local-node
   (***)  -d self -e service
   ```

   *For "*ApplicationName-of-userApplication*," specify the application name of the user application that was set to "(4)" of "CRM.htm."

   **For "*resource-ID-of-cluster-service*," specify the resource ID of the cluster service that was registered in Step 2.

   ***For "*node-identification-number-of-local-node*," specify the node identification number that was obtained in Step 4 for the local node.

   Example) If the application name of userApplication is "generic," the resource ID of the cluster service is "100," and the node identification number of the local node is "0."

   ```
   # /etc/opt/FJSVcluster/sys/claddrsc -c AndBinding -k generic -p 100 -t nocontroler -s 0 -d self
   -e service
   101
   ```

6. Log in with the system administrator authority to any one of the nodes that make up the cluster system.

7. Set up the operation sequence of the cluster service instances.

   Execute the clsetinorder(1M) command in the format shown below, and set the operation sequence of the cluster service instances.

   ```
   # /etc/opt/FJSVcluster/bin/clsetinorder -s resource-ID-of-cluster-service(*) -r resource-ID-of-
   cluster-service-instance ...(**) -S
   ```

   *For "*resource-ID-of-cluster-service*," specify the resource ID of the cluster service that was registered in Step 2.

   **For "*resource-ID-of-cluster-service-instance*," specify the resource ID of the cluster service instance that was registered in Step 5 for each node. Specify the resource IDs of the cluster service instances in the "Machines" sequence that was set in "(5)" of "CRM.htm."

   Example) If the resource ID of the cluster service is "100," and the Machines sequence is as follows: The resource ID of the cluster service instance in the narcissus node is "101" and the resource ID of the cluster service instance in the sweetpea node is "102."

   ```
   Machines[0]  narcissusRMS
   ```

```
Machines[1]    sweetpeaRMS
```

```
# /etc/opt/FJSVcluster/bin/clsetinsorder -s 100 -r 101,102 -S
```

8. Check if cluster services and cluster service instance resources are registered using the clgettree(1) command as follows:

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

Check the following:

- The cluster services are displayed.

  The cluster service resource class is "Standby" or "Scalable".

- The cluster service instances are displayed.

  The cluster service instance resource class is "AndBinding".

- The order of operating cluster service instances is displayed.

  The order is "()" in the cluster service instance resources. The first cluster service instance will be "RUN", and the remaining instances will be "WAIT".

Example) In the following example, the cluster service resource ID is "100", and the cluster service instance resource IDs are "101 and "102".

```
# /etc/opt/FJSVcluster/bin/clgettree -s
Standby 100 generic STOP REJECT
       AndBinding 101 genelic STOP UNDEFINED(RUN) narcissus
       AndBinding 102 genelic STOP UNDEFINED(WAIT) sweetpea
```

G.2 Associating Resources to a Cluster Service Instance

The SynfinityCluster-compatible cluster service instance needs to be associated with the following product:

- Interstage Traffic Director

Note

For products other than the above, the procedure for associating the product to a cluster service instance is unnecessary.

The procedure for associating the above product with a SynfinityCluster-compatible cluster service instance is described below.

Operation Procedure:

1. Log in with the system administrator authority.

   Log in with the system administrator authority to one of the nodes of the cluster system.

Note

Perform Step 2 only if userApplication was set up with the CUI and IP address takeover was set during userApplication setup. If userApplication was set up with the GUI or if IP address takeover was not set during userApplication setup, Step 2 is unnecessary.

2. Register the IP address takeover resources.

   - Execute the "clgettree(1)" command in the format shown below, and check the resource IDs of the PublicLAN resources in all nodes that execute IP address takeover. The resource class of PublicLAN resources is "Ethernet."

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

Example) In the following case, the resource ID of the PublicLAN resource in the narcissus node is "25" and the resource ID of the PublicLAN resource in the sweetpea node is "26."

```
# /etc/opt/FJSVcluster/bin/clgettree
Cluster 1 /
          Domain 2 RC9
          Shared 7 SHD_RC9
          Node 3 narcissus ON
                  Psd 21 hme1 ON
                  Psd 22 hme2 ON
                  Ethernet 25 hme0 ON
                  DISK 19 c0t0d0 UNKNOWN
                  SytemState2 40 FJSVndbs UNKNOWN
          Node 5 sweetpea ON
                  Psd 23 hme1 ON
                  Psd 24 hme2 ON
                  Ethernet 26 hme0 ON
                  DISK 20 c0t0d0 UNKNOWN
                  SytemState2 41 FJSVndbs UNKNOWN
```

- Execute the "clcreate-iprsc(1M)" command in the format shown below, and register the IP address takeover resources.

```
# /etc/opt/FJSVcluster/sys/clcreate-iprsc -p resource-ID-of-PublicLAN-resource...(*1)
takeover-IP-address(*2)
```

*1For "*resource-ID-of-PublicLAN-resource...*," specify the resource IDs of the PublicLAN resources of all nodes that were confirmed with the previous "clgettree(1)" command.

*2For "*takeover-IP-address*," specify the takeover IP address that was specified during userApplication setup.

Example) If the resource IDs of the PublicLAN resources are "25," "26" and the takeover IP address is "192.168.246.1"

```
# /etc/opt/FJSVcluster/sys/clcreate-iprsc -p 25,26 192.168.246.1
```

3. Associate the resources to cluster service instances.

   Associate the following resources to cluster service instances:

   - IP address takeover resources (*1)

     *1If IP address takeover resources were registered in Step 2, associated them to cluster service instances.

   - "Interstage Traffic Director" resources

   1. Execute the clgettree(1) command in the format shown below, and check the resource IDs of the resources to be associated to cluster service instances. The resource class of IP address takeover resources is "IP_Address." The resource name of the Interstage Traffic Director resource is "FJSVndbs."

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

   Example) For the "clgettree(1) command execution example" described below, the resource IDs for the IP address takeover and Interstage Traffic Director resources are as follows:

Node	Resource ID of IP address takeover resource	Resource ID of Interstage Traffic Director resource
narcissus	27	40
sweetpea	28	41

   [Execution example of clgettree(1) command]

```
# /etc/opt/FJSVcluster/bin/clgettree
Cluster 1 /
          Domain 2 RC9
          Shared 7 SHD_RC9
                  SHD_Host 29 192.168.246.1 UNKNOWN
                          IP_Address 27 hme0:1 UNKNOWN narcissus
                          IP_Address 28 hme0:1 UNKNOWN sweetpea
```

```
                        Node 3 narcissus ON
                                Psd 21 hme1 ON
                                Psd 22 hme2 ON
                                Ethernet 25 hme0 ON
                                        IP_Address 27 hme0:1 UNKNOWN
                                DISK 19 c0t0d0 UNKNOWN
                                SytemState2 40 FJSVndbs UNKNOWN
                        Node 5 sweetpea ON
                                Psd 23 hme1 ON
                                Psd 24 hme2 ON
                                Ethernet 26 hme0 ON
                                        IP_Address 28 hme0:1 UNKNOWN
                                DISK 20 c0t0d0 UNKNOWN
                                SytemState2 41 FJSVndbs UNKNOWN
```

2. Using the format shown below, execute the "clrepparent(1M)" command for each node in the userApplication configuration, and associate the IP address takeover and Interstage Traffic Director resource to cluster service instances.

```
# /etc/opt/FJSVcluster/sys/clrepparent -s resource-IDs-of-IP-address-takeover-and-
Interstage-Traffic Director-resources(*1) -t resource-ID-of-cluster-service-instance(*2)
```

*1Specify the resource IDs of the IP address takeover and Interstage Traffic Director resources that were obtained with the clgettree(1) command in each node.

*2

If userApplication was set up with the CUI

Specify the resource ID of the cluster service instance that was registered in each node according to Step 5 "Register the cluster service instance" in "G.1 Registering a Cluster Service."

If userApplication was set up with the GUI."

Execute the "clgettree(1)" command in the format shown below, and check the resource ID of the cluster service instance in each node. The resource class for cluster service instances is "AndBinding." The node name to which the cluster service instance belongs is displayed at the end of the line for the "AndBinding" resource class.

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

In the following example, the resource ID of the cluster service instance in the narcissus node is "101," and the resource ID of the cluster service instance in the sweetpea node is "102."

```
  # /etc/opt/FJSVcluster/bin/clgettree -s
  Standby 68 Service1 STOP AVAILABLE
     AndBinding 101 Instance1_Of _narcissus STOP STOP(RUN) narcissus
     AndBinding 102 Instance1_Of _sweetpea STOP STOP(RUN) sweetpea
```

Example) When the resource ID of the cluster service instance in the narcissus node is "101" and the resource ID of the cluster service instance in the sweetpea node is "102" and the following resources are registered in each node:

Node	Resource ID of IP address takeover resource	Resource ID of Interstage Traffic Director resource
narcissus	27	40
sweetpea	28	41

```
# /etc/opt/FJSVcluster/sys/clrepparent -s 27,40 -t 101
# /etc/opt/FJSVcluster/sys/clrepparent -s 28,41 -t 102
```

G.3 Deleting a Cluster Service

This section describes the procedure to be performed in the following cases:

- If userApplication is deleted

- If the following product is installed:

    - Interstage Traffic Director

After carrying out the procedure indicated in this section, perform the "G.1 Registering a Cluster Service" and "G.2 Associating Resources to a Cluster Service Instance" operations in the following cases:

- If the number of nodes configuring userApplication is changed

- If the ApplicationName of userApplication is changed

- If a new IP address takeover is to be executed or if an IP address takeover is to be discontinued

- If a takeover IP address is changed

Operation Procedure:

1. Log in with the system administrator authority.

   Log in with the system administrator authority to one of the nodes of the cluster system.

2. Check the resource ID of the cluster service.

   Execute the "clgettree(1)" command in the format shown below, and check the cluster service resource. The resource class of the cluster service is "Standby" or "Scalable." The cluster service is output in the first line.

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

   Example) In the example below, the resource ID of the cluster service is "68."

```
# /etc/opt/FJSVcluster/bin/clgettree -s
Standby 68 Service1 STOP AVAILABLE
     AndBinding 101 Instance1_Of _narcissus STOP STOP(RUN) narcissus
     AndBinding 102 Instance1_Of _sweetpea STOP STOP(RUN) sweetpea
```

3. Delete the cluster service resource.

   Execute the "cldelrsc(1M)" command in the following format, and delete the cluster service resource.

```
# /etc/opt/FJSVcluster/bin/cldelrsc -r resource-ID-of-cluster-service(*1)
```

   *1 For "resource-ID-of-cluster-service," specify the resource ID of the cluster service that was obtained in Step 2.

   Example) To delete the cluster service that has resource ID "68."

```
# /etc/opt/FJSVcluster/bin/cldelrsc -r 68
```

4. Delete the IP address takeover resources.

   - Execute the "clgettree(1)" command in the format shown below, and check the resource IDs of the IP address takeover resources. The resource classes of the IP address takeover resources are "SHD_Host" and "IP_Address." If resources of the "SHD_Host" and "IP_Address" resource classes are not output, the execution of the subsequent steps is unnecessary because there are no IP address takeover resources.

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

   Example) In the following case, the resource IDs of the IP address takeover resources are "29," "27" and "28."

```
# /etc/opt/FJSVcluster/bin/clgettree
Cluster 1 /
            Domain 2 RC9
            Shared 7 SHD_RC9
                    SHD_Host 29 192.168.246.1 UNKNOWN
                            IP_Address 27 hme0:1 UNKNOWN narcissus
                            IP_Address 28 hme0:1 UNKNOWN sweetpea
            Node 3 narcissus ON
```

```
                        Psd 21 hme1 ON
                        Psd 22 hme2 ON
                        Ethernet 25 hme0 ON
                                IP_Address 27 hme0:1 UNKNOWN
                        DISK 19 c0t0d0 UNKNOWN
                        SytemState2 40 FJSVndbs UNKNOWN
                Node 5 sweetpea ON
                        Psd 23 hme1 ON
                        Psd 24 hme2 ON
                        Ethernet 26 hme0 ON
                                IP_Address 28 hme0:1 UNKNOWN
                        DISK 20 c0t0d0 UNKNOWN
                        SytemState2 41 FJSVndbs UNKNOWN
```

- If the previous procedure indicates that there is an IP address takeover resource, execute the "cldelrsc(1)" command in the format shown below, and delete the IP address takeover resource. Execute this step only if there is an IP address takeover resource.

```
# /etc/opt/FJSVcluster/bin/cldelrsc -r resource-ID-of-IP-address-takeover-resource(*1)
```

*1For "resource-ID-of-IP-address-takeover-resource," specify all resource IDs of IP address takeover resources that were obtained with the previous "clgettree(1)" command.

Example) If the resource IDs of the IP address takeover resources are "29," "27" and "28."

```
# /etc/opt/FJSVcluster/bin/cldelrsc -r 29,27,28
```

Appendix H Using PRIMECLUSTER in Oracle VM Server for SPARC Environments

This appendix provides explanations on applying PRIMECLUSTER in Oracle VM Server for SPARC (formerly referred to as Logical Domains or LDoms) environments.

- Cluster systems in Oracle VM Server for SPARC environments

- Extent of PRIMECLUSTER support in Oracle VM Server for SPARC environments

- Procedure for configuration of PRIMECLUSTER in Oracle VM Server for SPARC environments

- Precautions on using cluster systems in Oracle VM Server for SPARC environments

- Maintenance of cluster systems in Oracle VM Server for SPARC environments

- Recommended configuration

🗃️ See
....
For further details on Oracle VM Server for SPARC, please refer to the "Oracle VM Server for SPARC Guide".
....

H.1 Cluster Systems in Oracle VM Server for SPARC Environments

This section provides explanations on cluster systems in Oracle VM Server for SPARC environments.

H.1.1 Oracle VM Server for SPARC

Oracle VM Server for SPARC is a virtual hardware environment which runs by using Hypervisor. It splits a platform into multiple server environments, each of which can run their separate operating systems.
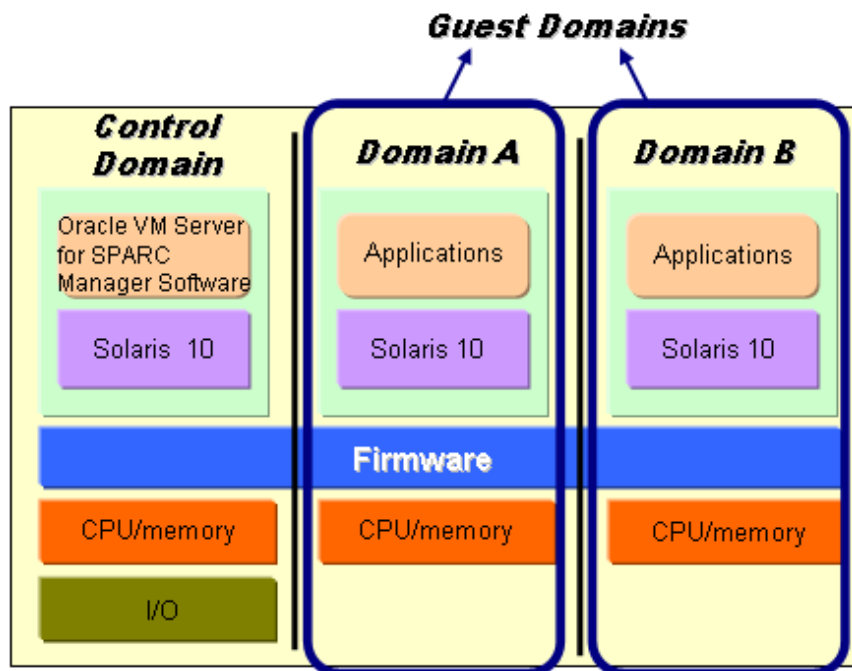
H.1.2 Cluster Systems in Oracle VM Server for SPARC Environments

In Oracle VM Server for SPARC environments, if an error occurs in the control domain, virtual in- and output are suspended, and any applications in guest domains cannot run anymore.

PRIMECLUSTER applied to the control domain monitors below errors that cause applications in guest domains to stop.

When such an error occurs, the guest domains in the standby system are booted, thus securing a highly reliable guest domain environment.

- OS failure in the control domain

- Hardware (network/disk) failure in the control domain

- Failure of shared disks and their transmission paths
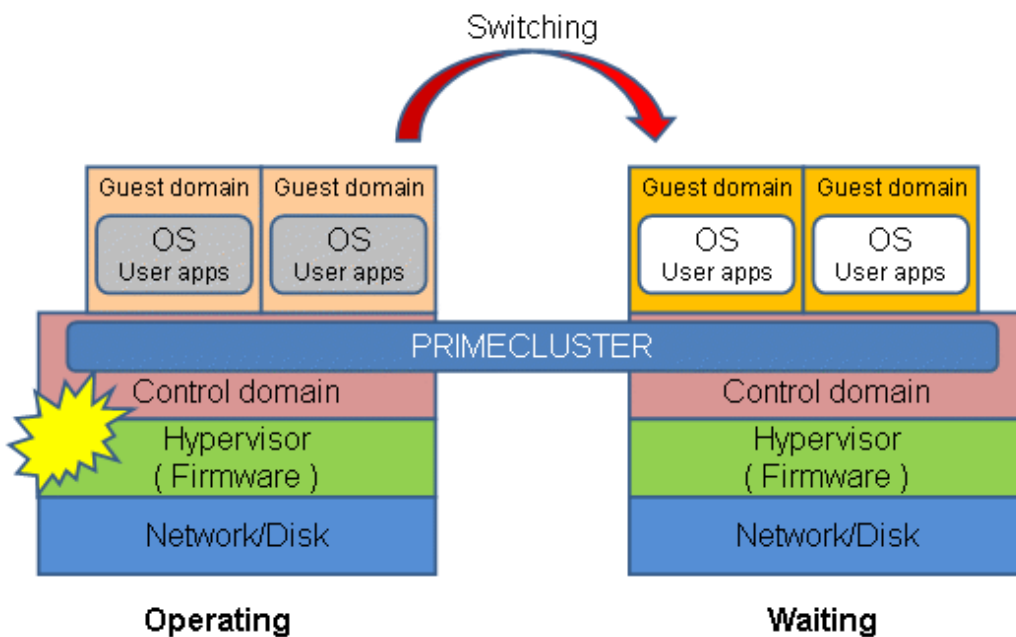
 Note

....

PRIMECLUSTER monitors and controls guest domains from control domain. It does not monitor inside guest domains, including operating systems and applications in guest domains.

For details on statuses that are monitored by PRIMECLUSTER, please refer to "H.2.2 Supported Monitoring Functions". Also note that there are a number of precautions on operating cluster systems in Oracle VM Server for SPARC environments. For details, please refer to "H.4 Precautions on Using Cluster Systems in Oracle VM Server for SPARC Environments".

....

H.2 Extent of PRIMECLUSTER Support in Oracle VM Server for SPARC Environments

This section explains the scope of PRIMECLUSTER support in Oracle VM Server for SPARC environments.

H.2.1 Supported Configurations

With PRIMECLUSTER in Oracle VM Server for SPARC environments, the following configurations are supported:

- Operation modes of cluster system

  - 1:1 setup (one node operational, one on standby)

  - Mutual standby setup

- Cluster resource types configurable in control domain

  - GDS resource

  - GLS resource

  - Cmdline resource

- Duplication method supported by GLS
  In control domain clusters, the only duplication method supported by GLS is NIC switching.

Note

- GLS must be installed in both control and guest domains.

- Guest domains must be configured as all disk/network in- and output from/to guest domains are done through control domain.

- Multiple guest domains on the same control domain cannot share a GDS shared class. When configuring multiple guest domains, please create shared classes separately for each guest domain.

H.2.2 Supported Monitoring Functions

With PRIMECLUSTER in Oracle VM Server for SPARC environments, the following monitoring functions are supported:

- Hardware monitoring in the control domain

- Monitoring of shared disks

- Monitoring of networks

- Monitoring of guest domain statuses
  PRIMECLUSTER monitors if guest domains are "active", "bound", or "inactive".
  For further details on the statuses, please refer to the "Oracle VM Server for SPARC Guide".

Note

With PRIMECLUSTER in Oracle VM Server for SPARC environments, the following monitoring purposes are not supported:

- Statuses of guest domain operating systems

- User applications in guest domains

H.3 Procedure for Configuration of PRIMECLUSTER in Oracle VM Server for SPARC Environments

Below section explains the procedure for configuring control domain clusters in Oracle VM Server for SPARC environments.

H.3.1 Software Installation and Configuration of Cluster Environment

Install the Solaris software and the software required for PRIMECLUSTER in the control and guest domains of all nodes. Then make the OS and hardware settings required for actual implementation and operation.

For further details on configuring Oracle VM Server for SPARC, please refer to the "Oracle VM Server for SPARC Guide".
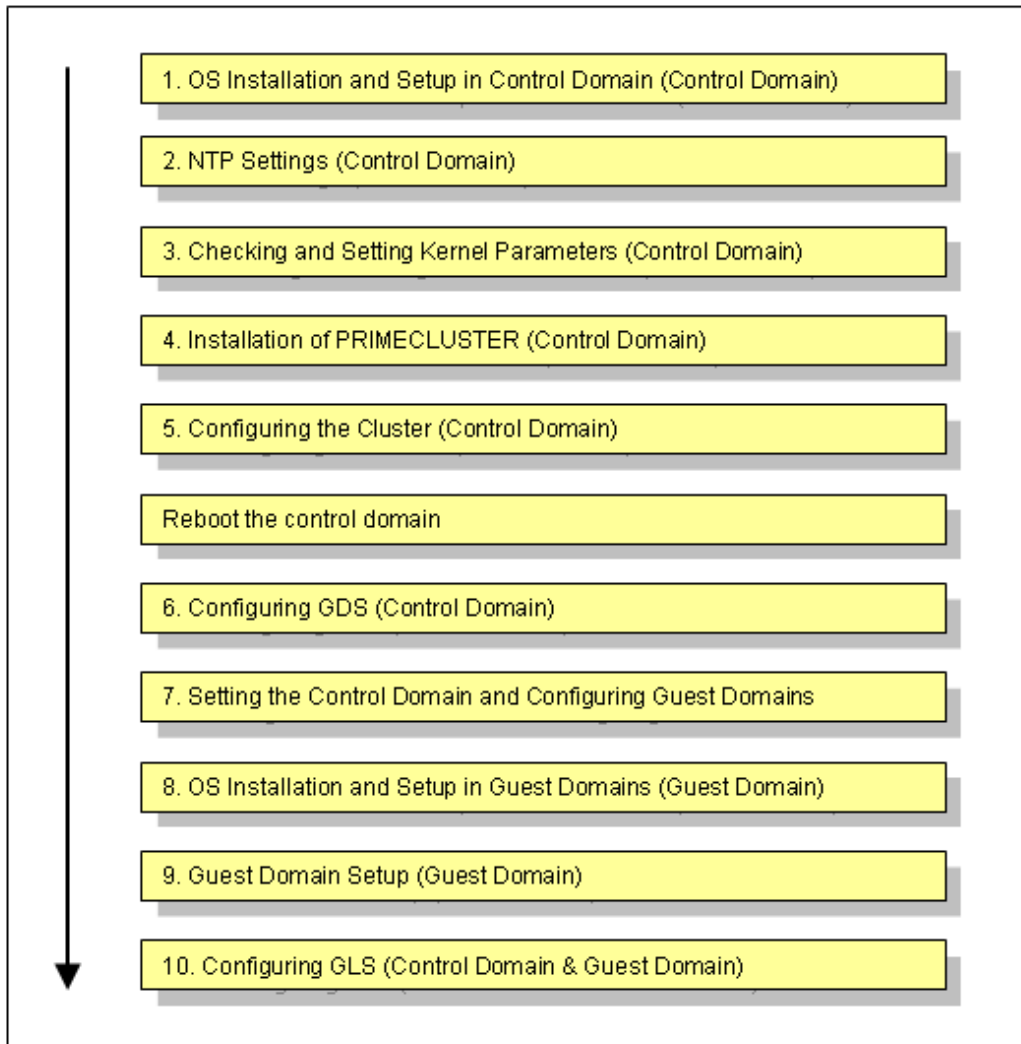
Carry out the following work as necessary.

1. OS Installation and Setup in Control Domain (Control Domain)

   Please complete all preparations for configuring your Oracle VM Server for SPARC environment by installing the operating system and the Oracle VM Server for SPARC package in the control domain. Referring to the "Oracle VM Server for SPARC Guide" for the details, complete all preparations through to the files for the virtual disk server devices.

   ### Note

   To enable SST settings, refer to the "PRIMECLUSTER Security Guide" and open the necessary ports.

   Disk-Related Settings

   For using disk-sharing devices, installation and setup of related software products are required.
   Please carry out these settings in the control domain before installing PRIMECLUSTER.
   For details, refer to "3.2.2 Setting Up Disk Units".

2. NTP Settings (Control Domain)

   This setup serves to synchronize the clocks of every node in the cluster system. Please make sure to make this setting when configuring a cluster.
   Please carry out these settings in the control domain before installing PRIMECLUSTER.

3. Checking and Setting Kernel Parameters (Control Domain)

For running PRIMECLUSTER-related software, the OS kernel parameters need to be adjusted to the environment.
Please carry out these settings in the control domain before rebooting the system after "Step 4 Installation of PRIMECLUSTER (Control Domain)".
For details, refer to "3.2.3 Checking the Kernel Parameters".

4. Installation of PRIMECLUSTER (Control Domain)

For details on installing PRIMECLUSTER, please refer to "3.1 PRIMECLUSTER Installation".
After installing PRIMECLUSTER, please make the settings for the OBP auto-boot? variable.

Setting the OBP auto-boot? Variable

In order to prevent the control domain from rebooting automatically after an enforced stop, set the OBP auto-boot? variable to "false". Otherwise, the suspended control domain may be automatically rebooted, possibly causing data loss due to restored virtual in- and output from guest domains that were stopped.

Example:

```
{1} ok setenv auto-boot? False
              auto-boot? = false
          {1} ok reset-all
```

After setting the OBP auto-boot? variable, reboot the node to make the OBP settings effective.

5. Configuring the Cluster (Control Domain)

Referring to "Chapter 4 Preparation Prior to Building a Cluster" and "Chapter 5 Building a Cluster", make the initial cluster settings in the control domain.

6. Configuring GDS (Control Domain)

Referring to "6.2 Initial GDS Setup", make the GDS settings in the control domain.
For further details on how to make GDS settings, please refer to the "PRIMECLUSTER Global Disk Services 4.2 Configuration and Administration Guide".

Note

- For enhanced availability, we recommend allocating virtual disks to mirror volumes if they are to be used as guest domain system disks.
  When configuring GDS, create a root-class or local-class mirror volume on which the files for the virtual disks shall be located. Do not deploy any files for virtual disks on shared-class volumes.

- When using multiple guest domains, please create shared classes separately for each guest domain.

7. Setting the Control Domain and Configuring Guest Domains

Referring to the "Oracle VM Server for SPARC Guide", complete all configurations for the control and guest domains.
After that, create the virtual network (administrative/user LAN).

Sample Configuration of Control and Guest Domains

As the virtual disk used as the system disk, specify a file in the format /LDoms/Vol1/vdisk0.img.
As the shared-class volume to be used as the switching disk in the cluster, specify a block device like /dev/sfdsk/class0001/dsk/volume0001.
In such cases, make sure to specify the "slice" option.

Example: Configuration Script for Control Domain

```
# Defining virtual disk server.
ldm add-vdiskserver primary-vds0 primary
# Defining virtual disk server and specifying hardware.
ldm add-vdiskserverdevice /LDoms/Vol1/vdisk0.img vo1@primary-vds0
ldm add-vdiskserverdevice options=slice /dev/sfdsk/class0001/dsk/
volume0001 vol2@primary-vds0
```

```
ldm add-vdiskserverdevice options=slice /dev/sfdsk/class0001/dsk/
volume0002 vol3@primary-vds0
```

Example: Configuration Script for Guest Domain

```
VDISK0=vol1@primary-vds0
VSHDISK0=vol2@primary-vds0
VSHDISK1=vol3@primary-vds0
[...]
ldm add-vdisk vdisk0 $VDISK0 $DOMAIN
ldm add-vdisk timeout=360 vshdisk0 $VSHDISK0 $DOMAIN
ldm add-vdisk timeout=360 vshdisk1 $VSHDISK1 $DOMAIN
```

Creation of Virtual Network

For administrative/user LAN communication via virtual network (vnet) devices in the guest domain, create a virtual switch service (vsw) in the control domain and add the vnet devices to the guest domain.

Besides this, for creating network redundancy and for switching clusters after a network failure in environments using virtual network devices for user LAN communication of guest domains, GLS must be installed in both control and guest domains. For using GLS, please create a virtual switch service and a virtual network device for multiple physical network devices.

See

For details on how to create a virtual network, please refer to the "Oracle VM Server for SPARC Guide".

Note

In control domain clusters, the only duplication method supported by GLS is NIC switching.

8. OS Installation and Setup in Guest Domains (Guest Domains)

With PRIMECLUSTER, in order to enable cluster switching, an operating system must be installed in each guest domain. Referring to the "Oracle VM Server for SPARC Guide", complete installation of operating systems in all guest domains.
After installing the operating systems, install GLS in the guest domains.

See

For details on installing GLS, please refer to the "PRIMECLUSTER GLS Installation Guide".

Note

The guest domains recognize the GDS volume created in the control domain as a virtual disk. At this point, the number "0" is allocated as the slice, rendering the device name into a format like, e.g., /dev/dsk/c0d1s0. As this name is different from the device name in the control domain, please be careful not to confuse the disk to be used with another one.

```
# prtvtoc /dev/rdsk/c0d1s0
* /dev/rdsk/c0d1s0 (volume "vdisk") partition map
*
* Dimensions:
*     512 bytes/sector
*    2048 sectors/track
*      16 tracks/cylinder
*   32768 sectors/cylinder
*     874 cylinders
*     874 accessible cylinders
*
```

```
* Flags:
*   1: unmountable
*  10: read-only
*
* Unallocated space:
*       First     Sector    Last
*       Sector    Count     Sector
*            0         1          0
*
*                          First     Sector    Last
* Partition  Tag  Flags    Sector    Count     Sector   Mount Directory
  0     0    00              1  28639232  28639232
# newfs /dev/dsk/c0d1s0
 ...
# vi /etc/vfstab
```

9. Guest Domain Setup (Guest Domain)

   In order to reboot the OS automatically after switching guest domains, set the OBP auto-boot? variable to "true".

   Example:

```
{1} ok setenv auto-boot? true
           auto-boot? = true
        {1} ok reset-all
```

   After setting the OBP auto-boot? variable, reboot the guest domain to make the OBP settings effective.

10. Configuring GLS (Control Domain & Guest Domain)

    For information on the initial GLS setup, please refer to "Chapter 5 Operation on Cluster System" in the "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function".

    This section uses below example to explain the setup procedure for GLS (redundant line control function) for a 1:1 operational/standby setup of a control domain cluster with IPv4 NIC switching in an Oracle VM Server for SPARC environment.

Note

In control domain clusters, the only duplication method supported by GLS is NIC switching.

The setting values for each category correspond to those in the "GLS Setup Worksheet".
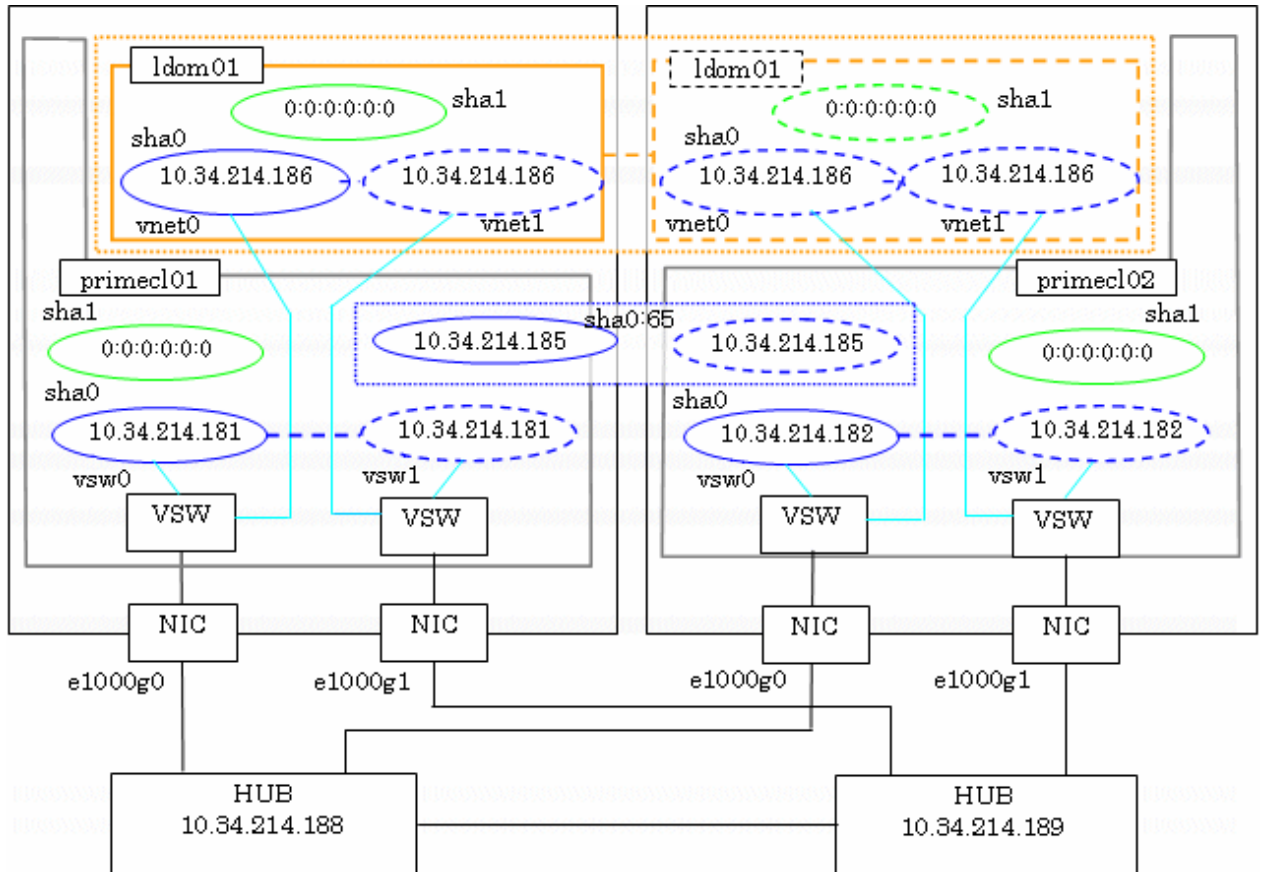
In this sample configuration, control domains primecl01 and primecl02 are forming a cluster, whereas guest domain ldom01 runs as the cluster resource. The control and guest domains share the same user LAN network lines via a virtual switch service.

When user LAN transmissions on the operational node are disabled, GLS in the control domains detects a failure in the transmission route and switches clusters. When clusters are switched, the guest domains in the operational node are stopped, and, instead, the guest domains in the standby node are started.

The IP addresses of each control domains (10.34.214.181 for primecl01, 10.34.214.182 for primecl02) can be accessed independently of their operational statuses; however, setting takeover IP addresses allows you to automatically access the control domains currently in use, even if you are not sure which one it is.

Operating Procedure

- For [HOST-primecl01] as the operational node

1. Make the system settings.

    1. Define the IP addresses and host names to be used in the /etc/inet/hosts file.

    ```
    10.34.214.185    takeoverIP    # virtual IP
    10.34.214.181    primecl01     # physical IP of primecl01
    10.34.214.182    primecl02     # physical IP of primecl02
    10.34.214.186    ldom01        # physical IP of ldom01
    10.34.214.188    swhub01       # IP of primary monitoring target hub
    10.34.214.189    swhub02       # IP of secondary monitoring target hub
    ```

    2. Define the subnet masks in the /etc/inet/netmasks file.

    ```
    10.34.214.0      255.255.255.0
    ```

3. Add the host names from above definitions to the /etc/hostname.vsw0 file.

   Contents of /etc/hostname.vsw0:

   primecl01

4. Create the virtual network to be used in the user LAN.

```
# ldm add-vsw net-dev=e1000g0 primary-vsw0 primary
# ldm add-vsw net-dev=e1000g1 primary-vsw1 primary
# ldm add-vnet vnet0 primary-vsw0 ldom01
# ldm add-vnet vnet1 primary-vsw1 ldom01
# touch /reconfigure
```

2. Reboot the system.

   Execute below command to reboot the system. After rebooting, check that vsw0 is active.

```
# /usr/sbin/shutdown -y -i6 -g0
```

3. Make the virtual interface settings.

   The underlined portion specifies the physical IP address.

```
# /opt/FJSVhanet/usr/sbin/hanetconfig create -n sha0 -m d -i 10.34.214.185 -e 10.34.214.181
-t vsw0,vsw1
```

   After executing above command, check that the settings are correct.

```
# /opt/FJSVhanet/usr/sbin/hanetconfig print
```

4. Make the settings for HUB monitoring.

   The underlined portions specify the IP addresses of the HUB to be monitored.

```
# /opt/FJSVhanet/usr/sbin/hanetpoll create -n sha0 -p 10.34.214.188,10.34.214.189 -b off
```

   After executing above command, check that the settings are correct.

```
# /opt/FJSVhanet/usr/sbin/hanetpoll print
```

5. Make the settings for the GLS standby patrol function.

```
# /opt/FJSVhanet/usr/sbin/hanetconfig create -n sha1 -m p -a 0:0:0:0:0:0 -t sha0
```

   After executing above command, check that the settings are correct.

```
# /opt/FJSVhanet/usr/sbin/hanetconfig print
```

6. Create a takeover IP address (takeover virtual interface).

```
# /opt/FJSVhanet/usr/sbin/hanethvrsc create -n sha0
```

   After executing above command, check that the settings are correct.

```
# /opt/FJSVhanet/usr/sbin/hanethvrsc print
```

- For [HOST-primecl02] as the standby node

  1. Make the system settings.

     1. Define the IP addresses and host names to be used in the /etc/inet/hosts file.

        The contents to be defined are the same as for primecl01.

     2. Add the host names from above definitions to the /etc/hostname.vsw0 file.

        Contents of /etc/hostname.vsw0:

        primecl02

3. Define the IP addresses and host names to be used in the /etc/inet/hosts file.

    The contents to be defined are the same as for primecl01.

4. Define the subnet masks in the /etc/inet/netmasks file.

    The contents to be defined are the same as for primecl01.

5. Add the host names from above definitions to the /etc/hostname.vsw0 file.

    Contents of /etc/hostname.vsw0:

    primecl02

6. Create the virtual network to be used in the user LAN.

```
# ldm add-vsw net-dev=e1000g0 primary-vsw0 primary
# ldm add-vsw net-dev=e1000g1 primary-vsw1 primary
# ldm add-vnet vnet0 primary-vsw0 ldom01
# ldm add-vnet vnet1 primary-vsw1 ldom01
# touch /reconfigure
```

2. Reboot the system.

    Execute below command to reboot the system. After rebooting, check that vsw0 is active.

```
# /usr/sbin/shutdown -y -i6 -g0
```

3. Make the virtual interface settings.

    The underlined portion specifies the physical IP address.

```
# /opt/FJSVhanet/usr/sbin/hanetconfig create -n sha0 -m d -i 10.34.214.185 -e 10.34.214.182
-t vsw0,vsw1
```

    After executing above command, check that the settings are correct.

```
# /opt/FJSVhanet/usr/sbin/hanetconfig print
```

4. Make the settings for HUB monitoring.

    The underlined portions specify the IP addresses of the HUB to be monitored.

```
# /opt/FJSVhanet/usr/sbin/hanetpoll create -n sha0 -p 10.34.214.188,10.34.214.189 -b off
```

    After executing above command, check that the settings are correct.

```
# /opt/FJSVhanet/usr/sbin/hanetpoll print
```

5. Make the settings for the GLS standby patrol function.

```
# /opt/FJSVhanet/usr/sbin/hanetconfig create -n sha1 -m  -a 0:0:0:0:0:0 -t sha0
```

    After executing above command, check that the settings are correct.

```
# /opt/FJSVhanet/usr/sbin/hanetconfig print
```

6. Create a takeover IP address (takeover virtual interface).

```
# /opt/FJSVhanet/usr/sbin/hanethvrsc create -n sha0
```

    After executing above command, check that the settings are correct.

```
# /opt/FJSVhanet/usr/sbin/hanethvrsc print
```

- For guest domain [HOST-ldom01]

Make the GLS settings for guest domains on one node after another, first on the nodes that are designated to be operational, and then on the nodes designated to be standby nodes, making sure not to assign duplicate IP addresses.

1. Make the system settings.

   1. Define the IP addresses and host names to be used in the /etc/inet/hosts file.

   ```
   10.34.214.186   ldom01        # physical IP of ldom01
   10.34.214.188   swhub1        # IP of primary monitoring target hub
   10.34.214.189   swhub2        # IP of secondary monitoring target hub
   ```

   2. Define the subnet masks in the /etc/inet/netmasks file.

   ```
   10.34.214.0     255.255.255.0
   ```

   3. Add the host names from above definitions to the /etc/hostname.vnet0 file.

   Contents of /etc/hostname.vnet0:

   ldom01

2. Reboot the system.

   Execute below command to reboot the system. After rebooting, check that vnet0 is active.

   ```
   # /usr/sbin/shutdown -y -i6 -g0
   ```

3. Make the virtual interface settings.

   The underlined portion specifies the physical IP address.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetconfig create -n sha0 -m e -i 10.34.214.186 -t vnet0,vnet1
   ```

   After executing above command, check that the settings are correct.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetconfig print
   ```

4. Make the settings for HUB monitoring.

   The underlined portions specify the IP addresses of the HUB to be monitored.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetpoll create -n sha0 -p 10.34.214.188,10.34.214.189 -b off
   ```

   After executing above command, check that the settings are correct.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetpoll print
   ```

5. Make the settings for the GLS standby patrol function.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetconfig create -n sha1 -m p -a 0:0:0:0:0:0 -t sha0
   ```

   After executing above command, check that the settings are correct.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetconfig print
   ```

6. Start hub monitoring.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetpoll on
   ```

7. Activate the virtual interface.

   ```
   # /opt/FJSVhanet/usr/sbin/strhanet
   ```

H.3.2 Configuration of Cluster Applications

In control domain clusters in Oracle VM Server for SPARC environments, cluster applications for the following resources can be configured.

Configure the cluster application with one Cmdline resource, at least one GLS resource and at least one GDS resource.

For details on cluster applications and Cmdline resources, please refer to "Chapter 6 Building Cluster Applications".

- Cmdline resource for monitoring guest domain statuses

Starts, stops and monitors the corresponding guest domain.

- GLS resource

Monitors the networks used by the corresponding guest domain via virtual in- and output from the control domain.

- GDS resource

Starts, stops, and monitors the volumes used by the corresponding guest domain via virtual in- and output from the control domain.
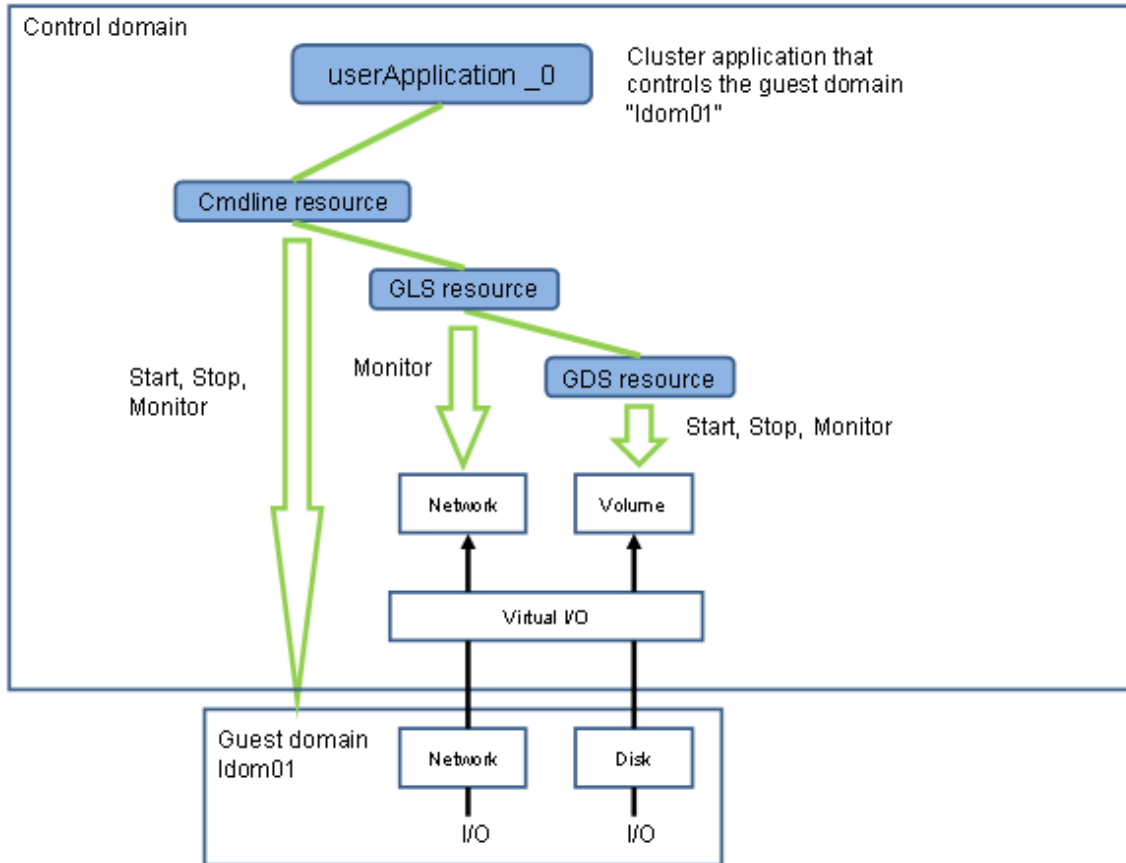
The procedure for configuring cluster applications in the control domain cluster is as follows:

1. Creation of Cmdline Resources

2. Setup of Cluster Applications

H.3.2.1 Creation of Cmdline Resources for Monitoring Guest Domain Statuses

Set up Online, Offline and Check scripts so that the corresponding guest domain starts/stops in conjunction with userApplication.
Set the commands to start the corresponding guest domain to the Online script. Set the commands to stop the corresponding guest domain to the Offline script. In the Check script, set the commands to verify the corresponding guest domain's ability to start its OS.

Cmdline resource that controls the guest domain "ldom01"
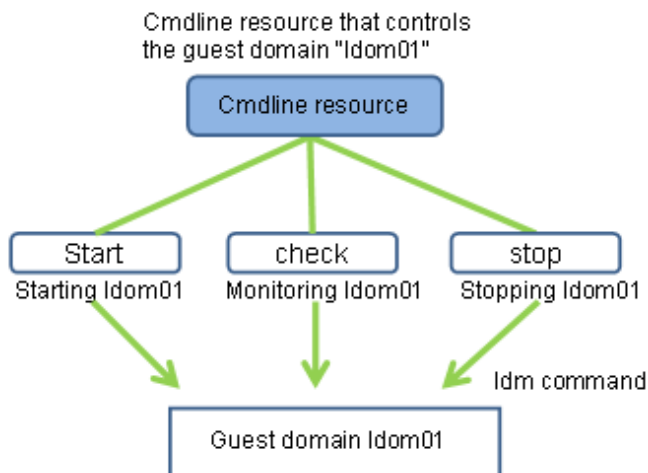
Sample Scripts

This section presents sample settings for Online (Start), Offline (Stop) and Check scripts.

Set a Cmdline resource with the scripts to userApplication. For details, refer to "6.6.1.1 Creating Cmdline Resources".

- Start script: starts the corresponding guest domains. Returns "0" when execution was successful, or any other value when failed.

```
   #!/bin/sh
#
#  Sample online script for a guest domain
#      Exit with  0: Success   1: Fail
#
#  Copyright(c) 2008 FUJITSU LIMITED. All rights reserved.
#

LDOMSNAME=$1
LDM="/opt/SUNWldm/bin/ldm"

$LDM bind-domain $LDOMSNAME

$LDM start-domain $LDOMSNAME

# Exits on "0" (success), only when logical domain is active
$LDM list-domain $LDOMSNAME | tail -1 | awk '{print $2}' | grep "^active" && exit 0

exit 1
```

- Stop script: stops the corresponding guest domains. Returns "0" when execution was successful, or any other value when failed.

```
   #!/bin/sh
#
#  Sample offline script for a guest domain
#      Exit with  0: Success   1: Fail
#
#  Copyright(c) 2008 FUJITSU LIMITED. All rights reserved.
#

LDOMSNAME=$1
LDM="/opt/SUNWldm/bin/ldm"

$LDM stop-domain $LDOMSNAME
[ 0 -ne $? ] && $LDM stop-domain -f $LDOMSNAME

$LDM unbind-domain $LDOMSNAME
```

```
# Exits on "0" (success), only when logical domain is inactive
$LDM list-domain $LDOMSNAME | tail -1 | awk '{print $2}' | grep "^inactive"
&& exit 0

exit 1
```

- Check script: monitors "active" statuses of corresponding guest domains. Returns "0" when domain is active, or any other value otherwise.

```
     #!/bin/sh
#
#  Sample check script for a guest domain
#       Exit with  0: Online   1: Offline
#
#  Copyright(c) 2008 FUJITSU LIMITED. All rights reserved.
#

LDOMSNAME=$1
LDM="/opt/SUNWldm/bin/ldm"

# Exits on "0" (success), only when logical domain is active
$LDM list-domain $LDOMSNAME | tail -1 | awk '{print $2}' | grep "^active" &&
exit 0

exit 1
```

H.3.2.2 Setup of Cluster Applications

Make the settings for Cmdline resource, Gds resource and Gls resource cluster applications.

For details on Gds and Gls resource settings, please refer to "Creating Gds Resources" and "Creating Gls Resources" in "6.6.1 Setting Up Resources". For details on cluster application settings, please refer to "6.6 Setting Up Cluster Applications".

H.4 Precautions on Using Cluster Systems in Oracle VM Server for SPARC Environments

For operating cluster systems in Oracle VM Server for SPARC environments, please be careful about the following matters:

H.4.1 During Installation

Installation of PRIMECLUSTER

After installing PRIMECLUSTER, in order to prevent the control domain from rebooting automatically after an enforced stop, set the OBP auto-boot? variable to "false". For details, please refer to "4. Installation of PRIMECLUSTER (Control Domain)" in "H.3.1 Software Installation and Configuration of Cluster Environment".

Setting of Guest Domain

Operating systems and applications in guest domains cannot be controlled by PRIMECLUSTER in the control domain. Set up your guest domains so that operating systems and applications are automatically booted in extension of each guest domain's startup. For details on how to automatically start operating systems in guest domains, please refer to "9. Guest Domain Setup (Guest Domain)" in "H.3.1 Software Installation and Configuration of Cluster Environment".

H.4.2 During Operation

Controlling Guest Domains

Never use the ldm(1M) command immediately to start or stop a guest domain or to bind its resources.

Guest domains have to be started and stopped from PRIMECLUSTER.

Otherwise, you risk loss or damage of data.

Stopping the System While Guest Domains Are Operational

- Control Domains

Run the shutdown(1M) or the init(1M) command to stop the system. Using reboot(1M), uadmin(1M) or similar commands may possibly cause data loss as the control domain is rebooted without suspended guest domains.

- Guest Domains

To stop guest domains, run the shutdown(1M) command with the "i0" option.

After that, put the cluster applications of the suspended guest domains in the control domain offline.

When Control Domain Is Stopped

Sometimes, guest domains on a forcibly stopped node, which have been monitored by the cluster system, are not stopped although the control domain has been forced to stop. This may possibly cause data loss when restoring the control domain while any of the guest domains are not stopped.

In order to stop absolutely all input and output from guest domains, please make sure to interrupt the electrical power in the server rack by either of the following methods before restoring the stopped control domain. After that, reboot the disconnected nodes. Also note that no kernel crash dump is collected for operating systems in guest domains.

- Press and hold the power button for a while.

- On the ILOM remote control, select "Immediate Power Off".

See
....
For details on the remote control, please refer to the "Integrated Lights Out Manager User's Guide".
....

Manual Switching of Cluster Applications

When guest domains are switched to a standby system, PRIMECLUSTER uses the ldm(1M) command to stop those guest domains.

Therefore, whenever PRIMECLUSTER suspends a guest domain, the OS and all applications in that guest domain stop abruptly.

In order to avoid such abrupt stops when switching cluster applications manually, log in to the guest domain OS and stop the OS by running the shutdown(1M) or a similar command before switching the cluster application.

H.5 Maintenance of Cluster Systems in Oracle VM Server for SPARC Environments

This section describes maintenance methods for cluster systems in Oracle VM Server for SPARC environments.

H.5.1 Maintenance of the Control Domain

Carry out maintenance of the control domain in the same way as for normal clusters.

For further details, refer to "Chapter 10 Maintenance of the PRIMECLUSTER System".

H.5.2 Maintenance of Guest Domains

For doing maintenance work on guest domains, the statuses of cluster applications which correspond to the guest domains have to be "online". Follow the following procedure:

1. Check that the status of each cluster application corresponding to a guest domain in which to carry out maintenance is "online". If there are any cluster applications that are not online, shutdown the corresponding guest domains with shutdown(1M) command, and then switch the guest domains to the intended node with hvswitch(1M) command.

See
. . . .
For details on how to determine whether cluster applications are online, please refer to "7.1.3.1 RMS Tree".
. . . .

2. Log in to the guest domain in which to carry out maintenance and boot it in single-user mode.

3. Implement maintenance of the guest domain as necessary.

4. When you have finished all maintenance work, switch back any cluster applications that you switched in above Step 1. If you have not switched any cluster applications in Step 1, boot the guest domain in multi-user mode.

See
. . . .
For details on how to switch back, please refer to "7.2.2.3 Switching a Cluster Application".
. . . .

H.5.3 Collection of Diagnostic Data

After any trouble in the PRIMECLUSTER system, collect the following data, necessary for diagnosis, from all clustered control domains, from all guest domains specified for cluster applications, and from the cluster management server.
After that, please contact our customer support.

- Control Domains

  Referring to "C.1 Collecting Troubleshooting Information", collect the following data:

  - PRIMECLUSTER diagnostic data

  - If the malfunction is reproducible, documentation describing procedures for reproduction

- Guest Domains

  Run fjsnap to collect the data needed for error diagnosis.
  For details, refer to "C.1.1 Executing the fjsnap Command".

Information
. . . .
- For reporting a problem, first collect the necessary error diagnosis data. If you do not provide information enabling us to confirm the problem and reproduce the error, it may take us longer or even be impossible to simulate and diagnose the problem, hence, making it impossible to help you at all with troubleshooting.

- Please collect the diagnostic material from all nodes in the PRIMECLUSTER System as soon as possible. Especially in the case of data to be collected by fjsnap, necessary information may get lost if too much time elapses after the error.

- For nodes that have been suspended by force, run "sync" in OBP mode in order to collect the system dump.
. . . .

H.6 Recommended Configuration

This section presents a sample layout for a PRIMECLUSTER system in an Oracle VM Server for SPARC environment. Please use it for reference when designing your system.

System Configuration

Figure H.1 Sample System Configuration
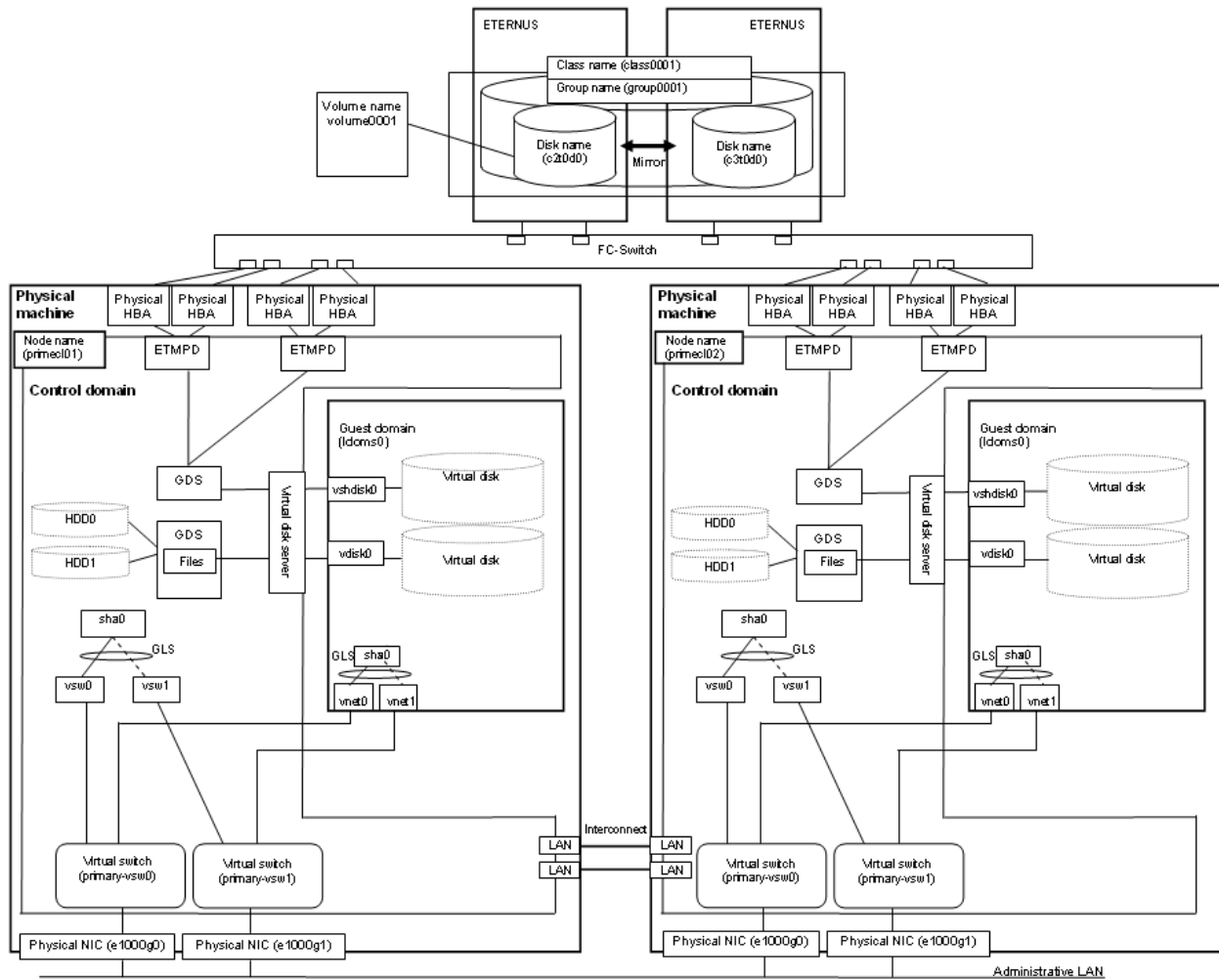
Utilized Hardware

- SPARC Enterprise T5120

Utilized Software

- Solaris 10

- Enhanced Support Facility 3.0A30

- PRIMECLUSTER 4.2A00

- Oracle VM Server for SPARC Manager Software 1.0.3

- ETERNUS MultiPath Driver 2.0.3

- Fujitsu PCI Fibre Channel 4.0

- Internet Explorer 7 (installed on administrator PC)

- Java(TM) Runtime Environment Standard Edition v1.5.0-14

Shared Disk Storage System Settings

Device Name	LUs	RAID Groups	RAID Level	Size
ETERNUS 4000 Model 300 (FC)	4	1	5	248.2GB
Add-On File Unit (FC) x2	-	-	-	18.2GB

Note

For using the disk storage system with UNIX, the disks must be formatted beforehand with the "format(1M)" command.
When you plan to format the disk storage system as part of your cluster system setup works, please take sufficient time for formatting into consideration.

Sample Settings for Virtual Network

Node	Name of Control Domain	Name of Physical NIC	Name of Virtual Switch	Name of Guest Domain	Name of Virtual NIC
primecl01	primary	e1000g0	primary-vsw0	ldoms0	vnet0
		e1000g1	primary-vsw1	ldoms0	vnet1
primecl02	primary	e1000g0	primary-vsw0	ldoms0	vnet0
		e1000g1	primary-vsw1	ldoms0	vnet1

In the control domain, create virtual switches primary-vsw0 and primary-vsw1 using physical NICs e1000g0 and e1000g1, and add vnet0 and vnet1 as virtual NICs to be used in the guest domains. Use GLS to create redundant virtual NICs in each domain. Specify vsw0 and vsw1 in the control domain as well as vnet0 and vnet1 in the guest domains as virtual NICs to be made redundant by GLS.

Note

Settings for the virtual network are made by running the ldm(1M) command in the control domain.
Network adapters used in virtual networks must be compliant with GLDv3.

H.6.1 Cluster Configuration Worksheet

Below section presents a sample worksheet for cluster configuration.

Item			Setting	Remarks
Shared disk unit: Type 1	Machine model		ETERNUS4000 model 300	
	Device name		ETERNUS4000#003	
	IP address		10.33.66.215	
	Subnet mask		255.255.255.0	
	Interface		fibre channel	
	LUN number		0 - 0F	
	Ports		CM0CA2P0, CM1CA2P0	
	Connection paths		2	
	FC-SW	Use of fibre channnel switching hub	■ yes □ no	
		Machine model	ETERNUS SN200 model 210	
		Device name	ETERNUSSN200#003	
		IP address	10.33.66.203	
		Subnet mask	255.255.255.0	
	File system selection (type of file system used in cluster)		UFS	
NTP	Operation mode		□ NTP server (integrated in cluster) ■ NTP client	
	Protocol (only when broadcast is specified)			
	Server	Primary	Host name	pclntp1	
			IP address	10.124.95.11	
			Subnet mask	255.255.255.0	
		Secondary	Host name	pclntp2	
			IP address	10.21.8.3	
			Subnet mask	255.255.255.0	
Web-Based Admin View	Operation mode		□ 1 ■ 2 □ 3 layers	
	Server	Primary	Host name	primecl01	
			IP address	10.34.214.181	
			Subnet mask	255.255.255.0	
			URL	http://10.34.214.181:8081/Plugin.html	
		Secondary	Host name	primecl02	
			IP address	10.34.214.182	
			Subnet mask	255.255.255.0	
			URL	http://10.34.214.182:8081/Plugin.html	
	User ID	wvroot group	pclwv	
		clroot group	pclcl	
		cladmin group	pcladm	
		clmon group	pclmon	
		sdxroot group	pclsdx	

Cluster Foundation (CF)	Cluster name	SLV002	
	Cluster nodes	2	
	Cluster node names	primecl01, primecl02	
	Subnets	1	
	Subnet IP (network address)	192.168.1.0	
	Subnet mask	255.255.255.0	
	Use of CF remote service	■ cfcp (CF file copy) ■ cfsh (CF remote command execution)	
	Use by RMS	◆ [yes] ◇ no	
Shutdown facility (SF)	Shutdown agent	□ RCI Panic (SA_rcip) □ RCI Reset (SA_rcir) □ XSCF Panic (SA_xscfp) □ XSCF Reset (SA_xscfr) □ XSCF Console Break (SA_rccu) □ RCCU Console Break (SA_rccu) □ ALOM Console Break (SA_sunF) ■ ILOM Panic (SA_ilomp)	
Cluster resource management facility (CRM)	Hardware devices stored in resource database	□ Network unit	Required for using takeover of public LAN (IP) and node names
		□ Line switching unit	Required for using line takeover function
		■ Disk unit	Required when using GDS and GFS

Node 1		Machine model		SPARC Enterprise T5120	
		Device name		DBSV#001	
	Control domain	Domain ID / Domain name		—	
		Node name (#uname -n)		primecl01	
		CF node name		primecl01	
		Interconnect	IP address	192.168.1.1	
			Subnet mask	255.255.255.0	
			Device name	e1000g3, e1000g4	
		Node in CF quorum set		◆[yes] ◇ No	Do not change the default setting
		Use by RMS		◆ [yes] ◇ no Suffix : primecl01RMS	
		Weight		2	Node weight (priority for any partitioned clusters) used by SF
		RCI	Address	—	Only when using RCI asynchronous monitoring
		System console	IP address	10.33.66.71	RCCU, XSCF, ALOM and ILOM data
			Subnet mask	255.255.255.0	
			User ID	root	
			User authority	—	
			Connection type	■ SSH □telnet	
		Use of shared disk		■ yes □ no	
		scsi-initiator-id (OBP variable #eeprom command)		7	
		SCSI / fibre channel	Circuits (paths)	2	
			Device name 1	fjpfca#0	
			sd device name	sd11 - sd42	
			Device name 2	fjpfca#1	
			sd device name	sd43 - sd74	
		Management LAN	IP address	10.33.66.91	
			Subnet mask	255.255.255.0	
			Device name	e1000g0	
		Public LAN 1	IP address	10.34.214.181	
			Subnet mask	255.255.255.0	
			Device name	e1000g1	
		Public LAN 2	IP address	10.34.215.21	
			Subnet mask	255.255.255.0	
			Device name	e1000g2	
	Guest domain	Domain ID / Domain name		ldoms0	
	Guest domain	Domain ID / Domain name			
	Guest domain	Domain ID / Domain name			

		Machine model		SPARC Enterprise T5120	
Node 2		Device name		DBSV#002	
	Control domain	Domain ID / Domain name		—	
		Node name (#uname -n)		primecl02	
		CF node name		primecl02	
		Interconnect	IP address	192.168.1.2	
			Subnet mask	255.255.255.0	
			Device name	e1000g3, e1000g4	
		Node in CF quorum set		◆[yes] ◇ No	Do not change the default setting
		Use by RMS		◆ [yes] ◇ no Suffix : primecl02RMS	
		Weight		2	Node weight (priority for any partitioned clusters) used by SF
		RCI	Address	—	Only when using RCI asynchronous monitoring
		System console	IP address	10.33.66.72	RCCU, XSCF, ALOM and ILOM data
			Subnet mask	255.255.255.0	
			User ID	root	
			User authority	—	
			Connection type	■ SSH □telnet	
		Use of shared disk		■ yes □ no	
		scsi-initiator-id (OBP variable #eeprom command)		8	
		SCSI / fibre channel	Circuits (paths)	2	
			Device name 1	fjpfca#0	
			sd device name	sd11 - sd42	
			Device name 2	fjpfca#1	
			sd device name	sd43 - sd74	
		Management LAN	IP address	10.33.66.92	
			Subnet mask	255.255.255.0	
			Device name	e1000g0	
		Public LAN 1	IP address	10.34.214.182	
			Subnet mask	255.255.255.0	
			Device name	e1000g1	
		Public LAN 2	IP address	10.34.215.22	
			Subnet mask	255.255.255.0	
			Device name	e1000g2	
	Guest domain	Domain ID / Domain name		ldoms0	
	Guest domain	Domain ID / Domain name			
	Guest domain	Domain ID / Domain name			

Edit the /etc/inet/ntp.conf file for every node as follows:

```
server 10.34.214.100
server 10.34.214.101
server 127.127.1.0

fudge 127.127.1.0 stratum 9

enable auth monitor
disable pll
driftfile /etc/ntp/ntp.drift
statsdir /var/ntp/ntpstats/
filegen peerstats file peerstats type day enable
```

```
filegen loopstats file loopstats type day enable
filegen clockstats file clockstats type day enable
```

For details on how to edit the /etc/inet/ntp.conf file, run the following command to find online reference on xntpd(1M):

```
# man xntpd
```

H.6.2 GLS Setup Worksheet

This worksheet serves to organize necessary input items beforehand for using GLS (Global Link Services) in Oracle VM Server for SPARC environments.

This section presents a sample worksheet for the following duplication method:

- **NIC Switching** (logical IP takeover in control domain, physical IP takeover in guest domains)

NIC Switching (Logical IP Takeover in Control Domain, Physical IP Takeover in Guest Domains)

Item			Setting
GLS settings	Takeover virtual interface name		sha0:65
	Takeover virtual IP address (or host name)		10.34.214.185
	Subnet mask		255.255.255.0
	Node name (1)		primecl01
	Control domain configuration information	Virtual interface name	sha0
		Switching mode	◆NIC switching mode (logical IP takeover) ◇NIC switching mode (physical IP takeover)
		Primary physical interface name	vsw0
		Secondary physical interface name	vsw1
		Physical IP address (or host name)	10.34.214.181
		Logical IP address (or host name)	10.34.214.185
	Guest domain configuration information	Virtual interface name	sha0
		Switching mode	◇NIC switching mode (logical IP takeover) ◆NIC switching mode (physical IP takeover)
		Primary physical interface name	vnet0
		Secondary physical interface name	vnet1
		Physical IP address (or host name)	10.34.214.186
	Monitoring destination information	IP address (or host name) of primary monitoring destination	10.34.214.188
		IP address (or host name) of secondary monitoring destination	10.34.214.189
	Optional function	Standby NIC patrol	◇Disable ◆Enable Set to standby NIC MAC address (0:0:0:0:0:0)
	Node name (2)		primecl02
	Control domain configuration	Virtual interface name	sha0
		Switching mode	◆NIC switching mode (logical IP takeover) ◇NIC switching mode (physical IP takeover)
		Primary physical interface name	vsw0
		Secondary physical interface name	vsw1
		Physical IP address (or host name)	10.34.214.182
		Logical IP address (or host name)	10.34.214.185
	Guest domain configuration information	Virtual interface name	sha0
		Switching mode	◇NIC switching mode (logical IP takeover) ◆NIC switching mode (physical IP takeover)
		Primary physical interface name	vnet0
		Secondary physical interface name	vnet1
		Physical IP address (or host name)	10.34.214.186
	Monitoring destination information	IP address (or host name) of primary monitoring destination	10.34.214.188
		IP address (or host name) of secondary monitoring destination	10.34.214.189
	Optional function	Standby NIC patrol	◇Disable ◆Enable Set to standby NIC MAC address (0:0:0:0:0:0)

P Point

For creating redundancy on virtual NICs (vsw, vnet) of Oracle VM Server for SPARC, the standby patrol MAC addresses must be set to "0:0:0:0:0:0". This enables standby patrol to use the virtual MAC addresses that are allocated to the virtual NICs.

H.6.3 GDS Configuration Worksheet

Below section presents a sample worksheet for GDS configuration.

Item				Setting
GDS configuration	Class 1	Class name		class0001
		Class scope (node name)	Node 1	primecl01
			Node 2	primecl02
		Guest domain name	Node 1	ldoms0
			Node 2	ldoms0
		Spare disk 1	SDX disk name	
			Physical disk name in node 1	
			Physical disk name in node 2	
		Single disk 1	SDX disk name	
			Physical disk name in node 1	
			Physical disk name in node 2	
			Single volume 1	Volume name	
				Size	
				Virtual disk server device	
				Virtual disk	
			Single volume 2	Volume name	
				Size	
				Virtual disk server device	
				Virtual disk	
		Single disk 2	SDX disk name	
			Physical disk name in node 1	
			Physical disk name in node 2	
			Single volume 1	Volume name	
				Size	
				Virtual disk server device	
				Virtual disk	
			Single volume 2	Volume name	
				Size	
				Virtual disk server device	
				Virtual disk	

		Disk 1 to be connected to group	SDX disk name		diskmplb0001
			Physical disk name in node 1		mplb2048
			Physical disk name in node 2		mplb2048
		Disk2 to be connected to group	SDX disk name		diskmplb0002
			Physical disk name in node 1		mplb2049
			Physical disk name in node 2		mplb2049
		Highest-order group 1	Group name		group0001
			Group type		mirror
			Stripe width (*6)		
			Disk / low order group name	Disk / low-order group 1	
				Disk / low-order group 2	
			Volume 1	Volume name	volume0001
				Size	8380799
				Virtual disk server device	vol00@primery-vds0
				Virtual disk	vshdisk00
			Volume 2	Volume name	
				Size	
				Virtual disk server device	
				Virtual disk	

	Class 2	Class name			class0002
		Guest domain (node name)	Node 1		primecl01
			Node 2		primecl02
		Class scope (node name)	Node 1		ldoms1
			Node 2		ldoms1
		Spare disk 1	SDX disk name		
			Physical disk name in node 1		
			Physical disk name in node 2		
		Single disk 1	SDX disk name		diskmplb0003
			Physical disk name in node 1		mplb2050
			Physical disk name in node 2		mplb2050
			Single volume 1	Volume name	volume0001
				Size	8380799
				Virtual disk server device	vol01@primery-vds0
				Virtual disk	vshdisk01
			Single volume 2	Volume name	
				Size	
				Virtual disk server device	
				Virtual disk	
		Single disk 2	SDX disk name		
			Physical disk name in node 1		
			Physical disk name in node 2		
			Single volume 1	Volume name	
				Size	
				Virtual disk server device	
				Virtual disk	
			Single volume 2	Volume name	
				Size	
				Virtual disk server device	
				Virtual disk	

		Disk 1 to be connected to group	SDX disk name		
			Physical disk name in node 1		
			Physical disk name in node 2		
		Disk2 to be connected to group	SDX disk name		
			Physical disk name in node 1		
			Physical disk name in node 2		
		Highest-order group 1	Group name		
			Group type		
			Stripe width　(*6)		
			Disk / low order group name	Disk / low-order group 1	
				Disk / low-order group 2	
			Volume 1	Volume name	
				Size	
			Volume 2	Volume name	
				Size	

Appendix I   Release information

This appendix describes the main contents for change of this manual.

Number	Edition	Part and Chapter	Details
1	2	5.1.2 Configuring the Shutdown Facility	The setting method of SPARC Enterprise M3000, T1000, T2000, T5120, T5220, or T5440 was added.
2	2	6.2.2 GDS Configuration Setup	Notes when two or more single nodes were shifted to the cluster system were added.
3	2	6.6.1.2 Creating Fsystem Resources	The attention when the file lock was used was added. The explanation of the Timeout value used by the Fsystem resource was corrected.
4	2	9.1 Adding a Node	"(7) Setting of GDS" was added to the procedure.
5	2	9.1.3 Connecting a Node	The procedure when the GFS shared file system was used was added.
6	2	9.1.7 Setting Up GDS	The paragraph was newly added.
7	2	9.1.9 GFS Shared Settings	The operation procedures were corrected.
8	2	Chapter 12 PRIMECLUSTER Product List	The version levels of the corresponding products were updated.
9	2	A.9 Cluster Application Worksheets	The value of the default of "CLUSTEREXCLUSIVE" and "AUTORECOVER" was corrected.
10	2	A.10.1 Cluster Configuration Worksheet	"ALOM" and "ILOM" were added to "Shutdown Facility".
11	2	A.5.1 CF Configuration	Added setting method for asynchronous RCI monitoring.
12	2	Glossary	Added definitions for "domain", "control domain", "guest domain", "control domain cluster", and "Logical Domains Manager (LDoms)".
13	2	Appendix H Using PRIMECLUSTER in LDoms (Logical Domains) Environments	Added new appendix.
14	3	Throughout document	Changed company name to "Fujitsu Technology Solutions".
15	4	5.1.2.4.1 Checking Console Configuration	Added information on ILOM 3.0.
16	4	6.6.1.6.1 Prerequisites	Added precaution on the state transition procedure.
17	4	A.5 Kernel Parameter Worksheet	Added precaution on the state transition procedure.
18	4	F.1 Registering a Procedure Resource	Added precaution on the state transition procedure.
19	5	5.1.1 Setting Up CF and CIP	Added description for CIP nodenames to the note.

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.

See also *Meta Data Server*.

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 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 Wizard kit to create object definitions for a specific type of application.

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 configuration (GDS)

Function that automatically recognizes the physical connection configuration of shared disk units and registers the units to the resource database.

automatic power control

This function is provided by the Enhanced Support Facility (ESF), and it automatically switches the PRIMEPOWER power on and off.

automatic switchover (RMS)

The procedure by which RMS automatically switches control of userApplication over to another host after specified conditions are detected.

See also *directed switchover*, *failover*, *switchover*, and *symmetrical switchover*.

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*.

BM (base monitor) (RMS)

The RMS module that maintains the availability of resources. The base monitor is supported by daemons and detectors. Each host 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.*

ccbr.conf

The environment configuration file that is used for backup and restore operations, and is placed in the "/opt/SMAW/ccbr" directory. This file is used in the "$CCBRHOME" variable setting. For details, see the manual pages for the "cfbackup(1M)" and "cfrestore(1M)" commands and the comments in the "ccbr.conf" file.

ccbr.gen

The file that stores the generation number and is placed in the "/opt/SMAW/ccbr" directory. A value of 0 or higher is stored in this file. For details, see the manual pages for the "cfbackup(1M)" and "cfrestore(1M)" commands.

CCBRHOME variable

The variable that identifies the directory in which backup data is stored. The initial value is the "/var/spool/pcl4.1/ccbr" directory. This variable can be set only in the "ccbr.conf" file.

CF (Cluster Foundation or Cluster Framework)

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*, *object*, *parent*, and *leaf object*.

CIM

Cluster Integrity Monitor

CIP

Cluster Interconnect Protocol

CLI

command-line interface

CLM

Cluster Manager

class (GDS)

See *disk class*.

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 Configuration Backup and Restore (CF)

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.

Cluster Foundation

The set of PRIMECLUSTER modules that provides basic clustering communication services.

See also *base cluster foundation*.

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.

Cluster Resource Management facility

Facility that manages hardware units that are shared among multiple 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 Technology Solutions products only.)

See also *mirror virtual disk, simple virtual disk, striped virtual disk, virtual disk*.

concatenation (GDS)

The linking of multiple physical disks. This setup allows multiple disks to be used as one virtual disk that has a large capacity.

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.

console

See *single console*.

Console Break agent

The Console Break agent is used for the Shutdown Facility to eliminate a node by sending a break signal from RCCU.

control domain

Domain in which the Oracle VM Server for SPARC is installed. All platforms that are using Oracle VM Server for SPARC must contain a control domain. By using the ldm command within this domain, other domains can be created and controlled.

control domain cluster

Cluster that is configured between different control domains

CRM

Cluster Resource Management

custom detector (RMS)

See *detector*.

custom type (RMS)

See *generic type*.

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*, *service node*, *Scalable Internet Services*.

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 userApplication over to another host.

See also *automatic switchover*, *failover*, *switchover*, and *symmetrical switchover*.

disk class (GDS)

Collection of SDX objects. The shared type disk class is also a resource unit that can be used by the PRIMECLUSTER system. A disk class is sometimes simply called a "class."

disk group (GDS)

A collection of disks or low-order groups that become the unit for mirroring, striping, or concatenation. Disk and low-order groups that belong to the same disk group are mutually mirrored, striped, or concatenated according to the type attribute (mirror, stripe, or concatenation) of that disk group.

A disk group is sometimes simply called a "group."

DLPI

Data Link Provider Interface

Domain

A set of one or more system boards that function as an independent system. While the server is shared, an operating system can be installed in each domain to enable each domain to operate as an independent system.

Each domain consists of a logical system board assigned to it. Each domain is electrically insulated by each hardware partition. Therefore, if one domain fails, it does not affect the other domains in the server.

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*, *LEFTCLUSTER*, *node state*.

EE

Enterprise Edition

ENS (CF)

See *Event Notification Services*.

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.

Ethernet

LAN standard that is standardized by IEEE 802.3. Currently, except for special uses, nearly all LANs are Ethernets. Originally the expression Ethernet was a LAN standard name for a 10 megabyte per second type LAN, but now it is also used as a general term that includes high-speed Ethernets and gigabyte Ethernets.

Event Notification Services (CF)

This PRIMECLUSTER module provides an atomic-broadcast facility for events.

failover (RMS, SIS)

The process by which a user application automatically transfers processes and data inherited from an operating system to a standby system because some failure has occurred.
With SIS, this process switches a failed node to a backup node. With RMS, this process is known as switchover.

See also *automatic switchover*, *directed switchover*, *switchover*, *symmetrical switchover*.

Fast switching mode

One of the LAN duplexing modes presented by GLS.

This mode uses a multiplexed LAN simultaneously to provide enhanced communication scalability between Solaris servers and high-speed switchover if a LAN failure occurs.

fault tolerant network

A network with the ability to withstand faults (fault tolerant). Fault tolerant is the ability to maintain and continue normal operation even if a fault occurs in part of the computer system. A fault tolerant network is therefore a network that can continue normal communication even if a flat occurs in part of the network system.

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*, *database node*, *Scalable Internet Services*.

GDS

See *Global Disk Services*.

generation number

Data generation management is enabled in the PRIMECLUSTER backup and restore operations. The current generation number is added as part of the backup and restore data name. Integers of 0 or higher are used as generation numbers, and the generation number is incremented each time backup is successful. The generation number is stored in the "ccbr.gen" file and can be specified as an optional argument in the "cfbackup(1M)" and "cfrestore(1M) " commands.
For details, see the manual pages for the "cfbackup(1M)" and "cfrestore(1M)" commands.

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*.

GFS

See *Global File Services*.

GFS local file system

A local file system that features high-speed file access and file system switchover after a system failure. High-speed file access is available through the serial block allocation to file data, and file system switchover is assured through a file system consistency recovery function.

A GFS local file system is used by one node or by two nodes, in which case the file system is used as a switchover file system.

GFS shared file system

A shared file system that allows simultaneous access from multiple Solaris systems that are connected to shared disk units, while maintaining data consistency, and allows processing performed by a node to be continued by other nodes even if the first node fails.

A GFS shared file system can be mounted and used concurrently from multiple nodes.

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.

GLS

See *Global Link Services*.

graph (RMS)

See *system graph*.

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.

group (GDS)

See *disk group*.

GS/SURE linkage mode

One of the LAN duplexing modes presented by GLS.

This mode uses a duplexed LAN simultaneously and high reliance communication with Global server or SURE system is realized.

GUI

See *graphical user interface*.

guest domain

Virtualized hardware environment in which an independent operating system is running. It can be started and stopped without any influence on other domains.

HA

high availability

high availability

This concept applies to the use of redundant resources to avoid single points of failure.

highest-order group (GDS)

Group that does not belong to another group. A volume can be created in the highest-order group.

hub

Star-type wiring device used for LAN or fibre channels.

ICF

Internode Communication Facility

I/F

Interface

installation server

System having a Solaris CD image on the disk or CD-ROM drive to distribute the Solaris CD image to other systems over the network.

interconnect (CF)

See *cluster interconnect*.

Internet Protocol address

A numeric address that can be assigned to computers or applications.

See also *IP aliasing*.

internode communication facility

Communication function between cluster nodes that are used by PRIMECLUSTER CF. Since this facility is designed especially for communication between cluster nodes, the overhead is less than that of TCP/IP, and datagram communication services that also guarantee the message arrival sequence can be carried out.

I/O

input/output

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 host.

See also *Internet Protocol address*.

JOIN (cluster join services module) (CF)

See *Cluster Join Services*.

keyword (reserved words)

A word that has special meaning in a programming language. For example, in the configuration file, the keyword node identifies the kind of definition that follows.

LAN

local area network

latency (RMS)

Time interval from when a data transmission request is issued until the actual response is received.

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 purpose for the intermediate LEFTCLUSTER state is to avoid the network partition problem.

See also *UP*, *DOWN*, *network partition*, *node state*.

line switching unit

This device connects external lines to more than one node and switches the connected nodes by the RCI.

link (RMS)

Designates a child or parent relationship between specific resources.

local area network

See *public LAN*.

local host

The host from which a command or process is initiated.

See also *remote host*.

local MAC address

MAC address that the system administrator of a local area network (LAN) system guarantees to be unique within that system.

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.

logical volume (GDS)

General term for a virtual disk device that the user can access directly. The user can access a logical volume in the same way as accessing a physical disk slice (partition). A logical volume is sometimes simply called a "volume."

low-order group (GDS)

Group that belongs to another group. A volume cannot be created in a low-order group.

MA

Monitoring Agents

MAC address

Address that identifies the office or node that is used by the MAC sublayer of a local area network (LAN).

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)

GFS daemon that centrally manages the control information of a file system (meta-data).

MIB

Management Information Base

MIPC

Mesh Interprocessor Communication

mirrored disk (RCFS)

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 Technology Solutions products only.)

See also *mirrored piece*.

mirrored piece (RCFS)

Physical pieces that together comprise a mirrored virtual disk. These pieces include mirrored disks and data disks. (Applies to transitioning users of existing Fujitsu Technology Solutions products only.)

See also *mirrored disk*.

mirrored volume (GDS)

A volume that is created in a mirror group. Data redundancy is created by mirroring.

mirror group (GDS)

A disk group of the mirror type. This a collection of mutually mirrored disks or low-order groups.

mirroring (GDS)

A setup that maintains redundancy by writing the same data to multiple slices. Even if an error occurs in some of the slices, this setup allows access to the volume to continue as long as a normal slice remains.

mirror virtual disk (VM)

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 Technology Solutions products only.)

See also *concatenated virtual disk, simple virtual disk, striped virtual disk,* and *virtual disk*.

mixed model cluster

A cluster system that is built from different PRIMEPOWER models. For example, one node is a PRIMEPOWER 1000 machine, and another node is a PRIMEPOWER 200 machine.

The models are divided into two groups, which are represented by the PRIMEPOWER 200, 400, and 600 machines and the PRIMEPOWER 800, 1000, and 2000 machines.

monitoring agent

Component that monitors the state of a remote cluster node and immediately detects if that node goes down. This component is separate from the SA function.

mount point

The point in the directory tree where a file system is attached.

multihosting

Same disk via multiple controllers. (Applies to transitioning users of existing Fujitsu Technology Solutions products only.)

native operating system

The part of an operating system that is always active and translates system calls into activities.

network adapter

A LAN network adapter.

network interface card

See *network adapter*.

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.

NIC

network interface card

NIC switching mode

One of the LAN duplexing modes presented by GLS. The duplexed NIC is used exclusively, and LAN monitoring between the Solaris server and the switching HUB, and switchover if an error is detected are implemented.

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, DOWN, LEFTCLUSTER*.

NSM

Node State Monitor

object (RMS)

In the configuration file or a system graph, this is a representation of a physical or virtual resource.

See also *leaf object, object definition, node state, object type*.

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*, *object type*.

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*.

online maintenance

The capability of adding, removing, replacing, or recovering devices without shutting or powering off the host.

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 the database to users and applications in a clustered or MPP (massively parallel processing) platform.

Oracle VM Server for SPARC

Virtualization function using Hypervisor, which is provided as part of the firmware.

OSD (CF)

See *operating system dependent*.

parent (RMS)

An object in the configuration file or system graph that has at least one child.

See also *child*, *configuration file*, and *system graph*.

PAS

Parallel Application Services

patrol diagnosis

A function that periodically diagnoses hardware faults.

physical IP address

IP address that is assigned directory to the interface (for example, hme0) of a network interface card. See also *logical IP address*. For information about the logical interface, see the explanation of logical interface in ifconfig(1M).

primary host (RMS)

The default host on which a user application comes online when RMS is started. This is always the hostname of the first child listed in the userApplication object definition.

PRIMECLUSTER services (CF)

Service modules that provide services and internal interfaces for clustered applications.

private network address

Private network addresses are a reserved range of IP addresses specified by RFC1918. 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 host and not accessible to other RMS hosts.

See also *resource*, *shared resource*.

PS

Parallel Server

public LAN

The local area network (LAN) by which normal users access a machine.

See also *Administrative LAN*.

queue

See *message queue*.

quorum

State in which integrity is maintained among the nodes that configure the cluster system. Specifically, the CF state in all nodes that configure the cluster system is either UP or DOWN (there is no LEFTCLUSTER node).

RAO

RMS-Add on

RC2000

Software (Java application) that provides functions for operating the operating system console of each node.

RCCU

Abbreviation for Remote Console Connection Unit.

See also *remote console connection unit*.

RCI

Remote Cabinet Interface

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.

Reliant Monitor Services (RMS)

The package that maintains high availability of user-specified resources by providing monitoring and switchover capabilities.

remote console connection unit

Device that converts an RS232C interface and a LAN interface. This device allows another device (personal computer) that is connected to the LAN to use the TTY console functions through the Telnet function.

remote host

A host that is accessed through a telecommunications line or LAN.

See also *local host*.

remote node

See *remote host*.

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 host.

See also *private resource*, *shared resource*.

resource database (CF)

Database that manages information on hardware units that are shared among multiple nodes.

The resource database is managed by the cluster resource management facility.

resource definition (RMS)

See *object definition*.

resource label (RMS)

The name of the resource displayed in a system graph.

resource state (RMS)

Current state of a resource.

RMS

See *Reliant Monitor Services*.

RMS command

Commands that enable RMS resources to be administered from the command line.

RMS configuration

A configuration in which two or more nodes are connected to shared resources. Each node has its own copy of operating system and RMS software, as well as its own applications.

RMS Wizard kit

Each component of the RMS Wizard Kit adds new menu items to the RMS Wizard Tools for a specific application.

See also *RMS Wizard Tools*, *Reliant Monitor Services (RMS)*.

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 Wizard kit*, *Reliant Monitor Services*.

route

In the PRIMECLUSTER Concept Guide, this term refers to the individual network paths of the redundant cluster interfaces that connect the nodes to each other.

Rolling update

Update method used to fix an application or maintenance within the cluster system. Fix application is enabled by applying fixes to each node sequentially without stopping jobs.

SA

Shutdown Agent

SAN (Storage Area Network)

See *Storage Area Network*.

SC

Scalability Cluster

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.

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.

SCON (single console software)

See *single console*.

scope (GDS)

The range of nodes that can share objects in the shared type disk class.

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.

SD

Shutdown Daemon

SDX disk (GDS)

General term for disks that GDS manages. Depending on its use, an SDX disk may be called a single disk, a keep disk, a spare disk, or an undefined disk. An SDS disk is sometimes simply called a "disk."

SDX object (GDS)

General term for resources that GDS manages. The resources include classes, groups, SDX disks, and volumes.

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*, *gateway node*, and *Scalable Internet Services*.

SF

Shutdown Facility

shared disk connection confirmation

Function that checks whether that all shared disk units are turned on and all cable connections are correct when a node is started.

shared resource

A resource, such as a disk drive, that is accessible to more than one node.

See also private resource, resource.

Shutdown Facility

A facility that forcibly stops a node in which a failure has occurred. When PRIMECLUSTER decides that system has reached a state in which the quorum is not maintained, it uses the Shutdown Facility (SF) to return the cluster system to the quorum state.

shutdown request

Instruction that forcibly stops the specified node so that the quorum is restored.

simple virtual disk

Simple virtual disks define either an area within a physical disk partition or an entire partition.

See also *concatenated virtual disk*, *striped virtual disk*, and *virtual disk*.

single console

The workstation that acts as the single point of administration for hosts being monitored by RMS. The single console software, SCON, is run from the single console.

single disk (GDS)

SDX disk that does not belong to a group and can be used to create a single volume.

single volume (GDS)

A volume that is created in a single disk that does not belong to a group. There is no data redundancy.

SIS

See *Scalable Internet Services*.

spare disk (GDS)

A spare disk for restoring the mirroring state in place of a failed disk.

state

See *resource state*.

state transition procedure

The state transition procedure receives a state transition instruction from the cluster control and controls activation and deactivation of the resource (start and stop of the application).

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 group (GDS)

A disk group of the stripe type. This is a collection of disks or low-order groups that become striping units.

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).

See also *concatenated virtual disk*, *mirror virtual disk*, *simple virtual disk*, *virtual disk*.

striped volume (GDS)

A volume that is created in a striped group. Striping allows the I/O load to be distributed among multiple disks. There is no data redundancy.

stripe width (GDS)

The size is which data is divided when striping takes place.

striping (GDS)

Dividing data into fixed-size segments, and cyclically distributing and writing the data segments to multiple slices. This method distributes I/O data to multiple physical disks and issues I/O data at the same time.

switching mode

LAN duplexing mode presented by GLS.

There is a total of six switching mode types: fast switching mode, NIC switching mode, GS/SURE linking mode, multipath mode, and multilink Ethernet mode:

switchover

The process by which a user application transfers processes and data inherited from an operating node to a standby node, based on a user request.

switchover (RMS)

The process by which RMS switches control of userApplication over from one monitored host to another.

See also *automatic switchover*, *directed switchover*, *failover*, and *symmetrical switchover*.

symmetrical switchover (RMS)

This means that every RMS host is able to take on resources from any other RMS host.

See also *automatic switchover*, *directed switchover*, *failover*, and *switchover*.

synchronized power control

When the power of one node is turned in the cluster system configured with PRIMEPOWER, this function turns on all other powered-off nodes and disk array unit that are connected to nodes through RCI cables.

system disk (GDS)

The disk in which the operating Solaris is installed. This term refers to the entire disk, including slices that are currently operating as one of the following file systems or swap area:

/, /usr, /var, or swap area

system graph (RMS)

A visual representation (a map) of monitored resources used to develop or interpret the configuration file.

See also *configuration file*.

template

See *application template*.

type

See *object type*.

UP (CF)

A node state that indicates that the node can communicate with other nodes in the cluster.

See also *DOWN*, *LEFTCLUSTER*, *node state*.

user group

A group that limits the environment setup, operation management, and other operations presented by Web-Based Admin View and the Cluster Admin GUI. There are four user groups: wvroot, clroot, cladmin, and clmon. Each user ID is registered in an appropriate user group by the operation system administrator of the management server.

VIP

Virtual Interface Provider

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.

See also *concatenated virtual disk*, *mirror virtual disk*, *simple virtual disk*, *striped virtual disk*.

virtual interface (VIP)

Virtual IP address or host name that SIS uses to make multiple nodes of the cluster appear as a single-system image.

volume (GDS)

See *logical volume (GDS)*.

watchdog timer monitoring

Timer value that measures operating system hangs and boot failures.

Web-Based Admin View

This is a common base enabling use of 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.

WK

Wizard Kit

WT

Wizard Tools

XSCF

Abbreviation for eXtended System Control Facility. XSCF is a system monitoring facility that consists of dedicated processors that are independent from a main CPU. XSCF performs integrated management of the cooling system (FAN unit), power supply unit, system monitoring, and power on/off and monitoring system of peripherals. This is enabled from remote places, providing functions to monitor a main unit, notify a system administrator of a system failure, and perform console input/output from remote places via serial port or Ethernet port.

Index