PRIMECLUSTER

Installation and Administration Guide 4.3

Linux

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. Maintaining 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 4.3 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 Modifications

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 items related to the maintenance 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 Registering, Changing, and Deleting State Transition Procedure Resources for PRIMECLUSTER Compatibility

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

Appendix F Registering PRIMECLUSTER-compatible Resources

Audience: All users who use PRIMECLUSTER-compatible resources
Contents: This appendix describes procedures for registering PRIMECLUSTER-compatible resources.

Appendix G Configuration Update Service for SA

Audience: All users who use PRIMECLUSTER systems
Contents: This appendix descries Configuration Update Service for SA.

Appendix H Using PRIMECLUSTER in a VMware Environment

Audience: All users who use PRIMECLUSTER systems in a VMware environment
Contents: This appendix describes the installation procedures for using the PRIMECLUSTER system in a VMware environment.

Appendix I Startup Scripts and Startup Daemons, and Port Numbers in PRIMECLUSTER

Audience: System administrators who build PRIMECLUSTER systems
Contents: This appendix provides explanations on scripts and daemons that are started by PRIMECLUSTER, and the port numbers being used.

Appendix J Using Firewall

Audience: All users who use PRIMECLUSTER systems
Contents: This appendix describes the procedure when using Firewall in the PRIMECLUSTER system.

Appendix K Changes in Each Version

Audience: All users who use PRIMECLUSTER 4.0A20, 4.1A20, 4.1A30, 4.2A00, 4.2A30 or 4.3A00.
Contents: This appendix describes the changes made to the specifications of PRIMECLUSTER 4.3A01.

Appendix L Release Information

Audience: All users who use PRIMECLUSTER systems
Contents: This appendix describes the main changes 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) Reference 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 Web-Based Admin View Operation Guide*

- *PRIMECLUSTER Messages*

- *PRIMECLUSTER Wizard for Oracle Configuration and Administration Guide*

- *Interstage Application Server Reference Manual - Command Edition*

- *Interstage Application Server Operator's Guide*

- *PRIMEQUEST 1000 Series User Interface Operating Instructions*

- *PRIMEQUEST 1000 Series Message Reference*

- *PRIMEQUEST 1000 Series Tool Reference*

- *PRIMEQUEST 1000 Series Installation Manual*

- *PRIMEQUEST 1000 Series Administration Manual*

Note

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

- Software Release Guide/Installation Guide for PRIMECLUSTER

  This Software Release Guide and Installation Guide are provided with each PRIMECLUSTER product package.

  The data is stored on "CD" of each package. For details on the file names, see the documentations.

Manual Series
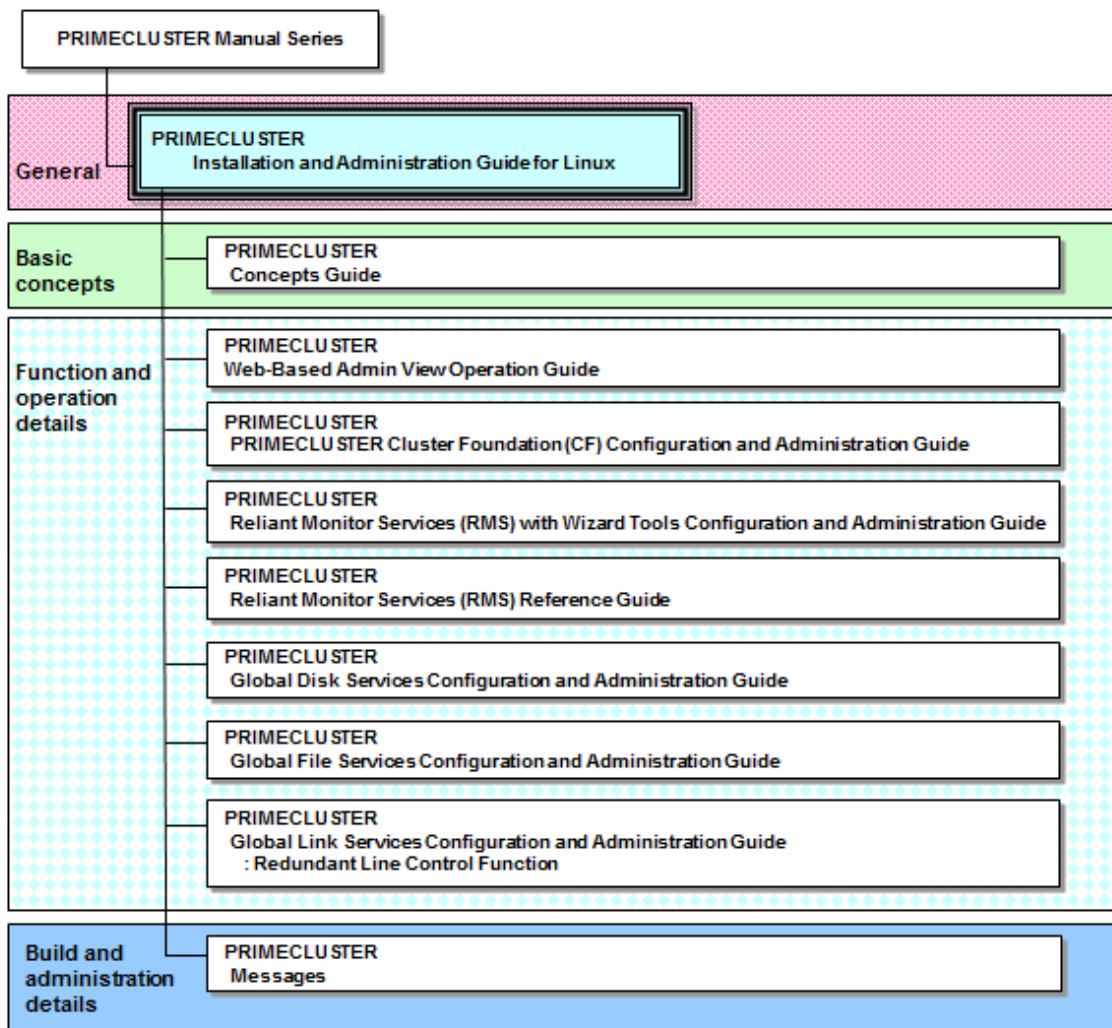
Manual Printing

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 Software Release Guide for PRIMECLUSTER that comes with the product.

Adobe Reader is required to read and print this PDF file. To get Adobe Reader, see Adobe Systems Incorporated's website.

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 Linux(R) 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:0:0:root:/root:/bin/bash

bin:x:1:1:bin:/bin:/bin/bash

daemon:x:2:2:daemon:/sbin:/bin/bash

lp:x:4:7:lp daemon:/var/spool/lpd:/bin/bash

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:

 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.
. . . .

Abbreviations

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

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

- Microsoft(R) Windows(R) 7 Professional operating system is abbreviated as Windows(R) 7.

- Windows(R) XP, Windows(R) Vista, and Windows(R) 7 are generically referred to as Microsoft(R) Windows.

- Red Hat Enterprise Linux is abbreviated as RHEL.

- Red Hat Enterprise Linux AS is abbreviated as RHEL-AS.

- Red Hat Enterprise Linux ES is abbreviated as RHEL-ES.

- RHEL-AS and RHEL-ES are abbreviated as Linux(R).

- VMware(R) ESX(R) is abbreviated as ESX.

- VMware(R) ESXi(TM) is abbreviated as ESXi.

- If you use ESXi, read "ESX" as "ESXi."

- VMware vSphere(R) is abbreviated as VMware vSphere.

- VMware vSphere(R) High Availability is abbreviated as VMware HA.

- VMware vSphere(R) Fault Tolerance is abbreviated as VMware FT.

- VMware vSphere(R) Distributed Resource Scheduler(TM) is abbreviated as VMware DRS.

- VMware vSphere(R) vMotion(R) is abbreviated as VMware vMotion.

- VMware vSphere(R) Storage vMotion(R) is abbreviated as VMware Storage vMotion.

- VMware(R) vCenter(TM) Converter(TM) is abbreviated as VMware vCenter Converter.

Export Controls

Exportation/release of this document may require necessary procedures in accordance with the regulations of your resident country and/or US export control laws.

Trademarks

Red Hat is a registered trademark of Red Hat, Inc. in the U.S. and other countries.

Linux is a registered trademark of Linus Torvalds.

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

Microsoft, Windows, WindowsXP, WindowsVista, Windows7 and Internet Explorer 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.

VMware is a registered trademark or trademark of VMware, Inc. in the United States and/or other jurisdictions.

PRIMECLUSTER is a registered trademark of Fujitsu Limited.

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

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.

Revision History

Revision	Location	Edition
Added the setting procedure for the guest OS in the host OS.	3.2.1.1 Host OS setup 3.2.2.1 Host OS setup 3.2.3.1 Installation and Setup of Related Software (Host OS)	6.1
Added "6.8 Setting Up the RMS Environment."	Chapter 6 Building Cluster Applications	
Added the item of RMS shutdown wait time.	A.8.1 RMS Setup Worksheet	
Added the notes for the maximum numbers of resources and detectors.	A.10 Cluster Application Worksheets	
Changed the description in Note.	H.1 Cluster Systems in a VMware Environment	
Changed the step 1.	H.2.4.1 Setting Up I/O Fencing	
Added the notes to the explanation of the test for forced shutdown of cluster nodes.	1.4 Test	6.2
Changed the use when building a cluster system between guest OSes on one host OS.	2.2.1 Virtual Machine Function	
Changed "Attached Driver Guide" to "PRIMEQUEST 1000 SERIES HBA blockage function USER'S GUIDE".	3.1.4.2 PRIMEQUEST	
Changed the setting procedure of virtual SCSI devices in a KVM environment.	3.2.1.1 Host OS setup 3.2.2.1 Host OS setup	
Changed the setting procedure of virtual bridges in a Xen environment.	3.2.1.1 Host OS setup 3.2.2.1 Host OS setup	
Changed the procedure of Host OS failover function.	3.2.3 When building a cluster system between guest OSes on multiple host OSes using Host OS failover function	
Added the notes on the operation check after setting up the shutdown agent.	5.1.2.3 Setting up the shutdown agent in PRIMERGY	

Revision	Location	Edition
	5.1.2.4 Setting up the shutdown agent in PRIMEQUEST 5.1.2.5 Setting up the shutdown agent in virtual machine environment	
Added the explanations of examples and notes on setting the configuration file for the shutdown agent.	5.1.2.3.2 IPMI 5.1.2.3.3 Blade 5.1.2.5.1 vmSP (Virtual Machine Service Provider) 5.1.2.5.2 libvirt 5.1.2.5.3 vmchkhost	
Added the notes on use of maintenance mode.	7.2.2.6 Entering maintenance mode for Cluster Application	
Added the reference when changing the settings in /etc/sysconfig/libvirt-guests.	3.2.1.1 Host OS setup 3.2.2.1 Host OS setup 3.2.3.1 Installation and Setup of Related Software (Host OS)	6.3
Added the information to be checked when "TestFailed" or "InitFailed" is displayed in Information "About the displayed results."	5.1.2.5.2 libvirt	
Added "8.9.2.3 Changing the Settings in /etc/sysconfig/libvirt-guests."	8.9.2 KVM environment	
Changed the explanation for executing failover of the cluster application.	9.3.2.2 Procedure for Applying Correction by Rolling Update	
Changed the contents in "Setting value" of "Use" in "Item."	A.10.3 Gds Resource Worksheet	
Added the notes on environments and functions which PRIMECLUSTER supports.	H.1 Cluster Systems in a VMware Environment	
Changed the descriptions of the setting for SCSI controllers.	H.2.1.1 Installation and Configuration of Related Software	
Added "I.4 Changes in PRIMECLUSTER 4.3A10 from 4.1A40." Added "I.6 Changes in PRIMECLUSTER 4.3A10 from 4.2A30."	Appendix I Changes in Each Version	
Added "HV_CONNECT_TIMEOUT" in all versions as an incompatible function.	Appendix I Changes in Each Version	
Added "Single-node Cluster Operation."	2.3 Determining the Cluster System Operation Mode 2.3.3 Single-node Cluster Operation	6.4
Added the notes on an immediate cluster failover when an error occurs on an I/O device that a system volume is allocated.	3.1.2 Setting Up Disk Units 3.2.1.1 Host OS setup 3.2.2.1 Host OS setup 3.2.3 When building a cluster system between guest OSes on multiple host OSes using Host OS failover function 3.2.3.1 Installation and Setup of Related Software (Host OS)	
Added the notes on the single-node cluster operation.	4.3.3.1 Initial setup of the operation management server Chapter 5 Building a Cluster 5.1.1 Setting Up CF and CIP 6.7.2.1 Creating Standby Cluster Applications 6.7.3.4 Setting Up Gds Resources	
Added the note in the case of creating a redundant administrative LAN used in the shutdown facility by using GLS.	5.1.2 Setting Up the Shutdown Facility	

Revision	Location	Edition
Changed the setting of "weight" of "Node information" found on the Cluster Configuration Worksheet in System Design Examples.	A.11.1.1 Cluster Configuration Worksheet A.11.2.1 Cluster Configuration Worksheet A.12.1.1 Cluster Configuration Worksheet A.12.2.1 Cluster Configuration Worksheet A.13.1.1 Cluster Configuration Worksheet A.13.2.1 Cluster Configuration Worksheet H.7.1.1 Cluster Configuration Worksheet H.7.2.1 Cluster Configuration Worksheet	
Changed the note on the path policy for the Native Multipathing.	H.1 Cluster Systems in a VMware Environment	
Changed the note that try not to target a virtual machine that installs PRIMECLUSTER as cluster functions of VMware.	H.1 Cluster Systems in a VMware Environment	
Changed the descriptions on the installation of VMware vSphere.	H.2.1.1 Installation and Configuration of Related Software H.3.1.1 Installation and Configuration of Related Software	
Added the section "7.5 Notes on Operation."	Chapter 7 Operations	6.5
Added the description about the setting of kdump.	3.3 PRIMECLUSTER Installation	6.6
Added the note in the case of creating a redundant administrative LAN used in the shutdown facility by using GLS.	5.1.2 Setting Up the Shutdown Facility	
Added the RX500 series for the supported hardware device.	5.1.2 Setting Up the Shutdown Facility 5.1.2.3.4 kdump	
Described the procedure to encrypt the password.	5.1.2.3.2 IPMI	
Added the notes on characters allowed in passwords.	5.1.2.4.1 MMB	
Added the notes when nodes were forcibly stopped by the SA_mmbr shutdown agent.	5.1.2.4.1 MMB	
Changed the procedure to start the shutdown facility.	5.1.2.5.1 vmSP (Virtual Machine Service Provider) 5.1.2.5.2 libvirt 5.1.2.5.3 vmchkhost	
Described the information about the maximum concurrent connections for SSH.	5.1.2.5.1 vmSP (Virtual Machine Service Provider) 5.1.2.5.2 libvirt 5.1.2.5.3 vmchkhost	
Added the note on the hvw command.	6.7.2 Setting Up userApplication	
Added the notes on the virtual machine function.	2.2.1 Virtual Machine Function	6.7
Added the notes on umask.	Chapter 3 Software Installation	
Deleted the descriptions for "Behavior setup of recovery function."	3.1.3 Setting Up Hardware Monitoring with ServerView (PRIMERGY)	
Added the notes on setting up the virtual bridge for the administrative LAN in a KVM environment.	3.2.1.1 Host OS setup 3.2.2.1 Host OS setup	
Added the notes when setting /etc/sysconfig/libvirtguests.	3.2.1.1 Host OS setup 3.2.2.1 Host OS setup 3.2.3.1 Installation and Setup of Related Software (Host OS)	
Changed the setting procedure of virtual SCSI devices in a KVM environment.	3.2.1.1 Host OS setup 3.2.2.1 Host OS setup	
Added the notes on Firewall.	5.1 Initial Cluster Setup	
Modified the descriptions of the CF node name.	5.1.1 Setting Up CF and CIP	
Changed the procedure to start the shutdown facility.	5.1.2.3.5 Starting up the shutdown daemon	

Revision	Location	Edition
Added the notes on SA_libvirtgp.	5.1.2.5.2 libvirt	
Changed the operation procedure to register hardware devices.	5.1.3.2 Registering Hardware Devices	
Changed the descriptions for "Preparation prior to displaying fault resource identification and operator intervention request messages."	5.2 Setting Up Fault Resource Identification and Operator Intervention Request	
Added the notes on GDS configuration setup.	6.3 GDS Configuration Setup	
Added the notes on setting up shared disks.	6.3.2 Setting Up Shared Disks	
Changed the explanation of the environment variable HV_AUTORECOVER.	6.6 Setting Up Online/Offline Scripts	
Added the reference to set up Online/Offline scripts.	6.6 Setting Up Online/Offline Scripts	
Added "PartialCluster" to the attribute of the userApplication.	6.7.2.1 Creating Standby Cluster Applications	
Deleted the reference.	6.7.2.1 Creating Standby Cluster Applications	
Modified the display example of the hvw command.	6.7.2.1 Creating Standby Cluster Applications 6.7.2.2 Creating Scalable Cluster Applications 6.7.3 Setting Up Resources 6.7.3.1 Setting Up Cmdline Resources 6.7.3.2 Setting Up Fsystem Resources 6.7.3.4 Setting Up Gds Resources 6.7.3.5 Setting Up Gls Resources 6.7.3.6 Setting Up Takeover Network Resources 6.7.3.7 Setting Up Procedure Resources 6.7.4 Generate and Activate 8.1 Changing the Cluster Configuration 8.5.1 Changing Operation Attributes (CUI) 8.7.1 Deleting a Resource 8.7.3 Deleting a userApplication	
Changed the example of how to execute a command.	6.7.3.3 Preliminary Setup for Gds Resources	
Changed the operation procedure to set up takeover network resources.	6.7.3.6 Setting Up Takeover Network Resources	
Added the notes on Configuration-Activate.	6.7.4 Generate and Activate	
Added the state icon to the table of "State display of other objects."	7.1.3.1 RMS Tree	
Changed the operation procedure.	7.2.2.4 Bringing Faulted Cluster Application to Online State	
Changed the explanation of maintenance mode.	7.2.2.6 Entering maintenance mode for Cluster Application	
Added the state icon to the node state table.	7.3.1 Monitoring the State of a Node	
Added the symbols that indicate the state of a cluster application.	7.3.2 Monitoring the State of a Cluster Application	
Changed the explanation of how to change PRIMECLUSTER operation attributes.	8.6 Changing PRIMECLUSTER Operation Attributes	
Changed the structure of this section to add explanations to change the user name and password for iRMC.	8.10 Changing IPMI Settings 8.10.1 Changing the IPMI IP Address 8.10.2 Changing the User Name and Password for iRMC	
Changed the structure of this section to add explanations to change the slot number of server blades.	8.11 Changing Blade Settings 8.11.1 Changing the IP Address of the Management Blade 8.11.2 Changing the Slot Number of Server Blades	

Revision	Location	Edition
Added the information of how to stop RMS.	8.12.1.1 Adding the shared disk device	
Added the explanations of when setting Gds resources is not necessary.	8.12.1.1 Adding the shared disk device	
Added the procedure to add the shared disk device.	8.12.1.1 Adding the shared disk device	
Added the procedure "5. Change the device name of the resource of the shared disk device."	8.12.2.1 Deleting the shared disk device	
Described the procedure required after execution of the cldelrsc command.	8.12.2.1 Deleting the shared disk device 8.12.3.1 Changing the shared disk device	
Added the notes on the GDS configuration file.	8.12.2.1 Deleting the shared disk device 8.12.3.1 Changing the shared disk device	
Added the procedure to change the shared disk device.	8.12.3.1 Changing the shared disk device	
Changed the descriptions of how to change the hardware.	8.12.3.2 Changing a network interface card used for the public LAN and the administrative LAN	
Added the notes when using message queues on applications.	A.6.2 RMS Configuration	
Changed the Gds Resource worksheet.	A.10.3 Gds Resource Worksheet	
Changed the notes on configuration on the virtual machine function.	A.12.1 When building a cluster system between guest OSes on one host OS A.12.2 When building a cluster system between guest OSes on multiple host OSes A.13.1 When building a cluster system between guest OSes on one host OS A.13.2 When building a cluster system between guest OSes on multiple host OSes	
Deleted the following commands from the list of the online manual page for RMS: hvdist hvgdmake hvrclev hvreset hvthrottle	B.6 RMS	
Changed the table of the search procedure.	D.1 Searching for a Message	
Changed the description of the operation environment for Configuration Update Service for SA.	G.2 Operation Environment	
Changed the notes on the cluster system in a VMware environment.	H.1 Cluster Systems in a VMware Environment	
Changed the descriptions of how to configure the /etc/hostid file to run I/O fencing properly.	H.2.1.1 Installation and Configuration of Related Software	
Added the following appendixes: "Appendix I Startup Daemons, and Port Numbers in PRIMECLUSTER" "Appendix J Using Firewall"	Appendixes	
Changed the descriptions of information on incompatibility.	K.1.2 hvdisp command K.1.10 Node state K.1.11 Operation Procedures and Displayed Items for Cluster Application Setup and Modification K.2.2 hvdisp command	

Revision	Location	Edition
	K.2.4 Node state	
	K.2.5 Operation Procedures and Displayed Items for Cluster Application Setup and Modification	
	K.3.2 Operation Procedures and Displayed Items for Cluster Application Setup and Modification	
Changed the description of the settings to enable the fault resource identification and the operator intervention request.	5.2 Setting Up Fault Resource Identification and Operator Intervention Request	6.8
Added S06clonltrc to the startup script list.	I.2 Startup Script List	
Changed the note when allowing the port number used by PRIMECLUSTER.	Appendix J Using Firewall	

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
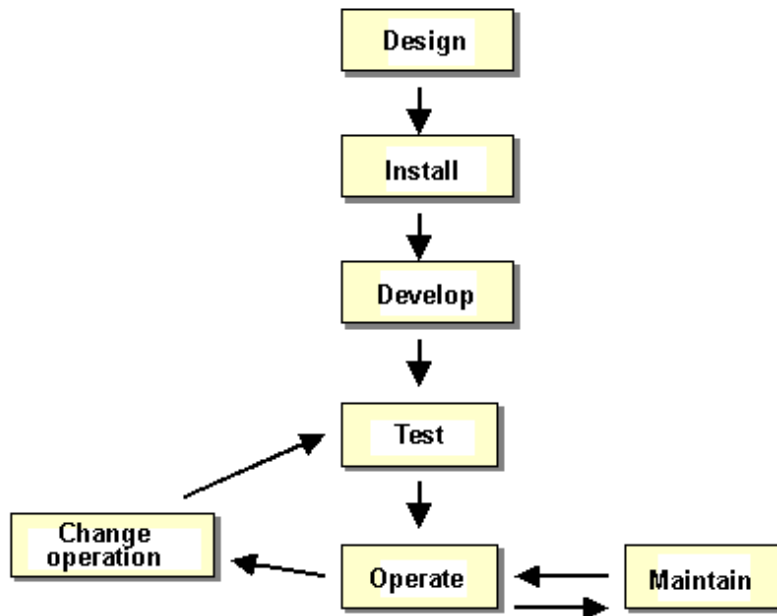
1.1 Planning

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 and whether to use the virtual machine function. This includes selecting the applications to be used and determining the required hardware resources, such as the number of hosts, networks, the number of cluster interconnect paths, and disk size.

    Up to 4 nodes can be added to one cluster system.

    We recommend that you use 2 or more cluster interconnects.

    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 4 nodes.

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

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) and the takeover address.

- Determine whether a user-defined RMS configuration script is to be used. Determine whether there are other items to be used as resources.

- 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 Determining 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
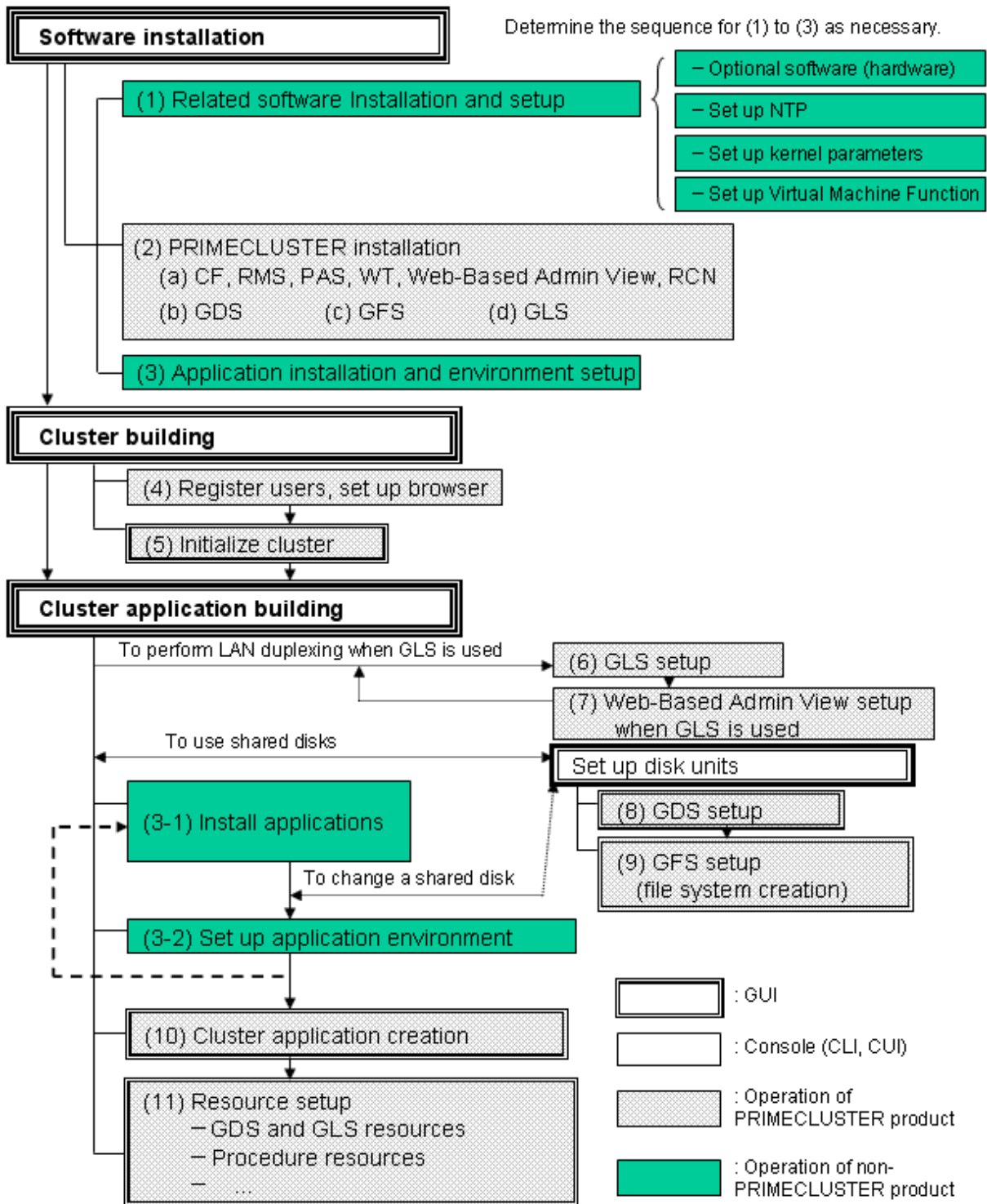
The abbreviations in the flowchart for PRIMECLUSTER system installation are explained below.

CF: Cluster Foundation

RMS: Reliant Monitor Services

WT: Wizard Tools

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) Related software Installation and setup	3.1 Installation and Setup of Related Software (for the Virtual Machine Function not used) 3.2 Installation and Setup of Related Software (for the Virtual Machine Function used)	Optional
(2) PRIMECLUSTER installation	3.3 PRIMECLUSTER Installation	Required
(3) Application installation and environment setup	3.4 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	6.2 Initial GLS Setup	Optional
(7) Web-Based Admin View setup when GLS is used		Note that it is required when IP address takeover and redundant line control are used.
(8) GDS setup	6.3 GDS Configuration Setup	Optional (required to use GDS)
(9) GFS setup (file system creation)	6.4 Initial GFS Setup	Optional (required to use GFS)
(10) Cluster application creation	6.7 Setting Up Cluster Applications	Required
(11) Resource setup	6.7.3 Setting Up Resources	Optional

1.3 Development

To monitor a user application using 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.

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
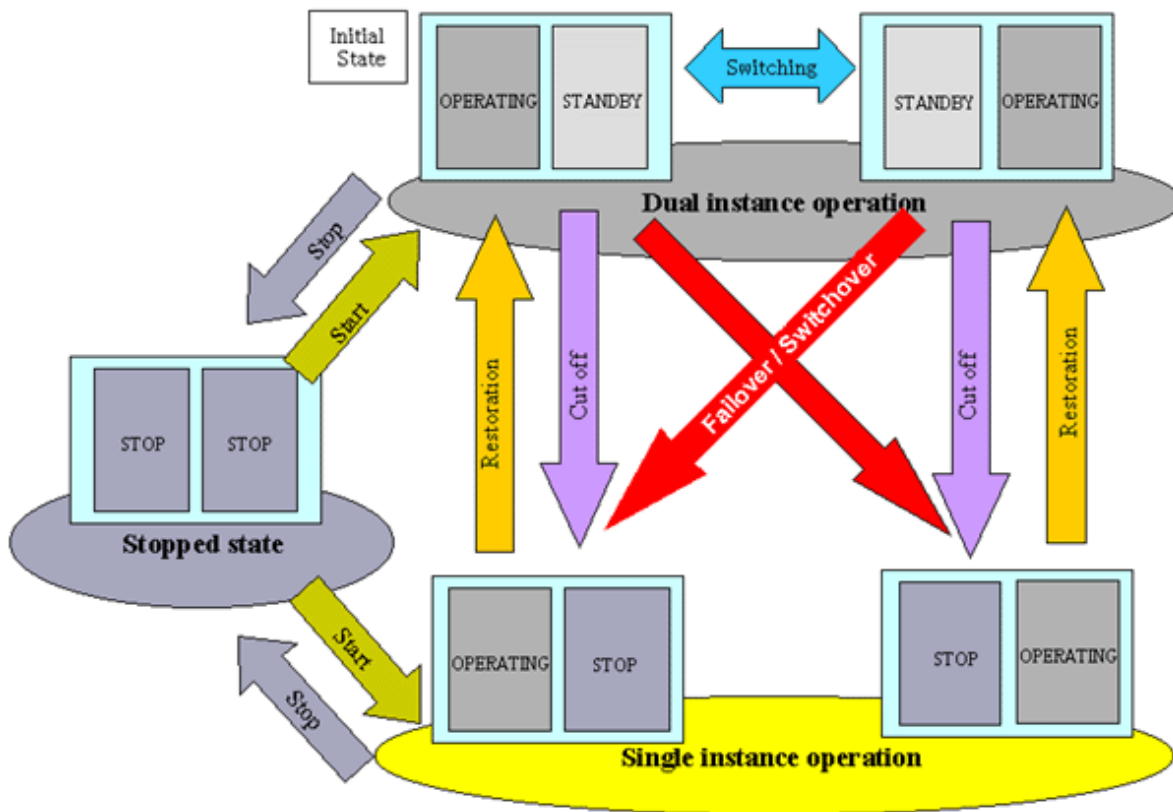
PRIMECLUSTER System State

	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.

	Description
	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 "6.7.2 Setting Up userApplication."

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 failoveror switchovertest 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 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(1)" 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(8)" command.

  - 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(1)" 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.

Test for forced shutdown of cluster nodes

Check that the settings of the shutdown facility work correctly.

Conduct a test to check that every node in the cluster is shut down at least once with the following viewpoints:

- Induce an OS error to check that the cluster node in which a failure has occurred is forcibly shut down.

- Disconnect the cluster interconnect to check that the cluster node with the lowest priority is forcibly shut down.

📒 Note

So as to detect an NIC linkdown event on both paths, disconnect the cluster interconnect.

For example, if the both nodes are connected through a switch instead of being connected directly, disconnect the two cluster interconnects from the same node side. If you perform a method of disconnection that does not allow for the detection of an NIC linkdown event on both paths, there will be time differences in detecting an error for each route and the node that detected the error first will have priority and stop peer node forcibly.

In addition, check that crash dumps for the cluster node which has been forcibly shut down are collected.

📘 See

- For information on the operation procedures for start, clear fault, failover, switchover, and stop, see "7.2 Operating the PRIMECLUSTER System."

- For information on IP address takeover, see "6.7.3.5 Setting Up Gls Resources."

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.

📘 See

For details for collecting information required for an investigation, see "Appendix C Troubleshooting."

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.

📘 See

For details on changing the operation mode, see "Chapter 8 Changing the Operation Configuration."

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 Oracle 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.

- **PRIMECLUSTER Clustering Base (CB)**

  Cluster foundation software that realizes a switchover-type cluster system.

  This product is only for PRIMERGY.

- **PRIMECLUSTER Lite Pack (LP)**

  Cluster foundation software that runs in a two-node configration only and on specific models only. It provides superior cost performance, supporting switching (HA) cluster and volume management through the operation mode, 1:1 standby and mutual standby.

  This product is only for PRIMERGY.

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

Components		Products			
Names	**Features**	**EE**	**HA**	**CB**	**LP**
PCLsnap	Refers to the function that collects information on a system or cluster that is needed to investigate the failures.	Y	Y	Y	Y
Web-Based Admin View	Refers to the function for realizing PRIMECLUSTER operations and monitoring with the GUI (management view).	Y	Y	Y	Y
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	Y	Y
Reliant Monitor Services (RMS)	Refers to the software monitoring function that is used to realize high-availability (HA) of the application that is to be executed within the cluster.	Y	Y	Y	Y
Wizard Tools	Refers to the function that is used to create an application that is to be controlled with RMS.	Y	Y	Y	Y
RAO	Refers to the function that is used to manage resources that run on PRIMECLUSTER.	Y	Y	Y	Y
SA	Refers to the shutdown agent function for which BMC, iRMC, Blade, and MMB are used.	Y	Y	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 shared disk device.	Y	Y	-	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.*"

2.2 System Design

You can use the following configuration of the cluster system. Use the "PRIMECLUSTER System Design Worksheet" to design the system in either case.

- Virtual Machine function not used

- Virtual Machine function used

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

📖 See
··
For details on the operation environment, see "2. Operation Environment" in the Installation Guide for PRIMECLUSTER.
··

2.2.1 Virtual Machine Function

The virtual machine function is a function for the realization of virtual machine environment. This function enables you to create multiple independent virtual machines on one physical machine by virtualizing resources such as the CPU, memory, network and disk that are installed on a physical machine.

The function is made up of one host OS and multiple guest OSes.
PRIMECLUSTER makes it possible to build a cluster system between guest OSes on multiple host OSes or between host OSes.

 Note

 ....

- In a KVM environment, read the "host OS" as "hypervisor," in a VMware environment, read "host OS" as "ESX host."

- When installing PRIMECLUSTER in a virtual machine environment, do not perform the following procedures:

    - Temporary stopping the Guest OS

    - Restart the Guest OS from a temporary stopped state

    - Restart or stop the hypervisor (KVM environment), ESX hosts (VMware), without stopping the Guest OS

 ....

 See

 ....

- For details on the virtual machine function in a Xen environment, see "PRIMEQUEST Virtual Machine Function User's Manual."

- For details on the virtual machine function in a KVM environment, see "Red Hat Enterprise Linux 6 Virtualization Getting Started Guide."

- For details on the virtual machine function in a VMware environment, see the documentation for VMware.

 ....

Virtual machine function configuration
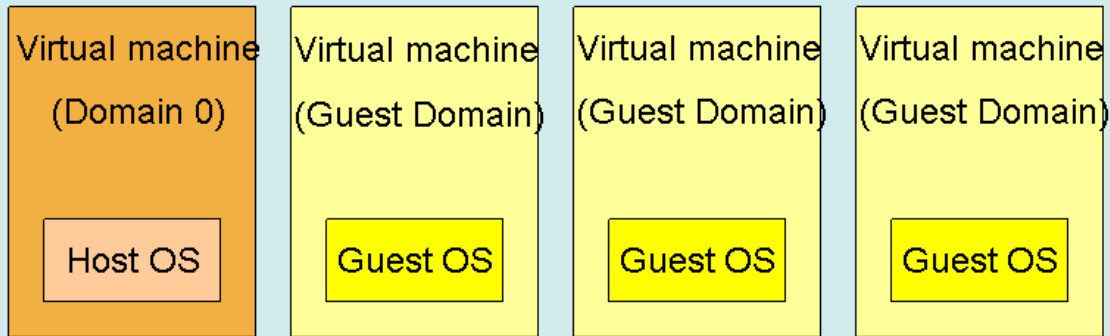
Cluster system in the virtual machine function

The virtual machine function provides the following methods to build a cluster system:

- Building a cluster system between guest OSes on one host OS

- Building a cluster system between guest OSes on multiple host OSes without the Host OS failover function

- Building a cluster system between guest OSes on multiple host OSes with the Host OS failover function

- Building a cluster system between host OSes

The table below shows uses and notes for each cluster system.

Method	Use	Note

Building a cluster system between guest OSes on one host OS	- In a single physical server, build a cluster environment between guest OSes with the same cluster application configuration as that of the physical environment.	- Since this cluster system is built with a single physical server, all cluster nodes will be shut down and operations will be suspended when the physical server fails. Therefore, this is not suitable to operate services.
		
		- Do not perform the following procedures.
		  - Temporary stopping the Guest OS
		  - Restart from a temporary stopped state.
		  - Restart or stop the hypervisor without stopping the guest OS.

Building a cluster system between guest OSes on multiple host OSes without the Host OS failover function	- To build a cluster environment between guest OSes with the same cluster application configuration as that of the physical environment and to use it as a development and test environment for cluster applications, or operate services	- If the host OS fails, manual switchover of cluster applications is required because guest OSes cannot be forcibly shut down.
		
		- Do not perform the following procedures.
		  - Temporary stopping the Guest OS
		  - Restart from a temporary stopped state.
		  - Restart or stop the hypervisor without stopping the guest OS.

Building a cluster system between guest OSes on	- To build a cluster environment between guest OSes with the same cluster application configuration as	- The ShutdownPriority attribute of RMS cannot be set on guest OSes.

Method	Use	Note

multiple host OSes with the Host OS failover function	that of the physical environment and operate services
	
	- To switch the cluster application when the host OS fails
		- The active node may be stopped depending on the settings of survival priority.
		
		- When a host OS failure is detected, the host OS is forcibly shut down. Then, all guest OSes on the failed host OS will stop.
		
		- Do not perform the following procedures.
		
		  - Temporary stopping the Guest OS
		
		  - Restart from a temporary stopped state.
		
		  - Restart or stop the hypervisor without stopping the guest OS.

Building a cluster system between host OSes	- To configure the following cluster environment and operate services: - Monitors the states of both guest and host OSes and starts guest OSes on other host OSes if an error occurs on guest OSes or host OSes.	- Hot-standby and scalable operations cannot be used. - When a host OS failure is detected, the host OS is forcibly shut down. Then, all guest OSes on the failed host OS will stop.

Note

When installing PRIMECLUSTER in a KVM environment and building a cluster system between guest OSes, do not perform the following procedures.

- Temporary stopping the Guest OS

- Restart from a temporary stopped state.

- Restart or stop the hypervisor without stopping the guest OS.

When building a cluster system between guest OSes on one host OS

This configuration allows you to run a cluster system on a physical machine. This is effective when you verify that the userApplication runs on PRIMECLUSTER.
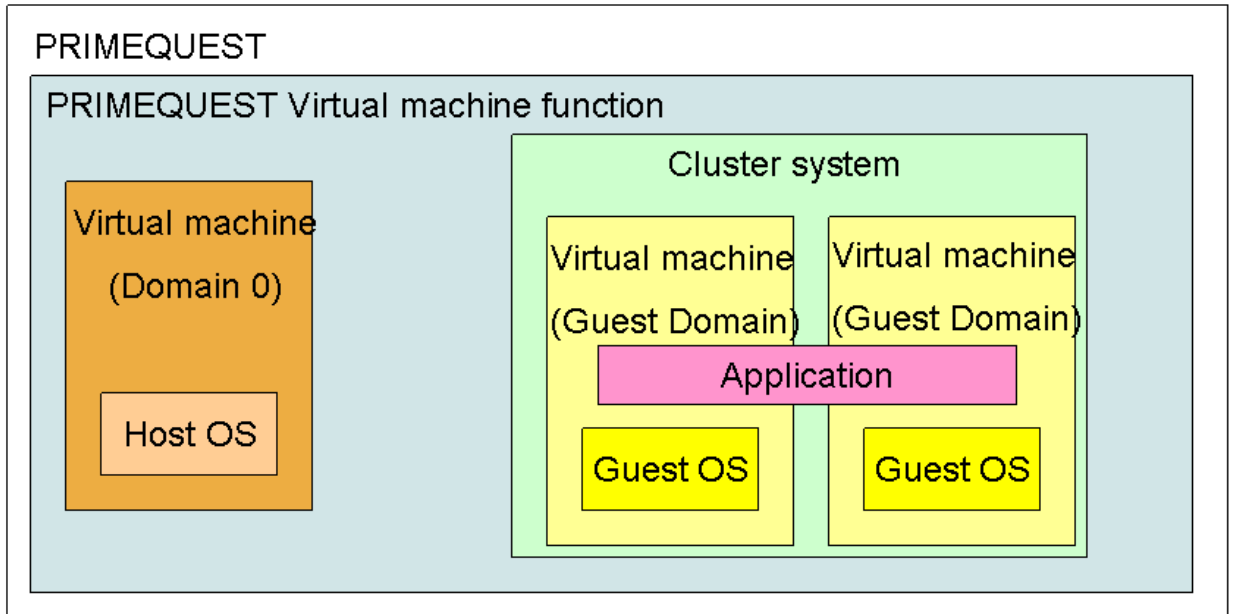
When building a cluster system between guest OSes on multiple host OSes

This configuration allows you to continue work processes by a failover even if hardware such as a network or a disk fails.
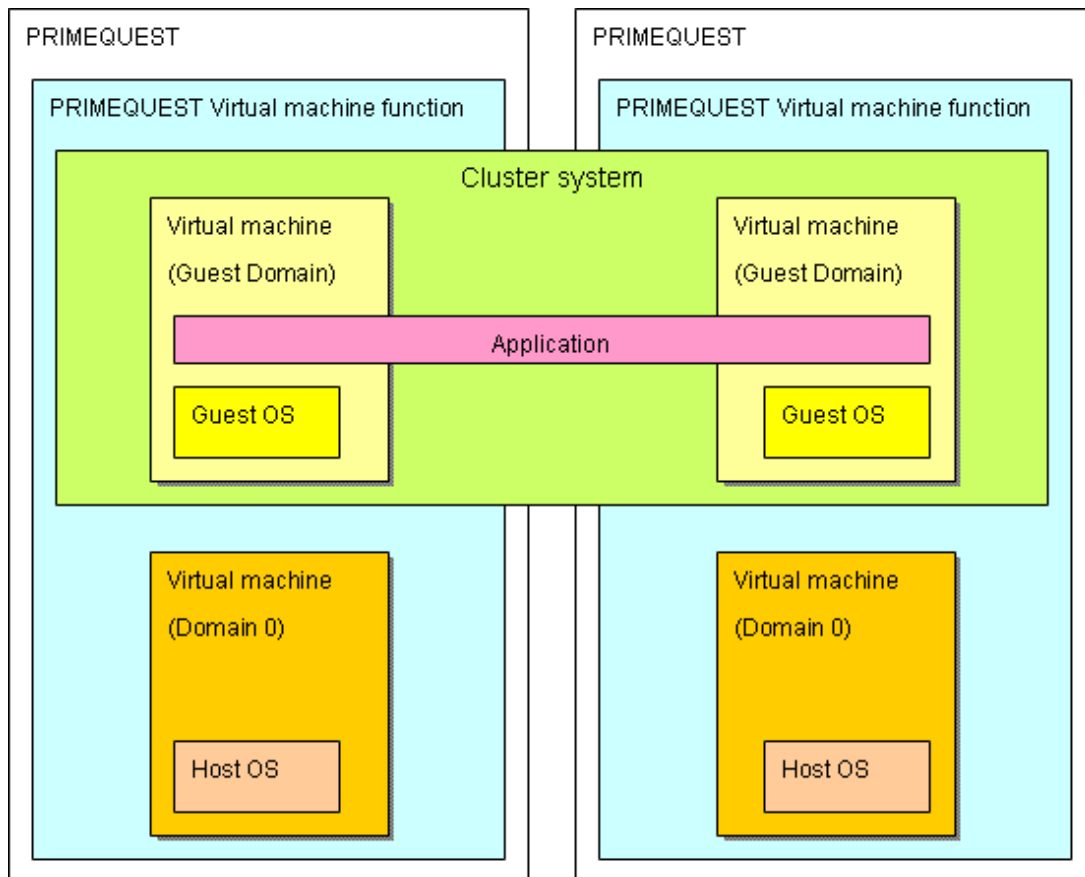
Note

....

If the host OS cannot run, the node enters the LEFTCLUSTER state and manual switching by an operator may be required. For details, see "7.4.2 Corrective Action in the event of the LEFTCLUSTER state when the virtual machine function is used" or "7.2 Operating the PRIMECLUSTER System."

....

When building a cluster system between guests on multiple host OSes in a Xen environment or a KVM environment, you can use a function that automatically perform a failover when the host OS fails (Host OS failover function).

Host OS failover function

When building a cluster between guests in different units on a virtual machine, if an error occurs in the host OS, nodes in the cluster may enter the LEFTCLUSTER state. Host OS failover function allows for automatically switching cluster applications on the guest OSes in the case of the following errors in a cluster system between guests in different units in a Xen environment or a KVM environment.

- Panic of the host OS

- Hang-up of the host OS (slowdown)

This function is achieved by linking PRIMECLUSTER installed on the host OS with guest OSes.

Note that there are some precautions for operations, for example, setting the priority of RMS is not available by using this function. Then, you should take these precautions into consideration when designing the system.

Note

....

- When creating a cluster application for a guest OS, do not set the ShutdownPriority attribute of RMS.

- When a host OS failure is detected, the host OS is forcibly shut down. Then, all guest OSes on that host OS with a failure will stop regardless of whether they are clusters or not.

- Do not build any cluster applications on the host OS.

....

Figure 2.1 Cluster system using the Host OS failover function on the virtual machine
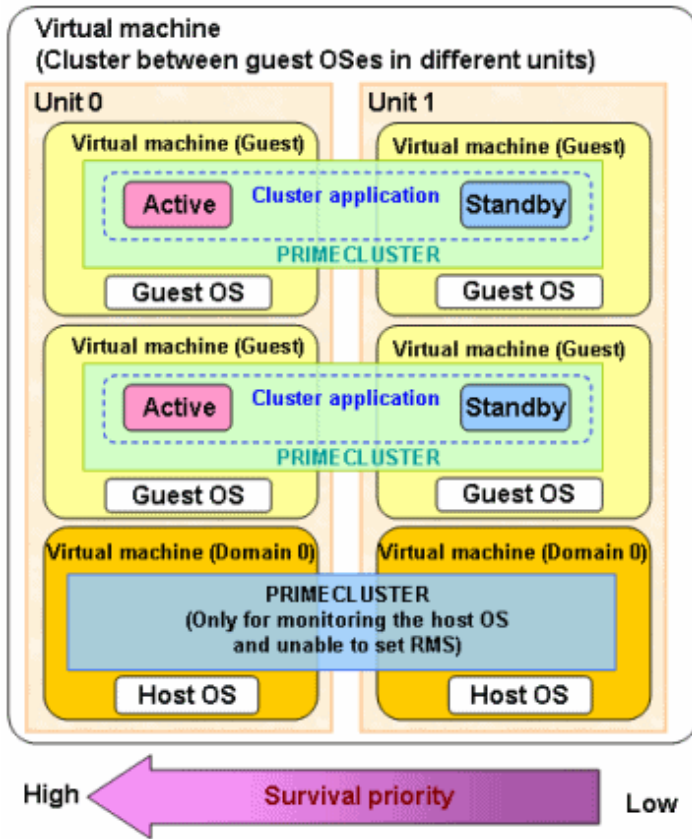
Figure 2.2 Failover image in the case of host OS failure
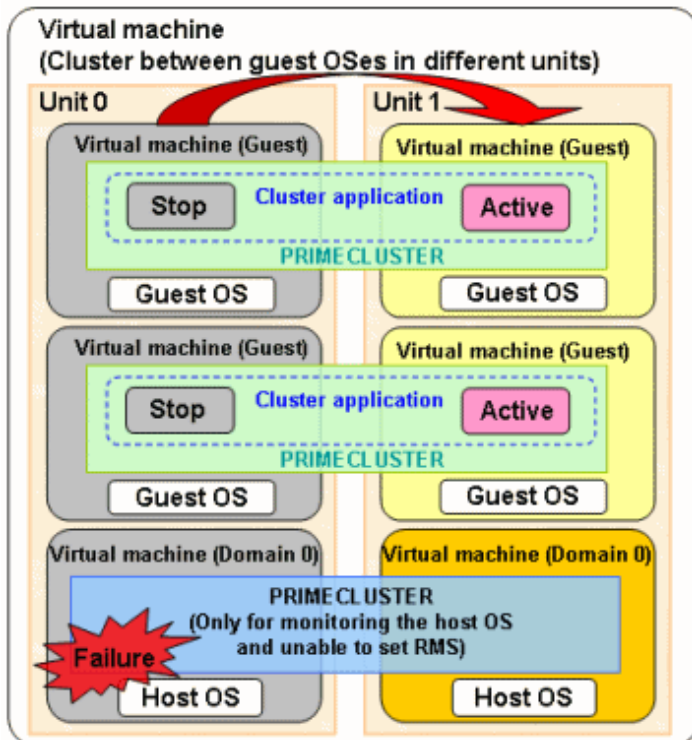
Moreover, in a RHEL5-Xen environment on PRIMEQUEST, you can replicate the cluster system by doing live migration of guest OSes in which PRIMECLUSTER is installed or by copying the virtual machine image.

When building a cluster system between host OSes

This configuration allows you to continue work processes with starting up other guest OSes on other host OSes by a failover of the host OSes even if hardware such as a network or a disk fails.
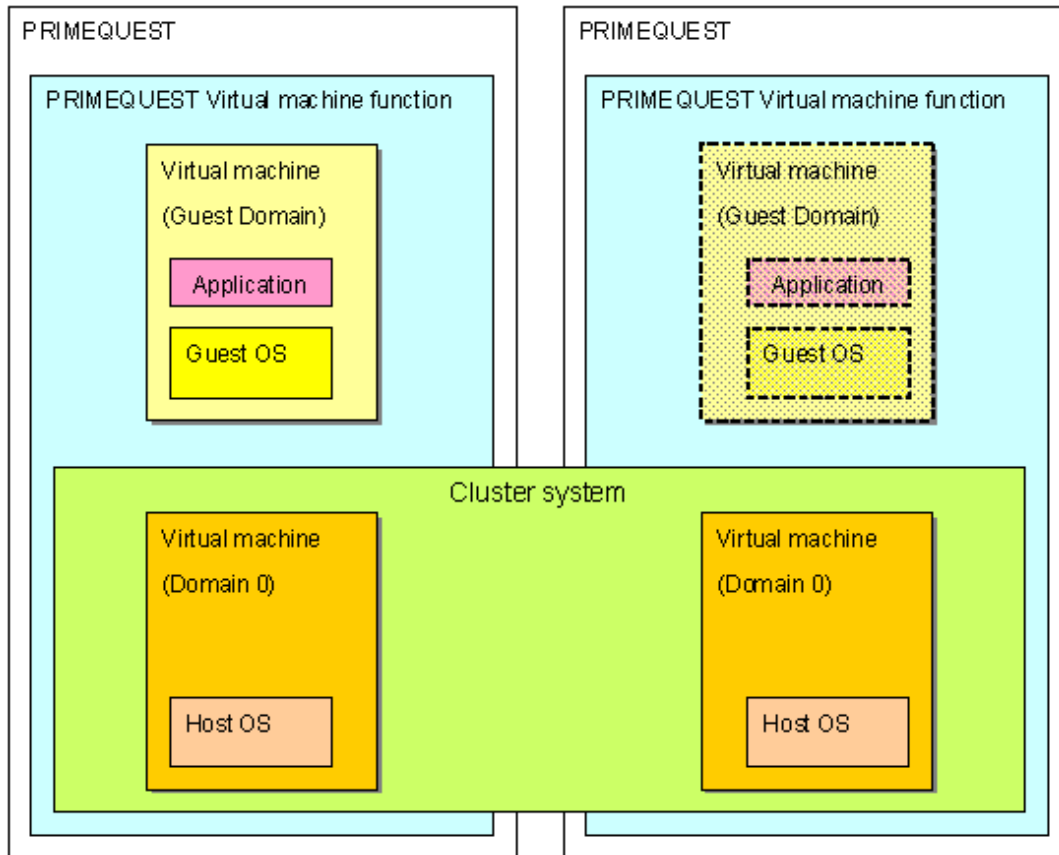
2.3 Determining the Cluster System Operation Mode

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)
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)

Classification	Operation mode	Number of cluster applications	Number of nodes
Single-node cluster operation	-	1	1

Note

If an operating node is powered off, the failover function may fail to work. Take corrective action as follows.

- Connect all nodes to UPS.

- Connect the node, and BMC or iRMC to different power sources to prevent simultaneous power-off.

2.3.1 Standby Operation

The topologies for standby operation are as shown below.

Information

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
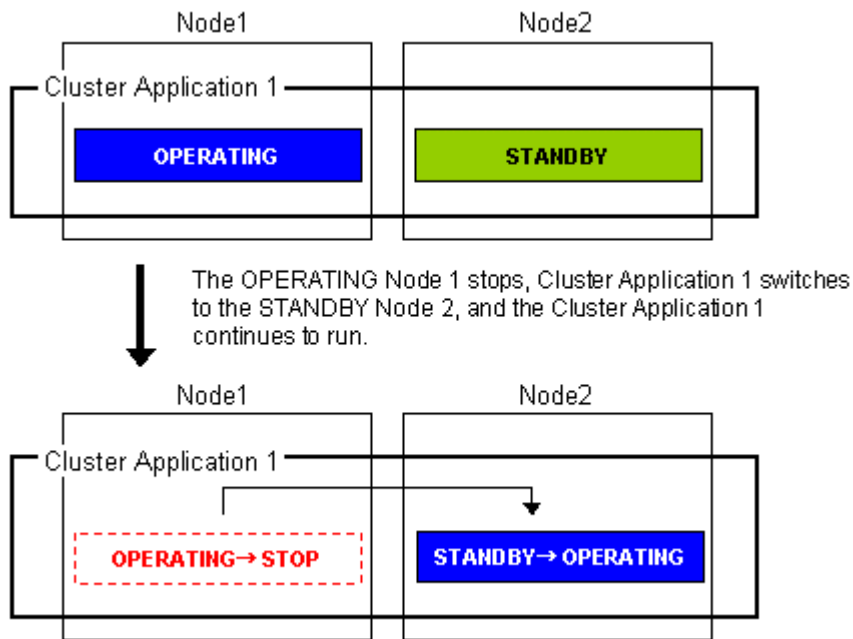
Mutual standby

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 cluster applications. The operating cluster application has its own standby in each other's node.

Advantage

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

Disadvantage

- 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
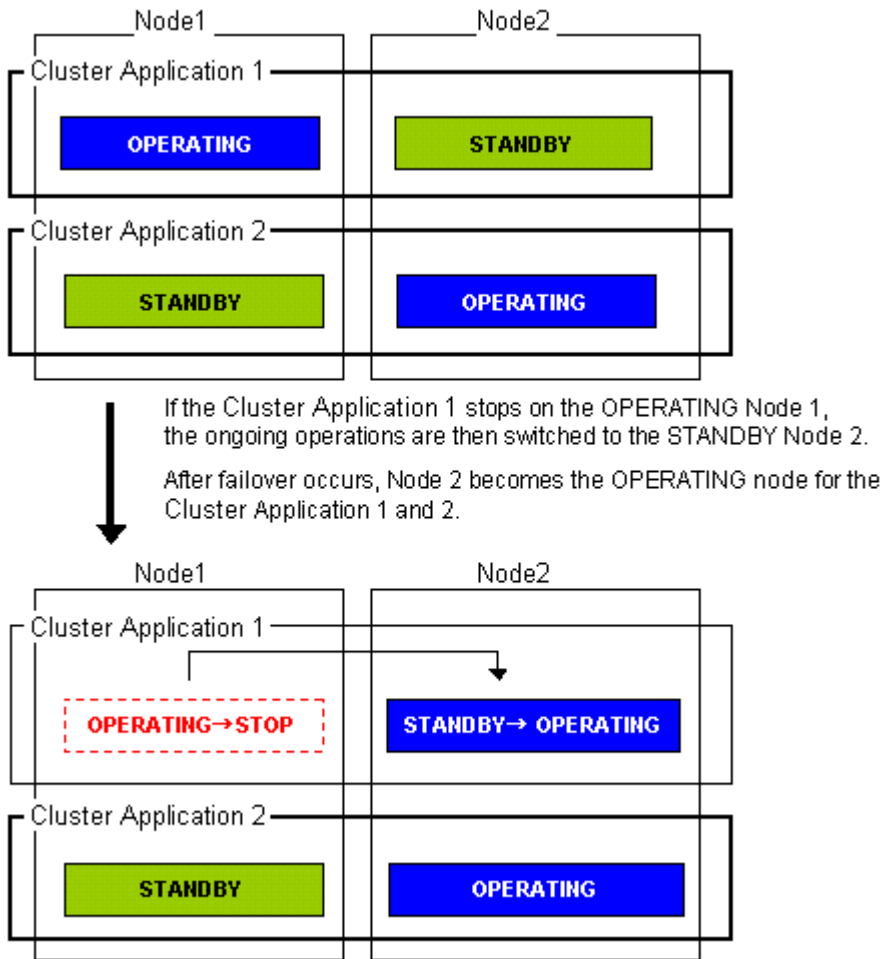
N:1 standby

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 cluster applications, the STANDBY cost can be reduced when the number of cluster applications is large.

Disadvantages

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

Failover image

The OPERATING Node 1 stops, Cluster Application 1 switches to the STANDBY Node 2, and the Cluster Application 1 continues to run.
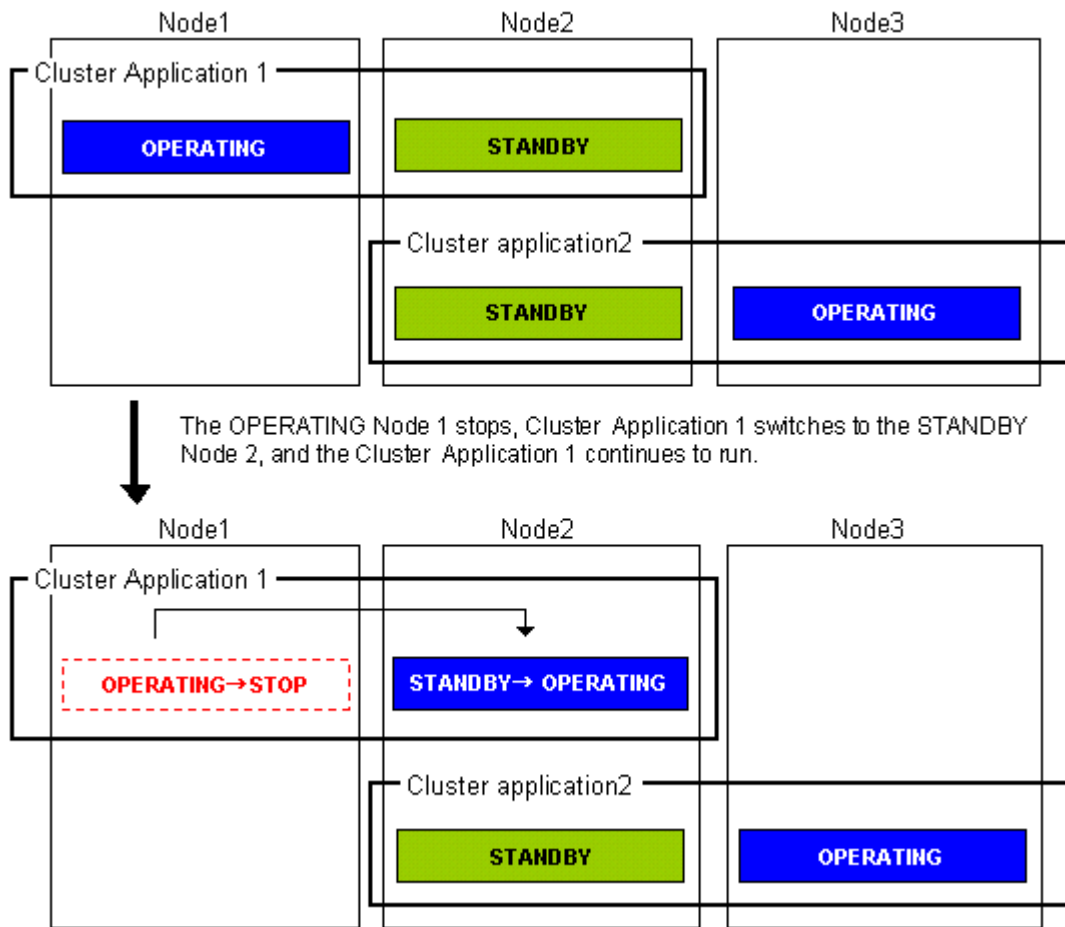
Cascade (using one cluster application)

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 cluster 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.

Disadvantage

- 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.
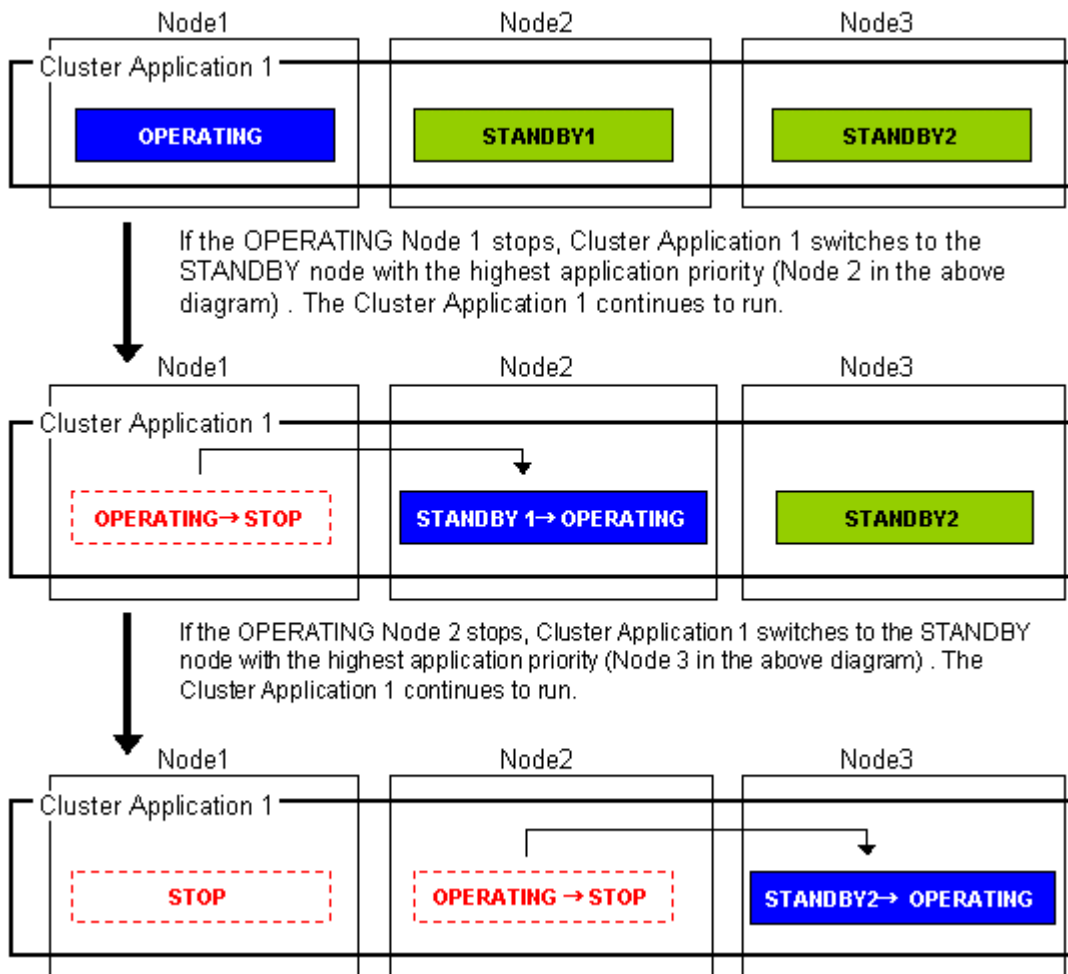
Priority transfer (application of N:1 standby)

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.

Disadvantages

- 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
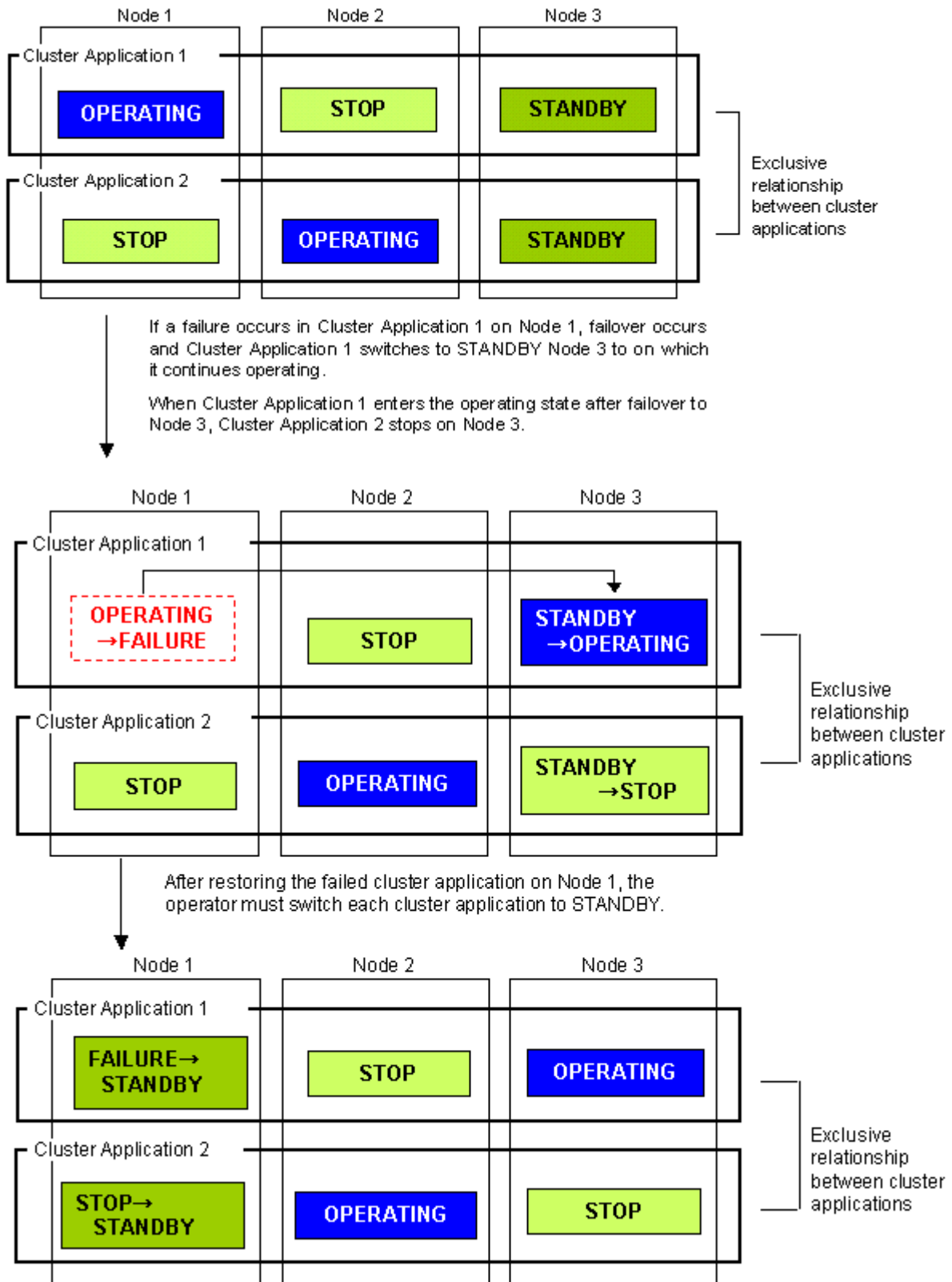
2.3.2 Scalable Operation

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.

Advantage

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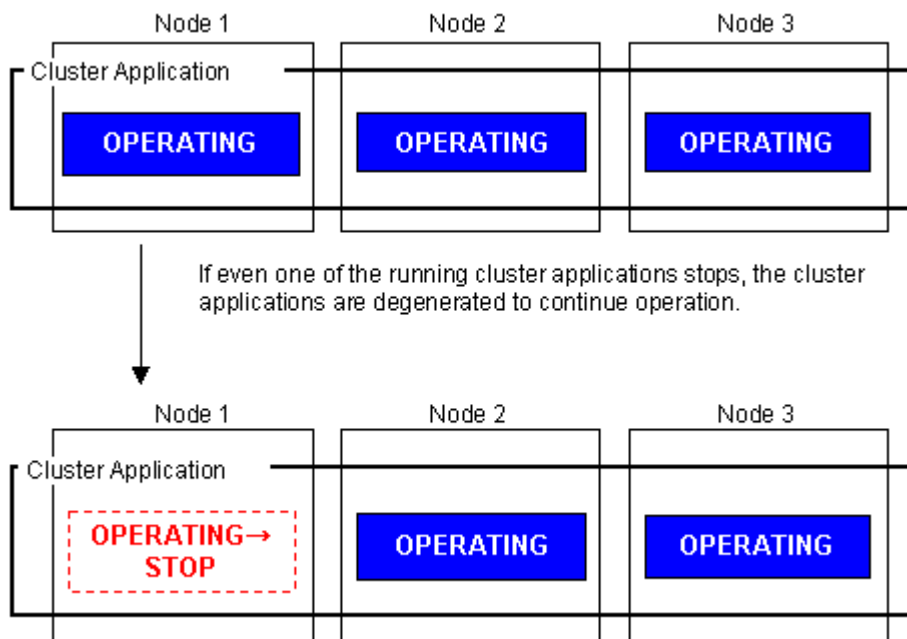
If even one of the running cluster applications stops, the cluster applications are degenerated to continue operation.

Note

Scalable operation can be used in combination with some PRIMECLUSTER-related products. For information on the related products, see the "Table 10-1 PRIMECLUSTER Product List."

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.

- 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.
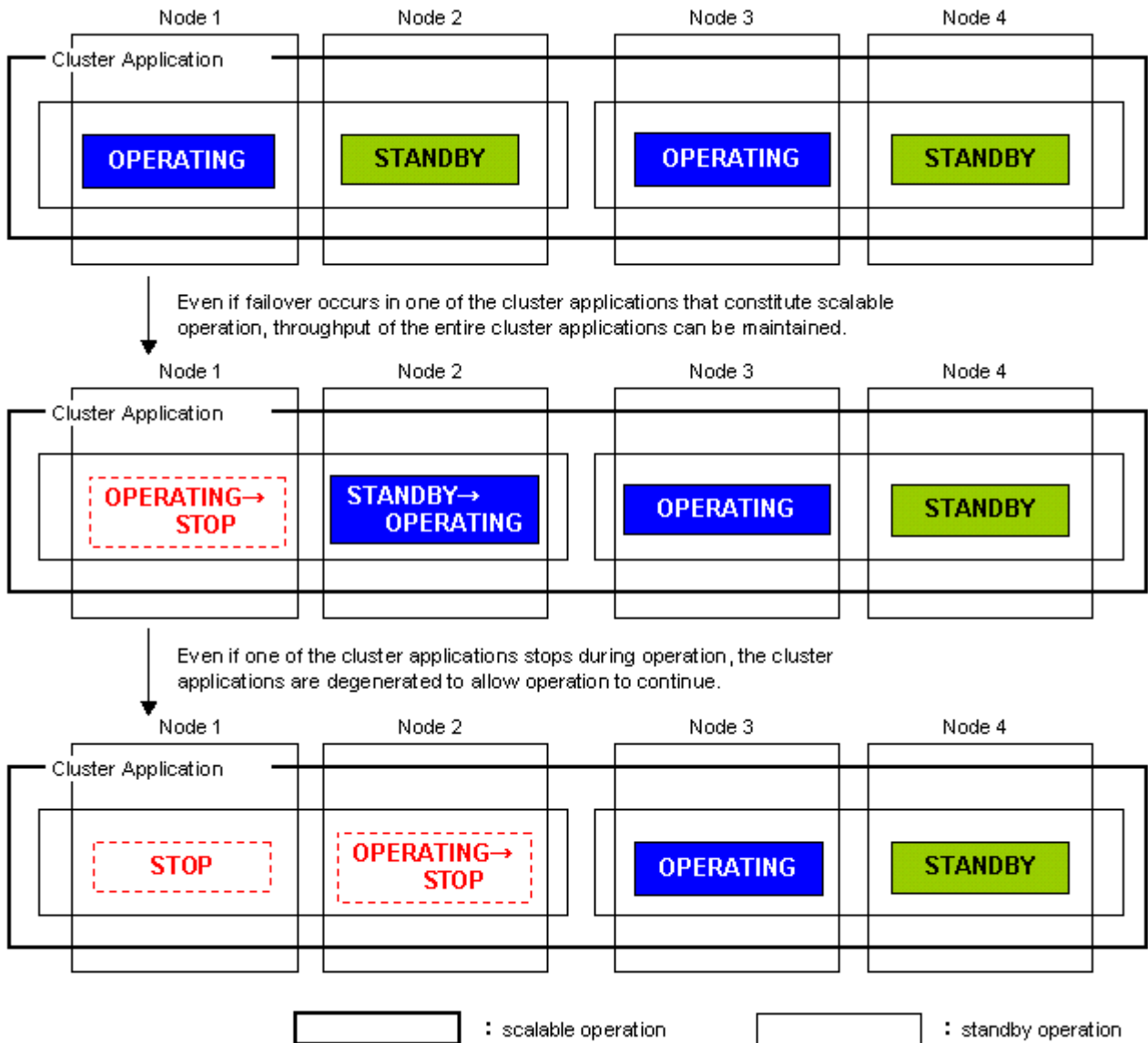
⚑ Note

Scalable operation can be used in combination with some PRIMECLUSTER-related products. For information on the related products, see the "Table 10-1 PRIMECLUSTER Product List."

2.3.3 Single-Node Cluster Operation

This section explains the topologies used for single-node cluster operation:

Definition

- It is an operation mode in which a cluster system consists of one node.

Advantages

- This operation mode enables monitoring and control jobs on the node in a single node configuration.

- If an error occurs in the resource to which the AUTORECOVER attribute is set, the availability can be improved by automatically rebooting the system for restoration.

- You can also use this mode as a development environment for creating and testing cluster applications.

Disadvantage

- If a hardware error occurs, jobs will be stopped.

Failover image

No failover occurs in the single-node cluster operation.
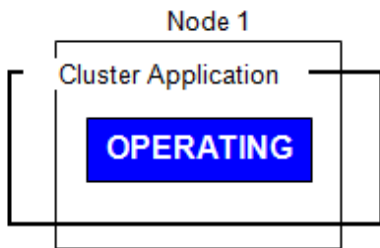
Automatically restored in the case of resource error.
(When the AUTORECOVER attribute is set.)

![Note]

Note

. . . .

You need at least one network interface card for cluster interconnect that is used in PRIMECLUSTER also in the single-node cluster operation.

. . . .

2.4 Determining the Web-Based Admin View Operation Mode

Determine the operation mode of Web-Based Admin View according to your PRIMECLUSTER configuration.

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]

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*."

. . . .

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.

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 Linux(R) Workstation (with Intel CPU).

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 on the Linux (R) servers.

Cluster nodes

Cluster nodes construct the PRIMECLUSTER system.
Cluster nodes run on the Linux (R) servers.

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 servers.

This model supports configurations where the number of nodes is relatively small and which does not require a specific cluster management server.

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.
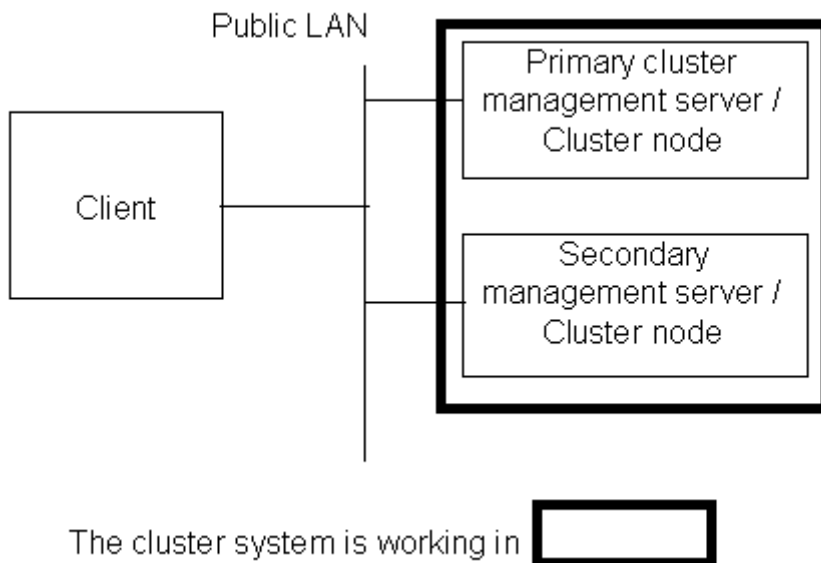
Topology where separate LANs are used

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."

Client

Public network

Primary cluster
management server /
Cluster node

Secondary cluster
management server /
Cluster node

Public LAN
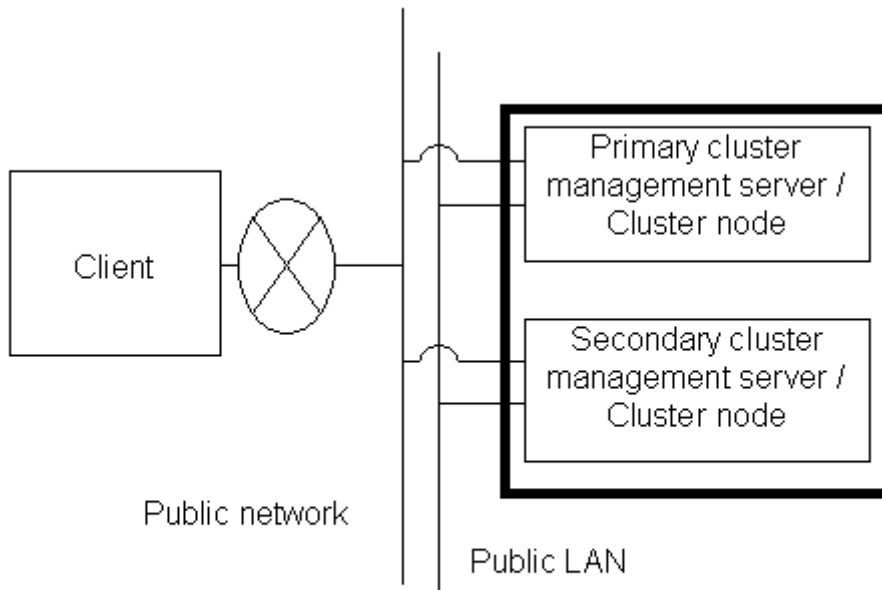
The cluster system is working in ☐

3-tier model (PRIMERGY)

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.

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.

Public LAN

Client

Primary cluster
management server

Secondary cluster
management server

Cluster node

Cluster node

The cluster system is working in
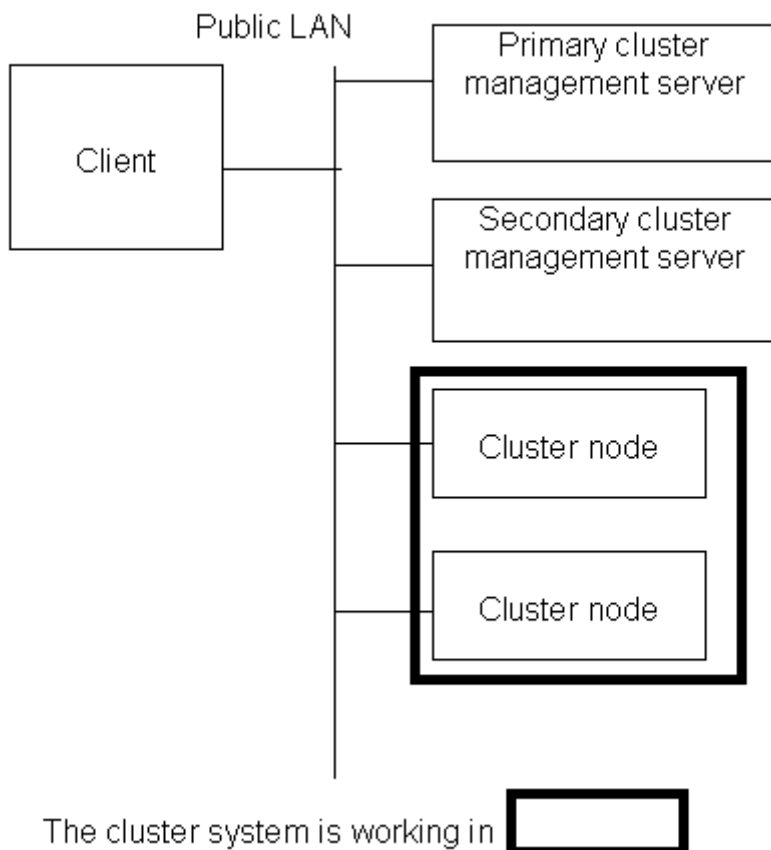
Topology where separate LANs are used

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 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 [ ]
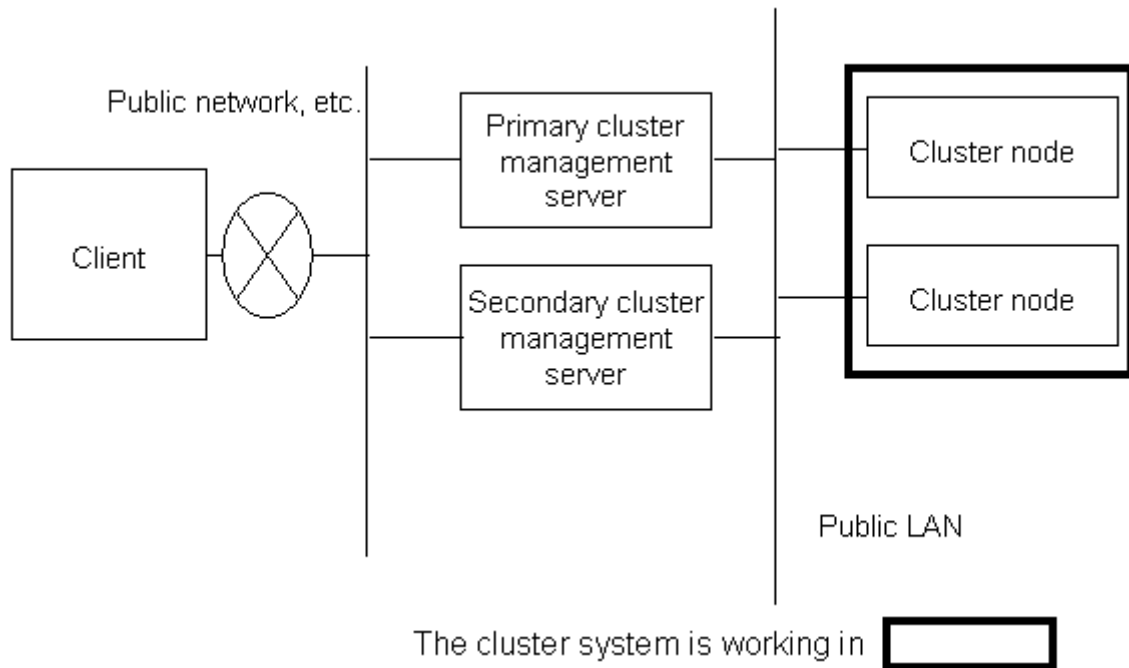
2.5 Determining the Failover Timing of a Cluster Application

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 "6.7.2 Setting Up userApplication."

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 software products related to PRIMECLUSTER on each node.

📝 Note
. . . .

- For the security, set "No Firewall" when a Red Hat Enterprise Linux is installed or when the setup command is executed. If Firewall has already been set for the security, change the setting to "No Firewall" If the "Firewall" setting is left as is, the clsetup (setting of the resource database) command will operate abnormally.

- PRIMECLUSTER is guaranteed to work with a maximum umask value of "022." Do not modify the umask value.

. . . .

This chapter describes the following topics:

- Installation and configuration of related software

- PRIMECLUSTER installation (overview)

- Application installation and environment configuration

3.1 Installation and Setup of Related Software (for the Virtual Machine Function not used)

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.
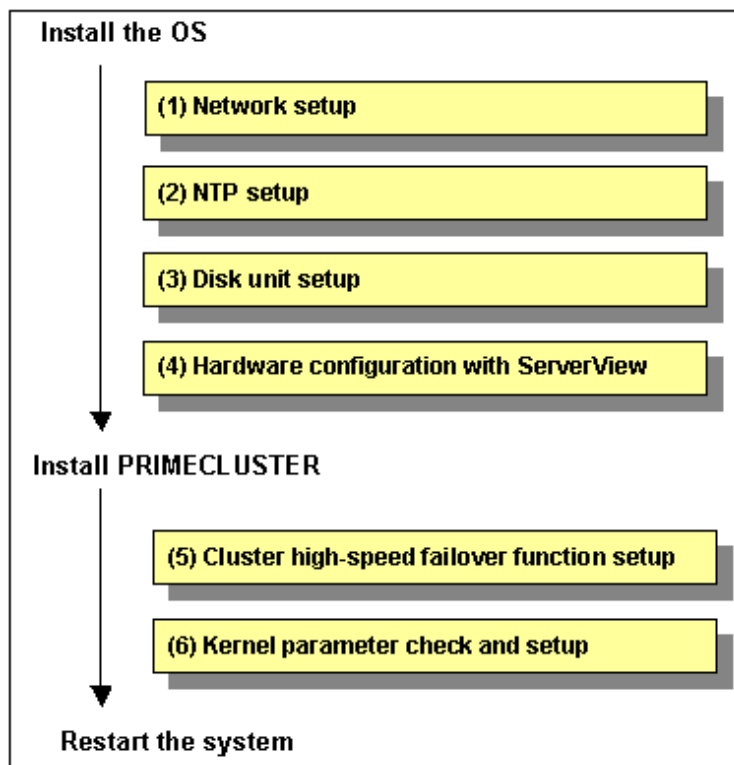
1. Network setup

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

   For details, see "3.1.1 Setting Up the Network."

2. 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.

3. Disk device setup

When using the shared disk device, you need to install and set up the software product. This disk unit setup should be done before installing the OS.

For details, see "3.1.2 Setting Up Disk Units."

4. Hardware configuration with ServerView

You need to configure software that monitors hardware before installing the OS.

For details, see "3.1.3 Setting Up Hardware Monitoring with ServerView (PRIMERGY)."

5. Cluster high-speed failover function setup

You need to configure software and hardware that enables cluster high-speed failover after installing the OS and PRIMECLUSTER.

For details, see "3.1.4 Setting Up the Cluster High-Speed Failover Function."

6. 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 restarting after installing PRIMECLUSTER.

For details, see "3.1.5 Checking the Kernel Parameters."

3.1.1 Setting Up the Network

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

 See
. . . .
For information on changing the public LAN and administrative LAN that the PRIMECLUSTER system uses, see "8.2 Changing an IP Address on the Public / Administrative LAN."
. . . .

 Information
. . . .
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.
. . . .

3.1.2 Setting Up Disk Units

When using the shared disk device, you need to install and set up the software product.

Installation and Setup of Related Software

Install and set up the software products (ETERNUS Multipath driver) required for using the shared disk device. For details on the installation and setup procedure, see the "Software Information" for ETERNUS Multipath Driver.

 Note
. . . .
For immediate cluster failover if an I/O device where the system volume is placed fails

In the default setting of the ext3 file system, even if an I/O device where the system volume is placed fails, a cluster failover does not occur and the system operation may continue based on the data stored on the memory.

If you want PRIMECLUSTER to trigger a cluster failover immediately in the event that an I/O device where the system volume is placed fails, perform the following setting:

Setting

For each partition included in the system volume, specify "errors=panic" to the mount option of ext3.

Example: To set it in /etc/fstab (when /, /var, and /home exist in one system volume)

```
LABEL=/      /     ext3 errors=panic 1 1
LABEL=/boot /boot ext3 errors=panic 1 2
LABEL=/var  /var  ext3 errors=panic 1 3
LABEL=/home /home ext3 errors=panic 1 4
```

However, an immediate cluster failover may not become available due to taking time for an I/O error to reach theext3 file system. The regularly writing to the system volume enhances the detection frequency of I/O error.

3.1.3 Setting Up Hardware Monitoring with ServerView (PRIMERGY)

ServerView supports monitoring of the fan, temperature sensor, and power supply in addition to watchdog monitoring of the OS.

ServerView allows you to specify the action which will be done in the event of failure for each monitored target. For example, if "Shut down the server immediately" is selected and an error is detected, the failed node will be shut down. The ongoing operations on the failed node are then quickly switched to the standby node.

Note

No failover will be triggered by PRIMECLUSTER even if the operating system hangs up as long as communication with cluster interconnect is performed normally.

This state can be avoided by enabling watchdog timer monitoring.

For information about behavior setup, see the ServerView Operations Manager manual.

3.1.4 Setting Up the Cluster High-Speed Failover Function

3.1.4.1 PRIMERGY

Overview

If heartbeat monitoring fails because of a node failure, PRIMECLUSTER shutdown facility removes the failed node. If this occurs during crash dump collection, you might not be able to acquire information for troubleshooting.

The cluster high-speed failover function prevents node elimination during crash dump collection, and at the same time, enables the ongoing operations on the failed node to be quickly moved to the another node.

kdump
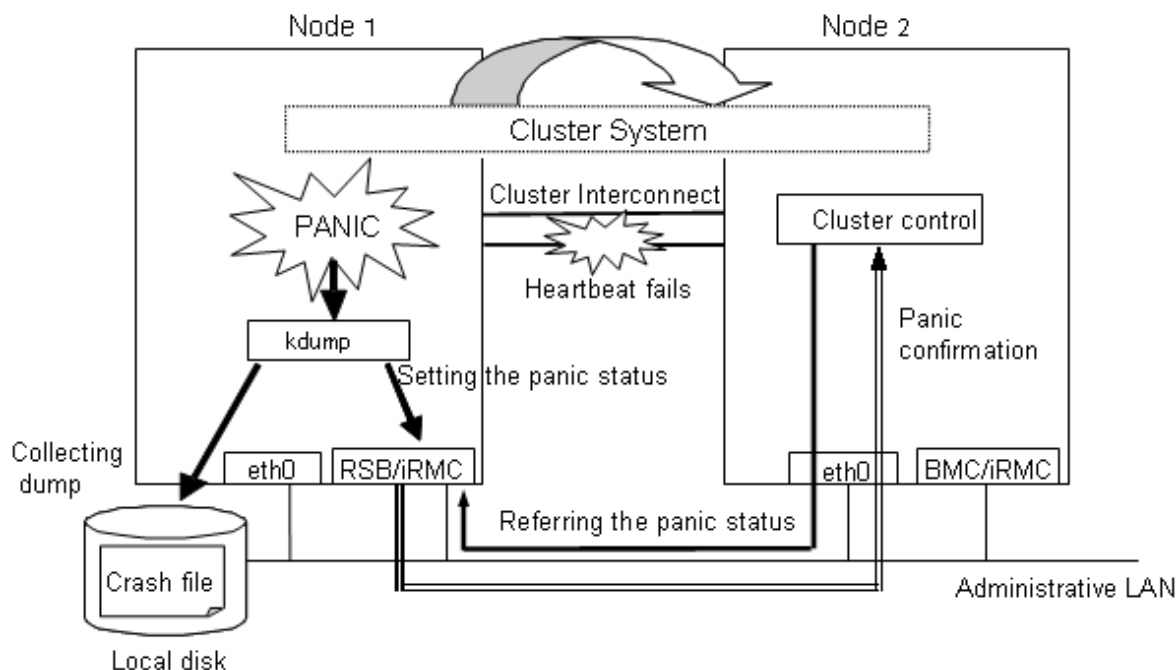
As shown in the above figure, the cluster fast switching function allows for panic status setting and reference through BMC (Baseboard Management Controller) or iRMC when a heartbeat monitoring failure occurs. The node that detects the failure can consider that the other node is stopped and takes over ongoing operation without eliminating the node that is collecting crash dump.

![Note]

- If you reboot the node that is collecting crash dump, collection of the crash dump will fail.

- When the node completes collecting the crash dump after it gets panicked, the behavior of the node follows the setting of kdump.

Required setting for the kdump shutdown agent

1. Configure kdump

   When using kdump, it is necessary to configure the kdump.

2. Check kdump

   Check if the kdump is available. If not, enable the kdump using the "runlevel(8)" and "chkconfig(8)" commands.

   - Check the current run level using the "runlevel(8)" command.

     Example)

     ```
     # /sbin/runlevel
     
     N 3
     ```

     The above example shows that the run level is 3.

   - Check if the kdump is available using the "chkconfig(8)" command.

     Example)

     ```
     # /sbin/chkconfig --list kdump
     
     kdump  0:off 1: off 2: off 3: off 4: off 5: off 6: off
     ```

     The above example shows that the kdump of the runlevel 3 is currently off.

- If the kdump is off, enable it by executing the "chkconfig(8)" command.

> `# ` **`/sbin/chkconfig kdump on`**

Then, start it by executing the service command.

> `# ` **`/sbin/service kdump start`**

Prerequisites for the other shutdown agent settings

After you completed configuring the kdump shutdown agent, set the IPMI (Inteligent Platform Management Interface) or BLADE server.

Information

....

The IPMI shutdown agent is used with the hardware device in which BMC or iRMC is installed.

....

Prerequisites for the IPMI shutdown agent settings

Set the following for the IPMI user.

- User ID

- Password

- IP address

For details, see the "*User Guide*" provided with the hardware and the "*ServerView Operations Manager manual.*"

Prerequisites for the Blade shutdown agent settings

Set the following for the BLADE server:

- Install ServerView

- Set SNMP community

- Set an IP address of the management blade

For details, see the operation manual provided with the hardware and the "*ServerView Operations Manager manual.*"

3.1.4.2 PRIMEQUEST

When PRIMEQUEST is used, if an error occurs in one of the nodes of the cluster system, the PRIMECLUSTER shutdown facility uses the two methods described below to detect that error. For details, see "3.3.1.7 PRIMECLUSTER SF" in the "*PRIMECLUSTER Concept Guide.*"

(1) Node status change detection through MMB units (asynchronous monitoring)

(2) Heartbeat failure between cluster nodes (NSM: node status monitoring) (fixed-cycle monitoring)

Asynchronous monitoring of (1) allows node errors to be detected immediately, and failover occurs at a higher speed than when detected by fixed-cycle monitoring.
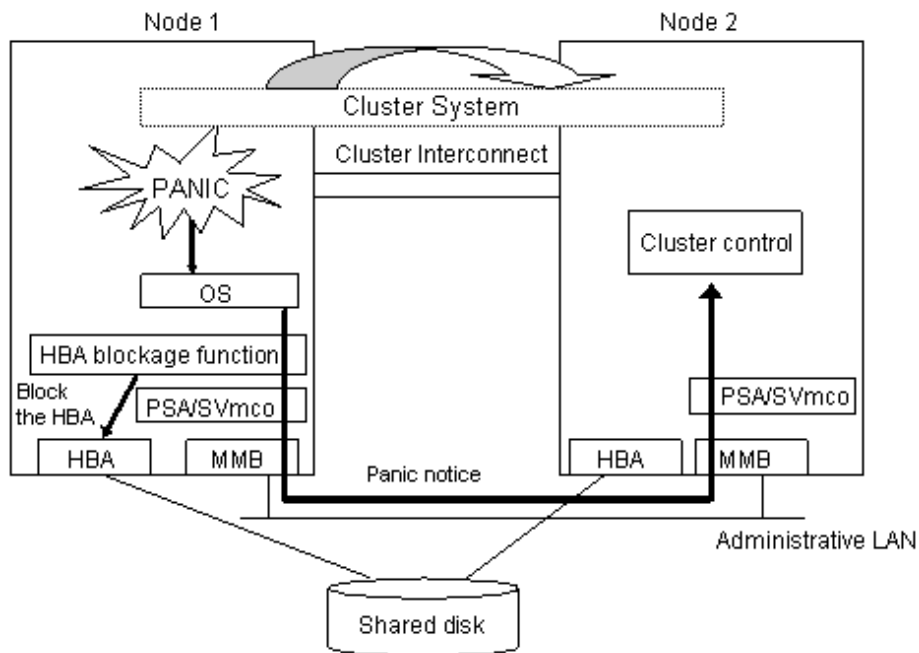
As shown in the above figure, if a panic occurs, the cluster control facility uses the MMB units to receive the panic notice. This allows the system to detect the node panic status faster than it would be a heartbeat failure.

📑 See

. . . .

PRIMEQUEST allows you to set the panic environment so that a crash dump is collected if a panic occurs.

For details about the PRIMEQUEST dump function, setup method, and confirmation method, refer to the following manuals:

- "PRIMEQUEST 1000 Series Installation Manual"

- "PRIMEQUEST 1000 Series ServerView Mission Critical Option User Manual"

. . . .

To use asynchronous monitoring (1), you must install software that controls the MMB and specify appropriate settings for the driver. This section describes procedures for installing the MMB control software and setting up the driver, which are required for realizing high-speed failover.

1. Installing the HBA blockage function and the PSA/SVmco

   The HBA blockage function and the PSA/SVmco report node status changes through the MMB units to the shutdown facility. Install the HBA blockage function and the PSA/SVmco before setting up the shutdown facility. For installation instructions, refer to the following manuals:

   - "PRIMEQUEST 1000 SERIES HBA blockage function USER'S GUIDE"

   - "PRIMEQUEST 1000 Series Installation Manual"

   - "PRIMEQUEST 1000 Series ServerView Mission Critical Option User Manual"

2. Setting up the PSA/SVmco and the MMB units

   The PSA/SVmco and MMB must be set up so that node status changes are reported properly to the shutdown facility through the MMB units. Set up the PSA/SVmco units before setting up the shutdown facility. For setup instructions, refer to the following manuals:

   - "PRIMEQUEST 1000 Series Installation Manual"

   - "PRIMEQUEST 1000 Series ServerView Mission Critical Option User Manual"

   You must create an RMCP user so that PRIMECLUSTER can link with the MMB units.

In all PRIMEQUEST instances that make up the PRIMECLUSTER system, be sure to create a user who uses RMCP to control the MMB. To create a user who uses RMCP to control the MMB, log in to MMB Web-UI, and create the user from the "Remote Server Management" window of the "Network Configuration" menu. Create the user as shown below.

- Set [Privilege] to "Admin".

- Set [Status] to "Enabled".

For details about creating a user who uses RMCP to control the MMB, Refer to "PRIMEQUEST 1000 Series Tool Reference" provided with the computer.

The user name created here and the specified password are used when the shutdown facility is set up. Record the user name and the password.

Note

The MMB units have two types of users:

- User who controls all MMB units

- User who uses RMCP to control the MMB

The user created here is the user who uses RMCP to control the MMB. Be sure to create the correct type of user.

3. Setting up the HBA blockage function

Note

Be sure to carry out this setup when using shared disks.

If a panic occurs, the HBA units that are connected to the shared disks are closed, and I/O processing to the shared disk is terminated. This operation maintains data consistency in the shared disk and enables high-speed failover.

On all nodes, specify the device paths of the shared disks (GDS device paths if GDS is being used) in the HBA blockage function command, and add the shared disks as targets for which the HBA function is to be stopped. If GDS is being used, perform this setup after completing the GDS setup. For setup instructions, see the "PRIMEQUEST 1000 SERIES HBA blockage function USER'S GUIDE".

4. Setting the I/O completion wait time

To maintain consistent I/O processing to the shared disk if a node failure (panic, etc.) occurs and failover takes place, some shared disk devices require a fixed I/O completion wait time, which is the duration after a node failure occurs until the new operation node starts operating.

The initial value of the I/O completion wait time is set to 0 second. However, change the value to an appropriate value if you are using shared disk devices that require an I/O completion wait time.

Information

ETERNUS Disk storage systems do not require an I/O completion wait time. Therefore this setting is not required.

Specify this setting after completing the CF setup. For setting instructions, see "Setting the I/O Completion Wait Time".

Note

If an I/O completion wait time is set, the failover time when a node failure (panic, etc.) occurs increases by that amount of time.

3.1.5 Checking the Kernel Parameters

You need to edit the values of the kernel parameters, depending on 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.6 Kernel Parameter Worksheet."
....

Note
....
To enable modifications, you need to restart the system after installation.
....

3.2 Installation and Setup of Related Software (for the Virtual Machine Function used)

After installing the PRIMECLUSTER-related software, you need to set up the operating system, hardware, etc. that will be used and administered.

See
....
When using the virtual machine function in a VMware environment, see "Appendix H Using PRIMECLUSTER in a VMware Environment."
....

3.2.1 When building a cluster system between guest OSes on one host OS

This section describes how to install and set up related software when building a cluster system between guest OSes on one host OS.

Perform the following as needed.

1. Host OS setup

   If you plan to operate a guest OS as part of a cluster, set up the required disk devices, virtual bridges, virtual SCSI devices, user accounts, and guest OS initializations on the host OS.

   Perform this setup after installing the operating system.

   For details, see "3.2.1.1 Host OS setup."

2. NTP setup (host OS/guest OS)

   This setup is for synchronizing the time on each node comprising the cluster system, which is necessary when creating a cluster.

   This setup should be performed on the host OS and guest OS before installing PRIMECLUSTER.

   ![See] See
   ....
   For details on settings, see "Red Hat Enterprise Linux 6 Virtualization Host Configuration and Guest Installation Guide."
   ....

3. Guest OS Setup

   On the guest OSes comprising the cluster system, you need to set up the network, including IP addresses of the public LAN and administrative LAN.

   Perform this setup after installing the operating system.

   For details, see "3.2.1.2 Guest OS setup."

4. Kernel parameter check and setup

   To operate the PRIMECLUSTER-related software, you need to edit the values of the kernel parameters based on the environment.

   Perform this setup before rebooting the installed PRIMECLUSTER.

   For details, see "3.2.1.3 Kernel parameter check."

3.2.1.1 Host OS setup

If you plan to operate a guest OS as part of a cluster, you need to set up the host OS.

Perform the following setup after installing the operating systems on the host OS and guest OS.

1. Setting up disks and related devices

   You need to complete the setup of disks and related devices beforehand.

   Installing and setting up related software

   Install and set up the software product (ETERNUS Multipath Driver) required for using system disk of the guest OS on the host OS. For how to install and set up the related software, see "Software Information" for ETERNUS Multipath Driver.

 Note
. . . .

   For immediate cluster failover if an I/O device where the system volume is placed fails

   In the default setting of the ext3 file system, even if an I/O device where the system volume is placed fails, a cluster failover does not occur and the system operation may continue based on the data stored on the memory.

   If you want PRIMECLUSTER to trigger a cluster failover immediately in the event that an I/O device where the system volume is placed fails, perform the following setting:

   Setting

   For each partition included in the system volume, specify "errors=panic" to the mount option of ext3.

   Example: To set it in /etc/fstab (when /, /var, and /home exist in one system volume)

   ```
   LABEL=/      /     ext3 errors=panic 1 1
   LABEL=/boot /boot ext3 errors=panic 1 2
   LABEL=/var  /var  ext3 errors=panic 1 3
   LABEL=/home /home ext3 errors=panic 1 4
   ```

   However, an immediate cluster failover may not become available due to taking time for an I/O error to reach theext3 file system. The regularly writing to the system volume enhances the detection frequency of I/O error.
. . . .

   Setting up virtual SCSI devices (Xen environment)

   For using a shared disk on a guest OS, you need to set up a virtual SCSI device.

   Setting up virtual SCSI devices (KVM environment)

   The following shows the setup procedure for virtual SCSI devices in a KVM environment:

   1. Stop the guest OS.

   2. Select the stopped guest OS with the Virtual Machine Manager and click the [Open] button in the toolbar.

   3. Click in the toolbar to display the detailed information of hardware.

   4. Select a virtual disk (VirtIO Disk) from the hardware list in the left.

   5. In the [Virtual disk] window, perform the following settings and click [Apply].

      - Select the Shareable check box.

      - Select 'none' for the cache model.

6. Check the version of the libvirt package on the host OS by using the rpm(8) command.

```
# rpm -qi libvirt
```

7. If the version of the libvirt package is libvirt-0.9.4-23.el6_2.4 or later, change the device attribute from disk to lun, which is set in the guest setting file (/etc/libvirt/qemu/guestname.xml) on the host OS.

```
# virsh edit guestname
```

Example before change

```
     :
  <disk type='block' device='disk'>
    <driver name='qemu' type='raw'/>
    <source dev='/dev/disk/by-id/scsi-1FUJITSU_30000085002B'/>
    <target dev='vdb' bus='virtio'/>
    <shareable/>
    <address type='pci' domain='0x0000' bus='0x00' slot='0x06' function='0x0'/>
  </disk>
     :
```

Example after change

```
     :
  <disk type='block' device='lun'>
    <driver name='qemu' type='raw'/>
    <source dev='/dev/disk/by-id/scsi-1FUJITSU_30000085002B'/>
    <target dev='vdb' bus='virtio'/>
    <shareable/>
    <address type='pci' domain='0x0000' bus='0x00' slot='0x06' function='0x0'/>
  </disk>
```

8. Start the guest OS.

🖅 Note
....

- The "by-id" name should be specified when creating the virtual disk in the guest OS.

- The virtual disk cannot be partitioned when creating it in the guest OS.
....

Mirroring the guest OS system disk

To mirror the guest OS system disk, set up mirror volumes in the local class or the shared class created on the host OS for the guest OS.

🖳 See
....
For details on settings, see "Red Hat Enterprise Linux 6 Virtualization Administration Guide."
....

2. Setting up the virtual bridge (administrative LAN/public LAN/cluster interconnect)

   a. Xen environment

   For the network interfaces, including the administrative LAN, public LAN and cluster interconnect, that are used by virtual domains, you need to set up virtual bridges for the virtual networks beforehand.

   The following virtual bridge configurations are used in cluster systems.

   - gextbr:

      Virtual bridge for the public LAN

- xenbr:

   Virtual bridge for the administrative LAN

- gintbr:

   Virtual bridge for the cluster interconnect

Note

. . . .

Use only the gintbr configuration for the virtual bridge of the cluster interconnect.

. . . .

Example: Define the virtual bridges (gintbr4 and gintbr5) for the cluster interconnect in the script file (network-bridge-pcl) of the previously set up virtual bridge.

gintbr4,gintbr5:

   Virtual bridge for the cluster interconnect

xenbr0,xenbr1:

   Virtual bridge for the administrative LAN

gextbr2,gextbr3:

   Virtual bridge for the public LAN

network-bridge-pcl:

   The file name of the script to call the virtual bridge creation script

(1) Check that the script (/etc/xen/scripts/gint-network-bridge) exists. If not, create it.

See

. . . .

For details, see the manual of "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function".

. . . .

(2) Edit the script that defines the virtual bridge (network-bridge-pcl) which is located in "/etc/xen/scripts" on the host OS.

```
#!/bin/sh
#
# Sample of Create/Delete virtual bridge
#
# $1 start : Create virtual bridge
#    stop  : Delete virtual bridge
#    status: Display virtual bridge information

# Exit if anything goes wrong
set -e

command=$1

glsxenscript=/opt/FJSVhanet/local/sbin/hanetxen
xenscript=/etc/xen/scripts/network-bridge
xenscriptgext=/etc/xen/scripts/gext-network-bridge
xenscriptgint=/etc/xen/scripts/gint-network-bridge  ***added

# op_start:subscript for start operation #
op_start () {
        $xenscript $command vifnum=0 netdev=eth0
        $xenscript $command vifnum=1 netdev=eth1
        $xenscriptgext $command extnum=2 netdev=eth2
        $xenscriptgext $command extnum=3 netdev=eth3
        $xenscriptgint $command intnum=4  ***added
```

```
        $xenscriptgint $command intnum=5  ***added

}

# op_stop:subscript for stop operation #
op_stop () {
        op_start $command
}

case "$command" in
        start)
                # Create your virtual bridge
                $glsxenscript stop
                op_start
                $glsxenscript start
        ;;

        stop)
                # Delete virtual bridge
                $glsxenscript stop
                op_stop
        ;;

        status)
                # display virtual bridge information
                $xenscript status
        ;;

        *)
                echo "Unknown command: $command" >&2
                echo 'Valid commands are: start, stop, status' >&2
                exit 1

esac
```

(3) Set execute permissions

Execute the command below and set the execute permissions to the script (network-bridge-pcl).

```
# cp  network-bridge-pcl /etc/xen/scripts
# cd /etc/xen/scripts
# chmod +x network-bridge-pcl
```

(4) Register with the xend service

Check the "network-script" parameter of the xend service configuration file (/etc/xen/xend-config.sxp). Set the "network-bridge-pcl" if it is not set.

```
# Your default ethernet device is used as the outgoing interface, by default.
# To use a different one (e.g. eth1) use
#
# (network-script 'network-bridge netdev=eth1')
#
# The bridge is named xenbr0, by default.  To rename the bridge, use
#
# (network-script 'network-bridge bridge=<name>')
#
# It is possible to use the network-bridge script in more complicated
# scenarios, such as having two outgoing interfaces, with two bridges, and
# two fake interfaces per guest domain.  To do things like this, write
# yourself a wrapper script, and call network-bridge from it, as appropriate.
#
(network-script network-bridge-pcl)
```

b. KVM environment

For the network interfaces, including the administrative LAN, public LAN and cluster interconnect, that are used by virtual domains, you need to set up virtual bridges for the virtual networks beforehand.

(1) Setting up a virtual bridge for the administrative LAN

Edit the /etc/sysconfig/network-scripts/ifcfg-ethX file as follows:

```
DEVICE=ethX
HWADDR=XX:XX:XX:XX:XX:XX
BOOTPROTO=none
ONBOOT=yes
BRIDGE=brX
```

📔 **Note**

. . . .

For HWADDR, set the MAC address of the network interface card you are using.

. . . .

Create the interface setting file, /etc/sysconfig/network-scripts/ifcfg-brX, for the virtual bridge.

```
DEVICE=brX
TYPE=Bridge
BOOTPROTO=static
IPADDR=xxx.xxx.xxx.xxx
NETMASK=xxx.xxx.xxx.xxx
ONBOOT=yes
```

📔 **Note**

. . . .

For IPADDR and NETMASK, set IP addresses and netmasks to connect to the external network.

. . . .

(2) Setting up a virtual bridge for the public LAN

Edit the /etc/sysconfig/network-scripts/ifcfg-ethX file as follows:

```
DEVICE=ethX
HWADDR=XX:XX:XX:XX:XX:XX
BOOTPROTO=none
ONBOOT=yes
BRIDGE=brX
```

📔 **Note**

. . . .

For HWADDR, set the MAC address of the network interface card you are using.

. . . .

Create the interface setting file, /etc/sysconfig/network-scripts/ifcfg-brX, for the virtual bridge.

```
DEVICE=brX
TYPE=Bridge
ONBOOT=yes
```

(3) Setting up a virtual bridge for the cluster interconnect

Create the interface setting file, /etc/sysconfig/network-scripts/ifcfg-brX, for the virtual bridge.

```
DEVICE=brX
TYPE=Bridge
BOOTPROTO=static
ONBOOT=yes
```

3. Setting the guest OS in the host OS (in a KVM environment)

In a KVM environment, perform the following settings to stop the guest OS normally if the host OS is shut down by mistake while the guest OS running.

Define the following two values in /etc/sysconfig/libvirt-guests. When values are already defined, change them to the following values:

- ON_SHUTDOWN=shutdown

- SHUTDOWN_TIMEOUT=300

Specify the timeout duration (seconds) for shutdown of the guest OS to SHUTDOWN_TIMEOUT. Estimate the length of time for shutting down the guest OS and set the value. When multiple guest OSes are set, set the time whichever is greater. The above is an example when the time is 300 seconds (5 minutes).

🗒 Note

- When setting /etc/sysconfig/libvirt-guests, do not describe the setting values and comments on the same line.

- When changing the settings in /etc/sysconfig/libvirt-guests during operation, make sure to follow the procedure in "8.9.2.3 Changing the Settings in /etc/sysconfig/libvirt-guests."

4. Creating a user ID

🅿 Point

This user ID will be the one used by the shutdown facility to log in to the host OS to force shut down the nodes. This user ID and password are used for configuring the shutdown facility.

a. Xen environment

In a Xen environment, create the required general user ID (FJSVvmSP) on the host OS for the guest OS control by PRIMECLUSTER.

```
# useradd FJSVvmSP
```

b. KVM environment

In a KVM environment, you need to set up a user for the shutdown facility for the guest OS control by PRIMECLUSTER.

(1) Creating a general user ID (optional)

Create a general user ID (optional) for the shutdown facility in the hypervisor.

```
# useradd <User ID>
```

(2) Setting up the "sudo" command

You need to set up the "sudo" command so that the general user ID (optional) for the shutdown facility can execute the command as the root user.

Use the "visudo" command to set up the general user ID created in step (1) so that the user can execute the command without entering the password.

```
# visudo
```

Example

```
<User ID>   ALL=(root) NOPASSWD: ALL
```

5. Initializing the guest OS

After completing the installation of a guest OS and the installation of the bundled software on the guest OS, initialize the guest OS.

3.2.1.2 Guest OS setup

On the guest OSes comprising the cluster system, you need to set up the network, including IP addresses of the public LAN and administrative LAN.

Perform this setup on all guest OSes of a cluster.

3.2.1.3 Kernel parameter check

You need to edit the values of the kernel parameters based on the environment.

Target node:

 All nodes on 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.

3.2.2 When building a cluster system between guest OSes on multiple host OSes without using Host OS failover function

This section describes how to install and set up related software when building a cluster system between guest OSes on multiple host OSes without using Host OS failover function.

Perform the following as needed.

```
Install the OS

    (1) Host OS setup

    (2) NTP setup

    (3) Guest OS setup

Install PRIMECLUSTER

    (4) Kernel parametar check and setup

Restart the system
```

1. Host OS setup

   If you plan to operate a guest OS as part of a cluster, set up the required disk devices, virtual bridges, virtual SCSI devices, user accounts, and guest OS initializations on the host OS.

   Perform this setup after installing the operating system.

   For details, see "3.2.1.1 Host OS setup."

2. NTP setup (host OS/guest OS)

   This setup is for synchronizing the time on each node comprising the cluster system, which is necessary when creating a cluster.

   This setup should be performed on the host OS and guest OS before installing PRIMECLUSTER.

   ### 🗐 See
   ....
   For details on settings, see "Red Hat Enterprise Linux 6 Virtualization Host Configuration and Guest Installation Guide".
   ....

3. Guest OS Setup

   On the guest OSes comprising the cluster system, you need to set up the network, including IP addresses of the public LAN and administrative LAN.

   Perform this setup after installing the operating system.

   For details, see "3.2.1.2 Guest OS setup."

4. Kernel parameter check and setup

To operate the PRIMECLUSTER-related software, you need to edit the values of the kernel parameters based on the environment.

Perform this setup before rebooting the installed PRIMECLUSTER.

For details, see "3.2.1.3 Kernel parameter check."

See

• •

For details on the configuration in a Xen environment, see "A.12.2 When building a cluster system between guest OSes on multiple host OSes".
For details on the configuration in a KVM environment, see "A.13.2 When building a cluster system between guest OSes on multiple host OSes".

• •

3.2.2.1 Host OS setup

If you plan to operate a guest OS as part of a cluster, you need to set up the host OS.

Perform the following setup after installing the operating systems on the host OS and guest OS.

1. Setting up disks and related devices

   You need to complete the setup of disks and related devices beforehand.

   - Installing and setting up related software

   Install and set up the software product (ETERNUS Multipath Driver) required for using system disk of the guest OS on the host OS. For how to install and set up the related software, see "Software Information" for ETERNUS Multipath Driver.

Note

• •

For immediate cluster failover if an I/O device where the system volume is placed fails

In the default setting of the ext3 file system, even if an I/O device where the system volume is placed fails, a cluster failover does not occur and the system operation may continue based on the data stored on the memory.

If you want PRIMECLUSTER to trigger a cluster failover immediately in the event that an I/O device where the system volume is placed fails, perform the following setting:

Setting

   For each partition included in the system volume, specify "errors=panic" to the mount option of ext3.

   Example: To set it in /etc/fstab (when /, /var, and /home exist in one system volume)

```
LABEL=/     /     ext3 errors=panic 1 1
LABEL=/boot /boot ext3 errors=panic 1 2
LABEL=/var  /var  ext3 errors=panic 1 3
LABEL=/home /home ext3 errors=panic 1 4
```

However, an immediate cluster failover may not become available due to taking time for an I/O error to reach theext3 file system. The regularly writing to the system volume enhances the detection frequency of I/O error.

• •

Setting up virtual SCSI devices (Xen environment)

   For using a shared disk on a guest OS, you need to set up a virtual SCSI device.

Setting up virtual SCSI devices (KVM environment)

   The following shows the setup procedure for virtual SCSI devices in a KVM environment:

   1. Stop the guest OS.

   2. Select the stopped guest OS with the Virtual Machine Manager and click the [Open] button in the toolbar.

   3. Click in the toolbar to display the detailed information of hardware.

4. Select a virtual disk (VirtIO Disk) from the hardware list in the left.

5. In the [Virtual disk] window, perform the following settings and click [Apply].

   - Select the Shareable check box.

   - Select 'none' for the cache model.

6. Check the version of the libvirt package on the host OS by using the rpm(8) command.

```
# rpm -qi libvirt
```

7. If the version of the libvirt package is libvirt-0.9.4-23.el6_2.4 or later, change the device attribute from disk to lun, which is set in the guest setting file (/etc/libvirt/qemu/*guestname*.xml) on the host OS.

```
# virsh edit guestname
```

Example before change

```
    :
  <disk type='block' device='disk'>
    <driver name='qemu' type='raw'/>
    <source dev='/dev/disk/by-id/scsi-1FUJITSU_30000085002B'/>
    <target dev='vdb' bus='virtio'/>
    <shareable/>
    <address type='pci' domain='0x0000' bus='0x00' slot='0x06' function='0x0'/>
  </disk>
    :
```

Example after change

```
    :
  <disk type='block' device='lun'>
    <driver name='qemu' type='raw'/>
    <source dev='/dev/disk/by-id/scsi-1FUJITSU_30000085002B'/>
    <target dev='vdb' bus='virtio'/>
    <shareable/>
    <address type='pci' domain='0x0000' bus='0x00' slot='0x06' function='0x0'/>
  </disk>
```

8. Start the guest OS.

Note

- The "by-id" name should be specified when creating the virtual disk in the guest OS.

- The virtual disk cannot be partitioned when creating it in the guest OS.

2. Setting up the virtual bridge (administrative LAN/public LAN/cluster interconnect)

   a. Xen environment

   For the network interfaces, including the administrative LAN, public LAN and cluster interconnect, that are used by virtual domains, you need to set up virtual bridges for the virtual networks beforehand.

   The following virtual bridge configurations are used in cluster systems.

   - gextbr:

     Virtual bridge for the public LAN/cluster interconnect

   - xenbr:

     Virtual bridge for the administrative LAN

   Example: Define the virtual bridges (gextbr4 and gextbr5) for the cluster interconnect in the script file (network-bridge-pcl) of the previously set up virtual bridge.

gextbr4,gextbr5:

    Virtual bridge for the cluster interconnect

xenbr0,xenbr1:

    Virtual bridge for the administrative LAN

gextbr2,gextbr3:

    Virtual bridge for the public LAN

network-bridge-pcl:

    The file name of the script to call the virtual bridge creation script

(1) Check that the script (/etc/xen/scripts/gext-network-bridge) exists. If not, create it.

See

. . . .

For details, see the manual of "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function".

. . . .

(2) Edit the script that defines the virtual bridge (network-bridge-pcl) which is located in "/etc/xen/scripts" on the host OS.

```sh
#!/bin/sh
#
# Sample of Create/Delete virtual bridge
#
# $1 start : Create virtual bridge
#    stop  : Delete virtual bridge
#    status: Display virtual bridge information

# Exit if anything goes wrong
set -e

command=$1

glsxenscript=/opt/FJSVhanet/local/sbin/hanetxen
xenscript=/etc/xen/scripts/network-bridge
xenscriptgext=/etc/xen/scripts/gext-network-bridge

# op_start:subscript for start operation #
op_start () {
        $xenscript $command vifnum=0 netdev=eth0
        $xenscript $command vifnum=1 netdev=eth1
        $xenscriptgext $command extnum=2 netdev=eth2
        $xenscriptgext $command extnum=3 netdev=eth3
        $xenscriptgext $command extnum=4 netdev=eth4.10 ***added
        $xenscriptgext $command extnum=5 netdev=eth5.20 ***added
}

# op_stop:subscript for stop operation #
op_stop () {
        op_start $command
}

case "$command" in
        start)
                # Create your virtual bridge
                $glsxenscript stop
                op_start
                $glsxenscript start
        ;;
```

```
        stop)
                # Delete virtual bridge
                $glsxenscript stop
                op_stop
        ;;

        status)
                # display virtual bridge information
                $xenscript status
        ;;

        *)
                echo "Unknown command: $command" >&2
                echo 'Valid commands are: start, stop, status' >&2
                exit 1

esac
```

(3) Set execute permissions

Execute the command below and set the execute permissions to the script (network-bridge-pcl).

```
# cp  network-bridge-pcl /etc/xen/scripts
# cd /etc/xen/scripts
# chmod +x network-bridge-pcl
```

(4) Register with the xend service

Check the "network-script" parameter of the xend service configuration file (/etc/xen/xend-config.sxp). Set the "network-bridge-pcl" if it is not set.

```
# Your default ethernet device is used as the outgoing interface, by default.
# To use a different one (e.g. eth1) use
#
# (network-script 'network-bridge netdev=eth1')
#
# The bridge is named xenbr0, by default.  To rename the bridge, use
#
# (network-script 'network-bridge bridge=<name>')
#
# It is possible to use the network-bridge script in more complicated
# scenarios, such as having two outgoing interfaces, with two bridges, and
# two fake interfaces per guest domain.  To do things like this, write
# yourself a wrapper script, and call network-bridge from it, as appropriate.
#
(network-script network-bridge-pcl)
```

b. KVM environment

For the network interfaces, including the administrative LAN, public LAN and cluster interconnect, that are used by virtual domains, you need to set up virtual bridges for the virtual networks beforehand.

(1) Setting up a virtual bridge for the administrative LAN

Edit the /etc/sysconfig/network-scripts/ifcfg-ethX file as follows:

```
DEVICE=ethX
HWADDR=XX:XX:XX:XX:XX:XX
BOOTPROTO=none
ONBOOT=yes
BRIDGE=brX
```

Create the interface setting file, /etc/sysconfig/network-scripts/ifcfg-brX, for the virtual bridge.

```
DEVICE=brX
TYPE=Bridge
BOOTPROTO=static
IPADDR=xxx.xxx.xxx.xxx
NETMASK=xxx.xxx.xxx.xxx
ONBOOT=yes
```

(2) Setting up virtual bridges for the public LAN and cluster interconnect

Edit the /etc/sysconfig/network-scripts/ifcfg-ethX file as follows:

```
DEVICE=ethX
HWADDR=XX:XX:XX:XX:XX:XX
BOOTPROTO=none
ONBOOT=yes
BRIDGE=brX
```

Create the interface setting file, /etc/sysconfig/network-scripts/ifcfg-brX, for the virtual bridge.

```
DEVICE=brX
TYPE=Bridge
ONBOOT=yes
```

3. Setting the guest OS in the host OS (in a KVM environment)

In a KVM environment, perform the following settings to stop the guest OS normally if the host OS is shut down by mistake while the guest OS running.

Define the following two values in /etc/sysconfig/libvirt-guests. When values are already defined, change them to the following values:

- ON_SHUTDOWN=shutdown

- SHUTDOWN_TIMEOUT=300

Specify the timeout duration (seconds) for shutdown of the guest OS to SHUTDOWN_TIMEOUT. Estimate the length of time for shutting down the guest OS and set the value. When multiple guest OSes are set, set the time whichever is greater. The above is an example when the time is 300 seconds (5 minutes).

4. Creating a user ID

Point

This user ID will be the one used by the shutdown facility to log in to the host OS to force shut down the nodes. This user ID and password are used for configuring the shutdown facility.

a. Xen environment

Create the required general user ID (FJSVvmSP) on the host OS for the guest OS control by PRIMECLUSTER.

```
# useradd FJSVvmSP
```

b. KVM environment

In a KVM environment, you need to set up a user for the shutdown facility for the guest OS control by PRIMECLUSTER.

(1) Creating a general user ID (optional)

Create a general user ID (optional) for the shutdown facility in the hypervisor.

```
# useradd <User ID>
```

(2) Setting up the "sudo" command

You need to set up the "sudo" command so that the general user ID (optional) for the shutdown facility can execute the command as the root user.

Using the "visudo" command, set up the general user ID created in step (1) so that it can execute the command without entering the password.

```
# visudo
```

Example

```
<User ID>    ALL=(root) NOPASSWD: ALL
```

5. Initializing the guest OS

After completing the installation of a guest OS and the installation of the bundled software on the guest OS, initialize the guest OS.

Note

This setup should be performed for all guest OSes comprising the cluster system.

See

For details on settings, see "Red Hat Enterprise Linux 6 Virtualization Host Configuration and Guest Installation Guide".

3.2.2.2 Guest OS setup

On the guest OSes comprising the cluster system, you need to set up the network, including IP addresses of the public LAN and administrative LAN.

Perform this setup on all guest OSes of a cluster.

Note

Guest domain names set on installation of the guest OSes are used when setting up the Shutdown Facility.

To check guest domain names, see "Red Hat Enterprise Linux 6 Virtualization Host Configuration and Guest Installation Guide".

Information

The Web-Based Admin View automatically configures the settings so that the interface, whose IP address host name is equivalent to the node name created when the PRIMECLUSTER was installed, can be used as a transfer route between cluster nodes and cluster management servers or between cluster management servers and clients.

3.2.2.3 Kernel parameter check

You need to edit the values of the kernel parameters based on the environment.

Target node:

All nodes on 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.

Note

To enable modifications, you need to restart the system after installation.

3.2.3 When building a cluster system between guest OSes on multiple host OSes using Host OS failover function

This section describes how to install and set up related software when building a cluster system between guest OSes on multiple host OSes using Host OS failover function.

Figure 3.1 Flow for building a cluster system when using Host OS failover function
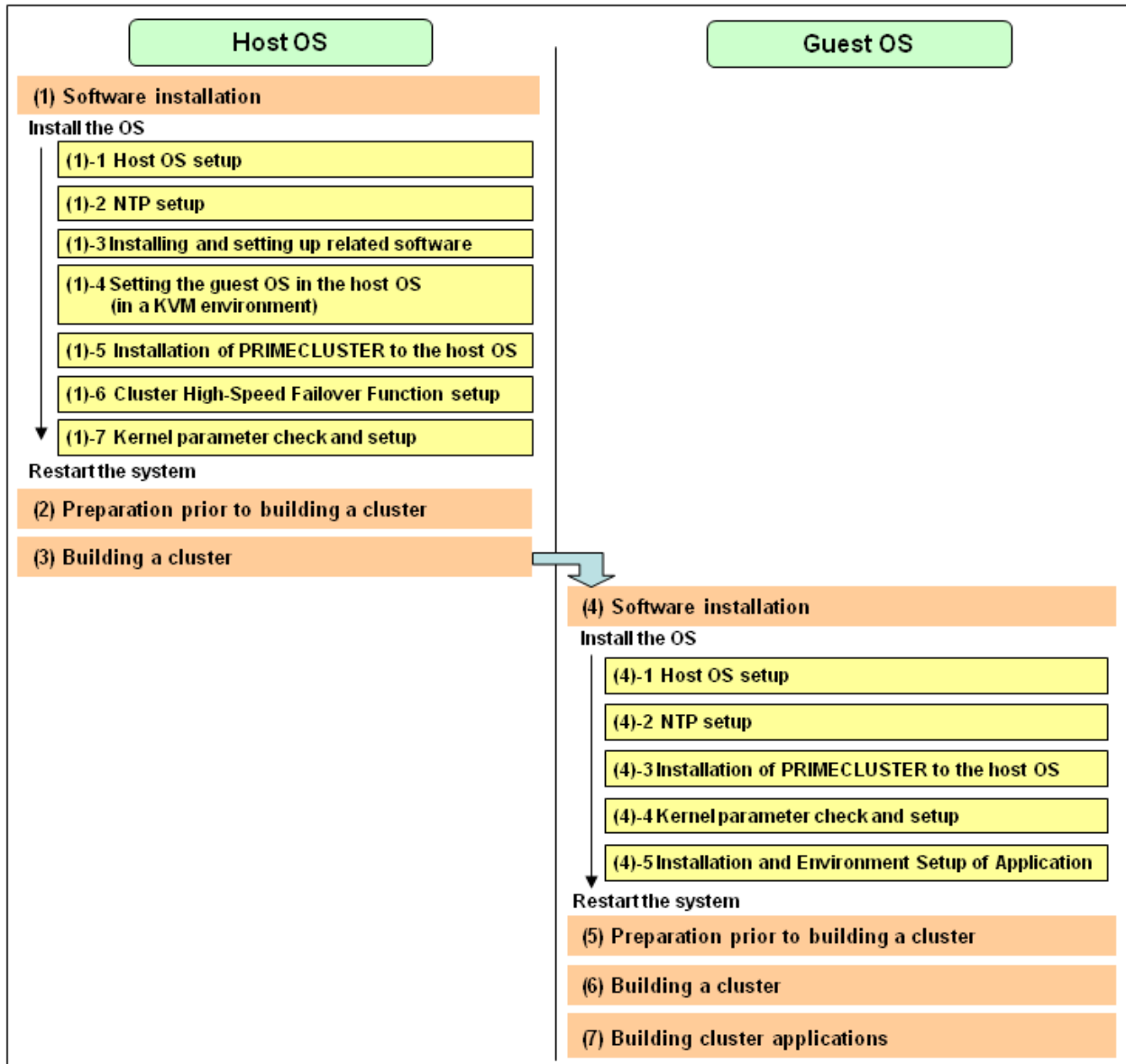
1. Software installation (Host OS)

   After installing the PRIMECLUSTER-related software, you need to set up the operating system, hardware, etc. that will be used and administered.

   For details, see "3.2.3.1 Installation and Setup of Related Software (Host OS)".

2. Preparation prior to building a cluster (Host OS)

   Before building a cluster, preparation work is required in the host OS, such as starting up the Web-Based Admin View screen. For details, see "Chapter 4 Preparation Prior to Building a Cluster".

3. Building a cluster (Host OS)

   Build a cluster of PRIMECLUSTER on the host OS. For details, see "Chapter 5 Building a Cluster." To build a cluster, perform the procedures described in "5.1.1 Setting Up CF and CIP" and "5.1.2 Setting Up the Shutdown Facility". Also, for the shutdown facility, set shutdown agent in the same way as the setting between natives. See "5.1.2 Setting Up the Shutdown Facility", and check the hardware model/configuration to set up the appropriate shutdown agent.

Note

- Set the timeout value of the cluster system on the host OS at the CF setting as follows:

  Timeout value: 20 seconds (Change the timout value with the cfset command.)

- Share the cluster interconnect LAN of the host OS with other guest OSes, and separate networks for each cluster system with Virtual LAN.

- Do not build a cluster application on the host OS in a Xen environment.

4. Software installation (Guest OS)

   After building a cluster on the host OS, install the PRIMECLUSTER-related software, and set up the OS and hardware for installing and operating PRIMECLUSTER.

   For details, see "3.2.2 When building a cluster system between guest OSes on multiple host OSes without using Host OS failover function."

Note

In a Xen environment, after a guest OS shuts down due to PANIC (crash), the guest OS is automatically rebooted in the standard settings. In the virtual machine environment, however, after automatic collection of the guest OS memory dump, it is necessary to collect the host OS system information by fjsnap before the guest OS is rebooted. Therefore, modify the settings in the domain configuration file, and make sure not to reboot after a panic occurred.

In a KVM environment, the guest OS is not rebooted after a panic occurred by default. You do not need to make any settings.

5. Preparation prior to building a cluster (Guest OS)

   Before building a cluster, preparation work is required in the host OS, such as starting up the Web-Based Admin View screen. For details, see "Chapter 4 Preparation Prior to Building a Cluster".

6. Building a cluster (Guest OS)

   Build a cluster of PRIMECLUSTER on the guest OS. For details, see "3.2.3.2 Building a Cluster (Guest OS)".

7. Building cluster applications (Guest OS)

   Create cluster applications on the guest OS. For details, see "Chapter 6 Building Cluster Applications".

Note

When creating a cluster application for a guest OS, do not set the ShutdownPriority attribute of RMS.

3.2.3.1 Installation and Setup of Related Software (Host OS)

After installing the PRIMECLUSTER-related software, you need to set up the operating system, hardware, etc. that will be used and administered.

Perform the following as needed.

1. Host OS setup

   In order for the host OS to work as the cluster, set up the network setting.

2. NTP setup

   This setup is for synchronizing the time on each node comprising the cluster system, which is necessary when creating a cluster.

   This setup should be performed before installing PRIMECLUSTER.

3. Installing and setting up related software

   Install and set up the software product (ETERNUS Multipath Driver) required for using system disk of the guest OS on the host OS. For how to install and set up the related software, see "Software Information" for ETERNUS Multipath Driver.

📝 Note

For immediate cluster failover if an I/O device where the system volume is placed fails

In the default setting of the ext3 file system, even if an I/O device where the system volume is placed fails, a cluster failover does not occur and the system operation may continue based on the data stored on the memory.

If you want PRIMECLUSTER to trigger a cluster failover immediately in the event that an I/O device where the system volume is placed fails, perform the following setting:

Setting

For each partition included in the system volume, specify "errors=panic" to the mount option of ext3.

Example: To set it in /etc/fstab (when /, /var, and /home exist in one system volume)

```
LABEL=/     /     ext3 errors=panic 1 1
LABEL=/boot /boot ext3 errors=panic 1 2
LABEL=/var  /var  ext3 errors=panic 1 3
LABEL=/home /home ext3 errors=panic 1 4
```

However, an immediate cluster failover may not become available due to taking time for an I/O error to reach theext3 file system. The regularly writing to the system volume enhances the detection frequency of I/O error.

4. Setting the guest OS in the host OS (in a KVM environment)

In a KVM environment, perform the following settings to stop the guest OS normally if the host OS is shut down by mistake while the guest OS running.

Define the following two values in /etc/sysconfig/libvirt-guests. When values are already defined, change them to the following values:

- ON_SHUTDOWN=shutdown

- SHUTDOWN_TIMEOUT=300

Specify the timeout duration (seconds) for shutdown of the guest OS to SHUTDOWN_TIMEOUT. Estimate the length of time for shutting down the guest OS and set the value. When multiple guest OSes are set, set the time whichever is greater. The above is an example when the time is 300 seconds (5 minutes).

📝 Note

- When setting /etc/sysconfig/libvirt-guests, do not describe the setting values and comments on the same line.

- When changing the settings in /etc/sysconfig/libvirt-guests during operation, make sure to follow the procedure in "8.9.2.3 Changing the Settings in /etc/sysconfig/libvirt-guests."

5. Installation of PRIMECLUSTER to the host OS

Install PRIMECLUSTER to the host OS. For details on the installation procedures, see the Installation Guide for PRIMECLUSTER.

6. Cluster High-Speed Failover Function setup

Make various settings for the Cluster High-Speed Failover Function related software and hardware.

Set up the setting after installing OS and PRIMECLUSTER.

For details, see "3.1.4 Setting Up the Cluster High-Speed Failover Function".

7. Kernel parameter check and setup

To operate the PRIMECLUSTER-related software, you need to edit the values of the kernel parameters based on the environment.

Perform this setup before rebooting the installed PRIMECLUSTER.

For details, see "3.1.5 Checking the Kernel Parameters".

3.2.3.2 Building a Cluster (Guest OS)

Build a cluster on the guest OS. For details on each item, see " Chapter 5 Building a Cluster".

> 🛑 **Note**
>
> - Share the cluster interconnect LAN of the guest OS with other guest OSes and the host OS, and separate networks for each cluster system with Virtual LAN.
>
> - Do not change a timeout value of the guest OS from 10 seconds at the CF setting.
>
> - For setup policy for survival priority, see "Survival scenarios" in "5.1.2 Setting Up the Shutdown Facility."

3.3 PRIMECLUSTER Installation

You can install PRIMECLUSTER using an installation script.

The installation script is also called the CLI Installer.

It is used to install PRIMECLUSTER on a system in which Linux(R) software and related Fujitsu software have been installed. It is also used for installation of the cluster management server.

> 📖 **See**
>
> For details on the installation procedures, see the Installation Guide for PRIMECLUSTER.

After PRIMECLUSTER was installed, perform the following settings for RHEL6 environments so that the CF modules are not incorporated to an initial RAM disk (initramfs) for kdump:

1. Add CF modules (cf, symsrv) to the setting of blacklist for /etc/kdump.conf.

   > 📝 **Example**
   >
   > - If the setting of blacklist has been described to /etc/kdump.conf
   >
   >   Before change:
   >
   >   ```
   >   blacklist kvm-intel
   >   ```
   >
   >   After change:
   >
   >   ```
   >   blacklist kvm-intel cf symsrv
   >   ```
   >
   > - If the setting of blacklist has not been described to /etc/kdump.conf
   >
   >   Add the following line.
   >
   >   ```
   >   blacklist cf symsrv
   >   ```

2. Restart the system or the service of kdump.

   > 📖 **See**
   >
   > For details on kdump, see the Linux documentation.

3.4 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 10-1 PRIMECLUSTER Product List".

- For details on installing applications, see the manuals, Software Release Guides 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 "Operating Environment" in the Installation Guide for PRIMECLUSTER.
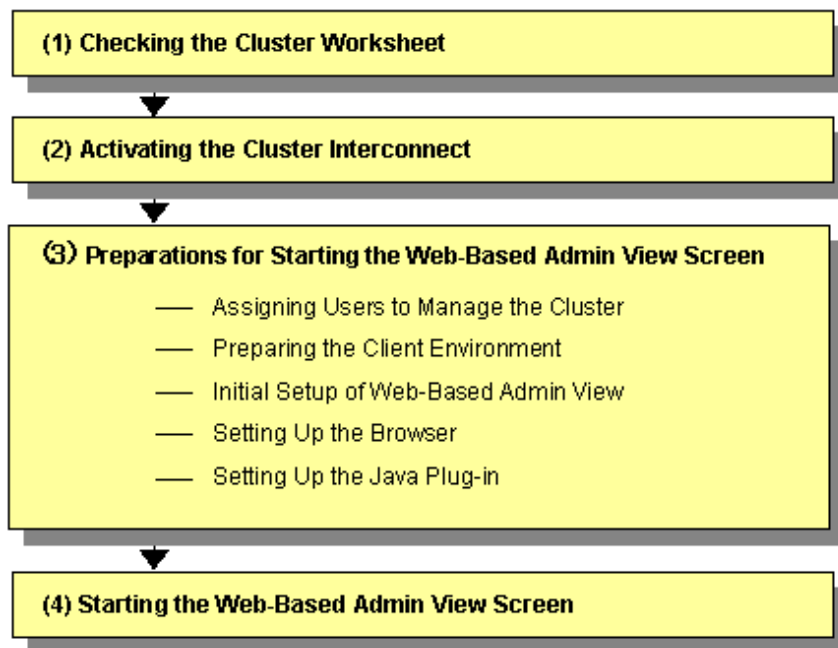
Table 4.1 Operation procedure and manual reference location for starting the Web-Based Admin View screen

	Work item	Execution Node	Required/ Optional	Manual reference location*
(1)	4.1 Checking the Cluster Worksheet	-	Required	This manual: Appendix A
(2)	4.2 Activating the Cluster Interconnect	All nodes	Required	
(3)	4.3 Preparations for Starting the Web-Based Admin View Screen			
	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:

- **WEB:** PRIMECLUSTER Web-Based Admin View Operation Guide

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 Activating the Cluster Interconnect

As preparation for setting up CF, execute the following operation:

1. Edit the /etc/sysconfig/network-scripts/ifcfg-ethX file.

   Edit "ONBOOT" of the /etc/sysconfig/network-scripts/ifcfg-ethX file as follows:

   ```
   ONBOOT=yes
   ```

 Note
. . . .
ethX indicates a network interface that is used for the cluster interconnect.

A number is specified in X.
. . . .

2. Confirm the cluster interconnect.

   Confirm the state of the interface for interconnect with the following command.

   ```
   # ifconfig <relevant interface> <Return>
   ```

3. If the beginning of the second row output with the above command is not "UP" execute the following command to confirm if "UP" is output.

   ```
   # ifconfig <relevant interface> up <Return>
   ```

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	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.
	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 allow 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 Linux(R) 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 register the above group to a user, you should register the group as a Supplemental Group. To register a group as a Supplemental Group, use the usermod(8) or useradd(8) command.

- To add a user group to a registered user

```
# usermod -G wvroot username
```

- To register a new user

```
# useradd -G wvroot username
```

📖 Note

. . . .

When you register a new user, use the passwd(8) command to set a password.

```
# passwd username
```

. . . .

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.

See

For details, see "3.1.2 Preparing the Client Environment" in the "*PRIMECLUSTER Web-Based Admin View Operation Guide.*"

Note

If Microsoft(R) Windows is used for a client that is supported with Web-Based Admin View, the following color pallet problem may occur:

- When the GUI screen of Web-Based Admin View is displayed with Microsoft(R) Windows, 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.

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.

```
# /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>
```

   - In the case of multiple-node cluster operation
     Execute the command above on all the management servers and nodes, referring to the example.
     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 10.20.30.40
# /etc/opt/FJSVwvbs/etc/bin/wvSetparam secondary-server 10.20.30.41
```

   - In the case of the single-node cluster operation
     For each IP address or host name of the primary management server and secondary management server, specify the node's own IP address or host name.
     Example:

```
# /etc/opt/FJSVwvbs/etc/bin/wvSetparam primary-server 10.20.30.40
# /etc/opt/FJSVwvbs/etc/bin/wvSetparam secondary-server 10.20.30.40
```

   In addition, no value is displayed in Web-Based Admin View on the secondary management server.

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

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

 Note
....

For making entries to /etc/hosts in Japanese, use EUC encoding and set "ja" for the system requirements variable "lang" for Web-Based Admin View. For further details on the Web-Based Admin View language settings, refer to "4.3.3.3 Setting the Web-Based Admin View Language."
....

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" 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  3m41s
primaryServer   Sessions: 0
primaryServer   Nodes: 2
        10.20.30.40 node1      Linux-2.4.9-e.8enterprise      3m36s
        10.20.30.41 node2      Linux-2.4.9-e.8enterprise      2m58s
secondaryServer 10.20.30.41 node2  http=10.20.30.41      Run  2m46s
secondaryServer Sessions: 0
secondaryServer Nodes: 2
        10.20.30.40 node1      Linux-2.4.9-e.8enterprise      2m41s
        10.20.30.41 node2      Linux-2.4.9-e.8enterprise      2m23s
#
```

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" 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.

Attribute	Variable	Possible values	Meaning
			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> <setting_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 Linux(R) 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 "3.1.2 Prerequisite client environment" in the "*PRIMECLUSTER Web-Based Admin View Operation Guide.*" 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.*"

- When a URL is to be registered for the bookmark of a browser, register "Plugin.cgi."

. . . .

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*.
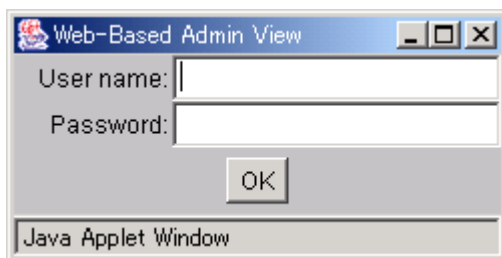
Note

. . . .

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 "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 cluster 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 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
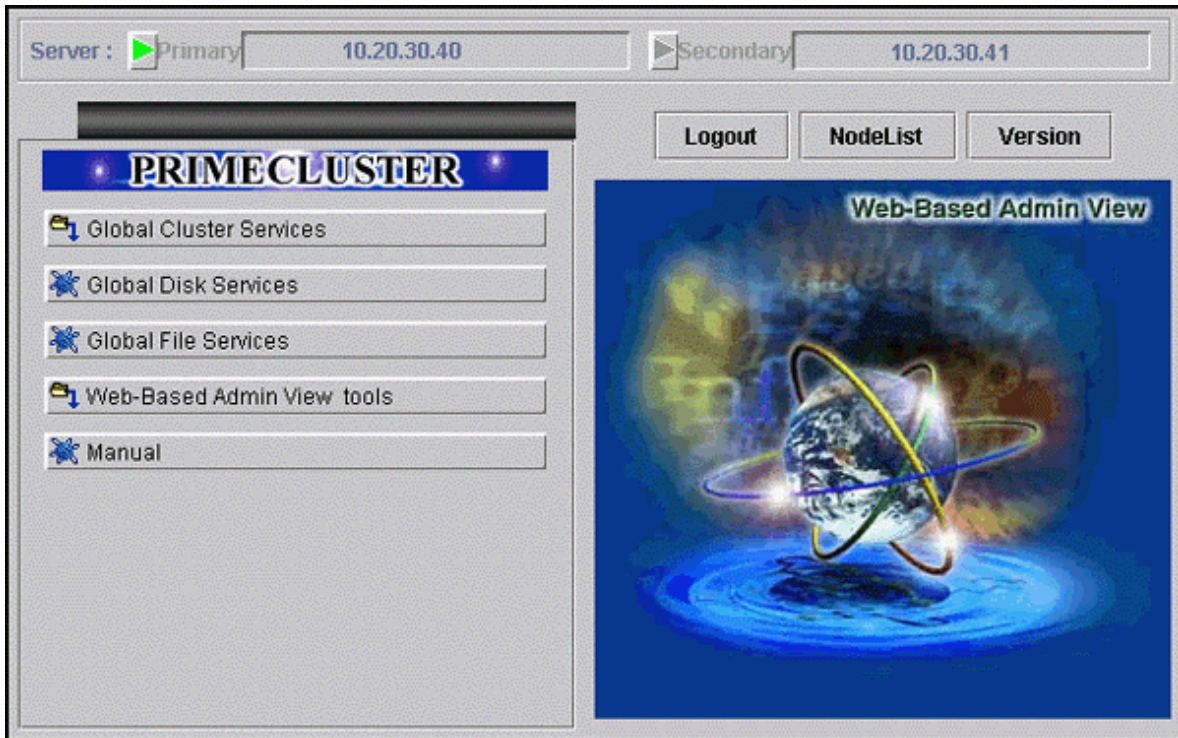
Menu Outline

The operation menus are divided into the following two categories:

1. **PRIMECLUSTER category**

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

2. **MISC category**

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

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)

  - **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. For example, in an environment in which Symfoware is installed, this menu category will be displayed.

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.

  - Symfoware

- **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.
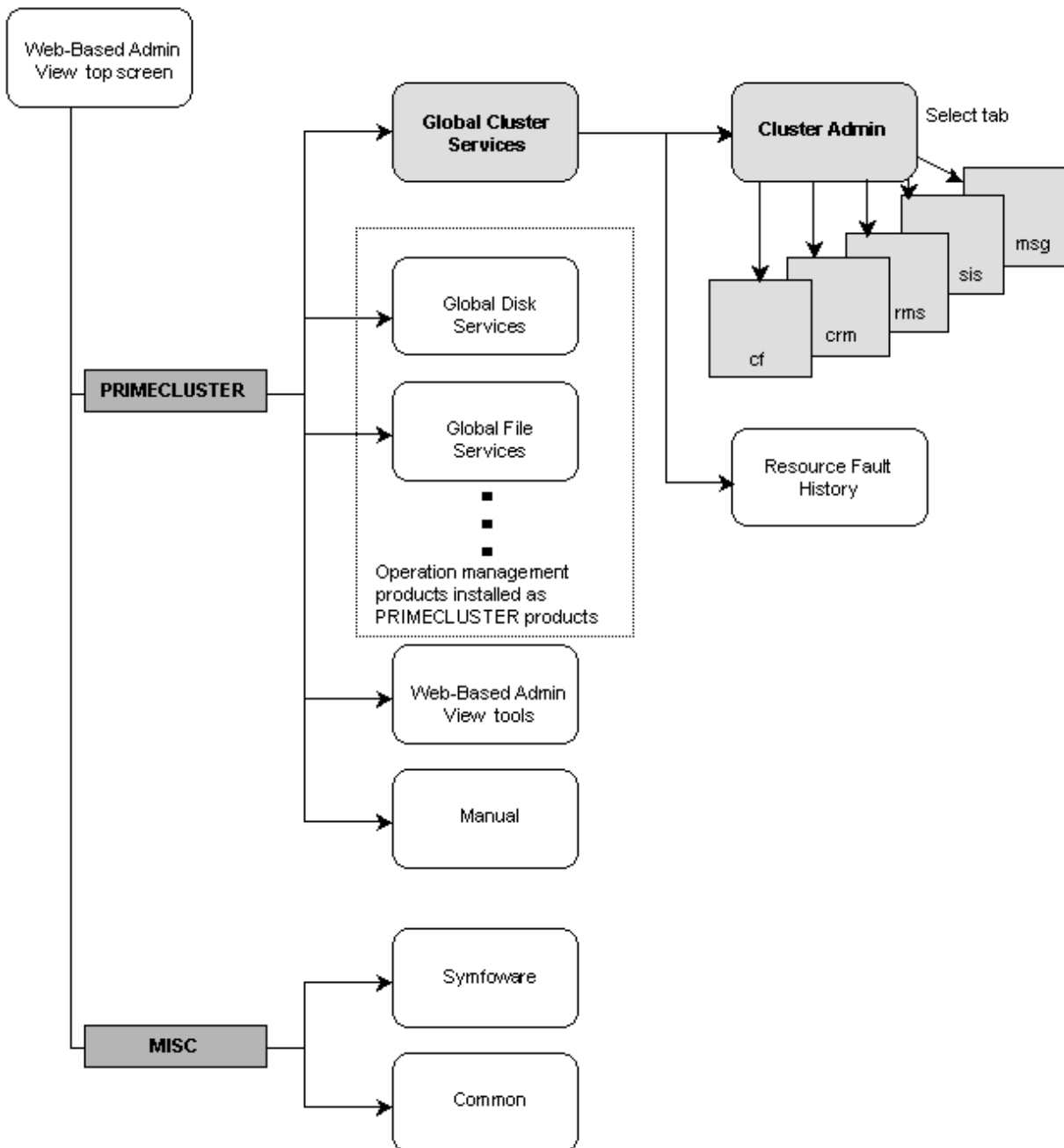
At the Cluster Admin screen, you can switch the window by clicking the following tabs:

- 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.

![Note icon] **Note**

. . . .

SIS cannot be used with PRIMECLUSTER 4.3.

. . . .

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)
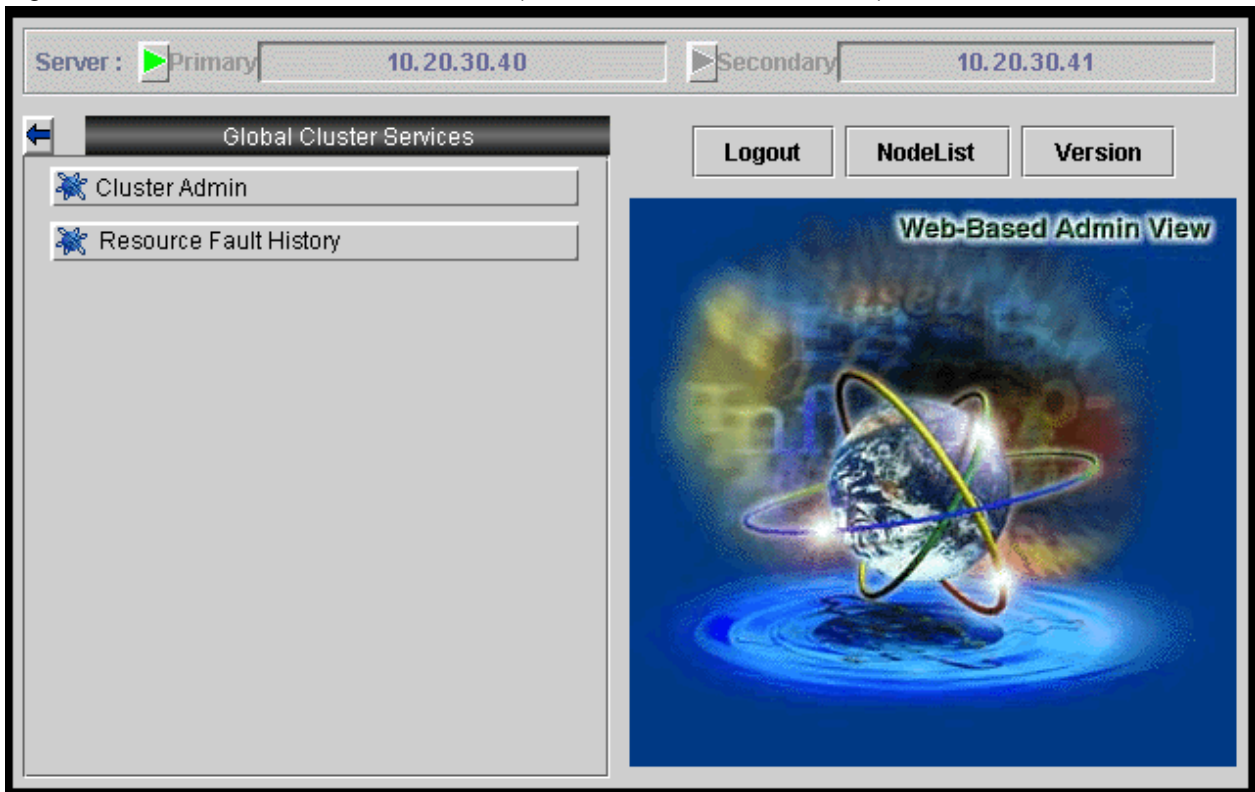
Overview of the Global Cluster Services menu

- **Cluster Admin**

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

- **Resource Fault History**

  This function allows you to display the resource fault history. For details, see "C.2.2 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

Figure 4.3 Web-Based Admin View screen (Cluster Admin)
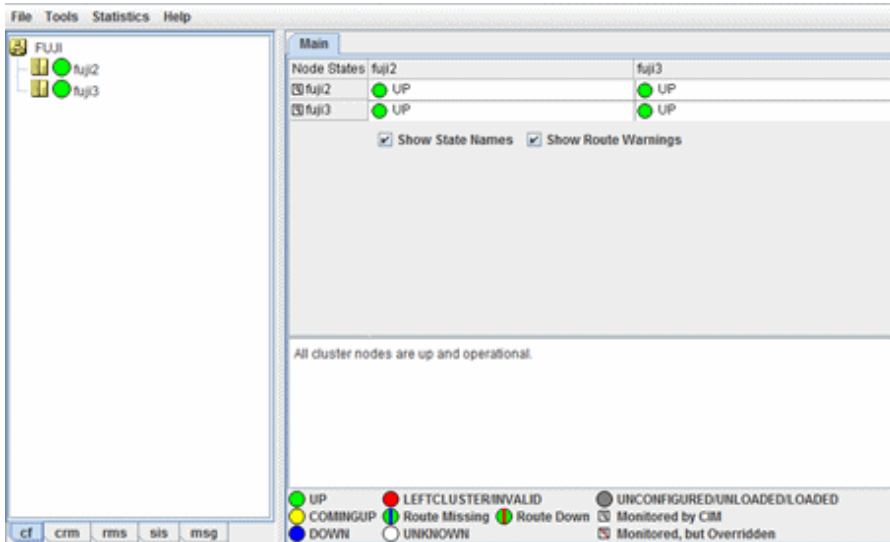
Cluster Admin supports the functions described below.

The manual reference locations are indicated in "Overview of 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 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)**

  PRIMECLUSETR 4.3, this function is not available.

- **msg (Message)**

  Cluster control messages are displayed.

  Reference location: "Chapter 7 Operations"

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.

- 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 * Only in the multiple-node operation.
- Initial Setup of the Cluster Resource Management Facility

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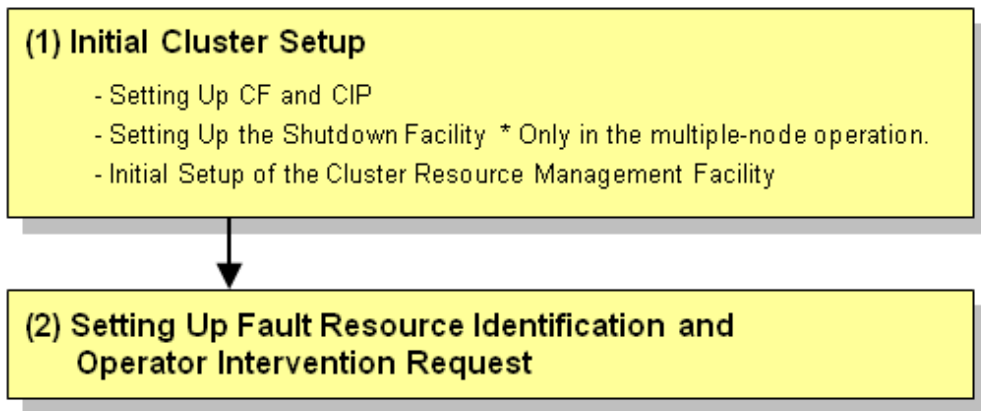
Table 5.1 Cluster building procedure and manual reference locations

	Work item	Execution Node	Required/ Optional	Manual reference location*1
(1)	5.1.1 Setting Up CF and CIP	All nodes	Required	**CF**
	5.1.2 Setting Up the Shutdown Facility	All nodes	Required**2**	**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 Fault Resource Identification and Operator Intervention Request	1 node	Required	

*1 The PRIMECLUSTER manual names are abbreviated as follows:

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

*2 When configuring a single node cluster, it is not necessary to setup the shutdown facility.

Note
....
Execute the configuration setting of GDS after initializing the cluster.
....

5.1 Initial Cluster Setup

This section describes the initial cluster setup for PRIMECLUSTER.

If the virtual machine function is used, you need to set up a virtual network first.

For details, refer to "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function".

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

	Contents	Manual reference location*
1	5.1.1 Setting Up CF and CIP (setting up cluster configuration information and IP addresses)	**CF** "2.1 CF, CIP, and CIM configuration"
2	5.1.2 Setting Up the Shutdown Facility	**CF** "8 Shutdown Facility"

	Contents	Manual reference location*
3	5.1.3 Initial Setup of the Cluster Resource Management Facility	**CF** "4.3 Resource Database configuration"

* The PRIMECLUSTER manual names are abbreviated as follows:

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

Note

. . . .

When Firewall is enabled, disable it before the initial cluster setup.

When enabling Firewall after completing the installation of the cluster, see "Appendix J Using Firewall."

. . . .

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.7 Cluster Configuration Worksheet."

For details on the setting, see "2.1.5 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", "Node 2", "Node 3" and "Node 4"
CF node names	The beginning character of each CF node name must be a lower-case alphabet. The length of each node name must be not more than 11 characters which consist of lower-case alphabets, numbers, or symbols (- and _).	"CF node name" for "Node 1", "Node 2", "Node 3" and "Node 4"
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. However, IP addresses must not be assigned to the interfaces.	"Path 0 NIC name" and "Path 1 NIC name" for "Node 1", "Node 2", "Node 3" and "Node 4"
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)	"CF remote service usage"

Setup item	Description	Described in worksheet
	With the default settings, these services are "Not selected." To use RMS, you need to select at least one function.	
Cluster Integrity Monitor (CIM) configuration	Set the nodes to be monitored by CIM.	"Node in CF quorum set" for "Node 1", "Node 2", "Node 3" and "Node 4"

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

- Hereinafter, the CF remote services (CFCP and CFSH) must be enabled. To enable the CFCP and CFSH, add the following definition to the "/etc/default/cluster.config" file before configuring the CF.

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

- If the CF and CIP configuration fails, check the following:

    - The cluster interconnect is connected.

    - The network interface that is used for the cluster interconnect is activated.

    - An IP address is configured for the network interface.

- 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.

- After the CF setup is completed, the "SF Wizard Startup Check" pop-up window is displayed. Click the No button. Set up the SF according to the instructions in "5.1.2 Setting Up the Shutdown Facility."

Note

In the case of the single-node cluster operation

- When constructing multiple single-node clusters, specify a different name per each node, such as including the node' own name in the name.

- For the network interface specified to the cluster interconnect, select a dedicated network interface device listed in the above table.

- Select [No] to the message of "Do you wish to invoke the Shutdown Facility Wizard to configure this cluster?" output after configuring CF in ClusterAdmin since the setting of the shutdown facility is unnecessary.

- Following messages of the shutdown facility and RMS are output, however, this is no problem, since the setting of the shutdown facility is not performed.

    - Messages of the shutdown facility:

    ```
    fopen of /etc/opt/SMAW/SMAWsf/rcsd.cfg failed, errno 2
    Could not correctly read the rcsd.cfg file
    ```

    - Messages of RMS:

    ```
    (SCR,26): ERROR The sdtool notification script has failed with status 1 after dynamic
    modification.
    ```

5.1.2 Setting Up the Shutdown Facility

This section describes the setup procedure of the shutdown facility for the PRIMERGY, PRIMEQUEST, and virtual machine environment.

The setup procedure for the shutdown facility is different depending on the model/configuration. Check the hardware model/configuration to set up the appropriate shutdown agent.

The following shows the shutdown agents required for each hardware model/configuration.

Table 5.2 Shutdown agent for PRIMERGY

Server model	SA_lkcd	SA_ipmi	SA_blade
BX620/920/960 series	Y	Y (*1)(*2)	Y
RX200/300/500/600 series	Y	Y	-
TX200/300 series	Y	Y	-

*1) For use in combination with ServerView Resource Orchestrator Virtual Edition in the BX620/920/960 series, set SA_ipmi.

*2) Set the same BMC or iRMC user name and password on all blades.

Table 5.3 Shutdown agent for PRIMEQUEST

Server model		MMB	
		Panic	Reset
PRIMEQUEST	1000 series	SA_mmbp	SA_mmbr

Table 5.4 Shutdown agent for virtual machine environment (Xen environment)

Server model		Cluster configuration			
		Virtual machine function (Xen environment)			
		Guests in a unit		Guests in other units	
		vmSP			
		Panic	Reset	Panic	Reset
PRIMEQUEST	1000 series	SA_vmSPgp	SA_vmSPgr	SA_vmSPgp	SA_vmSPgr

Table 5.5 Shutdown agent for virtual machine environment (KVM environment)

Server model		Cluster configuration			
		Virtual machine function (KVM environment)			
		Guests in a unit		Guests in other units	
		libvirt			
		Panic	Reset	Panic	Reset
PRIMERGY		SA_libvirtgp	SA_libvirtgr	SA_libvirtgp	SA_libvirtgr
PRIMEQUEST	1000 series	SA_libvirtgp	SA_libvirtgr	SA_libvirtgp	SA_libvirtgr

Also, when using Host OS failover function, set the following shutdown agents. The SA_vmSPg, SA_vmSPgr, SA_libvirtgp, and SA_libvirtgr of the shutdown agent that are set on the guest OS are the same as those used in the virtual machine function. For details on SA_vmSPgp, SA_vmSPgr, SA_libvirtgp, and SA_libvirtgr, see "5.1.2.3.1 Setting up the shutdown daemon" and "5.1.2.5.2 libvirt". Set SA_vmchkhost according to the procedures described in "5.1.2.5.3 vmchkhost".

Table 5.6 Shutdown agent for Host OS failover function (Xen environment)

Server model		Cluster configuration				
		Virtual machine function (Xen environment)				
		Guests in other units (Host OS failover function)				
		MMB		vmSP		vmchkhost
		Panic	Reset	Panic	Reset	Checking the status
PRIMEQUEST 1000 series	Host OS	SA_ mmbp	SA_mmbr	-	-	-
	Guest OS	-	-	SA_vmSPgp	SA_vmSPgr	SA_vmchkhost

Table 5.7 Shutdown agent for Host OS failover function (KVM environment)

Server model		Cluster configuration			
		Virtual machine function (KVM environment)			
		Guests in other units (Host OS failover function)			
			libvirt		vmchkhost
		Depending on server model	Panic	Reset	Checking the status
PRIMERGY	Host OS	See Table 5.2.	-	-	-
	Guest OS	-	SA_libvirtgp	SA_libvirtgr	SA_vmchkhost
PRIMEQUEST 1000 series	Host OS	See Table 5.3	-	-	-
	Guest OS	-	SA_libvirtgp	SA_libvirtgr	SA_vmchkhost

🔔 Note

When creating a redundant administrative LAN used in the shutdown facility by using GLS, use the logical IP address takeover function of the NIC switching mode, and for the administrative LAN used in the shutdown facility, set a physical IP address.

📘 See

For details on the shutdown facility, see the following manuals:

- "3.3.1.7 PRIMECLUSTER SF" in the "*PRIMECLUSTER Concept Guide*"

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

5.1.2.1 Checking the Shutdown Agent Information

Check the information of the shutdown agent to be used.

🔔 Note

Check the shutdown agent information before cluster initialization.

MMB check items

If an MMB is being used, check the following settings:

- The "Privilege" setting of the user is set to "Admin" so that the user can control the MMB with RMCP.

- The "Status" setting of the user is set to "Enabled" so that the user can control the MMB with RMCP.

Check the settings for the user who uses RMCP to control the MMB. Log in to MMB Web-UI, and check the settings from the "Remote Server Management" window of the "Network Configuration" menu.

If the above settings have not been set, set up the MMB so that the above settings are set.

Jot down the following information related to the MMB:

- User's name for controlling the MMB with RMCP (*1)

- User's password for controlling the MMB with RMCP.

    *1) The user must be granted the Admin privilege.

Note

The MMB units have two types of users:

- User who controls all MMB units

- User who uses RMCP to control the MMB

The user to be checked here is the user who uses RMCP to control the MMB. Be sure to check the correct type of user.

Virtual machine check items (Xen environment)

When setting up vmSP (Virtual Machine Service Provider) to the shutdown facility in a Xen environment, log in to the host OS using SSH in order to force stop the guest OS. To do this, you need to set up the following information:

- Host OS IP address

- User name for logging in to the host OS (FJSVvmSP)

- User password for logging in to the host OS

For the information about User name and password for logging in to the host OS, record it you have set up in "3.2.1.1 Host OS setup."

Virtual machine check items (KVM environment)

When setting up the shutdown facility in a KVM environment, log in to the hypervisor using SSH in order to force stop the guest OS. To do this, you need to set up the following information.

- IP address for the hypervisor

- User for logging in to the hypervisor (*2)

- User password for logging in to the hypervisor

    *2) In order to execute the command as a root user, you must specify the user who sets the "sudo" command.

For the information about User name and password for logging in to the hypervisor, record it you have set up in "3.2.1.1 Host OS setup."

5.1.2.2 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 "8.5 Changing the Operation Attributes of a userApplication."

See

See

. . . .

For details on the ShutdownPriority attribute of userApplication, see "11.1 Attributes available to the user" in the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide".

. . . .

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).
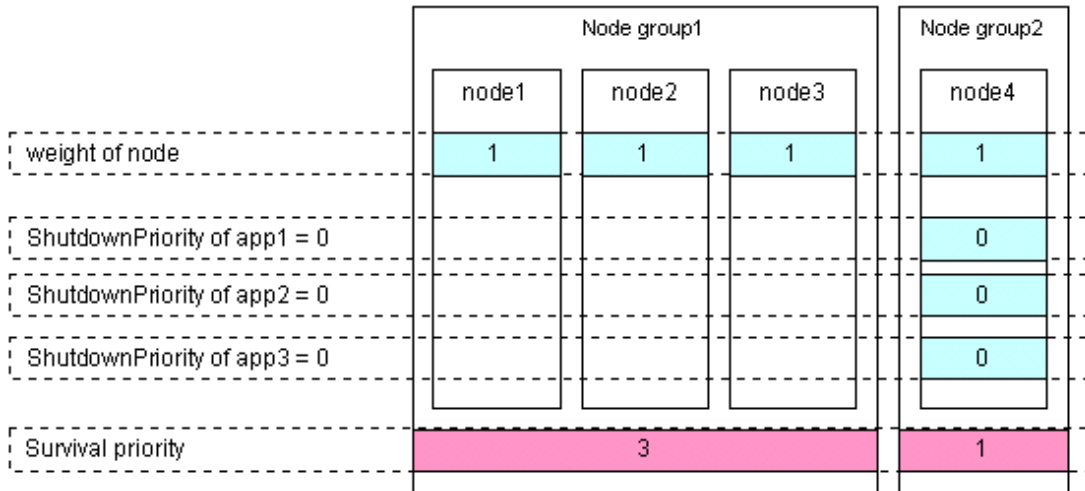
[Specific node survival]

- 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:

	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
Survival 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
Survival priority	21	5		

[Node survival in a specific order]

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

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

In the following example, node1, node2, node3, and node4 are to survive in this order:

	Node group1	Node group2		
	node1	node2	node3	node4
weight of node	18	6	2	1
ShutdownPriority of app1 = 0		0		
ShutdownPriority of app2 = 0			0	
ShutdownPriority of app3 = 0				0
Survival priority	18	9		

[Host OS failover function]

- Set the "weight" of nodes to a power-of-two value (1,2,4,8,16,...) in ascending order of survival priority in each cluster system.

- The "weight" set to a guest OS should have the same order relation with a corresponding host OS.

  For example, when setting a higher survival priority to host1 than host2 between host OSes, set a higher survival priority to node1 (corresponding to host1) than node2-4 (corresponding to host2) between guest OSes.

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

In the following example, node1, node2, node3, and node4 are to survive in this order:

	Unit0	Unit1		
	Node group1	Node group2		
	Virtual machine (guest)	Virtual machine (guest)	Virtual machine (guest)	Virtual machine (guest)
	node1 Guest OS	node2 Guest OS	node3 Guest OS	node4 Guest OS
weight of node	8	4	2	1
ShutdownPriority of app1 = 0		0		
ShutdownPriority of app2 = 0			0	
ShutdownPriority of app3 = 0				0
Survival priority	8	7		

	Node group1	Node group2
	Virtual machine (domain0)	Virtual machine (domain0)
	host1 Host OS	host2 Host OS
weight of node	2	1
Survival priority	2	1

5.1.2.3 Setting up the shutdown agent in PRIMERGY

This section describes the procedure for setting up the shutdown agent in PRIMERGY. To use this in a virtual machine environment, see "5.1.2.5 Setting up the shutdown agent in virtual machine environment."

Note

....

After setting up the shutdown agent, conduct a test for forced shutdown of cluster nodes to check that the correct nodes can be forcibly stopped. For details of the test for forced shutdown of cluster nodes, see "1.4 Test".

....

5.1.2.3.1 Setting up the shutdown daemon

Create /etc/opt/SMAW/SMAWsf/rcsd.cfg on all nodes as shown below:

```
CFNameX,weight=weight,admIP=myadmIP. agent=SA_xxx,timeout=20
CFNameX,weight=weight,admIP=myadmIP. agent=SA_xxx,timeout=20
```

```
CFNameX          : CF node name of the cluster host.
weight           : Weight of the SF node.
```

```
myadmIP          : IP address of the administrative LAN on the local node.
agent=SA_XXX     : Specify the name of the Shutdown Agent.
                     For the IPMI Shutdown Agent, specify SA_ipmi.
                     For the Blade Shutdown Agent, specify SA_blade.
```

Example) IPMI Shutdown Agent

```
node1, weight=1, admIP=10.20.30.100: agent=SA_ipmi, timeout=25
node2, weight=1, admIP=10.20.30.101: agent=SA_ipmi, timeout=25
```

Example) Blade Shutdown Agent

```
node1,weight=1,admIP=10.20.30.100: agent=SA_blade,timeout=20
node2,weight=1,admIP=10.20.30.101: agent=SA_blade,timeout=20
```

Note

- For IPMI shutdown agent, set timeout to 25.

- For using STP (Spanning Tree Protocol) in PRIMERGY, it is necessary to set the SF timeout value to the current value plus (+) 50 (seconds), taking into account the time STP needs to create the tree and an extra cushion. This setting also causes delays in failover times.

Information

When the "/etc/opt/SMAW/SMAWsf/rcsd.cfg" file is to be created, the "/etc/opt/SMAW/SMAWsf/rcsd.cfg.template" file can be used as a prototype.

5.1.2.3.2 IPMI

For the server with the BMC (Baseboard Management Controller) or iRMC (integrated Remote Management Controller) installed, configure the IPMI shutdown agent. You must configure the IPMI shutdown agent before you configure the kdump shutdown agent.

Create /etc/opt/SMAW/SMAWsf/SA_ipmi.cfg on all nodes as shown below:

```
CFName1 ip-address:user:passwd {cycle | leave-off}
CFName2 ip-address:user:passwd {cycle | leave-off}
```

```
CFNameX          : CF node name of the cluster host.
ip-address       : Ip address for IPMI (BMC or iRMC).
user             : User name defined when IPMI (BMC or iRMC) was setup.
passwd           : Password defined when IPMI (BMC or iRMC) was setup.
                     Alphanumeric characters and the following symbols are available for passwords.
                       # % * + - . @ [ ] ^ _ { } ~
cycle            : Reboot the node after forcibly stopping the node.
leave-off        : Power-off the node after forcibly stopping the node.
```

Example)

When the IP address of iRMC of node1 is 10.20.30.50, the IP address of iRMC of node2 is 10.20.30.51.

```
node1 10.20.30.50:root:bmcpwd cycle
node2 10.20.30.51:root:bmcpwd cycle
```

Information

When the "/etc/opt/SMAW/SMAWsf/SA_ipmi.cfg" file is to be created, the "/etc/opt/SMAW/SMAWsf/SA_ipmi.cfg.template" file can be used as a prototype.

Note

- Check if the setting contents of the /etc/opt/SMAW/SMAWsf/SA_ipmi.cfg file are correct. If there is an error in the setting contents, the shutdown facility cannot be performed normally.

- Check if the IP address (ip-address) of IPMI (BMC or iRMC) corresponding to the cluster host's CF node name (CFNameX) of the /etc/opt/SMAW/SMAWsf/SA_ipmi.cfg file are set. If there is an error in the setting, a different node may be forcibly stopped.

- If you have used symbols in the password specified when making the IPMI (BMC or iRMC) settings, it is unnecessary to add backslashes ("\"), indicating escape characters in front of the symbols.

  Example: If the password specified when making the IPMI (BMC or iRMC) settings is "bmcpwd#"

  ```
  node1 10.20.30.50:root:bmcpwd# cycle
  node2 10.20.30.51:root:bmcpwd# cycle
  ```

- Change the permission of the /etc/opt/SMAW/SMAWsf/SA_ipmi.cfg file to 600 by executing the following command.

  ```
  # chmod 600 /etc/opt/SMAW/SMAWsf/SA_ipmi.cfg
  ```

5.1.2.3.3 Blade

For the Blade server, configure the Blade shutdown agent. You must configure the Blade shutdown agent before you configure the kdump shutdown agent.

Create /etc/opt/SMAW/SMAWsf/SA_blade.cfg on all nodes as shown below:

(1) Cluster configuration within a single chassis

```
management-blade-ip IPaddress
community-string SNMPcommunity
CFName1 slot-no {cycle | leave-off}
CFName2 slot-no {cycle | leave-off}
```

(2) Cluster configuration across multiple chassis

```
community-string SNMPcommunity
management-blade-ip IPaddress
CFName1 slot-no {cycle | leave-off}
management-blade-ip IPaddress
CFName2 slot-no {cycle | leave-off}
```

```
IPaddress          : IP address of the management blade.
SNMPcommunity      : SNMP community.
CFNameX            : CF node name of the cluster host.
slot-no            : Slot No. of the server blade.
cycle              : Reboot the node after forcibly stopping the node.
leave-off          : Power-off the node after forcibly stopping the node.
```

Example 1)

When the IP address of the management blade of node1 and node2 is 10.20.30.50, the slot number of node1 is 1 and the slot number of node2 is 2.

```
management-blade-ip 10.20.30.50
community-string public
node1 1 cycle
node2 2 cycle
```

Example 2)

When the IP address of the management blade of node1 is 10.20.30.50, and the slot number of node1 is 1.
Moreover, when the IP address of the management blade of node2 is 10.20.30.51, and the slot number of node2 is 2.

```
community-string public
management-blade-ip 10.20.30.50
node1 1 cycle
management-blade-ip 10.20.30.51
node2 2 cycle
```

Information

When the "/etc/opt/SMAW/SMAWsf/SA_blade.cfg" file is to be created, the "/etc/opt/SMAW/SMAWsf/SA_blade.cfg.template" file can be used as a prototype.

Note

- Check if the setting contents of the /etc/opt/SMAW/SMAWsf/SA_blade.cfg file are correct. If there is an error in the setting contents, the shutdown facility cannot be performed normally.

- Check if the IP address (IPaddress) of the management blade and the slot number (slot-no) of the server blade corresponding to the cluster host's CF node name (CFNameX) of the /etc/opt/SMAW/SMAWsf/SA_blade.cfg file are set. If there is an error in the setting, a different node may be forcibly stopped.

- Change the permission of the /etc/opt/SMAW/SMAWsf/SA_blade.cfg file to 600 by executing the following command.

```
# chmod 600 /etc/opt/SMAW/SMAWsf/SA_blade.cfg
```

Note

The rcsd.cfg, SA_ipmi.cfg and SA_blade.cfg files must be the same on all nodes. If not, operation errors might occur.

5.1.2.3.4 kdump

Set up the kdump shutdown agent when kdump is used to collect the crash dump.

- Initializing the configuration file for the kdump

Execute the following command on any one of the cluster nodes.

```
# /etc/opt/FJSVcllkcd/bin/panicinfo_setup
```

Note

To execute the command, CF and CF services (CFSH and CFCP) must be activated. For details, see "5.1.1 Setting Up CF and CIP."

- Setting crash dump collection

The procedures for setting up may differ depending on the hardware used for the node.

- PRIMERGY RX200/300/500/600 series, TX200/300 series, and BX620/920/960 series (used in combination with ServerView Resource Orchestrator Virtual Edition)

Change /etc/opt/FJSVcllkcd/etc/SA_lkcd.tout as follows on all nodes.

Before change

```
PANICINFO_TIMEOUT 5
RSB_PANIC 0
```

After change

```
PANICINFO_TIMEOUT 10
RSB_PANIC 3
```

Change the timeout value of SA_lkcd in the /etc/opt/SMAW/SMAWsf/rcsd.cfg file as follows on all nodes.

Before change

```
agent=SA_lkcd,timeout=20
```

After change

```
agent=SA_lkcd,timeout=25
```

- BLADE servers (BX620/920/960 series)

Change RSB_PANIC of /etc/opt/FJSVcllkcd/etc/SA_lkcd.tout as follows on all nodes.

Before change

```
RSB_PANIC 0
```

After change

```
RSB_PANIC 2
```

5.1.2.3.5 Starting up the shutdown facility

Start or restart the shutdown facility on all nodes.

- If the shutdown daemon (rcsd) has not yet been started

Start the shutdown daemon (rcsd) with sdtool -b.

```
# sdtool -b
```

- If the shutdown daemon (rcsd) is active

Stop the shutdown daemon (rcsd) with sdtool -e and then start it with sdtool -b.

```
# sdtool -e
# sdtool -b
```

Use sdtool -s to confirm whether the shutdown daemon (rcsd) is active.

```
# sdtool -s
```

By executing sdtool -s on all nodes, the composition of the shutdown facility can be confirmed.

Note

Confirm the shutdown facility operates normally by the display result of the sdtool -s command.

There is a possibility that the mistake is found in the configuration setting of the agent or hardware when displayed as follows though the setting of the shutdown facility is completed.

- "InitFailed" is displayed as the initial status.

- "Unknown" or "TestFailed" is displayed as the test status.

Confirm whether the error message is output to/var/log/messages file. Then, take corrective actions according to the content of the output message.

・・・

5.1.2.4 Setting up the shutdown agent in PRIMEQUEST

This section describes the procedure for setting up the shutdown agent in PRIMEQUEST. To use this in a virtual machine environment, see "5.1.2.5 Setting up the shutdown agent in virtual machine environment."

Note

・・・

After setting up the shutdown agent, conduct a test for forced shutdown of cluster nodes to check that the correct nodes can be forcibly stopped. For details of the test for forced shutdown of cluster nodes, see "1.4 Test".

・・・

5.1.2.4.1 MMB

This section describes the procedure for setting up the MMB in the shutdown facility.

Check the information of the shutdown agent before setting up the shutdown facility.

Setting up the MMB Shutdown Facility

Registering MMB information

Note

・・・

Carry out the MMB information registration described here after "5.1.1 Setting Up CF and CIP" and before "Setting Up the Shutdown Daemon", which is described later.

・・・

1. Execute the "clmmbsetup -a" command on all nodes, and register the MMB information.

   For instructions on using the "clmmbsetup" command, see the "clmmbsetup" manual page.

   ```
   # /etc/opt/FJSVcluster/bin/clmmbsetup -a mmb-user
   Enter User's Password:
   Re-enter User's Password:
   ```

   For *mmb-user* and User's Password, enter the following values that were checked in "5.1.2.1 Checking the Shutdown Agent Information."

   mmb-user

   　　User's name for controlling the MMB with RMCP

   User's Password

   　　User's password for controlling the MMB with RMCP.

   ## Note

   ・・・

   Only alphanumeric characters can be used for User's Password, but not symbols.

   ・・・

2. Execute the "clmmbsetup -l" command on all nodes, and check the registered MMB information.

   If the registered MMB information was not output on all nodes in Step 1, start over from Step 1.

   ```
   # /etc/opt/FJSVcluster/bin/clmmbsetup -l
   cluster-host-name   user-name
   ----------------------------------
   ```

```
node1              mmb-user
node2              mmb-user
```

Setting up the shutdown daemon

On all nodes, create /etc/opt/SMAW/SMAWsf/rcsd.cfg with the following information:

```
CFNameX,weight=weight,admIP=myadmIP: agent=SA_xxx,timeout=20
CFNameX,weight=weight,admIP=myadmIP: agent=SA_xxx,timeout=20
```

```
CFNameX         :  Specify the CF node name of the cluster host.
weight          :  Specify the weight of the SF node.
myadmIP         :  Specify the IP address of the administrative LAN for the local node.
SA_xxx          :  Specify the name of the shutdown agent.
                       To set the node to panic status through the MMB
                       Specify "SA_mmbp".
                       To reset the node through the MMB
                       Specify "SA_mmbr".
```

Example) Shown below is a setup example for a 2-node configuration.

```
# cat /etc/opt/SMAW/SMAWsf/rcsd.cfg
node1,weight=2,admIP=fuji2:agent=SA_mmbp,timeout=20:agent=SA_mmbr,timeout=20
node2,weight=2,admIP=fuji3:agent=SA_mmbp,timeout=20:agent=SA_mmbr,timeout=20
```

Note

. . . .

- For the shutdown agents to be specified in the rcsd.cfg file, set both the SA_mmbp and SA_mmbr shutdown agents in that order.

- Set the same contents in the rcsd.cfg file on all nodes. Otherwise, a malfunction may occur.

. . . .

Information

. . . .

When creating the /etc/opt/SMAW/SMAWsf/rcsd.cfg file, you can use the /etc/opt/SMAW/SMAWsf/rcsd.cfg.mmb.template file as a template.

. . . .

Starting the MMB asynchronous monitoring daemon

1. Starting the MMB asynchronous monitoring daemon

   Check that the MMB asynchronous monitoring daemon has been started on all nodes.

   ```
   # /etc/opt/FJSVcluster/bin/clmmbmonctl
   ```

   If "The devmmbd daemon exists." is displayed, the MMB asynchronous monitoring daemon has been started.

   If "The devmmbd daemon does not exist." is displayed, the MMB asynchronous monitoring daemon has not been started. Execute the following command to start the MMB asynchronous monitoring daemon.

   ```
   # /etc/opt/FJSVcluster/bin/clmmbmonctl start
   ```

2. Starting the shutdown facility.

   Check that the shutdown facility has been started on all nodes.

   ```
   # sdtool -s
   ```

   If the shutdown facility has already been started, execute the following command to restart the shutdown facility on all nodes.

   ```
   # sdtool -e
   # sdtool -b
   ```

If the shutdown facility has not been started, execute the following command to start the shutdown facility on all nodes.

```
# sdtool -b
```

3. Checking the status of the shutdown facility

Check the status of the shutdown facility on all nodes.

```
# sdtool -s
```

Information

Display results of the sdtool -s command

- If "InitFailed" is displayed as the initial status, it means that a problem occurred during initialization of that shutdown agent.

- If "TestFailed" is displayed as the test status, it means that a problem occurred while the agent was testing whether or not the node displayed in the cluster host field could be stopped. Some sort of problem probably occurred in the software, hardware, or network resources being used by that agent.

- If "Unknown" is displayed as the stop or initial status, it means that the SF has still not executed node stop, path testing, or SA initialization. "Unknown" is displayed temporarily until the actual status can be confirmed.

- If "TestFailed" or "InitFailed" is displayed, check the SA log file or /var/log/messages. The log file records the reason why SA testing or initialization failed. After the failure-causing problem is resolved and SF is restarted, the status display changes to InitWorked or TestWorked.

Note

- If "sdtool -s" is executed immediately after the OS is started, "TestFailed" may be displayed as the test status in the local node. However, this status is displayed because the snmptrapd daemon is still being activated and does not indicate a malfunction. If "sdtool -s" is executed 10 minutes after the shutdown facility is started, TestWorked is displayed as the test status.

In the following example, "TestFailed" is displayed test status for the local node (node1).

```
# sdtool -s
Cluster Host    Agent      SA State    Shut State   Test State  Init State
------------    -----      --------    ----------   ----------  ----------
node1           SA_mmbp.so Idle        Unknown      TestFailed  InitWorked
node1           SA_mmbr.so Idle        Unknown      TestFailed  InitWorked
node2           SA_mmbp.so Idle        Unknown      TestWorked  InitWorked
node2           SA_mmbr.so Idle        Unknown      TestWorked  InitWorked
```

The following messages may be displayed right after the OS is started by same reason as previously described.

```
3084: Monitoring another node has been stopped.
SA SA_mmbp.so to test host nodename failed
SA SA_mmbr.so to test host nodename failed
```

These messages are also displayed because the snmptrapd daemon is being activated and does not indicate a malfunction. The following message is displayed 10 minutes after the shutdown facility is started.

```
3083: Monitoring another node has been started.
```

- If a node is forcibly shut down by the SA_mmbr shutdown agent, the following message may be output. This message indicates that it takes time to shut down the node, not a malfunction.

```
Fork SA_mmbp.so(PID pid) to shutdown host nodename
        :
SA SA_mmbp.so to shutdown host nodename failed
        :
Fork SA_mmbr.so(PID pid) to shutdown host nodename
```

```
          :
SA SA_mmbr.so to shutdown host nodename failed
          :
MA SA_mmbp.so reported host nodename leftcluster, state MA_paniced_fsnotflushed
          :
MA SA_mmbr.so reported host nodename leftcluster, state MA_paniced_fsnotflushed
          :
Fork SA_mmbp.so(PID pid) to shutdown host nodename
          :
SA SA_mmbp.so to shutdown host nodename succeeded
```

If "sdtool -s" is executed immediately after the message above was output, KillWorked is displayed in Shut State of SA_mmbp.so. Then, KillFailed is displayed in Shut State of SA_mmbr.so.

The following indicates the example when shutting down nodes from node1 to node2 and executing "sdtool -s" after the message above.

```
# sdtool -s
Cluster Host    Agent        SA State    Shut State   Test State  Init State
------------    -----        --------    ----------   ----------  ----------
node1           SA_mmbp.so   Idle        Unknown      TestWorked  InitWorked
node1           SA_mmbr.so   Idle        Unknown      TestWorked  InitWorked
node2           SA_mmbp.so   Idle        KillWorked   TestWorked  InitWorked
node2           SA_mmbr.so   Idle        KillFailed   TestWorked  InitWorked
```

To restore KillFailed that was displayed by "sdtool -s," follow the procedure below.

```
# sdtool -e
# sdtool -b
# sdtool -s
Cluster Host    Agent        SA State    Shut State   Test State  Init State
------------    -----        --------    ----------   ----------  ----------
node1           SA_mmbp.so   Idle        Unknown      TestWorked  InitWorked
node1           SA_mmbr.so   Idle        Unknown      TestWorked  InitWorked
node2           SA_mmbp.so   Idle        Unknown      TestWorked  InitWorked
node2           SA_mmbr.so   Idle        Unknown      TestWorked  InitWorked
```

Setting the I/O Completion Wait Time

Set the wait time until I/O completion (WaitForIOComp) during failover triggered by a node failure (panic, etc.) according to the procedure described below.

1. Prechecking the shared disk

    The standard setting for the I/O completion wait time during failover triggered by a node failure (for example, if a panic occurs during MMB asynchronous monitoring) is 0 seconds. However, if a shared disk that requires an I/O completion wait time is being used, this setting must be set to an appropriate value.

    ### Information

    ETERNUS Disk storage systems do not require an I/O completion wait time. Therefore, this setting is not required.

    ### Note

    If an I/O completion wait time is set, the failover time when a node failure (panic, etc.) occurs increases by that amount of time.

2. Setting the I/O completion wait time

    Execute the following command, and set the wait time until I/O completion (WaitForIOComp) during failover triggered by a node failure (panic, etc.). For details about the "cldevparam" command, see the "cldevparam" manual page.

Execute the command in any node that is part of the cluster system.

```
# /etc/opt/FJSVcluster/bin/cldevparam -p WaitForIOComp value
```

Alternatively, execute the following command and check the setting of the wait time until I/O processing is completed (WaitForIOComp).

```
# /etc/opt/FJSVcluster/bin/cldevparam -p WaitForIOComp
 value
```

3. Starting the shutdown facility

   Check that the shutdown facility has been started on all nodes.

```
# sdtool -s
```

   If the shutdown facility has already been started, execute the following command to restart the shutdown facility on all nodes:

```
# sdtool -r
```

   If the shutdown facility has not been started, execute the following command to start the shutdown facility on all nodes.

```
# sdtool -b
```

4. Checking the status of the shutdown facility

   Check the status of the shutdown facility on all nodes.

```
# sdtool -s
```

5.1.2.5 Setting up the shutdown agent in virtual machine environment

This section describes the procedure for setting up the shutdown agent in a virtual machine environment.

📖 **Note**

. . . .

After setting up the shutdown agent, conduct a test for forced shutdown of cluster nodes to check that the correct nodes can be forcibly stopped. For details of the test for forced shutdown of cluster nodes, see "1.4 Test".

. . . .

5.1.2.5.1 vmSP (Virtual Machine Service Provider)

This section describes the procedure for setting vmSP (Virtual Machine Service Provider) as the shutdown facility in a Xen environment.

Be sure to perform "5.1.2.1 Checking the Shutdown Agent Information" before setting up the shutdown facility.

📖 **Note**

. . . .

Be sure to perform the following operations from 1. to 6. on all guest OSes (nodes).

. . . .

1. Encrypt the password

   Execute the sfcipher command to encrypt passwords for the account FJSVvmSP of all host OSes, where guest OSes set as cluster nodes exist.

   For details on how to use the sfcipher command, see the manual page of "sfcipher."

```
# sfcipher -c
Enter User's Password:
Re-enter User's Password:
D0860AB04E1B8FA3
```

2. Setup the shutdown agent

   Set the shutdown agent. When the shutdown agent for PANIC (SA_vmSPgp) is used, create /etc/opt/SMAW/SMAWsf/
   SA_vmSPgp.cfg, and when the shutdown agent for RESET (SA_vmSPgr) is used, create /etc/opt/SMAW/SMAWsf/
   SA_vmSPgr.cfg as below.

   ```
   CFNameX domainX ip-address user passwd
   CFNameX domainX ip-address user passwd
   ```

   ```
   CFNameX     : Specify the CF node name of the cluster host.
   domainX     : Specify the guest OS domain name.
   ip-address : Specify the IP address of the host OS.
   user        : Specify the account FJSVvmSP of the host OS.
   passwd      : A login password of the account FJSVvmSP of the host OS.
                 Specify the encrypted password that you have checked in 1.
   ```

   Example) The following is a setup example.

   When the guest OS domain name of node1 is domain1, and the IP address of the host OS on which node1 operates is 10.20.30.50.
   Moreover, when the guest OS domain name of node2 is domain2, and the IP address of the host OS on which node2 operates is
   10.20.30.51.

   ```
   # cat /etc/opt/SMAW/SMAWsf/SA_vmSPgp.cfg
   node1 domain1 10.20.30.50 FJSVvmSP D0860AB04E1B8FA3
   node2 domain2 10.20.30.51 FJSVvmSP D0860AB04E1B8FA3
   ```

   ```
   # cat /etc/opt/SMAW/SMAWsf/SA_vmSPgr.cfg
   node1 domain1 10.20.30.50 FJSVvmSP D0860AB04E1B8FA3
   node2 domain2 10.20.30.51 FJSVvmSP D0860AB04E1B8FA3
   ```

   ### 📌 Note

   ・

   - Check if the setting contents of the /etc/opt/SMAW/SMAWsf/SA_vmSPgp.cfg file and the /etc/opt/SMAW/SMAWsf/
     SA_vmSPgr.cfg file are correct. If there is an error in the setting contents, the shutdown facility cannot be performed normally.

   - Check if the domain name (domainX) of the guest OS and the IP address (ip-address) of the host OS corresponding to the cluster
     host's CF node name (CFNameX) of the /etc/opt/SMAW/SMAWsf/SA_vmSPgp.cfg file and the /etc/opt/SMAW/SMAWsf/
     SA_vmSPgr.cfg file are set. If there is an error in the setting, a different node may be forcibly stopped.

   ・

3. Log in to the host OS

   The shutdown facility accesses the target node with SSH. Therefore, you need to authenticate yourself (create the RSA key) in
   advance, which is required when using SSH for the first time.
   On all guest OSes (nodes), log in to each host OS IP address (ip-address) set in the step 2. using each host OS user name (user) set
   in the step 2.

   ```
   # ssh -l FJSVvmSP XXX.XXX.XXX.XXX
   The authenticity of host 'XXX.XXX.XXX.XXX (XXX.XXX.XXX.XXX)' can't be established.
   RSA key fingerprint is xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx.
   Are you sure you want to continue connecting (yes/no)? yes <- "Enter yes."
   ```

4. Setup the shutdown daemon

   Create /etc/opt/SMAW/SMAWsf/rcsd.cfg as below.

   ```
   CFNameX,weight=weight,admIP=myadmIP:agent=SA_xxxx,timeout=35
   CFNameX,weight=weight,admIP=myadmIP:agent=SA_xxxx,timeout=35
   ```

   ```
   CFNameX        : Specify the CF node name of the cluster host.
   weight         : Specify the weight of the SF node.
   myadmIP        : Specify the IP address of the administrative LAN for your guest OS (node).
   SA_xxxx        : Specify the name of the shutdown agent.
                    Here, "SA_vmSPgp" or "SA_vmSPgr" is specified.
   ```

Example) The following is a setup example.

```
# cat /etc/opt/SMAW/SMAWsf/rcsd.cfg
node1,weight=2,admIP=fuji2:agent=SA_vmSPgp,timeout=35:agent=SA_vmSPgr,timeout=35
node2,weight=2,admIP=fuji3:agent=SA_vmSPgp,timeout=35:agent=SA_vmSPgr,timeout=35
```

 Note
. . . .

- For the shutdown agent set by the rcsd.cfg file, set both shutdown agents in the order of SA_vmSPgp to SA_vmSPgr.

- The contents of the SA_vmSPgp.cfg, SA_vmSPgr.cfg, and rcsd.cfg files of all guest OSes (nodes) should be identical. If not, a malfunction will occur.

. . . .

5. Start the shutdown facility

    Check that the shutdown facility has already been started on all the nodes.

```
# sdtool -s
```

    If the shutdown facility has already been started, execute the following to restart it on all the nodes.

```
# sdtool -e
# sdtool -b
```

    If the shutdown facility has not been started, execute the following to start it on all the nodes.

```
# sdtool -b
```

6. Check the state of the shutdown facility

    Check the state of the shutdown facility.

```
# sdtool -s
```

 Information
. . . .

About the displayed results

- If "InitFailed" is displayed as the initial status, it means that a problem occurred during initialization of that shutdown agent.

- If "TestFailed" is displayed as the test status, it means that a problem occurred while the agent was testing whether or not the node displayed in the cluster host field could be stopped. Some sort of problem probably occurred in the software, hardware, network resources, or the host OS being used by that agent.

- When the maximum concurrent connections for SSH are "the number of cluster nodes" or less, the status of the shutdown facility may be displayed as InitFailed or TestFailed. Change the configuration to set up the maximum concurrent connections for SSH to be "the number of cluster nodes + 1"or more.

- If "Unknown" is displayed as the stop or initial status, it means that the SF has still not executed node stop, path testing, or SA initialization. "Unknown" will be displayed temporarily until the actual status can be confirmed.

- If "TestFailed" or "InitFailed" is displayed, check the SA log file or /var/log/messages. The log file records the reason why SA testing or initialization failed. After the failure-causing problem is resolved and SF is restarted, the status display changes to InitWorked or TestWorked.

. . . .

5.1.2.5.2 libvirt

This section describes the procedure for setting libvirt as the shutdown facility in a KVM environment.

Be sure to perform "5.1.2.1 Checking the Shutdown Agent Information" before setting up the shutdown facility.

> 📝 **Note**
> . . . .
> Be sure to perform the following operations from 1. to 6. on all guest OSes (nodes).
> . . . .

1. Encrypt the password

   Execute the sfcipher command to encrypt the password for a user for the shutdown facility.

   For details on how to use the sfcipher command, see the manual page of "sfcipher."

   ```
   # sfcipher -c
   Enter User's Password:
   Re-enter User's Password:
   D0860AB04E1B8FA3
   ```

2. Setup the shutdown agent

   Set the shutdown agent. When the shutdown agent for PANIC (SA_libvirtgp) is used, create /etc/opt/SMAW/SMAWsf/ SA_libvirtgp.cfg, and when the shutdown agent for RESET (SA_libvirtgr) is used, create /etc/opt/SMAW/SMAWsf/ SA_libvirtgr.cfg as below.

   ```
   CFNameX domainX ip-address user passwd
   CFNameX domainX ip-address user passwd
   ```

   ```
   CFNameX     : Specify the CF node name of the cluster host.
   domainX     : Specify the guest OS domain name.
   ip-address  : Specify the IP address of the hypervisor.
   user        : Account of the hypervisor. Specify the user for the shutdown facility.
   passwd      : A login password of the account specified by "user".
                    Specify the encrypted password that you have checked in 1.
   ```

   Example) The following is a setup example.

   When the guest OS domain name of node1 is domain1, and the IP address of the hypervisor on which node1 operates is 10.20.30.50. Moreover, when the guest OS domain name of node2 is domain2, and the IP address of the hypervisor on which node2 operates is 10.20.30.51.

   ```
   # cat /etc/opt/SMAW/SMAWsf/SA_libvirtgp.cfg
   node1 domain1 10.20.30.50 user D0860AB04E1B8FA3
   node2 domain2 10.20.30.51 user D0860AB04E1B8FA3
   ```

   ```
   # cat /etc/opt/SMAW/SMAWsf/SA_libvirtgr.cfg
   node1 domain1 10.20.30.50 user D0860AB04E1B8FA3
   node2 domain2 10.20.30.51 user D0860AB04E1B8FA3
   ```

   > 📝 **Note**
   > . . . .
   > - Check if the setting contents of the /etc/opt/SMAW/SMAWsf/SA_libvirtgp.cfg file and the /etc/opt/SMAW/SMAWsf/ SA_libvirtgr.cfg file are correct. If there is an error in the setting contents, the shutdown facility cannot be performed normally.
   >
   > - Check if the domain name (domainX) of the guest OS and the IP address (ip-address) of the hypervisor corresponding to the cluster host's CF node name (CFNameX) of the /etc/opt/SMAW/SMAWsf/SA_libvirtgp.cfg file and the /etc/opt/SMAW/ SMAWsf/SA_libvirtgr.cfg file are set. If there is an error in the setting, a different node may be forcibly stopped.
   >
   > - After forcibly stopping the node (guest OS) by SA_libvirtgp, the state of the guest OS may remain temporarily stopped (when there is no sufficient disk space in /var/crash on the host OS, and so on.). In this case, forcibly stop the guest OS with the virsh destroy command.
   > . . . .

3. Log in to the hypervisor

   The shutdown facility accesses the target node with SSH. Therefore, you need to authenticate yourself (create the RSA key) in advance, which is required when using SSH for the first time.

On all guest OSes (nodes), log in to each hypervisor IP address (ip-address) set in the step 2. using each user for the shutdown facility.

```
# ssh -l user XXX.XXX.XXX.XXX
The authenticity of host 'XXX.XXX.XXX.XXX (XXX.XXX.XXX.XXX)' can't be established.
RSA key fingerprint is xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx.
Are you sure you want to continue connecting (yes/no)? yes <- "Enter yes."
```

4. Setup the shutdown daemon

   Create /etc/opt/SMAW/SMAWsf/rcsd.cfg as below.

```
CFNameX,weight=weight,admIP=myadmIP:agent=SA_xxxx,timeout=35
CFNameX,weight=weight,admIP=myadmIP:agent=SA_xxxx,timeout=35
```

```
CFNameX       : Specify the CF node name of the cluster host.
weight        : Specify the weight of the SF node.
myadmIP       : Specify the IP address of the administrative LAN for your guest OS (node).
SA_xxxx       : Specify the name of the shutdown agent.
                 Here, "SA_libvirtgp" or "SA_libvirtgr" is specified.
```

   Example) The following is a setup example.

```
# cat /etc/opt/SMAW/SMAWsf/rcsd.cfg
node1,weight=1,admIP=10.20.30.100:agent=SA_libvirtgp,timeout=35:agent=SA_libvirtgr,timeout=35
node2,weight=1,admIP=10.20.30.101:agent=SA_libvirtgp,timeout=35:agent=SA_libvirtgr,timeout=35
```

📑 **Note**

. . . .

- For the shutdown agent set by the rcsd.cfg file, set both shutdown agents in the order of SA_libvirtgp to SA_libvirtgr.

- The contents of the SA_libvirtgp.cfg, SA_libvirtgr.cfg, and rcsd.cfg files of all guest OSes (nodes) should be identical. If not, a malfunction will occur.

. . . .

5. Start the shutdown facility

   Check that the shutdown facility has already been started on all the nodes.

```
# sdtool -s
```

   If the shutdown facility has already been started, execute the following to restart it on all the nodes.

```
# sdtool -e
# sdtool -b
```

   If the shutdown facility has not been started, execute the following to start it on all the nodes.

```
# sdtool -b
```

6. Check the state of the shutdown facility

   Check the state of the shutdown facility.

```
# sdtool -s
```

📖 **Information**

. . . .

About the displayed results

- If "InitFailed" is displayed as the initial status, it means that a problem occurred during initialization of that shutdown agent.

- If "TestFailed" is displayed as the test status, it means that a problem occurred while the agent was testing whether or not the node displayed in the cluster host field could be stopped. Some sort of problem probably occurred in the software, hardware, network resources, or the host OS being used by that agent.

- When the maximum concurrent connections for SSH are "the number of cluster nodes" or less, the status of the shutdown facility may be displayed as InitFailed or TestFailed. Change the configuration to set up the maximum concurrent connections for SSH to be "the number of cluster nodes + 1"or more.

- If "Unknown" is displayed as the stop or initial status, it means that the SF has still not executed node stop, path testing, or SA initialization. "Unknown" will be displayed temporarily until the actual status can be confirmed.

- If "TestFailed" or "InitFailed" is displayed, check the SA log file, /var/log/messages, or /etc/sysconfig/libvirt-guests. The log file records the reason why SA testing or initialization failed. For /etc/sysconfig/libvirt-guests, check whether settings in "3.2.1.1 Host OS setup" have been made. After the failure-causing problem is resolved and SF is restarted, the status display changes to InitWorked or TestWorked.

5.1.2.5.3 vmchkhost

This section describes the procedure for setting up vmchkhost (Host OS check) as the shutdown agent in a virtual machine environment.

Set up this setting after setting vmSP (Virtual Machine Service Provider) or libvirt in the shutdown facility.

Note

Be sure to perform the following operations from 1. to 6. on all guest OSes (nodes).

Information

About a log file

A log file of the vmchkhost shutdown agent is output at the following:

- /var/opt/SMAWsf/log/SA_vmchkhost.log

1. Encrypt the password

   In a Xen environment, use the encrypted passwords for the account FJSVvmSP of all host OSes, which were used when you set vmSP (Virtual Machine Service Provider) to the shutdown facility.

   In a KVM environment, use the encrypted passwords of general users for the shutdown facility, which were used when you set libvirt to the shutdown facility.

2. Set up the shutdown agent.

   Create /etc/opt/SMAW/SMAWsf/SA_vmchkhost.cfg as described in the following:

   ```
   guest-cfname host-cfname ip-address user password
   guest-cfname host-cfname ip-address user password
   ```

   ```
   guest-cfname      : CF node name of the guest OS.
   host-cfname       : CF node name of the host OS.
   ip-address        : An IP address of the host OS.
   user              : An account of the host OS.
                         For a Xen environment, FJSVvmSP is fixed.
                         For a KVM environment, specify the user which were created when you set
                         libvirt for the shutdown facility.
   password          : A login password of the account specified by "user".
                         Specify the encrypted password that you have checked in 1.
   ```

   Example) The following is a setup example.

When the CF node name of the host OS on which node1 (CF node name of the guest OS) operates is hostos1, the IP address of the host OS is 10.20.30.50, the CF node name of the host OS on which node2 (CF node name of the guest OS) operates is hostos2, and the IP address of the host OS is 10.20.30.51.

For Xen environment:

```
# cat /etc/opt/SMAW/SMAWsf/SA_vmchkhost.cfg
node1 hostos1 10.20.30.50 FJSVvmSP 3CA1wxVXKD8a93077BaEkA==
node2 hostos2 10.20.30.51
```

For KVM environment:

```
# cat /etc/opt/SMAW/SMAWsf/SA_vmchkhost.cfg
node1 hostos1 10.20.30.50 user D0860AB04E1B8FA3
node2 hostos2 10.20.30.51 user D0860AB04E1B8FA3
```

📝 **Note**

. . . .

- Check if the setting contents of the /etc/opt/SMAW/SMAWsf/SA_vmchkhost.cfg file are correct. If there is an error in the setting contents, the shutdown facility cannot be performed normally.

- Check if the domain name (domainX) of the guest OS and the IP address (ip-address) of the host OS corresponding to the cluster host's CF node name (CFNameX) of the /etc/opt/SMAW/SMAWsf/SA_vmchkhost.cfg file are set. If there is an error in the setting, the shutdown facility cannot be performed normally.

. . . .

3. Log in to the host OS

The shutdown facility accesses the target node with SSH. Therefore, you need to authenticate yourself (create the RSA key) in advance, which is required when using SSH for the first time.

Check that you have already authenticated yourself (created the RSA key) when setting up vmSP (Virtual Machine Service Provider) or libvirt in the shutdown facility.

4. Set up the shutdown daemon

Create /etc/opt/SMAW/SMAWsf/rcsd.cfg as below.

```
CFNameX,weight=weight,admIP=myadmIP:agent=SA_xxxx,timeout=35
CFNameX,weight=weight,admIP=myadmIP:agent=SA_xxxx,timeout=35
```

```
CFNameX          :Specify the CF node name of the cluster host.
weight           :Specify the weight of the SF node.
myadmIP          :Specify the IP address of the administrative LAN for your guest OS (node).
SA_xxxx          :Specify the name of the shutdown agent.
                    Here, "SA_vmchkhost" is specified.
```

Example) The following is a setup example.

For Xen environment:

```
# cat /etc/opt/SMAW/SMAWsf/rcsd.cfg
node1,weight=2,admIP=fuji2:agent=SA_vmSPgp,timeout=35:agent=SA_vmSPgr,timeout=35:agent=SA_vmchk
host,timeout=35
node2,weight=1,admIP=fuji3:agent=SA_vmSPgp,timeout=35:agent=SA_vmSPgr,timeout=35:agent=SA_vmchk
host,timeout=35
```

For KVM environment:

```
# cat /etc/opt/SMAW/SMAWsf/rcsd.cfg
node1,weight=2,admIP=fuji2:agent=SA_libvirtgp,timeout=35:agent=SA_libvirtgr,timeout=35:agent=SA
_vmchkhost,timeout=35
node2,weight=1,admIP=fuji3:agent=SA_libvirtgp,timeout=35:agent=SA_libvirtgr,timeout=35:agent=SA
_vmchkhost,timeout=35
```

5. Start the shutdown facility

   Check that the shutdown facility has already been started on all the nodes.

   ```
   # sdtool -s
   ```

   If the shutdown facility has already been started, execute the following to restart it on all the nodes.

   ```
   # sdtool -e
   # sdtool -b
   ```

   If the shutdown facility has not been started, execute the following to start it on all the nodes.

   ```
   # sdtool -b
   ```

6. Check the state of the shutdown facility

   Check the state of the shutdown facility.

   ```
   # sdtool -s
   ```

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. Registering Hardware Devices

   Register the connected hardware devices (shared disks and network interface cards) to the resource database that CRM manages.

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.
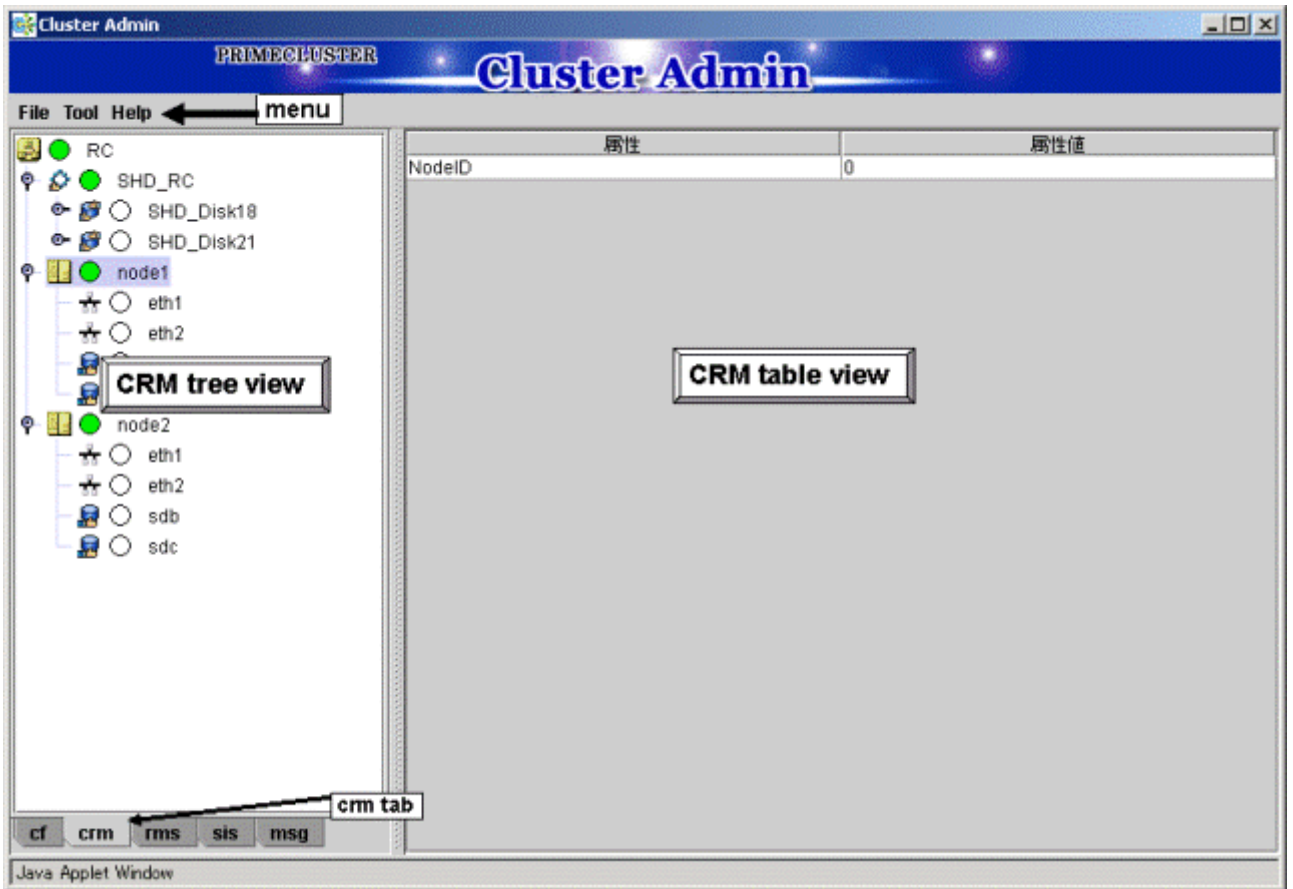
The areas shown in the screen are described below.

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 5.1 Screen for cluster resource management facility
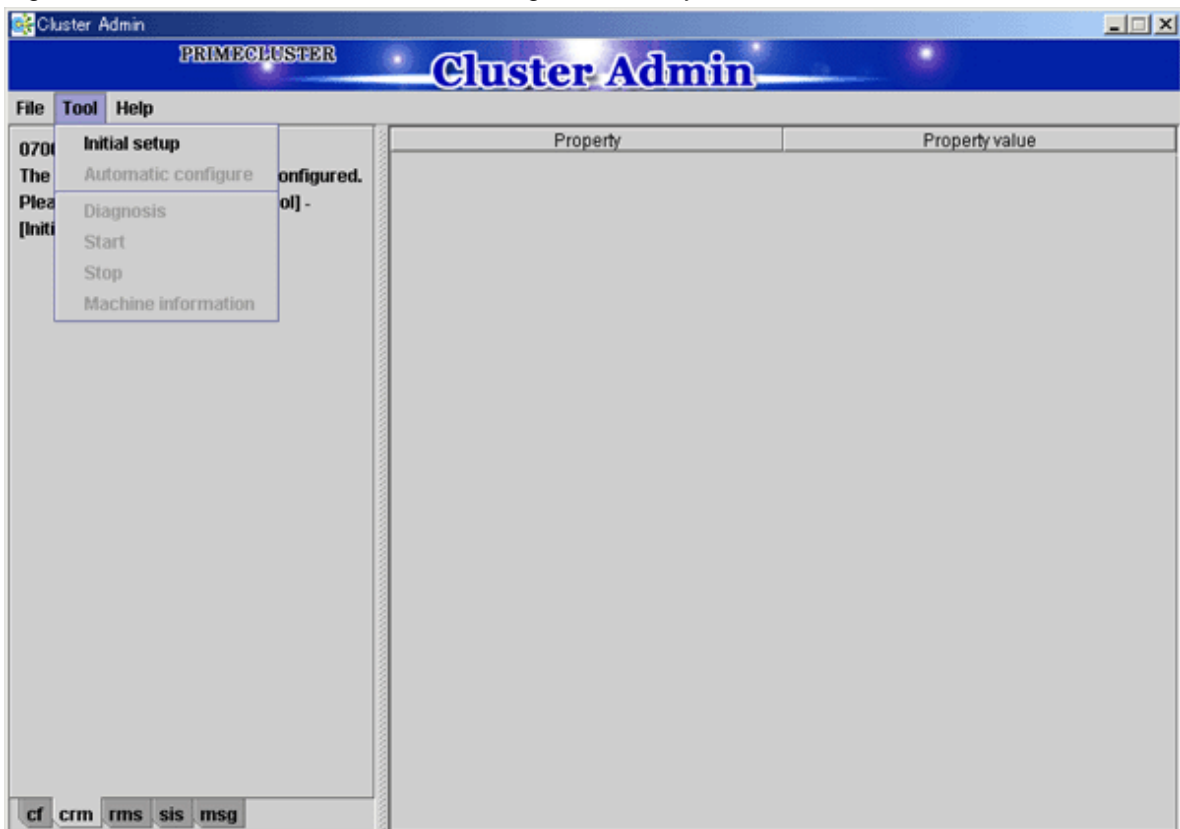
![Note icon] **Note**

The *Initial setup* can be selected only if the resource database has not been set.

2. The screen for initial setup is displayed.

Figure 5.2 Screen for initial setup
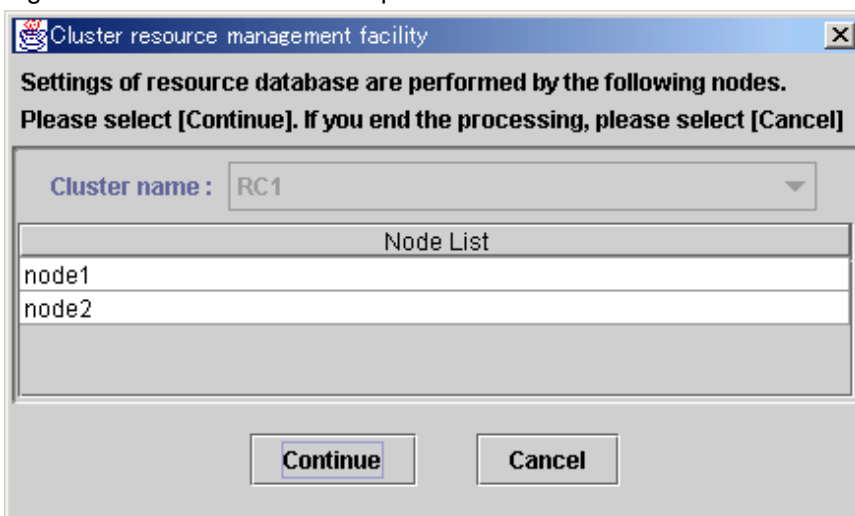
Cluster name

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 icon] **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.
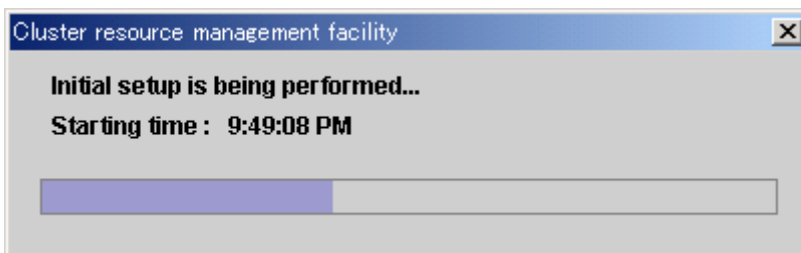
5. When initial setup ends, the following message is displayed.
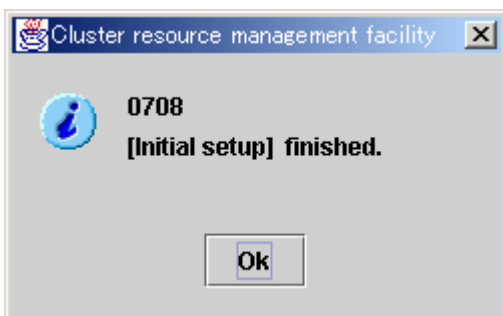
![Note icon] **Note**

. . . .

- 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*"

        - "11.10 Cluster Resource Management Facility Messages"

- Appendix D in this manual

  - "D.5.2 Failed Resource and Operator Intervention Messages (GUI)"

- 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 Registering Hardware Devices

The automatic configure process registers the connected hardware devices (shared disks and network interface cards) to the resource database.

Operation Procedure:

1. Registering the network interface card

   1. Confirm that all nodes have been started in multi-user mode.

   2. Log in to any of nodes using system administrator access privileges and execute the "clautoconfig" command. Then, the network interface is automatically recognized and registered in the resource database.

      ```
      #/etc/opt/FJSVcluster/bin/clautoconfig -r -n [RETURN]
      ```

   For details on this command, see the manual pages of "clautoconfig".

   ### 📒 Note

   - Do not execute the "clautoconfig" command on the node in which the "clautoconfig" command is being executed or on any other node while the "clautoconfig" command is being executed. If you execute it, the shared disk device cannot be registered correctly. If you have executed it, execute the following operation on all the nodes that constitute the cluster system to re-execute "5.1.3 Initial Setup of the Cluster Resource Management Facility" described in this chapter:

     1. Reset the resource database using the "clinitreset" command. For details on this command, see the manual pages of "clinitreset".

     2. Reboot the node.

   - Only an active network interface card is automatically detected. Confirm the state of the network interface card using the "ifconfig(8)" command.
     Execute the following command to activate the network interface:

     ```
     # ifconfig network interface card up
     ```

     (Example) Enable the network interface card eth1

     ```
     # ifconfig eth1 up
     ```

2. Registering a shared disk

   When you use GDS, register a shared disk in the resource database using the following steps on any one of the nodes of the cluster system. For details on the procedure, see "Appendix H" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*"

   1. Log in any one of the nodes of the cluster system using system administrator access privileges.

   2. Create the shared disk definition file in the following format. The configuration file defines settings of a shared disk connected to all the nodes of the cluster system. Create the shared disk definition file with an arbitrary name.

      ```
      <Resource key name> <device name> <node identifier>
      <Resource key name> <device name> <node identifier>
                          :
      ```

      - Define "resource key name device name node identifier" for each shared disk in one row.

      - "resource key name", "device name", and "node identifier" are delimited by a single space.

- Set up resource key name, device name and node identifier as follows;

Resource key name

Specify a resource key name that indicates the sharing relationship for each shared disk. You need to specify the same name for each disk shared between nodes. The resource key name should be specified in the "shd number" format. "shd" is a fixed string. For "number", you can specify any four-digit numbers. If multiple shared disks are used, specify unique numbers for each shared disk.

```
(Example) When /dev/sdb and /dev/sdc are shared between nodes

  Resource key name of /dev/sdb: shd0001
  Resource key name of /dev/sdc: shd0002
```

Device name

Specify a device name by the full device path of the shared disk.

```
(Example) In the case of /dev/sdb

  /dev/sdb
```

Node identifier

Specify a node identifier for which the shared disk device is available. Confirm the node identifier by executing the "clgettree" command. For details on this command, see the manual pages of "clgettree".

```
(Example) node1 and node2 are node identifiers in the following case:

  # /etc/opt/FJSVcluster/bin/clgettree
  Cluster 1 cluster
      Domain 2 PRIME
          Shared 7 SHD_PRIME
          Node 3 node1 ON
          Node 5 node2 ON
```

The following example shows the configuration file of the shared disk when shared disks /dev/sdb and /dev/sdc are shared between node1 and node2.

```
shd0001 /dev/sdb node1
shd0001 /dev/sdb node2
shd0002 /dev/sdc node1
shd0002 /dev/sdc node2
```

- When adding the shared disk device and registering the added shared disk device on the resource database, define only the information of the added shared disk device.

```
Example: When registering the added disk device /dev/sdd (*1) on the resource database after
shd0001 and shd0002 are already registered on the resource database:
```

```
shd0003 /dev/sdd node1
shd0003 /dev/sdd node2
```

(*1) Note
The device name of the added shared disk device may not follow the device name of the registered device in alphabetical order. Make sure to check the device name of the added shared disk device before defining the information of the added disk device.

3. Execute the "clautoconfig" command to register the settings of the shared disk device that is stored in the shared disk definition file in the resource database.

Specify the "clautoconfig" command in the following format:

```
(Format)
```

```
/etc/opt/FJSVcluster/bin/clautoconfig -f [full path of the shared disk definition file]
```

(Example)

```
# /etc/opt/FJSVcluster/bin/clautoconfig -f /var/tmp/diskfile
```

Note

∙∙

- If the "clautoconfig" command ends abnormally, take corrective action according to the error message. For details on the messages of this command, see the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*."

- This command does not check whether the shared disk defined in the configuration file is physically connected.

- If you found an error in the shared disk definition file after executing the "clautoconfig" command, reset the resource database by executing the "clinitreset" command and reboot the node.

∙∙

4. When the initial setup and automatic configure are completed, the following screen appears.
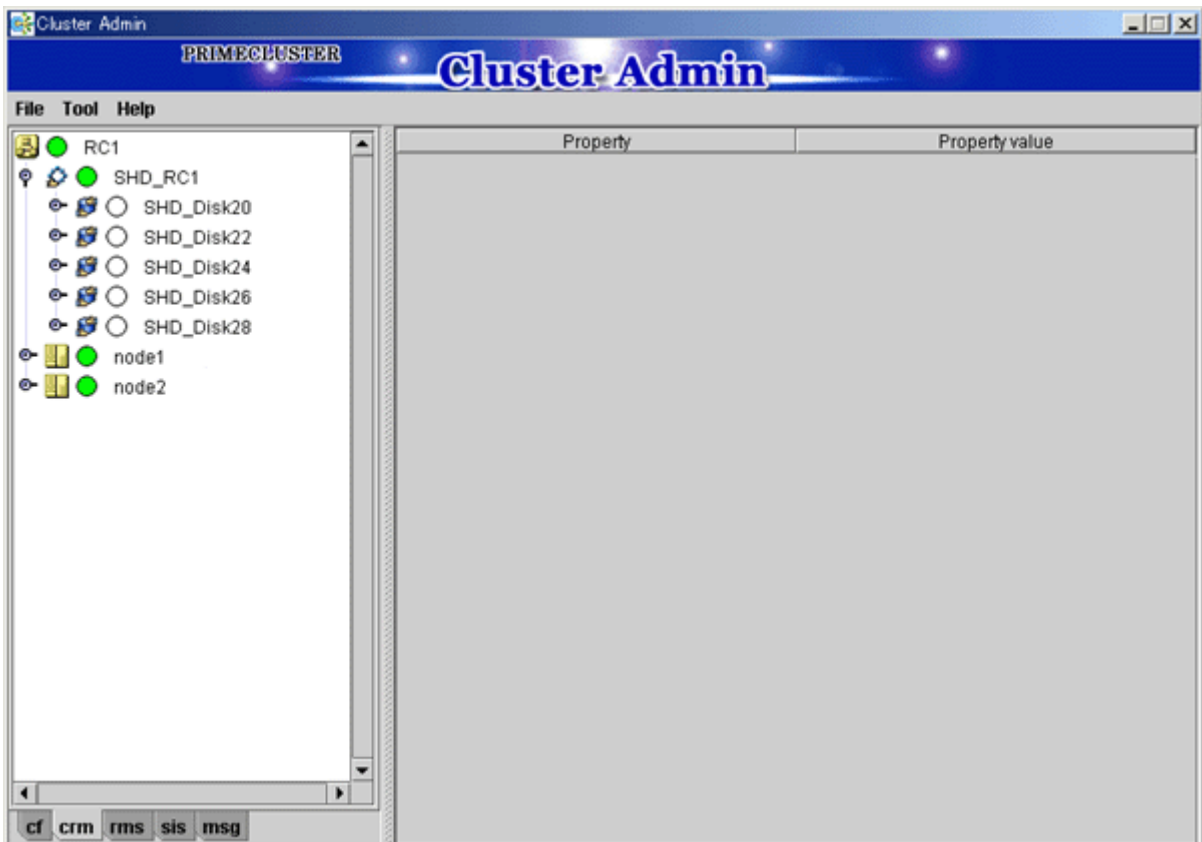
5. When automatic configuration is completed, go to the CRM main window and confirm that the resource registration is completed by checking the following.

- Whether the disk configuration is different among the nodes.

- Whether the number of disks in each node differs from the number of shared disk devices.

- Whether the number of shared disk device resources is less than the actual device configuration.

- Whether any disks other than shared disk device are registered as the shared disk device.

- Whether the number of public LAN resources is less than the actual device configuration.

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.

If the resources are not registered correctly, first review the above causes.

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"

    - "11.10 Cluster Resource Management Facility Messages"

  - Appendix D in this manual

    - "D.5.2 Failed Resource and Operator Intervention Messages (GUI)"

- 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.

5.2 Setting Up Fault Resource Identification and Operator Intervention Request

The fault resource identification is a function that outputs a message to syslogd(8) 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. 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(8) 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.
```

After setting the initial configuration of the resource database, specify the settings for enabling the fault resource identification and the operator intervention request.

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.5.2 Failed Resource and Operator Intervention Messages (GUI)" and "D.4 Operator Intervention Messages."

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.

Take the following procedures to enable the fault resource identification and the operator intervention request:

- Procedure to display the messages of the fault resource identification and the operator intervention request on the console, and to enable these functions

    1. Execute "Preparation prior to displaying fault resource identification and operator intervention request messages" on all nodes.

    2. Apply the settings described in "Enabling the operation of fault resource identification and operator intervention request."

- Procedure to enable the fault resource identification and the operator intervention request without displaying the messages of these functions on the console

    1. Apply the settings described in "Enabling the operation of fault resource identification and operator intervention request."

Preparation prior to displaying fault resource identification and operator intervention request messages

The fault resource identification and operator intervention request messages are displayed by using syslogd(8) / rsyslogd(8). daemon.err is specified to determine the priority (facility.level) of the fault resource identification and operator intervention request messages. For details on the priority, see the manual page describing syslog.conf(5) / rsyslogd.conf(5). To display the fault resource identification and operator intervention request messages on the console, execute the following procedure on all the nodes.

Procedure:

1. Log in the node using system administrator access privileges.

2. Check the setting of syslogd / rsyslogd in /etc/syslog.conf to see that daemon.err is set to be displayed on the console.

    - RHEL5

        1. Check the setting of syslogd in /etc/syslog.conf to see that daemon.err is set to be displayed on the console.

            (Example) Daemon.err is set to be displayed on the console.

            ```
            daemon.err          /dev/console
            ```

            For further details on /etc/syslog.conf, see the manual pages of syslog.conf(5).

        2. If daemon.err is not set to be displayed on the console, change the setting of syslogd in /etc/syslog.conf.

            To enable this change, restart the system log daemon by executing the following command.

            ```
            # /etc/init.d/syslog restart
            ```

    - RHEL6

        1. Check the setting of rsyslogd in /etc/rsyslog.conf to see that daemon.err is set to be displayed on the console.

            (Example) Daemon.err is set to be displayed on the console.

            ```
            daemon.err          /dev/console
            ```

            For further details on /etc/rsyslog.conf, see the manual pages of rsyslog.conf(5).

        2. If daemon.err is not set to be displayed on the console, change the setting of rsyslogd in /etc/rsyslog.conf.

            To enable this change, restart the system log daemon by executing the following command.

            ```
            # /etc/init.d/syslog restart
            ```

3. Starting the console.

    Start the console.

    (Example)

    ```
    # kterm –C or # xterm –C
    ```

> 📝 **Note**
> ....
> If a Japanese-language system is being used, use a console that supports Japanese.
>
> By setting the operator intervention request messages, messages other than the operator intervention request messages will be displayed on the console. If a console that does not support Japanese is used, Japanese messages will not be displayed.
> ....

Enabling the operation of fault resource identification and operator intervention request

1. Execute the "clsetparam" 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" 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 cluster nodes.

   The functions of the fault resource identification and the operator intervention request will work after starting the node.

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" 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" 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 cluster nodes.

   The functions of the fault resource identification and the operator intervention request will not work after starting the node.

Chapter 6 Building Cluster Applications

The procedure for building a cluster application is shown below.
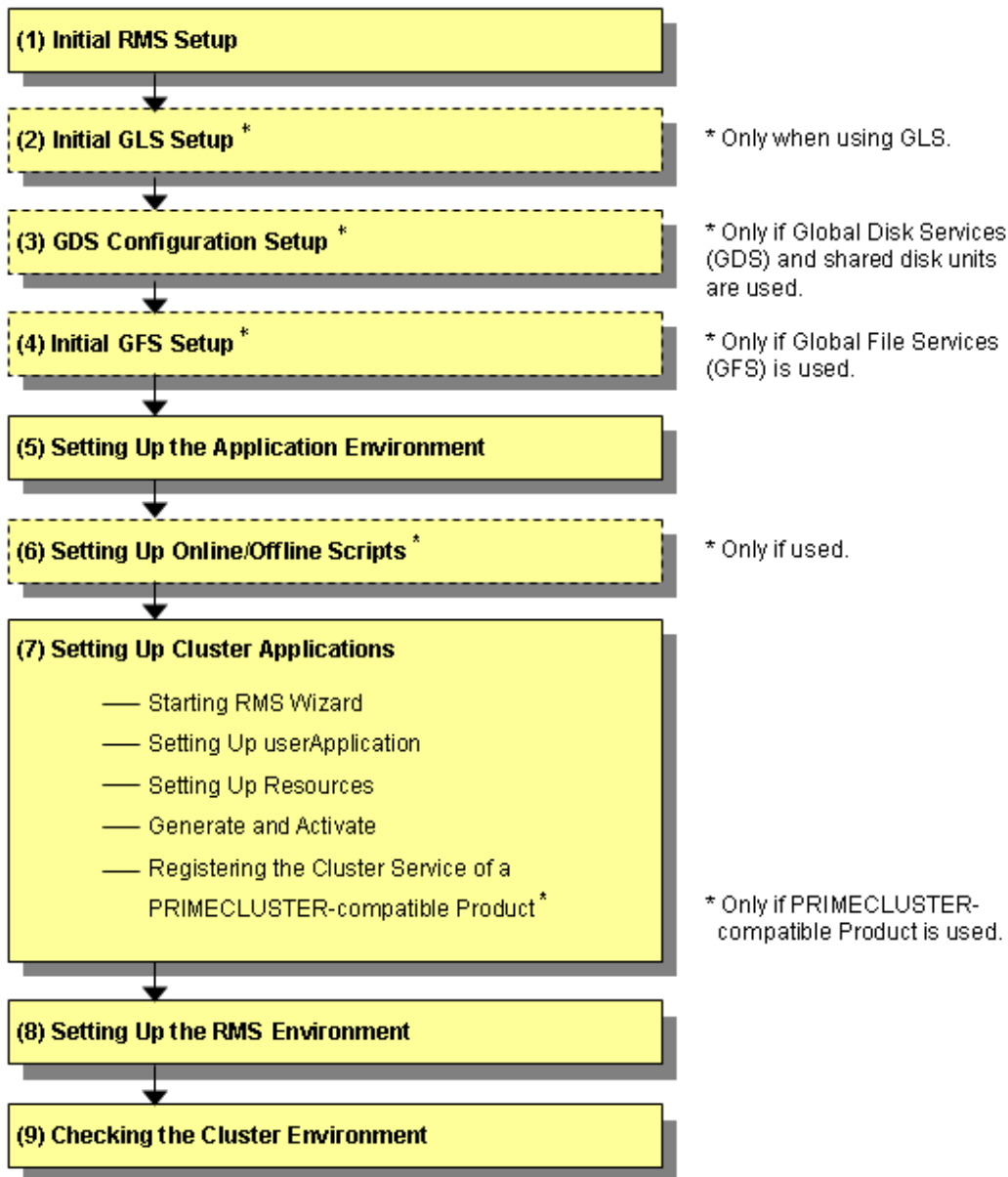
Table 6.1 Application building procedure and manual reference locations

	Work item	Execution Nodes	Required/ optional	Manual reference location*
(1)	6.1 Initial RMS Setup	All nodes	Required	**RMS** "7.1.2 Starting RMS automatically at boot time"
(2)	6.2 Initial GLS Setup	All nodes	Optional (Required when the GLS redundant line control function is used)	GLSR
(3)	6.3 GDS Configuration Setup	All nodes	Optional (required when GDS is used)	**GDSG** " Chapter 5 Operations"

	Work item	Execution Nodes	Required/ optional	Manual reference location*
(4)	6.4 Initial GFS Setup	All nodes	Optional (required when GFS is used)	GFSG
(5)	6.5 Setting Up the Application Environment	All nodes	Required	Manuals for each application
(6)	6.6 Setting Up Online/Offline Scripts	All nodes	Optional	**RMS** "2.9 Environment variables", "12 Appendix - Environment Variables"
(7)	6.7.1 Starting RMS Wizard	All nodes	Required	-
	6.7.2 Setting Up userApplication			
	6.7.3 Setting Up Resources			
	6.7.4 Generate and Activate			
	6.7.5 Registering the Cluster Service of a PRIMECLUSTER-compatible product	All nodes	Optional (required when a PRIMECLUSTER-compatible product is used)	
(8)	6.8 Setting Up the RMS Environment	All nodes	Required	**RMS** "2.9 Environment variables", "12 Appendix - Environment Variables"
(9)	6.9 Checking the Cluster Environment	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*

 - **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*

6.1 Initial RMS Setup

When RMS is to be used, you must first check the "A.8.1 RMS Setup Worksheet" and change the following environment variable as required:

 - Automatic startup of RMS (HV_RCSTART)

In this version, "Start up automatically" is set as default.

If you want to set RMS not to be started automatically along with node startup, select "Does not start up automatically."

When RMS is to be used, we recommend that you set "Start up automatically."

6.2 Initial GLS Setup

This section outlines the steps for configuring Global Link Services (GLS).

6.2.1 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.*"

This section describes how to set up "Example of the Cluster system" of "Example of configuring NIC switching mode (IPv4)" that GLS (redundant line control function) provides. This procedure is described in the example below.

![Note icon] **Note**

Although it is possible to have "takeover network" for PRIMECLUSTER and "IP address takeover" for GLS on the same cluster system, you must not configure them on the same interface. If you do so, the communication through "takeover IP address" will be disabled.

For example, when you select 'eth1' for the interface when you set "takeover network" for PRIMECLUSTER, do not use 'eth1' for GLS environment settings (do not specify 'eth1' by using the '-t' option for "hanetconfig create" command).

When you need to duplex the interface for a takeover network, use "IP address takeover" for GLS. You cannot set "takeover network" for the bonding interface.

The setup values correspond to those on the "A.8.3 GLS Setup Worksheet."
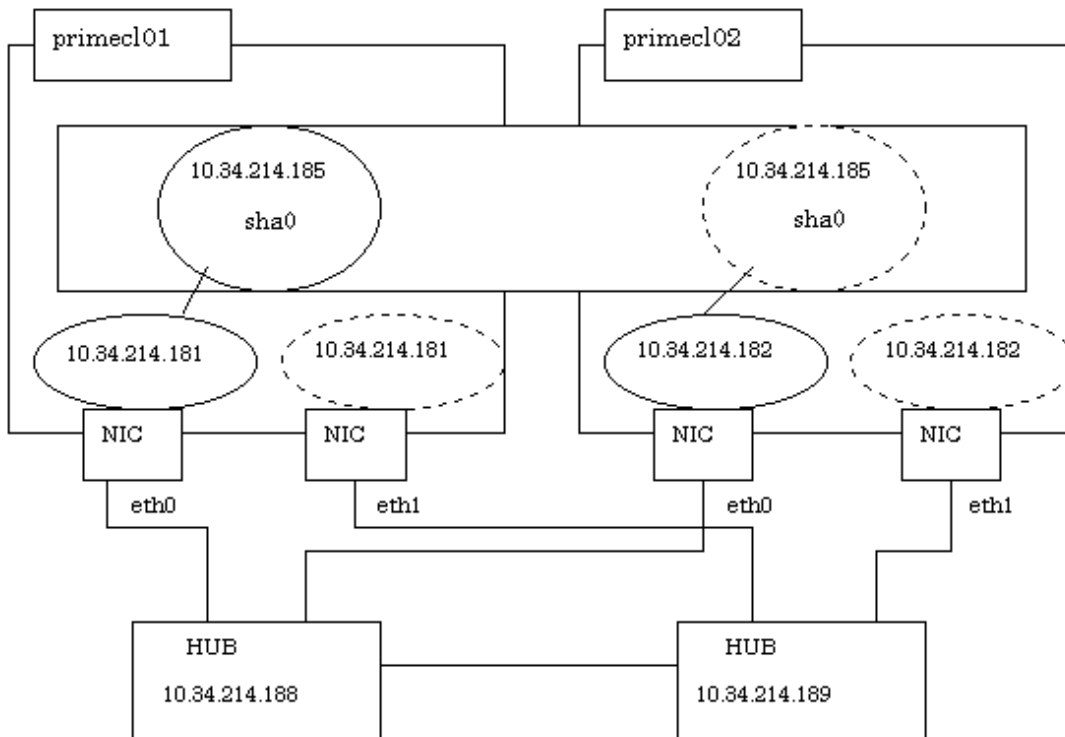
Operation Procedure:

If the OPERATING node is [HOST-primecl01]

1. Setting up the system

    1. Define the IP address and Host name in /etc/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 physical interface definition file as follows.

        - Contents of /etc/sysconfig/network-scripts/ifcfg-eth0

        ```
        DEVICE=eth0
        BOOTPROTO=static
        HWADDR=XX:XX:XX:XX:XX:XX
        HOTPLUG=no
        BROADCAST=10.34.214.255
        IPADDR=10.34.214.181
        NETMASK=255.255.255.0
        NETWORK=10.34.214.0
        ONBOOT=yes
        TYPE=Ethernet
        ```

        - Contents of /etc/sysconfig/network-scripts/ifcfg-eth1

        ```
        DEVICE=eth1
        BOOTPROTO=static
        HWADDR=XX:XX:XX:XX:XX:XX
        HOTPLUG=no
        ONBOOT=yes
        TYPE=Ethernet
        ```

    ## 📝 Note

    ....

    If the basic software is RHEL5 or RHEL6, and you specify "HWADDR=XX:XX:XX:XX:XX:XX" in the settings for the physical interfaces bundled by GLS (/etc/sysconfig/network-scripts/ifcfg-ethX file), add "HOTPLUG=no." In the RHEL5 environment, disable the hot plug function settings for the virtual interface (sha) by editing the /etc/udev/rules.d/60-net.rules file.

    ....

    ## 📖 Information

    ....

    "HOTPLUG=no" used when the basic system (OS) is RHEL5 or RHEL6 does not disable the PCI hot plug function.

    You can perform hot maintenance for NIC (PCI card) to the physical interfaces with "HOTPLUG=no."

    ....

    3. When the system is RHEL, on the /etc/sysconfig/network file, define a statement which enables the network configuration.

    ```
    NETWORKING=yes
    NETWORKING_IPV6=no
    ```

    4. Setting Hot Plug

    Edit the /etc/udev/rules.d/60-net.rules file for RHEL5 environments, and nullify the Hot Plug to virtual interface (sha). The setting of this file is unnecessary in RHEL6 environments.

```
SUBSYSTEM=="net", ENV{INTERFACE}=="sha*", GOTO="skipgls"
ACTION=="add", SUBSYSTEM=="net", IMPORT{program}="/lib/udev/rename_device"
SUBSYSTEM=="net", RUN+="/etc/sysconfig/network-scripts/net.hotplug"
LABEL="skipgls"
```

2. Rebooting

   Run the following command and reboot the system. After rebooting the system, verify eth0 is enabled using "ifconfig" command.

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

3. Setting a subnet mask

   For the underlined parameter, specify the network address and the subnet mask of the virtual interface.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetmask create -i 10.34.214.0 -m 255.255.255.0
   ```

   Check that the facility has been set up correctly.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetmask print
   ```

Note

. . . .

For details on the subnet mask value, see "hanetmask command" in the "PRIMECLUSTER Global Link Services Configuration and Administration Guide : Redundant Line Control Function."

. . . .

4. 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
   eth0,eth1
   ```

   Check that the virtual interface has been set up correctly.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetconfig print
   ```

5. 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
   ```

6. Setting up the Standby patrol monitoring function

   ```
   # /opt/FJSVhanet/usr/sbin/hanetconfig create -n sha1 -m p -t sha0
   ```

   Check that the facility has been set up correctly.

   ```
   # /opt/FJSVhanet/usr/sbin/hanetconfig print
   ```

7. 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/hosts file. Defined content is same as HOST-primecl01.

    2. Configure physical interface definition file as follows.

        - Contents of /etc/sysconfig/network-scripts/ifcfg-eth0

        ```
        DEVICE=eth0
        BOOTPROTO=static
        HWADDR=XX:XX:XX:XX:XX:XX
        HOTPLUG=no
        BROADCAST=10.34.214.255
        IPADDR=10.34.214.182
        NETMASK=255.255.255.0
        NETWORK=10.34.214.0
        ONBOOT=yes
        TYPE=Ethernet
        ```

        - Contents of /etc/sysconfig/network-scripts/ifcfg-eth1

        ```
        DEVICE=eth1
        BOOTPROTO=static
        HWADDR=XX:XX:XX:XX:XX:XX
        HOTPLUG=no
        ONBOOT=yes
        TYPE=Ethernet
        ```

    3. When the system is RHEL, on the /etc/sysconfig/network file, define a statement which enables the network configuration.

        ```
        NETWORKING=yes
        NETWORKING_IPV6=no
        ```

    4. Setting Hot Plug

        Edit the /etc/udev/rules.d/60-net.rules file for RHEL5 environments, and nullify the Hot Plug to virtual interface (sha). The setting of this file is unnecessary in RHEL6 environments.

        ```
        SUBSYSTEM=="net", ENV{INTERFACE}=="sha*", GOTO="skipgls"
        ACTION=="add", SUBSYSTEM=="net", IMPORT{program}="/lib/udev/rename_device"
        SUBSYSTEM=="net", RUN+="/etc/sysconfig/network-scripts/net.hotplug"
        LABEL="skipgls"
        ```

2. Rebooting

    Run the following command and reboot the system. After rebooting the system, verify eth0 is enabled using "ifconfig" command.

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

3. Setting a subnet mask

    For the underlined parameter, specify the network address and the subnet mask of the virtual interface.

    ```
    # /opt/FJSVhanet/usr/sbin/hanetmask create -i 10.34.214.0 -m 255.255.255.0
    ```

    Check that the facility has been set up correctly.

    ```
    # /opt/FJSVhanet/usr/sbin/hanetmask print
    ```

📎 **Note**

. . . .

For details on the subnet mask value, see "hanetmask command" in the "PRIMECLUSTER Global Link Services Configuration and Administration Guide : Redundant Line Control Function."

. . . .

4. 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
eth0,eth1
```

Check that the virtual interface has been set up correctly.

```
# /opt/FJSVhanet/usr/sbin/hanetconfig print
```

5. Setting up the HUB monitoring function

In 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
```

6. Setting up the Standby patrol monitoring function

```
# /opt/FJSVhanet/usr/sbin/hanetconfig create -n sha1 -m p -t sha0
```

Check that the facility has been set up correctly.

```
# /opt/FJSVhanet/usr/sbin/hanetconfig print
```

7. 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 "6.7.3.5 Setting Up Gls Resources" and "6.7 Setting Up Cluster Applications."

You also need to reboot the system after registering the cluster application. Then, start RMS and check the RMS tree to confirm whether the Gls resources are displayed correctly. For details, see "7.1.3.1 RMS Tree."

The Gls resource name is displayed as GlsX (X is integer).

 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*."
....

6.2.2 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*."
....

6.3 GDS Configuration Setup

The GDS (Global Disk Services) setup operations are described below.

- "6.3.1 Setting Up System Disk Mirroring"

  Set up system disk mirroring in both nodes, and then restart both nodes.

- "6.3.2 Setting Up Shared Disks"

  Set up the shared disk volumes.

Note

- If you are using the shared disk device, you must use GDS to manage that unit.

- Execute the configuration setting of GDS after initializing the cluster.

- When using EC or REC function of an ETERNUS disk array without using PRIMECLUSTER GDS Snapshot, do not register the GDS class to the cluster application if the copy destination disk of EC or REC is registered in the class.
  When EC and REC are being synchronized or their equivalence is maintained, the access from the server to the destination disk may fail. Therefore, if the class in which the copy destination disk is registered is added to the cluster application, the access from the server to the copy destination disk fails. In this case, the cluster application fails over.

6.3.1 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 "System Disk Mirror Setup Worksheet". In the operation procedure, the "A.9 GDS Setup Worksheets" is abbreviated as "worksheet".

See

For setup details, see "5.2.1 System Disk Settings [EFI]" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*."

Note

When using the virtual machine function to perform the system disk mirroring in the guest OS by GDS, you need to set up mirror volumes of the local class that were created in the host OS to the guest OS. For information on how to set up the host OS, see "1. Setting up disks and related devices" in "3.2.1.1 Host OS setup."

For details on settings, see 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.
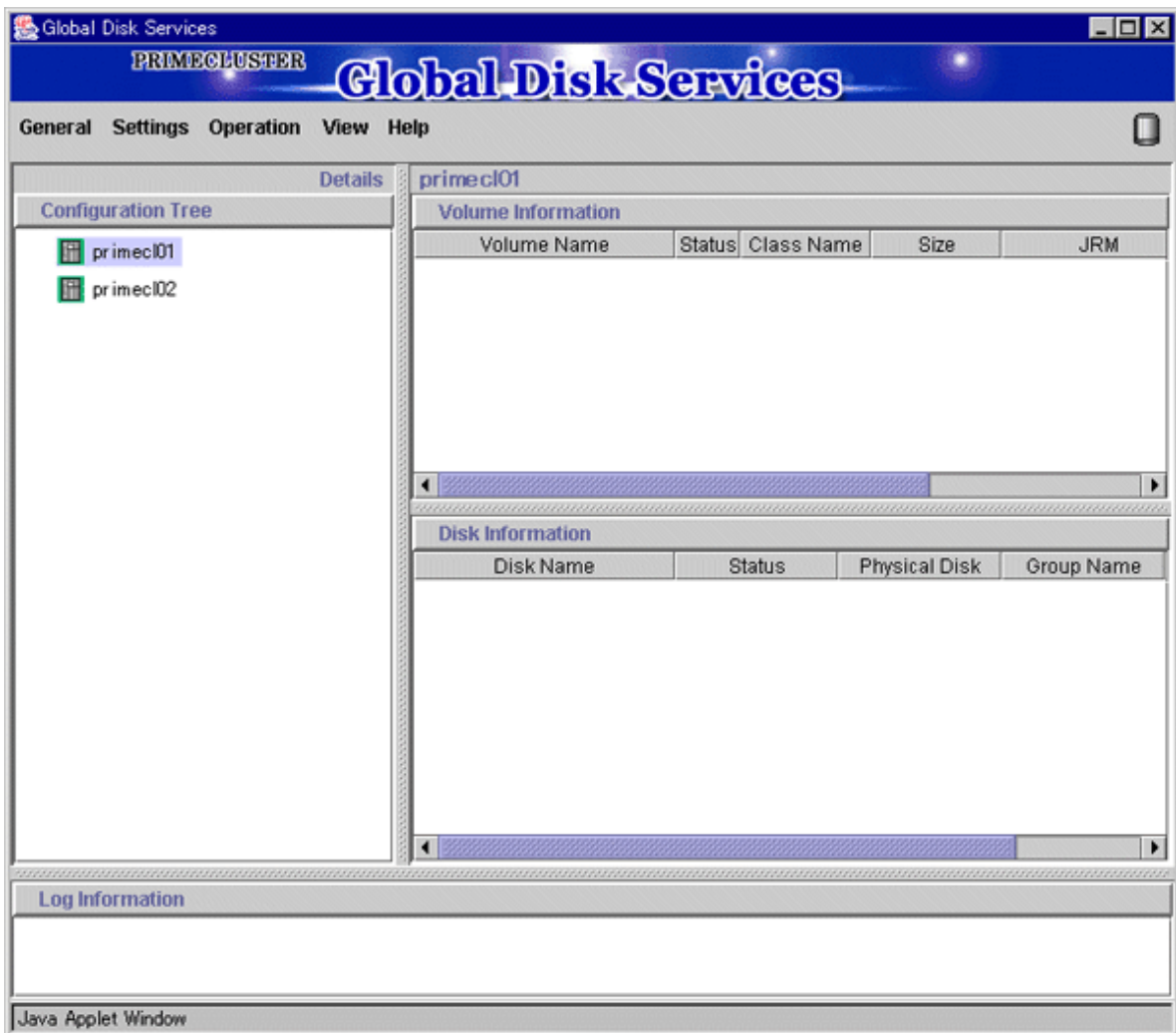
2. From the GDS configuration tree, select the node in which the system disk mirror is to be set, click the Settings menu, and select System Disk Settings.

   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.
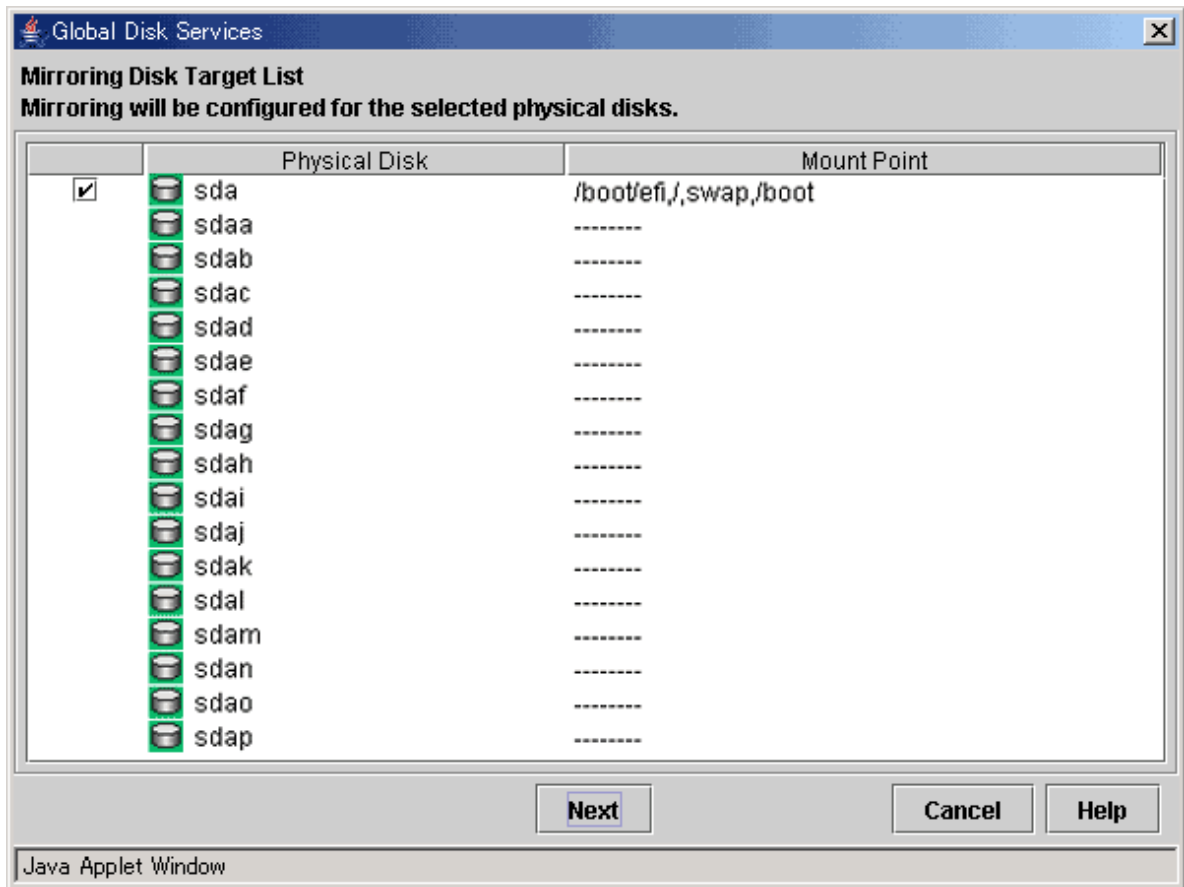
3. Specify class name of the root class.

Enter the class name ("Class name" on the worksheet) of the root class, and click Next.
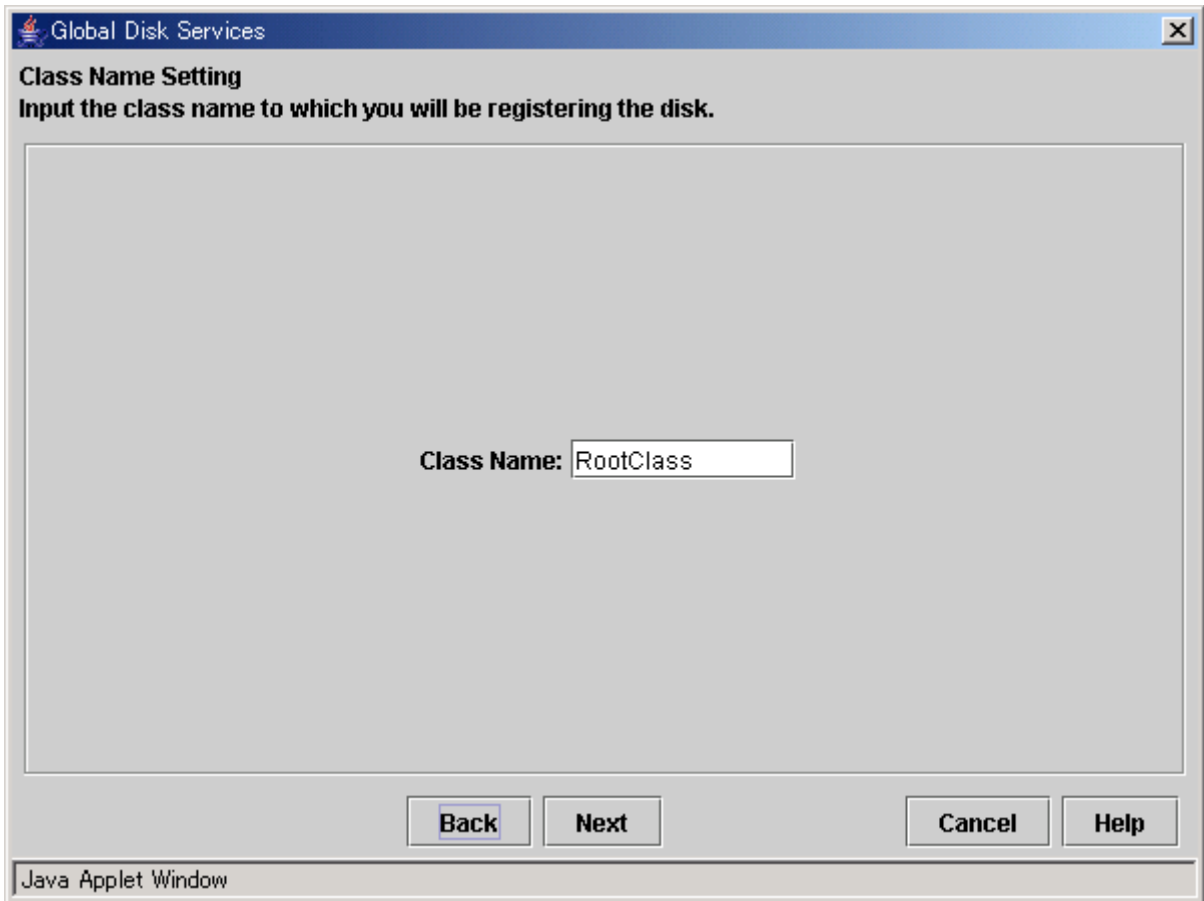
Note

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.
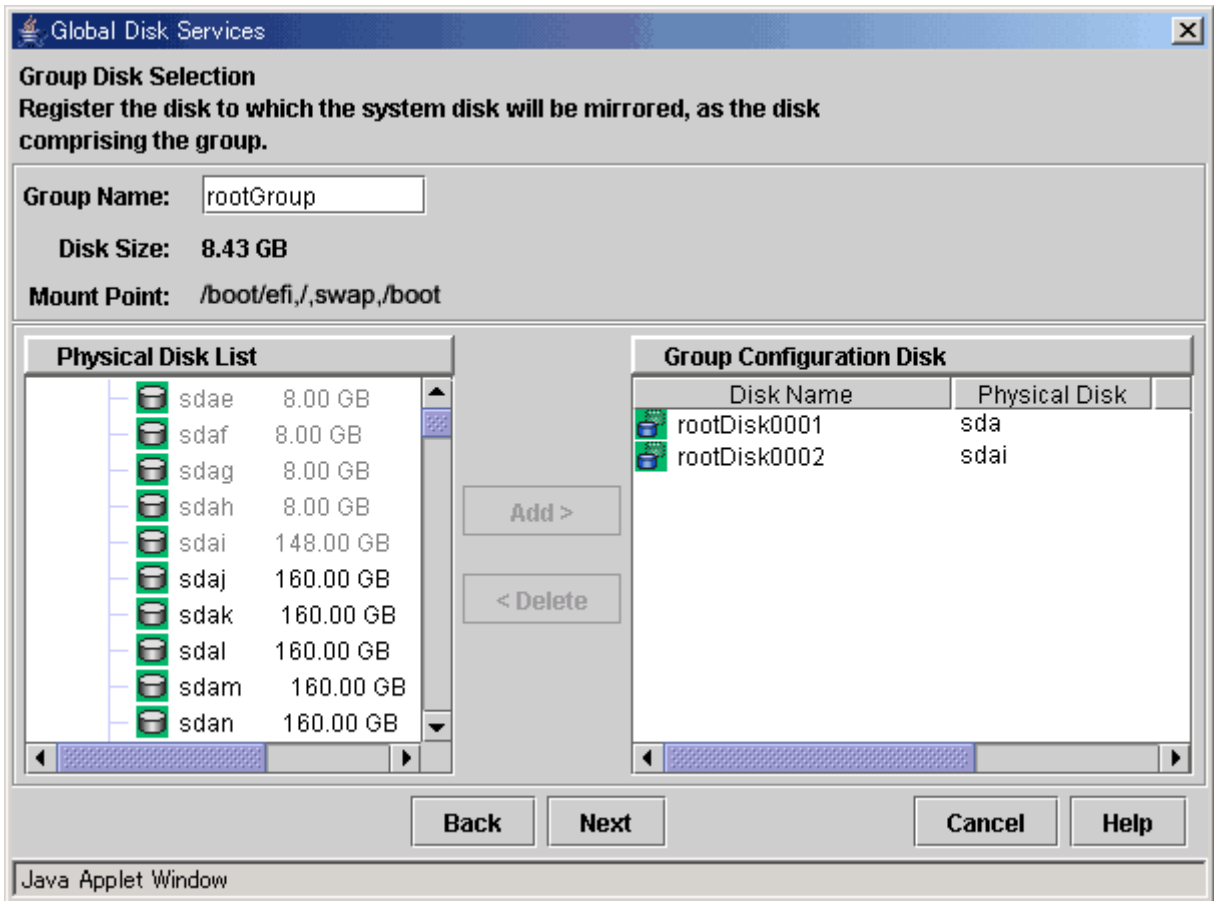
5. Select a Spare Disk ("Spare disk name" on the worksheet) from the "Physical Disk List," and click Add.

   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.

Global Disk Services

Spare Disk Selection
Register the spare disk. - 122 -

Class Name: RootClass

Size: 8.43 GB(Minimum:8.43 GB)

Physical Disk List

- linux04
  - sd
    - sda   34.17 GB
    - sdaa   8.00 GB
    - sdab   8.00 GB
    - sdac   8.00 GB
    - sdad   8.00 GB
    - sdae   8.00 GB
    - sdaf   8.00 GB
    - sdag   8.00 GB
    - sdah   8.00 GB
    - sdai   148.00 GB
    - sdaj   160.00 GB

Add >

< Delete

Spare Disk

Disk Name	Physical Disk

Back   Next   Cancel   Help

Java Applet Window
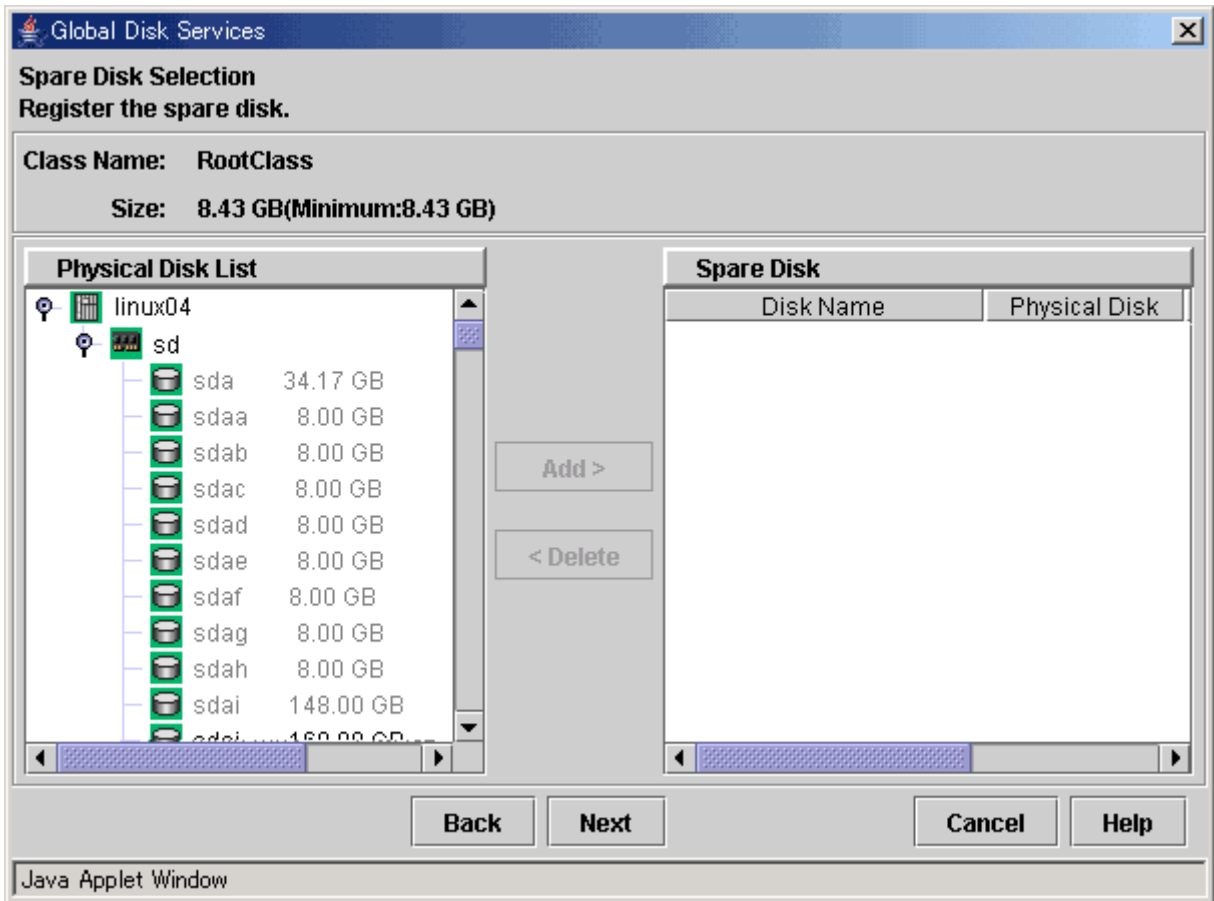
6. Check the system disk configuration.

Check the physical disk name and the mirror disk name, and then click Create.
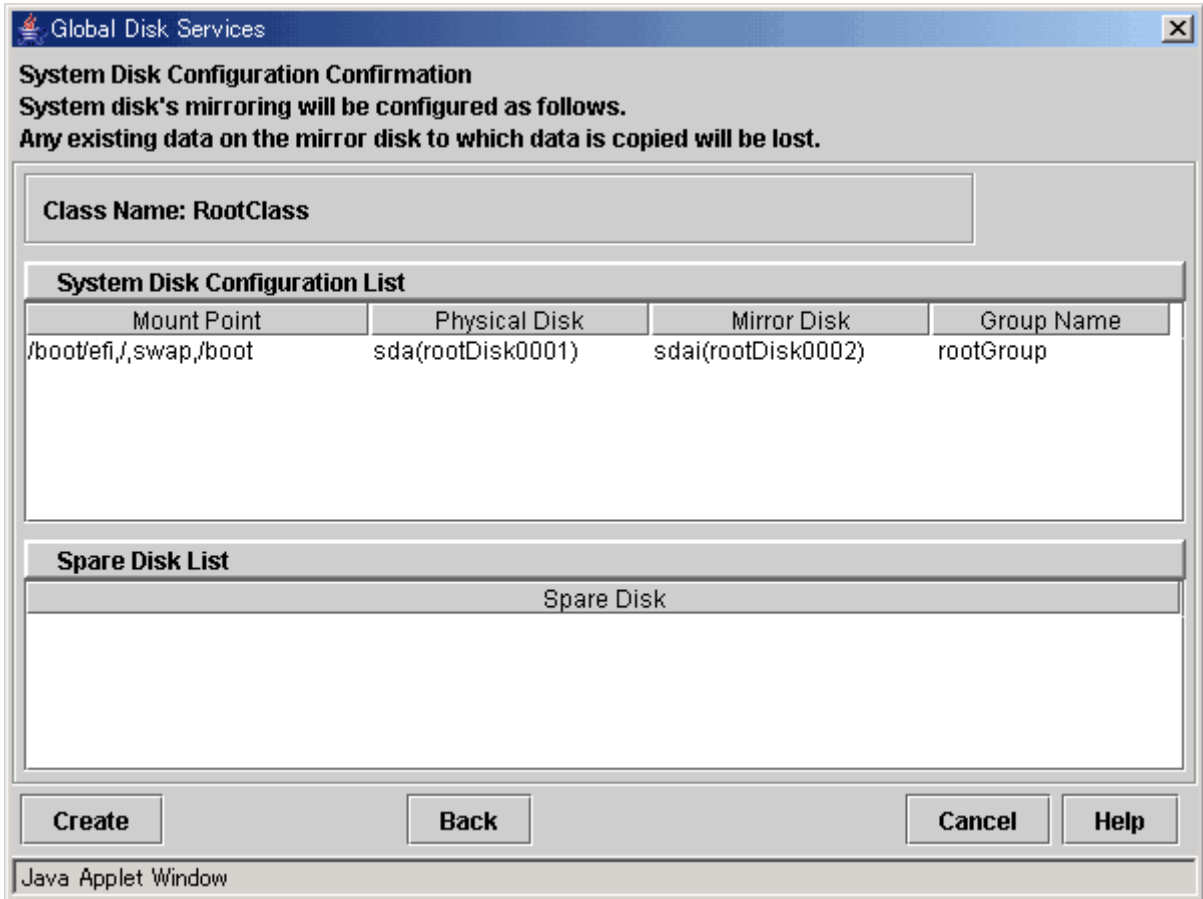
After creation of the system disk is completed, the following screen is displayed.

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.
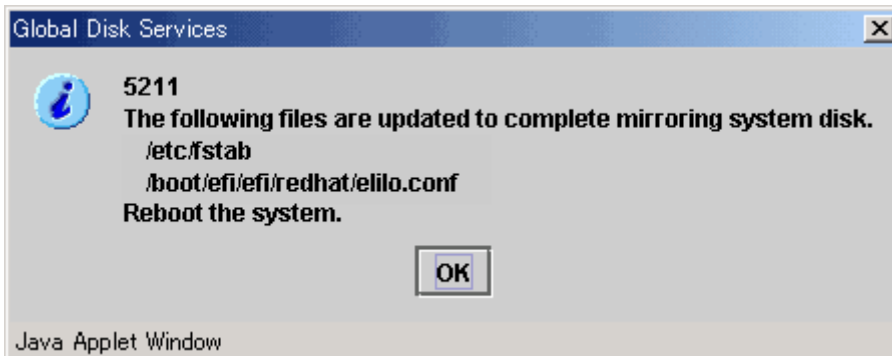
6.3.2 Setting Up Shared Disks

Set up the shared disk as described below:

- **When the GFS Shared File System is to be used**

    1. Execute "Volume setup."

    2. Execute "6.4 Initial GFS Setup."

- **When the ext3 file system is to be used**

    1. Execute "Volume setup."

    2. Execute "File system setup."

3. Create a Gds resource and register it to a cluster application.
   For details, see "6.7.3.4 Setting Up Gds Resources" and "6.7 Setting Up Cluster Applications."

- **When the file system is not to be used**

  1. Execute "Volume setup."

  2. For use as a RAW device, make the settings according to the manual pages related to the "raw(8)" command and rawdevices.

  3. Create a Gds resource and register it to a cluster application.
     For details, see "6.7.3.4 Setting Up Gds Resources" and "6.7 Setting Up Cluster Applications."

Note

- "When the GFS Shared File System is to be used" and "When the file system is not to be used," "File system setup" is not necessary.

- The setup procedures for "When the ext3 file system is to be used" and "When the file system is not to be used" must be carried out before the Gds resources are set up. For details, see "6.7.3.3 Preliminary Setup for Gds Resources."

- "When the GFS Shared File System is to be used", "6.7.3.4 Setting Up Gds Resources" must not be carried out.

- The local class disks or shared class disks used by GDS on the guest OS should be configured as the following virtual disks if they are used in the virtual machine environment.

    - Xen environment
      Virtual SCSI devices

    - KVM environment
      virtio block devices

Volume setup

There are four types of volumes:

  a. Single volume

  b. Mirror volume

  c. Stripe 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" of "A.9.2 Shared Disk Setup 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.

- When neither the system nor the GDS Management screen are reactivated after "2. Registering a shared disk" of "5.1.3.2 Registering Hardware Devices", the registered shared disk might not be correctly recognized to GDS. In this case, setup the volume after updating physical disk information. Physical disk information can be updated by selecting *Update Physical Disk Information* from *Operation* menu of the main screen.

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.
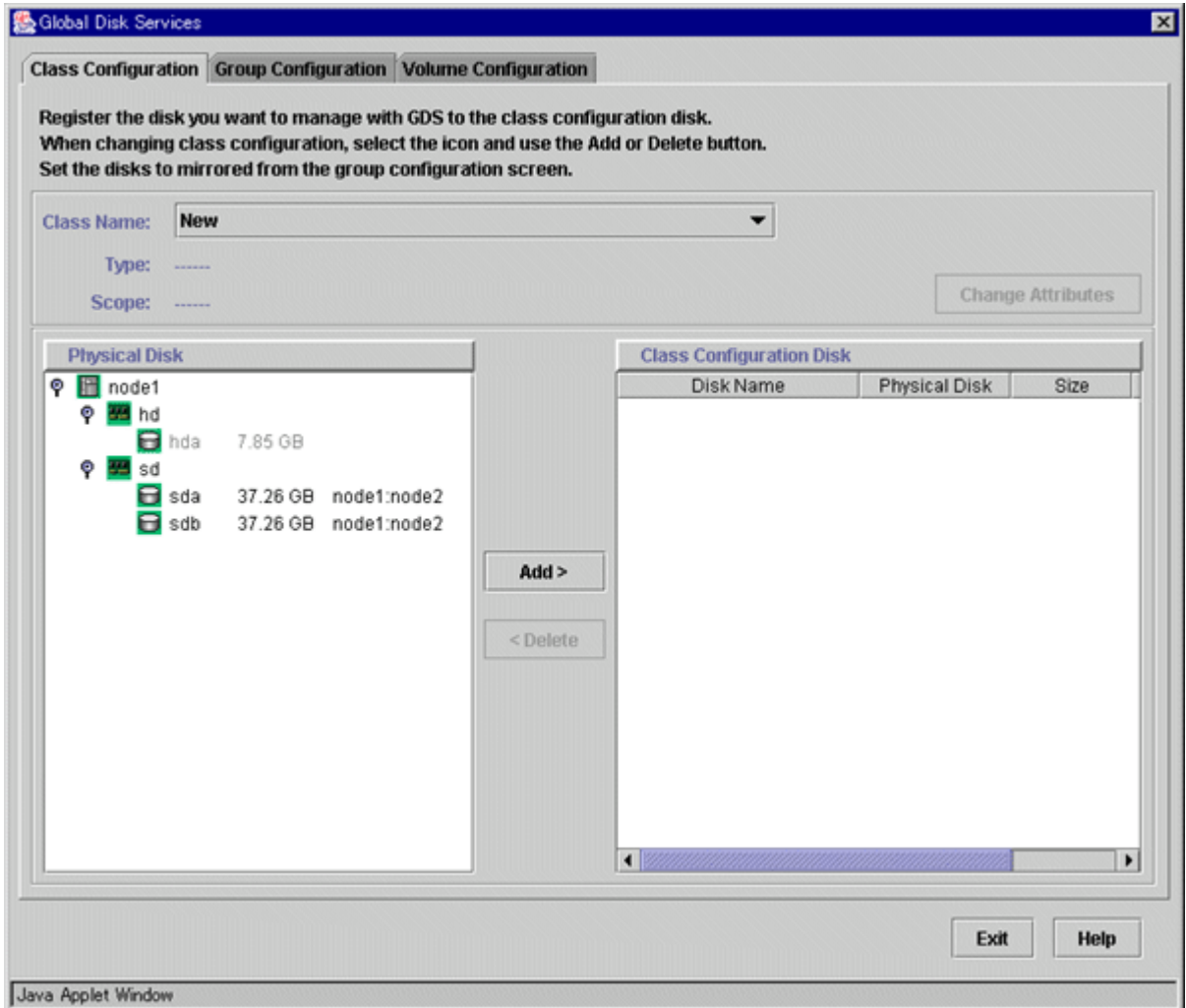
   At the above screen, select the physical disk to be registered from the *Physical Disk* list, and then click *Add*. When *Add* is clicked, the Class Attributes Definition screen opens. Enter the *Class Name* but do not change the *Type* value (leave the value as "shared").

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.
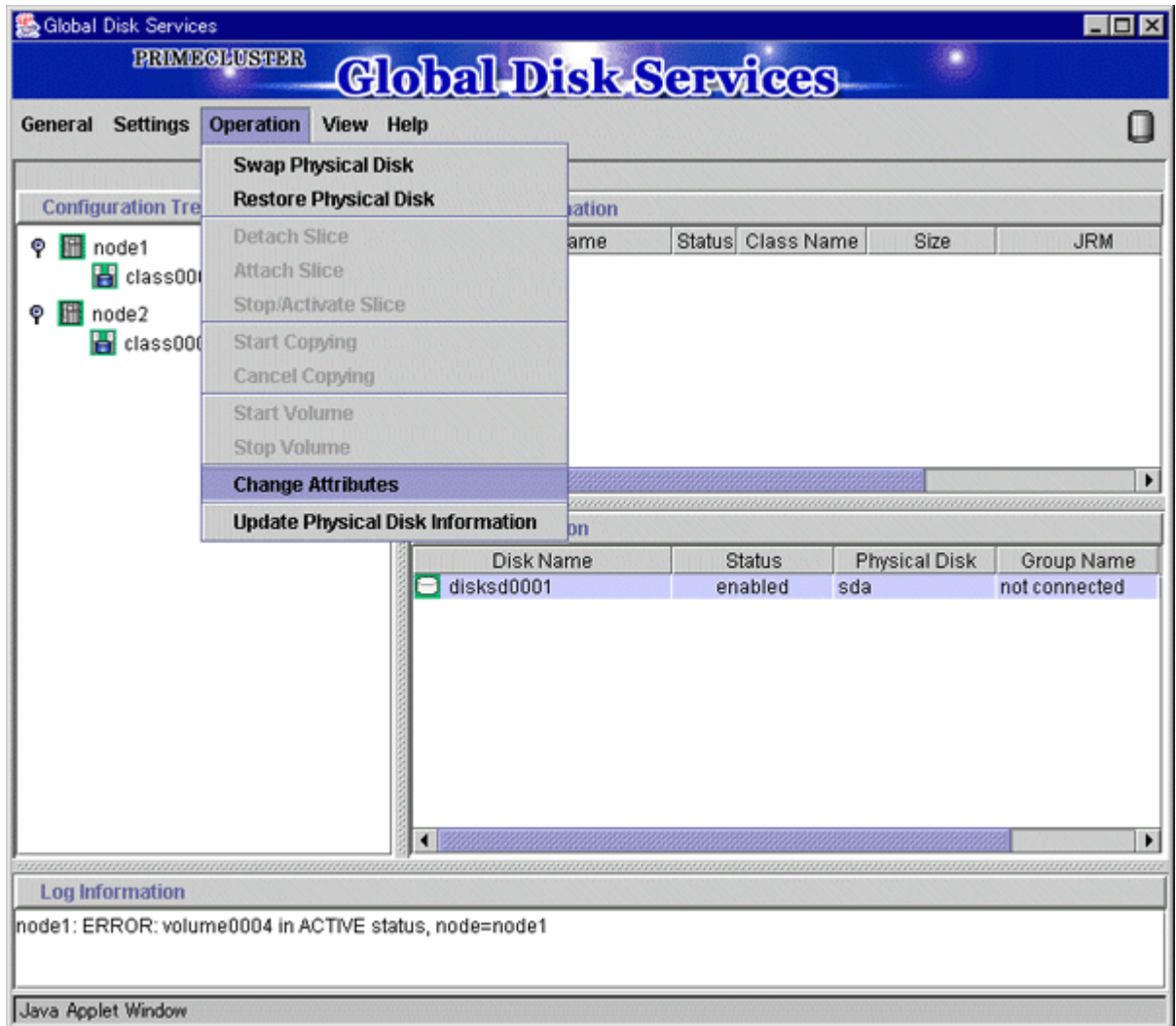
Set *Disk Type* to "single," and then click *OK*.
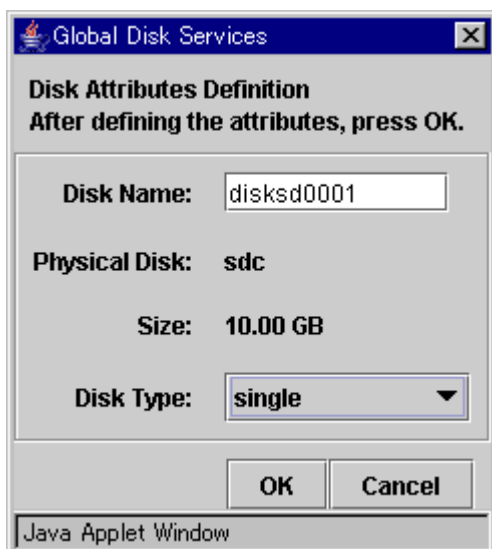
4. Volume creation

   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*.
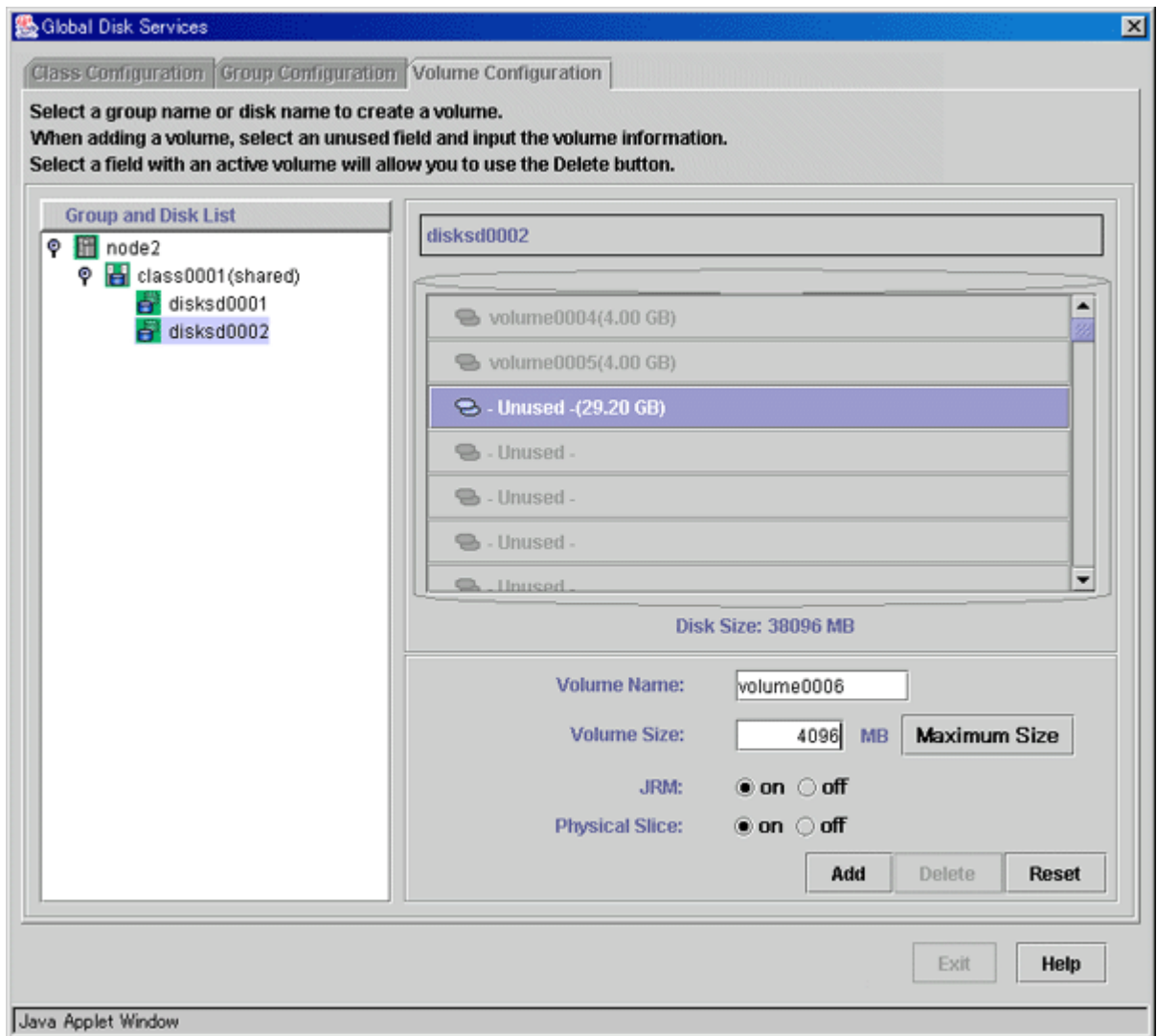
Setup for other volumes

If you are using only "single" volumes, this setup is unnecessary.

Operation Procedure:

1. Creating a disk class and registering physical disks

   At the main screen, select *Class Configuration* from the *Settings* menu.
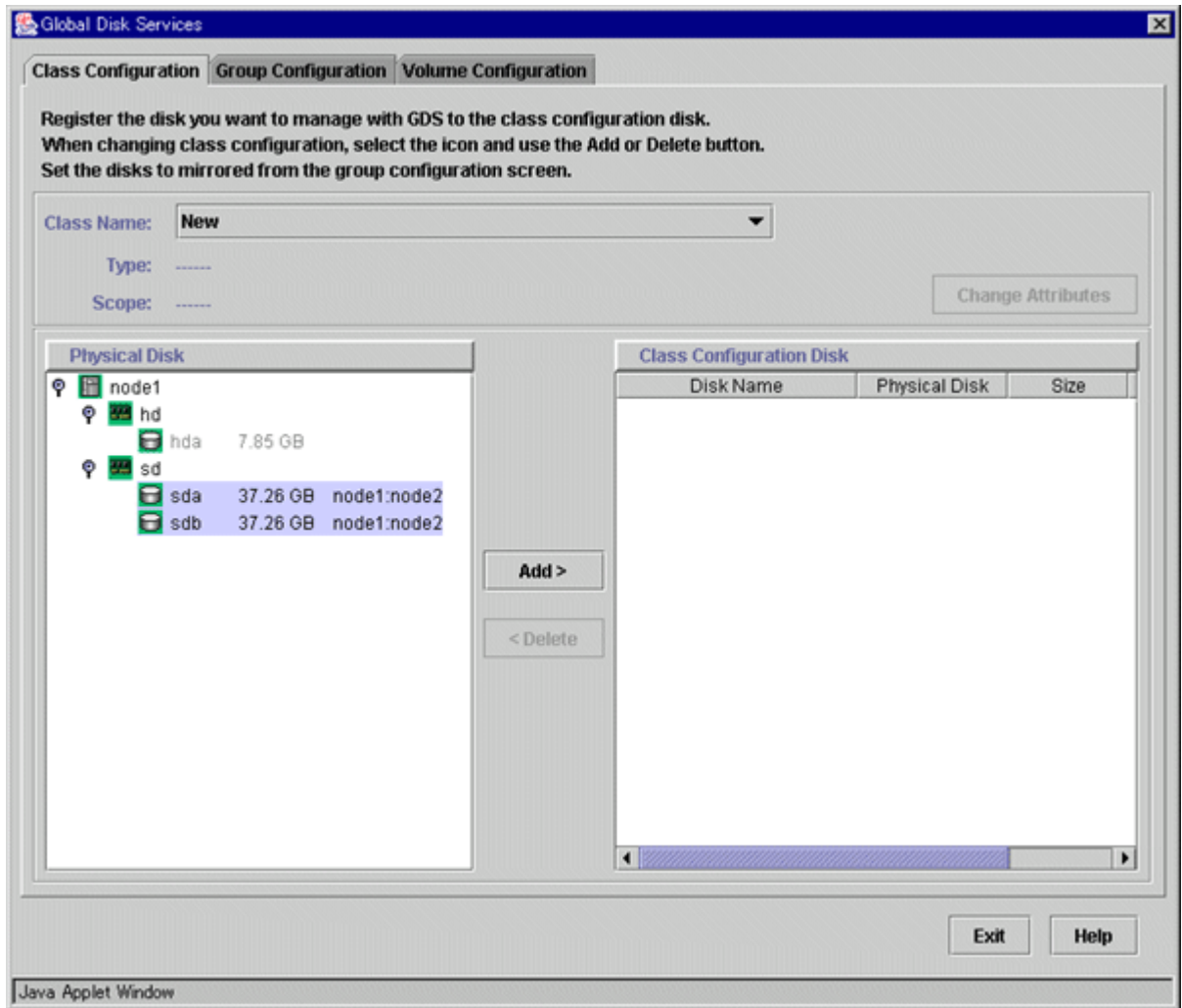
At the above screen, select the physical disks to be registered from the *Physical Disk* list, and then click *Add*. When *Add* is clicked, the Class Attributes Definition screen opens. Enter the *Class Name* but do not change the *Type* value (leave the value as "shared"). Then click *Exit*.

2. Setting up the disk group configuration

    Click the Group Configuration tab.
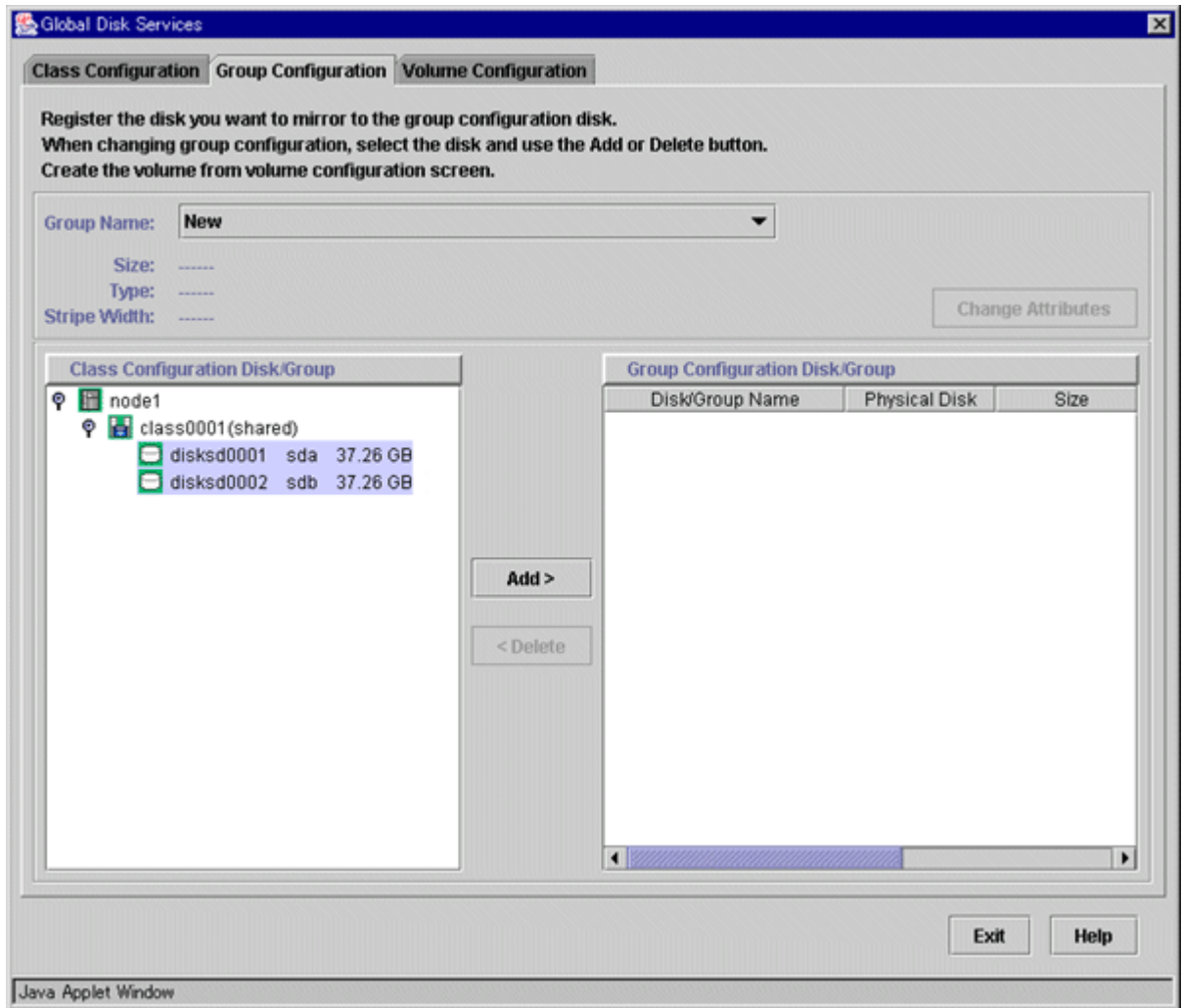
At the above screen, select the disks to be added to the group from the *Class Configuration Disk/Group* list, and then click *Add*.

Enter the *Group Name* and *Type* in the Group Attributes Definition screen, and then click *OK*.
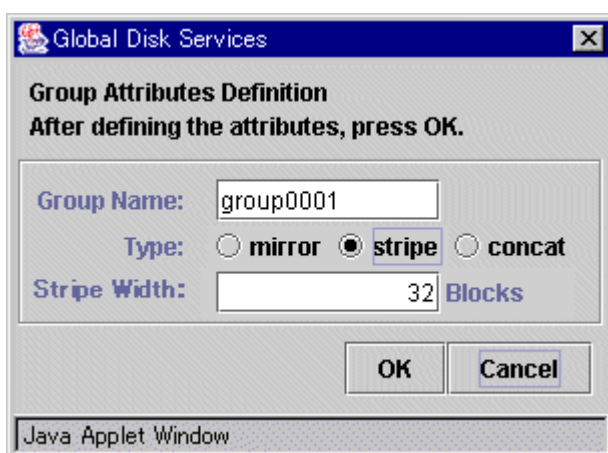
3. Creating a volume

   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*.
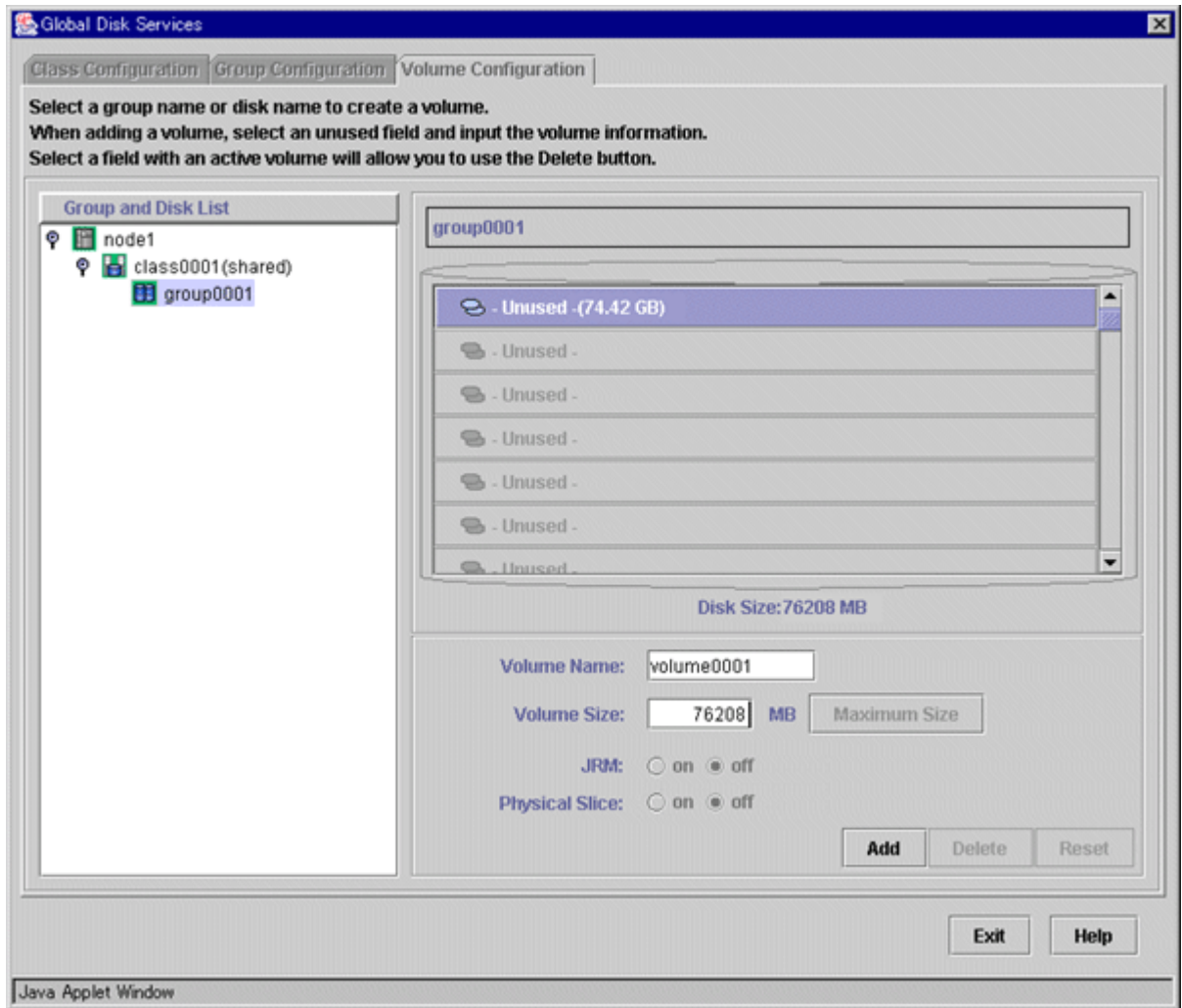
4. Checking the configuration

The disk configuration is displayed as shown below.
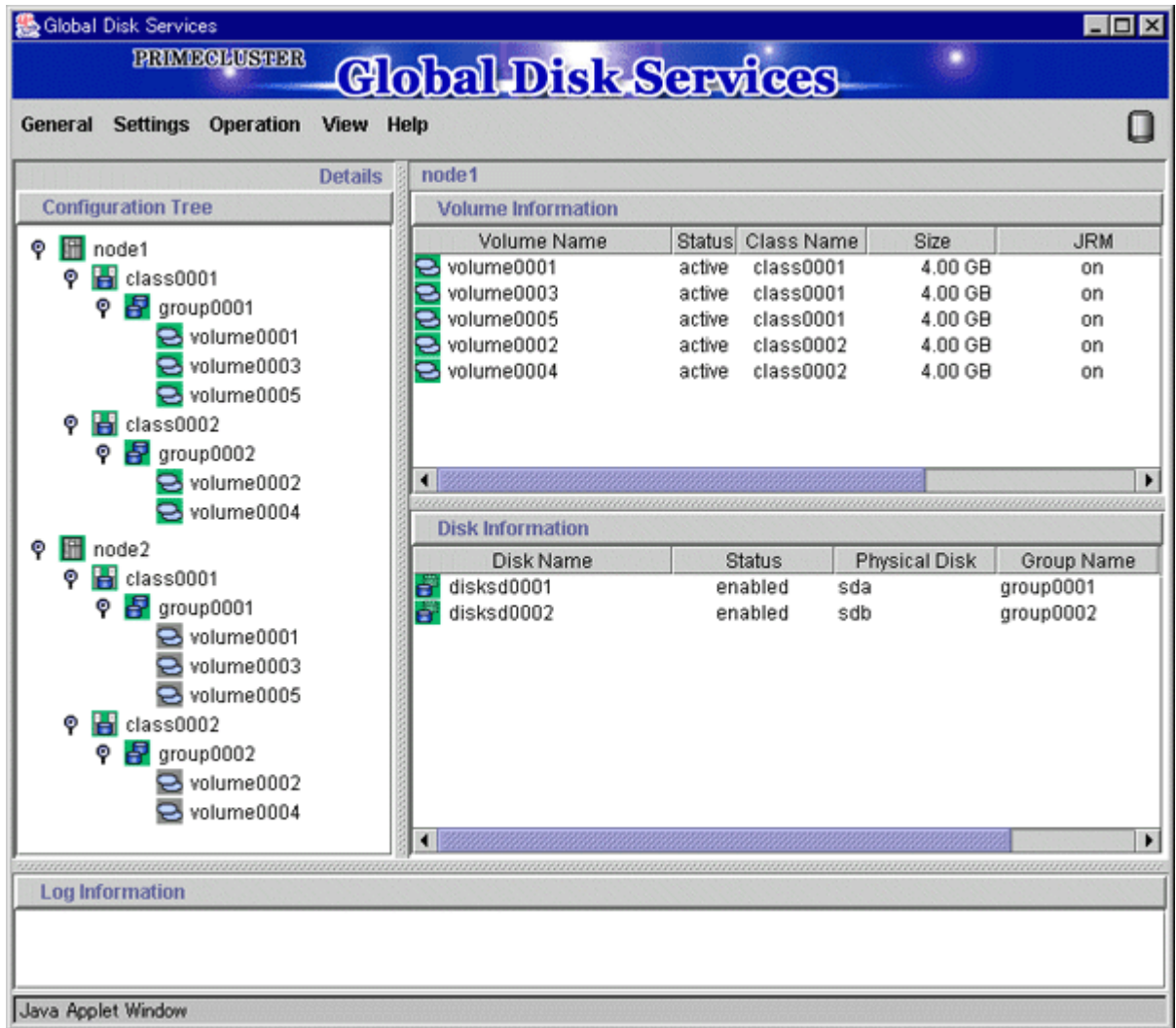
File system setup

You can set up a file system on the File System Configuration screen for the created volume. For Red Hat Enterprise Linux, you can set up the "ext3" file system.

🖐 **Note**

If you set up a file system according to this procedure, an entry of the created file system is added to the "/etc/fstab" file. To use the created file system as an Fsystem resource, you must modify the entry that was added to the "/etc/fstab" file. For details, see "6.7.3.2 Setting Up 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*.
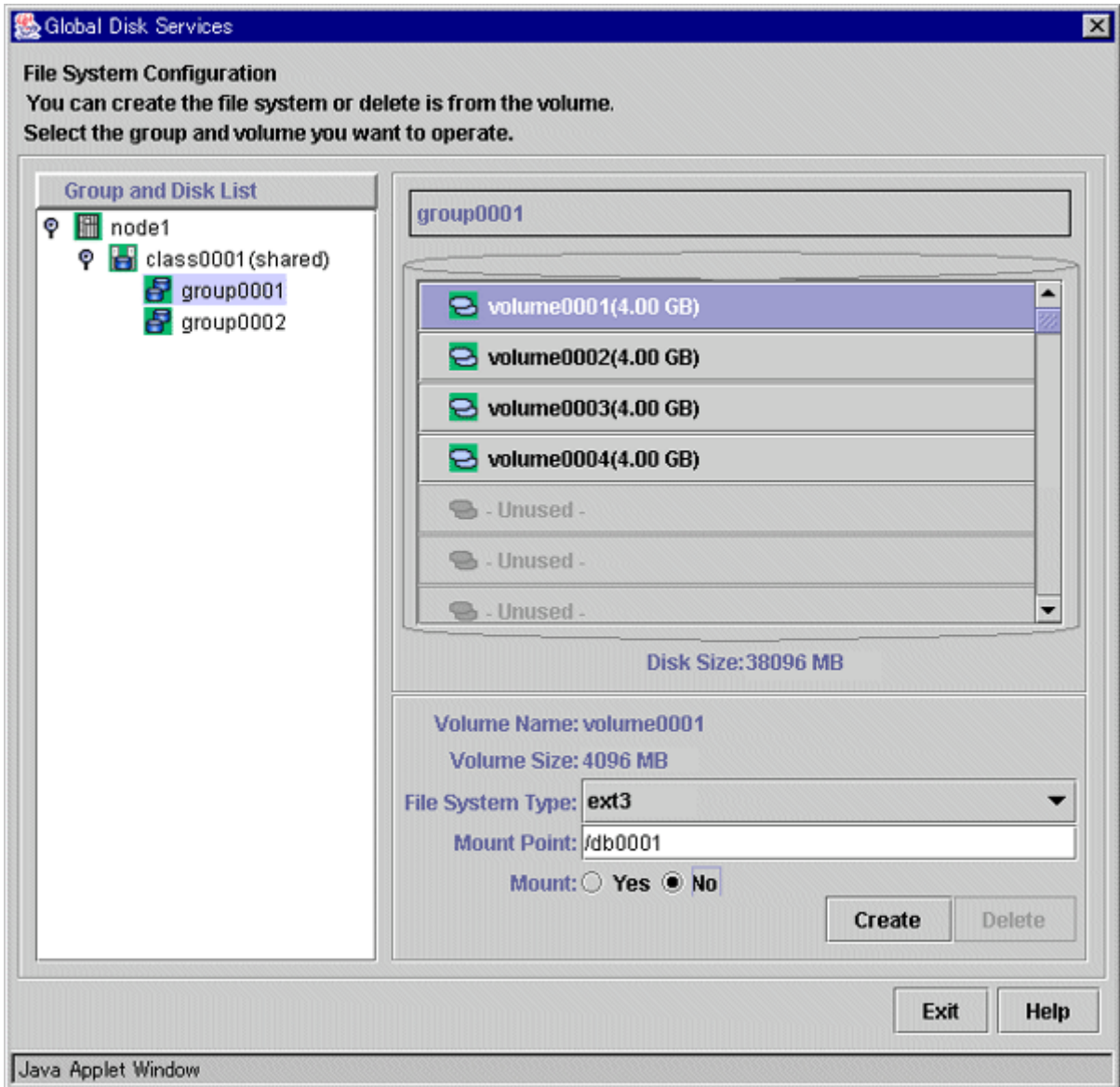
![Note icon] **Note**

....

For the "Mount" setting, do not select "Yes." Also, do not delete "noauto" from the entries that were added to the /etc/fstab file. For details, see "Automatic mounting of the file system" in "A.2 Important Points" of the "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

....

2. To exit the *File System Configuration*, click *Exit*.

6.4 Initial GFS Setup

For a Global File Services (GFS) shared 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.

To use the GFS shared file system in RMS cluster operation, you need to set up GFS according to the flow below:
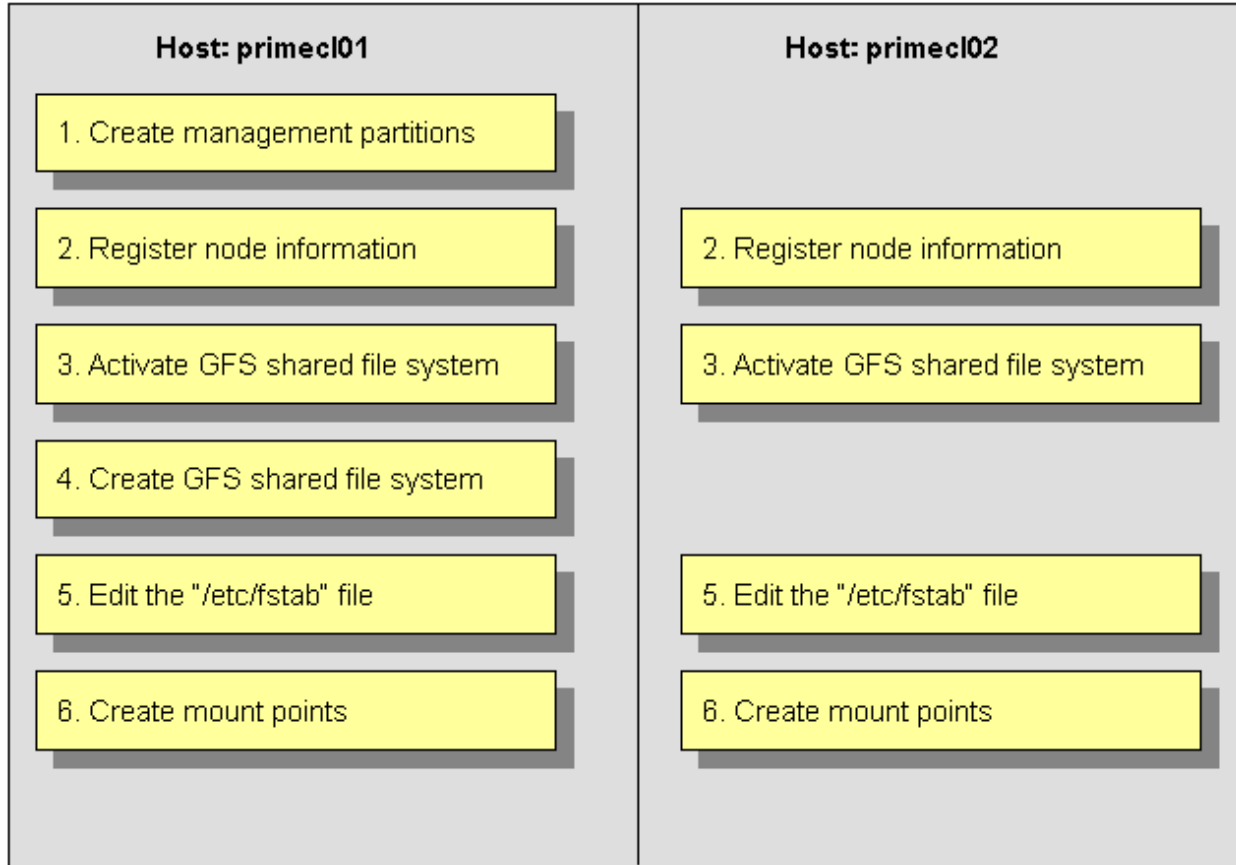
The device name and mount points that are specified here correspond to the values on the "A.8.2 GFS Shared File System Setup Worksheet."

Note
. . . .
- You need to prepare a management partition that is exclusive to the GFS shared file system. The GDS volume disk class is used for a switching file system and non-switching file system. For the management partition, non-switching file system must be allocated.

- If you are using a GFS shared file system, you must not carry out "6.7.3.4 Setting Up Gds Resources."
. . . .

Operation Procedure:

1. Create a management partition for the GFS shared file system on any one of the nodes.

```
# sfcsetup -c /dev/sfdsk/class0001/dsk/GFSctl
```

2. Register the node information in the management partition on each node.

```
primecl01# sfcsetup -a /dev/sfdsk/class0001/dsk/GFSctl
primecl02# sfcsetup -a /dev/sfdsk/class0001/dsk/GFSctl
```

3. Activate the GFS shared file system on each node.

```
primecl01# sfcfrmstart
primecl02# sfcfrmstart
```

![Note icon] Note

....

If sfcfrmstart ends abnormally, confirm that sfcprmd is started with the "ps" command. If sfcfrmstart has not been started, execute the following command on the node on which sfcprmd is not started:

```
# service sfcfsrm start
```

....

4. Create a GFS shared file system on any one of the nodes.

```
# sfcmkfs -o node=primecl01,primecl02 /dev/sfdsk/class0002/dsk/volume0001
```

5. Add the mount information of the GFS shared file system to /etc/fstab on each node. Specify "noauto" in the "mount options" filed of the mount information. Do not specify "noatrc" in the same field.

```
/dev/sfdsk/class0002/dsk/volume0001 /sfcfs1 sfcfs rw,noauto 0 0
```

6. Create a mount point on each node.

```
primecl01# mkdir /sfcfs1
```

```
primecl02# mkdir /sfcfs1
```

![See icon] See

....

The operations described in procedures 4, 5, and 6 can be set up by using the GUI management view. For details, see "6.4.1 File System Creation."

....

6.4.1 File System Creation

This section explains how to create a file system.

Operation Procedure:

1. Start the GFS management view.

   Choose Global File Services on the Web-Based Admin screen, select a node from the node list, and then display the main screen of Global File Services.

2. Set up shared information.

   Select *Operation* -> *Create* and then open the "Create File System Wizard (1)."

   Set up the node information that is to be shared within the file system and then mount the information using the "Create File System Wizard (1)."
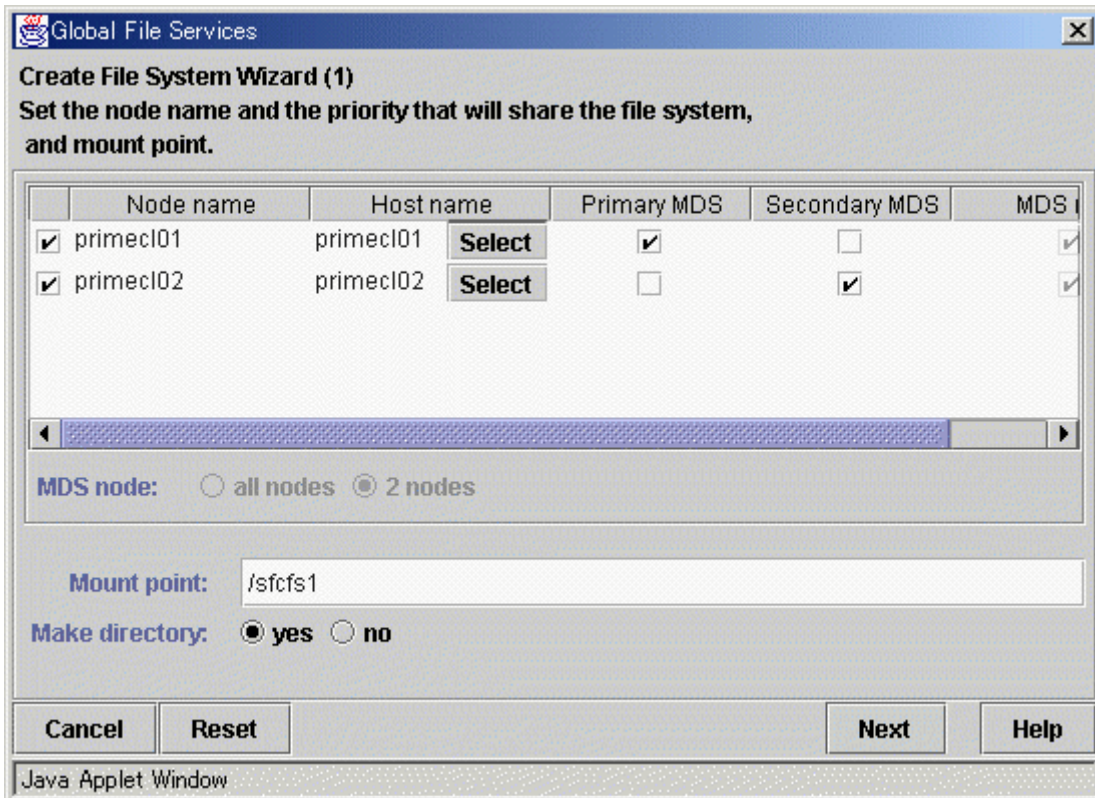
Selecting "Node name"

Select the node names to be shared with "Node Names." You must select two nodes.
Note that the selection of the local node (displayed node) cannot be canceled.

Selecting a "Host name"

To select a host name other than that which is currently displayed, click the *Select* button and specify the host name of the LAN to be used on each node. Note that two or more host names cannot be specified.

Setting the "Primary MDS" and "Secondary MDS"

Specify the nodes that boot the management server of the shared file system in "Primary MDS" and "Secondary MDS."

Setting the "Mount point" and "Make directory"

Specify the full path for the "Mount point." If you select "yes" from "Make directory," creates a directory with the following attributes:

- Owner: root

- Group: sys

- Access authority: 775

After setting or changing this information, click the *Next* button to open the "Create File System Wizard (2)."

To return each setup item to its default value, click the *Reset* button.

To stop the processing of the file system creation, click the *Cancel* button.

3. Select the configuration partition.

Using the "Create File System Wizard (2)," register the partition that is to be used as the file system.
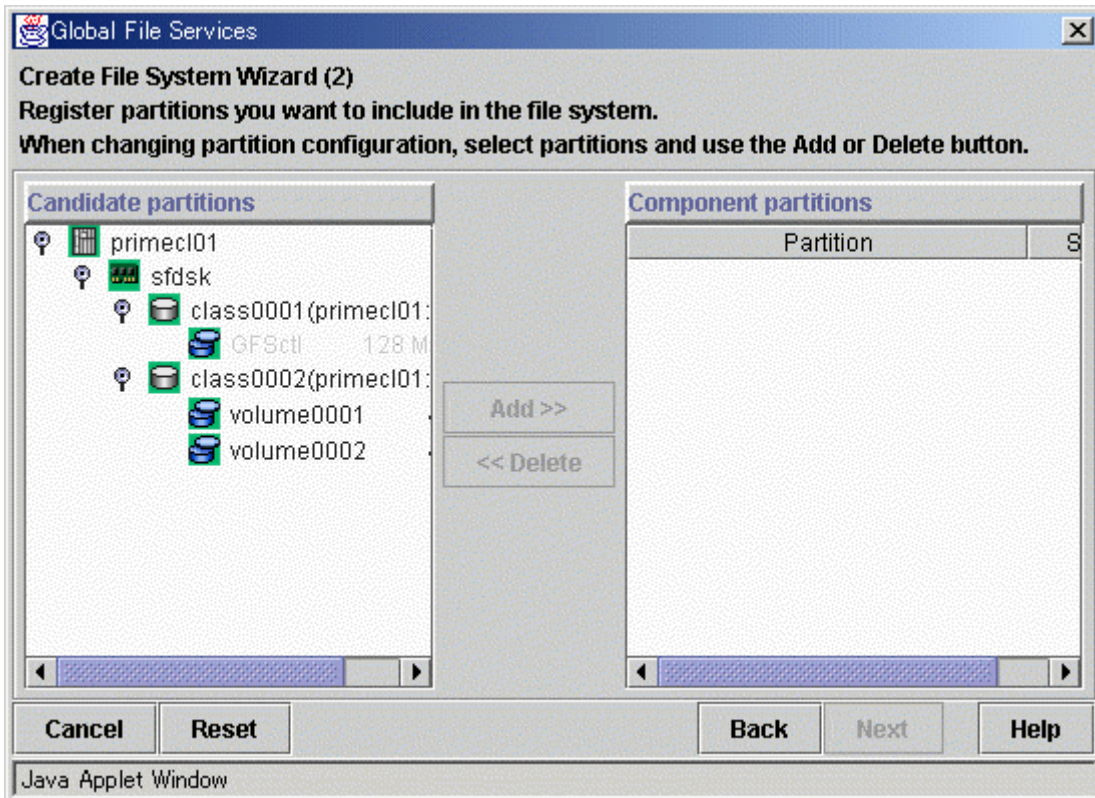
Select the partition to be used from the [*Candidate partitions*] list and then click the *Add* button.

Only one partition can be selected at a time. A partition that is already being used as a file system or as a management partition cannot be selected.

After the partition has been selected, click the *Next* button to open the "Create File System Wizard (3)."

To return to the "Create File System Wizard (1)," click the *Back* button.

To abandon file system creation, click the *Cancel* button.

4. Set up the partition information.

Using the "Create File System Wizard (3)," select the meta, log, and data areas that are to be allocated to each of the partitions selected with the "Create File System Wizard (2)."

The partition to which the meta data area is allocated is used as the representative partition.
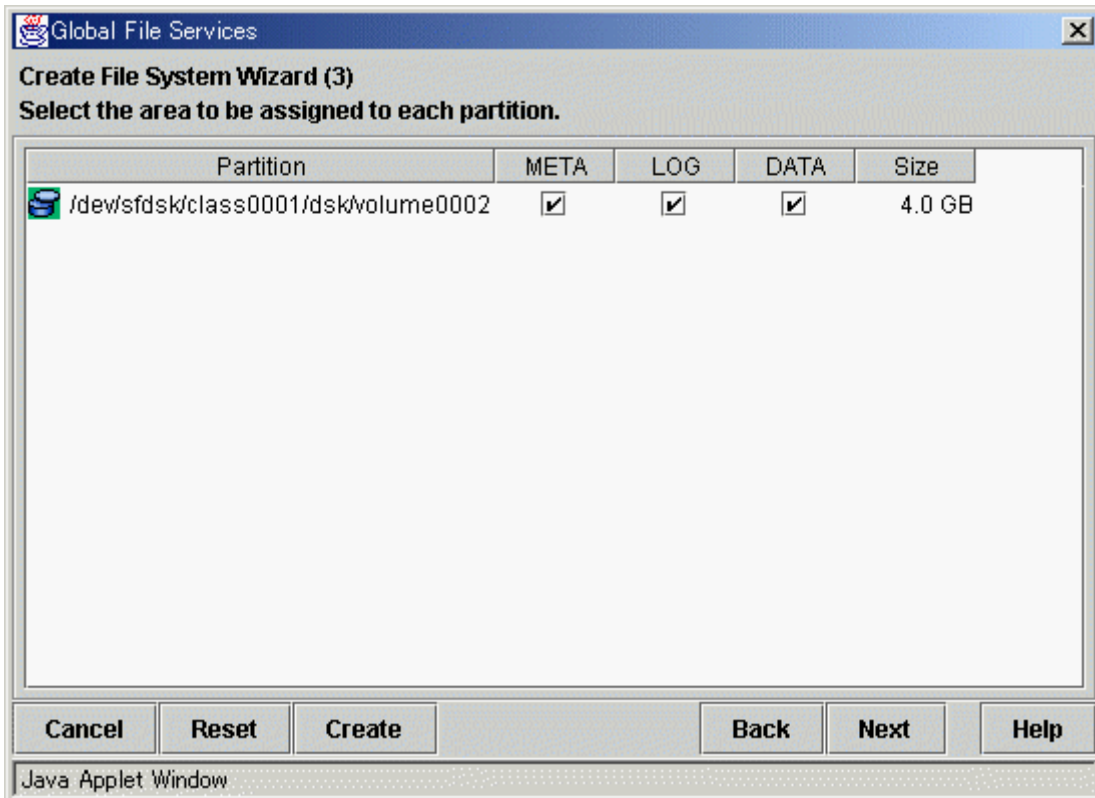
After setting the above information, click the *Next* button to open the "Create File System Wizard (4)."

No information can be set with the "Create File System Wizard (4)." Go to the "Create File System Wizard (5)."

To return each setup item to its default value, click the *Reset* button.

To return to "Create File System Wizard (2)," click the *Back* button.

To abandon file system creation, click the *Cancel* button.

To create the file system while leaving the default settings of the extended, detailed, and mount information as is, click the *Create* button.

5. Set up the detailed information.

Set up the "Detailed information" by using the "Create File System Wizard (5)."
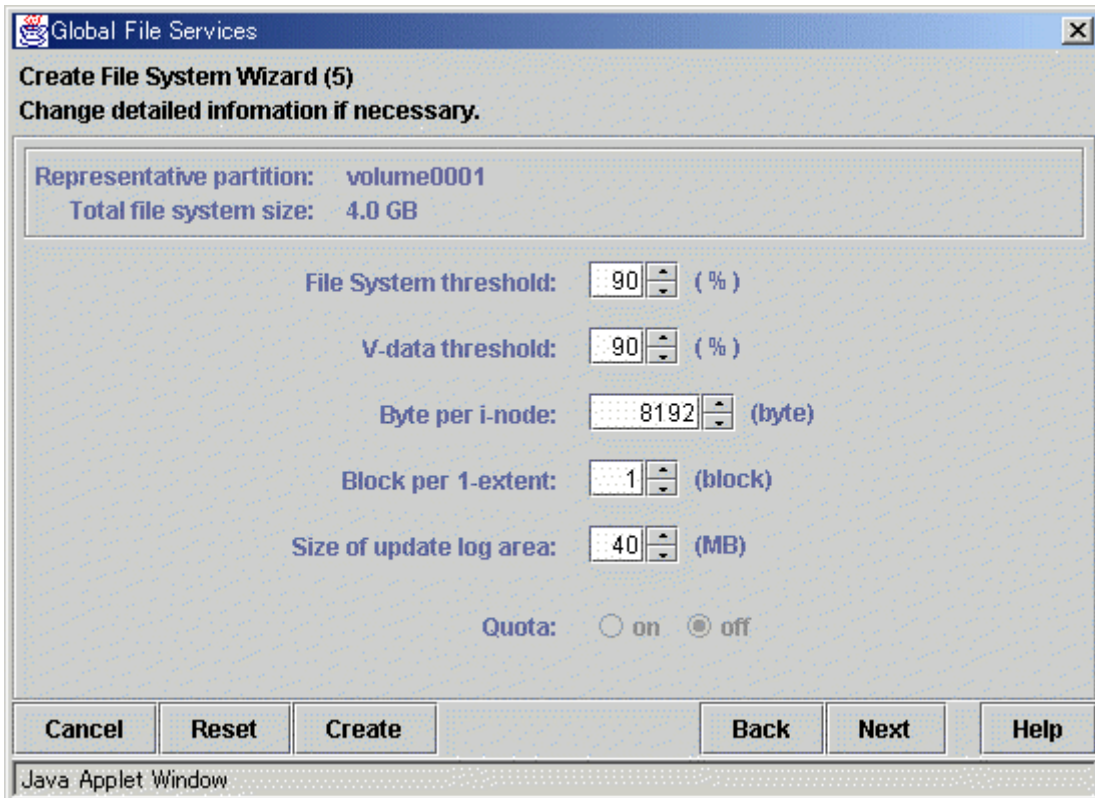
After setting the above information, click the *Next* button to open the "Create File System Wizard (6)."

To return each setup item to its default value, click the *Reset* button.

To return to the "Create File System Wizard (4)," click the *Back* button.

To abandon file system creation, click the *Cancel* button.

To create the file system while leaving the default setting of the mount information as is, click the *Create* button.

6. Set up the mount information.

   Set up the "Mount information" by using the "Create File System Wizard (6)."
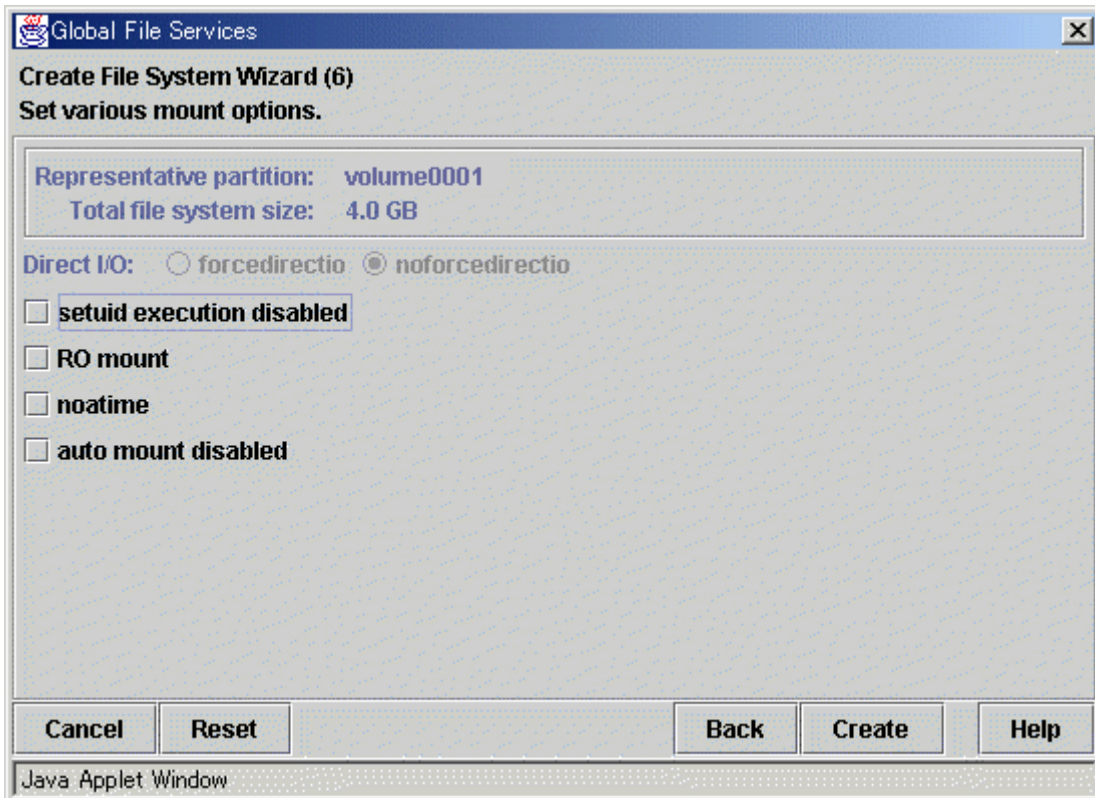
After setting the above information, click the *Create* button to create the file system. To return each setup item to its default value, click the *Reset* button.

To return to the "Create File System Wizard (5)," click the *Back* button.

To abandon file system creation, click the *Cancel* button.

6.5 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 device and takeover networks that were set in this chapter.

See
. . . .
See the manuals for the individual applications.
. . . .

6.6 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 "6.7.3.1 Setting Up 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.

Start script/Stop script

```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
      # Start your application
elif [[ $1 = "-u" ]]; then
      # Stop your application
else
      # Default operation
      exit 1 # Error
fi
exit 0
```

The above script sample covers both the Start script and the Stop script.
An example of Check script is shown below:

Check script

```sh
#!/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 "6.7.3.1 Setting Up 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.

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.

Check script (hot-standby operation)

```sh
#!/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
```

1. Setting attributes for the Cmdline resources

   Enable the STANDBYCAPABLE and the ALLEXITCODES attributes.

   For details, see "6.7.3.1 Setting Up 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 Fault 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 "Change the attribute" of "6.7.3.1 Setting Up 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 Changing the Cluster Configuration."

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	The value of this variable indicates whether the script is triggered by AutoRecover or not (1 or 0). For details on AutoRecover, see "11 Appendix - Attributes" in "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide." 0: Not triggered by AutoRecover. 1: Triggered by AutoRecover.
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. 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.

![Note icon] **Note**

- -

RMS environment variables displayed by the "hvdisp" command (ENV and ENVL option) are the only environment variables that are set when each script (Online/Offline/Check script) is executed.

The environment variables, /etc/profile, set to individual servers (systems) are not inherited.

You must therefore define environment variables that are required for script execution in each script.

- -

6.7 Setting Up Cluster Applications

This section explains how to set up a cluster application.

You can use any one of the nodes of the cluster system for the settings described in this section.

Cluster application setup flow

The setup flow for a cluster application is explained for each topology, below.

For information on these topologies, see "2.3 Determining the Cluster System Operation Mode."

1) 1:1 standby for standby operation

Set up 1:1 standby for standby operation as follows.
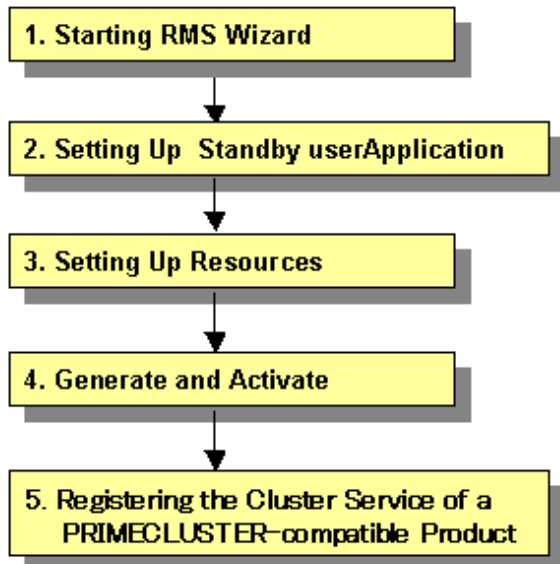
2) Mutual standby for standby operation

Set up mutual standby for standby operation as follows.
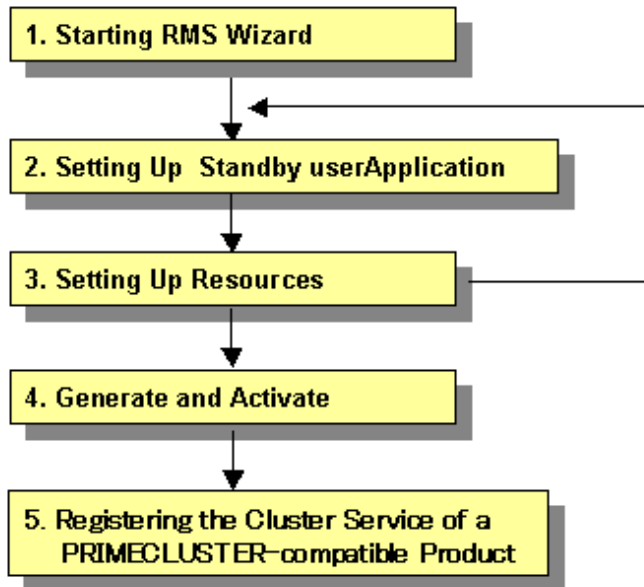
- To create two cluster applications, repeat steps 2. to 3.

3) N:1 standby for standby operation

Set up N:1 standby for standby operation as follows.
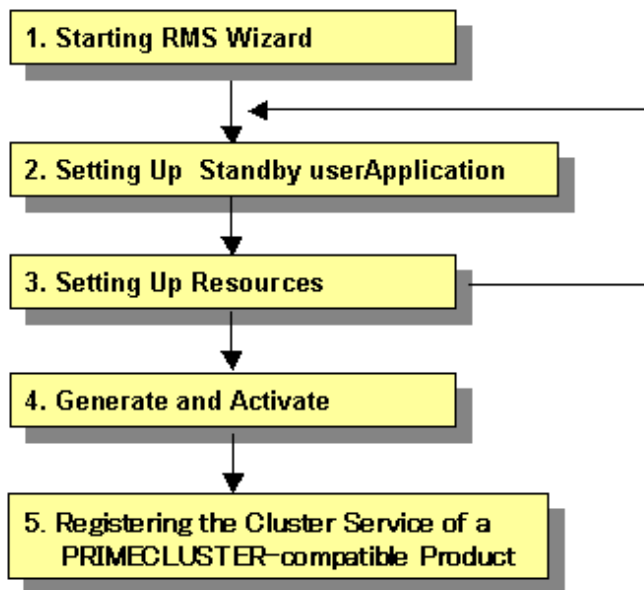
- To create two or more cluster applications, repeat steps 2. and 3.

   Example) For 2:1 standby, repeat steps 2. and 3. two times, to create two cluster applications.

4) Cascaded standby operation

Set up cascaded standby operation as follows.
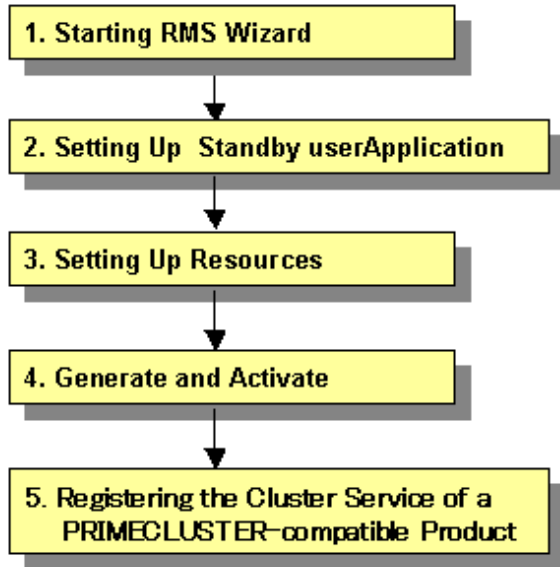
5) Priority transfer of standby operation

Set up the priority transfer of standby operation as follows.
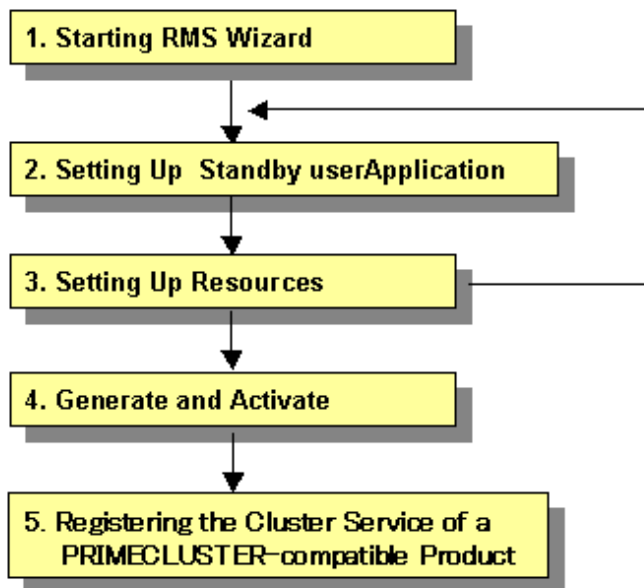
- To create two or more cascade cluster applications, repeat steps 2. and 3.

- Establish an exclusive relationship between the cluster applications.

  For details on exclusive relationships, see "6.7.7 Exclusive Relationships Between Cluster Applications."

  Example) For priority transfer with two cascade cluster applications, repeat steps 2. and 3. two times to create two cascade cluster applications.

6) Scalable operation

Set up a scalable operation as follows.

1. Starting RMS Wizard
2. Setting Up Standby userApplication
3. Setting Up Resources
4. Setting Up Scalable userApplication
5. Generate and Activate
6. Registering the Cluster Service of a PRIMECLUSTER-compatible Product
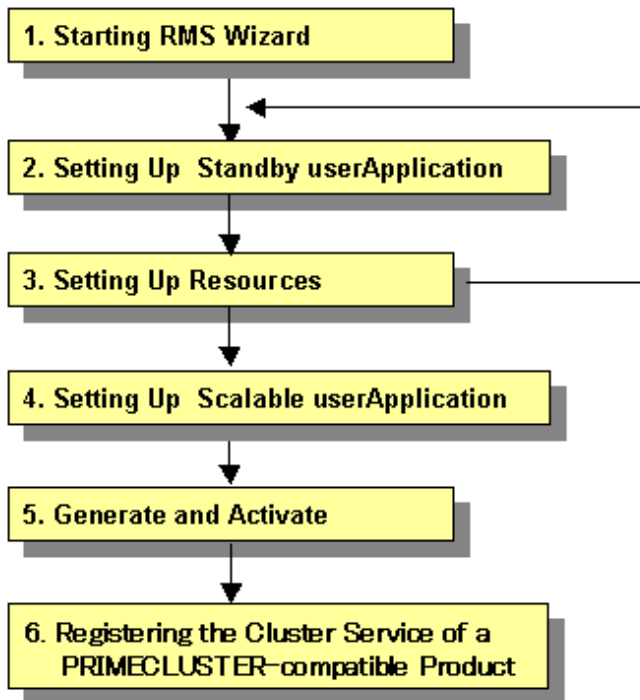
- Before you create cluster applications as part of scalable operation, create cluster applications in standby operation that act as the constituent factors of the cluster applications in scalable operation. To create cluster applications in standby operation, repeat steps 2. to 3.

  Example 1) For scalable operation with three nodes, repeat steps 2. and 3. three times to create three cluster applications of standby operation.

  Example 2) For high-availability scalable 1:1 standby (standby operation), repeat steps 2. and 3. once to create 1:1 standby cluster applications.

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 "8.7 Deleting a Cluster Application."

Note

- Set up the cluster application and resources based on the cluster application and resource information in the "A.10 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.

- Generate and Activate process fail if RMS is running. Using Cluster Admin or hvdisp, you need to confirm that RMS has not started before creating a cluster application. If RMS has already started, stop RMS from Cluster Admin or execute the "hvshut" command to stop RMS on all the nodes of the cluster system. For details on "hvdisp" and "hvshut", see the pages of these commands respectively.

- Set up "remote file copy" and "remote command execution" for the RMS Wizard. See the notes on "5.1.1 Setting Up CF and CIP." If the cluster interconnect is not protected by security, cancel the "remote file copy" and "remote command execution" settings on all the cluster nodes after setting up the cluster applications.

6.7.1 Starting RMS Wizard

Execute the "hvw" command.

The following is an example of starting the RMS Wizard with the configuration file name (testconf):

```
# /opt/SMAW/SMAWRrms/bin/hvw -n testconf
```

6.7.2 Setting Up userApplication

This section explains how to configure a cluster application.

There are two types of cluster applications, namely, standby operation and scalable operation.

Note that the term "userApplication" has the same meaning as "cluster application."

🖑 Note
. . . .

About the name of userApplication

Cluster application name must start with a letter and may have no more than 14 characters, either alphanumeric or "_" (underscore).

Cluster application name cannot be set to the same name of the configuration file name specified for the argument of hvw command described in "6.7.1 Starting RMS Wizard."

Only uppercase (capital) letters can be used in cluster application name.
. . . .

6.7.2.1 Creating Standby Cluster Applications

This section explains how to configure a cluster application for standby operation.

Operation Procedure:

1. Select "Application-Create" from the "Main configuration menu."

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                        10) Configuration-Remove
 2) QUIT                        11) Configuration-Freeze
 3) Application-Create          12) Configuration-Thaw
 4) Application-Edit            13) Configuration-Edit-Global-Settings
 5) Application-Remove          14) Configuration-Consistency-Report
 6) Application-Clone           15) Configuration-ScriptExecution
 7) Configuration-Generate      16) RMS-CreateMachine
 8) Configuration-Activate      17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 3
```

2. Select "STANDBY" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP                    7) SCALABLE
 2) QUIT                    8) STANDBY
 3) RETURN
 4) OPTIONS
 5) DEMO
 6) GENERIC
Application Type: 8
```

3. Next, "turnkey wizard STANDBY" will be output. Select "Machines+Basics" and then set up userApplication.

```
Consistency check ...
Yet to do: process the basic settings using Machines+Basics
Yet to do: process at least one of the non-basic settings

Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                          4) REMOVE+EXIT
 2) NO-SAVE+EXIT                  5) ApplicationName=APP1
 3) SAVE+EXIT                     6) Machines+Basics(-)
Choose the setting to process: 6
```

4. The userApplication setup page will appear. Set up the following for the userApplication:

   - Nodes that constitute the userApplication

   - Attributes of the userApplication

Set up the nodes that constitute userApplication by selecting "Machines[number]" and then a SysNode name on the subsequent screen that is displayed.

The procedure for setting up the nodes that constitute a userApplication is explained for each topology, below.

Topology	How to set up userApplication configuration nodes
1:1 standby	In "Machines[0]," specify a SysNode that is online when the userApplication first starts up.
	In "Machines[1]," specify a SysNode that is in standby status or offline when the userApplication first starts up
Mutual standby	**(For the first userApplication)**
	In "Machines[0]," specify a SysNode that is online when the userApplication first starts up.
	In "Machines[1]," specify a SysNode that is in standby status or offline when the userApplication first starts up.
	(For the second userApplication)
	In "Machines[0]," specify a SysNode that is online when the userApplication first starts up. For this SysNode, specify the SysNode specified for "Machines[1]" when the first userApplication was set up.
	In "Machines[1]," specify a SysNode that is in standby status or offline when the userApplication first starts up. Specify the SysNode specified in "Machines[0]" when the first userApplication was set up.
N:1 standby	**(For the first userApplication)**
	In "Machines[0]," specify a SysNode that is online when the userApplication first starts up.
	In "Machines[1]," specify a SysNode that is in standby status or offline when the userApplication first starts up.
	(For the second or subsequent userApplication)
	In "Machines[0]," specify a SysNode that is online when the userApplication first starts up. For this, specify a SysNode other than that previously specified for "Machines[0]" or "Machines[1]" when the userApplication was set up.
	In "Machines[1]," specify a SysNode that is in standby status or offline when the userApplication first starts up. For this, specify the same SysNode as that previously specified in "Machines[1]" when the userApplication was set up.

Topology	How to set up userApplication configuration nodes
Cascaded	In "Machines[0]," specify a SysNode that is online when the userApplication first starts up. For "Machines[1]" or later, specify a SysNode that is in standby status or offline when the userApplication first starts up. State transition occurs in ascending order of the numbers specified for "Machines[number]." Example) When there are four nodes, state transition occurs in the order shown below: "Machines[0]" -> "Machines[1]" -> "Machines[2]" -> "Machines[3]"
Priority transferring	**(For the first userApplication)** In "Machines[0]," specify a SysNode that is online when the userApplication first starts up. For "Machines[1]" or later, specify a SysNode that is in standby status or offline when the userApplication first starts up. **(For the second or subsequent userApplication)** In "Machines[0]," specify a SysNode that is online when the userApplication first starts up. For this, specify a SysNode other than that previously specified in "Machines[0]" when the userApplication was set up. For "Machines[1]" or later, specify a SysNode that is in standby status or offline when the userApplication first starts up. State transition occurs in ascending order of the numbers specified in "Machines[number]." Example) When there are four nodes, state transition occurs in the order shown below: "Machines[0]" -> "Machines[1]" -> "Machines[2]" -> "Machines[3]"

Set up the attributes of the userApplication as follows:

Attribute	Setup value	Contents	Remark
AutoStartUp	yes	Automatically starts up the cluster application when RMS is started.	To create a cluster application in standby operation that constitutes scalable operation, set "AutoStartUp" to "no." For information on how to create scalable cluster applications, see "6.7.2.2 Creating Scalable Cluster Applications."
AutoSwitchOver	HostFailure| ResourceFailure| Shutdown	Automatically performs failover if a node or resource fails or when the node is stopped.	Do not set a value in the single-node cluster operation.
HaltFlag	yes	The shutdown facility forcibly stops the application if another error (double errors) occurs during failover.	To ensure safe operation, always set "yes." Set [No] in the single-node cluster operation.

Attribute	Setup value	Contents	Remark
StandbyTransitions	ClearFaultRequest| StartUp| SwitchRequest	Monitor the states of the resources on the standby node performing standby operation.	This setting must be made when you are monitoring the states of the Gls resources on the standby node by using the redundant line control function of GLS.
LicenseToKill	yes	Sets up an exclusive relationship between cluster applications.	For information on exclusive relationships, see the following.
AutoBreak	yes	Cancels an exclusive relationship between cluster applications.	
PartialCluster	0	Allows userApplication to start even if RMS on all the nodes configuring userApplication is not started.	On cluster applications for a standby operation, "0" is set unless otherwise directed.

Note

In the case of the single-node cluster operation

- Do not set a value to "AutoSwitchOver".

- Set [No] to "HaltFlag".

- Do not set a value to "ShutdownPriority".

- Do not set a value to "OnlinePriority".

To set up exclusive relationships between cluster applications, you must set up the following.
For details on exclusive relationships between applications, see "6.7.7 Exclusive Relationships Between Cluster Applications."

Create multiple cluster application groups between which an exclusive relationship can be established. Exclusive control is established between the cluster applications within a single group.

Up to 52 groups of A to Z or a to z can be specified. "20X" and "10X" are fixed values. Therefore, you must always specify either "20X" or "10X" after the group.

- Example) When the cluster application is included in group A and the job priority is high

    A20X

- Example) When the cluster application is included in group A and the job priority is low

    A10X

Note

Exclusive relationships between cluster applications can be established only when the operation is being performed with two or more cluster applications. When the operation is to be performed with a single cluster application, do not set up any relationships between cluster applications.

- When a cluster application with a high job priority is to be used

    Select "LicenseToKill" and then enter the following into ">>" that is output after "FREECHOICE" has been selected.

```
        Group 20X
```

- When a cluster application with a low job priority is to be used

   Select "LicenseToKill" and then enter the following into ">>" that is output after "FREECHOICE" has been selected.

```
        Group 10X
```

   When an exclusive relationship is to be cancelled, set up the following:

```
        "LicenseToKill" : "no"

        "AutoBreak"     : "yes"
```

....

After the completion of setup, select "SAVE+EXIT."

![Note icon] **Note**
....

Operator intervention requests and error resource messages are displayed only when the AutoStartUp and PersistentFault attributes are set to yes(1). When the operator intervention and error resource messages are to be displayed, set yes(1) for the AutoStartUp and PersistentFault attributes. For information on the operator intervention and error resource messages, see "D.4 Operator Intervention Messages."

....

5. Confirm if the setup information is correct, and then select "SAVE+EXIT."

```
Machines+Basics (app1:consistent)
  1) HELP
  2) -
  3) SAVE+EXIT
  4) REMOVE+EXIT
  5) AdditionalMachine
  6) AdditionalConsole
  7) Machines[0]=fuji01RMS
  8) Machines[1]=fuji02RMS
  9) (PreCheckScript=)
 10) (PreOnlineScript=)
 11) (PostOnlineScript=)
 12) (PreOfflineScript=)
 13) (OfflineDoneScript=)
 14) (FaultScript=)
 15) (AutoStartUp=no)
 16) (AutoSwitchOver=HostFailure|ResourceFailure|ShutDown)
 17) (PreserveState=no)
 18) (PersistentFault=0)
 19) (ShutdownPriority=)
 20) (OnlinePriority=)
 21) (StandbyTransitions=ClearFaultRequest|StartUp|SwitchRequest)
 22) (LicenseToKill=no)
 23) (AutoBreak=yes)
 24) (HaltFlag=no)
 25) (PartialCluster=0)
 26) (ScriptTimeout=)
Choose the setting to process: 3
```

6. "turnkey wizard STANDBY" is output. Specify the settings for each resource.

```
Consistency check ...
Yet to do: process at least one of the non-basic settings

Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                          9) Procedure:BasicApplication(-)
 2) -                            10) Symfoware(-)
 3) SAVE+EXIT                     11) Procedure:SystemState3(-)
 4) -                            12) Procedure:SystemState2(-)
 5) ApplicationName=APP1         13) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)        14) IpAddresses(-)
 7) CommandLines(-)              15) LocalFileSystems(-)
 8) Procedure:Application(-)     16) Gds:Global-Disk-Services(-)
Choose the setting to process:
```

6.7.2.2 Creating Scalable Cluster Applications

This section explains how to register a cluster application in scalable operation.

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.

Example 1) Preparing for scalable operation

When you create a cluster application in a scalable operation, you must first create a cluster application in a standby operation, which is a prerequisite for scalable operation.

If the cluster application of scalable operation is to run on three nodes, create a cluster application of standby operation on each of those nodes (the node is for operation only and has no standby).

When you create a cluster application for standby operation, which is a prerequisite for scalable operation, set up only "Machines[0]."

Example 2) Preparing for high-availability scalable operation

To create a high-availability scalable cluster application, you must first create a cluster application for standby operation, which is a prerequisite for high-availability scalable operation.

If the cluster application for high-availability scalable operation is 1:1 standby, create a cluster application for 1:1 standby.
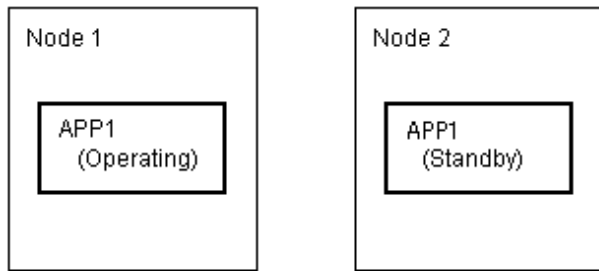
: Cluster application of 1:1 standby that is a prerequisite
for high-availability scalable operation

![Note icon] **Note**

. . . .

To create a cluster application in standby operation that constitutes scalable operation, set "AutoStartUp" to "no." To start the cluster applications automatically when you start RMS, set the value of "AutoStartUp" to "yes" when you create a cluster application as part of scalable operation.

. . . .

The procedure for setting up the node of a cluster application in a standby operation, which is a prerequisite for scalable operation, is as shown below.

Topology	How to set up userApplication configuration nodes
Scalable	In "Machines[0]," specify a SysNode that is online when the userApplication first starts up. Since standby is not included, you do not need to set up "Machines[1]" and any subsequent items.
High-availability scalable	Note that the set-up method varies depending on the topology of the standby operation that is a constituent factor of the cluster application in scalable operation. For information on making this setting, see how to set up the topology of each standby operation.

For information on how to create standby cluster applications, see "6.7.2.1 Creating Standby Cluster Applications."

After you complete the setup of standby operation, which is a prerequisite for scalable operation, you must create the cluster application of scalable operation as explained below.

Creating scalable cluster applications

Operation Procedure:

1. Select "Application-Create" from the "Main configuration menu."

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                          10) Configuration-Remove
 2) QUIT                          11) Configuration-Freeze
 3) Application-Create            12) Configuration-Thaw
 4) Application-Edit              13) Configuration-Edit-Global-Settings

 5) Application-Remove            14) Configuration-Consistency-Report
 6) Application-Clone             15) Configuration-ScriptExecution
 7) Configuration-Generate        16) RMS-CreateMachine
 8) Configuration-Activate        17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 3
```

2. Select "SCALABLE" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP                          7) SCALABLE
 2) QUIT                          8) STANDBY
 3) RETURN
 4) OPTIONS
 5) DEMO
 6) GENERIC
Application Type: 7
```

3. "turnkey wizard "SCALABLE"" is output. Select "Machines+Basics" and set up the userApplication.

```
Consistency check ...
Yet to do: process the basic settings using Machines+Basics
Yet to do: process at least one of the non-basic settings

Settings of turnkey wizard "SCALABLE" (APP2:not yet consistent)
 1) HELP                          4) REMOVE+EXIT
 2) NO-SAVE+EXIT                  5) ApplicationName=APP2
 3) SAVE+EXIT                     6) Machines+Basics(-)
Choose the setting to process: 6
```

4. The userApplication setup screen is output. Specify the following settings for the userApplication:

   - Nodes where the userApplication is configured

   - userApplication attributes

   Set up the nodes where the userApplication is configured as follows:

   - Specify all SysNode names where the cluster application is configured (standby operation) in "Machines[*number*]".

   Refer to the following when setting the userApplication attributes:

Attribute	Setting	Description	Remarks
AutoStartUp	yes	Automatically starts the cluster application when RMS is started.	

5. For "Settings of application type," select "SAVE+EXIT."

```
Consistency check ...

Machines+Basics (app1:consistent)
 1) HELP
 2) -
 3) SAVE+EXIT
 4) REMOVE+EXIT
 5) AdditionalMachine
 6) AdditionalConsole
 7) Machines[0]=fuji01RMS
 8) Machines[1]=fuji02RMS
 9) (PreCheckScript=)
10) (PreOnlineScript=)
11) (PostOnlineScript=)
12) (PreOfflineScript=)
13) (OfflineDoneScript=)
14) (FaultScript=)
15) (AutoStartUp=no)
16) (AutoSwitchOver=HostFailure|ShutDown)
17) (PreserveState=yes)
18) (PersistentFault=0)
19) (ShutdownPriority=)
20) (OnlinePriority=0)
21) (StandbyTransitions=)
22) (LicenseToKill=no)
23) (AutoBreak=yes)
24) (HaltFlag=no)
25) (PartialCluster=1)
26) (ScriptTimeout=)
Choose the setting to process: 3
```

6. "turnkey wizard "SCALABLE"" is output. Select "Controllers."

```
Consistency check ...
Yet to do: process at least one of the non-basic settings

Settings of turnkey wizard "SCALABLE" (APP2:not yet consistent)
 1) HELP                  4) -                      7) Controllers(-)
 2) -                     5) ApplicationName=APP2
 3) SAVE+EXIT             6) Machines+Basics(app2)
Choose the setting to process: 7
```

7. "Settings of application type" is output. Select "AdditionalAppToControl."

```
Consistency check ...
Yet to do: assign at least one application to control
Yet to do: configure at least one controlled application without the M flag

Settings of application type "Controller" (not yet consistent)
 1) HELP                  4) REMOVE+EXIT            7) (FaultScript=)
 2) -                     5) ControlPolicy=SCALABLE 8) (ApplicationSequence=)
 3) SAVE+EXIT             6) AdditionalAppToControl 9) (StateChangeScript=)
Choose the setting to process: 6
```

8. Select a cluster application (standby operation) that allows scalable operation.

```
 1) HELP
 2) RETURN
 3) FREECHOICE
 4) app1
 5) app2
Choose an application to control: 4
```

9. Select "SAVE+RETURN" from "Set global flags for scalable."

```
Set *global* flags for all scalable (sub) applications: app1
Currently set: TIMEOUT (T180)
 1) HELP                           5) MONITORONLY(M)
 2) -                              6) TIMEOUT(T)
 3) SAVE+RETURN
 4) DEFAULT
Choose one of the flags: 3
```

10. To allow scalable operation with multiple cluster applications (standby operation), repeat steps 7. to 9.

11. Set up the order in which cluster applications are started up (standby operation). When you start the cluster applications, start from the one with the smallest startup sequence number. When stopping, from the one with the largest startup sequence number. Cluster applications with the same startup sequence number must start up or stop in parallel.

Note

If you do not need to set up a startup sequence number, you do not have to perform the procedure described below.

1. Select "(ApplicationSequence=)" from "Settings of application type."

```
Settings of application type "Controller" (consistent)
 1) HELP                           7) Controllers[0]=T180:app1
 2) -                              8) Controllers[1]=T180:app2
 3) SAVE+EXIT                      9) (FaultScript=)
 4) REMOVE+EXIT                   10) (ApplicationSequence=)
 5) ControlPolicy=SCALABLE        11) (StateChangeScript=)
 6) AdditionalAppToControl
Choose the setting to process: 10
```

2. Select "FREECHOICE."

```
 1) HELP
 2) RETURN
 3) NONE
 4) FREECHOICE
Set the application sequence: 4
```

3. Enter the startup sequence number, and then press the return key.

   - Enter the cluster application with the highest startup sequence number first.

   - If the startup sequence numbers are different, input a single colon (:) between the cluster applications.

   - If the startup priority numbers are the same, input a single space between the cluster applications.

📖 **Note**

. . . .

The cluster application for standby operation must be entered entirely in lowercase characters.

The following is an example in which the startup sequence of app1 is the first, followed by app2 and then app3 (app2 and app3 have the same startup sequence number).

. . . .

```
1) HELP
2) RETURN
3) NONE
4) FREECHOICE
Set the application sequence: 4
     >> app1:app2 app3
```

12. Select "SAVE+EXIT" from "Settings of application type."

```
Settings of application type "Controller" (consistent)
 1) HELP                          7) Controllers[0]=T180:app1
 2) -                             8) Controllers[1]=T180:app2
 3) SAVE+EXIT                     9) (FaultScript=)
 4) REMOVE+EXIT                  10) (ApplicationSequence=app1:app2)
 5) ControlPolicy=SCALABLE       11) (StateChangeScript=)
 6) AdditionalAppToControl
Choose the setting to process: 3
```

When two or more cluster applications for scalable operation are to be created, repeat steps 1. to 12.

6.7.3 Setting Up Resources

This section explains how to register resources to the userApplication that was set up in the previous section.

You can register the following resources:

- **Cmdline resources**

  You can use Cmdline resources to set up script files or commands as resources. The Cmdline resources are required to generate the state transition of userApplication along with the stop of user applications, and conversely, to start or stop ISV applications or user applications along with the state transition of the userApplication.

- **Fsystem resources**

  Used when you mount a file system along with userApplication startup.

📖 **Note**

. . . .

To use a file system in a class created by GDS as an Fsystem resource, you must register the Gds resource to the same userApplication.

. . . .

- **Gds resources**

  Used when you start and stop a disk class to be defined by GDS (Global Disk Services) by linking it with the userApplication.

- **Gls resources**

  Used when you set up a takeover IP address that is to be defined in a userApplication with the redundant line control function, or when you set a takeover IP address in a userApplication with the single line control function.

- **Takeover network resource**

  Used when you set up a takeover IP address on a single network interface without using GLS.

- **Procedure resources**

  Used when you register a state transition procedure in the userApplication.

Resource setup flow
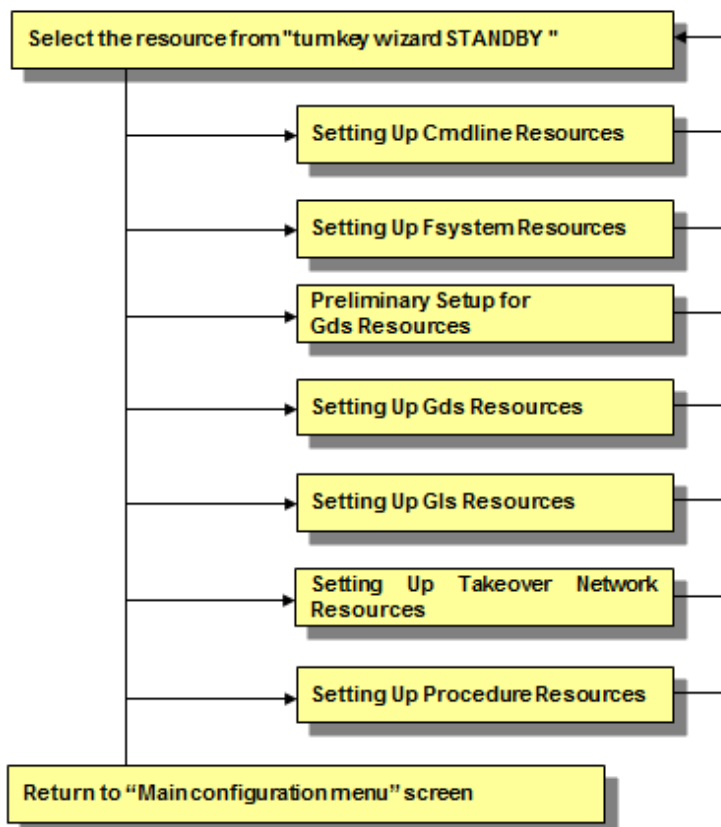
6.7.3.1 Setting Up Cmdline Resources

This section describes the procedure for creating Cmdline resources.

Operation Procedure:

1. Select "CommandLines" from "turnkey wizard STANDBY."

```
Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                          9) Procedure:BasicApplication(-)
 2) -                            10) Symfoware(-)
 3) SAVE+EXIT                    11) Procedure:SystemState3(-)
 4) -                            12) Procedure:SystemState2(-)
 5) ApplicationName=APP1         13) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)        14) IpAddresses(-)
 7) CommandLines(-)              15) LocalFileSystems(-)
 8) Procedure:Application(-)     16) Gds:Global-Disk-Services(-)
Choose the setting to process:7
```

2. "CommandLines" will appear. Select "AdditionalStartCommand."

```
CommandLines (Cmd_APP1:not yet consistent)
1) HELP                          4) REMOVE+EXIT
2) -                             5) AdditionalStartCommand
3) SAVE+EXIT                     6) (ReturnCodes)
Choose the setting to process: 5
```

3. Select "FREECHOICE" and then enter the full path of the StartCommand. If you need to specify arguments, delimit them with blanks.

   StartCommand is executed during Online processing to start user applications.

   [StartCommand exit codes]

   StartCommand has the following exit codes:

   0: Normal exit. The Online processing is successfully done.

   Other than 0: Abnormal exit. The Online processing fails. When the script exits with the cord other than 0, the resource will enter Faulted.

```
1) HELP
2) RETURN
3) FREECHOICE
Set a start command: 3
        >>
```

4. "CommandLines" will appear. If you need to stop the user programs, select "StopCommands."

   StopCommand is executed during Offline processing to stop user applications.

   You do not always have to set up the StopCommand.

   [StopCommand exit codes]

   StopCommand has the following exit codes:

   0: Normal exit. The Offline processing is successfully done.

   Other than 0: Abnormal exit. The Offline processing fails. When the script exits with the cord other than 0, the resource will enter Faulted.

   If you do not use StopCommand, start from step 6.

```
CommandLines (Cmd_APP1:consistent)
 1) HELP                              7) StopCommands[0]=none
 2) -                                 8) CheckCommands[0]=none
 3) SAVE+EXIT                         9) CheckCommandTimeouts[0]=none
 4) REMOVE+EXIT                      10) Flags[0]=DT300
 5) AdditionalStartCommand           11) (ReturnCodes)
 6) StartCommands[0]='xxxx'
Choose the setting to process:7
```

5. Select "FREECHOICE" and then enter the full path of StopCommand. If you need to specify arguments, delimit them with blanks.

```
1) HELP
2) RETURN
3) FREECHOICE
4) NONE
Set a start command: 3
        >>
```

6. "CommandLines" will appear. Select "CheckCommands."

   CheckCommand is executed periodically to notify RMS of the state of the user applications.

   [CheckCommand exit codes]

       CheckCommand has the following exit codes:

           0: Online.

           Other than 0: Offline.

   If you enable the ALLEXITCODES attribute, the script that is defined in CheckCommand provides more detailed state of the user applications. Change the attribute in step 8.

```
CommandLines (Cmd_APP1:consistent)
 1) HELP                              7) StopCommands[0]='yyyy'
 2) -                                 8) CheckCommands[0]=none
 3) SAVE+EXIT                         9) CheckCommandTimeouts[0]=none
 4) REMOVE+EXIT                      10) Flags[0]=DT300
 5) AdditionalStartCommand           11) (ReturnCodes)
 6) StartCommands[0]='xxxx'
Choose the setting to process:8
```

7. Select "FREECHOICE" and then enter the full path of the CheckCommand. If you need to specify arguments, delimit them with blanks.

```
1) HELP
2) RETURN
3) FREECHOICE
4) NONE
Set a start command: 3
        >>
```

8. Change the attribute.

   Change the attribute to suit the purpose. To change the attribute, select "Flags[0]." For details on the attribute, see "11 Appendix - Attributes" and "HELP" in the "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide*."

Note

If you enable the "NULLDETECTOR" attribute, CheckCommand is not started from RMS. For hot-standby operation, enable the following two attributes;

- STANDBYCAPABLE

  RMS executes Standby processing of the resources on all nodes where the userApplication is Offline.

- ALLEXITCODES

  Check script provides the detailed state of the resource with the exit code.

  For further details about the hot-standby operation settings, see "6.6 Setting Up Online/Offline Scripts."

9. Finally, select "SAVE+EXIT."

```
CommandLines (Cmd_APP1:consistent)
 1) HELP                            7) StopCommands[0]='yyyy'
 2) -                              8) CheckCommands[0]=none
 3) SAVE+EXIT                      9) CheckCommandTimeouts[0]=none
 4) REMOVE+EXIT                   10) Flags[0]=DT300
 5) AdditionalStartCommand        11) (ReturnCodes)
 6) StartCommands[0]='xxxx'
Choose the setting to process:3
```

6.7.3.2 Setting Up Fsystem Resources

An Fsystem resource must be set up 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.

[Prerequisites]

You need to work on the following prerequisites to use an Fsystem resource.

1. Reserving monitoring disk area

   Reserve the dedicated monitoring disk area in shared disk device so that shared disk failures and cable failures are detected on each shared disk device.

   Reserve the monitoring area in either 1or 2 unit shown in the table below. Configuration based on the No. 2 is recommended.

No	Reserve monitoring area for each	Range of failure detection	Prerequisites
1	Shared disk device	Failures including cable disconnection are detected in unit of shared disk device.	- Reserve one monitoring disk area (slice or GDS volume) per shared disk device when configuring the shared disk after OS is installed. e.g.) /dev/sdf1 - Allow about 10 MB for monitoring disk area.
2	Disk area (1LU)	In addition to the above, failures such as LUN errors on shared disk device are detected in unit of shared disk area (operation unit).	- Reserve one monitoring disk area (slice or GDS volume) per LU used for operations when configuring the shared disk after OS is installed. e.g.) If /dev/sdd1 and /dev/sde1 are used for disk operation, specify /dev/sdd2 and /dev/sde2 as monitoring disks. - Allow about 10 MB for monitoring disk area per LU as monitoring area.

2. Defining mount points

   Define mount points that are used as resources in /etc/fstab.pcl on all the nodes where userApplication is configured. Each line must begin with "#RMS#."

   ```
   Example) /etc/fstab.pcl
   ```

   ```
   #RMS#/dev/sdd1          /mnt/swdsk1          ext3    defaults      1 2
   #RMS#/dev/sdd2          /mnt/swdsk2          ext3    defaults      1 2
   #RMS#/dev/sdd3          /mnt/monitor         ext3    defaults,ro   1 2
   ```

   If a file system is created in the GDS setting screen, the mount point is also defined in /etc/fstab. Copy the line of the mount point defined by GDS onto /etc/fstab.pcl, and each line of the mount point that is set with GDS must begin with "#RMS#."

   ```
   Example) /etc/fstab.pcl
   ```

   ```
   #RMS#/dev/sfdsk/class0001/dsk/volume0001 /mnt/swdsk1 ext3 noauto 0 0
   #RMS#/dev/sfdsk/class0001/dsk/volume0002 /mnt/swdsk2 ext3 noauto 0 0
   #RMS#/dev/sfdsk/class0001/dsk/volume0003 /mnt/monitor ext3 noauto,ro 0 0
   ```

   📓 Note

   - You must set the shared classes that were used to create the file system from the GDS setting screen as Gds resources. For details, see "6.7.3.4 Setting Up Gds Resources."

   - Read Only must be set for the mount attribute of the monitoring-only disk area.

     Example:

     ```
     Example)
     ```

     ```
     #RMS#/dev/sdf1          /mnt/monitor   ext3    defaults,ro   1 2
     ```

   - You cannot define different volumes on the same mount point.

3. Setting file systems

   It is necessary to create file systems on disk partitions that are used as resources.

   ```
   Example) When creating the ext3 file system
   ```

   ```
   # /sbin/mkfs -t ext3 /dev/sdd1
   ```

   Confirm on each node that the file system can be mounted.

   ```
   Example) When checking if the file systems can be mounted
   ```

   ```
   # /bin/mount -t ext3 /dev/sdd1 /mnt/swdsk1
   # /bin/umount /mnt/swdsk1
   ```

   📓 Note

   - Set file systems for the monitoring disk area as well.

   - Forcible file system check prevention (recommended)

     If ext3 is used for a file system, the file system might forcibly be checked during online processing of a switching file system. It is part of the ext3 specification that file systems are checked when a certain number of mounting has been executed since the last file system check, or a certain period of time has passed.

     When the file systems are forcibly checked along with node startup or failover, timeout occurs due to file system online processing, and PRIMECLUSTER startup or failover might fail.

It is necessary to prevent the file systems from being checked by executing the following command for all the ext3 switching files.

```
# tune2fs -c0 -i0 <device name>
```

Example)

```
# tune2fs -c0 -i0 /dev/sdi1
```

After executing the above command, check if "Maximum mount count :-1","Check interval:0"is displayed using the following command:

```
# tune2fs -l /dev/sdi1
[snip]
Mount count: 10
Maximum mount count: -1
[snip]
Check interval: 0 (<none>)
[snip]
```

If the forcible file system check is prevented, file systems might corrupt due to failures such as disk errors and kernel bug. These failures cannot be detected through file system logging and journaling. The file system corruption might cause data corruption. To prevent this, execute the "fsck - f" command to enable the file system forcible check during periodic maintenance.

Operation Procedure:

1. Select "LocalFileSystems" from "turnkey wizard STANDBY."

```
Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                           9) Procedure:BasicApplication(-)
 2) -                             10) Symfoware(-)
 3) SAVE+EXIT                     11) Procedure:SystemState3(-)
 4) -                             12) Procedure:SystemState2(-)
 5) ApplicationName=APP1          13) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)         14) IpAddresses(-)
 7) CommandLines(-)               15) LocalFileSystems(-)
 8) Procedure:Application(-)      16) Gds:Global-Disk-Services(-)
Choose the setting to process:15
```

2. Select "AdditionalMountPoint."

```
File systems (Lfs_APP1:consistent)
 1) HELP                  4) REMOVE+EXIT           7) (Timeout=180)
 2) -                     5) AdditionalMountPoint
 3) SAVE+EXIT             6) (Filter=)
Choose the setting to process: 5
```

3. The mount point, which is defined in /etc/fstab.pcl, will appear. Select mount points for monitoring-only disks.

Note

Set up mount points for monitoring-only disks before setting disk mount points.

If you register multiple mount points, repeat steps 3 to 7 for each mount point.

```
1) HELP                        6) /mnt/swdskl
2) RETURN                      7) /mnt/monitor2
3) FREECHOICE                  8) /mnt/swdsk2
4) ALL
5) /mnt/monitorl
Choose a mount point: 5
```

4. Select "SAVE+RETURN."

```
Set flags for mount point: /mnt/monitorl Currently set: LOCAL,AUTORECOVER (LA)
1) HELP                   4) DEFAULT              7) SHARE(S)
2) -                      5) SYNC(Y)              8) MONITORONLY(M)
3) SAVE+RETURN            6) NOT:AUTORECOVER(A)
Choose one of the flags: 3
```

5. The mount point, which is defined in /etc/fstab.pcl, will appear. Select mount points for disk operation.

```
1) HELP                        6) /mnt/swdskl
2) RETURN                      7) /mnt/monitor2
3) FREECHOICE                  8) /mnt/swdsk2
4) ALL
5) /mnt/monitorl
Choose a mount point: 6
```

6. Select "MONITORONLY", so the "MONITORONLY" attribute will be ON.

```
Set flags for mount point: /mnt/swdskl Currently set: LOCAL,AUTORECOVER (LA)
1) HELP                   4) DEFAULT              7) SHARE(S)
2) -                      5) SYNC(Y)              8) MONITORONLY(M)
3) SAVE+RETURN            6) NOT:AUTORECOVER(A)
Choose one of the flags: 8
```

7. Select "SAVE+RETURN."

```
Set flags for mount point: /mnt/swdskl Currently set:
LOCAL,AUTORECOVER,MONITORONLY (LAM)
1) HELP                   4) DEFAULT              7) SHARE(S)
2) -                      5) SYNC(Y)              8) NOT:MONITORONLY(M)
3) SAVE+RETURN            6) NOT:AUTORECOVER(A)
Choose one of the flags: 3
```

8. Select "SAVE+EXIT."

```
File systems (Lfs_APPl:consistent)
1) HELP                        6) MountPoints[0]=LA:/mnt/monitorl
2) -                           7) MountPoints[1]=LAM:/mnt/swdskl
3) SAVE+EXIT                   8) (Filter=)
4) REMOVE+EXIT                 9) (Timeout=360)
5) AdditionalMountPoint
Choose the setting to process: 3
```

6.7.3.3 Preliminary Setup for Gds Resources

[Prerequisites]

If you need to set up a Gds resource, you must first set up a shared volume.
Also, before you make the settings required for the Gds resources, execute the following command on either node.

This command operation is required for linking the specified Gds resources with the status of the userApplication and to start and stop the GDS volume. In addition, this command sets all nodes sharing the volume to offline status.

```
# /opt/SMAW/SMAWRrms/bin/hvgdsetup -a [class-name]
...
Do you want to continue with these processes ? [yes/no] y
```

📘 Information
....

To check the setup status of a shared volume, execute the following command:

```
# /opt/SMAW/SMAWRrms/bin/hvgdsetup -l
```
....

📙 Note
....

- If the preliminary setup is not performed, the cluster application is set to Inconsistent status. For details, see "(3) Cluster applications become "Inconsistent"." in "F.1.9 Cluster System Related Error" of the *"PRIMECLUSTER Global Disk Services Configuration and Administration Guide."*

- This operation must not be performed when a GFS shared file system is used.
....

6.7.3.4 Setting Up Gds Resources

After completing the preliminary setup for the Gds resources, register the Gds resources to the userApplication. The procedure for setting Gds resources to a userApplication is described below.

Operation Procedure:

1. Select "Gds:Global-Disk-Services" from "turnkey wizard STANDBY."

```
Settings of turnkey wizard "STANDBY"  (APP1:not yet consistent)
 1) HELP                              9) Procedure:BasicApplication(-)
 2) -                                10) Symfoware(-)
 3) SAVE+EXIT                        11) Procedure:SystemState3(-)
 4) -                                12) Procedure:SystemState2(-)
 5) ApplicationName=APP1             13) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)            14) IpAddresses(-)
 7) CommandLines(-)                  15) LocalFileSystems(-)
 8) Procedure:Application(-)         16) Gds:Global-Disk-Services(-)
Choose the setting to process:16
```

2. Select "AdditionalDiskClass" from "Volume management."

   If you want to register multiple resources, repeat steps 2 to 3 for each resource.

   Example) When you register the class [CL] of the shared volume in userApplication:

```
Volume management (Gds_APP1:not yet consistent)
 1) HELP                4) REMOVE+EXIT           7) (StandbySupport=no)
 2) -                   5) AdditionalDiskClass   8) (AutoRecover=no)
 3) SAVE+EXIT           6) (ClassNameFilter=)    9) (Timeout=1800)
Choose the setting to process:5
```

3. A list of the cluster resources that can be registered in userApplication appears.

   - In the case of the multiple-node cluster operation
     Select the cluster resource.

   ```
   1) HELP
   2) RETURN
   3) FREECHOICE
   4) CL
   Choose the setting to process:4
   ```

   - In the case of the single-node cluster operation
     Enter the created class name of GDS by selecting "FREECHOICE", since the list of the cluster resources is not displayed.

4. Select "SAVE+EXIT."

   ```
   Volume management (Gds_APP1:consistent)
    1) HELP                5) AdditionalDiskClass    9) (AutoRecover=no)
    2) -                   6) DiskClasses[0]=CL     10) (Timeout=1800)
    3) SAVE+EXIT           7) (ClassNameFilter=)
    4) REMOVE+EXIT         8) (StandbySupport=no)
   Choose the setting to process:3
   ```

6.7.3.5 Setting Up Gls Resources

[Prerequisites]

Before setting up the Gls resources, you must first set up the virtual interfaces and takeover virtual interfaces. For details, see "*PRIMECLUSTER Global Link Services Configuration and Administration Guide : Redundant Line Control Function.*"

Operation Procedure:

1. Select "Gls:Global-Link-Services" from "turnkey wizard STANDBY."

   ```
   Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
    1) HELP                            9) Procedure:BasicApplication(-)
    2) -                              10) Symfoware(-)
    3) SAVE+EXIT                      11) Procedure:SystemState3(-)
    4) -                              12) Procedure:SystemState2(-)
    5) ApplicationName=APP1           13) Gls:Global-Link-Services(-)
    6) Machines+Basics(app1)          14) IpAddresses(-)
    7) CommandLines(-)                15) LocalFileSystems(-)
    8) Procedure:Application(-)       16) Gds:Global-Disk-Services(-)
   Choose the setting to process:13
   ```

2. Select "AdditionalTakeoverIpaddress" and then set the takeover IP address.

   If you need to register multiple resources, repeat steps 2 to 4 for each resource.

   ```
   Gls (Gls_APP1:not yet consistent)
    1) HELP                    4) REMOVE+EXIT
    2) -                       5) AdditionalTakeoverIpaddress
    3) SAVE+EXIT               6) (Timeout=60)
   Choose the setting to process: 5
   ```

3. The takeover IP addresses are displayed as options.

   Select one.

```
 1) HELP
 2) RETURN
 3) FREECHOICE
 4) 10.34.214.185
Choose a takeover IP address for Gls: 4
```

4. Set up the operation mode of the resource.

   The operation modes of the resources are "AUTORECOVER(A)" (selective value) and "NOT:AUTORECOVER(N)" (default value). If you select "SAVE+RETURN," the default value "NOT:AUTORECOVER(N)" will be set. Select "AUTORECOVER(A)" if an attempt to recover the resource is to be made for a given duration (default: 60 seconds) when a resource error occurs, or select "NOT:AUTORECOVER(N)" if switchover to another node is to be performed immediately.

```
Set a flag for takeover IP address: 10.34.214.185
Currently set:
 1) HELP                            5) AUTORECOVER(A)
 2) -
 3) SAVE+RETURN
 4) DEFAULT
Choose additionally one of the flags: 3
```

5. To save the Gls resource settings and then exit, select "SAVE+EXIT."

   You can change the timeout value of the Gls resource by selecting "(Timeout=60)" and setting any value (seconds).

```
Gls (Gls_APP1:consistent)
 1) HELP                           5) AdditionalTakeoverIpaddress
 2) -                              6) TakeoverIpaddress[0]=N,10.34.214.185
 3) SAVE+EXIT                      7) (Timeout=60)
 4) REMOVE+EXIT
Choose the setting to process: 3
```

See

By setting up the value in the StandbyTransition attribute when the cluster application is created, Gls resources on the standby node can be switched to the "Standby" state and the state of the Gls resources on the standby node can be monitored. For information on how to make this setting, see "6.7.2.1 Creating Standby Cluster Applications."

6.7.3.6 Setting Up Takeover Network Resources

[Prerequisites]

You need to configure the following information before using takeover network resources.

1. Define an IP address and a host name to be taken over between nodes

   First, define a host address to be taken over between nodes. Next, allocate a host name to this IP address. The allocated host name should be defined in the /etc/hosts file for the node which uses the takeover IP address.

2. Define a target host

   If you want to monitor the reachability of your network using ICMP while a takeover IP address is active, define a host name and an IP address of the device (e.g. router or layer 3 hub) which enables ICMP communication between more than one node in the /etc/hosts file. This process can be skipped if you do not need to monitor the reachability of the network.

3. Confirming a network interface

   Confirm the network interface (e.g. eth2) on which the takeover IP address will be active is being recognized by the operating system and available.

4. Define a setting file for the cluster

   Add the takeover IP address to the following file:
   /usr/opt/reliant/etc/hvipalias
   Enter entries according to the following rules.

   ```
   <node name> <takeover> <interface> <netmask>
   ```

   ```
   <node name> : Host name of the node which uses the takeover IP address (The result of executing
                 "uname -n")
   <takeover>  : Host name of the takeover IP address
   <interface> : Network interface name on which the takeover IP address will be active
   <netmask>   : Netmask for the takeover IP address
   ```

   ### 📔 Example

   ....

   When an IP address for the host "takeover" (netmask 255.255.255.0) is taken over between two nodes (node0 and node1) on the
   network interface eth2, define as follows:

   ```
   node0 takeover eth2 0xffffff00
   node1 takeover eth2 0xffffff00
   ```

   ....

Operation Procedure:

1. Select "IpAddresses" of turnkey wizard "STANDBY"

   ```
   Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
   1) HELP                            9) Procedure:BasicApplication(-)
   2) -                              10) Symfoware(-)
   3) SAVE+EXIT                      11) Procedure:SystemState3(-)
   4) -                              12) Procedure:SystemState2(-)
   5) ApplicationName=APP1           13) Gls:Global-Link-Services(-)
   6) Machines+Basics(app1)          14) IpAddresses(-)
   7) CommandLines(-)                15) LocalFileSystems(-)
   8) Procedure:Application(-)       16) Gds:Global-Disk-Services(-)
   Choose the setting to process:14
   ```

2. When you have previously specified the target host to monitor its network reachability using ICMP, select "AdditionalPingHost"
   and specify that target host.

   ```
   Ipaddresses and ipaliases (Adr_APP1:not yet consistent)
   1) HELP                4) REMOVE+EXIT            7) (NeedAll=yes)
   2) -                   5) AdditionalInterface    8) (Timeout=60)
   3) SAVE+EXIT 6) AdditionalPingHost 9) (InterfaceFilter=)
   Choose the setting to process:6
   ```

   The target host name registered in the process of prerequisites will be shown as an option. Select the host name you have previously
   specified.

   ```
   1) HELP
   2) RETURN
   3) FREECHOICE
   4) router
   5) l3hub
   6) takeover
   Choose another trusted host to ping:4
   ```

   When you finish specifying the target host, you will be brought back to the previous screen. Since you are required to specify more
   than one target host, you need to select "AdditionalPingHost" again to add another target host on the previous screen.

3. Select "AdditionalInterface" to set up the takeover IP address.

   When you have more than one IP address, you need to repeat this process for each IP address.

```
Ipaddresses and ipaliases (Adr_APP1:not yet consistent)
1) HELP                            7) PingHostPool[0]=router
2) -                               8) PingHostPool[1]=l3hub
3) SAVE+EXIT                       9) (NeedAll=yes)
4) REMOVE+EXIT                    10) (Timeout=60)
5) AdditionalInterface           11) (InterfaceFilter=)
6) AdditionalPingHost
Choose the setting to process:
```

1. Takeover IP address registered in the process of prerequisites will be shown as an option.

   Select the host name for the takeover IP address you have previously specified.

```
1) HELP
2) RETURN
3) FREECHOICE
4) router
5) l3hub
6) takeover
Choose an interface name:6
```

2. Specify the operation mode for the takeover IP address.

```
Currently set: VIRTUAL,AUTORECOVER (VA)
1) HELP                4) DEFAULT              7) MONITORONLY(M)
2) -                   5) BASE(B)              8) PING(P)
3) SAVE+RETURN         6) NOT:AUTORECOVER(A)
Choose one of the flags:
```

- DEFAULT

  If you choose "DEFAULT", all values will revert back to their default values.

- BASE, VIRTUAL

  The default value is "VIRTUAL".

  - BASE

  If you specify "BASE", activation/deactivation of the takeover IP address and activation/deactivation of the physical interface (for example, eth2) are performed at the same time. "BASE" will be shown on "Currently set" and "5) VIRTUAL" is shown on the menu page.

  - VIRTUAL

  If you specify "VIRTUAL", activation/deactivation of the takeover IP address and activation/deactivation of the logical interface (for example, eth2:1) are performed at the same time. "VIRTUAL" will be shown on "Currently set" and "5) BASE" is shown on the menu page.

  You must activate the IP address on the physical interface (for example, eth2) where the logical interface will be created beforehand because the takeover IP address with this attribute specifies the IP address for the logical interface. To activate the IP address on the physical interface beforehand, make settings so that the IP address is activated on the physical interface at startup of the operating system, or register the takeover IP address with "BASE" attribute with the same takeover network resource.

- AUTORECOVER, NOT:AUTORECOVER

  If you reactivate the takeover IP address, specify this attribute. The default value is "AUTORECOVER".

  - AUTORECOVER

  If you specify "AUTORECOVER" and the network interface goes down or becomes unreachable due to an error, it will try to activate the takeover IP address only once. "AUTORECOVER" will be shown on "Currently set"and "6) NOT:AUTORECOVER" is shown on the menu page. When the activation of the takeover IP address fails, it will be notified to the cluster.

  - NOT:AUTORECOVER

If you specify "NOT:AUTORECOVER", the "AUTORECOVER" setting will be disabled. "NOT:AUTORECOVER" will be shown on "Currently set" and "AUTORECOVER" is shown on the menu page."

- BASIC-IF

You cannot use this attribute. Do not change.

- MONITORONLY, NOT:MONITORONLY

- MONITORONLY

If you specify "MONITORONLY" and the network interface goes down or becomes unreachable due to an error, the error will not be notified to the cluster. "MONITORONLY" will be shown on "Currently set" and "7) NOT:MONITORONLY" is shown on the menu page. If you specify this attribute, a switchover due to a takeover IP address failure will not occur; however, even though "MONITORONLY" is specified for all takeover IP addresses, it is not possible to completely deter switching.

- NOT:MONITORONLY

If you specify "NOT:MONITORONLY", the "MONITORONLY" setting will be disabled. "NOT:MONITORONLY" will be shown on "Currently set" and "7) MONITORONLY" is shown on the menu page. At least one out of all takeover IP addresses you have registered should be set to "NOT:MONITORONLY".

- PING

By setting this attribute, you can specify the previously configured target host for the takeover IP address. Select the target host name to be monitored which you have set in the process of prerequisites.

```
1) HELP
2) RETURN
3) router(000)
4) l3hub
Choose a ping host of the pool ():3
```

4. Change "Timeout" if needed.

Use "Timeout" to specify the timeout value (in seconds) for the system to wait until all registered takeover IP addresses become active or inactive. When any takeover IP address does not become active or inactive after the timeout value elapses, the resource will notify the error to the cluster. The default value is 60 seconds. Specify the value bigger than 45 seconds.

Note

- NeedAll, InterfaceFilter

You cannot use these attributes. Do not change.

6.7.3.7 Setting Up Procedure Resources

The procedure resource setting is used to register the state transition procedure resource of the products supported by PRIMECLUSTER in userApplication.

[Prerequisites]

To create a procedure resource, you must first create a state transition procedure and register the procedure to the resource database.

For details, see "E.1 Registering a Procedure Resource."

Operation Procedure:

1. Select "Procedure:XXXXXXXXXX" from "turnkey wizard STANDBY."

Example of registering cluster resources of the BasicApplication class to a userApplication:

```
Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                            9) Procedure:BasicApplication(-)
 2) -                              10) Symfoware(-)
 3) SAVE+EXIT                      11) Procedure:SystemState3(-)
 4) -                              12) Procedure:SystemState2(-)
 5) ApplicationName=APP1           13) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)          14) IpAddresses(-)
 7) CommandLines(-)                15) LocalFileSystems(-)
 8) Procedure:Application(-)       16) Gds:Global-Disk-Services(-)
Choose the setting to process:9
```

2. Select "AdditionalProcedureResource" from "Application detail Resource wizard."

   To register multiple resources, execute steps 2 to 4 for each resource.

```
BasicApplication Procedure (ProBApp_APP1:not yet consistent)
1) HELP                            3) REMOVE+EXIT
2) -                              4) AdditionalProcedureResource
Choose the setting to process:4
```

3. The list of cluster resources will appear. Select one.

📘 **Note**

··

If a cluster resource does not appear on this screen, it indicates that the cluster resource has not been registered in the resource database. Confirm whether the cluster resource has been registered on each node of the userApplication, which is designed with "6.7.2 Setting Up userApplication." Register cluster resources if they are not registered. For details on the "clgettree" command, see the manual pages of this command. For details on registering the cluster resource in the resource database, see "E.1 Registering a Procedure Resource."

··

```
1) HELP
2) RETURN
3) -
4) rscl
Choose the resource: 4
```

4. You can change the following on this screen. If necessary, select "SAVE+RETURN" from "Application detail Resource wizard" after that.

   - Timeout value of the state transition procedure

     The default value is 1,800 seconds. If you use a machine that requires more than 1800 seconds for timeout, you need to change the value by selecting "TIMEOUT."

   - Priority within the resource class

     The default value is specified by -p option when the state transition procedure resource is registered with the "claddprocrsc" command. If the -p option is not specified, 65535 is used as the default value. If you register multiple resources of the same class in the cluster application and specify the order of online and offline, change this value by selecting "PRIORITY" from the following screen. The resources will be online in the ascending order, and will be offline in the descending order.

```
Set flags for Procedure resource :
Currently set: TIMEOUT (T1800), PRIORITY (P1)
1) HELP
2) -
3) SAVE+RETURN
4) TIMEOUT
5) PRIORITY
Choose one of the flags:
```

6.7.4 Generate and Activate

This section explains how to execute Generate and Activate. You need to confirm first that the cluster application has been correctly created.

Operation Procedure:

1. Select "Configuration-Generate" from the "Main configuration menu."

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                          10) Configuration-Remove
 2) QUIT                          11) Configuration-Freeze
 3) Application-Create            12) Configuration-Thaw
 4) Application-Edit              13) Configuration-Edit-Global-Settings
 5) Application-Remove            14) Configuration-Consistency-Report
 6) Application-Clone             15) Configuration-ScriptExecution
 7) Configuration-Generate        16) RMS-CreateMachine
 8) Configuration-Activate        17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 7
```

2. Select "Configuration-Activate" from the "Main configuration menu."

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                          10) Configuration-Remove
 2) QUIT                          11) Configuration-Freeze
 3) Application-Create            12) Configuration-Thaw
 4) Application-Edit              13) Configuration-Edit-Global-Settings
 5) Application-Remove            14) Configuration-Consistency-Report
 6) Application-Clone             15) Configuration-ScriptExecution
 7) Configuration-Generate        16) RMS-CreateMachine
 8) Configuration-Activate        17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 8
```

Note

Do not execute "Configuration-Activate" simultaneously on multiple nodes which constitute the cluster.

6.7.5 Registering the Cluster Service of a PRIMECLUSTER-compatible product

If the resources registered to a userApplication are for a PRIMECLUSTER-compatible product, register the resources to the cluster service according to the procedure described below.

Operation Procedure

1. Register the cluster service of the PRIMECLUSTER-compatible product.
   Execute the following command in a node that is part of the cluster system:

   ```
   # /etc/opt/FJSVcluster/bin/clrwzconfig
   ```

2. Register the resources to the cluster service of the PRIMECLUSTER-compatible product. For details, see "F.1 Associating Resources to a Cluster Service Instance."

Note

- If the cluster service and resources for the PRIMECLUSTER-compatible product are not registered, the PRIMECLUSTER-compatible product will not operate correctly. Therefore be sure to register the cluster service and the resources.

- To find out the PRIMECLUSTER-compatible products, see "Table 10-1 PRIMECLUSTER Product List."

6.7.6 Attributes

See

For information on the attributes, see "11 Appendix - Attributes" in the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide."

6.7.7 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.

Exclusive control is established between cluster applications within a single group.

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.

- For information on setting up an exclusive relationship, see "6.7.2.1 Creating Standby Cluster Applications."

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

The example below explains the exclusive relationship between cluster applications within a single group. Even when there are multiple groups, the operation within a single group is the same.

When the job priorities of the cluster applications with an exclusive relationship are different

Cluster applications with the highest job priority take the top priority for startup on the nodes on which the cluster applications with high job priority are running or on the nodes to which the cluster applications with high job priority are failed over. Therefore, cluster applications running with low priorities will be forcibly exited.

The states indicated in the following figure are as follows:

- OPERATING: The cluster application is in the Online state.

- STANDBY: The cluster application is in the Standby state.

- STOP: The cluster application is in the Offline state.

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.
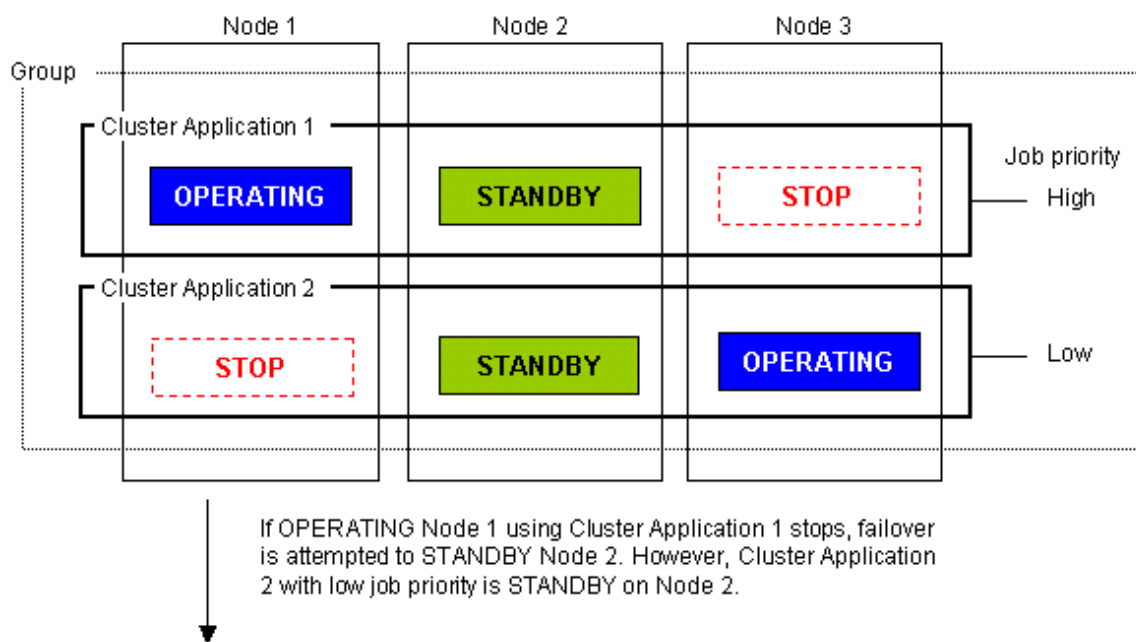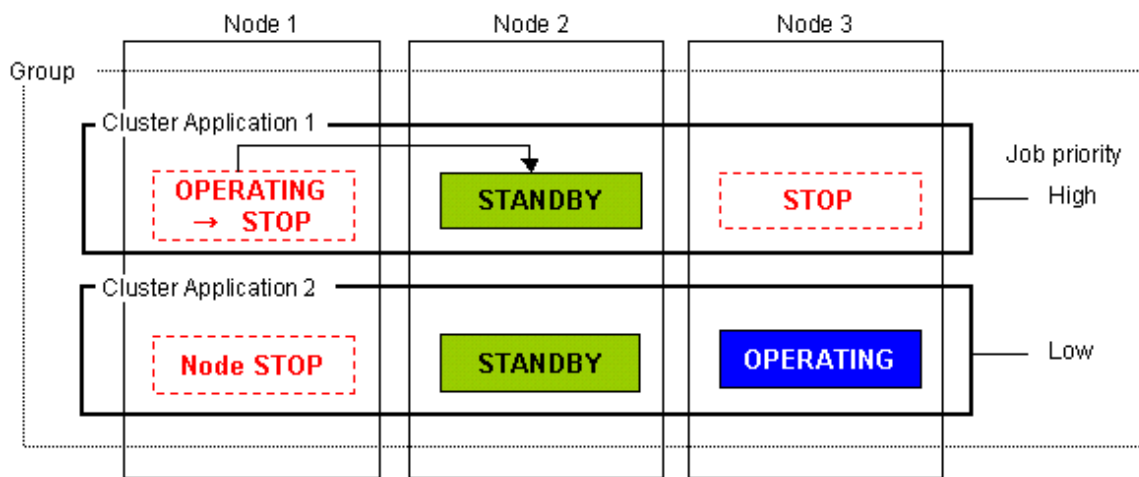
Before failover occurs for Cluster Application 1 whose job priority is high, Cluster Application 2, whose job priority is low, stops.
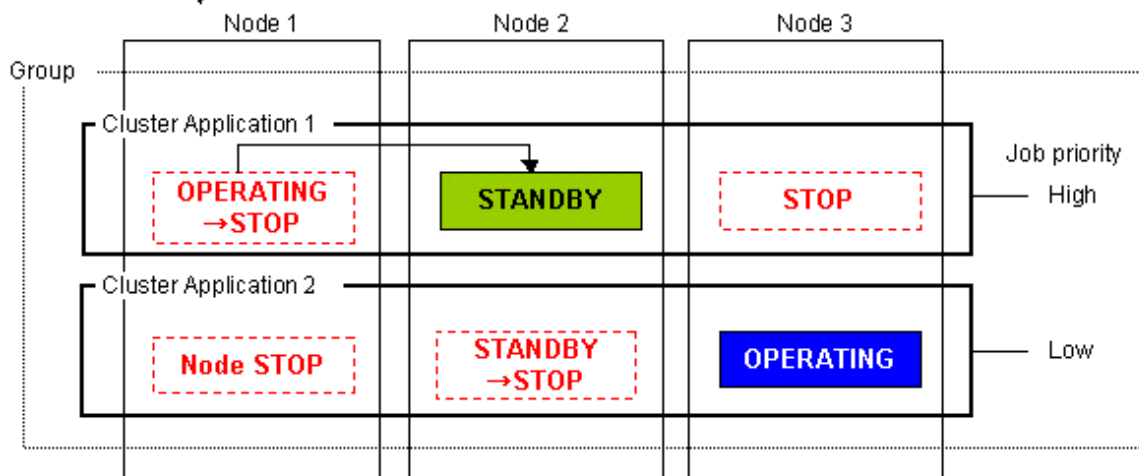
After Cluster Application 2 stops on Node 2, failover occurs for Cluster Application 1 whose job priority is high.
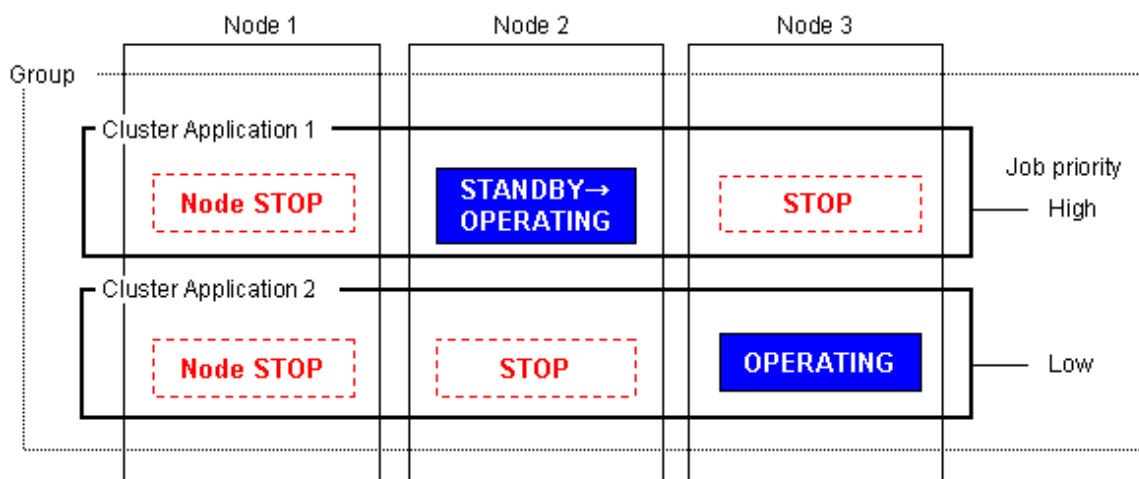
Moreover if Node 2 using Cluster Application 1 stops, failover is attempted to Node 3. However, Cluster Application 2 is OPERATING on Node 3.
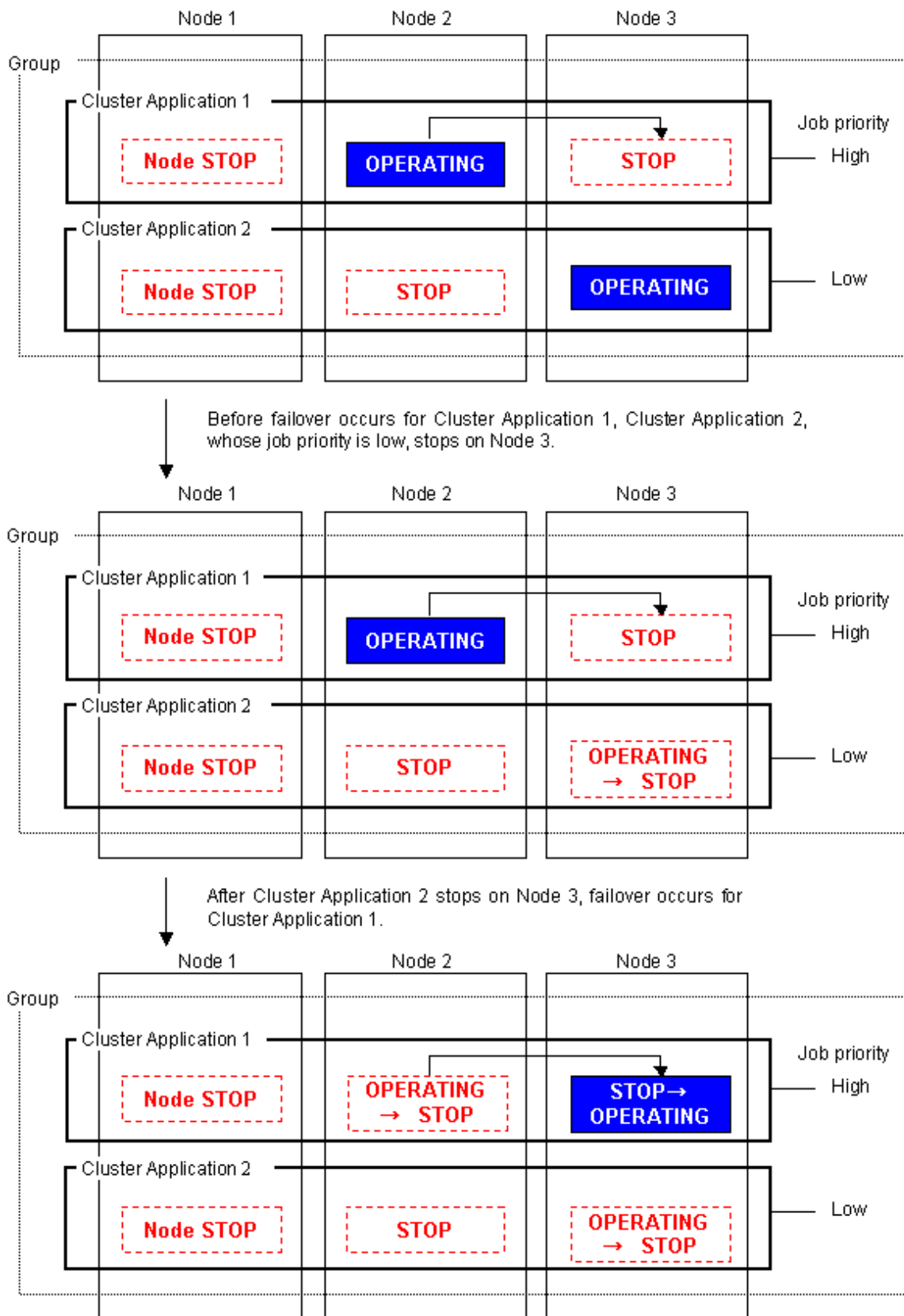
Failover of the cluster application with a low job priority

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.
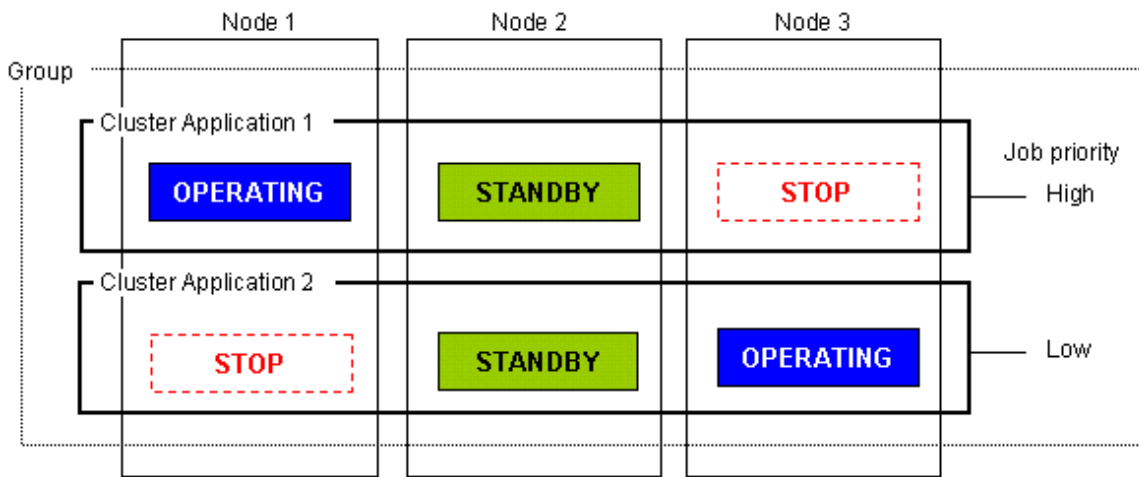
If OPERATING Node 3 using Cluster Application 2 stops, failover is attempted to STANDBY Node 2. However, Cluster Application 1 with high job priority is STANDBY on Node 2.
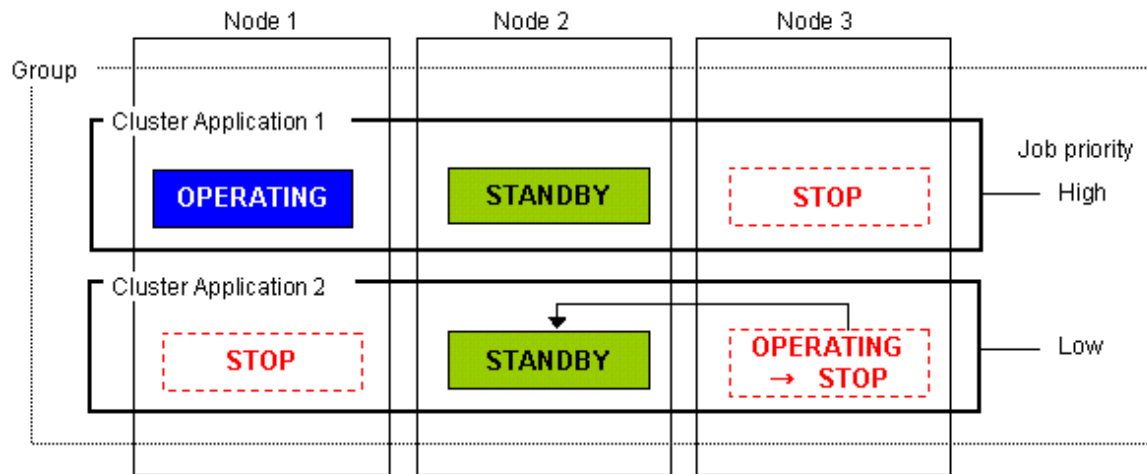
Before failover occurs for Cluster Application 2 whose job priority is low, Cluster Application 1, whose job priority is high, stops on Node 2.
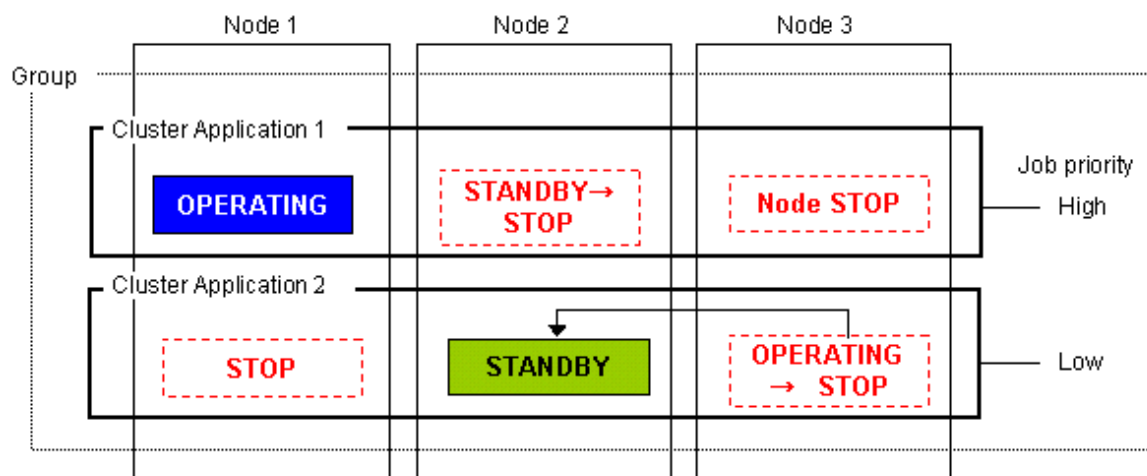
After Cluster Application 1 stops on Node 2, failover occurs for Cluster Application 2 whose job priority is low.

	Node 1	Node 2	Node 3	
Group				
Cluster Application 1	OPERATING	STOP	Node STOP	Job priority — High
Cluster Application 2	STOP	OPERATING	Node STOP	Low
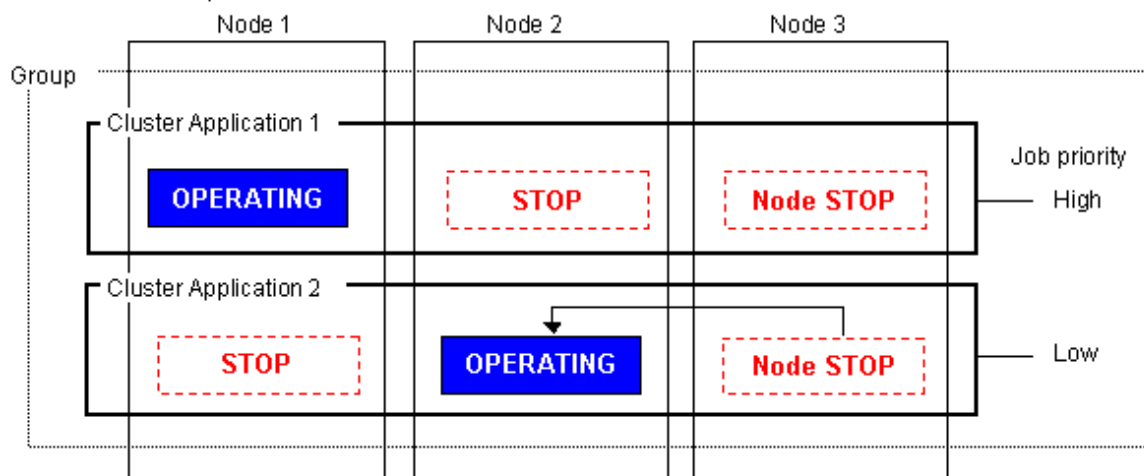
Moreover if Node 2 using Cluster Application 2 stops, failover is attempted to Node 1. However, Cluster Application 1, whose job priority is high, is OPERATING on Node 1.

	Node 1	Node 2	Node 3	
Group				
Cluster Application 1	OPERATING	Node STOP	Node STOP	Job priority — High
Cluster Application 2	STOP	OPERATING →STOP	Node STOP	Low
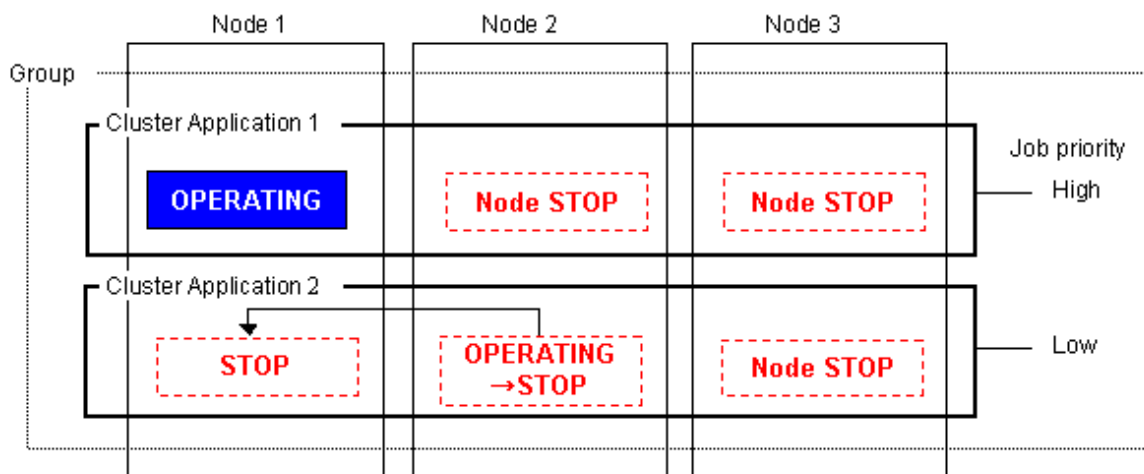
Since Cluster Application 1, whose job priority is high, is OPERATING on Node 1,Cluster Application 2, whose job priority is low, does not execute failover.

	Node 1	Node 2	Node 3	
Group				
Cluster Application 1	OPERATING	Node STOP	Node STOP	Job priority — High
Cluster Application 2	STOP	Node STOP	Node STOP	Low
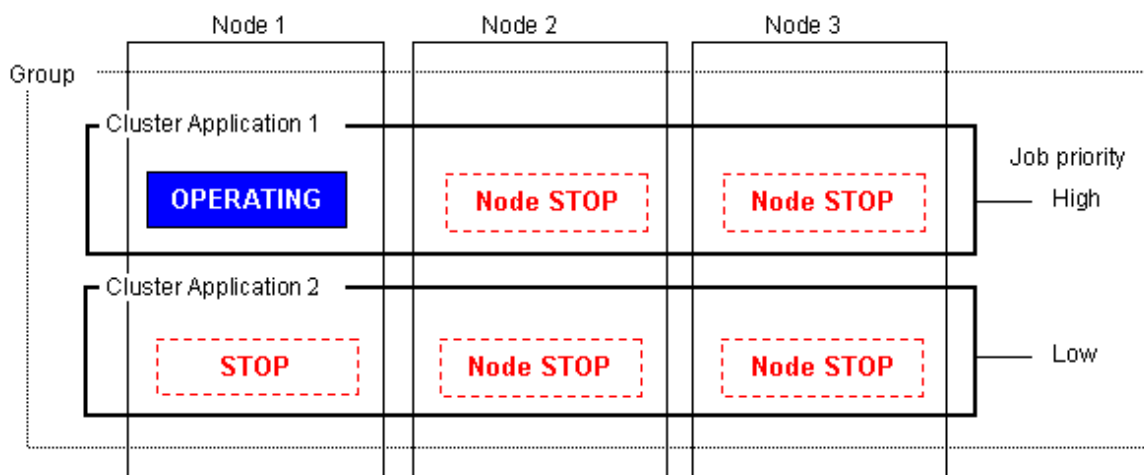
When the job priorities of cluster applications with an exclusive relationship are the same

The operation of the cluster applications that are already running will be continued. On the node on which cluster applications are already running, cluster applications that subsequently start up will be stopped.

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
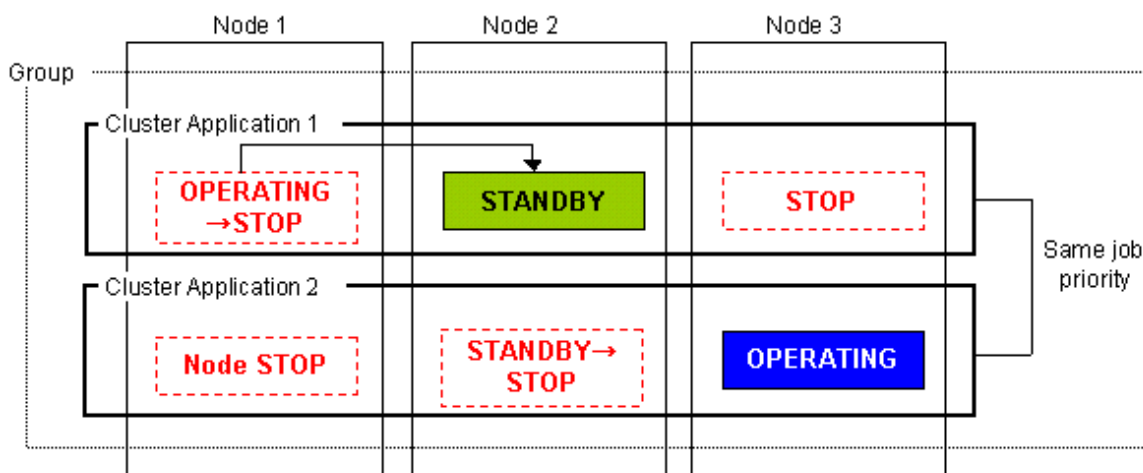
Cluster Application 2 continues to run, and Cluster Application 1
does not execute failover.

	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
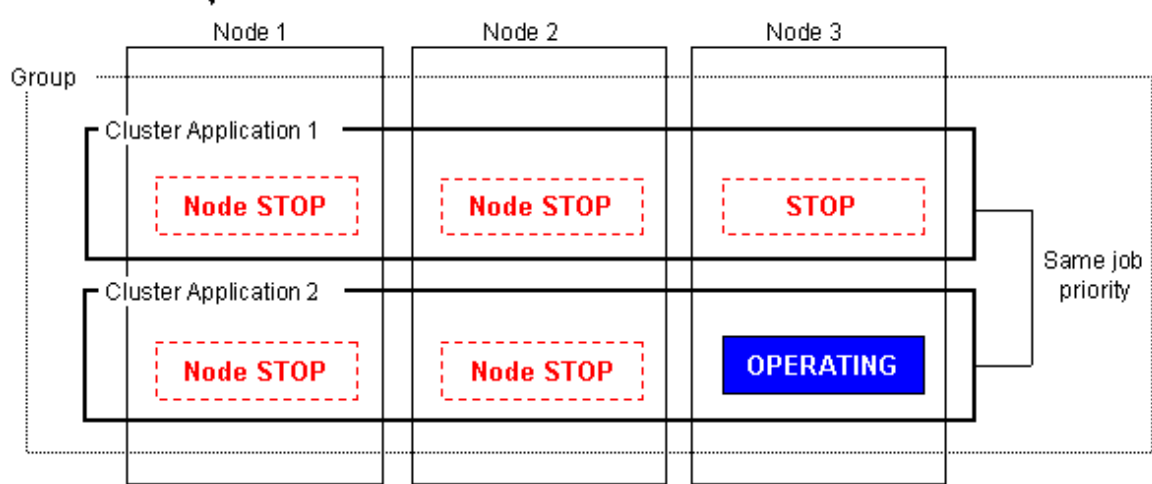
6.8 Setting Up the RMS Environment

When using RMS, you need to check the "A.8.1 RMS Setup Worksheet" and change the following environment variable to the value corresponding to the configuration setup.

- RMS shutdown wait time (RELIANT_SHUT_MIN_WAIT)

  The default value of the environment variable is 900 (seconds) in this version.

See

For information on how to check and change the RMS environment variables, see "2.9 Environment variables" and "12 Appendix - Environment variables" in "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide*."

6.9 Checking the Cluster Environment

After making the settings for the PRIMECLUSTER environment, use the PRIMECLUSTER environment checking tool to make sure there are no mistakes in the settings.

For information on usage of and contents checked by the PRIMECLUSTER environment checking tool, refer to the following file on CD of the PRIMECLUSTER product package:

Tool/clchkcluster/readme_en.txt

Part 3   Operations

Chapter 7 Operations

This chapter describes the functions managing PRIMECLUSTER system operations. They monitor operation statuses for PRIMECLUSTER system and operate PRIMECLUSTER system according to its operation statuses and so on. Also, notes for operating PRIMECLUSTER system are described.

The following user groups are allowed to do each specific operation:

Operation	Target
Referring 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

  The states of the resources that are managed with the cluster resource management facility are displayed.

- 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] **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.
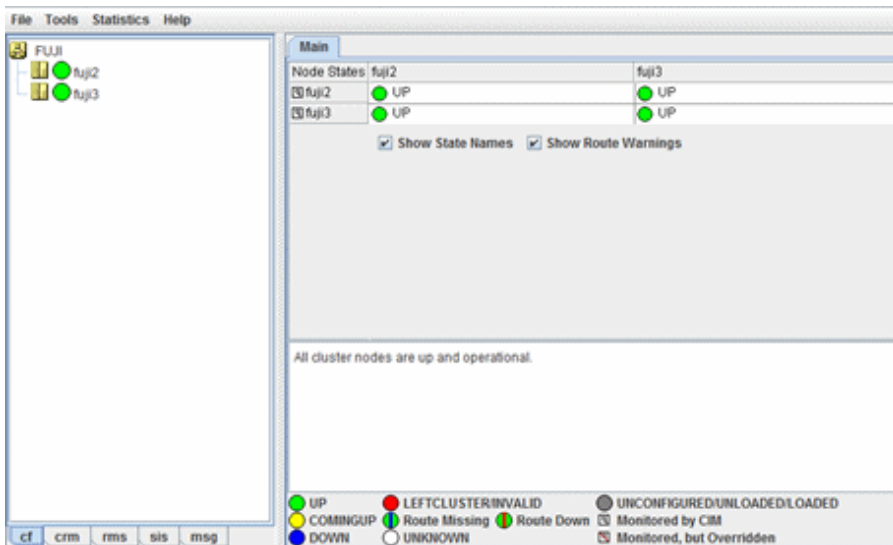
See

. . . .

For details, see "5 GUI administration" in "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

. . . .

Note

. . . .

The node states may be displayed as Unknown. In this case, exit the Web-Based Admin View screen and restart. If the node states in the CF main window of Cluster Admin are still displayed as Unknown, check the node states by using cftool -n.

. . . .

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.

In the screenshot: window title "Cluster Admin", "PRIMECLUSTER Cluster Admin", menu "File Tool Help" with "menu" label arrow. Tree view showing:
- RC
- SHD_RC
  - SHD_Disk18
  - SHD_Disk21
- node1
  - eth1
  - eth2
  - "CRM tree view"
- node2
  - eth1
  - eth2
  - sdb
  - sdc

Right panel columns 属性 / 属性値, NodeID / 0, "CRM table view".

Tabs: cf crm rms sis msg with "crm tab" label. "Java Applet Window".
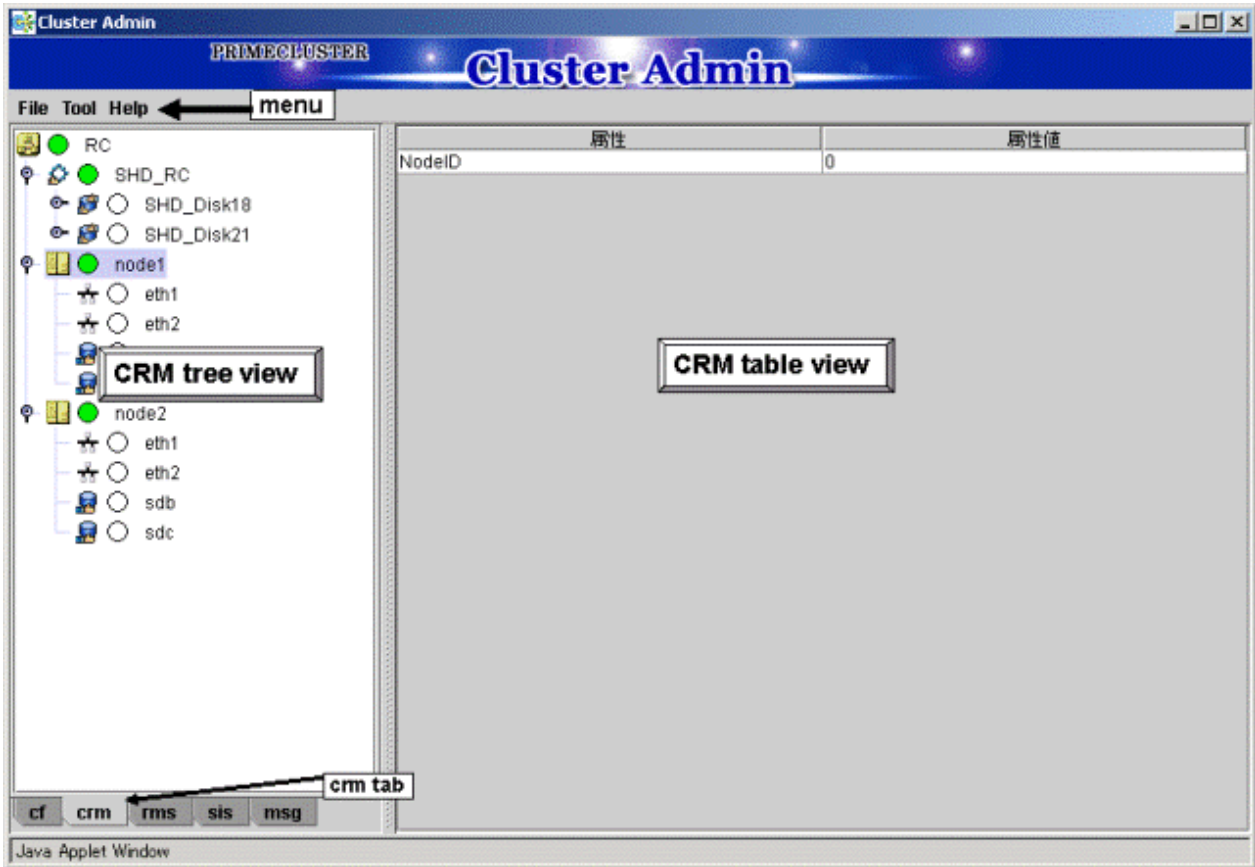
See

The CRM main window is a screen of the cluster resource monitoring facility. See "crm" in "4.5.3 Cluster Admin Functions."

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
	Shared disk device
	IP address
	Network interface
	Takeover network

Icon	Resource
	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.

🛈 Note

. . . .

- If a node is LEFTCLUSTER in CF, it becomes ON in the CRM tree view.

- After you execute the "clinitreset" command to initialize the resource database, the resource states displayed in the CRM main window differ according to the actual resource states. Execute "clinitreset", reboot the nodes, and then close the browser and redisplay the screen. Execute the "clinitreset" 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 on 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
Request Resource activation	*Tool - Start*	SDX_DC (*2)	wvroot clroot cladmin
Request Resource deactivation	*Tool - Stop*	SDX_DC (*2)	wvroot clroot cladmin
Exit Cluster Admin screen	*File - Exit*	All No selection	All
View Help	*Help - Content* (*3)	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 Only the disk resources that are registered to Global Disk Services are enabled.

*3 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."

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.

 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"

  - "11.10 Cluster Resource Management Facility Messages"

7.1.2.2 Detailed Resource Information

This section describes the resource attributes that are defined in the CRM main window.

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).
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.
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.
	WebView	This item indicates the network interface to be used by Web-Based Admin View.
		If Web-Based Admin View is being used, USE is set. If not, UNUSE is set.
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.

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
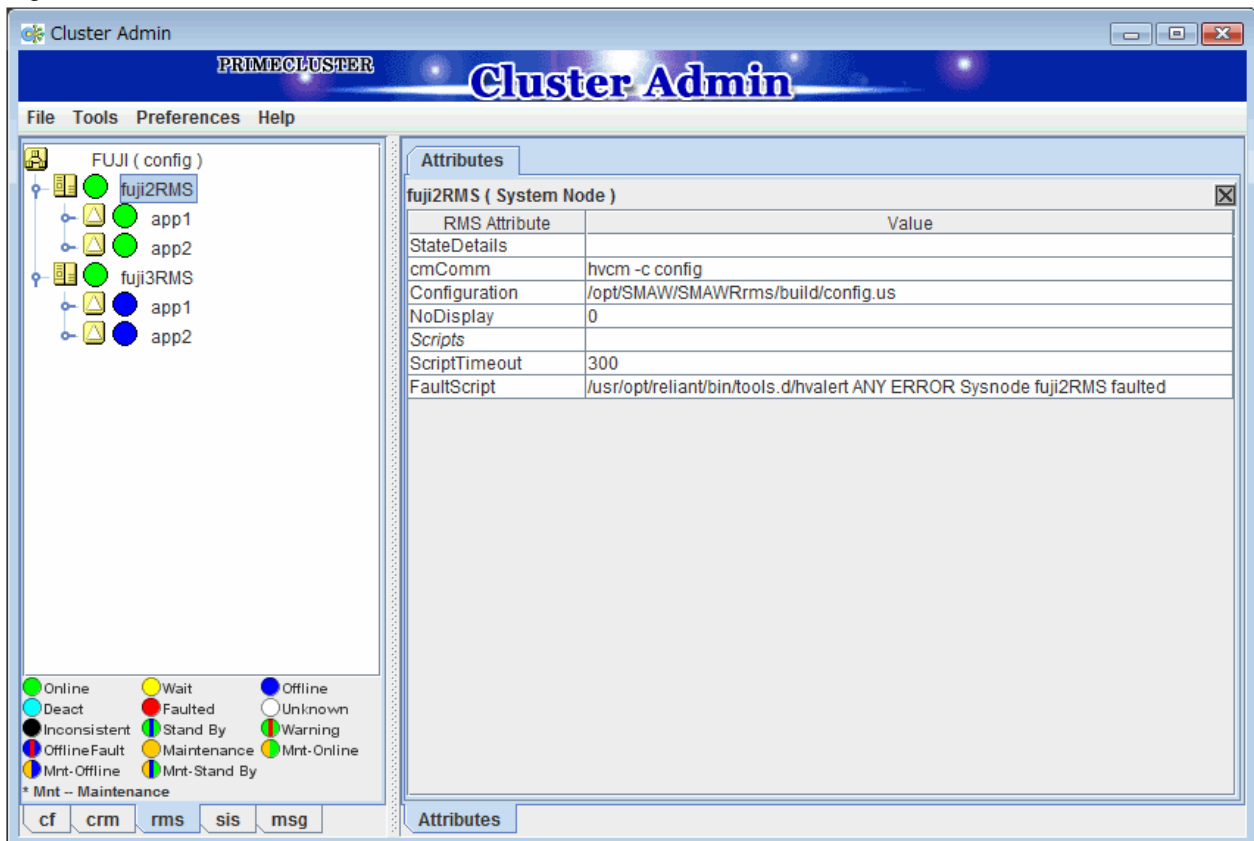
Meanings of object icons

Icon	Meaning
	Represents the cluster.

Icon	Meaning
	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."

Node state display

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.

Note

The node states in the RMS main window of Cluster Admin may be displayed as Unknown. In this case, exit the Web-Based Admin View screen and restart. If the node states in the RMS main window of Cluster Admin are still displayed as Unknown, check the node states by using hvdisp -a.

State display of other objects

The following state icons are displayed with object icons that fall under parent objects.

Icon	Icon color	Outline	Details
Parent object	Green	Online	Object is enabled and ready for use.
	Green with vertical red lines	Warning	Object has exceeded some warning threshold.

Icon	Icon color	Outline	Details
Child object **Controller object** **Leaf object**	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 red lines	OfflineFault	Object is Offline, but a fault has occurred before and is not cleared yet.
	Orange	Maintenance	Object is in maintenance mode.
	Orange in the left and green in the right	Maintenance-Online	Object is in maintenance mode and must be Online when exiting maintenance mode.
	Orange in the left and blue in the right	Maintenance-Offline	Object is in maintenance mode and must be Offline when exiting maintenance mode.
	Orange in the left and green in the right with vertical blue lines	Maintenance-Stand By	Object is in maintenance mode and must be Stand By when exiting maintenance mode.

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.
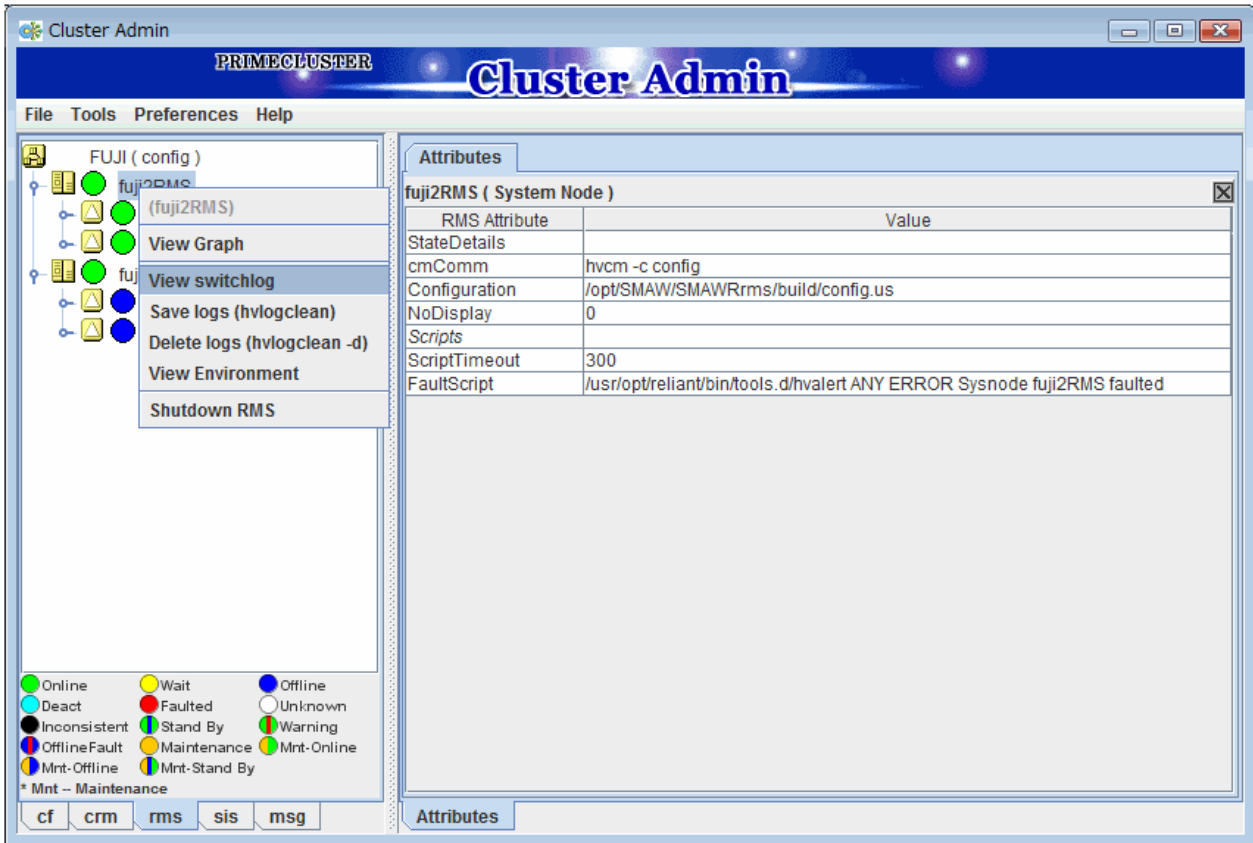
7.1.3.2 Configuration information or object attributes

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*.
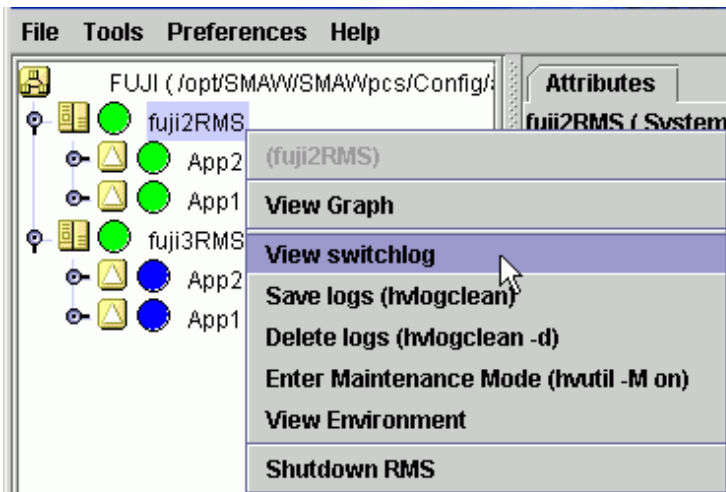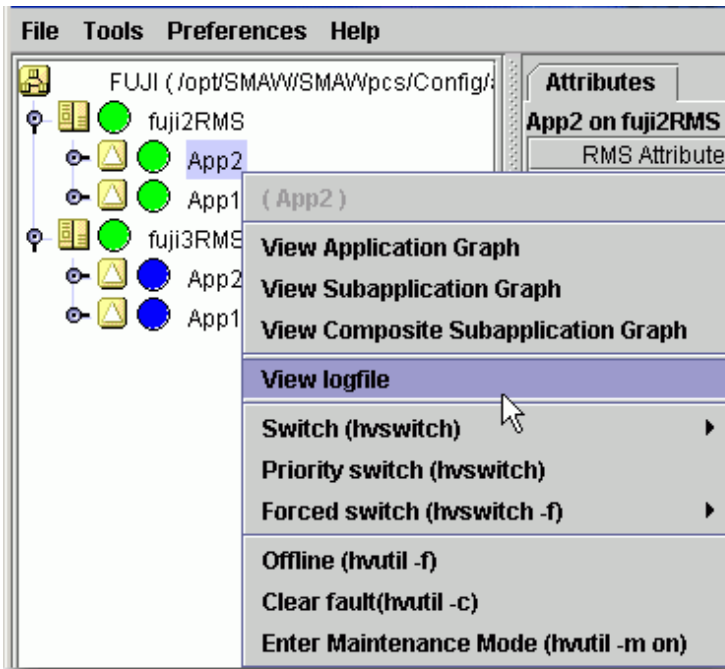
7.2 Operating the PRIMECLUSTER System

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
. . . .

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

See "7.1.1 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 *Tool* pull-down menu on the RMS main window or right-click the system node, and then select the shutdown mode on the screen that appears next.

    1. Choose either a specific node or all the nodes.

    2. Choose to stop all cluster applications, leave the applications running, or forcibly shutdown the node(s).

See

See "7.1.3 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

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 Resource Information" and "7.3.3 Concurrent Viewing of Node and Cluster Application States."

Note

To start a cluster application manually, check that the cluster application and resources under it are stopped on all nodes other than the node on which the cluster application is to be started. You can check whether they are stopped by the Offline or Standby state. With the state other than Offline or Standby, they may be running. In this case, stop them and then start the cluster application on the target node.

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.

📖 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 Resource Information" and "7.3.3 Concurrent Viewing of Node and Cluster Application States."

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.

📖 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 Resource Information" and "7.3.3 Concurrent Viewing of Node and Cluster Application States."

7.2.2.4 Bringing Faulted Cluster Application to Available State

The procedure for bringing a Faulted cluster application to available state is described below.

Operation Procedure:

Right-click on the cluster application object in the RMS tree, and select *Clear Fault*.

📖 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 Resource Information" and "7.3.3 Concurrent Viewing of Node and Cluster Application States."

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 "Clear Wait & shutdown (hvutil -u)" from the menu.

. . . .

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.
. . . .

📘 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 Resource Information" and "7.3.3 Concurrent Viewing of Node and Cluster Application States."
. . . .

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, it cannot be switched.

Note that cluster nodes and resources are monitored during maintenance mode. In this case, when the resource state is changed, the resource state of the cluster application that is viewed on the RMS tree is also changed.

Maintenance mode is only allowed when PRIMECLUSTER products specified.

For using maintenance mode, see "7.3 Using maintenance mode" in "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide".

📝 Note
. . . .

Please note the following for using maintenance mode.

- Perform maintenance mode to the cluster application of the standby operation containing resources for which the maintenance is necessary.

- Since the resources for which the maintenance is necessary during the operation are not contained, it is not necessary to make the cluster application of the scalable operation into maintenance mode.

- To start maintenance mode, a cluster application must be in the Online, Standby, or Offline state.

- To exit maintenance mode, a cluster application and each resource must be returned in the same state before starting maintenance mode.

- Do not stop RMS or the system with cluster applications in maintenance mode. Be sure to exit maintenance mode of all cluster applications before stopping RMS or the system.

For details, see "7.3.2 Maintenance mode operating notes" in "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide" or "Restrictions during maintenance mode" in "PRIMECLUSTER Reliant Monitor Services (RMS) Reference 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.
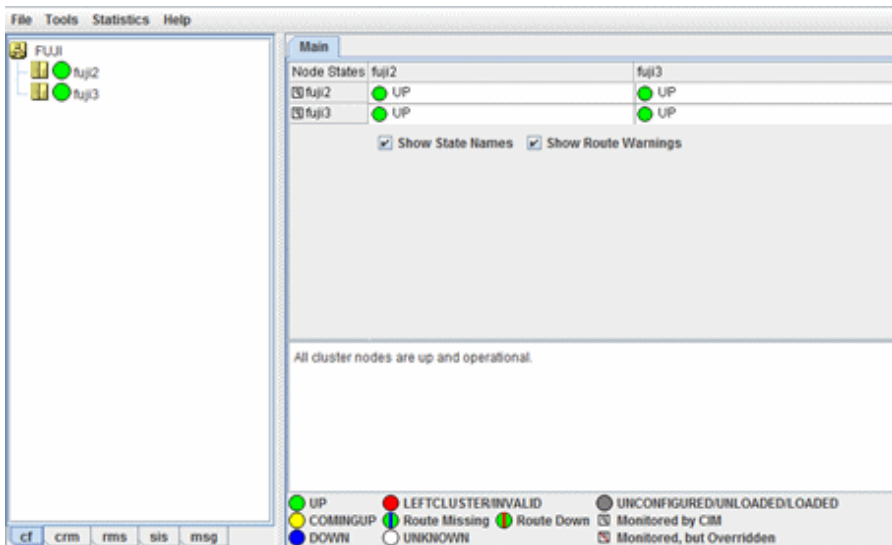
The node indicates one of the following states:

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.
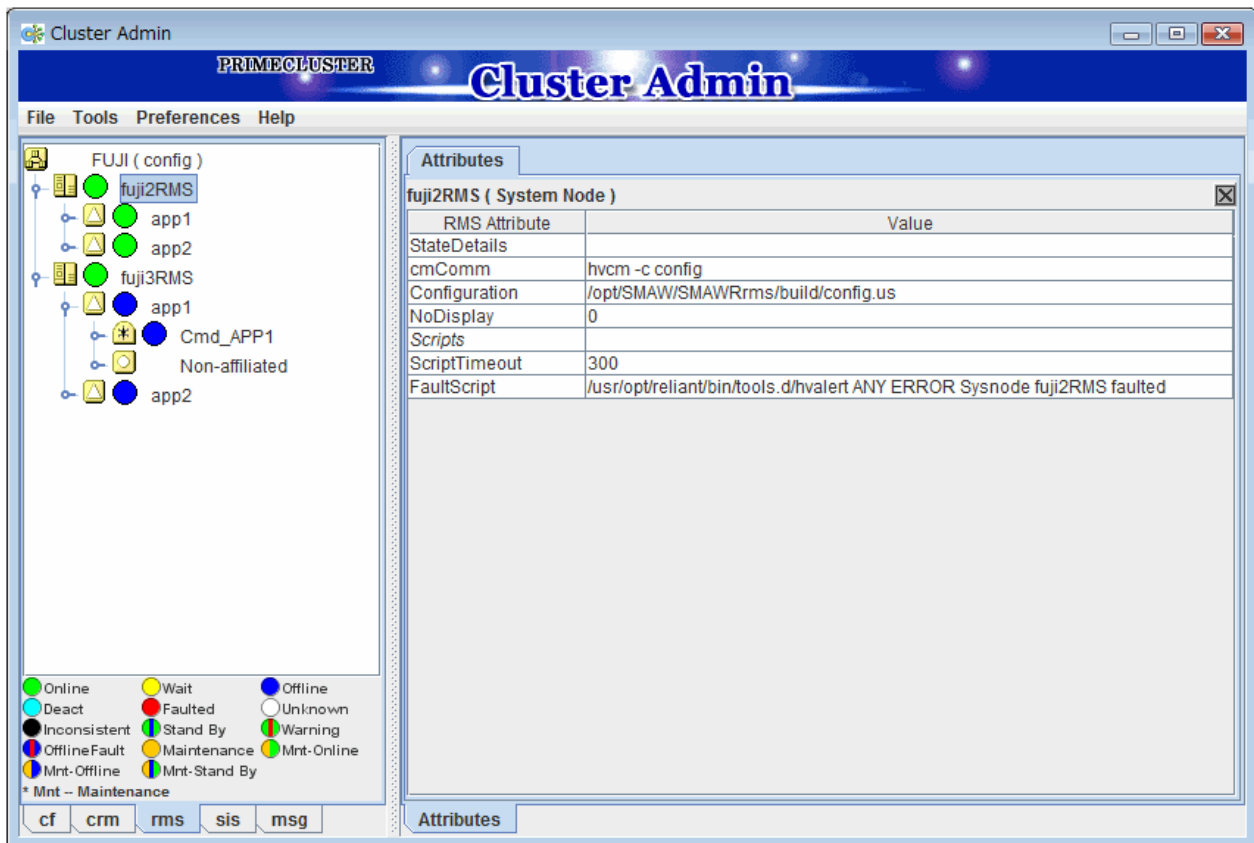
The cluster application shows one of the following states:

- Online

- Offline

- Unknown

- Wait

- Deact

- Inconsistent

- Faulted

- Stand By

- Warning

- OfflineFault

- Maintenance

- Maintenance-Online

- Maintenance-Offline

- Maintenance-Stand By

See

See "State display 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
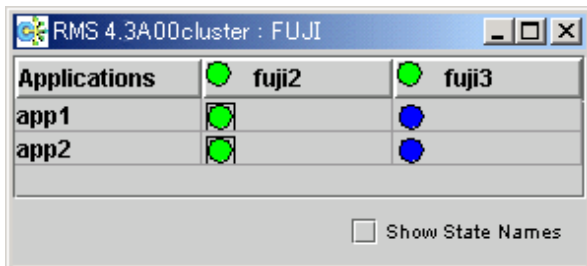
The first line shows the names of the nodes that RMS is managing (fuji2 and fuji3 in the example above). To the left of each node name is a state display icon that shows the state of that node.

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 "6.1 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.
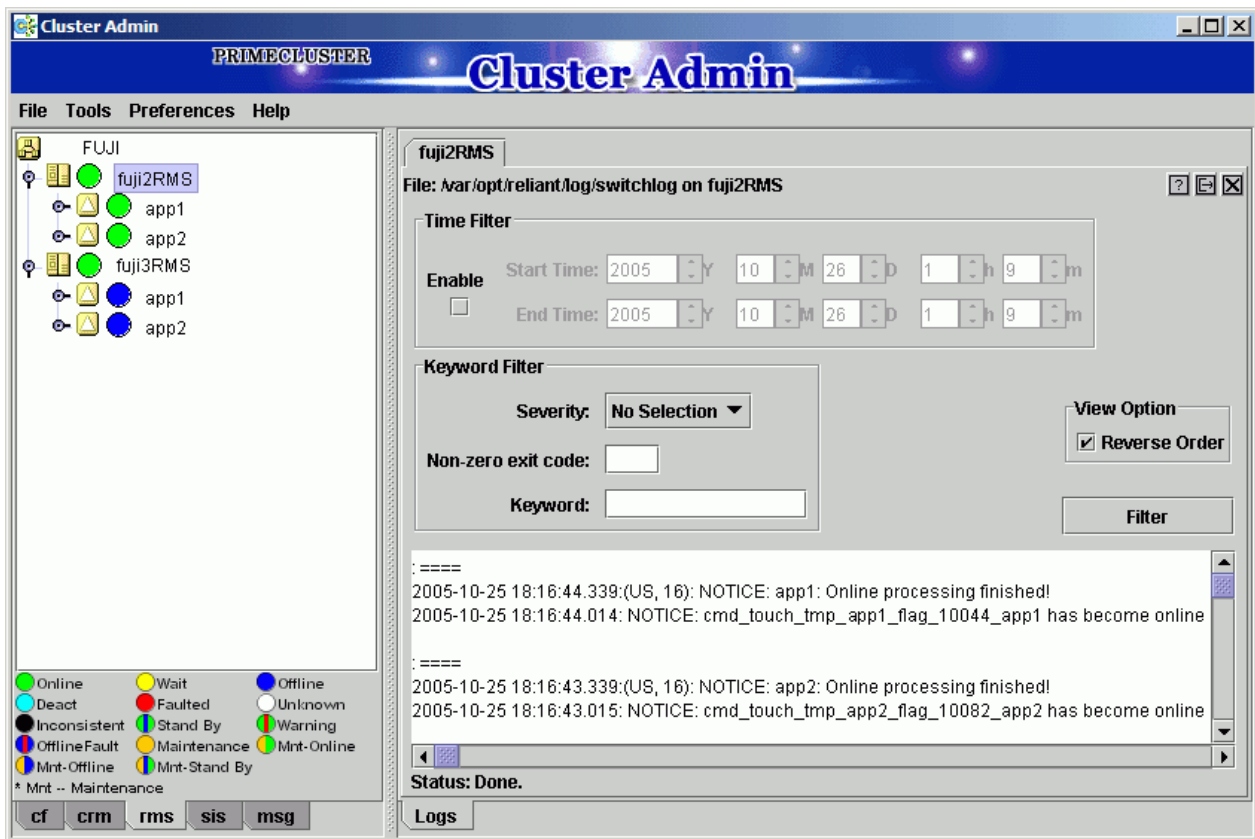
7.3.4.2 Viewing application logs

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.
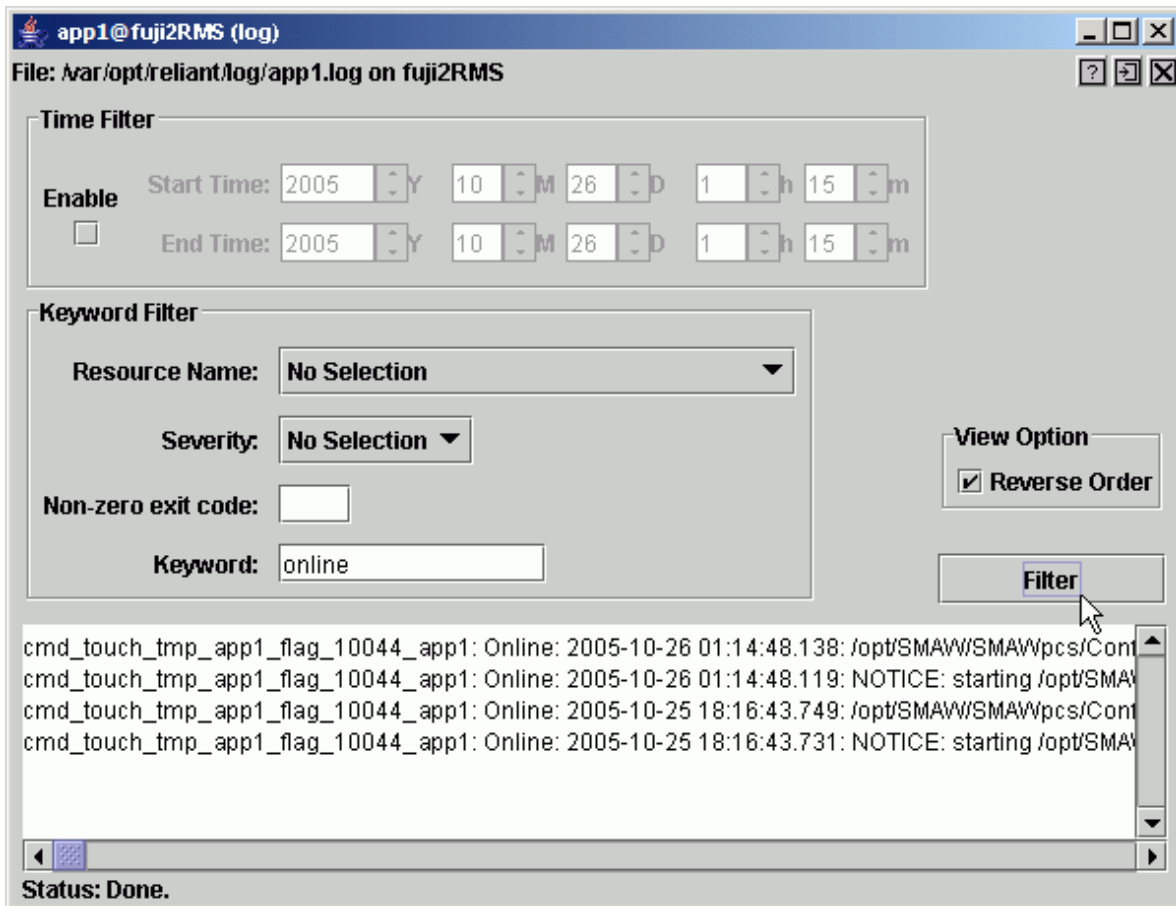
Information

The following display formats are enabled for the log. For details, see "6.4 Viewing RMS log messages" 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 Resource Information

Use RMS graphs to display detailed resource 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
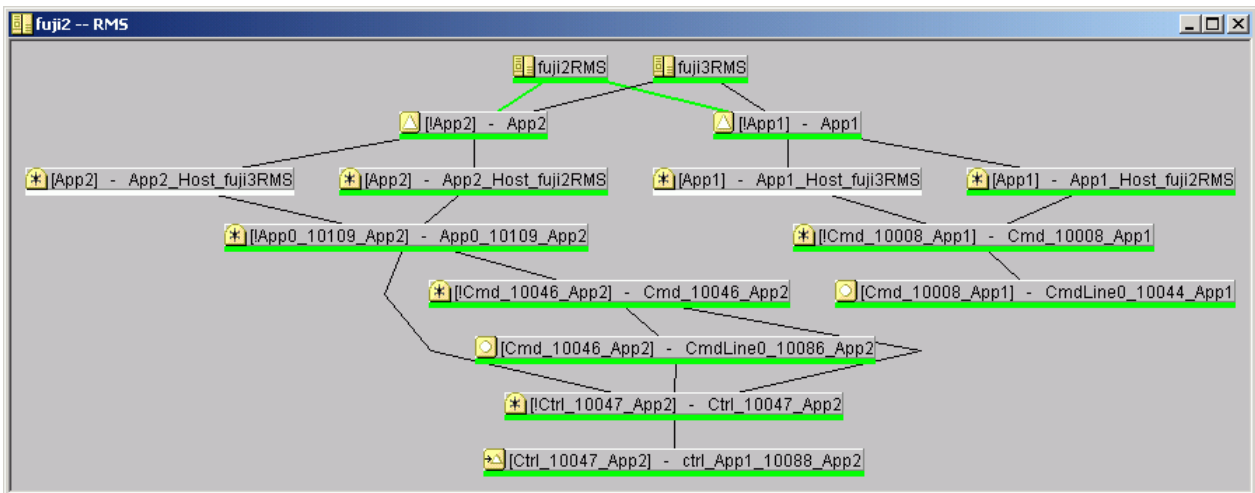
If you left-click the target object, the attributes of the object will be displayed on a pop-up screen.
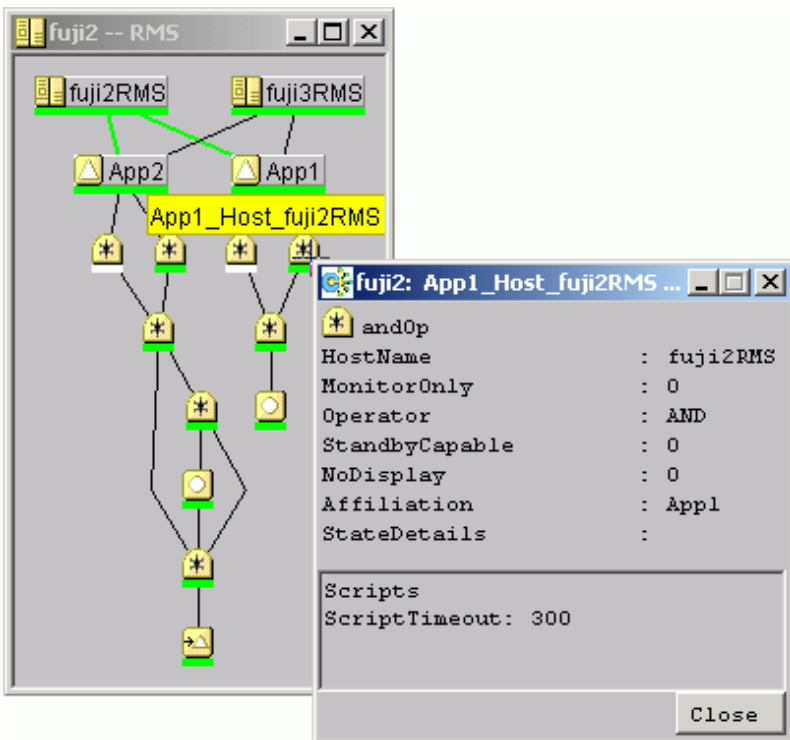
 See
. . . .

See "6.2 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.
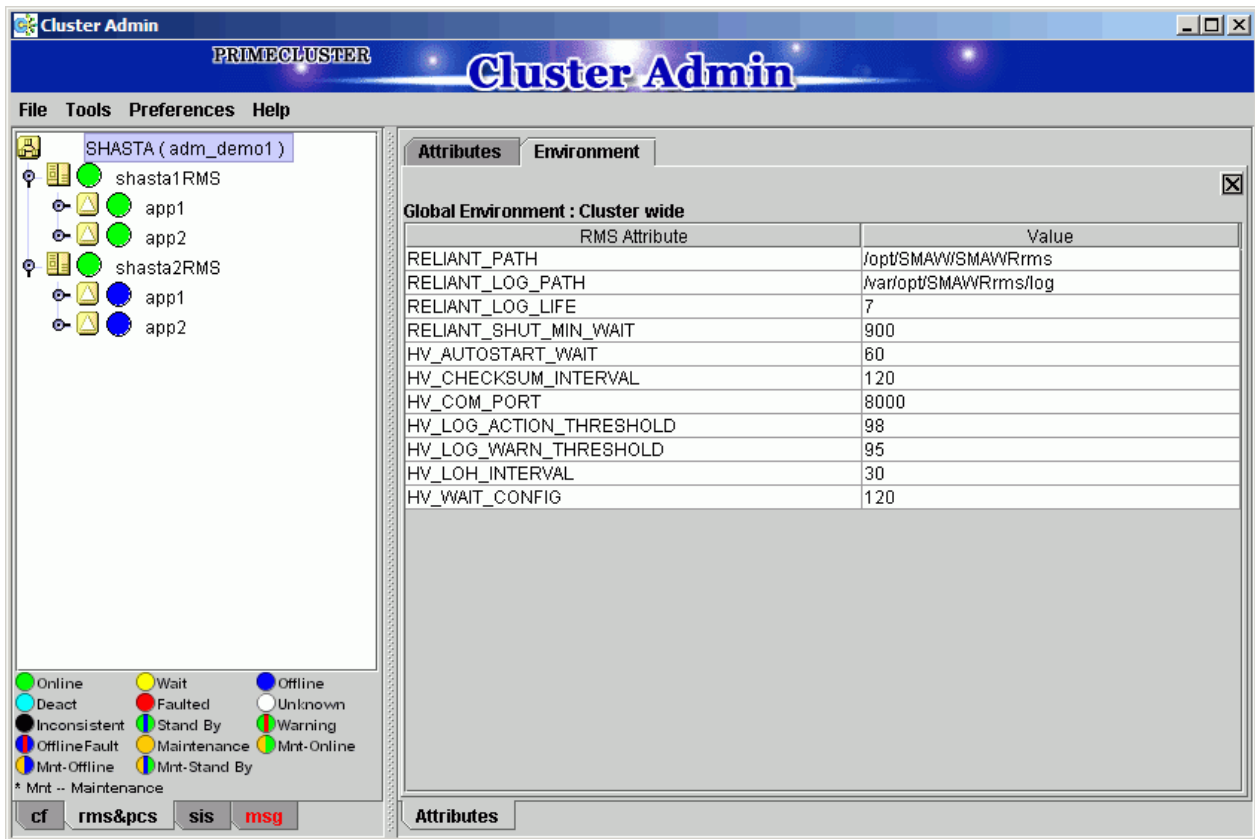
Right-click a node in the RMS tree, and select *View Environment*. The local variables are displayed.
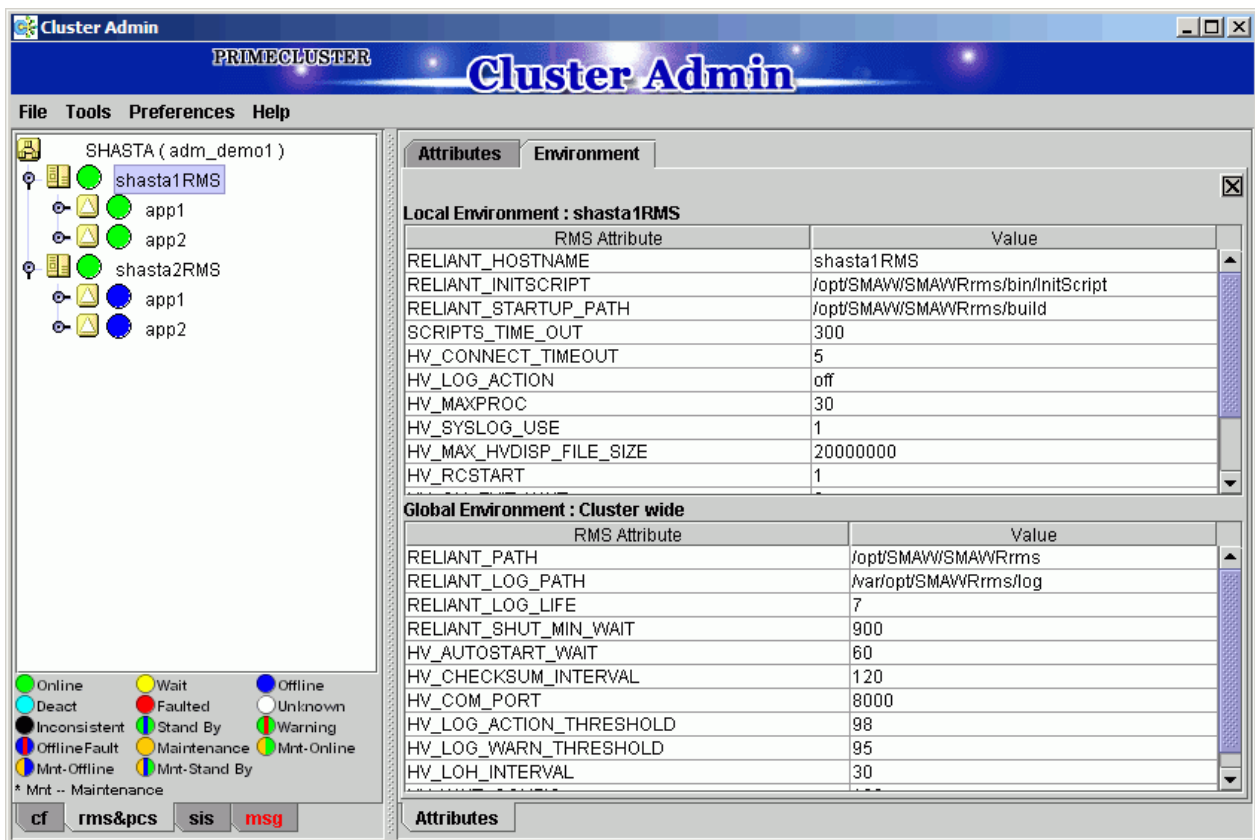
7.3.7 Monitoring Cluster Control Messages

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 console screen. Based on this message, you need to identify the faulted resource using the CF 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 failed.

7.4.1 Corrective Action in the event of a resource failure

This section describes the corrective actions to take when a resource failure occurs.
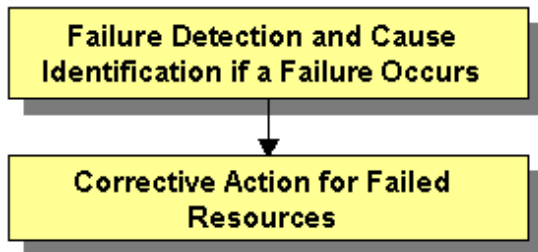
7.4.1.1 Failure Detection and Cause Identification if a Failure Occurs

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 (g) below are relevant to the "Failure confirmation features list" given below:

Failure detection

Normally, the RMS main window (a) 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:

    - The color of the icons in the RMS main window (a) changes.

    - A message is output to the msg main window (c), Syslog(f), and the console (g).

- 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:

    - A message is output to Syslog(f) and the console (g).

- 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(f) and the console (g). By responding to the operator intervention request message, you can start the userApplication.

    For further details, see "D.4 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 list" to detect the failure.

Cause identification

You can also use the function that detected the failure and the features listed in "Failure confirmation features list" below to identify the faulted resource that caused the failure.

Failure confirmation features list

	Failure confirmation features	Manual reference
(a)	**RMS main window** The RMS tree and the RMS cluster table can be used from this screen.	7.1.3 RMS Main Window
(b)	**CF main window** The CF tree can be used from this screen.	7.1.1 CF Main Window
(c)	**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.	-
(d)	Application log	7.3.4.2 Viewing application logs
(e)	switchlog	7.3.4.1 Viewing switchlogs
(f)	Syslog	-
(g)	**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
(h)	GDS GUI	"PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

📝 Note

. . . .

Console

- The operator intervention request messages (message numbers: 1421, 1423), incurred when RMS is not started on all nodes, are displayed only when yes(1) is set for the AutoStartUp attribute of the userApplication. For information on the userApplication attribute, see"11 Appendix - Attributes" in the *"PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide."*

- The operator intervention request messages (message numbers: 1422, 1423) and the error resource messages incurred after a resource or system error occurs are displayed only when yes(1) is set for the PersistentFault attribute of the userApplication. For information on the userApplication attribute, see "11 Appendix - Attributes" in the "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide.*"

- The operator intervention request and error resource messages are displayed by using the "clwatchlogd" daemon to monitor switchlog. You need to send the SIGHUP signal to clwatchlogd when you change the value of RELIANT_LOG_PATH that is defined in the "hvenv.local" file. When clwatchlogd receives this signal, clwatchlogd acquires the latest value of RELIANT_LOG_PATH. After you change RELIANT_LOG_PATH, you must start RMS.

Note

When you check the message of a resource failure, a resource with the "MONITORONLY" attribute may be in the fault state even if the cluster application is in the Offline state. Check whether there are any resources in the fault state. Especially, check that Fsystem resources are not in the fault state.

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 Reliant Monitor Services (RMS) Reference Guide."*

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.

(Example) Symfoware

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 Available State."

7.4.1.3 Recovery of Failed Cluster Interconnect

The following problems can cause cluster interconnect failures.

- Hardware error

  - Error on LAN card, hub, or cable

  - Connection error

- Network configuration error

  - Configuration error on IP address, netmask, or routing information, etc.

Contact your system administrator on the network configuration error. The following section describes how to fix hardware related errors.

If any heartbeat error on the cluster interconnect is detected, either of the following messages will be output to the /var/log/messages file.

```
"CF: Problem detected on cluster interconnect NIC_NAME to node NODE_NAME: missing heartbeat replies.
(CODE)"
"CF: Problem detected on cluster interconnect NIC_NAME to node NODE_NAME: ICF route marked down.
(CODE)"
```

"*NIC_NAME*" indicates the network interface card on which the error is detected.

"*NODE_NAME*" indicates the CF node name on which the error is detected.

"*CODE*" indicates the necessary information to determine the cause.

When either of the above messages is output to the file, follow the steps below.

Corrective action

1. Determining the failed node

   Confirm that each device is working properly. You can also use the ping command to determine the failed node and its location.

   📝 **Note**

   When an error on the entire cluster interconnects (all interconnects for every node) occurs, the cluster system forcibly shut down all nodes except one which has the highest survival priority.

   For details on survival priority, see "5.1.2 Setting Up the Shutdown Facility."

   If an error on an active node (e.g. LAN card error of a node on which an active cluster application resides) occurs, you must stop the node before fixing it. To minimize the down time, make sure to follow the steps below before performing "Step 2. Performing maintenance tasks".

   1. Stopping a node in the "Online" state

      Before performing the maintenance task, stop the node on which "Online" cluster application resides.

   2. Starting the forcefully terminated node

      Start the node which was forcefully terminated by the cluster system and make the cluster application back to the "Online" state. For details on how to start a cluster application, see "7.2.1.1 Starting RMS."

      Be sure to check that the node, which is described in Step 1. Stopping a node in the "Online" state, is completely stopped before performing this step.

2. Performing maintenance tasks

   After determining the cause of the error, perform the following maintenance task depending on the category of error.

   📝 **Note**

   For a LAN card error, the failed node must be stopped to perform the maintenance task.

   For an error on cables or hubs, you can perform the maintenance task with the node being active.

   - When the error was caused by your LAN card or cable

     If the cable is unplugged, plug in properly.
     If the cable is properly plugged, your LAN card might be the cause. Contact your Fujitsu systems engineer (SE).

   - When the error was caused by a hub

     If the power is off , push the power button.
     If the power is on, there is a possibility the hub is broken down. Contact your Fujitsu systems engineer (SE).

3. Recovery

   To recover the partial failure of the cluster interconnect, skip to "Step 2. Cluster interconnect recovery" below.

   1. Starting all nodes

      Start all nodes.

   2. Cluster interconnect recovery

      Use the ping command to confirm if nodes can communicate each other through the failed cluster interconnect.

After confirming that the cluster interconnect is recovered successfully, clear the "Faulted" state of the cluster application as necessary. For details on the operation, see "7.2.2.4 Bringing Faulted Cluster Application to Available State."

7.4.2 Corrective Action in the event of the LEFTCLUSTER state when the virtual machine function is used

If the host OS enters the panic state or hangs up when the virtual machine is used, the LETCLUSTER state may occur. This section describes the corrective actions in this case.

7.4.2.1 When the host OS enters the panic state

1. When the host OS enters the panic state, and the host is restored after that, go to the step 3.

2. Check the state of the host OS and restore the host OS.

3. If a failed node (guest OS) is in the LEFTCLUSTER state, perform the procedure in "6.2 Recovering from LEFTCLUSTER" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*."

4. Check that the node joins the cluster.

7.4.2.2 When the host OS hangs up

1. Check that the node (guest OS) is actually stopped. If it is running, stop it.

2. Check the state of the host OS and restore the host OS.

3. If a failed node is in the LEFTCLUSTER state, perform the procedure in "6.2 Recovering from LEFTCLUSTER" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*."

4. Check that the node joins the cluster.

7.5 Notes on Operation

This section describes notes when operating PRIMECLUSTER system.

Do not stop RMS while RMS is being started

Heartbeats between nodes are interrupted and the node where RMS is stopped may be forcibly shut down.

Stop RMS after completing its startup processing (completing the state transition processing of a cluster application).

Use hvshut -a to stop RMS on all nodes simultaneously

When executing the hvshut -l command on all nodes simultaneously, RMS will not be stopped and occasionally the timeout and hvshut command times out or hangs up.
When stopping RMS on all nodes, execute the hvshut -a command on any one of the nodes that configures a cluster system.
When stopping RMS on each node, execute the hvshut -l command on the node which stops RMS.

If mistakenly executing the hvshut -l command on all nodes simultaneously and the hvshut command times out, stop or reboot all the nodes. In addition, if the hvshut command hangs up, stop RMS forcibly using the hvshut -f command, and then stop or reboot all the nodes.

Do not stop operating system services after stopping RMS

Even if RMS is stopped using the hvshut command, other PRIMECLUSTER services (CF, SF, CRM, and so on) run.

Therefore, if you stop or reboot operating system services to modify its information (such as network information), heartbeat monitoring by CF fails and unexpected switchover will be occurred.

When modifying operating system information, be sure to do it after stopping all PRIMECLUSTER services (unloading CF) or in a single-user mode.

Create cluster applications used in RMS before starting RMS

If starting RMS without creating cluster applications, an error message (CML,14) will be output and RMS will not start.

The overview and the methods for creating cluster applications, "Chapter 6 Building Cluster Applications."

If operating systems hang up or slow down on a node in a cluster, a healthy node may be forcibly stopped.

If operating systems hang up or slow down on a node in a cluster due to system load, and so on, CF or RMS detects LEFTCLUSTER and stop the Shutdown Facility stops the node forcibly.

The Shutdown Facility forcibly stops a node according to the survival priority. Therefore, when the hang-up and slowdown of operating systems on the failed node are recovered before a healthy node forcibly stops the failed node, the healthy node may be forcibly stopped first.

When a system volume on a disk device cannot be referred to because all paths failed in a SAN boot /iSCSI boot configuration, the PRIMECLUSTER failure detection function cannot be operated depending on the status of the system.

Because the node which cannot refer to the system volume is unstable, set the node to panic status with the following method.

When you can log in cluster nodes other than the relevant node

Stop the relevant node using the sdtool command.

```
# sdtool -k <the relevant node>
```

When you cannot log in any nodes

Set the node to panic status manually with one of the following methods.

- Press <Alt> + <SysRq> + <C>.

- Press the NMI button.

For details, see "Linux user guide."

When the RMS environment variable HV_RCSTART is set to "0", stop RMS by executing the "hvshut" command before stopping the node.

If you set "0" to the RMS environment variable HV_RCSTART, RMS shutdown processing is not called in node stopping process.

Therefore, cluster applications shutdown processing is not executed.

When you start cluster applications manually or confirm the message of a resource failure, check whether a resource with the "MONITORONLY" attribute has been in the fault state.

If you start or switch over cluster applications before the failure of the resource with the "MONITORONLY" attribute is solved, cluster inconsistencies or data corruption may occur.

When you set Firewall and use the state module in Firewall, do not restart the iptables service or the ip6tables service during PRIMECLUSTER operation.

When using the state module in Firewall, restarting the iptables service or the ip6tables service triggers initializing information of the communication status, and subsequent communication may not work correctly. Neither applications nor PRIMECLUSTER can work correctly, when you change the setting of Firewall, restart the cluster node but the iptables service or the ip6tables service.

7.5.1 Notes on Switching a Cluster Application Forcibly

When Forced switch request (Forced startup) of an application is issued, RMS overrides all safety checks and starts the application. So if shared resources which require exclusive control between nodes become Online on the multiple nodes simultaneously, it could result in data corruption or other inconsistencies.

To perform forced startup of an application safely, check whether RMS is running on all nodes in the cluster before starting forced startup according to the following procedure, and if there are the nodes on which RMS is not running, then shut down the nodes.

1. Check the node state by one of the following methods:

   - Execute the cftool -n command on all nodes.

```
fuji2# cftool -n
Node  Number State        Os       Cpu
fuji2 1       UP          Linux    Pentium
fuji3 2       UP          Linux    Pentium
```

   - Check the CF tree of the Cluster Admin.
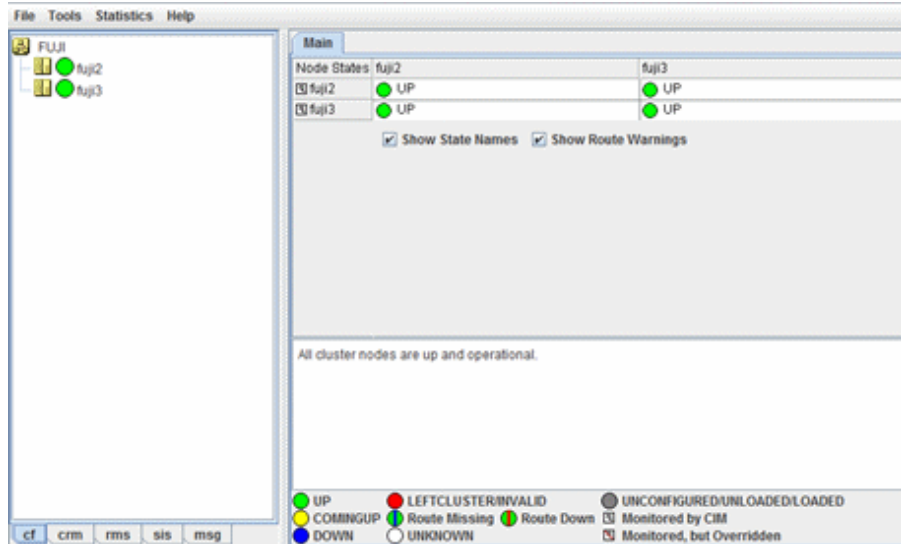
2. Check the following contents for the node states, and take corrective actions if necessary:

   - Check the node states are all UP.

   - If a LEFTCLUSTER node exists, recover CF from the LEFTCLUSTER state.

     For details, see "PRIMECLUSTER Cluster Foundation Configuration and Administration."

   - If a node with DOWN or UNKNOWN exists, or if a node for which the state is not displayed exists, check whether the operating system of the node has stopped. If the operating system is running, shut down the operating system or restart the system in single-user mode.

3. Check whether some nodes on which RMS is not running exist among the nodes on which cluster applications will be forcibly started by one of the following methods:

   - Execute the hvdisp -a command on nodes where cluster applications will be started and check that the state of objects whose Type is SysNode is Online.

```
fuji2# hvdisp -a

Local System:  fuji2RMS
Configuration: /opt/SMAW/SMAWRrms/build/config.us


Resource              Type      HostName          State          StateDetails
--------------------------------------------------------------------------------
fuji3RMS              SysNode                     Online
fuji2RMS              SysNode                     Online
app2                  userApp                     Offline
Machine001_app2       andOp     fuji3RMS
Machine000_app2       andOp     fuji2RMS          Offline
ManageProgram000_Cmd_APP2 gRes                        Offline
app1                  userApp                     Offline
Machine001_app1       andOp     fuji3RMS
Machine000_app1       andOp     fuji2RMS          Offline
ManageProgram000_Cmd_APP1 gRes                        Offline
```

- Check that the states of all SysNode displayed in the RMS tree of the Cluster Admin are Online.
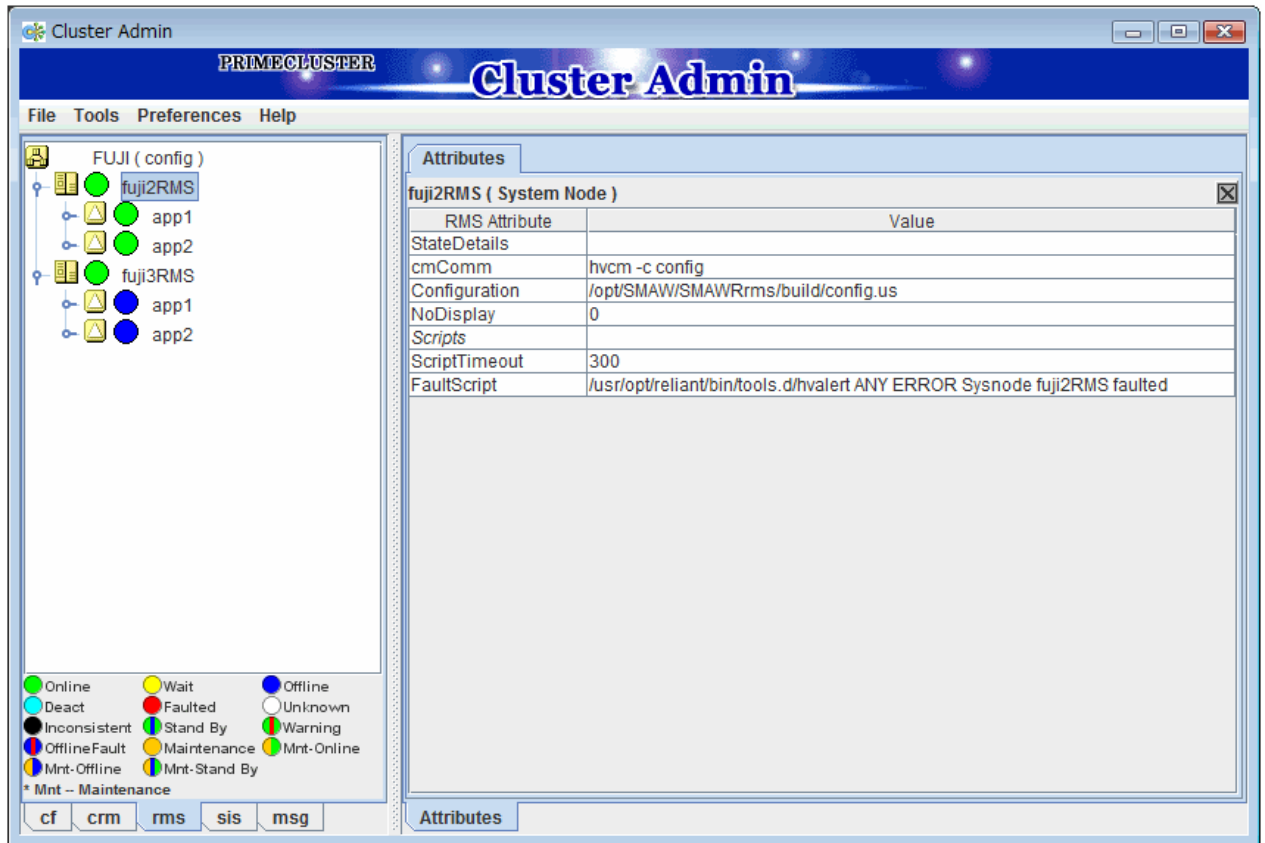
4. If nodes which satisfy the following conditions exist, shut down the operating system of the nodes, or restart the system in single-user mode.

   - The node state is UP, and

   - The state of SysNode is not Online.

5. Execute the Forced switch (hvswitch -f) to forcibly start the application.

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 how to modify a cluster configuration by following operations:

Operation flow
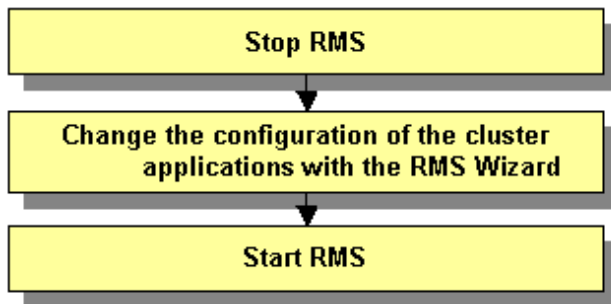
Operation Procedure:

1. Stop RMS.

   If RMS is running, see "7.2.1.2 Stopping RMS" and stop RMS.

2. Change the configuration of the cluster applications with the RMS Wizard.

   1. Log in to any one of the cluster nodes using system administrator access privileges.

   2. Start up the RMS Wizard.

      Execute the "hvw -n *configuration* file" command. Specify the name of the configuration file in which the configuration is defined.

      The following example shows how to start up RMS Wizard with the configuration file name "testconf."

      ```
      # /opt/SMAW/SMAWRrms/bin/hvw -n testconf
      ```

   3. Select "Application-Edit" from the "Main configuration menu."

      ```
      node1: Main configuration menu, current configuration: testconf
      No RMS active in the cluster
       1) HELP                          10) Configuration-Remove
       2) QUIT                          11) Configuration-Freeze
       3) Application-Create            12) Configuration-Thaw
       4) Application-Edit              13) Configuration-Edit-Global-Settings
       5) Application-Remove            14) Configuration-Consistency-Report
       6) Application-Clone             15) Configuration-ScriptExecution
       7) Configuration-Generate        16) RMS-CreateMachine
       8) Configuration-Activate        17) RMS-RemoveMachine
       9) Configuration-Copy
      Choose an action: 4
      ```

      When a cluster application that is performing standby operation is to be changed

- 215 -

1. Select the userApplication that needs modification of configuration from "Application selection menu." If more than one selection item is displayed, select userApplication written in capital letters. The following example shows how to select "APP1."

```
Edit: Application selection menu (restricted):
1) HELP
2) QUIT
3) RETURN
4) OPTIONS
5) APP1
Application Name: 5
```

2. When "turnkey wizard" appears, select what you want to change from the following table.

Contents to be changed	Selection item
Name of the userApplication	"ApplicationName"
Attribute of the userApplication	"Machines+Basics"
Cmdline resource configuration	"CommandLines"
Fsystem resource configuration	"LocalFileSystems"
Gds resource configuration	"Gds:Global-Disk-Services"
Gls resource configuration	"Gls:Global-Link-Services"
Procedure resource configuration	"Procedure:XXXXXX"

For details on the operation when you select above items, see "6.7 Setting Up Cluster Applications." Select "SAVE+EXIT" to return to the "Main configuration menu."

```
Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                          9) Procedure:BasicApplication(-)
 2) READONLY                     10) Symfoware(-)
 3) SAVE+EXIT                    11) Procedure:SystemState3(-)
 4) -                           12) Procedure:SystemState2(-)
 5) ApplicationName=APP1        13) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)       14) IpAddresses(-)
 7) CommandLines(Cmd_APP1)      15) LocalFileSystems(-)
 8) Procedure:Application(-)    16) Gds:Global-Disk-Services(-)
Choose the setting to process:
```

The following example shows how to change the attribute of "AutoStartUp" of the userApplication setting from "no" to "yes":

1. Select "Machines+Basics" from "turnkey wizard."

2. Select "AutoStartUp."

```
Machines+Basics (appl:consistent)
 1) HELP
 2) -
 3) SAVE+EXIT
 4) REMOVE+EXIT
 5) AdditionalMachine
 6) AdditionalConsole
 7) Machines[0]=fuji01RMS
 8) Machines[1]=fuji02RMS
 9) (PreCheckScript=)
10) (PreOnlineScript=)
11) (PostOnlineScript=)
12) (PreOfflineScript=)
13) (OfflineDoneScript=)
14) (FaultScript=)
15) (AutoStartUp=no)
16) (AutoSwitchOver=HostFailure|ResourceFailure|ShutDown)
17) (PreserveState=no)
18) (PersistentFault=0)
19) (ShutdownPriority=)
20) (OnlinePriority=)
21) (StandbyTransitions=)
22) (LicenseToKill=no)
23) (AutoBreak=yes)
24) (HaltFlag=no)
25) (PartialCluster=0)
26) (ScriptTimeout=)
Choose the setting to process:
```

3. Select "yes."

```
1) HELP
2) RETURN
3) yes
4) no
Set the AutoStartUp mode: 3
```

4. Confirm that "AutoStartUp" is changed to "yes," and then select "SAVE+EXIT."

```
Machines+Basics (appl:consistent)
 1) HELP
 2) -
 3) SAVE+EXIT
 4) REMOVE+EXIT
 5) AdditionalMachine
 6) AdditionalConsole
 7) Machines[0]=fuji01RMS
 8) Machines[1]=fuji02RMS
 9) (PreCheckScript=)
10) (PreOnlineScript=)
11) (PostOnlineScript=)
12) (PreOfflineScript=)
13) (OfflineDoneScript=)
14) (FaultScript=)
15) (AutoStartUp=yes)
16) (AutoSwitchOver=HostFailure|ResourceFailure|ShutDown)
17) (PreserveState=no)
18) (PersistentFault=0)
19) (ShutdownPriority=)
20) (OnlinePriority=)
21) (StandbyTransitions=)
22) (LicenseToKill=no)
23) (AutoBreak=yes)
24) (HaltFlag=no)
25) (PartialCluster=0)
26) (ScriptTimeout=)
Choose the setting to process:3
```

5. Select "SAVE+EXIT" from "turnkey wizard."

```
Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                              9) Procedure:BasicApplication(-)
 2) -                                10) Symfoware(-)
 3) SAVE+EXIT                        11) Procedure:SystemState3(-)
 4) -                                12) Procedure:SystemState2(-)
 5) ApplicationName=APP1            13) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)           14) IpAddresses(-)
 7) CommandLines(Cmd_APP1)          15) LocalFileSystems(-)
 8) Procedure:Application(-)        16) Gds:Global-Disk-Services(-)
Choose the setting to process:3
```

6. Select "RETURN" from the "Application selection menu."

```
Edit: Application selection menu (restricted):
1) HELP
2) QUIT
3) RETURN
4) OPTIONS
5) APP1
Application Name: 3
```

When a cluster application in a scalable operation is to be changed

For information on how to change a cluster application performing standby operation and which forms part of a cluster application in a scalable operation, see "When a cluster application that is performing standby operation is to be changed."

1. Select the userApplication to be reconfigured from the "Application selection menu." In the following example, "APP3" is selected.

```
Edit: Application selection menu (restricted):
 1) HELP
 2) QUIT
 3) RETURN
 4) OPTIONS
 5) APP1
 6) APP2
 7) APP3
Application Name: 7
```

2. Select "Controllers" from the "turnkey wizard SCALABLE" menu.

```
Settings of turnkey wizard "SCALABLE" (APP3:consistent)
 1) HELP              4) -                    7) Controllers(Ctl_APP3)
 2) -                 5) ApplicationName=APP3
 3) SAVE+EXIT         6) Machines+Basics(app3)
Choose the setting to process:7
```

3. "Settings of application type "Controller"" is displayed. Select one of the following according to the contents to be changed:

Contents to be changed	Resource to be selected
Start sequence of the cluster application (standby operation)	"ApplicationSequence"
Deletion of the application (standby operation)	"Controllers[*]"

[Supplement]

A number is specified in the "*" mark included in "Controllers[*]". Select the cluster application in a standby operation that you want to delete. You can delete a cluster application in a standby operation by specifying "NONE" on the screen after the selection.

For details on the operation to be performed after making the above selection, see "6.7 Setting Up Cluster Applications." After you change the configuration, select "SAVE+EXIT" and then return to the "Main configuration menu."

The following is an example in which the "AutoStartUp" attribute of the userApplication is changed to "yes" from "no."

1. Select "Machines+Basics" from the "turnkey wizard "SCALABLE"" menu.

```
Settings of turnkey wizard "SCALABLE" (APP3:consistent)
 1) HELP              4) -                    7) Controllers(Ctl_APP3)
 2) -                 5) ApplicationName=APP3
 3) SAVE+EXIT         6) Machines+Basics(app3)
Choose the setting to process:6
```

2. Select "(AutoStartUp=no)" from the "Machines+Basics" menu.

```
Machines+Basics (app3:consistent)
 1) HELP
 2) -
 3) SAVE+EXIT
 4) REMOVE+EXIT
 5) AdditionalMachine
 6) AdditionalConsole
 7) Machines[0]=fuji01RMS
 8) Machines[1]=fuji02RMS
 9) (PreCheckScript=)
10) (PreOnlineScript=)
11) (PostOnlineScript=)
12) (PreOfflineScript=)
13) (OfflineDoneScript=)
14) (FaultScript=)
15) (AutoStartUp=no)
16) (AutoSwitchOver=HostFailure|ShutDown)
17) (PreserveState=yes)
18) (PersistentFault=0)
19) (ShutdownPriority=)
20) (OnlinePriority=0)
21) (StandbyTransitions=)
22) (LicenseToKill=no)
23) (AutoBreak=yes)
24) (HaltFlag=no)
25) (PartialCluster=1)
26) (ScriptTimeout=)
Choose the setting to process:15
```

3. Select "yes."

```
 1) HELP
 2) RETURN
 3) yes
 4) no
Set the AutoStartUp mode:3
```

4. Check that "AutoStartUp" has been changed to "yes," and then select "SAVE+EXIT."

```
Machines+Basics (app3:consistent)
 1) HELP
 2) -
 3) SAVE+EXIT
 4) REMOVE+EXIT
 5) AdditionalMachine
 6) AdditionalConsole
 7) Machines[0]=fuji01RMS
 8) Machines[1]=fuji02RMS
 9) (PreCheckScript=)
10) (PreOnlineScript=)
11) (PostOnlineScript=)
12) (PreOfflineScript=)
13) (OfflineDoneScript=)
14) (FaultScript=)
15) (AutoStartUp=yes)
16) (AutoSwitchOver=HostFailure|ShutDown)
17) (PreserveState=yes)
18) (PersistentFault=0)
19) (ShutdownPriority=)
20) (OnlinePriority=0)
21) (StandbyTransitions=)
22) (LicenseToKill=no)
23) (AutoBreak=yes)
24) (HaltFlag=no)
25) (PartialCluster=1)
26) (ScriptTimeout=)
Choose the setting to process: 3
```

5. Select "SAVE+EXIT" from the "turnkey wizard "SCALABLE"" menu.

```
Settings of turnkey wizard "SCALABLE" (APP3:consistent)
 1) HELP            4) -                    7) Controllers(Ctl_APP3)
 2) -               5) ApplicationName=APP3
 3) SAVE+EXIT       6) Machines+Basics(app3)
Choose the setting to process:3
```

6. Select "RETURN" from the "Application selection menu."

```
Edit: Application selection menu (restricted):
 1) HELP
 2) QUIT
 3) RETURN
 4) OPTIONS
 5) APP1
 6) APP2
 7) APP3
Application Name:3
```

3. Select "Configuration-Generate" from the "Main configuration menu."

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                         10) Configuration-Remove
 2) QUIT                         11) Configuration-Freeze
 3) Application-Create           12) Configuration-Thaw
 4) Application-Edit             13) Configuration-Edit-Global-Settings
 5) Application-Remove           14) Configuration-Consistency-Report
 6) Application-Clone            15) Configuration-ScriptExecution
 7) Configuration-Generate       16) RMS-CreateMachine
 8) Configuration-Activate       17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 7
```

4. Select "Configuration-Activate" from the "Main configuration menu."

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                         10) Configuration-Remove
 2) QUIT                         11) Configuration-Freeze
 3) Application-Create           12) Configuration-Thaw
 4) Application-Edit             13) Configuration-Edit-Global-Settings
 5) Application-Remove           14) Configuration-Consistency-Report
 6) Application-Clone            15) Configuration-ScriptExecution
 7) Configuration-Generate       16) RMS-CreateMachine
 8) Configuration-Activate       17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 8
```

5. Select "QUIT" from the "Main configuration menu."

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                         10) Configuration-Remove
 2) QUIT                         11) Configuration-Freeze
 3) Application-Create           12) Configuration-Thaw
 4) Application-Edit             13) Configuration-Edit-Global-Settings
 5) Application-Remove           14) Configuration-Consistency-Report
 6) Application-Clone            15) Configuration-ScriptExecution
 7) Configuration-Generate       16) RMS-CreateMachine
 8) Configuration-Activate       17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 2
```

6. Check the cluster service for the PRIMECLUSTER-compatible product.
   Execute the following command in any node that is part of the cluster system:

```
# /etc/opt/FJSVcluster/bin/clrwzconfig -c
```

7. If the results of the cluster service check for the PRIMECLUSTER-compatible product shows that the "clrwzconfig" command output message 8050, re-register the cluster service for the PRIMECLUSTER-compatible product.
   Execute the following command in any node that is part of the cluster system:

```
# /etc/opt/FJSVcluster/bin/clrwzconfig
```

8. Change the resources of the PRIMECLUSTER-compatible product. For details, see "Appendix F Registering PRIMECLUSTER-compatible Resources."

9. Start RMS.

   Start RMS as described in "7.2.1.1 Starting RMS."

8.2 Changing an IP Address on the Public / Administrative 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.

📖 **Note**

⋯⋯⋯
If you use the virtual machine function, this section explains the Public / administrative LAN of the gesut OS.
⋯⋯⋯

8.2.1 Changing the IP address of the public LAN

The following describes how to change an IP address when the IP address of the public LAN has been changed.

Operation Procedure:

1. Execute the following command on one of the cluster nodes to stop RMS operation:

   ```
   # hvshut -a
   ```

2. Execute the following commands on all nodes to start the system in single-user mode.

   ```
   # /sbin/shutdown now
   ```

3. Mount the local file system on all nodes.

   ```
   (Example)
   ```

   ```
   # /bin/mount -a -t ext3
   ```

4. Edit the "/etc/hosts" file, and change the IP address on each node.

5. Change the IP address of the public LAN.

   For details on how to change the IP address, see the Linux documentation.

6. If an IP address of the Shutdown Facility or the Shutdown Agent needs to be changed along with public LAN changes, it is necessary to change "/etc/opt/SMAW/SMAWsf/rcsd.cfg" and "/etc/opt/SMAW/SMAWsf/SA_xxx.cfg" on each node.
   SA_xxx.cfg indicates the configuration file for the Shutdown Agent.

   📖 **See**

   ⋯⋯⋯
   For details, see "5.1.2 Setting Up the Shutdown Facility."
   ⋯⋯⋯

7. On each node, change the IP address of the public LAN used by Web-Based Admin View.

   📖 **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.*"
   ⋯⋯⋯

8. If a takeover IP address must be changed (when the takeover IP address is changed after installation, or when the takeover IP address is changed due to transfer of the node), correct the IP address being used as the takeover IP address in the "/etc/hosts" file of each node.

When you have created takeover network resources, and change the subnet mask value due to the change of the public LAN, you also need to edit the /usr/opt/reliant/etc/hvipalias file.

See
. . . .
For information on how to edit the /usr/opt/reliant/etc/hvipalias file, see "6.7.3.6 Setting Up Takeover Network Resources."
. . . .

9. If a Gls resource is created, edit the "/opt/SMAW/SMAWRrms/bin/hvenv.local" file of each cluster node as shown below so that RMS does not start automatically when the node is rebooted.

```
# vi /opt/SMAW/SMAWRrms/bin/hvenv.local <Return>
```

```
export HV_RCSTART=0
```

10. Restart the system on all nodes.

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

11. If the takeover IP address is defined by GLS, see "3.2.2 Changing configuration Information" in the "*PRIMECLUSTER Global Link Services Configuration and Administration Guide : Redundant Line Control Function.*"

Note
. . . .
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/hosts and the GLS environment settings, and re-register the Gls resource.

For details, see "6.2.1 GLS Setup" and "6.7.3.5 Setting Up Gls Resources" in this manual, and "5.2 Configuration for Cluster system" in the "PRIMECLUSTER Global Link Services Configuration and Administration Guide : Redundant Line Control Function."
. . . .

12. If you edited the "/opt/SMAW/SMAWRrms/bin/hvenv.local" file in Step 9, return the file to its previous setting.

8.2.2 Changing the IP address of the administrative LAN

The following describes how to change an IP address when the IP address of the administrative LAN has been changed.

Operation Procedure:

1. Execute the following command on one of the cluster nodes to stop RMS operation:

```
# hvshut -a
```

2. Execute the following commands on all nodes to start the system in single-user mode.

```
# /sbin/shutdown now
```

3. Mount the local file system on all nodes.

```
(Example)
```

```
# /bin/mount -a -t ext3
```

4. Edit the "/etc/hosts" file, and change the IP address on each node.

5. Change the IP address of the administrative LAN.

   For details on how to change the IP address, see the Linux documentation.

6. If an IP address of the Shutdown Facility or the Shutdown Agent needs to be changed along with administrative LAN changes, it is necessary to change "/etc/opt/SMAW/SMAWsf/rcsd.cfg" and "/etc/opt/SMAW/SMAWsf/SA_xxx.cfg" on each node. SA_xxx.cfg indicates the configuration file for the Shutdown Agent.

   🗂 See
   ....
   For details, see "5.1.2 Setting Up the Shutdown Facility."
   ....

7. Restart the system on all nodes.

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

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.

   If the nodes are already operating, you do not need to restart them.

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(5) 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 command to confirm that the IP address of CIP was changed.

   ```
    # /opt/SMAW/SMAWcf/bin/ciptool -a
   ```

🗂 See
....
For details on the "ciptool" command, see the manual page describing "ciptool".
....

8.4 Changing a Node Name

The following explains how to change the node name after building a PRIMECLUSTER system.

Operation Procedure:

1. Stop the CF on the node whose node name is to be changed.

   For information on how to stop CF, see "5.6 Starting and stopping CF" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide.*"

2. Change the old host name of the /etc/hosts file to a new host name.

   (Example)

   ```
   10.20.30.40    node1
   ```
   ↓ Correction
   ```
   10.20.30.40    nodeA
   ```

3. Change the old host name of the /etc/sysconfig/network file to a new host name.

   (Example)

   ```
   HOSTNAME=node1
   ```
   ↓ Correction
   ```
   HOSTNAME=nodeA
   ```

4. Reboot the system.

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

   After rebooting the system, execute the following procedure for the other node.

5. Change the old host name of the /etc/hosts file to a new host name.

6. Restart Web-Based Admin View.

8.5 Changing the Operation Attributes of a userApplication

PRIMECLUSTER allows operation attributes to be set by CUI operation, according to the needs of the user. Change the operation attributes to match the desired operation.

- Operation attribute types

  For further details about the operation attribute of the userApplication, see "11 Appendix - Attributes" in "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide.*"

- Change methods

  - ""

    Explains how to change the operation attributes of the userApplication.

  - ""

    Explains how to change the RMS environment variables.

Note

. . . .

Be sure to stop RMS before you change the operation attributes of userApplication. For instructions on stopping RMS, see "7.2.1.2 Stopping RMS"

. . . .

8.5.1 Changing Operation Attributes (CUI)

This section explains how to change the userApplication attributes with CUI.

For further details about the operation attribute specified in step 8, see "11 Appendix - Attributes" in "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide*."

Operation Procedure:

Note

. . . .

"Application" on the CUI screen indicates a cluster application.

. . . .

1. Log in to any one of the cluster nodes using system administrator access privileges.

2. Stop RMS.
   If RMS is running, see "7.2.1.2 Stopping RMS" and stop RMS.

3. Execute the "hvw" command.

   ```
   # /opt/SMAW/SMAWRrms/bin/hvw -n <configuration file>
   ```

   Specify an optional Configuration file name in <*configuration file*>.

4. Select "Application-Edit" from the main menu of CUI. Enter a number and then press the *Enter* key.

   Figure 8.1 Main menu

   ```
   apple: Main configuration menu, current configuration: config
   No RMS active in the cluster
    1) HELP                          10) Configuration-Remove
    2) QUIT                          11) Configuration-Freeze
    3) Application-Create            12) Configuration-Thaw
    4) Application-Edit              13) Configuration-Edit-Global-Settings
    5) Application-Remove            14) Configuration-Consistency-Report
    6) Application-Clone             15) Configuration-ScriptExecution
    7) Configuration-Generate        16) RMS-CreateMachine
    8) Configuration-Activate        17) RMS-RemoveMachine
    9) Configuration-Copy
   Choose an action:
   ```

5. Select the userApplication for which you want to change the operation attributes from the "Application selection menu."

   The following example shows how to select "APP1."

Figure 8.2 Cluster application selection menu

```
Edit: Application selection menu (restricted):
1) HELP
2) QUIT
3) RETURN
4) OPTIONS
5) APP1
Application Name: 5
```

6. When turnkey wizard STANDBY appears, select "Machines+Basics" and then change the operation attributes of the userApplication.

Figure 8.3 turnkey wizard "STANDBY"

```
Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                          9) Procedure:BasicApplication(-)
 2) -                            10) Symfoware(-)
 3) SAVE+EXIT                    11) Procedure:SystemState3(-)
 4) -                            12) Procedure:SystemState2(-)
 5) ApplicationName=APP1         13) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)        14) IpAddresses(-)
 7) CommandLines(Cmd APP1)       15) LocalFileSystems(-)
 8) Procedure:Application(-)     16) Gds:Global-Disk-Services(-)
Choose the setting to process:6
```

7. Select the operation attribute that you want to change from "Machines+Basics."

Figure 8.4 Machines+Basics

```
Machines+Basics {app1:consistent}
 1) HELP
 2) -
 3) SAVE+EXIT
 4) REMOVE+EXIT
 5) AdditionalMachine
 6) AdditionalConsole
 7) Machines[0]=fuji01RMS
 8) Machines[1]=fuji02RMS
 9) (PreCheckScript=)
10) (PreOnlineScript=)
11) (PostOnlineScript=)
12) (PreOfflineScript=)
13) (OfflineDoneScript=)
14) (FaultScript=)
15) (AutoStartUp=no)
16) (AutoSwitchOver=HostFailure|ResourceFailure|ShutDown)
17) (PreserveState=no)
18) (PersistentFault=0)
19) (ShutdownPriority=)
20) (OnlinePriority=)
21) (StandbyTransitions=)
22) (LicenseToKill=no)
23) (AutoBreak=yes)
24) (HaltFlag=no)
25) (PartialCluster=0)
26) (ScriptTimeout=)
Choose the setting to process:
```

To change the value:
Select the item that you want to change. Enter a number and then press the *Enter* key.
(Example)
Choose the setting to process: **20** <RETURN>
The "Value setup menu," as shown in step 8, will be displayed.

Select "SAVE+EXIT." Enter a number and then press the *Enter* key.
Choose the setting to process: **3** <RETURN>
"turnkey wizard "STANDBY"" of step 9 will be displayed.

8. Set up the value from the value setup menu.

Figure 8.5 Value setup menu (Example of OnlinePriority)

```
1) HELP
2) RETURN
3) 0
4) 1
Enable Online Priority (Active Standby) feature for this application:
```

Select a setup value. Enter a number and then press the *Enter* key.

Point

Select "RETURN" to return to the previous menu.

If there are multiple attributes to be changed, repeat steps 7 and 8 for each attribute.

If the attribute is other than "OnlinePriority," the menu number in step 8 will be different from that in this example.

9. Select "SAVE+EXIT" from the "Machines+Basics" screen to return to the "turnkey wizard "STANDBY"."

Figure 8.6 turnkey wizard "STANDBY"

```
Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                          9) Procedure:BasicApplication(-)
 2) -                            10) Symfoware(-)
 3) SAVE+EXIT                     11) Procedure:SystemState3(-)
 4) -                            12) Procedure:SystemState2(-)
 5) ApplicationName=APP1         13) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)        14) IpAddresses(-)
 7) CommandLines(Cmd APP1)       15) LocalFileSystems(-)
 8) Procedure:Application(-)     16) Gds:Global-Disk-Services(-)
Choose the setting to process:
```

Select "SAVE+EXIT" to return to the cluster application selection menu.

10. Select "RETURN" to return to the main menu.

Figure 8.7 Cluster application selection menu

```
Edit: Application selection menu (restricted):
1) HELP
2) QUIT
3) RETURN
4) OPTIONS
5) APP1
Application Name: 3
```

11. Select "Configuration-Generate" and then "Configuration-Activate" from the main menu.
    Content changes will be enabled on all the cluster nodes.

Figure 8.8 Main menu

```
apple: Main configuration menu, current configuration: config
No RMS active in the cluster
 1) HELP                          10) Configuration-Remove
 2) QUIT                          11) Configuration-Freeze
 3) Application-Create            12) Configuration-Thaw
 4) Application-Edit              13) Configuration-Edit-Global-Settings
 5) Application-Remove            14) Configuration-Consistency-Report
 6) Application-Clone             15) Configuration-ScriptExecution
 7) Configuration-Generate        16) RMS-CreateMachine
 8) Configuration-Activate        17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action:
```

```
apple: Main configuration menu, current configuration: config
No RMS active in the cluster
 1) HELP                          10) Configuration-Remove
 2) QUIT                          11) Configuration-Freeze
 3) Application-Create            12) Configuration-Thaw
 4) Application-Edit              13) Configuration-Edit-Global-Settings
 5) Application-Remove            14) Configuration-Consistency-Report
 6) Application-Clone             15) Configuration-ScriptExecution
 7) Configuration-Generate        16) RMS-CreateMachine
 8) Configuration-Activate        17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action:
```

Figure 8.9 Configuration distribution (Example of executing Configuration-Activate)

```
About to activate the configuration config ...
Testing for RMS to be up somewhere in the cluster ... done
Arranging sub applications topologically ... done.
Check for all applications being consistent ... done.
Running overall consistency check ... done.
Generating pseudo code [one dot per (sub) application]: .... done
Generating RMS resources.................................... done
hvbuild using /usr/opt/reliant/build/wizard.d/config/config.us
About    to   distribute   the   new   configuration   data   to   hosts:
test node1RMS,test node2RMS
The new configuration was distributed successfully.
About to put the new configuration in effect ... done.
The activation has finished successfully.
Hit CR to continue
```

Note

When the processing is successfully done, the message "The activation has finished successfully" appears. If this message is not displayed, the modified information contains incorrect settings. Check and correct the settings.

12. Press the *Enter* key to return to the main menu.

Figure 8.10 Main menu

```
apple: Main configuration menu, current configuration: config
No RMS active in the cluster
 1) HELP                          10) Configuration-Remove
 2) QUIT                          11) Configuration-Freeze
 3) Application-Create            12) Configuration-Thaw
 4) Application-Edit              13) Configuration-Edit-Global-Settings
 5) Application-Remove            14) Configuration-Consistency-Report
 6) Application-Clone             15) Configuration-ScriptExecution
 7) Configuration-Generate        16) RMS-CreateMachine
 8) Configuration-Activate        17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action:
```

13. Select "QUIT" to terminate the processing.

14. Check the cluster service for the PRIMECLUSTER-compatible product.
    Execute the following command in any node that is part of the cluster system:

    ```
    # /etc/opt/FJSVcluster/bin/clrwzconfig -c
    ```

15. If the results of the cluster service check for the PRIMECLUSTER-compatible product shows that the "clrwzconfig" command output message 8050, re-register the cluster service for the PRIMECLUSTER-compatible product.
    Execute the following command in any node that is part of the cluster system:

    ```
    # /etc/opt/FJSVcluster/bin/clrwzconfig
    ```

16. Change the resources of the PRIMECLUSTER-compatible product. For details, see "Appendix F Registering PRIMECLUSTER-compatible Resources."

17. Start up RMS and userApplication from Cluster Admin.

    ## Information

    ....

    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."

    ....

8.6 Changing PRIMECLUSTER Operation Attributes

PRIMECLUSTER Operation Attributes are configured by RMS environment variables.

You can change the environment variable by editing the /opt/SMAW/SMAWRrms/bin/hvenv.local file to add or change its value.

See

....

- Create the "hvenv.local" file as required.

- For details on hvenv.local, 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 "13 Appendix - Environment variables" in "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide."

....

8.7 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. For instructions on stopping RMS, see "7.2.1.2 Stopping RMS."

....

8.7.1 Deleting a Resource

This section explains how to delete a resource.

Operation Procedure:

1. Log in to any one of the cluster nodes using system administrator access privileges.

2. Start the RMS Wizard.

   Execute the "hvw -n *configuration file*" command. Specify the name of the configuration file in which the resource is defined.

   The following example shows how to start RMS Wizard with the configuration file name "testconf."

```
# /opt/SMAW/SMAWRrms/bin/hvw -n testconf
```

3. Select "Application-Edit" from the "Main configuration menu."

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                          10) Configuration-Remove
 2) QUIT                          11) Configuration-Freeze
 3) Application-Create            12) Configuration-Thaw
 4) Application-Edit              13) Configuration-Edit-Global-Settings
 5) Application-Remove            14) Configuration-Consistency-Report
 6) Application-Clone             15) Configuration-ScriptExecution
 7) Configuration-Generate        16) RMS-CreateMachine
 8) Configuration-Activate        17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 4
```

4. Select the userApplication in which the resource is registered from the "Application selection menu." The following example shows how to select "APP1."

```
Edit: Application selection menu (restricted):
1) HELP
2) QUIT
3) RETURN
4) OPTIONS
5) APP1
Application Name: 5
```

5. Use the "turnkey wizard" to select the resource.

   Select "REMOVE+EXIT" from the screen after selecting the resource.
   Then, select "SAVE+EXIT" to return to the "Main configuration menu."

```
Settings of turnkey wizard "STANDBY" (APP1:consistent)
 1) HELP                          9) Procedure:BasicApplication(-)
 2) -                            10) Symfoware(-)
 3) SAVE+EXIT                    11) Procedure:SystemState3(-)
 4) -                            12) Procedure:SystemState2(-)
 5) ApplicationName=APP1         13) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)        14) IpAddresses(-)
 7) CommandLines(Cmd APP1)       15) LocalFileSystems(-)
 8) Procedure:Application(-)     16) Gds:Global-Disk-Services(-)
Choose the setting to process:
```

6. Select "Configuration-Generate" from the "Main configuration menu."

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                              10) Configuration-Remove
 2) QUIT                              11) Configuration-Freeze
 3) Application-Create                12) Configuration-Thaw
 4) Application-Edit                  13) Configuration-Edit-Global-Settings
 5) Application-Remove                14) Configuration-Consistency-Report
 6) Application-Clone                 15) Configuration-ScriptExecution
 7) Configuration-Generate            16) RMS-CreateMachine
 8) Configuration-Activate            17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 7
```

7. Select "Configuration-Activate" from the "Main configuration menu."

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                              10) Configuration-Remove
 2) QUIT                              11) Configuration-Freeze
 3) Application-Create                12) Configuration-Thaw
 4) Application-Edit                  13) Configuration-Edit-Global-Settings
 5) Application-Remove                14) Configuration-Consistency-Report
 6) Application-Clone                 15) Configuration-ScriptExecution
 7) Configuration-Generate            16) RMS-CreateMachine
 8) Configuration-Activate            17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 8
```

8. Select "QUIT" from the "Main configuration menu" to exit from the RMS Wizard.

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                              10) Configuration-Remove
 2) QUIT                              11) Configuration-Freeze
 3) Application-Create                12) Configuration-Thaw
 4) Application-Edit                  13) Configuration-Edit-Global-Settings
 5) Application-Remove                14) Configuration-Consistency-Report
 6) Application-Clone                 15) Configuration-ScriptExecution
 7) Configuration-Generate            16) RMS-CreateMachine
 8) Configuration-Activate            17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 2
```

9. Check the cluster service for the PRIMECLUSTER-compatible product.
   Execute the following command in any node that is part of the cluster system:

```
# /etc/opt/FJSVcluster/bin/clrwzconfig -c
```

10. If the results of the cluster service check for the PRIMECLUSTER-compatible product shows that the "clrwzconfig" command output message 8050, re-register the cluster service for the PRIMECLUSTER-compatible product.
    Execute the following command in any node that is part of the cluster system:

```
# /etc/opt/FJSVcluster/bin/clrwzconfig
```

11. Change the resources of the PRIMECLUSTER-compatible product. For details, see "Appendix F Registering PRIMECLUSTER-compatible Resources."

Note

- - When deleting a Gds resource, you need to bring the Gds volume Online.
    See "8.7.1.1 Settings made when deleting a Gds resource."

- - When deleting a procedure resource, first delete the procedure resource from the cluster resource management facility after deleting the procedure resource from the cluster application. For details on how to delete a procedure resource from the cluster resource management facility, see "E.3 Deleting a Procedure Resource."

- - When deleting an Fsystem resource, delete the mount point that was being used as the resource (mount point of the line beginning with "#RMS#") from /etc/fstab.pcl on all nodes.

- - When deleting network resources, you need to change the environment files. To change the environment files, see "F.2 Deleting Takeover Network Resources."

8.7.1.1 Settings made when deleting a Gds resource

When the Gds resource is to be deleted, you must also notify the cluster resource management facility (CRM) that the Gds resources can no longer be used for the cluster application.

Operation Procedure:

Execute the following command on the node on which you want to delete resources.

```
# /opt/SMAW/SMAWRrms/bin/hvgdsetup -d [class-name]
```

8.7.2 Deleting the Hardware Resource

This section explains how to delete the resources of the following hardware.

- - Shared disk device

- - Network interface card

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 " cldelrsc" command to delete the hardware resource that was registered.

   See the relevant manual pages for details on this command.

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" command.

- - If you have deleted other than the correct network takeover resource, reregister the deleted network takeover resource. See "F.1 Associating Resources to a Cluster Service Instance."

- - If the shared disk for which resources are to be deleted is registered to a GDS class, first delete the shared disk from the GDS class, and then delete the resources of the shared disk. For instructions on how to delete a shared disk from a GDS class, refer to the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*."

8.7.3 Deleting a userApplication

This section explains how to delete the userApplication.

🔰 Note

- - If you delete a userApplication, all the resources registered to the userApplication will also be deleted.

- - If Gds resources are registered to the userApplication to be deleted, bring the Gds volume online. See "8.7.1.1 Settings made when deleting a Gds resource."

Operation Procedure:

1. Log in to any one of the cluster nodes using system administrator access privileges.

2. Start the RMS Wizard.

   Execute the "hvw -n *configuration file*" command. Specify a name of the configuration file in which the userApplication is defined.

   The following example shows how to start RMS Wizard with the configuration file name "testconf."

   ```
   # /opt/SMAW/SMAWRrms/bin/hvw -n testconf
   ```

3. Select "Application-Remove" from the "Main configuration menu."

   ```
   node1: Main configuration menu, current configuration: testconf
   No RMS active in the cluster
    1) HELP                          10) Configuration-Remove
    2) QUIT                          11) Configuration-Freeze
    3) Application-Create            12) Configuration-Thaw
    4) Application-Edit              13) Configuration-Edit-Global-Settings
    5) Application-Remove            14) Configuration-Consistency-Report
    6) Application-Clone             15) Configuration-ScriptExecution
    7) Configuration-Generate        16) RMS-CreateMachine
    8) Configuration-Activate        17) RMS-RemoveMachine
    9) Configuration-Copy
   Choose an action: 5
   ```

4. Select the userApplication that you want to delete from the "Application selection menu."

   The following example shows how to select APP2.

   ```
   Edit: Application selection menu (restricted):
   1) HELP
   2) QUIT
   3) RETURN
   4) OPTIONS
   5) APP1
   6) APP2
   Application Name: 6
   ```

   Enter "yes" in response to the following message.

   ```
      About to remove all data of APP2,
          Please confirm this by typing yes: yes
   ```

5. Select "Configuration-Generate" from the "Main configuration menu."

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                          10) Configuration-Remove
 2) QUIT                          11) Configuration-Freeze
 3) Application-Create            12) Configuration-Thaw
 4) Application-Edit              13) Configuration-Edit-Global-Settings
 5) Application-Remove            14) Configuration-Consistency-Report
 6) Application-Clone             15) Configuration-ScriptExecution
 7) Configuration-Generate        16) RMS-CreateMachine
 8) Configuration-Activate        17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 7
```

6. Select "Configuration-Activate" from the "Main configuration menu."

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                          10) Configuration-Remove
 2) QUIT                          11) Configuration-Freeze
 3) Application-Create            12) Configuration-Thaw
 4) Application-Edit              13) Configuration-Edit-Global-Settings
 5) Application-Remove            14) Configuration-Consistency-Report
 6) Application-Clone             15) Configuration-ScriptExecution
 7) Configuration-Generate        16) RMS-CreateMachine
 8) Configuration-Activate        17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 8
```

7. Select "QUIT" from the "Main configuration menu" to exit from the RMS Wizard.

```
node1: Main configuration menu, current configuration: testconf
No RMS active in the cluster
 1) HELP                          10) Configuration-Remove
 2) QUIT                          11) Configuration-Freeze
 3) Application-Create            12) Configuration-Thaw
 4) Application-Edit              13) Configuration-Edit-Global-Settings
 5) Application-Remove            14) Configuration-Consistency-Report
 6) Application-Clone             15) Configuration-ScriptExecution
 7) Configuration-Generate        16) RMS-CreateMachine
 8) Configuration-Activate        17) RMS-RemoveMachine
 9) Configuration-Copy
Choose an action: 2
```

8. Check the cluster service for the PRIMECLUSTER-compatible product.
   Execute the following command in any node that is part of the cluster system:

```
# /etc/opt/FJSVcluster/bin/clrwzconfig -c
```

9. If the results of the cluster service check for the PRIMECLUSTER-compatible product shows that the "clrwzconfig" command output message 8050, re-register the cluster service for the PRIMECLUSTER-compatible product.
   Execute the following command in any node that is part of the cluster system:

```
# /etc/opt/FJSVcluster/bin/clrwzconfig
```

10. Change the resources of the PRIMECLUSTER-compatible product. For details, see "Appendix F Registering PRIMECLUSTER-compatible Resources."

8.8 Changing MMB Settings

This section explains how to change MMB settings.

8.8.1 Changing the MMB IP Address

This section explains how to change the MMB IP address.

📝 **Note**

. . . .

- Change the MMB IP address on each node.

- Repeat the operation procedure described below to change more than one MMB IP address.

. . . .

Operation Procedure:

1. Execute the following command on all nodes to stop the operation of the shutdown facility (SF):

```
# sdtool -e
```

2. Execute the following command on the node in which IP address is changed to stop MMB asynchronous monitoring daemons.

```
# /etc/opt/FJSVcluster/bin/clmmbmonctl stop
```

3. Change the IP address according to the MMB procedure.

4. On the node in which the IP address was changed, execute the following command to start MMB asynchronous monitoring daemons and the shutdown facility.

```
# /etc/opt/FJSVcluster/bin/clmmbmonctl start
# sdtool -b
```

5. After the shutdown facility started in Step 4, start the shutdown facility on the remaining nodes.

```
# sdtool -b
```

6. Execute the following command on all nodes and check that the shutdown facility operates normally.

```
# sdtool -s
```

📝 **Note**

. . . .

There is a possibility that the mistake is found in the configuration setting of the hardware when displayed as follows though change of the setting of the shutdown facility is completed.

- "InitFailed" is displayed as the initial status.

- "Unknown" or "TestFailed" is displayed as the test status.

In this case, review the configuration setting of the hardware.

8.8.2 Changing the User Name and Password for Controlling the MMB with RMCP

This section explains how to change the user name and password for controlling the MMB with RMCP.

Operation Procedure

1. Execute the following command on all nodes to stop the operation of the shutdown facility (SF):

```
# sdtool -e
```

2. According to the procedures of MMB, change the user name and password to control MMB by RMCP. If you change the user name and password for several nodes, change them for all nodes.

3. By executing the following command, change the user name and password of MMB information for MMB shutdown facility. If the user name and the password are to be changed on multiple nodes, change the values on all nodes to be changed.

```
# /etc/opt/FJSVcluster/bin/clmmbsetup -m [-u user-name] [-p]
```

4. Execute the following command on all nodes to start the shutdown facility (SF).

```
# sdtool -b
```

8.9 Changing Virtual Machine Settings

This section describes how to change the VMGuest settings when you have changed the Host OS setting.

8.9.1 Xen environment

This section describes how to change the settings of the shutdown facility when changing the settings of the host OS in the environment where the Xen virtual machine function is used with PRIMEQUEST1000 series.

8.9.1.1 Changing the FJSVvmSP password

The following describes how to change the settings when you have changed the guest OS user password from the host OS after introducing the PRIMECLUSTER system into a Xen environment.

Operation procedure

1. Execute the following command on all guest OSes (nodes) to stop the operation of the shutdown facility (SF).

```
# sdtool -e
```

2. Encrypt the password.

```
# /opt/SMAW/SMAWsf/bin/sfcipher -c
Enter Password:
Re-Enter Password:
Xh+kSlJ8nlQ=
```

3. Define the encrypted password in the Shutdown Agent configuration file.

See

For details on how to define the configuration file, see "5.1.2.5.1 vmSP (Virtual Machine Service Provider)."

4. Execute the following command on all nodes to start the shutdown facility (SF).

```
# sdtool -b
```

8.9.1.2 Changing the IP address of the Host OS

The following describes how to change the settings when you have changed the Host OS IP address of the virtual machine after introducing the PRIMECLUSTER system into a Xen environment.

Operation procedure

1. Execute the following command on all nodes to stop the operation of the shutdown facility (SF).

```
# sdtool -e
```

2. Define the changed IP address in the Shutdown Agent configuration file.

See
. . . .
For details on how to define the configuration file, see "5.1.2.5.1 vmSP (Virtual Machine Service Provider)."
. . . .

3. Execute the following command on all nodes to start the shutdown facility (SF).

```
# sdtool -b
```

8.9.2 KVM environment

This section describes how to change the settings of the shutdown facility when changing the settings of the hypervisor in the environment where the KVM virtual machine function is used.

8.9.2.1 Changing the password for the user for the shutdown facility

The following describes how to change the settings when you have changed the password for the user for the shutdown facility set in the Shutdown Agent configuration file after introducing the PRIMECLUSTER system into a KVM environment.

Operation procedure

1. Execute the following command on all guest OSes (nodes) to stop the operation of the shutdown facility (SF).

```
# sdtool -e
```

2. Encrypt the password.

```
# /opt/SMAW/SMAWsf/bin/sfcipher -c
Enter Password:
Re-Enter Password:
Xh+kSlJ8nlQ=
```

3. Define the encrypted password in the Shutdown Agent configuration file.

See
. . . .
For details on how to define the configuration file, see "5.1.2.5.2 libvirt."
. . . .

4. Execute the following command on all nodes to start the shutdown facility (SF).

```
# sdtool -b
```

8.9.2.2 Changing the IP address of the hypervisor

The following describes how to change the settings when you have changed the hypervisor IP address of the virtual machine after introducing the PRIMECLUSTER system into a KVM environment.

Operation procedure

1. Execute the following command on all nodes to stop the operation of the shutdown facility (SF).

```
# sdtool -e
```

2. Define the changed IP address in the Shutdown Agent configuration file.

> 📚 **See**
> ....
> For details on how to define the configuration file, see "5.1.2.5.2 libvirt."
> ....

3. Execute the following command on all nodes to start the shutdown facility (SF).

```
# sdtool -b
```

8.9.2.3 Changing the Settings in /etc/sysconfig/libvirt-guests

This section explains the procedure for changing the settings in /etc/sysconfig/libvirt-guests after installing the PRIMUCLUSTER system in a KVM environment.

Operation procedure

1. Execute the following command on all nodes to stop the operation of the shutdown facility (SF).

```
# sdtool -e
```

2. Change the settings in /etc/sysconfig/libvirt-guests.

> 📚 **See**
> ....
> For details on the settings in /etc/sysconfig/libvirt-guests, see "Setting the guest OS in the host OS (in a KVM environment)" for each virtual environment shown below:
>
> - When building a cluster system between guest OSes on one host OS, see "3.2.1.1 Host OS setup."
>
> - When building a cluster system between guest OSes on multiple host OSes without using Host OS failover function, see "3.2.2.1 Host OS setup."
>
> - When building a cluster system between guest OSes on multiple host OSes using Host OS failover function, see "3.2.3.1 Installation and Setup of Related Software (Host OS)."
> ....

3. Execute the following command on all nodes to start the shutdown facility (SF).

```
# sdtool -b
```

8.10 Changing IPMI Settings

This section describes changing IPMI settings.

8.10.1 Changing the IPMI IP Address

This section explains how to change the IPMI IP address.

Operation Procedure:

1. Execute the following command on all nodes to stop the operation of the shutdown facility (SF):

```
# sdtool -e
```

2. Change the IP address according to the IPMI procedure.

3. Define the changed IP address in the Shutdown Agent configuration file.

See

For details on how to define the configuration file, see "5.1.2.3.2 IPMI."

4. Execute the following command on any node to apply changes of the configuration file.

```
# /etc/opt/FJSVcllkcd/bin/panicinfo_setup
```

After the following message is displayed, select "I."

```
panicinfo_setup: WARNING: /etc/panicinfo.conf file already exists.
(I)nitialize, (C)opy or (Q)uit (I/C/Q) ?
```

5. Start the shutdown daemon (rcsd) on all nodes.

```
# sdtool -b
```

6. Execute the following command on all nodes and check that the shutdown facility operates normally.

```
# sdtool -s
```

Note

There is a possibility that the mistake is found in the configuration setting of the agent or hardware when displayed as follows though changing the setting of the shutdown facility is completed.

- "InitFailed" is displayed as the initial status.

- "Unknown" or "TestFailed" is displayed as the test status.

In this case, review the configuration setting of the agent or hardware.

8.10.2 Changing the User Name and Password for iRMC

This section explains how to change the user name and password for iRMC.

Operation Procedure:

1. Execute the following command on all nodes to stop the shutdown facility.

```
# sdtool -e
```

2. Change the user name and password according to the procedure for IPMI.

3. Define the changed user name and password for iRMC in the Shutdown Agent configuration file.

See

For details on how to define the configuration file, see "5.1.2.3.2 IPMI."

4. Execute the following command on any node to apply changes of the configuration file.

```
# /etc/opt/FJSVcllkcd/bin/panicinfo_setup
```

After the following message is displayed, select "I."

```
panicinfo_setup: WARNING: /etc/panicinfo.conf file already exists.
(I)nitialize, (C)opy or (Q)uit (I/C/Q) ?
```

5. Execute the following command on all nodes to start the shutdown facility.

```
# sdtool -b
```

6. Execute the following command on all nodes and check that the shutdown facility operates normally.

```
# sdtool -s
```

Note

. . . .

There is a possibility that the mistake is found in the configuration setting of the agent or hardware when displayed as follows though changing the setting of the shutdown facility is completed.

- "InitFailed" is displayed as the initial status.

- "Unknown" or "TestFailed" is displayed as the test status.

In this case, review the configuration setting of the agent or hardware.

. . . .

8.11 Changing Blade Settings

This section describes changing Blade settings.

8.11.1 Changing the IP Address of the Management Blade

This section explains how to change the IP address of the management blade.

Operation Procedure:

1. Execute the following command on all nodes to stop the operation of the shutdown facility (SF):

```
# sdtool -e
```

2. Change the IP address according to the procedure of the management blade.

3. Define the changed IP address in the Shutdown Agent configuration file.

See

. . . .

For details on how to define the configuration file, see "5.1.2.3.3 Blade."

. . . .

4. Execute the following command on any node to apply changes of the configuration file.

```
# /etc/opt/FJSVcllkcd/bin/panicinfo_setup
```

After the following message is displayed, select "I."

```
panicinfo_setup: WARNING: /etc/panicinfo.conf file already exists.
(I)nitialize, (C)opy or (Q)uit (I/C/Q) ?
```

5. Execute the following command on all nodes to start the shutdown facility.

```
# sdtool -b
```

6. Execute the following command on all nodes and check that the shutdown facility operates normally.

```
# sdtool -s
```

8.11.2 Changing the Slot Number of Server Blades

This section explains how to change the slot number of server blades.

Operation Procedure:

1. Execute the following command on all nodes to stop the shutdown facility.

   ```
   # sdtool -e
   ```

2. Change the slot position according to procedure for the server blade.

3. Define the changed slot number of the server blade in the Shutdown Agent configuration file.

   See

   ·

   'For details on how to define the configuration file, see "5.1.2.3.3 Blade."

   ·

4. Execute the following command on any node to apply changes of the configuration file.

   ```
   # /etc/opt/FJSVcllkcd/bin/panicinfo_setup
   ```

   After the following message is displayed, select "I."

   ```
   panicinfo_setup: WARNING: /etc/panicinfo.conf file already exists.
   (I)nitialize, (C)opy or (Q)uit (I/C/Q) ?
   ```

5. Execute the following command on all nodes to start the shutdown facility.

   ```
   # sdtool -b
   ```

6. Execute the following command on all nodes and check that the shutdown facility operates normally.

   ```
   # sdtool -s
   ```

8.12 Adding, Deleting, and Changing Hardware

This section describes how to add, delete, and change the following hardware in the existing configuration:

- Shared disk device

- Network interface card used for the public LAN and the administrative LAN

Note

- When you change a system board, reconfigure BMC or iRMC used by the shutdown facility.

- When you change a system board or a network interface card, do not restart the network.

8.12.1 Adding Hardware

This section describes how to add hardware.

8.12.1.1 Adding the shared disk device

The following describes how to add the shared disk device.

Figure 8.11 Procedure to add the shared disk device
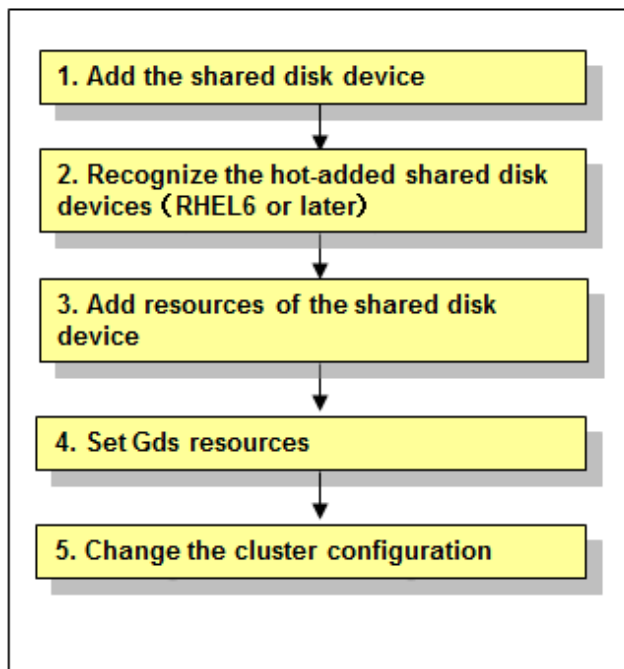
Information

You must stop RMS during performing "5. Change the cluster configuration."

However, you do not need to stop RMS if all the following conditions are met because performing "5. Change the cluster configuration" is not necessary under the condition:

- The added shared disk device is registered with the existing class of GDS.

- The added shared disk device is no used as Fsystem resource.

Operation Procedure:

1. Add the shared disk device.

   See "9.2 Maintenance Flow" and ask your Fujitsu customer engineer (CE) to add the shared disk device.

2. Recognize the hot-added shared disk devices (RHEL6 or later)

   When the shared disk devices are hot-added on RHEL6 or later version, the following setting is required for all the nodes that include the hot-added shared disk devices.

   ```
   # /etc/opt/FJSVsdx/bin/sdx_by_id replace device_name
   ```

   💠 Example
   ....
   When the names of the hot-added disk devices are sdf, sdg, and sdh

   ```
   # /etc/opt/FJSVsdx/bin/sdx_by_id replace sdf,sdg,sdh
   ```
   ....

3. Add resources of the shared disk device.

   Register resources corresponding to the added shared disk device to the resource database.

   📖 See
   ....
   To register resources, see "5.1.3.2 Registering Hardware Devices."
   ....

4. Set up Gds resources.

   To use GDS, set up GDS and create Gds resources.

   If you register the added shared disk device with the existing class of GDS, you do not need to set Gds resources.

   📖 See
   ....
   For information on how to set up GDS and create Gds resources, see "6.3 GDS Configuration Setup," "6.7.3.3 Preliminary Setup for Gds Resources," and "6.7.3.4 Setting Up Gds Resources."
   ....

5. Change the cluster configuration.

   Change the cluster configuration to add the following resources related to the added shared disk device to a cluster application.

   - Fsystem resource

   - Gds resource

   📖 See
   ....
   For information on how to change the cluster configuration, see "8.1 Changing the Cluster Configuration"
   ....

8.12.1.2 Adding a Network Interface Card Used for the Public LAN and the Administrative LAN

This section describes how to add a network interface card used for the public LAN and the Administrative LAN.

Figure 8.12 Procedure to add a network interface card
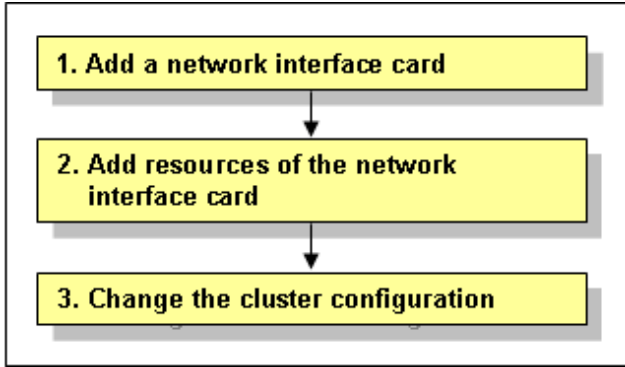
Operation Procedure:

1. Add a network interface card.

   See "9.2 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 to the resource database.

   ![See]
   . . . .
   To register resources, see "5.1.3.2 Registering Hardware Devices."
   . . . .

3. Change the cluster configuration.

   Change the cluster configuration to add the following resources related to the added network interface card to a cluster application.

   - Takeover network resource

   - Gls resource

   ![See]
   . . . .
   For information on how to change the cluster configuration, see "8.1 Changing the Cluster Configuration"
   . . . .

8.12.2 Deleting Hardware

This section describes how to delete hardware.

8.12.2.1 Deleting the shared disk device

To delete the shared disk device, you need to change the cluster configuration which includes resources of the shared disk device to be deleted beforehand.

Figure 8.13 Procedure to delete the shared disk device
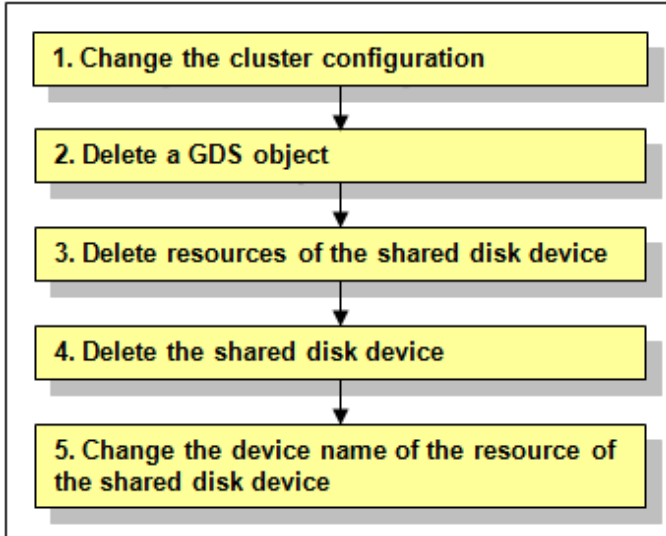
Operation Procedure:

1. Change the cluster configuration.

   Delete the following resources using the shared disk device to be deleted from the cluster configuration:

   - Fsystem resource

   - Gds resource

   ![See] **See**
   ....
   To change the configuration of a cluster application and delete resources, see "8.1 Changing the Cluster Configuration" and "8.7.1 Deleting a Resource."
   ....

2. Delete a GDS object.

   Delete a GDS object related to the shared disk device to be deleted.

   ![See] **See**
   ....
   To delete a GDS object, see "Removals" of "Chapter 5 Operation" in the "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."
   ....

3. Delete resources of the shared disk device.

   Delete resources of the registered shared disk device by using the "cldelrsc" command.

   For details on the "cldelrsc" command, see the manual page.

   After executing the "cldelrsc" command, execute the following command to inform that resources are deleted to GDS.

   Specify the full path of an empty file for filepath.

   ```
   # /etc/opt/FJSVcluster/bin/clautoconfig -f filepath
   ```

   ![Note] **Note**
   ....
   - When the shared disk device, from which resources are to be deleted, is 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 a GDS class, see the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide.*"

- When SDX_UDEV_USE=off is described in the GDS configuration file /etc/opt/FJSVsdx/sdx.cf, do not execute the clautoconfig command.

4. Delete the shared disk device.

   See "9.2 Maintenance Flow" and ask your Fujitsu customer engineer (CE) to delete the shared disk device.

5. Change the device name of the resource of the shared disk device.

   By deleting the shared disk device, any device name of the shared disk device which has not been deleted may be changed. To modify the device name of the resource of the shared disk device according to the correct device name, execute the following command. Specify the full path of an empty file for *filepath*.

   ```
   # /etc/opt/FJSVcluster/bin/clautoconfig -f filepath
   ```

   ![Note]

   When SDX_UDEV_USE=off is described in the GDS configuration file /etc/opt/FJSVsdx/sdx.cf, do not perform Step 5.

8.12.2.2 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 the cluster configuration which includes resources of the network interface card to be deleted beforehand.

Figure 8.14 Procedure to delete a network interface card
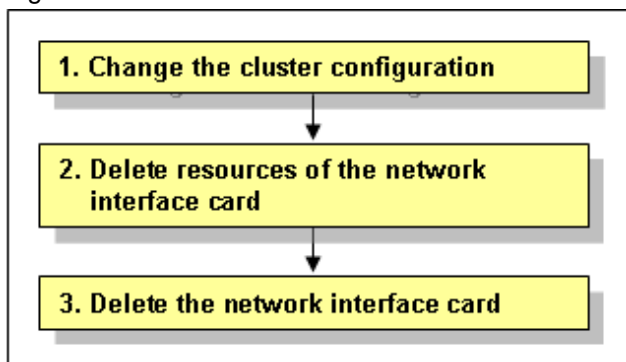
Operation Procedure:

1. Change the cluster configuration.

   Delete the following resources using the network interface card to be deleted from the cluster configuration:

   - Takeover network resource

   - Gls resource

   ![See]

   To change the configuration of a cluster application and delete resources, see "8.1 Changing the Cluster Configuration."

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 "9.2 Maintenance Flow" and ask your Fujitsu customer engineer (CE) to delete the network interface card.

8.12.3 Changing Hardware

This section describes how to change hardware.

8.12.3.1 Changing the shared disk device

To change the shared disk device, you need to delete resources of the target shared disk device beforehand. After the change, you need to add the resources of the changed shared disk device.

Figure 8.15 Changing the shared disk device
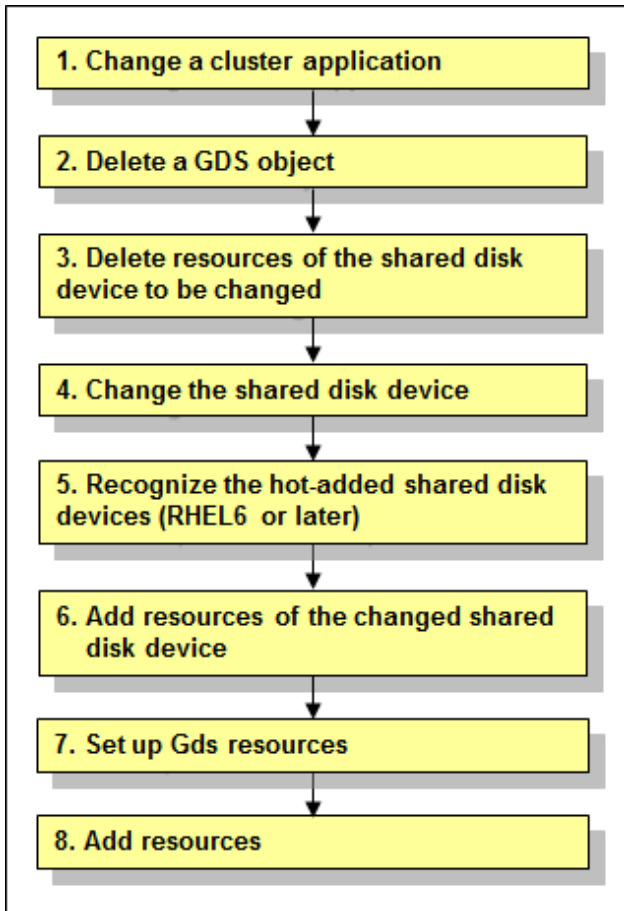
Operation Procedure:

1. Change a cluster application.

   Delete the following resources, which are using the shared disk device to be changed, from the cluster application:

   - Fsystem resource

   - Gds resource

   ### See
   . . . .
   For details on how to change the cluster application configuration and delete resources, see "8.1 Changing the Cluster Configuration" and "8.7.1 Deleting a Resource."
   . . . .

2. Delete a GDS object.

   Delete a GDS object related to the shared disk device to be changed.

To delete a GDS object, see "Removals" of "Chapter 5 Operation" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*."

3. Delete resources of the shared disk device to be changed.

   Delete resources of the registered shared disk device by using the "cldelrsc" command.

   For details on the "cldelrsc" command, see the manual page.

   After executing the "cldelrsc" command, execute the following command to inform that resources are deleted to GDS.

   Specify the full path of an empty file for *filepath*.

   ```
   # /etc/opt/FJSVcluster/bin/clautoconfig -f filepath
   ```

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 a GDS class, see the "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

- When SDX_UDEV_USE=off is described in the GDS configuration file /etc/opt/FJSVsdx/sdx.cf, do not execute the clautoconfig command.

4. Change the shared disk device.

   See "9.2 Maintenance Flow" and ask field engineers to change the shared disk device.

5. Recognize the hot-added shared disk devices (RHEL6 or later)

   When the shared disk devices are hot-added on RHEL6 or later version, the following setting is required for all the nodes that include the hot-added shared disk devices.

   ```
   # /etc/opt/FJSVsdx/bin/sdx_by_id replace device_name
   ```

Example

When the names of the hot-added disk devices are sdf, sdg, and sdh

```
# /etc/opt/FJSVsdx/bin/sdx_by_id replace sdf,sdg,sdh
```

6. Add resources of the changed shared disk device.

   Register resources corresponding to the changed shared disk device to the resource database.

See

For information on how to register the resource database, see "5.1.3.2 Registering Hardware Devices."

7. Set up 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 "6.3 GDS Configuration Setup" and "6.7.3.4 Setting Up Gds Resources."

8. Add resources.

   If you have deleted Fsystem resources in Step 1, add Fsystem resources.

   ![See icon] **See**

   ....
   To add resources, see "6.7.3 Setting Up Resources."
   ....

8.12.3.2 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 resources of the target network interface card beforehand. After the change, you need to add resources of the network interface card.

Figure 8.16 Procedure to change a network interface card
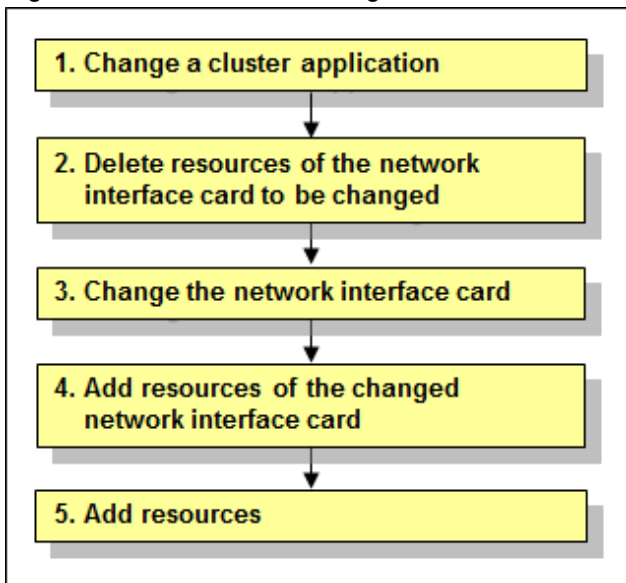
Operation Procedure:

1. Change a cluster application.

   Delete the following resources, which are using the network interface card to be changed, from the cluster application:

   - Takeover network resource

   - Gls resource

   ![See icon] **See**

   ....
   For details on how to change the cluster application configuration and delete resources, see "8.1 Changing the Cluster Configuration" and "8.7.1 Deleting a Resource."
   ....

2. 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.

3. Change the network interface card.

   See "9.2 Maintenance Flow" and ask field engineers to change the network interface card.

4. Add resources of the changed network interface card.

   Register resources corresponding to the changed network interface card to the resource database.

5. Add resources.

   If you have deleted takeover network resources and Gls resources in Step 1, add takeover network resources and Gls resources.

Part 5   Maintenance

This part explains the procedure for maintaining the PRIMECLUSTER system.

Chapter 9 Maintenance of the PRIMECLUSTER System

This chapter explains items and procedures related to maintenance of the PRIMECLUSTER system.

9.1 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.

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).

9.2 Maintenance Flow

Your Fujitsu customer engineer (CE) shall perform parts replacement. The flow of maintenance is as follows:

When stop maintenance is to be performed

1. All nodes of the running PRIMECLUSTER system shall be stopped by the 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 "9.2.1 Detaching Resources from Operation."

2. Your Fujitsu customer engineer (CE) shall confirm the target equipment and 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 "9.2.2 Executing Standby Restoration for an Operating Job."

9.2.1 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" 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" command.

💠 **See**
. . . .
For details on how to stop RMS, see "7.1.3 Stopping RMS" in "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide."
. . . .

Stopping a node

Execute the "shutdown(8)" command to stop the relevant node.

9.2.2 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."
. . . .

9.3 Software Maintenance

This section provides notes on intensive correction and formal repair, as well as the procedure for applying them to the PRIMECLUSTER system.

9.3.1 Notes on Applying Corrections to the PRIMECLUSTER System

Note the following when you apply intensive correction 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 an intensive 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).

9.3.2 Overview of the Correction Application Procedure

This section provides an overview of the procedure for applying corrections such as an intensive correction to a cluster system. The procedure explained here is a general procedure.

9.3.2.1 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
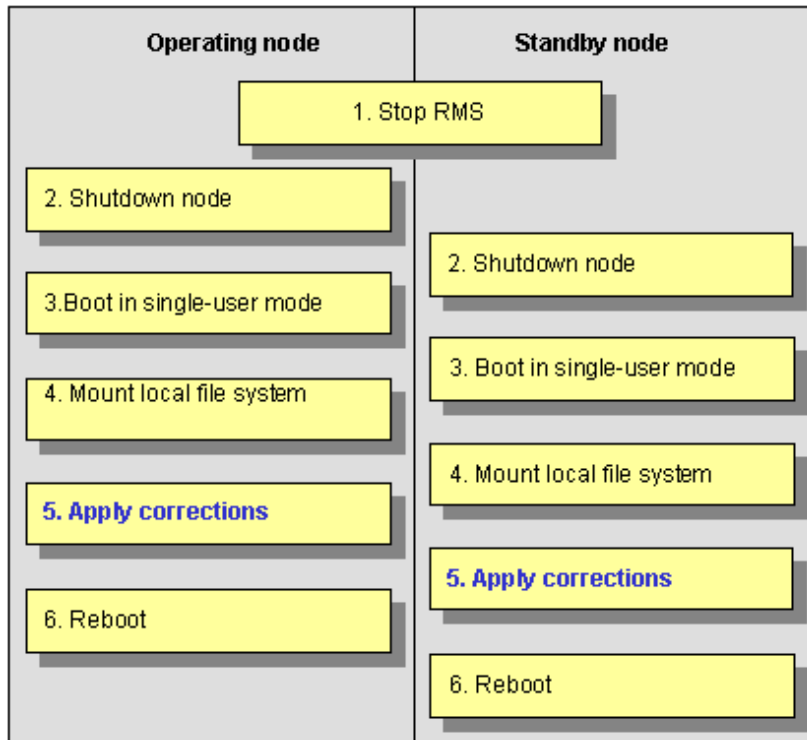
Procedure

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.

3. Boot in single-user mode.

   Boot all the nodes that were shut down in single-user mode.

4. Mount the local file system.

   Mount the required local file system on all the nodes.

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 -r**.

🖐 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."

. . . .

9.3.2.2 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
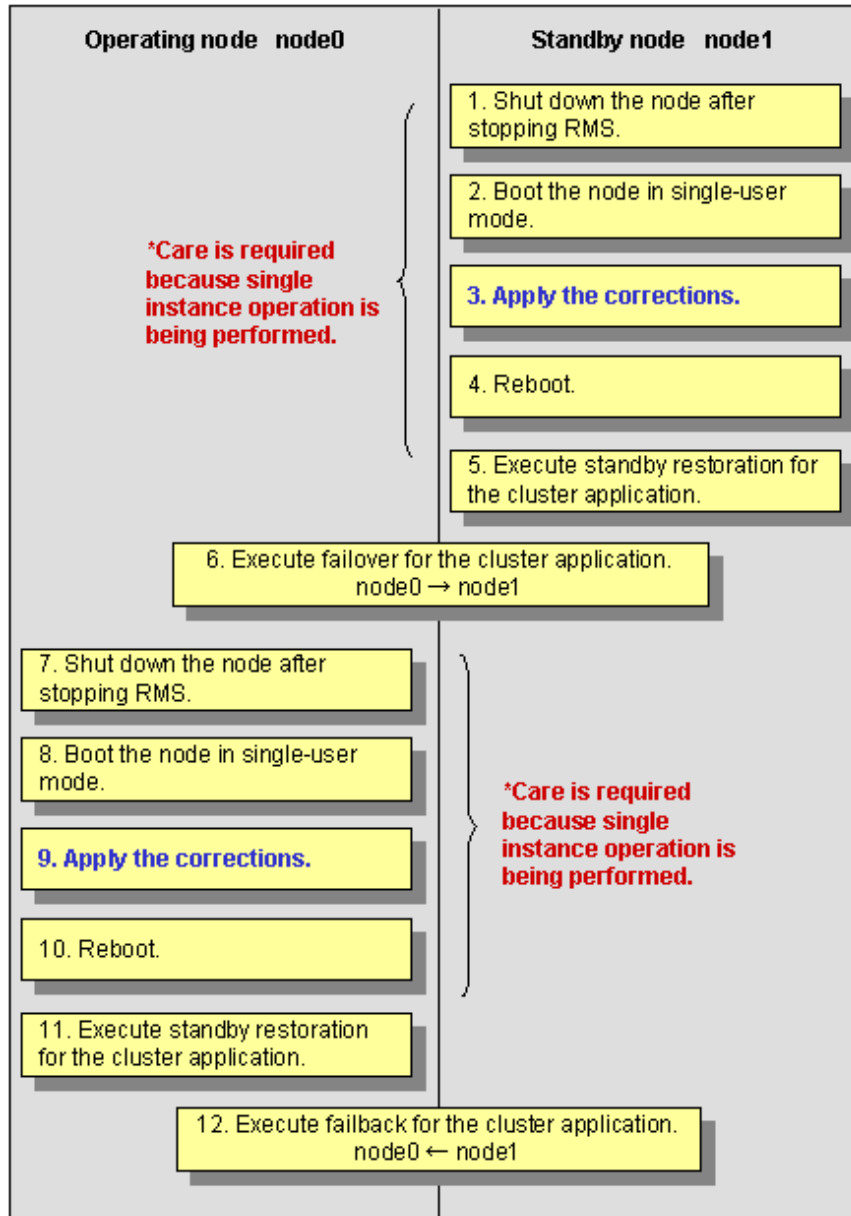
Procedure

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 all cluster applications to the standby node (node1) with the "**hvswitch**" command.

   If there is a scalable cluster application, switch all standby cluster applications under the scalable cluster application first, then switch the scalable cluster application.

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 shutdown 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."

Part 6   PRIMECLUSTER Products

This section explains the PRIMECLUSTER products.

- PRIMECLUSTER Product List

- PRIMECLUSTER Wizard for Oracle

- PRIMECLUSTER Wizard for NAS

- Interstage Application Server Enterprise Edition

- Systemwalker Centric Manager

- Systemwalker Operation Manager

Chapter 10 PRIMECLUSTER Product List

The following lists the version levels of PRIMECLUSTER products and the range of support:

Table 10-1 PRIMECLUSTER Product List

Product name	Operation mode	Standby operation					Scalable operation		Note
	System Configuration	Standby		Mutual standby	Cascade	Priority transfer	Scalble	High-availability acalable	
		1:1	N:1						
PRIMECLUSTER Wizard for Oracle	Y	Y	Y	Y	Y	Y	Y	N	For correspondence of Oracle Database, refer to the manual for PRIMECLUSTER Wizard for Oracle.
PRIMECLUSTER Wizard for NAS	Y	Y	Y	Y	Y	Y	Y	N	
PRIMECLUSTER Wizard for PowerGres Plus	Y	N	N	N	N	N	N	Unsupported on PRIMEQUEST
Interstage Application Server Enterprise Edition	Y	N	N	N	N	N	N	
Symfoware Server	Y	N	N	N	N	N	N	Unsupported on PRIMEQUEST
Symfoware Server Hot Standby Option	Y	N	N	N	N	N	N	
Systemwalker Centric Manager Enterprise Edition Server License (for manager)	Y	N	N	N	N	N	N	
Systemwalker Centric Manager Enterprise Edition Server License (for agent)	Y	N	N	Y	N	N	N	
Systemwalker Operation Manager Enterprise Edition Server License	Y	Y	Y	N	N	N	N	
Symfoware Server Enterprise Extended Edition	Y	Y	N	Y	Y	Y	Y	Unsupported on PRIMERGY
Symfoware Server Parallel Cluster Option	Y	Y	N	Y	Y	Y	Y	
Systemwalker Service Quality Coordinator Enterprise Edition	*1:Supported only in the manager	N	N	N	N	N	N	

Systemwalker Centric Manager Enterprise Edition	Y	N	N	*2:Supported only in the job server	N	N	N
Systemwalker Operation Manager Enterprise Edition	Y	Y	Y	N	N	N	N
Systemwalker Software Delivery Enterprise Edition	Y	N	N	N	N	N	N

Notes:

The "Y" and "N" marks in the above table indicate the range of PRIMECLUSTER-supported products.

For details, see the manuals for the product.

10.1 PRIMECLUSTER Operation Modes

This section explains the cluster operation mode.

System 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.
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 application operation mode

Cluster operation mode	Description
Standby operation mode	General names for the five operation modes are 1:1 standby, N:1 standby, mutual standby, cascade, and priority transfer.
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 of GLS	Refers to the function for taking over one or more IP addresses from an old operating node to a new operating node if failover occurs during standby operation (this function can also be applied to two or more target network interface cards).

Chapter 11 PRIMECLUSTER Wizard for Oracle

11.1 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 Oracle Real Application Clusters ("RAC" or "Oracle RAC 10g"hereafter). The standby operation does not require Oracle RAC, but requires Oracle cold-standby instead.

Startup and stop control

In line with the state transition of the userApplication, 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.

In the operation of Oracle RAC 10g, starting and stopping of Oracle services are also controlled.

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 table space, 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. Oracle ORA-XXXX errors are detected after the SQL command is executed, failover and restart of Oracle instance are automatically initiated according to the error level. 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.

In the operation of Oracle RAC 10g, monitoring of the Oracle service, Oracle instance, and Oracle listener are carried, while recovery and degeneration in the event of failures are performed by Oracle RAC 10g CRS (Cluster Ready Services).

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.

11.2 Operation Environment

11.2.1 System Configuration

System configuration figures are shown below:

Scalable operation

In scalable operation with RAC, Oracle is operated on all nodes. A client can use a database that is connected to either node.
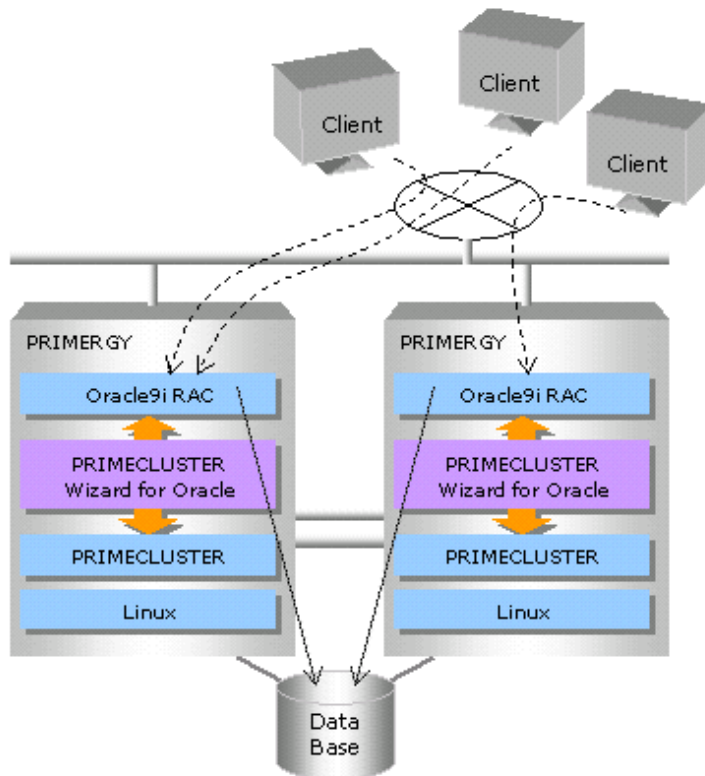
Standby operation

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.
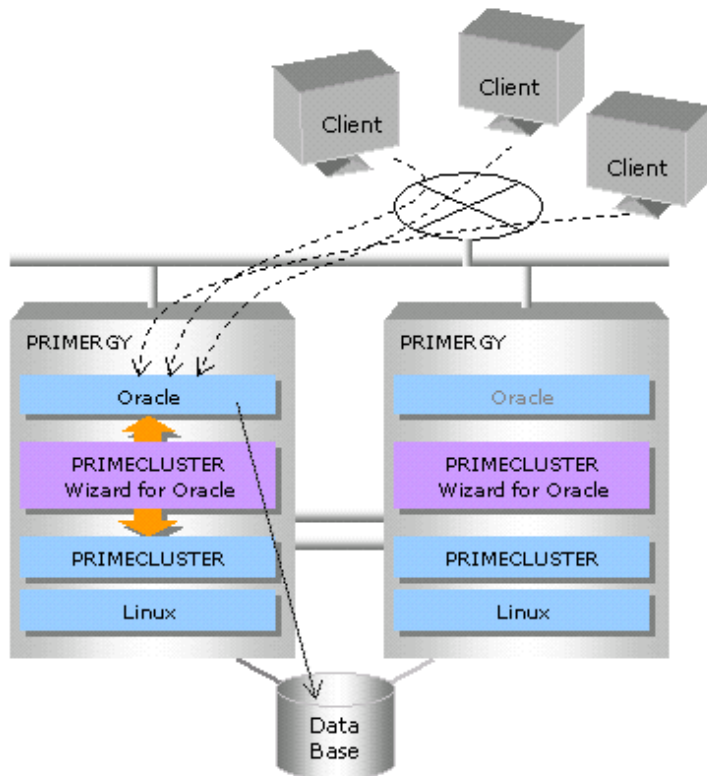
11.2.2 Supported Modes

For details, see the "Table 10-1 PRIMECLUSTER Product List."

11.2.3 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 RAC and Oracle RAC 10g standby operation.

11.3 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 12 PRIMECLUSTER Wizard for NAS

12.1 Functional Overview

PRIMECLUSTER Wizard for NAS is a software product that must be installed when the ETERNUS NR1000F series (NetApp(TM) Filer: Refers to the disk device manufactured by Network Appliance Inc. Hereinafter abbreviated to NAS device.) is being used as the PRIMECLUSTER shared disk device. This software releases the NFS file lock that can cause a problem upon the occurrence of a node down or application error, and then performs failover.

Startup and stop control

According to the state transition of userApplication, automatic control of export/unexport is allowed for the volume of the network attached storage (NAS) device. For standby operation, the file lock, retained by the old operation node, is released at startup.

Monitoring

PRIMECLUSTER Wizard for NAS does not monitor errors in the NAS device. Other resources, however, support error detection and failover control.

   - PRIMECLUSTER GLS monitors the access path to the NAS device.

   - The RemoteFileSystem resource, which allows remote file system mounting control, monitors accesses to the volume.

Environment setup

The environment setup wizard for running a NAS device as a shared disk on PRIMECLUSTER allows you to configure a cluster environment based on a NAS device easily and safely.

12.2 Operation Environment

12.2.1 System Configuration

System configuration figures are shown below:

Scalable operation

In scalable operation, the volumes of a NAS device can be mounted from any node. Also, the shared disk can be used from every node.
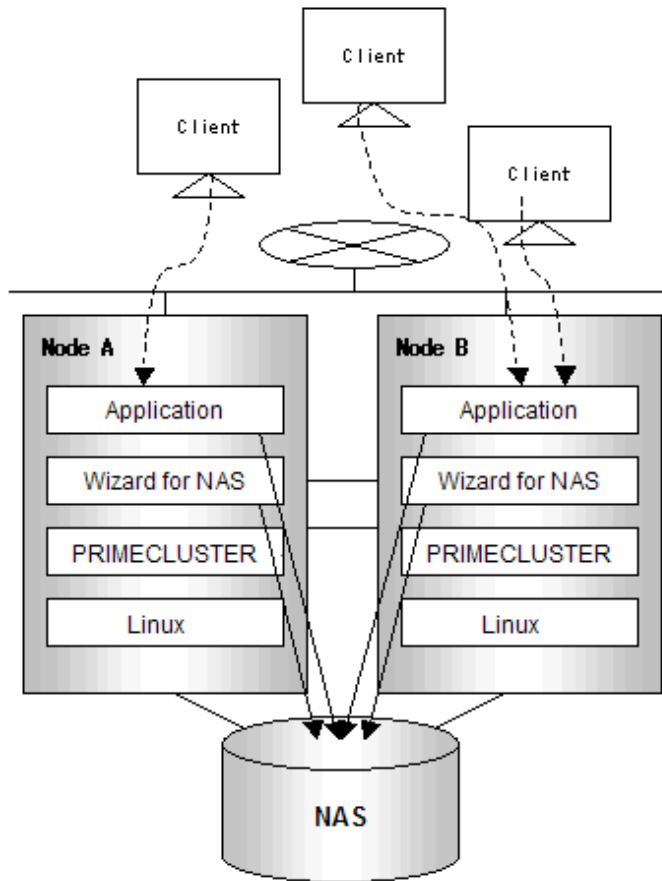
Standby operation (cold-standby)

In standby operation, a cluster system consists of an operational node and standby nodes. With a NAS device, however, a connection is established only on the operational node and that is inactive on the standby node. In the event of a failure, production operation is switched to one of the standby nodes, after which a connection to the standby node is established after the operational node lock on the NAS device is released.
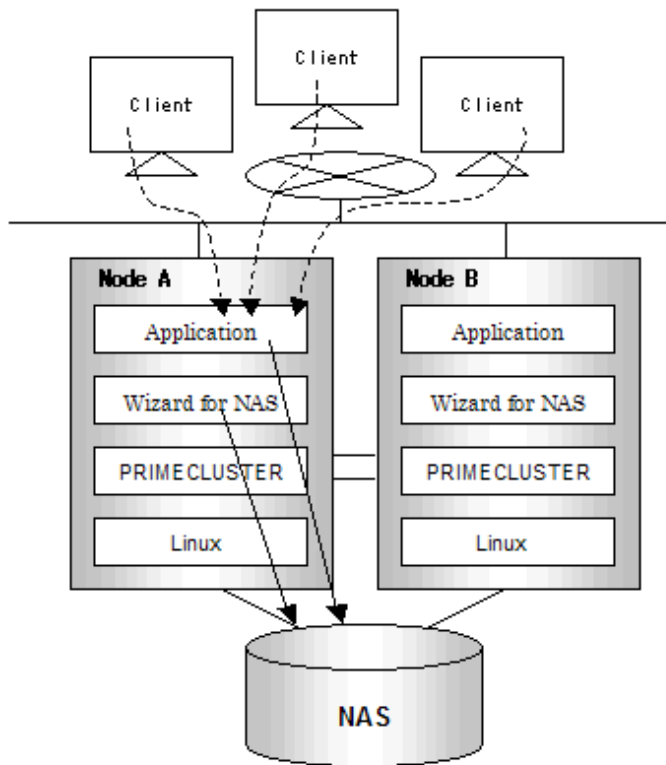
![Information]
Information

· ·

Hot-standby is not supported.

· ·

12.2.2 Supported Modes

For details, see the "Table 10-1 PRIMECLUSTER Product List."

12.2.3 Overview of Cluster System Support

The PRIMECLUSTER Wizard for NAS provides the environment setup wizard to enable the use of the NAS device as a shared disk of the cluster system, the start/stop script, the detector, etc. and supports an environment allowing the combined use of PRIMECLUSTER and NAS devices.

12.3 Notes

For details on the environment setup, etc., see the "*PRIMECLUSTER Wizard for NAS Configuration and Administration Guide*," which is attached to the PRIMECLUSTER Wizard for NAS.

Chapter 13 Interstage Application Server Enterprise Edition

13.1 Functional Overview

The Interstage Application Server is a high-reliability, high-performance application server that supports a simple Web system, large-scale E-commerce, and business-to-business enterprise systems.

Interstage Application Server Enterprise Edition is the most advanced version of Interstage Application Server. Interstage Application Server Enterprise Edition supports hot standby and load balancing, and is the most appropriate choice for large-scale systems that must provide stable responses and round-the-clock non-stop operation.

13.1.1 Overview of Cluster System Support

The Interstage Application Server realizes high availability (HA) by enabling the configuration of a cluster system.

This high availability is supported by an operation mode called 1:1 standby.

A server having the same configuration as the currently operating server is provided as a standby. If the operation server experiences a failure, the standby server takes over so that the transactions can be restarted swiftly.

13.1.2 Supported Modes

For details, see the "Table 10-1 PRIMECLUSTER Product List."

13.1.3 Notes

For details on the cluster system configuration, see the "*Interstage Application Server High-Reliability System Operation Guide.*"

Chapter 14 Systemwalker Centric Manager

14.1 Functional Overview

Systemwalker Centric Manager is an integrated operation management tool that has been designed to be used in an enterprise environment. Systemwalker Centric Manager provides an optimum means of enabling simple operation from anywhere, be it in a client/server environment, for which multiple servers are used, or a business environment, for which the Internet is used.

14.1.1 Overview of Cluster System Support

(1) Monitoring a cluster system

The Job Servers running on cluster system can be managed.

Multiple node cluster servers (Job Servers) configured using Microsoft Cluster Service and PRIMECLUSTER can be also monitored.

(2) Supporting of a cluster with the management server

The management facility of Systemwalker can operate on the cluster system.

High-availability monitoring can be realized because, even if an error occurs in either the Operation Management Server or the Section Management Server, the monitoring works can be taken over by another node.

![See icon] See
....
For details, see the cluster setup guides for Systemwalker Centric Manager.
....

14.1.2 Supported Modes

For details, see the "Table 10-1 PRIMECLUSTER Product List."

Chapter 15 Systemwalker Operation Manager

15.1 Functional Overview

Systemwalker Operation Manager enables the automatic operation of a server/client system that is distributed over a network, the scheduling of routine batch jobs and online jobs, the control of job execution, the monitoring of the state of execution, and the provision of an achievement management function to fully support a business.

Furthermore, the use of Systemwalker Operation Manager greatly reduces the system administrator' s workload.

15.1.1 Overview of Cluster System Support

High availability for batch job execution can be realized through the use of a cluster system.

See
. . . .

For details, see the cluster setup guides for Systemwalker Operation Manager.
. . . .

15.1.2 Supported Modes

For details, see the "Table 10-1 PRIMECLUSTER Product List."

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.

CE Instruction Worksheet

This worksheet compiles necessary information for installation and configuration, and is used to communicate instructions to the Fujitsu customer engineer (CE).

Cluster Installation Environment Setup Worksheet

Use this worksheet to configure the environment before you install Linux(R), PRIMECLUSTER, and other required or related software. Also use the worksheet to configure the PRIMECLUSTER system using Web-Based Admin View. This worksheet facilitates the installation of software products.

Kernel Parameter Worksheet

This worksheet allows you to determine the fundamental PRIMECLUSTER values for estimating the kernel parameters. After installing the Linux(R) 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. There are individual worksheets for configuring RMS, GFS, GLS, GDS, and cluster applications.

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.

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	<>Scalable operation
( )2 nodes	
( )3 or more nodes	
(5) Web-Based Admin View operation modes	
( )2-layer model	<>1 PC required
( )3-layer model	<>Operation management server and 1 PC required
(6) Shared disk units selection	
1)2-layer model	
<>RAID	[ ]Multipath disk control facility
<>Interface	<>ETMPD (ETERNUS multipath driver)
<>Fibre Channel	
<>Path	[ ]GDS (Global Disk Services)
<>Single path	
<>Path duplex switchover	
<>Manage disk cabinet	
2)Switchover file system selection	
<>ext2, ext3	
<>High reliability above ext2, ext3	[ ]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)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		

(8)Crash dump collection	
( )kdump	<>The local disk must have a dump collection area.

A.4 CE (FE for North America) Instruction Worksheet

The CE Instruction Worksheet indicates information required for hardware installation for a Fujitsu customer engineer (CE). This worksheet is also used for specifying information that was configured by the CE.

For information on hardware installation in the field, see the cluster field installation operation procedure.

- PRIMERGY
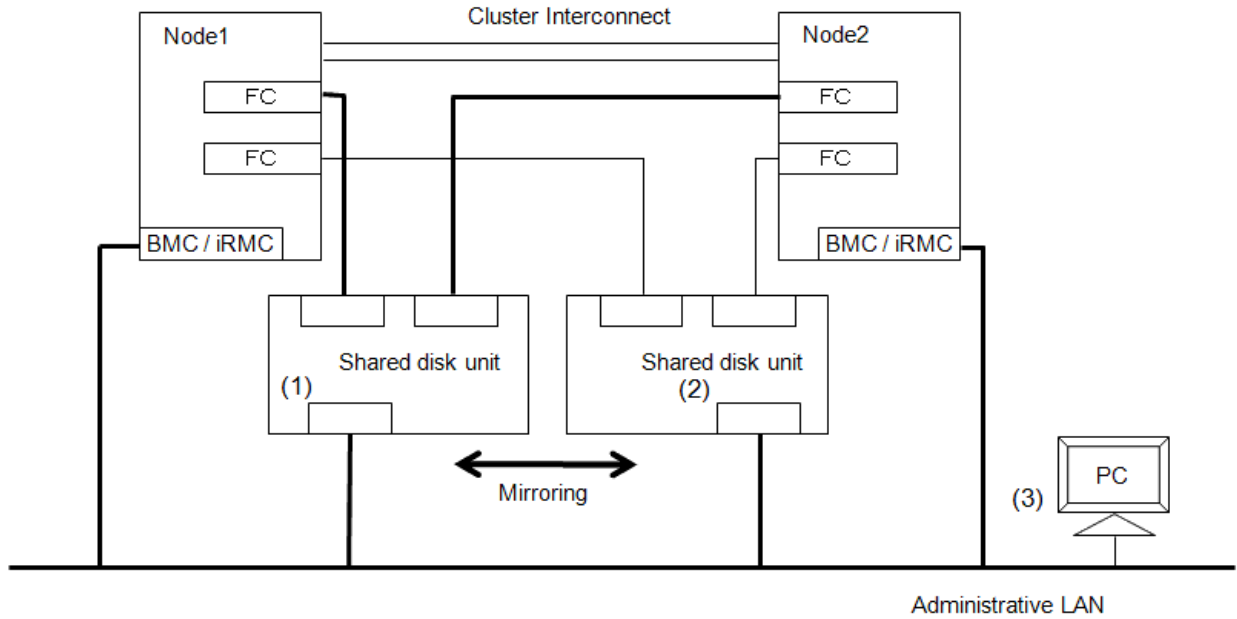
Shared RAID unit	IP address (1)	__ · __ · __ · __
	IP address (2)	__ · __ · __ · __
	Gateway	__ · __ · __ · __
	Subnet mask	__ · __ · __ · __
ServerView management machine	IP address (3)	__ · __ · __ · __

- PRIMEQUEST
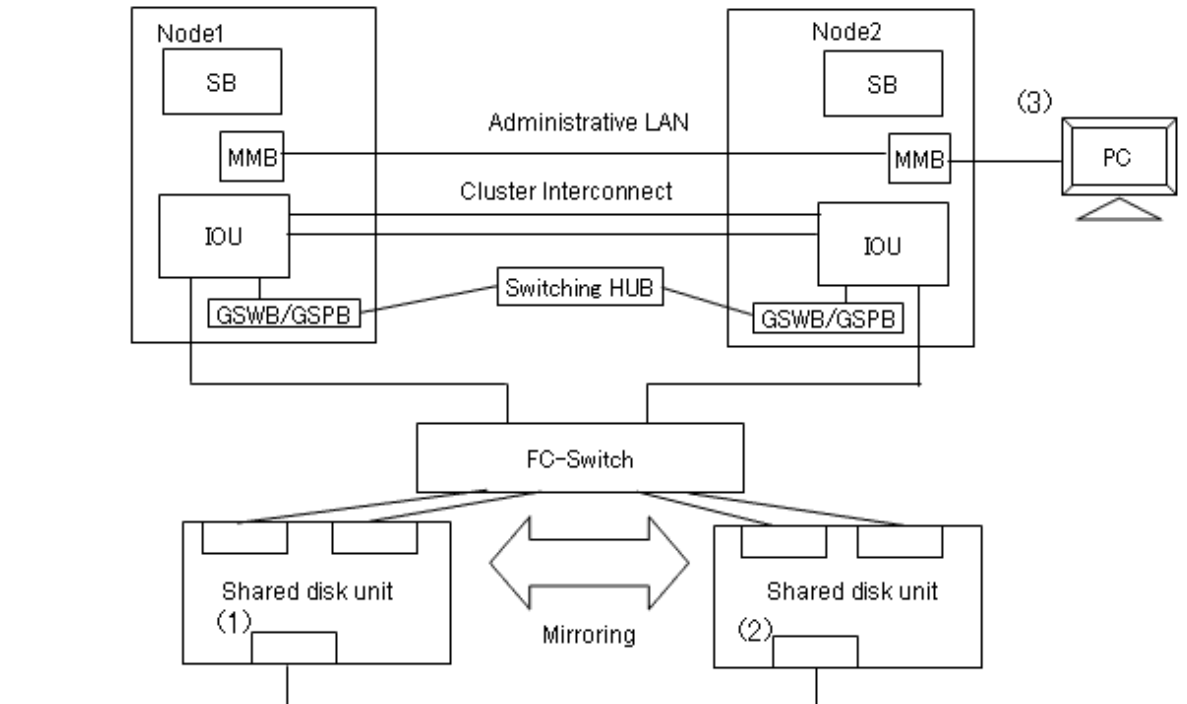
Shared RAID unit	IP address (1)	__ . __ . __ . __
	IP address (2)	__ . __ . __ . __
	Gateway	__ . __ . __ . __
	Subnet mask	__ . __ . __ . __
MMB console	IP address (3)	__ . __ . __ . __

A.5 Cluster Installation Environment Worksheet

Note
....
Use a single primary management server of Web-Based Admin View to monitor a single cluster system.
....

The Cluster Installation Environment Worksheet organizes items related to installation of Linux(R) software, PRIMECLUSTER, required and related software. Create the Cluster Installation Environment Worksheet.

The Cluster Installation Environment Worksheet contains initial configuration items related to the entire PRIMECLUSTER system. Configure a system on each node according to the information specified here.

| Cluster system name: | | Cluster Installation Environment Worksheet |
| Node name: | | |

Settings		Notes
(1) Software to be installed	[ ]Optional software	[ ]ETERNUS multipath driver
(2) Version of Linux	<> Red Hat Enterprise Linux 5.3 or later <> Red Hat Enterprise Linux 6.X	Use the same version on all nodes.

Settings			Notes		
(3) Partition information			Logical path	Size	Configuration should be synchronized on all nodes
	/boot				
	/var				
	/				
	/opt				
	/var/crash				
	/usr				
	/home				
	/var/opt				
	/etc/opt				
	/tmp				
(4) System locale	◇ C ◇ Ja ◇ Other ( )				
(5) Time zone	◇ Japan ◇ Other ( )				
(6) Node information Node name IP					
(7)Web-Based Admin View					
	Node name	IP	Access I/F	Http port number	RMI port number
Primary management server	_____ _____	_____ _____	_____ _____	_____ _____	_____ _____
Secondary management server	_____ _____	_____ _____	_____ _____	_____ _____	_____ _____

A.6 Kernel Parameter Worksheet

When you build the PRIMECLUSTER system, the default values for the kernel parameters of the Linux(R) may be inadequate. The Kernel Parameter Worksheet shows the kernel parameter values used for PRIMECLUSTER. If the values have been set in /etc/sysctl.conf, you need to edit the values (Defined values + kernel parameter values used for PRIMECLUSTER). If the values have not been set, you need to set the values (Linux(R) default values + the kernel parameter values used for PRIMECLUSTER.)

The kernel parameter values differ depending upon:

- CF configuration

- RMS configuration

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.

A.6.1 CF Configuration

To enable the resource database, you must change the default value of the Linux(R) kernel parameter.

The kernel parameters can be configured with the sysctl (8) utility when they are executed. The procedures for changing the kernel parameters and setting up a new kernel parameter are explained below.

The value of SHMMAX can be changed as follows:

- Edit the /etc/sysctl.conf file.

- Add the following entry. kernel.shmmax =value

Change the semaphore parameter values as follows:

```
kernel.sem = value1 value2 value3 value4
```

Define the original values as follows:

```
value1 = value of SEMMSL
```

```
value2 = value to SEMMNS
```

```
value3 = value of SEMOPM
```

```
value4 = value of SEMMNI
```

As the resource manager does not use SEMMSL or SEMOPM, you do not need to change these values. If you execute the following commands, the original parameter values will be displayed.

#**sysctl -a | grep sem**

#**sysctl -a | grep shm**

If you execute the following command, each parameter will be changed.

#**sysctl -p** *filename* (the default value of *filename* is /etc/sysctl.conf)

This command enables the value of /etc/sysctl.conf for the kernel. You do not have to reconfigure the kernel.
To change the kernel parameter while the OS is running, use sysctl (8). For details on the operation of sysctl, see the manual pages describing sysctl (8).

Kernel parameter	Minimum value	Comments (parameter description)
kernel.semmni	20	Value required for resource database
kernel.semmns	30	Value required for resource database
kernel.shmmax	1048576+ Value required for the resource database *	Value required for resource database
kernel.shmmni	30	Value required for resource database

*

Estimate the value required for resource database according to the following equation:

```
Value required for resource database = 2776 x number of resources
```

Estimate the number of resources according to the following equation:

```
Number of resources = Number of disks in shared system devices x (number of shared nodes +1) x 2
```

Specify the following in "Number of disks in shared system devices":

- For a disk array unit, specify the number of logical units (LUN).

- For other than a disk array unit, specify the number of physical units.

Specify the number of nodes connected to the shared disk in "number of shared nodes."

![Note icon] **Note**

....

For system expansion, if you increase the logical disks, you need to re-estimate the resources and restart each node in the cluster system. If you add disks to the cluster after installation, you must then calculate the resources required for the total number of logical disks after addition.

....

A.6.2 RMS Configuration

The table below shows the values of the kernel parameters required for the RMS configuration.

If the kernel parameters do not come up to the following minimum values when RMS is installed, they will be automatically changed to the minimum values. The adjustable kernel values shown below will be replaced by the value of /etc/system.

Kernel parameter	Minimum value	Comments (parameter description)
kernel.msgmnb	4194304	Byte size of the message queue
kernel.msgmax	16384	Maximum size of the message text
kernel.msgmni	8192	Maximum number of message queues for the entire system

![Note icon] **Note**

....

- If you uninstall the RMS package from /etc/sysctl.conf by executing the "rpm -e" command, you need to restart the system after uninstallation.

- In PRIMECLUSTER, message queues are used for interprocess communication.

  When RMS is running, 2076 message queues are reserved from 0x4d2.

  If you are using message queues for any applications, use the range other than above (0x4d2 to 0xcee).

....

A.6.3 Using GFS

When using a GFS shared file system, you must change the kernel parameters shown in the table below.

Table A.1 Kernel parameters that must be changed

Kernel parameter	Characteristics	Value	Parameter description
SEMMNI value	Addition	2	Maximum value for semaphore identifiers that can be used in the entire system. Add 2 to the current value.
SEMMNS value	Addition	11	Maximum number of shared memory units that can be used in the entire system. Add 11 to the current value.

During execution, you can use the sysctl(8) utility to change the kernel parameters.

![See icon] **See**

....

For details about the sysctl operation, see the sysctl(8) manual page.

....

Described below is the procedure for changing the kernel parameters and setting new values.

1. Check the current values of the kernel parameters.

   To check the current effective values of the kernel parameters, execute the following command:

```
# sysctl -a | grep sem
```

```
kernel.sem = 20 90 10 20
```

The displayed values represent the following kernel parameters:

```
kernel.sem = SEMMSL-value SEMMNS-value SEMOPM-value SEMMNI-value
```

If the kernel parameters have not been changed in the /etc/sysctl.conf file, the following default values are displayed:

```
kernel.sem = 250 32000 32 128
```

2. Determine the kernel parameter values.

   The kernel parameter values are determined by the current effective values that were checked in step 1 and the values in the above table. However, if the /etc/sysctl.conf file still contains the default values, indicating that the kernel parameters have not been changed, the values do not need to be changed because they are large enough to handle the parameters used by GFS.

   If the current effective values of the kernel parameters are those shown in step 1, the values to be set are as follows:

   ```
   SEMMSL value: 20
   SEMMNS value: 101
   SEMOPM value: 10
   SEMMNI value: 22
   ```

3. Change the kernel parameters.

   1. Edit the settings file.

      To set the kernel parameters, use an editor and edit the /etc/sysctl.conf file. Change the values in the following row to the values that were determined in step 2:

      ```
      kernel.sem = SEMMSL-value SEMMNS-value SEMOPM-value SEMMNI-value
      ```

      If the example displayed in step 1 shows the current effective values of the kernel parameters, the edited line becomes the following:

      ```
      kernel.sem = 20 101 10 22
      ```

   2. Apply the settings.

      To change the kernel parameter values to the values in the settings file, execute the following command:

      ```
      # sysctl -p
      ```

4. Check the setting changes in the kernel parameters.

   To check whether the kernel parameter values were changed correctly, execute the following command and display the current values:

   ```
   # sysctl -a | egrep sem
   ```

   ```
   kernel.sem = 20 101 10 22
   ```

   Check that the displayed values are the values that were determined in step 2.

A.7 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.11.1.1 Cluster Configuration Worksheet" in the "A.11 System Design Examples (for the Virtual Machine Function not used)", "A.12.1.1 Cluster Configuration

. . . .

Item				Setting		Remarks
Product type				<>PRIMECLUSTER Enterprise Edition <>PRIMECLUSTER HA Server <>PRIMECLUSTER Clustering Base <>PRIMECLUSTER Lite Pack		
Shared disk unit: Type 1	Interface					
	Path					
	File system selection					
Shared disk unit: Type 2	Interface					
	Path					
	File system selection					
NTP	Operation mode					
	Protocol (only when broadcast is specified)					
	Network configuration	External NTP servers	Host name			
			IP address			
		NTP server settings in cluster	Primary NTP server	Host name		
				IP address		
			Secondary NTP server	Host name		
				IP address		
Operation management topology	Web-Based Admin View operation mode				layer	
	Public LAN					
	Primary management server IP					
	Secondary management server IP					
User group				Group name	User name	
				<>wvroot		
				<>clroot		
				<>cladmin		
				<>clmon		

Cluster name					
Node information	CIP	Number of subnets			
		Subnet IP			
		Net mask			
	Node 1	Node name (uname -n)		- 282 -	
		CF node name			
		Public LAN	Network interface name		
			IP address		
			Net mask		
			Gateway		
		Administrative LAN	Network interface name		Specify the information of an administrative LAN. When shared with a public LAN, no entry is necessary.
			IP address		
			Net mask		
		Cluster interconnect	Path 0 NIC name		
			Path 1 NIC name		
		weight			Node weight (priority if Split-Brain occurs) used by SF Note: See setup policy for live priority if Split-Brain occurs.
		BMC/iRMC	IP address		Enter the information of BMC/iRM installed in S3 or S4 model.
			User ID		
			Password		
		Dump save area for kdump			when kdump is used.
		Memory dump save area for guest domains			When the Virtual Machine Function is used, they are collected in this area.
		Slot No. of server BLADE			When a server BLADE is used.
		Node in CF quorum set		<>[Yes] <>No	Do not change the default setting.
	Node 2	Node name (uname -n)			
		CF node name			
		Public LAN	Network interface name		
			IP address		
			Net mask		
			Gateway		
		Administrative LAN	Network interface name		Specify the information of an administrative LAN. When shared with a public LAN, no entry is necessary.
			IP address		
			Net mask		
		Cluster interconnect	Path 0 NIC name		
			Path 1 NIC name		
		weight			Node weight (priority if Split-Brain occurs) used by SF Note: See setup policy for live priority if Split-Brain occurs.
		BMC/iRMC	IP address		Enter the information of BMC/iRM installed in S3 or S4 model.
			User ID		
			Password		
		Dump save area for kdump			when kdump is used.
		Memory dump save area for guest domains			When the Virtual Machine Function is used, they are collected in this area.
		Slot No. of server BLADE			When a server BLADE is used.
		Node in CF quorum set		<>[Yes] <>No	Do not change the default setting.

	Node 3	Node name (uname -n)			
		CF node name			
		Public LAN	Network interface name		
			IP address		
			Net mask	- 283 -	
			Gateway		
		Administrative LAN	Network interface name		Specify the information of an administrative LAN. When shared with a public LAN, no entry is necessary.
			IP address		
			Net mask		
		Cluster interconnect	Path 0 NIC name		
			Path 1 NIC name		
		weight			Node weight (priority if Split-Brain occurs) used by SF Note: See setup policy for live priority if Split-Brain occurs.
		BMC/iRMC	IP address		Enter the information of BMC/iRM installed in S3 or S4 model.
			User ID		
			Password		
		Dump save area for kdump			when kdump is used.
		Memory dump save area for guest domains			When the Virtual Machine Function is used, they are collected in this area.
		Slot No. of server BLADE			When a server BLADE is used.
		Node in CF quorum set		<>[Yes] <>No	Do not change the default setting.
	Node 4	Node name (uname -n)			
		CF node name			
		Public LAN	Network interface name		
			IP address		
			Net mask		
			Gateway		
		Administrative LAN	Network interface name		Specify the information of an administrative LAN. When shared with a public LAN, no entry is necessary.
			IP address		
			Net mask		
		Cluster interconnect	Path 0 NIC name		
			Path 1 NIC name		
		weight			Node weight (priority if Split-Brain occurs) used by SF Note: See setup policy for live priority if Split-Brain occurs.
		BMC/iRMC	IP address		Enter the information of BMC/iRM installed in S3 or S4 model.
			User ID		
			Password		
		Dump save area for kdump			when kdump is used.
		Memory dump save area for guest domains			When the Virtual Machine Function is used, they are collected in this area.
		Slot No. of server BLADE			When a server BLADE is used.
		Node in CF quorum set		<>[Yes] <>No	Do not change the default setting.

	CF remote service usage	[ ]cfcp (CF file copy) [ ]cfsh (CF remote command execution)	
	IP interconnect setup	<>[Not used] <>Used (IP interconnect count)	
	CIP	<>[Used by RMS] <>Not used by RMS Suffix:	
	Management Blade	IP address1		When a server BLADE is used, IP address 2 is used for the clustering among cases.
		IP address2	
		SNMP community name	
Cluster resource management facility	Hardware device to be registered in the resource database	[ ]Network device *1 [ ]Disk unit *2		*1) Required when public LAN takeover is used *2) Required when GDS and GFS are used
Shutdown facility (SF)	Means of stopping a node forcibly	Means	Operation after stop	
		[ ]IPMI	<>[Stop] <>Reboot		When S3 or S4 model is used.
		[ ]BLADE server	<>[Stop] <>Reboot		When a BLADE server is used.
		[ ]MMB	User name		
			Password		
			User name		For clustering among cases
			Password		
		[ ]vmSP/libvirt	Host OS user name		When the Virtual Machine Function is used in a Xen or KVM environment
			Host OS password		
			Host OS IP address		

Note: Values enclosed in brackets ([ ]) are default values.

 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 Setting Up the Shutdown Facility."

A.8 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.8.1 RMS Setup Worksheet

- A.8.2 GFS Shared File System Setup Worksheet

- A.8.3 GLS Setup Worksheet

- A.9 GDS Setup Worksheets

- A.10 Cluster Application Worksheets

 See

For information on the setup items for this worksheet, see "Chapter 6 Building Cluster Applications."

A.8.1 RMS Setup Worksheet

You can follow the RMS setup worksheet to configure RMS.

 See

For information on the setup items in this worksheet, see "6.1 Initial RMS Setup" and "6.8 Setting Up the RMS Environment."

Item	Environment variable (Global)	Setup value	Remarks
RMS automatic startup	HV_RCSTART	<> No <> [Yes]	(Specified value using CLI) 0 [1]
RMS shutdown wait time	RELIANT_SHUT_MIN_WAIT	(Range: 0 to 2147483647) [900]	

Note: The value in brackets is the default.

[Setup items]

- RMS shutdown wait time (RELIANT_SHUT_MIN_WAIT)

  When using RMS, you need to change the environment variable to the value corresponding to the configuration setup.

  The default value of the environment variable is 900 (seconds) in this version.

  If RMS shutdown request ("hvshut" command) is executed, RMS performs shutdown processing after it performs offline processing of the active cluster applications. Add the maximum time required to complete offline processing of cluster applications and the maximum time required to shut down RMS (30 seconds). Then specify the added total time for this value.

See

For details on RELIANT_SHUT_MIN_WAIT, see "RELIANT_SHUT_MIN_WAIT" in "12.2 Global environment variables" in "*PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide.*"

A.8.2 GFS Shared File System Setup Worksheet

If you plan to use a GFS (Global File Services) Shared File System, use the GFS Shared File System Setup Worksheet to organize the input items beforehand.

See

For information on the setup items in this worksheet, see "6.4 Initial GFS Setup." Also see "A.11.1.2 GFS Shared File System Setup Worksheet" in "A.11 System Design Examples (for the Virtual Machine Function not used)."

		Item	Setting
GFS Shared File System settings		Device name	
		Mount point	
		Device name	
		Mount point	

A.8.3 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** (virtual IP takeover)

- **NIC switching mode** (physical IP takeover I)

- **NIC switching mode** (physical IP takeover II)

- **Fast switching mode**

- **GS linkage mode**

NIC switching mode (Virtual IP Takeover)

Item			Setting
GLS settings	Switching mode		`<Y>`NIC switching mode (virtual IP takeover) `< >` NIC switching mode (physical IP takeover I) `< >` NIC switching mode (physical IP takeover II) `< >` 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)	
		Virtual IP address (or host name)	
	Monitoring destination	IP address (or host name) of primary	
		IP address (or host name) of secondary	
	Optional function	Standby NIC patrol	`<>`Disable `<>`Enable
	Node name (2)		
	Configuration information	Virtual interface name	
		Primary physical interface name	
		Secondary physical interface name	
		Physical IP address (or host name)	
		Virtual IP address (or host name)	
	Monitoring destination	IP address (or host name) of primary	
		IP address (or host name) of secondary	
	Optional function	Standby NIC patrol	`<>`Disable `<>`Enable

[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 (virtual 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 virtual 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 for the takeover virtual IP address by using the "hanetmask create" command.

This value is applied to all physical IP address and virtual IP address.
Note that if the IPv6 address is used and if a subnet is not used for the IPv4 address, you do not need to specify this 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. eth0, eth1, 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(e.g. eth0, etc.). 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).

- Virtual IP address (or host name)

Specify the virtual 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).
At the virtual IP address, specify an IP address with the same network number as the physical IP address.

- 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.

NIC switching mode (physical IP takeover I)

Item			Setting
GLS settings	Switching mode		< >NIC switching mode (virtual IP takeover) <Y> NIC switching mode (physical IP takeover I) < > NIC switching mode (physical IP takeover II) < > 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)	
		Virtual IP address (or host name)	
	Monitoring destination	IP address (or host name) of primary	
		IP address (or host name) of secondary	
	Optional function	Standby NIC patrol	<>Disable <>Enable
	Node name (2)		
	Configuration information	Virtual interface name	
		Primary physical interface name	
		Secondary physical interface name	
		Physical IP address (or host name)	
		Virtual IP address (or host name)	
	Monitoring destination	IP address (or host name) of primary	
		IP address (or host name) of secondary	
	Optional function	Standby NIC patrol	<>Disable <>Enable

[Setup items]

- Switching mode

  Specify a redundant operation type for GLS; Fast switching mode, NIC (Network Interface Card) switching mode, or RIP mode.
  Add e (physical 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 virtual 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 for the takeover virtual IP address by using the "hanetmask create" command.

  This value is applied to all physical IP address and virtual IP address.
  Note that if the IPv6 address is used and if a subnet is not used for the IPv4 address, you do not need to specify this 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. eth0, 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 (e.g. eth0, etc.). 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).

- Virtual IP address (or host name)

  Specify the virtual 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).
  At the virtual IP address, specify an IP address with the same network number as the physical IP address.

- 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.

NIC switching mode (physical IP takeover II)

Item				Setting
GLS settings	Switching mode			< >NIC switching mode (virtual IP takeover) < > NIC switching mode (physical IP takeover I) <Y> NIC switching mode (physical IP takeover II) < > 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	
			Virtual IP address (or host name)	
		Monitoring destination	IP address (or host name) of primary	
			IP address (or host name) of secondary	
		Optional function	Standby NIC patrol	<>Disable <>Enable
	Node name (2)			
		Configuration information	Virtual interface name	
			Primary physical interface name	
			Secondary physical interface name	
			Virtual IP address (or host name)	
		Monitoring destination	IP address (or host name) of primary	
			IP address (or host name) of secondary	
		Optional function	Standby NIC patrol	<>Disable <>Enable

[Setup items]

- Switching mode

  Specify a redundant operation type for GLS; Fast switching mode, NIC (Network Interface Card) switching mode, or RIP mode.
  Add e (physical 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 virtual 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 for the takeover virtual IP address by using the "hanetmask create" command.

  This value is applied to virtual IP address.
  Note that if the IPv6 address is used and if a subnet is not used for the IPv4 address, you do not need to specify this 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. eth0, 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 (e.g. eth0, 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 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.

Fast switching mode

	Item			Setting
GLS settings	Switching mode			< >NIC switching mode (virtual IP takeover) < > NIC switching mode (physical IP takeover I) < > NIC switching mode (physical IP takeover II) <Y> 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.
  At the takeover IP address, specify an IP address with the same network number as the virtual IP address.

- Subnet mask

  Specify the subnet mask that is to be applied to the takeover virtual IP address. Specify this with the "hanetmask create" command.
  Note that if the IPv6 address is used and if a subnet is not used for the IPv4 address, you do not need to specify this 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. eth0, eth1, 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).
  At the virtual IP address, specify the IP address with a network number that is different from the physical IP address.

- Subnet mask

  Specify the value of the subnet mask that is to be applied to the virtual IP address. Specify this with the "hanetmask create" command.
  This value must be identical to the value of the subnet mask that is to be applied to the takeover virtual IP.
  Note that if the IPv6 address is used and if a subnet is not used for the IPv4 address, you do not need to specify this value.

- Physical IP address (1) and (2)

  Specify the physical interface name to be assigned to the physical interface to be duplexed.
  This value must be specified in the /etc/sysconfig/network-scripts/ifcfg-ethX file (ethX indicates a physical interface name).
  At physical IP addresses (1) and (2), specify IP addresses with different network numbers.

- Subnet mask

  Specify the value of the subnet mask that is to be applied to the physical IP address.
  This value must be specified in the /etc/sysconfig/network-scripts/ifcfg-ethX file (ethX indicates a physical interface name).
  Note that if the IPv6 address is used and if a subnet is not used for the IPv4 address, you do not need to specify this value.

GS linkage mode

Item				Setting
GLS settings	Switching mode			< >NIC switching mode (virtual IP takeover) < >NIC switching mode (physical IP takeoverI) < >NIC switching mode (physical IP takeoverII) < >Fast switching mode <Y>GS 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 Linkage mode.

  Add n(Physical interface setting) and c (Virtual interface setting) to the "-m" option of the "habetconfig 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 value of the subnet mask to be applied to takeover virtual IP addresses in the /etc/inet/netmasks file.

  This value is used for a physical IP and virtual 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).

- Physical interface name (1) and (2)

Specify the name of the interface to be first activated among the duplexed physical interfaces (e.g. eth1, eth2, 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)

- 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 "hanetobservcreate" 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.8.4 Setup Worksheet for Takeover Network Resources

Refer to this worksheet when you set up takeover network resources.

See

For information on the setup items in this worksheet, see "6.7.3.6 Setting Up Takeover Network Resources."

Item				Value
Takeover network resources			NeedAll	Yes (no change)
			Timeout	(Range: 45 to 3600) [60]
			InterfaceFilter	(No settings)
			Host name of the target host pingHostPool[0]	
			Host name of the target host pingHostPool[1]	
	Control Target 1	Object	Takeover IP address (Or host name can be specified)	
		Flag	VIRTUAL/BASE	<>[VIRTUAL] <>BASE
			AUTORECOVER	<>[Yes] <>No
			MONITORONLY	<>Yes <>[No]
			PING	<>Yes <>[No] *If "Yes" is specified, select one or more host names specified for pingHostPool.
	Control Target 2	Object	Takeover IP address (Or host name can be specified)	
		Flag	VIRTUAL/BASE	<>[VIRTUAL] <>BASE
			AUTORECOVER	<>[Yes] <>No
			MONITORONLY	<>Yes <>[No]
			PING	<>Yes <>[No] *If "Yes" is specified, select one or more host names specified for pingHostPool.

A.9　GDS Setup Worksheets

You can follow the GDS setup worksheets to configure GDS (Global Disk Services) with this worksheet. The GDS setup worksheets contain two worksheets for mirror setup of the system disk and for shared disk setup.

- 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 "6.3 GDS Configuration Setup."
. . . .

Note
. . . .
The class name should be unique in the cluster system.
. . . .

A.9.1 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 "6.3.1 Setting Up System Disk Mirroring." Also see "A.11.2.4 System Disk Mirror Setup Worksheet" in "A.11 System Design Examples (for the Virtual Machine Function not used)."

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	
	Name of connected guest domain (Only when the Virtual Machine Function is used.)			
System disk mirror settings	Node name			
		Class name		
			Group name	
				Mount point	
				Physical disk name	
				Mirror disk name	
				Spare disk name	
	Name of connected guest domain (Only when the Virtual Machine Function is used.)			

A.9.2 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 "6.3.2 Setting Up Shared Disks." Also see "A.11.1.4 GDS Configuration Worksheet" in "A.11 System Design Examples (for the Virtual Machine Function not used)."

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	
			Single volume 1 — Size	
			Single volume 2 — Volume name	
			Single volume 2 — 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	
			Single volume 1 — Size	
			Single volume 2 — Volume name	
			Single volume 2 — 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.10 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

- Procedure Resource Worksheet

Cluster application configuration:

- Cluster Application Setup Worksheet

Global Settings in Configuration Setup:

- Configuration Global Settings Setup Worksheet

Note

When setting up resources, if the number of objects of any resource exceeds 50, make sure that the numbers of resource objects and detectors do not exceed the maximum number. The maximum numbers of resource objects and detectors are the following:

- Resource objects: 500

- Detectors: 100

You cannot create the configuration of a cluster application which exceeds the values described above.

Calculate the required number of detectors according to the following procedure and make sure that the number of detectors does not exceed the maximum number.

(1) Calculate the required number of detectors for each resource based on the following table.

Resource name	Number of detectors (round up the decimal point)
Cmdline resource	Number of objects / 2 = (a)
Fsystem resource	Number of objects / 5 = (b)
NFS resource (the number of objects with the SHARE flag among the Fsystem resources)	Number of objects / 10 = (c)
Gds resource	Number of objects / 64 = (d)
Gls resource	Number of objects / 64 = (e)
Takeover network resource	Number of objects / 10 = (f)
Procedure resource	Number of objects / 1 = (g)

- Values such as 2, 5, 10, and 64 in the table indicate the number of the objects which is monitored by a detector of each resource.

- For details on objects corresponding to each resource, refer to setting worksheets of each resource.

- The number of objects indicates the total number of all objects of resources which are registered with all cluster applications. The number of detectors is calculated by using the total number of objects.

(2) Add up the number of detectors required for each resource.

```
                 Number of objects  Number of detectors
Cmdline              40             (a) 40 / 2 = 20
Fsystem             200             (b) 200 / 5 = 40
NFS                   0             (c) 0 / 10 = 0
Gds                 200             (d) 200 / 64 = 4 (approximately equal to 3.1)
Gls                  50             (e) 50 / 64 = 1 (approximately equal to 0.7)
Takeover network      0             (f) 0 / 10 = 0
Procedure            10             (g) 10 / 1 = 10
                    500
```

In the above example, the total number of detectors is "(a)20 + (b)40 + (c)0 + (d)4 + (e)1 + (f)0 + (g)10 = 75." This configuration is available because the value "75" does not exceed the maximum number of detectors "100."

A.10.1 Cmdline Resource Worksheet

The Cmdline Resource Setup Worksheet is used for configuring a Cmdline resource.

Item				Setting value	
Cmdline resource	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)

Note: Values enclosed in brackets ([ ]) are default values.

A.10.2 Fsystem Resource Worksheet

The Fsystem Resource Worksheet is used for configuring an Fsystem resource.

Item				Setting value	
Fsystem resource			Timeout	(Range: 45 to 3600) [180]	
	Mount point to be controlled 1	Object	Mount point		
		Flag	AUTORECOVER	<>Yes	<>[No]
			SHARE	<>Yes	<>[No]
	Mount point to be controlled 2	Object	Mount point		
		Flag	AUTORECOVER	<>Yes	<>[No]
			SHARE	<>Yes	<>[No]
	Mount point to be controlled 3	Object	Mount point		
		Flag	AUTORECOVER	<>Yes	<>[No]
			SHARE	<>Yes	<>[No]

Note: Values enclosed in brackets ([ ]) are default values.

A.10.3 Gds Resource Worksheet

The Gds Resource Worksheet is used for configuring a Gds (Global Disk Services) resource.

Item				Setting value	
Gds resource	Resource operation	Attribute	StandbySupport	<>Yes	<>[No]
			AutoRecover	<>Yes	<>[No]
			Timeout	(Range: 5 to 3600) [1800]	
	Disk class to be controlled 1	Object	Disk Class		
	Disk class to be controlled 2	Object	Disk Class		
	Disk class to be controlled 3	Object	Disk Class		

Note: Values enclosed in brackets ([ ]) are default values.

A.10.4 Gls Resource Worksheet

The Gls Resource Worksheet is used for configuring a Gls (Global Link Services) resource.

Item				Setting value	
Gls resource	Resource operation	Attribute	Timeout	(Range: 5 to 300) [60]	
	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: The values enclosed in brackets ([ ]) are the default values.

A.10.5 Procedure Resource Worksheet

The Procedure Resource Worksheet is used for configuring a procedure resource.

Item				Setting value
Procedure resource	Resource name			
	Resource operation	Attribute	SCRIPTTIMEOUT	(Range: 45 to 31535999) [1800]
	Procedure resource to be controlled	Procedure class		<>Application <>BasicApplication <>SystemState2 <>SystemState3 <>( )
			Procedure resource	

Note: Values enclosed in brackets ([ ]) are default values.

A.10.6 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					
	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.
		Exclusivity	Exclusivity setup	LicenseToKill	<>No exclusivity setup	[No]
					<>Cluster application with high job priority	Group 20X
					<>Cluster application with low job priority	Group 10X
						*Specify either A to Z or a to z in Group.
				AutoBreak	<>No exclusivity setup	[Yes]
					<>Cluster application with high job priority	Group 20X
					<>Cluster application with low job priority	Group 10X
						*If you set up a value in LicenseToKill, either Group 20X or Group 10X will be automatically set up in AutoBreak.
	Configuration	Resource				Specified resource name

Note: Values enclosed in brackets ([ ]) are default values.

A.10.7 Configuration Global Settings Worksheet

This worksheet is used for configuring the Global Settings in the 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]
	Detector Details	Monitoring Intervals	hvdet_execbin	(Range: 5 to 2147483647) [10]
			hvdet_ckhost	(Range: 5 to 2147483647) [10]
			hvdet_ddm	(Range: 5 to 2147483647) [19]
			hvdet_glbassrt	(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	(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 need to be changed, please contact Fujitsu systems engineer (SE). However, no changes may be necessary for the normal operation.

A.11 System Design Examples (for the Virtual Machine Function not used)

This section shows design examples of the PRIMECLUSTER system without the virtual machine function. Refer to these examples when designing the PRIMECLUSTER system.

A.11.1 PRIMERGY

This is a design example for PRIMERGY. Refer to the following example when designing the PRIMERGY.

System configuration

Figure A.1 System configuration example
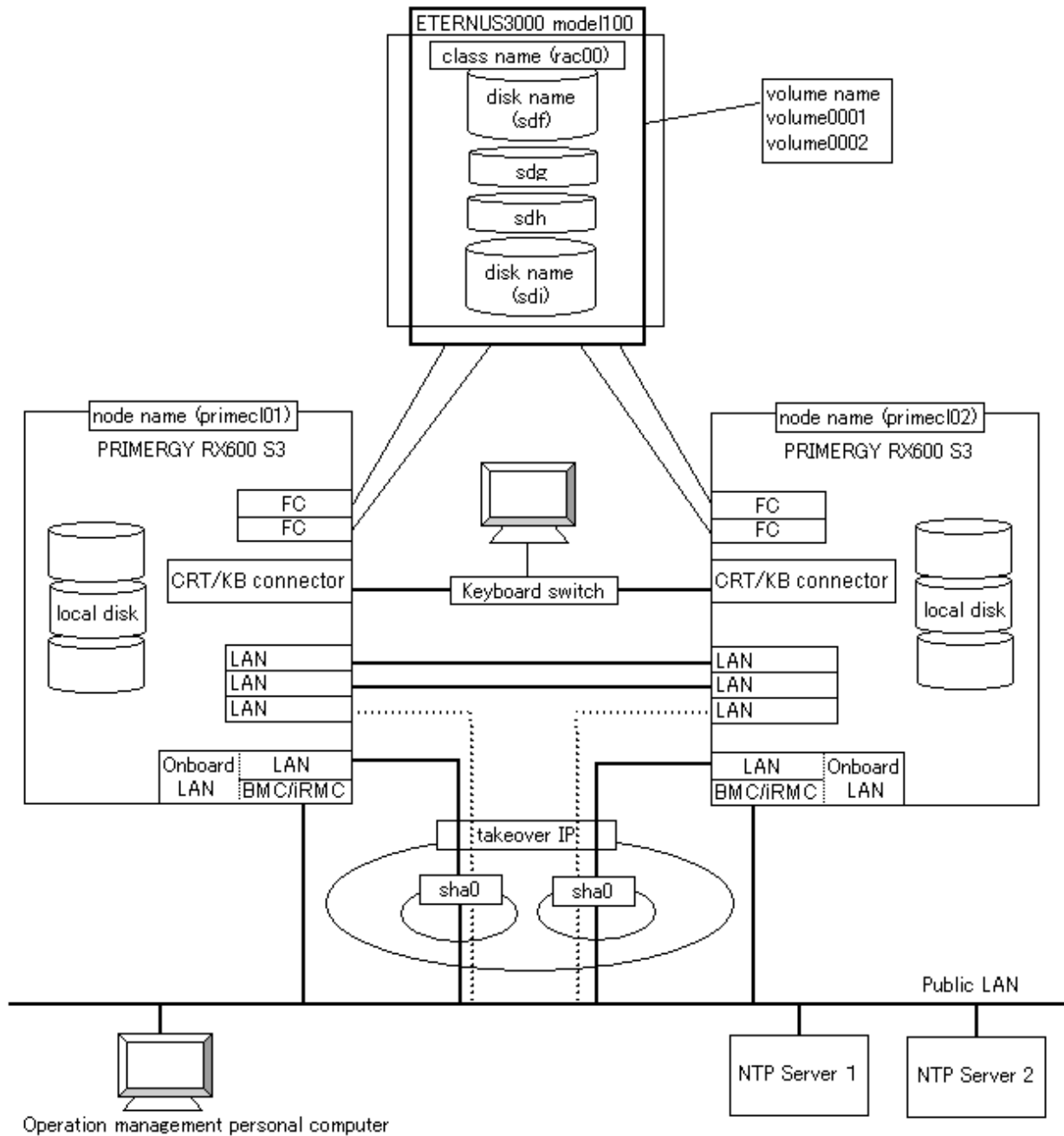
Note: The NTP server should be set to one of the following operation mode;

- 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

- PRIMERGY RX600 S3 x 2

- LAN card x 4

- Fibre channel card x 4

- Disk array unit (ETERNUS3000 model 100) x 1

Used software

- Red Hat Enterprise Linux

- PRIMECLUSTER 4.3

- ETERNUS Multipath driver

- Microsoft Internet Explorer 7 (installed in the operation management personal computer)

- JRE 1.5 (installed in the operation management personal computer)

Shared disk array settings

Device name: ETERNUS3000 model100 (FC)	Number of LU	Number of RAID groups	RAID level	Size
	4	1	5	584GB

A.11.1.1 Cluster Configuration Worksheet

Shown below is an example of creating the Cluster Configuration Worksheet.

Item				Setting		Remarks
Product type				<>PRIMECLUSTER Enterprise Edition <Y>PRIMECLUSTER HA Server <>PRIMECLUSTER Clustering Base <>PRIMECLUSTER Lite Pack		
Shared disk unit: Type 1				ETERNUS model100		
	Interface			Fibre channel		
	Path			Multipath		
	File system selection			GFS Shared File System		
Shared disk unit: Type 2						
	Interface					
	Path					
	File system selection					
NTP	Operation mode			-		
				-		
	Protocol (only when broadcast is specified)			-		
				-		
	Network configuration	External NTP servers	Host name	ntp-server		
			IP address	10.20.30.60		
		NTP server settings in cluster	Primary NTP server	Host name		
				IP address		
			Secondary NTP server	Host name		
				IP address		
Operation management topology	Web-Based Admin View operation mode			2	layers	
	Public LAN			10.20.30.40		
	Primary management server IP			10.20.30.40		
	Secondary management server IP			10.20.30.41		
User group				Group name	User name	
				<Y>wvroot	root	
				<Y>clroot	root	
				<>cladmin		
				<>clmon		
Cluster name				PRIMECLUSTER		

Node information	CIP		Number of subnets		
			Subnet IP	192.168.1.0	
			Net mask	255.255.255.0	
	Node 1		Node name (uname -n)	primecl01	
			CF node name	primecl01	
		Public LAN	Network interface name	eth0, eth3	
			IP address	10.20.30.40,10.20.30.41	
			Net mask	255.255.255.0	
			Gateway		
		Administrative LAN	Network interface name		Specify the information of an administrative LAN. When shared with a public LAN, no entry is necessary.
			IP address		
			Net mask		
		Cluster interconnect	Path 0 NIC name	eth2	
			Path 1 NIC name	eth4	
		weight		1	Node weight (priority for cluster partition) used by SF
		BMC/iRMC	IP address	10.20.30.50	Enter the information of BMC/iRMC installed in S3 or S4 model.
			User ID	bmcadmin	
			Password	bmcpasswd	
		Dump save area for kdump		/dev/sdd1 /var/crash	when kdump is used.
		Memory dump save area for guest domains			When the Virtual Machine Function is used, they are collected in this area.
		Slot No. of server BLADE			When a server BLADE is used.
		Node in CF quorum set		<Y>[Yes] <>No	Do not change the default setting.
	Node 2		Node name (uname -n)	primecl02	
			CF node name	primecl02	
		Public LAN	Network interface name	eth1, eth3	
			IP address	10.20.30.42,10.20.30.43	
			Net mask	255.255.255.0	
			Gateway		
		Administrative LAN	Network interface name		Specify the information of an administrative LAN. When shared with a public LAN, no entry is necessary.
			IP address		
			Net mask		
		Cluster interconnect	Path 0 NIC name	eth2	
			Path 1 NIC name	eth4	
		weight		2	Node weight (priority for cluster partition) used by SF
		BMC/iRMC	IP address	10.20.30.51	Enter the information of BMC/iRMC installed in S3 or S4 model.
			User ID	bmcadmin	
			Password	bmcpasswd	
		Dump save area for kdump		/dev/sdd1 /var/crash	when kdump is used.
		Memory dump save area for guest domains			When the Virtual Machine Function is used, they are collected in this area.
		Slot No. of server BLADE			When a server BLADE is used.
		Node in CF quorum set		<Y>[Yes] <>No	Do not change the default setting.
	Node 3		Node name (uname -n)		
			CF node name		
		Public LAN	Network interface name		
			IP address		
			Net mask		
			Gateway		
		Administrative LAN	Network interface name		Specify the information of an administrative LAN. When shared with a public LAN, no entry is necessary.
			IP address		
			Net mask		
		Cluster interconnect	Path 0 NIC name		
			Path 1 NIC name		
		weight			Node weight (priority for cluster partition) used by SF
		BMC/iRMC	IP address		Enter the information of BMC/iRMC installed in S3 or S4 model.
			User ID		
			Password		
		Dump save area for kdump			when kdump is used.
		Memory dump save area for guest domains			When the Virtual Machine Function is used, they are collected in this area.
		Slot No. of server BLADE			When a server BLADE is used.
		Node in CF quorum set		<>[Yes] <>No	Do not change the default setting.

	Node 4	Node name (uname -n)		
		CF node name		
		Public LAN	Network interface name	
			IP address	
			Net mask	
			Gateway	
		Administrative LAN	Network interface name	Specify the information of an administrative LAN. When shared with a public LAN, no entry is necessary.
			IP address	
			Net mask	
		Cluster interconnect	Path 0 NIC name	
			Path 1 NIC name	
		weight		Node weight (priority for cluster partition) used by SF
		BMC/iRMC	IP address	Enter the information of BMC/iRMC installed in S3 or S4 model.
			User ID	
			Password	
		Dump save area for kdump		when kdump is used.
		Memory dump save area for guest domains		When the Virtual Machine Function is used, they are collected in this area.
		Slot No. of server BLADE		When a server BLADE is used.
		Node in CF quorum set	<>[Yes] <>No	Do not change the default setting.
	CF remote service usage		[Y]cfcp (CF file copy) [Y]cfsh (CF remote command execution)	
	IP interconnect setup		<Y>[Not used] <>Used (IP interconnect count)	
	CIP		<Y>[Used by RMS] <>Not used by RMS Suffix:	
	Management Blade	IP address1		When a server BLADE is used, IP address 2 is used for the clustering among cases.
		IP address2		
		SNMP community name		
Cluster resource management facility	Hardware device to be registered in the resource database		[ ]Network device *1 [ ]Disk unit *2	*1) Required when public LAN takeover is used *2) Required when GDS and GFS are used
Shutdown facility (SF)	Means of stopping a node forcibly		Means	Operation after stop	
			[Y]IPMI	<Y>[Stop] <>Reboot	When S3 or S4 model is used.
			[ ]BLADE server	<>[Stop] <>Reboot	When a BLADE server is used.
			[ ]MMB	User name / Password / User name / Password	For clustering among cases
			[ ]vmSP/ libvirt	Host OS user name / Host OS password / Host OS IP address	When the Virtual Machine Function is used in a Xen or KVM environment
			[ ]icmp	value (timeout)	When the Virtual Machine Function is used in a VMware environment

Note: Values enclosed in brackets ([ ]) are default values.

In each node, specify /etc/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 /var/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
```

A.11.1.2 GFS Shared File System Setup Worksheet

Shown below is an example of creating the GFS Shared File System Setup Worksheet.

Item			Setting
GFS Shared File System		Device name	class0001:volume0001
		Mount point	/disk1
		Device name	class0001:volume0002
		Mount point	/disk2

A.11.1.3 GLS Setup Worksheet

Shown below are examples of creating the GLS Setup Worksheet.

NIC switching mode (Virtual IP Takeover)

Item			Setting
GLS settings	Switching mode		<Y>NIC switching mode (virtual IP takeover) < >NIC switching mode (physical IP takeover I) < >NIC switching mode (physical IP takeover II) < >Fast switching mode
	Takeover virtual interface name		sha0:65
		Takeover virtual IP address (or host name)	10.20.30.46
		Subnet mask	255.255.255.0
	Node name (1)		primecl01
	Configuration information	Virtual interface name	sha0
		Primary physical interface name	eth1
		Secondary physical interface name	eth3
		Physical IP address (or host name)	10.20.30.40
		Virtual IP address (or host name)	10.20.30.46
	Monitoring destination	IP address (or host name) of primary	10.20.30.61
		IP address (or host name) of secondary	10.20.30.62
	Optional function	Standby NIC patrol	<>Disable <Y>Enable
	Node name (2)		primecl02
	Configuration information	Virtual interface name	sha0
		Primary physical interface name	eth1
		Secondary physical interface name	eth3
		Physical IP address (or host name)	10.20.30.41
		Virtual IP address (or host name)	10.20.30.46
	Monitoring destination	IP address (or host name) of primary	10.20.30.61
		IP address (or host name) of secondary	10.20.30.62
	Optional function	Standby NIC patrol	<>Disable <Y>Enable

NIC switching mode (physical IP takeover I)

Item				Setting
GLS settings	Switching mode			< >NIC switching mode (virtual IP takeover) <Y>NIC switching mode (phygical IP takeover I) < >NIC switching mode (phygical IP takeover II) < >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	eth0
			Secondary physical interface name	eth1
			Physical IP address (or host name)	10.34.214.181
			virtual IP address (or host name)	10.34.214.185
		Monitoring destination	IP address (or host name) of primary	10.34.214.188
			IP address (or host name) of secondary	10.34.214.189
		Optional function	Standby NIC patrol	<>Disable <Y>Enable
	Node name (2)			primecl02
		Configuration information	Virtual interface name	sha0
			Primary physical interface name	eth0
			Secondary physical interface name	eth1
			Physical IP address (or host name)	10.34.214.182
			virtual IP address (or host name)	10.34.214.185
		Monitoring destination	IP address (or host name) of primary	10.34.214.188
			IP address (or host name) of secondary	10.34.214.189
		Optional function	Standby NIC patrol	<>Disable <Y>Enable

NIC switching mode (physical IP takeover II)

Item				Setting
GLS settings	Switching mode			< >NIC switching mode (virtual IP takeover) < >NIC switching mode (phygical IP takeover I) <Y>NIC switching mode (phygical IP takeover II) < >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	eth0
			Secondary physical interface name	eth1
			virtual IP address (or host name)	10.34.214.185
		Monitoring destination	IP address (or host name) of primary	10.34.214.188
			IP address (or host name) of secondary	10.34.214.189
		Optional function	Standby NIC patrol	<>Disable <Y>Enable
	Node name (2)			primecl02
		Configuration information	Virtual interface name	sha0
			Primary physical interface name	eth0
			Secondary physical interface name	eth1
			virtual IP address (or host name)	10.34.214.185
		Monitoring destination	IP address (or host name) of primary	10.34.214.188
			IP address (or host name) of secondary	10.34.214.189
		Optional function	Standby NIC patrol	<>Disable <Y>Enable

Fast switching mode

	Item			Setting
GLS settings	Switching mode			< >NIC switching mode (virtual IP takeover)
				< >NIC switching mode (phygical IP takeover I)
				< >NIC switching mode (phygical IP takeover II)
				<Y>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)	eth0
			Physical interface name (2)	eth1
			Virtual IP address (or host name)	10.34.214.181
			Subnet mask	255.255.255.0
			Physical IP address (1)	10.34.215.181
			Subnet mask	255.255.255.0
			Physical IP address (2)	10.34.216.181
			Subnet mask	255.255.255.0
	Node name (2)			primecl01
		Configuration information	Virtual interface name	sha0
			Physical interface name (1)	eth0
			Physical interface name (2)	eth1
			Virtual IP address (or host name)	10.34.214.182
			Subnet mask	255.255.255.0
			Physical IP address (1)	10.34.215.182
			Subnet mask	255.255.255.0
			Physical IP address (2)	10.34.216.182
			Subnet mask	255.255.255.0

A.11.1.4 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	disk0001
			Physical disk name in node 1	sde
			Physical disk name in node 2	sde
			Single volume 1	Volume name	volume0001
				Size	1024 MB
			Single volume 2	Volume name	volume0002
				Size	1024 MB
		Single disk 2	SDX disk name	disk0002
			Physical disk name in node 1	sdf
			Physical disk name in node 2	sdf
		Single disk 3	SDX disk name	disk0003
			Physical disk name in node 1	sdg
			Physical disk name in node 2	sdg
		Single disk 4	SDX disk name	disk0004
			Physical disk name in node 1	sdh
			Physical disk name in node 2	sdh

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 sdj.

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		sdi
		Physical disk name in node 2		sdj
		Single volume 1	Volume name	Volume1
			Size	1024 blocks
		Single volume 2	Volume name	Volume2
			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 sdi, 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	sdi
		Physical disk name in node 2	sdj
	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
			Size	1024 blocks
		Volume 2	Volume name	Volume2
			Size	1024 blocks

Example 3. If disk mirroring is to be executed:

In this example, mirror group Group1, which consists of physical disks sdk and sdl, is created, and the mirror 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	sdi
		Physical disk name in node 2	sdj
	Disk 1 connected to group	SDX disk name	Disk1
		Physical disk name in node 1	sdk
		Physical disk name in node 2	sdl
	Disk 2 connected to group	SDX disk name	Disk2
		Physical disk name in node 1	sdm
		Physical disk name in node 2	sdn
	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 sdi and sdj, 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	sdi
		Physical disk name in node 2	sdj
	Disk 2 connected to group	SDX disk name	Disk2
		Physical disk name in node 1	sdk
		Physical disk name in node 2	sdl
	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 name — Disk/low-order group 2	Disk2
		Volume 1 — Volume name	Volume1
		Volume 1 — Size	1024 blocks
		Volume 2 — Volume name	Volume2
		Volume 2 — Size	1024 blocks

Example 5. If disk concatenation is to be executed:

In this example, concatenation group Group1, which consists of physical disks sdi and sdk, 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	sdi
		Physical disk name in node 2	sdj
	Disk 2 connected to group	SDX disk name	Disk2
		Physical disk name in node 1	sdk
		Physical disk name in node 2	sdl
	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 name — Disk/low-order group 2	Disk2
		Volume 1 — Volume name	Volume1
		Volume 1 — Size	1024 blocks
		Volume 2 — Volume name	Volume2
		Volume 2 — 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 sdg and sdi, and striping group Group3 consists of physical disks sdk and sdm.

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	sde
		Physical disk name in node 2	sdf
	Disk 1 connected to group	SDX disk name	Disk1
		Physical disk name in node 1	sdg
		Physical disk name in node 2	sdh
	Disk 2 connected to group	SDX disk name	Disk2
		Physical disk name in node 1	sdi
		Physical disk name in node 2	sdj
	Disk 3 connected to group	SDX disk name	Disk3
		Physical disk name in node 1	sdk
		Physical disk name in node 2	sdl
	Disk 4 connected to group	SDX disk name	Disk4
		Physical disk name in node 1	sdm
		Physical disk name in node 2	sdn
	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 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 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 2	Group3
		Volume 1	Volume name	Volume1
			Size	1024 blocks
		Volume 2	Volume name	Volume2
			Size	1024 blocks

A.11.2 PRIMEQUEST

This is a design example for PRIMEQUEST. Refer to the following example when designing the PRIMEQUEST.

System configuration

Figure A.2 System configuration example
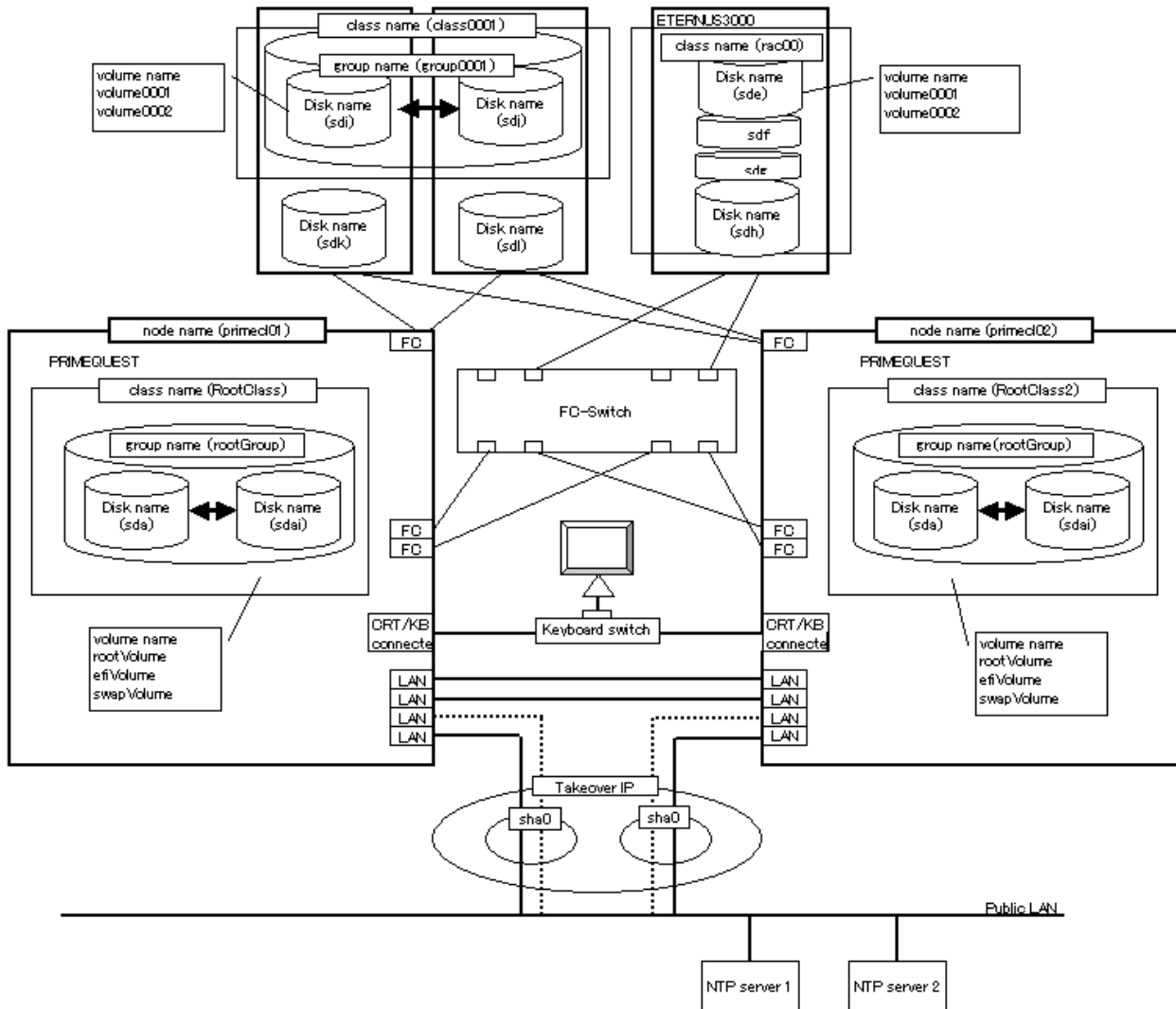
Note: The NTP server should be set to one of the following operation mode;

- 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

- PRIMEQUEST x 2

- FC-Switch x 1 unit

- LAN card x 8

- Fibre channel card x 4

- Disk array unit x 2

Used software

- Red Hat Enterprise Linux AS

- PRIMECLUSTER 4.3

- Multipath driver

- Microsoft Internet Explorer 7 (installed in the operation management personal computer)

- Java(TM) 2 Platform Standard Edition Runtime Environment Version 6 (installed in the operation management personal computer)

System disk settings

Node name	Physical disk name	Slice number	Slice name	Size (MB)
primecl01	sda	1	/	15374
		2	/boot/efi	256
		3	swap	2048
primecl02	sda	1	/	15374
		2	/boot/efi	256
		3	swap	2048

Note

When GDS is used to mirror the system disk, it uses one slice number and at least 32 MB of unused area (or swap area) on the system disk. If system disk mirroring is to be executed, consider the resources used by GDS when designing the system.

Shared disk array settings

Device name: ETERNUS3000	Number of LU	Number of RAID groups	RAID level	Size
	4	1	5	248.2GB

A.11.2.1 Cluster Configuration Worksheet

Shown below is an example of creating the Cluster Configuration Worksheet.

Item				Setting		Remarks
Product type				<>PRIMECLUSTER Enterprise Edition		
				<>PRIMECLUSTER HA Server		
Shared disk unit: Type 1				ETERNUS 3000		
	Interface			Fibre channel		
	Path			Multipath		
	File system selection			GFS Shared File System		
Shared disk unit: Type 2				Expansion file unit		
	Interface			Fibre channel		
	Path			single		
	File system selection			GFS Shared File System		
NTP	Operation mode					
	Protocol (only when broadcast is specified)					
	Network configuration	External NTP servers	Host name			
			IP address			
		NTP server settings in cluster	Primary NTP server	Host name	primecl01	
				IP address	10.20.30.40	
			Secondary NTP server	Host name	primecl02	
				IP address	10.20.30.41	
Operation management topology	Web-Based Admin View operation mode			2 layers		
	Public LAN					
	Primary management server IP			10.20.30.42		
	Secondary management server IP			10.20.30.43		
User group				Group name	User name	
				<Y>wvroot	root	
				<Y>clroot	root	
				<>cladmin		
				<>clmon		
Cluster name				PRIMECLUSTER		
Node information	CIP	Number of subnets				
		Subnet IP		192.168.1.0		
		Net mask		255.255.255.0		
	Node 1	Node name (uname -n)		prime01		
		CF node name		prime01		
		Public LAN	Network interface name	eth0, eth3		
			IP address	10.20.30.40,10.20.30.44		
			Net mask	255.255.255.0		
			Gateway			
		Administrative LAN	Network interface name			
			IP address			
			Net mask			
		Cluster interconnect	Path 0 NIC name	eth1, eth2		
			Path 1 NIC name			
		weight		1		Node weight (priority for cluster partition) used by SF
		Memory dump save area for guest domains		/var/lib/xen/dump on the host OS		When the Virtual Machine Function is used, they are collected in this area.
		Node in CF quorum set		<Y>[Yes] <>No		Do not change the default setting.

	Node 2	Node name (uname -n)		prime02	
		CF node name		prime02	
		Public LAN	Network interface name	eth0, eth3	
			IP address	10.20.30.41,10.20.30.45	
			Net mask	255.255.255.0	
			Gateway		
		Administrative LAN	Network interface name		
			IP address		
			Net mask		
		Cluster interconnect	Path 0 NIC name	eth1, eth2	
			Path 1 NIC name		
		weight		2	Node weight (priority for cluster partition) used by SF
		Memory dump save area for guest domains		/var/lib/xen/dump on the host OS	When the Virtual Machine Function is used, they are collected in this area.
		Node in CF quorum set		<Y>[Yes] <>No	Do not change the default setting.
	Node 3	Node name (uname -n)			
		CF node name			
		Public LAN	Network interface name		
			IP address		
			Net mask		
			Gateway		
		Administrative LAN	Network interface name		
			IP address		
			Net mask		
		Cluster interconnect	Path 0 NIC name		
			Path 1 NIC name		
		weight			Node weight (priority for cluster partition) used by SF
		Memory dump save area for guest domains			When the Virtual Machine Function is used, they are collected in this area.
		Node in CF quorum set		<>[Yes] <>No	Do not change the default setting.
	Node 4	Node name (uname -n)			
		CF node name			
		Public LAN	Network interface name		
			IP address		
			Net mask		
			Gateway		
		Administrative LAN	Network interface name		
			IP address		
			Net mask		
		Cluster interconnect	Path 0 NIC name		
			Path 1 NIC name		
		weight			Node weight (priority for cluster partition) used by SF
		Memory dump save area for guest domains			When the Virtual Machine Function is used, they are collected in this area.
		Slot No. of server BLADE			When a server BLADE is used.
		Node in CF quorum set		<>[Yes] <>No	Do not change the default setting.
	CF remote service usage			[Y]cfcp (CF file copy)	
				[Y]cfsh (CF remote command execution)	
	IP interconnect setup			<Y>[Not used]	
				<>Used (IP interconnect count)	
	CIP			<Y>[Used by RMS]	
				<>Not used by RMS	
				Suffix:	
Cluster resource management facility	Hardware device to be registered in the resource database			[ ]Network device *1	*1) Required when public LAN takeover is used
				[ ]Disk unit *2	*2) Required when GDS and GFS are used
Shutdown facility (SF)	Means of stopping a node forcibly			Means	Operation after stop
			[Y]MMB	User name	mmbrmcpaclmin
				Password	mmbrmcppasswd
				User name	
				Password	For clustering among cases
			[ ]vmSP/ libvirt	Host OS user name	
				Host OS password	When the Virtual Machine Function is used in a Xen or KVM environment
				Host OS IP address	
			[ ]icmp	value (timeout)	When the Virtual Machine Function is used in a VMware environment

Note: Values enclosed in brackets ([ ]) are default values.

In each node, specify /etc/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 /var/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
```

A.11.2.2 GFS Shared File System Setup Worksheet

Shown below is an example of creating the GFS Shared File System Setup Worksheet.

Item		Setting
GFS Shared File System	Device name	class0001:volume0001
	Mount point	/disk1
	Device name	class0001:volume0002
	Mount point	/disk2

A.11.2.3 GLS Setup Worksheet

Shown below are examples of creating the GLS Setup Worksheet.

NIC switching mode (virtual IP Takeover)

Item			Setting
GLS settings	Switching mode		<Y>NIC switching mode (virtual IP takeover) < >NIC switching mode (phygical IP takeover I) < >NIC switching mode (phygical IP takeover II) < >Fast switching mode < >GS linkage 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	eth0
			Secondary physical interface name	eth1
			Physical IP address (or host name)	10.34.214.181
			Virtual IP address (or host name)	10.34.214.185
		Monitoring destination	IP address (or host name) of primary	10.34.214.188
			IP address (or host name) of secondary	10.34.214.189
		Optional function	Standby NIC patrol	<>Disable <Y>Enable
	Node name (2)		primecl02
		Configuration information	Virtual interface name	sha0
			Primary physical interface name	eth0
			Secondary physical interface name	eth1
			Physical IP address (or host name)	10.34.214.182
			Virtual IP address (or host name)	10.34.214.185
		Monitoring destination	IP address (or host name) of primary	10.34.214.188
			IP address (or host name) of secondary	10.34.214.189
		Optional function	Standby NIC patrol	<>Disable <Y>Enable

NIC switching mode (physical IP takeover I)

Item				Setting
GLS settings	Switching mode			< >NIC switching mode (virtual IP takeover) <Y>NIC switching mode (phygical IP takeover I) < >NIC switching mode (phygical IP takeover II) < >Fast switching mode < >GS linkage 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	eth0
			Secondary physical interface name	eth1
			Physical IP address (or host name)	10.34.214.181
			Virtual IP address (or host name)	10.34.214.185
		Monitoring destination	IP address (or host name) of primary	10.34.214.188
			IP address (or host name) of secondary	10.34.214.189
		Optional function	Standby NIC patrol	<>Disable <Y>Enable
	Node name (2)			primecl02
		Configuration information	Virtual interface name	sha0
			Primary physical interface name	eth0
			Secondary physical interface name	eth1
			Physical IP address (or host name)	10.34.214.182
			Virtual IP address (or host name)	10.34.214.185
		Monitoring destination	IP address (or host name) of primary	10.34.214.188
			IP address (or host name) of secondary	10.34.214.189
		Optional function	Standby NIC patrol	<>Disable <Y>Enable

NIC switching mode (physical IP takeover II)

Item				Setting
GLS settings	Switching mode			< >NIC switching mode (virtual IP takeover) < >NIC switching mode (phygical IP takeover I) <Y>NIC switching mode (phygical IP takeover II) < >Fast switching mode < >GS linkage 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	eth0
			Secondary physical interface name	eth1
			Virtual IP address (or host name)	10.34.214.185
		Monitoring destination	IP address (or host name) of primary	10.34.214.188
			IP address (or host name) of secondary	10.34.214.189
		Optional function	Standby NIC patrol	<>Disable <Y>Enable
	Node name (2)			primecl02
		Configuration information	Virtual interface name	sha0
			Primary physical interface name	eth0
			Secondary physical interface name	eth1
			Virtual IP address (or host name)	10.34.214.185
		Monitoring destination	IP address (or host name) of primary	10.34.214.188
			IP address (or host name) of secondary	10.34.214.189
		Optional function	Standby NIC patrol	<>Disable <Y>Enable

Fast switching mode

Item				Setting
GLS settings	Switching mode			< >NIC switching mode (virtual IP takeover) < >NIC switching mode (phygical IP takeover I) < >NIC switching mode (phygical IP takeover II) <Y>Fast switching mode < >GS linkage 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)	eth0
			Physical interface name (2)	eth1
			Virtual IP address (or host name)	10.34.214.181
			Subnet mask	255.255.255.0
			Physical IP address (1)	10.34.215.181
			Subnet mask	255.255.255.0
			Physical IP address (2)	10.34.216.181
			Subnet mask	255.255.255.0
	Node name (2)			primecl01
		Configuration information	Virtual interface name	sha0
			Physical interface name (1)	eth0
			Physical interface name (2)	eth1
			Virtual IP address (or host name)	10.34.214.182
			Subnet mask	255.255.255.0
			Physical IP address (1)	10.34.215.182
			Subnet mask	255.255.255.0
			Physical IP address (2)	10.34.216.182
			Subnet mask	255.255.255.0

GS linkage mode

Item				Setting
GLS settings	Switching mode			< >NIC switching mode (virtual IP takeover) < >NIC switching mode (physical IP takeoverI) < >NIC switching mode (physical IP takeoverII) < >Fast switching mode <Y>GS linkage 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)	eth0
			Physical interface name (2)	eth1
			IP address (or host name)(1)	10.34.214.181
			Subnet mask	255.255.255.0
			IP address (or host name)(2)	10.34.214.182
			Subnet mask	255.255.255.0
	Node name (2)			primecl02
		Configuration information	Virtual interface name	sha0
			Physical interface name (1)	eth0
			Physical interface name (2)	eth1
			IP address (or host name)(1)	10.34.214.183
			Subnet mask	255.255.255.0
			IP address (or host name)(2)	10.34.214.184
			Subnet mask	255.255.255.0
	Remote host name			GS-1
		Configuration information	Virtual IP address	10.34.215.182
			IP address (1)	10.34.214.185
			IP address (2)	10.34.214.186
		Monitoring destination information	Remote virtual IP address monitoring	<Y>Enable <>Disable
		Optional function	RIP packet is sent from remote host	<Y>Wait <>Not Wait
			TCP relay	Network information of relay destinationNetwork IP address or host name ( )

A.11.2.4 System Disk Mirror Setup Worksheet

Shown below is an example of creating the System Disk Mirror Setup Worksheet.

Item					Setting
System disk mirror settings	Node name				primecl01
		Class name			RootClass
			Group name		rootGroup
				Mount point	/, /boot/efi, swap
				Physical disk name	sda
				Mirror disk name	sdai
				Spare disk name	-
	Node name				primecl02
		Class name			RootClass2
			Group name		rootGroup
				Mount point	/, /boot/efi, swap
				Physical disk name	sda
				Mirror disk name	sdai
				Spare disk name	-

A.11.2.5 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		disk0001
			Physical disk name in node 1		sde
			Physical disk name in node 2		sde
			Single volume 1	Volume name	volume0001
				Size	1024 MB
			Single volume 2	Volume name	volume0002
				Size	1024 MB
		Single disk 2	SDX disk name		disk0002
			Physical disk name in node 1		sdf
			Physical disk name in node 2		sdf
		Single disk 3	SDX disk name		disk0003
			Physical disk name in node 1		sdg
			Physical disk name in node 2		sdg
		Single disk 4	SDX disk name		disk0004
			Physical disk name in node 1		sdh
			Physical disk name in node 2		sdh

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 sdj.

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	sdi
		Physical disk name in node 2	sdj
		Single volume 1	Volume name	Volume1
			Size	1024 blocks
		Single volume 2	Volume name	Volume2
			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 sdi, 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	sdi
		Physical disk name in node 2	sdj
	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
			Size	1024 blocks
		Volume 2	Volume name	Volume2
			Size	1024 blocks

Example 3. If disk mirroring is to be executed:

In this example, mirror group Group1, which consists of physical disks sdk and sdl, is created, and the mirror 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	sdi
		Physical disk name in node 2	sdj
	Disk 1 connected to group	SDX disk name	Disk1
		Physical disk name in node 1	sdk
		Physical disk name in node 2	sdl
	Disk 2 connected to group	SDX disk name	Disk2
		Physical disk name in node 1	sdm
		Physical disk name in node 2	sdn
	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 name — Disk/low-order group 2	Disk2
		Volume 1 — Volume name	Volume1
		Volume 1 — Size	1024 blocks
		Volume 2 — Volume name	Volume2
		Volume 2 — Size	1024 blocks

Example 4. If disk striping is to be executed:

In this example, striping group Group1, which consists of physical disks sdi and sdj, 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	sdi
		Physical disk name in node 2	sdj
	Disk 2 connected to group	SDX disk name	Disk2
		Physical disk name in node 1	sdk
		Physical disk name in node 2	sdl
	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 sdi and sdk, 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	sdi
		Physical disk name in node 2	sdj
	Disk 2 connected to group	SDX disk name	Disk2
		Physical disk name in node 1	sdk
		Physical disk name in node 2	sdl
	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 sdg and sdi, and striping group Group3 consists of physical disks sdk and sdm.

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	sde
		Physical disk name in node 2	sdf
	Disk 1 connected to group	SDX disk name	Disk1
		Physical disk name in node 1	sdg
		Physical disk name in node 2	sdh
	Disk 2 connected to group	SDX disk name	Disk2
		Physical disk name in node 1	sdi
		Physical disk name in node 2	sdj
	Disk 3 connected to group	SDX disk name	Disk3
		Physical disk name in node 1	sdk
		Physical disk name in node 2	sdl
	Disk 4 connected to group	SDX disk name	Disk4
		Physical disk name in node 1	sdm
		Physical disk name in node 2	sdn
	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 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 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 2	Group3
		Volume 1	Volume name	Volume1
			Size	1024 blocks
		Volume 2	Volume name	Volume2
			Size	1024 blocks

A.12 System Design Examples (for the Virtual Machine Function used in Xen Environment)

In the following design example, the virtual machine function is used in a Xen environment on the PRIMECLUSTER system. Refer to this example when you design the system.

See

. . . .

When using the virtual machine function in a VMware environment, see "Appendix H Using PRIMECLUSTER in a VMware Environment."

. . . .

A.12.1 When building a cluster system between guest OSes on one host OS

System configuration

![Note icon] **Note**
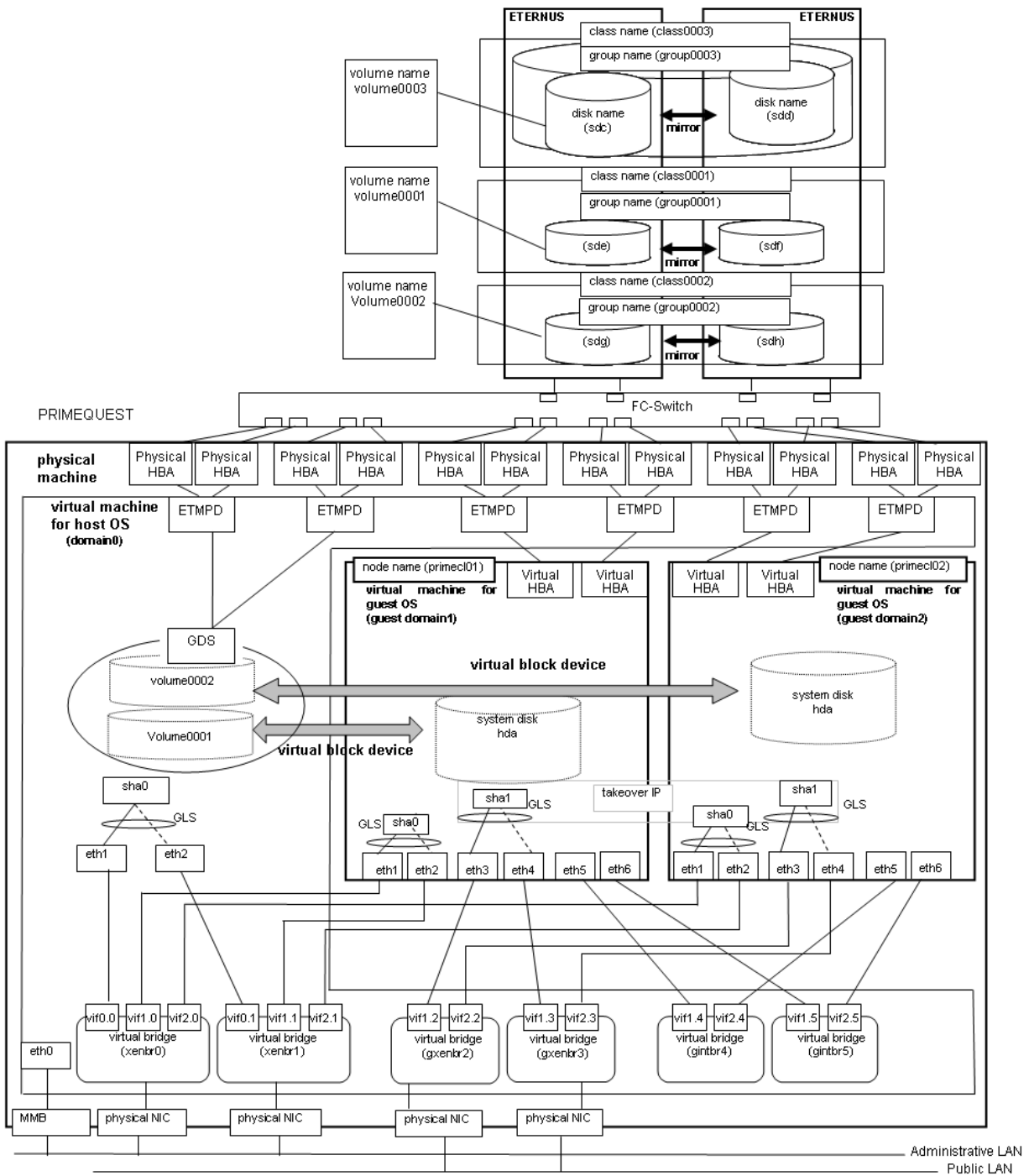
- Set the host OS as an NTP server and the guest OS as an NTP client, and then perform time synchronization.

- If you want to use the system as a test environment, you can configure it so that the redundant administrative LAN is created on the guest OS by GLS. For information on how to create a redundant administrative LAN, see "PRIMEQUEST Virtual Machine Function User's Manual".

- Assign more than two virtual CPUs to each guest OS which comprises the cluster system. For information on how to assign virtual CPUs to guest OSes, see "PRIMEQUEST Virtual Machine Function User's Manual".

Used hardware

- PRIMEQUEST 1800E x 1

- FC-Switch x 2

- LAN card x 4

- Fibre channel card x 8

- Disk array unit x 2

Used software

- Red Hat Enterprise Linux

- PRIMECLUSTER 4.3

- Multipath Driver

- Microsoft Internet Explorer 7 (installed on the operation management personal computer)

- Java(TM) 2 Platform Standard Edition Runtime Environment Version 6 (installed on the operation management personal computer)

System disk settings

Node name	Virtual disk name	Slice number	Slice name	Size (MB)
primecl01	hda	1	/	15374
		2	/boot/efi	256
		3	swap	2048
primecl02	hda	1	/	15374
		2	/boot/efi	256
		3	swap	2048

Note

If you want to mirror the system disk by using GDS, see the section about redundancy in "PRIMEQUEST Virtual Machine Function User's Manual".

Shared disk array settings

Device name: ETERNUS	Number of LUN	Number of RAID groups	RAID level	Size
	3	3	5	248.2GB

Note

The local class disks or shared class disks used by GDS on the guest OS should be created as virtual SCSI devices if they are used in the virtual machine environment.

A.12.1.1 Cluster Configuration Worksheet

Shown below is an example of the Cluster Configuration Worksheet.

Item					Setting			Remarks
Product type					<>PRIMECLUSTER Enterprise Edition			
					<>PRIMECLUSTER HA Server			
					<>PRIMECLUSTER Clustering Base			
					<>PRIMECLUSTER Lite Pack			
Shared disk unit: Type 1					ETERNUS3000			
	Interface				Fibre channel			
	Path				mutilpath			
	File system selection				GFS Shared File System			
Shared disk unit: Type 2					Expansion file unit			
	Interface				Fibre channel			
	Path				single			
	File system selection				GFS Shared File System			
NTP	Operation mode							
	Protocol (only when broadcast is specified)							
	NTP Network configuration	External NTP servers	Host name		host01			
			IP address		10.20.40.50			
		NTP server settings in cluster	Primary NTP server	Host name				
				IP address				
			Secondary NTP server	Host name				
				IP address				
Operation management topology	Web-Based Admin View operation mode				2		layers	
	Public LAN							
	Primary management server IP				10.20.40.40			
	Secondary management server IP				10.20.40.41			
User group					Group name		User name	
					<Y>wvroot		root	
					<Y>clroot		root	
					<>cladmin			
					<>clmon			
Cluster name					PRIMECLUSTER			
Node information	CIP	Number of subnets						
		Subnet IP			192.168.1.0			
		Net mask			255.255.255.0			
	Node 1	Node name (uname -n)			primecl01			
		CF node name			primecl01			
		Guest domain name			domain0			Virtual machine function used
		Public LAN	Network interface name		eth3,eth4			
			IP address		10.20.30.40, 10.20.30.44			
			Net mask		255.255.255.0			
			Gateway		-			
		Administrative LAN	Network interface name		eth1,eth2			
			IP address		10.20.40.40, 10.20.40.44			
			Net mask		255.255.255.0			
			Gateway		-			
		Cluster interconnect	Path 0 NIC		eth5			
			Path 1 NIC		eth6			
		Weight			1			Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area						When kdump is used and virtual machine function is not used
		Dump device for sadump and dump save area						When sadump is used and virtual machine function is not used
		Memory dump save area for guest domain			/var/lib/xen/dump on the Host OS			When virtual machine function is used, memory dump save this area.
		Node in CF quorum set			<Y>[Yes]　<>No			Do not change the default setting.

	Node 2	Node name (uname -n)	primecl02	
		CF node name	primecl02	
		Guest domain name	domain1	Virtual machine function used
		Public LAN	Network interface name	eth3,eth4	
			IP address	10.20.30.41, 10.20.30.45	
			Net mask	255.255.255.0	
			Gateway	-	
		Administrative LAN	Network interface name	eth1,eth2	
			IP address	10.20.40.41, 10.20.40.45	
			Net mask	255.255.255.0	
			Gateway	-	
		Cluster interconnect	Path 0 NIC	eth5	
			Path 1 NIC	eth6	
		Weight	2	Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area		When kdump is used and virtual machine function is not used
		Dump device for sadump and dump save area		When sadump is used and virtual machine function is not used
		Memory dump save area for guest domain	/var/lib/xen/dump on the Host OS	When virtual machine function is used, memory dump save this area.
		Node in CF quorum set	<Y>[Yes] <>No	Do not change the default setting.
	CF remote service usage		[ ]cfcp (CF file copy) [ ]cfsh (CF remote command execution)	
	IP interconnect setup		<Y>[Not used] <>Used (1P interconnect count)	
	CIP		< >[Used by RMS] < >Not used by RMS Suffix:	
Cluster resource management facility	Hardware device to be registered in the resource database		< > Network device *1 < > Disk units *2	*1) Required when public LAN takeover is used *2) Required when GDS and GFS are used
Shutdown facility (SF)	Means of stopping a node forcibly		Means	

		< >MMB	User name		
			Password		
			User name		Cluster between units used
			Password		
		<Y>vmSP	Host OS user name	FJSVvmSP (fixed)	
			Host OS user password		Virtual machine function used in a Xen environment
			Host OS IP address	10.20.40.50	

A.12.1.2 GLS Setup Worksheet

Shown below are examples of the GLS Setup Worksheet.

NIC switching mode (virtual IP Takeover)

	Item		Setting
GLS settings	Switching mode		<Y>NIC switching mode (virtual IP takeover) < >NIC switching mode (physical IP takeover I) < >NIC switching mode (physical IP takeover II) < >Fast switching mode
	Takeover virtual interface name		sha0:65
		Takeover virtual IP address (or host name)	10.20.30.46
		Subnet mask	255.255.255.0
	Node name (1)		primecl01
	Configuration information	Virtual interface name	sha0
		Primary physical interface name	eth1
		Secondary physical interface name	eth3
		Physical IP address (or host name)	10.20.30.40
		Virtual IP address (or host name)	10.20.30.46
	Monitoring destination	IP address (or host name) of primary	10.20.30.61
		IP address (or host name) of secondary	10.20.30.62
	Optional function	Standby NIC patrol	<>Disable <Y>Enable
	Node name (2)		primecl02
	Configuration information	Virtual interface name	sha0
		Primary physical interface name	eth1
		Secondary physical interface name	eth3
		Physical IP address (or host name)	10.20.30.41
		Virtual IP address (or host name)	10.20.30.46
	Monitoring destination	IP address (or host name) of primary	10.20.30.61
		IP address (or host name) of secondary	10.20.30.62
	Optional function	Standby NIC patrol	<>Disable <Y>Enable

A.12.1.3 GDS Configuration Worksheet

Shown below are examples of the GDS Configuration Worksheet.

Example. If disk mirroring is to be executed:

In this example, mirror group group0003, which consists of physical disks sdc and sdd, is created, and the mirror volume volume0003 is created in group0003.

Item				Setting
Class 1	Class name			class0003
	Class scope (node name)	Node 1		primecl01
		Node 2		primecl02
	Disk 1 connected to group	SDX disk name		Disk1
		Physical disk name in node 1		sdk
		Physical disk name in node 2		sdl
	Disk 2 connected to group	SDX disk name		Disk2
		Physical disk name in node 1		sdm
		Physical disk name in node 2		sdn
	Highest-order group 1	Group name		group0003
		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	volume0003
			Size	1024 blocks

A.12.2 When building a cluster system between guest OSes on multiple host OSes

System configuration

Inside of PRIMEQUEST

physical machine #2

physical machine #1

virtual machine for the Host OS (Domain 0)

Physical HBA Physical HBA Physical HBA Physical HBA Physical HBA Physical HBA Physical HBA Physical HBA

ETMPD ETMPD ETMPD ETMPD

PV SCSI PV SCSI PV SCSI PV SCSI

node name (primecl01)
virtual machine for the guest OS (Guest Domain 1)

PV SCSI PV SCSI

PV SCSI PV SCSI

node name (primecl11)
virtual machine for the guest OS (Guest Domain 2)

GDS

GDS

GDS

GDS

volume0002

volume0001

Shared Disk volume0005

Shared Disk volume0006

System Disk hda

System Disk hda

sha0

GLS

sha1

GLS

takeover IP

GLS sha0

sha0 GLS

sha1

GLS

takeover IP

eth1 eth2

eth1 eth2 eth3 eth4 eth5 eth6

eth1 eth2 eth3 eth4 eth5 eth6

eth0

vif0.1 vif1.1 vif2.1
virtual bridge (xenbr0)

vif0.2 vif1.2 vif2.2
virtual bridge (xenbr1)

vif1.3 vif2.3
virtual bridge (gextbr2)

vif1.4 vif2.4
virtual bridge (gextbr3)

vif1.5 vif2.5
virtual bridge (gextbr4)

vif1.6 vif2.6
virtual bridge (gextbr5)

MMB physical NIC physical NIC physical NIC physical NIC physical NIC physical NIC

Network configuration example

Administrative LAN

Public LAN

Cluster interconnect

Unit 1

peth1 peth2 peth3 peth4 eth5 eth6

eth1 eth2 eth3 eth4 VLAN 10 VLAN 11 VLAN 20 VLAN 21

xenbr1 xenbr2 xenbr3 xenbr4 gextbr5 gextbr51 VLAN 1 gextbr6 gextbr61 VLAN 2

eth1 eth2 eth3 eth4

Node 1 (Cluster 1)

Node 1 (Cluster 2)

Host OS

eth0

MMB

Unit 2

peth1 peth2 peth3 peth4 eth5 eth6

eth1 eth2 eth3 eth4 VLAN 10 VLAN 11 VLAN 20 VLAN 21

xenbr1 xenbr2 xenbr3 xenbr4 gextbr5 gextbr51 VLAN 1 gextbr6 gextbr61 VLAN 2

eth1 eth2 eth3 eth4

Node 2 (Cluster 1)

Node 2 (Cluster 2)

Host OS

eth0

MMB

Administrative LAN

FC-SW FC-SW

Eternus

> **Note**
>
> - Set the host OS as an NTP server and the guest OS as an NTP client, and then perform time synchronization.
>
> - If you plan to operate guest OSes on multiple host OSes as clusters, use NTP for time synchronization on host OSes as well.
>
> - If you want to use the system as a test environment, you can configure it so that the redundant administrative LAN is created on the guest OS by GLS.
>
> - Assign more than two virtual CPUs to each guest OS which comprises the cluster system.

Used hardware

- PRIMEQUEST 1800E x 2

- FC-Switch x 2

- LAN card x 14

- Fibre channel card x 16

- Disk array unit x 2

Used software

- Red Hat Enterprise Linux

- PRIMECLUSTER 4.3

- Multipath Driver

- Microsoft Internet Explorer 8 (installed on the operation management personal computer)

- J2SE (TM) Runtime Environment Version 6 (installed on the operation management personal computer)

System disk settings

Host OS	Node name	Virtual disk name	Slice number	Slice name	Size (MB)
host01	primecl01	hda	1	/	15374
			2	/boot/efi	256
			3	swap	2048
	primecl11	hda	1	/	15374
			2	/boot/efi	256
			3	swap	2048
host02	primecl02	hda	1	/	15374
			2	/boot/efi	256
			3	swap	2048
	primecl12	hda	1	/	15374
			2	/boot/efi	256
			3	swap	2048

Shared disk array settings

Device name: ETERNUS	Number of LUN	Number of RAID groups	RAID level	Size
	6	6	5	248.2GB

A.12.2.1 Cluster Configuration Worksheet

Shown below is an example of the Cluster Configuration Worksheet.

When the cluster name is PRIMECLUSTER1

Item				Setting		Remarks
Product type				<>PRIMECLUSTER Enterprise Edition		
				<>PRIMECLUSTER HA Server		
Shared disk unit: Type 1				ETERNUS3000		
	Interface			Fibre channel		
	Path			mutilpath		
	File system selection			ext3		
Shared disk unit: Type 2				Expansion file unit		
	Interface			Fibre channel		
	Path			mutilpath		
	File system selection			ext3		
NTP	Operation mode					
	Protocol (only when broadcast is specified)					
	NTP Network configuration	External NTP servers	Host name	host01,host02		
			IP address	10.20.40.50,10.20.40.51		
		NTP server settings in cluster	Primary NTP server	Host name		
				IP address		
			Secondary NTP server	Host name		
				IP address		
Operation management topology	Web-Based Admin View operation mode			2 layers		
	Public LAN					
	Primary management server IP			10.20.40.40		
	Secondary management server IP			10.20.40.41		
User group				Group name	User name	
				<Y>wvroot	root	
				<Y>clroot	root	
				<>cladmin		
				<>clmon		
Cluster name				PRIMECLUSTER1		
Node information	CIP	Number of subnets				
		Subnet IP		192.168.1.0		
		Net mask		255.255.255.0		
	Node 1	Node name (uname -n)		primecl01		
		CF node name		primecl01		
		Guest domain name		domain01		Virtual machine function used
		Public LAN	Network interface name	eth3,eth4		
			IP address	10.20.30.40,10.20.30.44		
			Net mask	255.255.255.0		
			Gateway	-		
		Administrative LAN	Network interface name	eth1,eth2		
			IP address	10.20.40.40,10.20.40.44		
			Net mask	255.255.255.0		
			Gateway	-		
		Cluster interconnect	Path 0 NIC	eth5		
			Path 1 NIC	eth6		
		Weight		1		Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area				When kdump is used and virtual machine function is not used
		Dump device for sadump and dump save area				When sadump is used and virtual machine function is not used
		Memory dump save area for guest domain		/var/lib/xen/dump on the Host OS		When virtual machine function is used, memory dump save this area.
		Node in CF quorum set		<Y>[Yes] <>No		Do not change the default setting.

	Node 2	Node name (uname -n)		primecl02	
		CF node name		primecl02	
		Guest domain name		domain02	Virtual machine function used
		Public LAN	Network interface name	eth3,eth4	
			IP address	10.20.30.41,10.20.30.45	
			Net mask	255.255.255.0	
			Gateway	-	
		Administrative LAN	Network interface name	eth1,eth2	
			IP address	10.20.40.41,10.20.40.45	
			Net mask	255.255.255.0	
			Gateway	-	
		Cluster interconnect	Path 0 NIC	eth5	
			Path 1 NIC	eth6	
		Weight		2	Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area			When kdump is used and virtual machine function is not used
		Dump device for sadump and dump save area			When sadump is used and virtual machine function is not used
		Memory dump save area for guest domain		/var/lib/xen/dump on the Host OS	When virtual machine function is used, memory dump save this area.
		Node in CF quorum set		<Y>[Yes] <>No	Do not change the default setting.
	CF remote service usage			[ ]cfcp (CF file copy) [ ]cfsh (CF remote command execution)	
	IP interconnect setup			<Y>[Not used] <>Used (1P interconnect count)	
	CIP			< >[Used by RMS] < >[Not used by RMS Suffix:	
Cluster resource management facility	Hardware device to be registered in the resource database			< > Network device *1 < > Disk units *2	*1) Required when public LAN takeover is used *2) Required when GDS and GFS are used
Shutdown facility (SF)	Means of stopping a node forcibly		Means		
			< >MMB	User name	
				Password	
				User name	
				Password	Cluster between units used
			<Y>FJSVvmSP	Host OS user name	FJSVvmSP (fixed)
				Host OS user password		Virtual machine function used in a Xen environment
				Host OS IP address	10.20.40.50
				Host OS user name	FJSVvmSP (fixed)	Cluster between units used
				Host OS user password	
				Host OS IP address	10.20.40.51

When the cluster name is PRIMECLUSTER2

Item				Setting		Remarks
Product type				<>PRIMECLUSTER Enterprise Edition		
				<>PRIMECLUSTER HA Server		
Shared disk unit: Type 1				ETERNUS3000		
	Interface			Fibre channel		
	Path			mutilpath		
	File system selection			ext3		
Shared disk unit: Type 2				Expansion file unit		
	Interface			Fibre channel		
	Path			mutilpath		
	File system selection			ext3		
NTP	Operation mode					
	Protocol (only when broadcast is specified)					
	NTP Network configuration	External NTP servers	Host name	host01,host02		
			IP address	10.20.40.50,10.20.40.51		
		NTP server settings in cluster	Primary NTP server	Host name		
				IP address		
			Secondary NTP server	Host name		
				IP address		
Operation management topology	Web-Based Admin View operation mode			2	layers	
	Public LAN					
	Primary management server IP			10.20.40.140		
	Secondary management server IP			10.20.40.141		
User group				Group name	User name	
				<Y>wvroot	root	
				<Y>clroot	root	
				<>cladmin		
				<>clmon		
Cluster name				PRIMECLUSTER2		
Node information	CIP	Number of subnets				
		Subnet IP		192.168.1.0		
		Net mask		255.255.255.0		
	Node 1	Node name (uname -n)		primecl11		
		CF node name		primecl11		
		Guest domain name		domain11		Virtual machine function used
		Public LAN	Network interface name	eth3,eth4		
			IP address	10.20.30.140,10.20.30.144		
			Net mask	255.255.255.0		
			Gateway	-		
		Administrative LAN	Network interface name	eth1,eth2		
			IP address	10.20.40.140,10.20.40.144		
			Net mask	255.255.255.0		
			Gateway	-		
		Cluster interconnect	Path 0 NIC	eth5		
			Path 1 NIC	eth6		
		Weight		1		Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area				When kdump is used and virtual machine function is not used
		Dump device for sadump and dump save area				When sadump is used and virtual machine function is not used
		Memory dump save area for guest domain		/var/lib/xen/dump on the Host OS		When virtual machine function is used, memory dump save this area.
		Node in CF quorum set		<Y>[Yes] <>No		Do not change the default setting.

	Node 2	Node name (uname -n)		primecl12	
		CF node name		primecl12	
		Guest domain name		domain12	Virtual machine function used
		Public LAN	Network interface name	eth3,eth4	
			IP address	10.20.30.141,10.20.30.145	
			Net mask	255.255.255.0	
			Gateway	-	
		Administrative LAN	Network interface name	eth1,eth2	
			IP address	10.20.40.141,10.20.40.145	
			Net mask	255.255.255.0	
			Gateway	-	
		Cluster interconnect	Path 0 NIC	eth5	
			Path 1 NIC	eth6	
		Weight		2	Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area			When kdump is used and virtual machine function is not used
		Dump device for sadump and dump save area			When sadump is used and virtual machine function is not used
		Memory dump save area for guest domain		/var/lib/xen/dump on the Host OS	When virtual machine function is used, memory dump save this area.
		Node in CF quorum set		<Y>[Yes] <>No	Do not change the default setting.
	CF remote service usage			[ ]cfcp (CF file copy) [ ]cfsh (CF remote command execution)	
	IP interconnect setup			<Y>[Not used] <>Used (1P interconnect count)	
	CIP			< >[Used by RMS] < >[Not used by RMS] Suffix:	
Cluster resource management facility	Hardware device to be registered in the resource database			< > Network device *1 < > Disk units *2	*1) Required when public LAN takeover is used *2) Required when GDS and GFS are used
Shutdown facility (SF)	Means of stopping a node forcibly	Means			
		< >MMB	User name		
			Password		
			User name		Cluster between units used
			Password		
		<Y>FJSVvmSP	Host OS user name	FJSVvmSP (fixed)	
			Host OS user password		Virtual machine function used in a Xen environment
			Host OS IP address	10.20.40.50	
			Host OS user name	FJSVvmSP (fixed)	Cluster between units used
			Host OS user password		
			Host OS IP address	10.20.40.51	

A.12.2.2 GLS Setup Worksheet

Shown below are examples of the GLS Setup Worksheet.

When the cluster name is PRIMECLUSTER1

Item				Setting
GLS settings	Switching mode			\<Y\>NIC switching mode (virtual IP takeover) < >NIC switching mode (physical IP takeover I) < >NIC switching mode (physical IP takeover II) < >Fast switching mode
	Takeover virtual interface name			sha1:65
		Takeover virtual IP address (or host name)		10.20.30.46
		Subnet mask		255.255.255.0
	Node name (1)			primecl01
		Configuration information	Virtual interface name	sha1
			Primary physical interface name	eth3
			Secondary physical interface name	eth4
			Physical IP address (or host name)	10.20.30.40
			Logical IP address (or host name)	10.20.30.46
		Monitoring destination	IP address (or host name) of primary	10.20.30.61
			IP address (or host name) of secondary	10.20.30.62
		Optional function	Standby NIC patrol	<>Disable \<Y\>Enable
	Node name (2)			primecl02
		Configuration information	Virtual interface name	sha1
			Primary physical interface name	eth3
			Secondary physical interface name	eth4
			Physical IP address (or host name)	10.20.30.41
			Logical IP address (or host name)	10.20.30.46
		Monitoring destination	IP address (or host name) of primary	10.20.30.61
			IP address (or host name) of secondary	10.20.30.62
		Optional function	Standby NIC patrol	<>Disable \<Y\>Enable

When the cluster name is PRIMECLUSTER2

Item				Setting
GLS settings	Switching mode			\<Y\>NIC switching mode (virtual IP takeover) < >NIC switching mode (physical IP takeover I) < >NIC switching mode (physical IP takeover II) < >Fast switching mode
	Takeover virtual interface name			sha1:65
		Takeover virtual IP address (or host name)		10.20.30.146
		Subnet mask		255.255.255.0
	Node name (1)			primecl11
		Configuration information	Virtual interface name	sha1
			Primary physical interface name	eth3
			Secondary physical interface name	eth4
			Physical IP address (or host name)	10.20.30.140
			Logical IP address (or host name)	10.20.30.146
		Monitoring destination	IP address (or host name) of primary	10.20.30.161
			IP address (or host name) of secondary	10.20.30.162
		Optional function	Standby NIC patrol	<>Disable \<Y\>Enable
	Node name (2)			primecl12
		Configuration information	Virtual interface name	sha1
			Primary physical interface name	eth3
			Secondary physical interface name	eth4
			Physical IP address (or host name)	10.20.30.141
			Logical IP address (or host name)	10.20.30.146
		Monitoring destination	IP address (or host name) of primary	10.20.30.161
			IP address (or host name) of secondary	10.20.30.162
		Optional function	Standby NIC patrol	<>Disable \<Y\>Enable

A.12.2.3 GDS Configuration Worksheet

Shown below are examples of the GDS Configuration Worksheet.

Item			Setting
Class 1	Class name		class0005
	Class scope (node name)	Node 1	primecl01
		Node 2	primecl02
	Disk 1 connected to group	SDX disk name	Disk1
		Physical disk name in node 1	sdc
		Physical disk name in node 2	sdc
	Disk 2 connected to group	SDX disk name	Disk2
		Physical disk name in node 1	sdd
		Physical disk name in node 2	sdd
	Highest-order group 1	Group name	group0005
		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	volume0005
			Size	1024 blocks
Class 2	Class name		class0006
	Class scope (node name)	Node 1	primecl11
		Node 2	primecl12
	Disk 1 connected to group	SDX disk name	Disk3
		Physical disk name in node 1	sde
		Physical disk name in node 2	sde
	Disk 2 connected to group	SDX disk name	Disk4
		Physical disk name in node 1	sdf
		Physical disk name in node 2	sdf
	Highest-order group 1	Group name	group0006
		Group type	mirror
		Stripe width	-
		Disk/low-order group name	Disk/low-order group 1	Disk3
			Disk/low-order group 2	Disk4
		Volume 1	Volume name	volume0006
			Size	1024 blocks

A.13 System Design Examples (for the Virtual Machine Function used in KVM Environment)

In the following design example, the virtual machine function is used in a KVM environment on the PRIMECLUSTER system. Refer to this example when you design the system.

 See

. . . .

When using the virtual machine function in a VMware environment, see "Appendix H Using PRIMECLUSTER in a VMware Environment."

. . . .

A.13.1 When building a cluster system between guest OSes on one host OS

System configuration

![Note icon] **Note**

. . . .

- Set the hypervisor as an NTP server and the guest OS as an NTP client, and then perform time synchronization.

- If you want to use the system as a test environment, you can configure it so that the redundant administrative LAN is created on the guest OS by GLS. For information on how to create a redundant administrative LAN, see " Configuration and Administration Guide: Redundant Line Control Function".

. . . .

Used hardware

- PRIMEQUEST 1800E x 1

- FC-Switch x 2

- LAN card x 4

- Fibre channel card x 8

- Disk array unit x 2

Used software

- Red Hat Enterprise Linux

- PRIMECLUSTER 4.3

- Multipath Driver

- Microsoft Internet Explorer 7 (installed on the operation management personal computer)

- Java(TM) 2 Platform Standard Edition Runtime Environment Version 6 (installed on the operation management personal computer)

System disk settings

Node name	Virtual disk name	Slice number	Slice name	Size (MB)
primecl01	hda	1	/	15374
		2	/boot/efi	256
		3	swap	2048
primecl02	hda	1	/	15374
		2	/boot/efi	256
		3	swap	2048

Note

If you want to mirror the system disk by using GDS, see the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*".

Shared disk array settings

Device name: ETERNUS	Number of LUN	Number of RAID groups	RAID level	Size
	3	3	5	248.2GB

Note

The local class disks or shared class disks used by GDS on the guest OS should be created as virtual disks (virtio block devices) if they are used in the virtual machine environment.

A.13.1.1 Cluster Configuration Worksheet

Shown below is an example of the Cluster Configuration Worksheet.

Item					Setting		Remarks
Product type					<>PRIMECLUSTER Enterprise Edition		
					<>PRIMECLUSTER HA Server		
					<>PRIMECLUSTER Clustering Base		
					<>PRIMECLUSTER Lite Pack		
Shared disk unit: Type 1					ETERNUS3000		
	Interface				Fibre channel		
	Path				mutilpath		
	File system selection				GFS Shared File System		
Shared disk unit: Type 2					Expansion file unit		
	Interface				Fibre channel		
	Path				single		
	File system selection				GFS Shared File System		
NTP	Operation mode						
	Protocol (only when broadcast is specified)						
	NTP Network configuration	External NTP servers	Host name		host01		
			IP address		10.20.40.50		
		NTP server settings in cluster	Primary NTP server	Host name			
				IP address			
			Secondary NTP server	Host name			
				IP address			
Operation management topology	Web-Based Admin View operation mode				2	layers	
	Public LAN						
	Primary management server IP				10.20.40.40		
	Secondary management server IP				10.20.40.41		
User group					Group name	User name	
					<Y>wvroot	root	
					<Y>clroot	root	
					<>cladmin		
					<>clmon		
Cluster name					PRIMECLUSTER		
Node information	CIP	Number of subnets					
		Subnet IP			192.168.1.0		
		Net mask			255.255.255.0		
	Node 1	Node name (uname -n)			primecl01		
		CF node name			primecl01		
		Guest domain name			domain0		Virtual machine function used
		Public LAN	Network interface name		eth3,eth4		
			IP address		10.20.30.40, 10.20.30.44		
			Net mask		255.255.255.0		
			Gateway		-		
		Administrative LAN	Network interface name		eth1,eth2		
			IP address		10.20.40.40, 10.20.40.44		
			Net mask		255.255.255.0		
			Gateway		-		
		Cluster interconnect	Path 0 NIC		eth5		
			Path 1 NIC		eth6		
		Weight			1		Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area					When kdump is used and virtual machine function is not used
		Dump device for sadump and dump save area					When sadump is used and virtual machine function is not used
		Memory dump save area for guest domain			/var/crash on the hypervisor		When virtual machine function is used, memory dump save this area.
		Node in CF quorum set			<Y>[Yes] <>No		Do not change the default setting.

	Node 2	Node name (uname -n)		primecl02	
		CF node name		primecl02	
		Guest domain name		domain1	Virtual machine function used
		Public LAN	Network interface name	eth3,eth4	
			IP address	10.20.30.41, 10.20.30.45	
			Net mask	255.255.255.0	
			Gateway	-	
		Administrative LAN	Network interface name	eth1,eth2	
			IP address	10.20.40.41, 10.20.40.45	
			Net mask	255.255.255.0	
			Gateway	-	
		Cluster interconnect	Path 0 NIC	eth5	
			Path 1 NIC	eth6	
		Weight		2	Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area			When kdump is used and virtual machine function is not used
		Dump device for sadump and dump save area			When sadump is used and virtual machine function is not used
		Memory dump save area for guest domain		/var/crash on the hypervisor	When virtual machine function is used, memory dump save this area.
		Node in CF quorum set		<Y>[Yes] <>No	Do not change the default setting.
	CF remote service usage			[ ]cfcp (CF file copy) [ ]cfsh (CF remote command execution)	
	IP interconnect setup			<Y>[Not used] <>Used (1P interconnect count)	
	CIP			< >[Used by RMS] < >Not used by RMS Suffix:	
Cluster resource management facility	Hardware device to be registered in the resource database			< > Network device *1 < > Disk units *2	*1) Required when public LAN takeover is used *2) Required when GDS and GFS are used
Shutdown facility (SF)	Means of stopping a node forcibly		Means		
			< >MMB	User name	
				Password	
				User name	Cluster between units used
				Password	
			<Y>libvirt	Host OS user name	user1	
				Host OS user password		Virtual machine function used in a KVM environment
				Host OS IP address	10.20.40.50	

A.13.1.2 GLS Setup Worksheet

Shown below are examples of the GLS Setup Worksheet.

NIC switching mode (virtual IP Takeover)

Item			Setting
GLS settings	Switching mode		\<Y>NIC switching mode (virtual IP takeover) \< >NIC switching mode (physical IP takeover I) \< >NIC switching mode (physical IP takeover II) \< >Fast switching mode
	Takeover virtual interface name		sha0:65
		Takeover virtual IP address (or host name)	10.20.30.46
		Subnet mask	255.255.255.0
	Node name (1)		primecl01
	Configuration information	Virtual interface name	sha0
		Primary physical interface name	eth1
		Secondary physical interface name	eth3
		Physical IP address (or host name)	10.20.30.40
		Virtual IP address (or host name)	10.20.30.46
	Monitoring destination	IP address (or host name) of primary	10.20.30.61
		IP address (or host name) of secondary	10.20.30.62
	Optional function	Standby NIC patrol	<>Disable \<Y>Enable
	Node name (2)		primecl02
	Configuration information	Virtual interface name	sha0
		Primary physical interface name	eth1
		Secondary physical interface name	eth3
		Physical IP address (or host name)	10.20.30.41
		Virtual IP address (or host name)	10.20.30.46
	Monitoring destination	IP address (or host name) of primary	10.20.30.61
		IP address (or host name) of secondary	10.20.30.62
	Optional function	Standby NIC patrol	<>Disable \<Y>Enable

A.13.1.3 GDS Configuration Worksheet

Shown below are examples of the GDS Configuration Worksheet.

Example. If disk mirroring is to be executed:

In this example, mirror group group0003, which consists of physical disks sdc and sdd, is created, and the mirror volume volume0003 is created in group0003.

Item			Setting
Class 1	Class name		class0003
	Class scope (node name)	Node 1	primecl01
		Node 2	primecl02
	Disk 1 connected to group	SDX disk name	Disk1
		Physical disk name in node 1	sdk
		Physical disk name in node 2	sdl
	Disk 2 connected to group	SDX disk name	Disk2
		Physical disk name in node 1	sdm
		Physical disk name in node 2	sdn
	Highest-order group 1	Group name	group0003
		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	volume0003
			Size	1024 blocks

A.13.2 When building a cluster system between guest OSes on multiple host OSes

System configuration
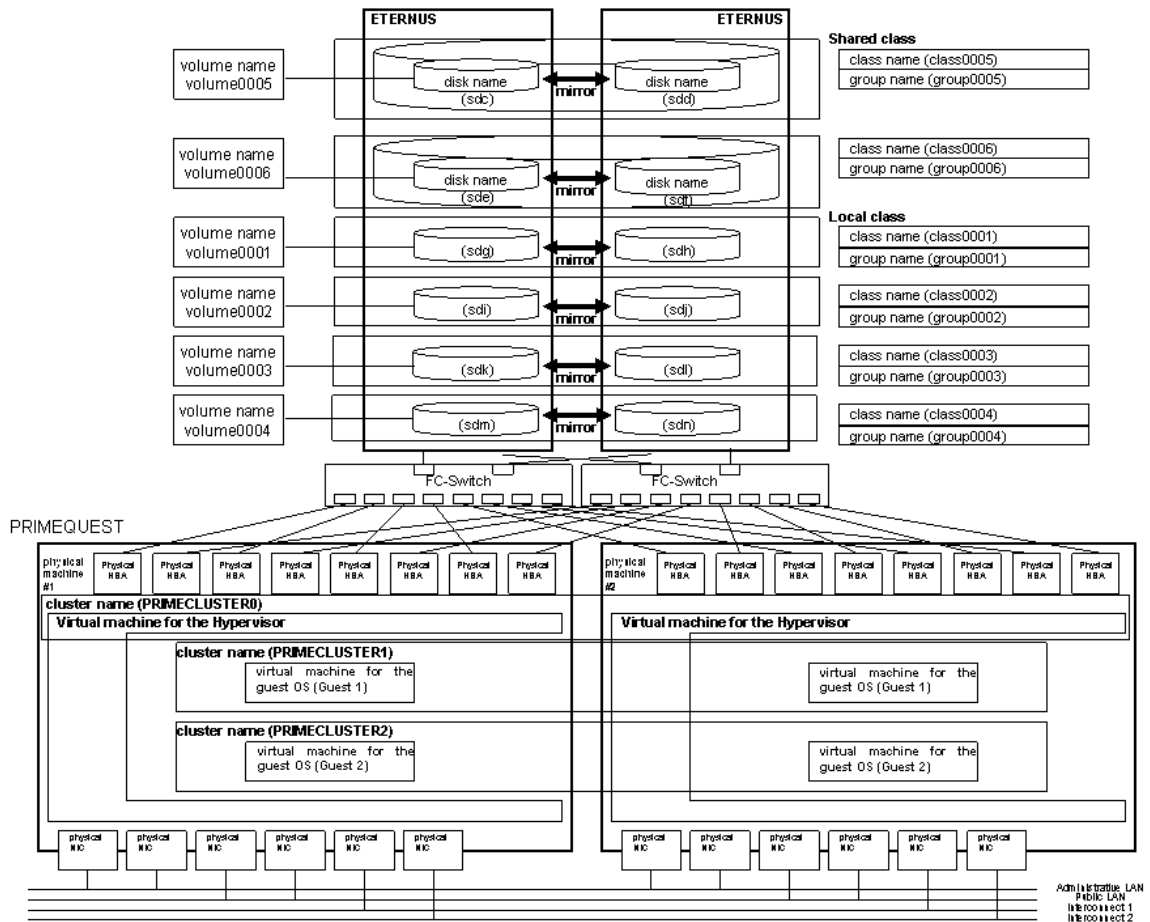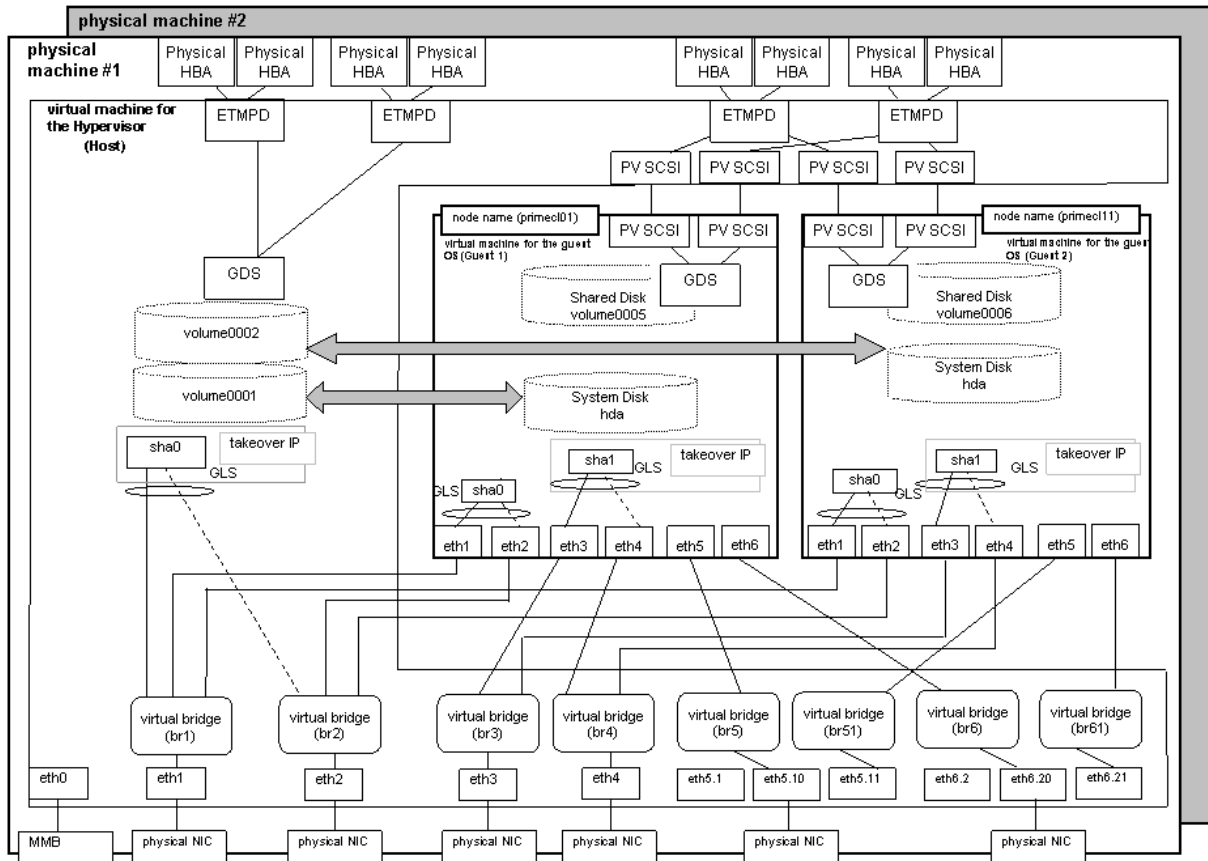
Network configuration example
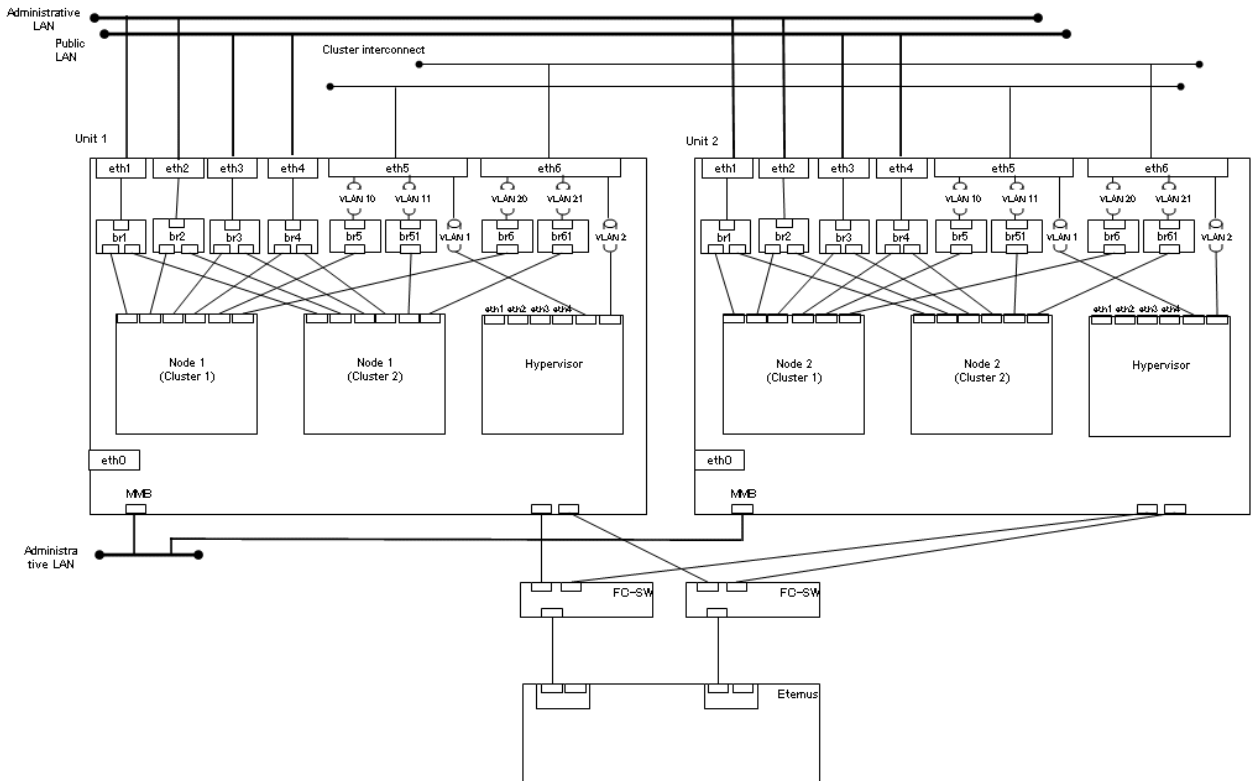
Note

- Set the hypervisor as an NTP server and the guest OS as an NTP client, and then perform time synchronization.

- If you plan to operate guest OSes on multiple host OSes as clusters, use NTP for time synchronization on the hypervisor as well.

- If you want to use the system as a test environment, you can configure it so that the redundant administrative LAN is created on the guest OS by GLS. For information on how to create a redundant administrative LAN, see " Configuration and Administration Guide: Redundant Line Control Function".

Used hardware

- PRIMEQUEST 1800E x 2

- FC-Switch x 2

- LAN card x 14

- Fibre channel card x 16

- Disk array unit x 2

Used software

- Red Hat Enterprise Linux

- PRIMECLUSTER 4.3

- Multipath Driver

- Microsoft Internet Explorer 8 (installed on the operation management personal computer)

- J2SE (TM) Runtime Environment Version 6 (installed on the operation management personal computer)

System disk settings

Host OS	Node name	Virtual disk name	Slice number	Slice name	Size (MB)
host01	primecl01	hda	1	/	15374
			2	/boot/efi	256
			3	swap	2048
	primecl11	hda	1	/	15374
			2	/boot/efi	256
			3	swap	2048
host02	primecl02	hda	1	/	15374
			2	/boot/efi	256
			3	swap	2048
	primecl12	hda	1	/	15374
			2	/boot/efi	256
			3	swap	2048

Note

If you want to mirror the system disk by using GDS, see the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide*".

Shared disk array settings

Device name: ETERNUS	Number of LUN	Number of RAID groups	RAID level	Size
	6	6	5	248.2GB

Note

The local class disks or shared class disks used by GDS on the guest OS should be created as virtual disks (virtio block devices) if they are used in the virtual machine environment.

A.13.2.1 Cluster Configuration Worksheet

Shown below is an example of the Cluster Configuration Worksheet.

When the cluster name is PRIMECLUSTER1

Item				Setting		Remarks
Product type				<>PRIMECLUSTER Enterprise Edition		
				<>PRIMECLUSTER HA Server		
Shared disk unit: Type 1				ETERNUS3000		
	Interface			Fibre channel		
	Path			mutilpath		
	File system selection			ext3		
Shared disk unit: Type 2				Expansion file unit		
	Interface			Fibre channel		
	Path			mutilpath		
	File system selection			ext3		
NTP	Operation mode					
	Protocol (only when broadcast is specified)					
	NTP Network configuration	External NTP servers	Host name	host01,host02		
			IP address	10.20.40.50,10.20.40.51		
		NTP server settings in cluster	Primary NTP server	Host name		
				IP address		
			Secondary NTP server	Host name		
				IP address		
Operation management topology	Web-Based Admin View operation mode			2	layers	
	Public LAN					
	Primary management server IP			10.20.40.40		
	Secondary management server IP			10.20.40.41		
User group				Group name	User name	
				<Y>wvroot	root	
				<Y>clroot	root	
				<>cladmin		
				<>clmon		
Cluster name				PRIMECLUSTER1		
Node information	CIP	Number of subnets				
		Subnet IP		192.168.1.0		
		Net mask		255.255.255.0		
	Node 1	Node name (uname -n)		primecl01		
		CF node name		primecl01		
		Guest domain name		domain01		Virtual machine function used
		Public LAN	Network interface name	eth3,eth4		
			IP address	10.20.30.40,10.20.30.44		
			Net mask	255.255.255.0		
			Gateway	-		
		Administrative LAN	Network interface name	eth1,eth2		
			IP address	10.20.40.40,10.20.40.44		
			Net mask	255.255.255.0		
			Gateway	-		
		Cluster interconnect	Path 0 NIC	eth5		
			Path 1 NIC	eth6		
		Weight		1		Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area				When kdump is used and virtual machine function is not used
		Dump device for sadump and dump save area				When sadump is used and virtual machine function is not used
		Memory dump save area for guest domain		/var/crash on the hypervisor		When virtual machine function is used, memory dump save this area.
		Node in CF quorum set		<Y>[Yes] <>No		Do not change the default setting.

	Node 2	Node name (uname -n)		primecl02		
		CF node name		primecl02		
		Guest domain name		domain02		Virtual machine function used
		Public LAN	Network interface name	eth3,eth4		
			IP address	10.20.30.41,10.20.30.45		
			Net mask	255.255.255.0		
			Gateway	-		
		Administrative LAN	Network interface name	eth1,eth2		
			IP address	10.20.40.41,10.20.40.45		
			Net mask	255.255.255.0		
			Gateway	-		
		Cluster interconnect	Path 0 NIC	eth5		
			Path 1 NIC	eth6		
		Weight		2		Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area				When kdump is used and virtual machine function is not used
		Dump device for sadump and dump save area				When sadump is used and virtual machine function is not used
		Memory dump save area for guest domain		/var/crash on the hypervisor		When virtual machine function is used, memory dump save this area.
		Node in CF quorum set		<Y>[Yes] <>No		Do not change the default setting.
	CF remote service usage			[ ]cfcp (CF file copy) [ ]cfsh (CF remote command execution)		
	IP interconnect setup			<Y>[Not used] <>Used (1P interconnect count)		
	CIP			< >[Used by RMS] < >[Not used by RMS Suffix:		
Cluster resource management facility	Hardware device to be registered in the resource database			< > Network device *1 < > Disk units *2		*1) Required when public LAN takeover is used *2) Required when GDS and GFS are used
Shutdown facility (SF)	Means of stopping a node forcibly			Means		
				< >MMB	User name	
					Password	
					User name	Cluster between units used
					Password	
				<Y>libvirt	Host OS user name	user1	Virtual machine function used in a KVM environment
					Host OS user password	
					Host OS IP address	10.20.40.50
					Host OS user name	user1	Cluster between units used
					Host OS user password	
					Host OS IP address	10.20.40.51

When the cluster name is PRIMECLUSTER2

Item				Setting			Remarks
Product type				<>PRIMECLUSTER Enterprise Edition			
				<>PRIMECLUSTER HA Server			
Shared disk unit: Type 1				ETERNUS3000			
	Interface			Fibre channel			
	Path			mutilpath			
	File system selection			ext3			
Shared disk unit: Type 2				Expansion file unit			
	Interface			Fibre channel			
	Path			mutilpath			
	File system selection			ext3			
NTP	Operation mode						
	Protocol (only when broadcast is specified)						
	NTP Network configuration	External NTP servers	Host name	host01,host02			
			IP address	10.20.40.50,10.20.40.51			
		NTP server settings in cluster	Primary NTP server	Host name			
				IP address			
			Secondary NTP server	Host name			
				IP address			
Operation management topology	Web-Based Admin View operation mode			2		layers	
	Public LAN						
	Primary management server IP			10.20.40.140			
	Secondary management server IP			10.20.40.141			
User group				Group name	User name		
				<Y>wvroot	root		
				<Y>clroot	root		
				<>cladmin			
				<>clmon			
Cluster name				PRIMECLUSTER2			
Node information	CIP	Number of subnets					
		Subnet IP		192.168.1.0			
		Net mask		255.255.255.0			
	Node 1	Node name (uname -n)		primecl11			
		CF node name		primecl11			
		Guest domain name		domain11			Virtual machine function used
		Public LAN	Network interface name	eth3,eth4			
			IP address	10.20.30.140,10.20.30.144			
			Net mask	255.255.255.0			
			Gateway	-			
		Administrative LAN	Network interface name	eth1,eth2			
			IP address	10.20.40.140,10.20.40.144			
			Net mask	255.255.255.0			
			Gateway	-			
		Cluster interconnect	Path 0 NIC	eth5			
			Path 1 NIC	eth6			
		Weight		1			Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area					When kdump is used and virtual machine function is not used
		Dump device for sadump and dump save area					When sadump is used and virtual machine function is not used
		Memory dump save area for guest domain		/var/crash on the hypervisor			When virtual machine function is used, memory dump save this area.
		Node in CF quorum set		<Y>[Yes] <>No			Do not change the default setting.

Node 2	Node name (uname -n)		primecl12	
	CF node name		primecl12	
	Guest domain name		domain12	Virtual machine function used
	Public LAN	Network interface name	eth3,eth4	
		IP address	10.20.30.141,10.20.30.145	
		Net mask	255.255.255.0	
		Gateway	-	
	Administrative LAN	Network interface name	eth1,eth2	
		IP address	10.20.40.141,10.20.40.145	
		Net mask	255.255.255.0	
		Gateway	-	
	Cluster interconnect	Path 0 NIC	eth5	
		Path 1 NIC	eth6	
	Weight		2	Node weight (priority if cluster partition occurs) used by SF
	Dump device for kdump and dump save area			When kdump is used and virtual machine function is not used
	Dump device for sadump and dump save area			When sadump is used and virtual machine function is not used
	Memory dump save area for guest domain		/var/crash on the hypervisor	When virtual machine function is used, memory dump save this area.
	Node in CF quorum set		<Y>[Yes] <>No	Do not change the default setting.
CF remote service usage			[ ]cfcp (CF file copy) [ ]cfsh (CF remote command execution)	
IP interconnect setup			<Y>[Not used] <>Used　　　　　(1P interconnect count)	
CIP			< >[Used by RMS] < >[Not used by RMS] Suffix:	
Cluster resource management facility	Hardware device to be registered in the resource database		< > Network device *1	*1) Required when public LAN takeover is used
			< > Disk units *2	*2) Required when GDS and GFS are used
Shutdown facility (SF)	Means of stopping a node forcibly	Means		
		< >MMB	User name		
			Password		
			User name		
			Password		Cluster between units used
		<Y>libvirt	Host OS user name	user1	Virtual machine function used in a KVM environment
			Host OS user password		
			Host OS IP address	10.20.40.50	
			Host OS user name	user1	Cluster between units used
			Host OS user password		
			Host OS IP address	10.20.40.51	

A.13.2.2 GLS Setup Worksheet

Shown below are examples of the GLS Setup Worksheet.

When the cluster name is PRIMECLUSTER1

Item			Setting
GLS settings	Switching mode		\<Y>NIC switching mode (virtual IP takeover) \< >NIC switching mode (physical IP takeover I) \< >NIC switching mode (physical IP takeover II) \< >Fast switching mode
	Takeover virtual interface name		sha1:65
		Takeover virtual IP address (or host name)	10.20.30.46
		Subnet mask	255.255.255.0
	Node name (1)		primecl01
	Configuration information	Virtual interface name	sha1
		Primary physical interface name	eth3
		Secondary physical interface name	eth4
		Physical IP address (or host name)	10.20.30.40
		Logical IP address (or host name)	10.20.30.46
	Monitoring destination	IP address (or host name) of primary	10.20.30.61
		IP address (or host name) of secondary	10.20.30.62
	Optional function	Standby NIC patrol	\<>Disable \<Y>Enable
	Node name (2)		primecl02
	Configuration information	Virtual interface name	sha1
		Primary physical interface name	eth3
		Secondary physical interface name	eth4
		Physical IP address (or host name)	10.20.30.41
		Logical IP address (or host name)	10.20.30.46
	Monitoring destination	IP address (or host name) of primary	10.20.30.61
		IP address (or host name) of secondary	10.20.30.62
	Optional function	Standby NIC patrol	\<>Disable \<Y>Enable

When the cluster name is PRIMECLUSTER2

Item			Setting
GLS settings	Switching mode		\<Y>NIC switching mode (virtual IP takeover) \< >NIC switching mode (physical IP takeover I) \< >NIC switching mode (physical IP takeover II) \< >Fast switching mode
	Takeover virtual interface name		sha1:65
		Takeover virtual IP address (or host name)	10.20.30.146
		Subnet mask	255.255.255.0
	Node name (1)		primecl11
	Configuration information	Virtual interface name	sha1
		Primary physical interface name	eth3
		Secondary physical interface name	eth4
		Physical IP address (or host name)	10.20.30.140
		Logical IP address (or host name)	10.20.30.146
	Monitoring destination	IP address (or host name) of primary	10.20.30.161
		IP address (or host name) of secondary	10.20.30.162
	Optional function	Standby NIC patrol	\<>Disable \<Y>Enable
	Node name (2)		primecl12
	Configuration information	Virtual interface name	sha1
		Primary physical interface name	eth3
		Secondary physical interface name	eth4
		Physical IP address (or host name)	10.20.30.141
		Logical IP address (or host name)	10.20.30.146
	Monitoring destination	IP address (or host name) of primary	10.20.30.161
		IP address (or host name) of secondary	10.20.30.162
	Optional function	Standby NIC patrol	\<>Disable \<Y>Enable

A.13.2.3 GDS Configuration Worksheet

Shown below are examples of the GDS Configuration Worksheet.

Item			Setting
Class 1	Class name		class0005
	Class scope (node name)	Node 1	primecl01
		Node 2	primecl02
	Disk 1 connected to group	SDX disk name	Disk1
		Physical disk name in node 1	sdc
		Physical disk name in node 2	sdc
	Disk 2 connected to group	SDX disk name	Disk2
		Physical disk name in node 1	sdd
		Physical disk name in node 2	sdd
	Highest-order group 1	Group name	group0005
		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	volume0005
		Size	1024 blocks
Class 2	Class name		class0006
	Class scope (node name)	Node 1	primecl11
		Node 2	primecl12
	Disk 1 connected to group	SDX disk name	Disk3
		Physical disk name in node 1	sde
		Physical disk name in node 2	sde
	Disk 2 connected to group	SDX disk name	Disk4
		Physical disk name in node 1	sdf
		Physical disk name in node 2	sdf
	Highest-order group 1	Group name	group0006
		Group type	mirror
		Stripe width	-
		Disk/low-order group name — Disk/low-order group 1	Disk3
		Disk/low-order group 2	Disk4
		Volume 1 — Volume name	volume0006
		Size	1024 blocks

Appendix B Manual Pages

This appendix provides online manual page lists for CF, CIP, operator intervention, PAS, cluster resource management facility, RMS, shutdown facility (SF), tracing failed resource, SIS, Web-Based Admin View, procedure resource, 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**

Note

In some cases, "(1M)" may be output as the section number of the manual page that is displayed with the man command.
Should this occur, assume the section number to be "(8)."

B.1 CF

System administrator

Command	Function
cfconfig	Configures or unconfigures a node for a PRIMECLUSTER cluster.
cfregd	CF registry synchronization daemon
cfset	Applies or modifies /etc/default/cluster.config entries into the CF module.
cftool	Prints the node communications state of a node or the cluster.
changeng	Replaces a node group definition.
deleteng	Deletes a node group.
descng	Replaces a node group explanation.
detailng	Displays the dynamic expansion of a node group.
newng	Creates a new node group.
rcqconfig	Configures or starts the quorum operation of a cluster system.
rcqquery	Acquires the state of consistency (quorum) of the cluster.
showng	Displays the name and definition of the node group.

B.2 CIP

System administrator

Command	Function
cipconfig	Starts or stops CIP 2.0.
ciptool	Retrieves CIP information about local and remote nodes in the cluster.

File format

File	Format
cip.cf	CIP configuration file format

B.3 Operator Intervention

System administrator

Command	Function
clreply	Responds to an operator intervention request message.

B.4 PAS

System administrator

Command	Function
mipcstat	MIPC statistics

B.5 Cluster Resource Management Facility

System administrator

Command	Function
clautoconfig	Executes automatic resource registration.
clbackuprdb	Saves the resource database.
clinitreset	Resets the resource database.
clrestorerdb	Restores the resource database.
clsetparam	Checks the connections of the shared disk device and sets up the operation for automatic resource registration.
clsetup	Sets up the resource database.
clstartrsc	Activates a resource (GDS only).
clstoprsc	Deactivates a resource (GDS only).
clsyncfile	Distributes a file between cluster nodes.

User command

P Point

. . . .

There is also a "clgettree" command in the Web-Based System Administration tool WSA.

. . . .

Command	Function
clgettree	Outputs tree information for the resource database.

B.6 RMS

System administrator

Command	Function
hvassert	Asserts (tests for) an RMS resource state.
hvcm	Starts the RMS configuration monitor.
hvconfig	Displays or saves the RMS configuration file.

Command	Function
hvdisp	Displays RMS resource information.
hvdump	Collects debugging information about RMS.
hvlogclean	Cleans the RMS log files.
hvlogcontrol	Controls the size of the log disk.
hvshut	Shuts down RMS.
hvswitch	Switches control of an RMS user application resource to another host.
hvutil	Manipulates the availability of an RMS resource.

File format

File	Format
config.us	RMS configuration file
hvenv.local	RMS local environment valuables file

B.7 Shutdown Facility (SF)

System administrator

Command	Function
cldevparam	Changes and displays the tunable operation environment for asynchronous monitoring.
clmmbmonctl	Starts, stops, restarts, and displays the operating system of the MMB asynchronous monitoring daemon.
clmmbsetup	Registers, changes, deletes, and displays MMB information.
clvmgsetup	Registers, changes, deletes, and displays host OS information.
sdtool	Interface tool for shutdown daemon
rcsd	Shutdown daemon for shutdown manager

File format

File	Format
rcsd.cfg	Configuration file for shutdown daemon
SA_ipmi.cfg	Configuration file for IPMI Shutdown Agent
SA_blade.cfg	Configuration file for blade Shutdown Agent

B.8 Tracing Failed Resource

System administrator

Command	Function
cldispfaultrsc	Outputs a list of the current failed resources

B.9 SIS

System administrator

Command	Function
dtcpadmin	Starts the SIS administration utility.
dtcpd	Starts the SIS daemon for configuring VIPs.
dtcpdbg	Displays SIS debugging information.
dtcpstat	Displays state information on SIS.

B.10 Web-Based Admin View

System administrator

Command	Function
fjsvwvbs	Stops Web-Based Admin View.
wvCntl	Starts, stops, or gets debugging information for Web-Based Admin View.
wvGetparam	Displays the Web-Based Admin View environment variables.
wvSetparam	Sets the Web-Based Admin View environment variables.
wvstat	Displays the operating state of Web-Based Admin View.

B.11 Procedure Resource

System administrator

Command	Function
claddprocrsc	Registers an application resource that uses a state transition procedure.
cldelproc	Deletes a state transition procedure.
cldelprocrsc	Deletes an application resource that uses state transition procedure.
clgetproc	Gets a state transition procedure.
clsetproc	Registers a state transition procedure.
clsetprocrsc	Changes the registered information of an application resource that uses a state transition procedure.

User command

Command	Function
cldspproc	Outputs information on the resource that uses the state transition procedure.

B.12 RMS Wizards

System administrator

Command	Function
clrwzconfig	Sets up the linking function between the PRIMECLUSTER resource manager and the middleware products after the RMS configuration definitions are activated.

RMS Wizards and RMS Application Wizard

The RMS Wizard manual will be saved in the following directory when the SMAWRhvdo package is installed.

`/usr/doc/packages/SMAWRhv-do/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.

    - 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*."

2. Investigation information for the failed application

3. Crash Dump

    If the failed node is restartable, manually collect a crash dump before restarting it. Crash dump will be useful for troubleshooting if the failure is OS related.

    Example) If the failover occurred due to an unexpected resource failure

    After the failover of the cluster application is complete, collect a crash dump on the node where the resource failure occurred.

    For details on the crash dump, see "C.1.2 Crash Dump."

4. 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.

. . . .

See

. . . .

When the virtual machine function is used in a Xen environment, see "PRIMEQUEST Virtual Machine Function User's Manual".

. . . .

C.1.1 Executing the fjsnap Command

The fjsnap command is a tool which gathers system information necessary for analyzing the trouble at a time. When the trouble occurs by the PRIMECLUSTER system, the cause can be pursued by collecting necessary error information by the fjsnap command.

You can execute this command as follows:

1. Log in with system administrator authority.

2. Execute the "fjsnap" command.

    /usr/sbin/fjsnap -a *output*

The file name which becomes an output destination of system information collected by using the fjsnap command for output is specified.

See
. . . .
For details on the "fjsnap" command, see 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 crash dump. Then start the system in single user mode, and execute the "fjsnap" command.
For information on how to collect a crash dump, see "C.1.2 Crash Dump."

- 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 crash dump.
. . . .

C.1.2 Crash Dump

Check the crash dump directory for a crash dump created after the switchover had occurred. The time the dump was written can be found by referring to the time stamp using, for example, "ls(1)" command.

- If a crash dump after the switchover is found

  Save the crash dump.

- If a crash dump after the switchover is not found

  If the failed node is restartable, manually collect a crash dump before restarting it.

Information
. . . .
Crash dump directory

A crash dump is stored as a file on the node in which the error occurred.

When you are using the virtual machine function in the KVM environment and your guest OS has been forcefully stopped by the shutdown facility, the crash dump will be stored in the following directory for the hypervisor.

```
/var/crash/<shutdown time of the guest OS (YYYYMMDDHHMMSS)>.<Domain name for the guest OS>.core
```

Example: node1 was forcefully stopped at 12:34:56 on 20th April, 2011

/var/crash/20110420123456.node1.core
. . . .

C.1.3 PSA Information

If a problem occurs in PRIMEQUEST shown below, collect the PSA information in addition to the PRIMECLUSTER failure investigation information.

Server model

  PRIMEQUEST 1400S/1400E/1400L/1800E/1800L

Execute "getopsa" command.

```
/opt/FJSVpsa/sh/getopsa <filename>
```

Example:

```
/opt/FJSVpsa/sh/getopsa /tmp/node1_getopsa
```

See
. . . .
For details on the "getopsa" command, see "PRIMEQUEST 1000 Series Tool Reference."
. . . .

C.1.4 SVmco Information

If a problem occurs in PRIMEQUEST shown below, collect the SVmco information in addition to the PRIMECLUSTER failure investigation information.

Server model

    PRIMEQUEST 1400S2/1400E2/1400L2/1800E2/1800L2

Execute "getosvmco" command.

```
/opt/fujitsu/SVmco/sh/getosvmco <filename>
```

Example:

```
/opt/fujitsu/SVmco/sh/getosvmco /tmp/node1_getosvmco
```

See
. . . .
For details on the "getosvmco" command, see "PRIMEQUEST 1000 Series ServerView Mission Critical Option User Manual."
. . . .

C.2 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 "5.2 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.2.1 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:
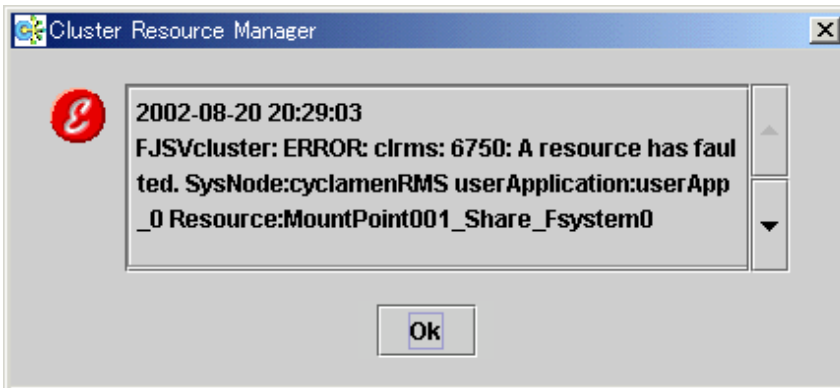
🍄 **Note**

If a message frame title says "Cluster resource management facility," refer to the following:

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

    - "11.10 Cluster Resource Management Facility Messages"

- Appendix D in this manual

    - "D.5.2 Failed Resource and Operator Intervention Messages (GUI)"

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 icon] **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.2.2 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*.
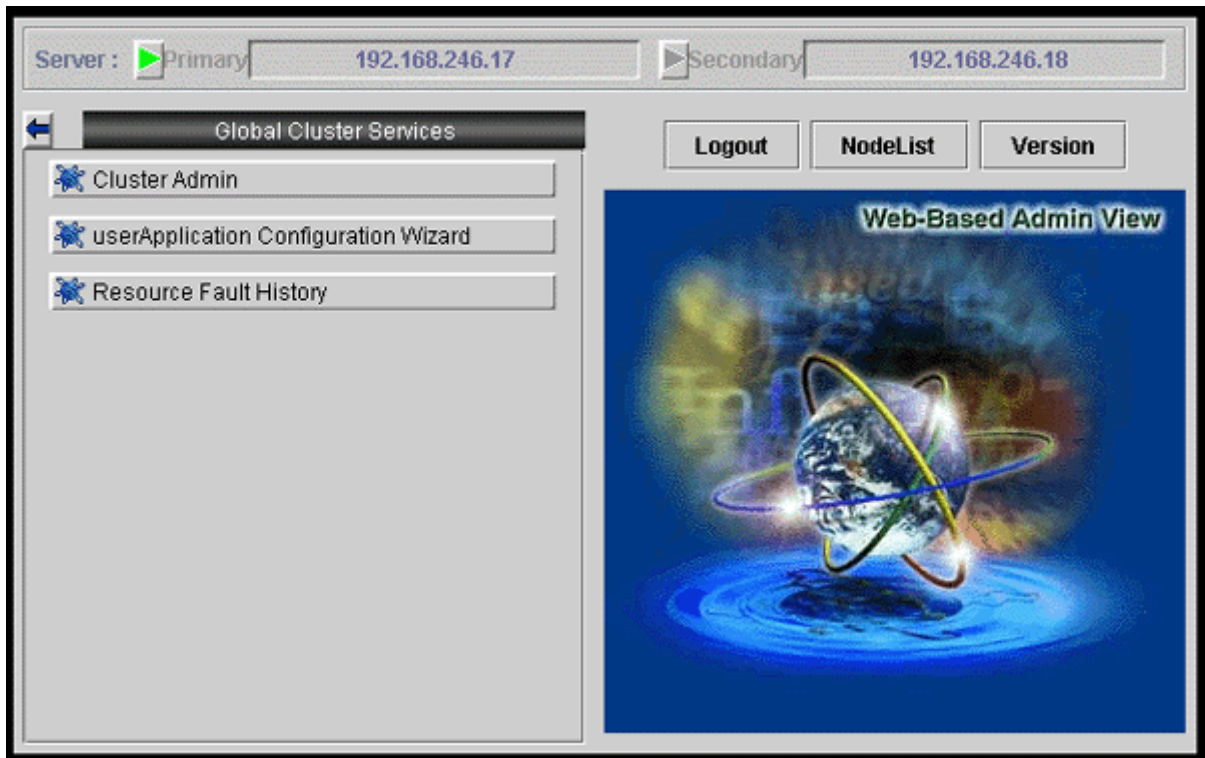
The "Resource Fault History" will be displayed.
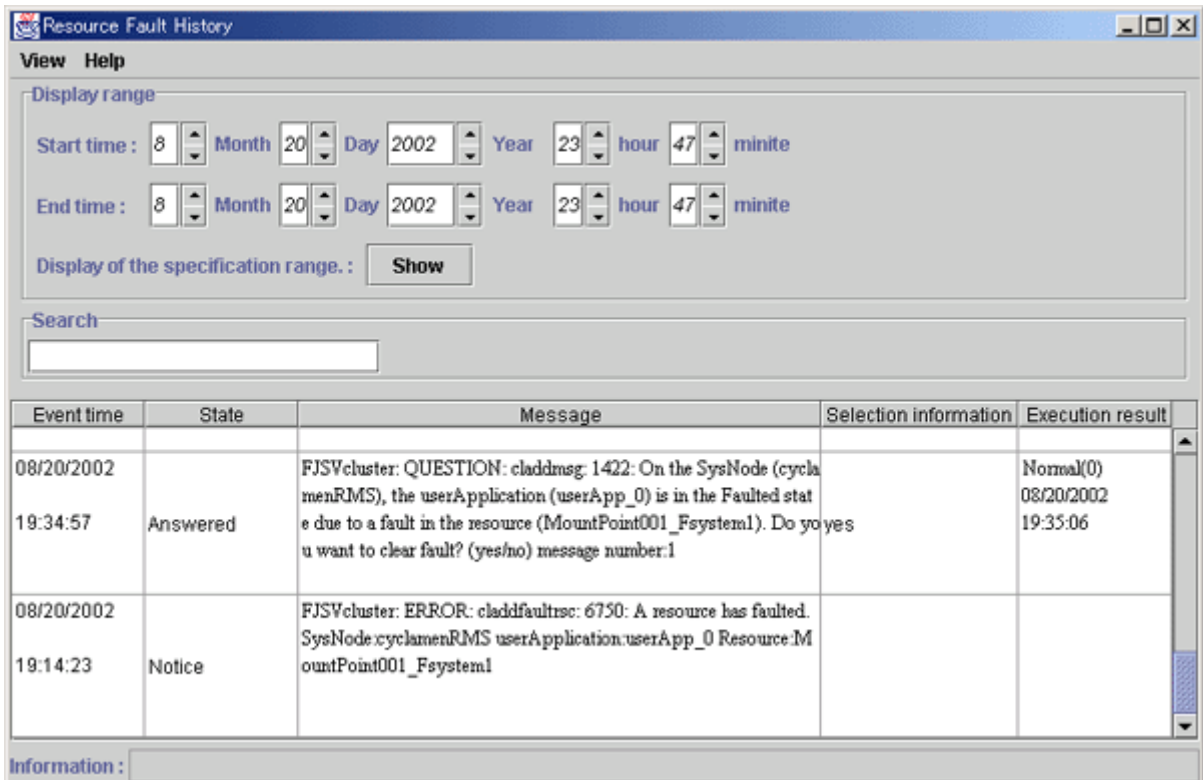
Note

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.2.3 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" command, nothing will be displayed.

- Execution result - The result and time of the response processing are displayed.
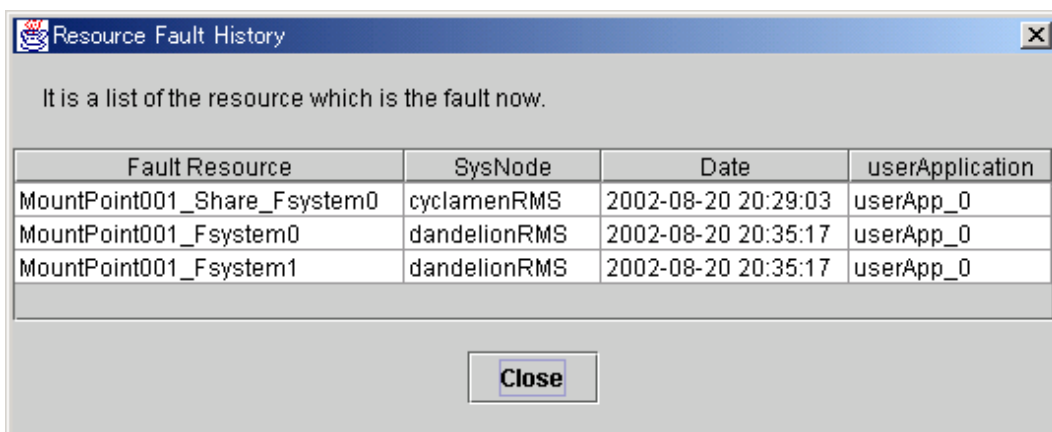
Information field

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.2.3 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" 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

- GUI Messages

    - Messages Related to CRM View

    - Failed resource and operator intervention messages

- Messages Related to the Setting of the Shutdown Facility

D.1 Searching for a Message

This section explains the procedure for searching for a message reference.

The PRIMECLUSTER display formats are classified into five types. Determine the reference according to the format of the displayed message.

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	CF Configuration and Administration Guide 11.10.2.1 HALT Messages
Information (INFO)	2620, 2621, 2622, 2700, 2701	D.2 Information Messages
	2100, 2200, 2201, 2202, 2203, 2204, 2205, 3200, 3201, 3202, 3203, 3204, 3205, 3206	CF Configuration and Administration Guide 11.10.2.2 INFORMATION messages

severity	message-number	Reference
	3080, 3081, 3082, 3083, 3084, 3085	CF Configuration and Administration Guide 11.13.1 INFORMATION messages
Warning (WARNING)	4250, 5021, 5200	CF Configuration and Administration Guide 11.10.2.3 WARNING messages
Error (ERROR)	6000, 6001, 6002, 6003, 6004, 6007, 6009, 6208, 6615, 6616, 6617, 6618, 6619, 6621, 6623, 6624, 6651, 6680, 6750, 6751, 6752, 6753, 6754, 6755, 6817	D.3 Error Messages
	6000-6001, 6003-6010, 6021, 6025 6200-6204, 6206-6218, 6300,6600-6604, 6606-6608, 6611, 6614-6616, 6653, 6900-6904 7500-7522, 7535-7543, 7545-7546	CF Configuration and Administration Guide 11.13.3 ERROR message
Question (QUESTION)	1421, 1422, 1423	D.4 Operator Intervention Messages

Display format 2

```
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 2 are described in the *PRIMECLUSTER Cluster Foundation (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	11.1 cfconfig message
CIP2.0 start/stop	cipconfig	11.2 cipconfig message
Outputting the node communication state	cftool	11.3 cftool message
Outputting and manipulating quorum information	rcqconfig	11.4 rcqconfig message
	rcqquery	11.5 rcqquery message
CF execution	CF	11.6 CF runtime messages

Display format 3

```
(SMAWsf,number1,number2):message
```

number1,number2:

Indicates the message number.

message:

Indicates the message.

Search procedure

Messages of display format 3 are described in *PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide*.

Confirm *number1* and *number2*, and then determine the reference by referring to the table below.

Operation	(*number1-number2*)	Reference location in CF Configuration and Administration Guide
Shutdown facility	(10, 2)-(10, 12) (10, 15)-(10, 17), (10, 19), (10, 20) (10, 23), (10, 101) (30, 2)-(30, 10), (30, 12)-(30, 17) (50, 3)-(50, 4), (50, 6), (50, 9)	11.11 Shutdown Facility

Display format 4

```
(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 4 are described in the "*PRIMECLUSTER Reliant Monitor Services (RMS) Reference Guide*." Confirm *err-code* and *message-number*, and then determine the reference by referring to the table below.

err-code	*message-number*	Reference location in RMS Reference Guide
ADC	1, 2, 3, 4, 5, 15, 17, 18, 20, 25, 27, 30, 31, 32, 33, 34, 35, 37, 38, 39, 41, 43, 44, 45, 46, 47, 48, 49, 57, 58, 59, 60, 61, 62, 63	5 Non-fatal error messages
ADM	55, 57, 58, 59, 60, 62, 63, 67, 70, 95, 96	
BM	8, 14, 52, 54, 68	
CML	11, 12, 17, 18, 19, 20, 21	
CRT	1, 2, 3, 4	
CUP	2, 3, 5, 7, 8	
DET	1, 3, 4, 5, 6, 7, 12, 13, 14, 24, 28, 33	
GEM	4	
INI	1, 9	
QUE	13	
SCR	8, 20	

err-code	message-number	Reference location in RMS Reference Guide
SWT	4, 20, 25, 26, 45	
SYS	1, 8, 13, 14, 15, 17, 48, 49, 50, 84, 90, 93, 94, 97, 98	
UAP	1, 5, 6, 7, 8, 9, 15, 16, 19, 21, 22, 23, 24, 27, 28, 29, 32, 33, 35, 41, 42	
US	5, 6, 31	
WLT	1, 3	
WRP	3, 4, 9, 11, 12, 14, 15, 16, 17	
ADC	16, 21	6 Fatal error messages
ADM	1, 2	
BM	3, 49, 51, 58, 67	
CML	14	
CMM	1, 2, 3	
DET	8, 9, 16	
HVC	14	
INI	7	
MIS	4	
QUE	1, 2, 5, 6, 11, 12	
SCR	4, 5, 10, 12, 13, 15, 18	
SYS	33, 52	
UAP	36	
US	1	

Display format 5 (other)

If the display format is other than display format 1 to 4, see the following:

- See "11.7 CF Reason Code Table" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide.*"

- See "11.8 Error messages for different systems" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide.*"

- See "7 Console error messages" in the "*PRIMECLUSTER Reliant Monitor Services (RMS) Reference Guide.*"

- See "Appendix E GDS Messages" in the "*PRIMECLUSTER Global Disk Services Configuration and Administration Guide.*"

- See "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.*"

📖 Information
. . . .
Messages that contain the character string "SDX:" or "sfdsk:" are GDS messages.
. . . .

D.2 Information Messages

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 are 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.

D.3 Error Messages

If an error message is output, analyze the log of the time zone in which the message was output from the "/var/log/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.

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 Installation Guide for PRIMECLUSTER, 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.6.1 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 Installation Guide for PRIMECLUSTER, 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.6.1 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.

6007 One of the required options (option) must be specified.

Corrective action:

Specify a correct option, and execute the command again.

option indicates an option.

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.

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.

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(1)" 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(1)" command. Correct the insufficient resource problem, and do the operation again.

If the problem cannot be solved even by applying 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" 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 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" command. For details on the "clgettree" command, see the manual page for clgettree.
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.

6680 The specified directory does not exist.

Corrective action:

Specify a correct directory, and then re-execute.

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" 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" 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" 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" 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.

6817 An error occurred during state transition procedure execution. (*error* procedure:*procedure* detail:*code1-code2-code3-code4-code5-code6-code7*)

Corrective action:

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.

Collect debugging information and contact your Fujitsu system engineer. For details on collecting debugging information, see "Appendix C Troubleshooting."

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" command. To register the state transition procedure, execute the "clsetproc" 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.

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

D.4 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.

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. For details on the "clreply" command, see the "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" command. For details on the "hvdisp" command, see the "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.

> 📄 **Note**
> ···
> If this message is output to the console and /var/log/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.
> ···

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.

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.

> 📄 **Note**
> ···
> If this message is output to the console and /var/log/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.
> ···

D.5 GUI Messages

The section explains GUI messages.

- Messages related to the CRM view.

- Failed resource and operator intervention messages

D.5.1 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:

- "D.5.2 Failed Resource and Operator Intervention Messages (GUI)"

D.5.2 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 "*PRIMECLUSTER 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.6　Messages Related to the Setting of the Shutdown Facility

D.6.1　Notes on the Setting of the Shutdown Facility

When you failed to execute the "panicinfo_setup" command and reboot any one of the cluster nodes, you need to confirm that the "/etc/opt/FJSVcllkcd/etc/SA_lkcd.idx" file is MYINDEX -1. If not, reboot the node after editing the file using an editor.

D.6.2　Error Messages

If panicinfo_setup fails, the following messages will be output to the standard error output according to the cause.

ERROR: *<command>* failed

Content:

Execution of the *<command>* command failed.

Corrective action:

Check the command and execute it again.

ERROR: *<command>* failed on *<node>*

Content:

The *<command>* command failed on the *<node>* node.

Corrective action:

Check the command and execute the *<command>* command on the *<node>* node again.

ERROR: *<command>* timeout

Content:

Execution of the *<command>* command timeout.

Corrective action:

Check the command and execute the *<command>* command on the *<node>* node again.

ERROR: failed to distribute index file to *<node>*

Content:

Distribution of the index file to *<node>* node failed.

Corrective action:

Confirm that the CF services (CFSH, CFCP) have started normally.

ERROR: failed to distribute /etc/panicinfo.conf file to *<node>*

Content:

Distribution of the panicinfo definition file to *<node>* node failed.

Corrective action:

Confirm that the CF services (CFSH, CFCP) have started normally.

ERROR: /etc/sysconfig/netdump is invalid on *<node>*
ERROR: Cannot find the Netdump client's IP address for *<device>* on *<node>*

Content:

Reading of the Netdump client settings on the *<node>* node failed.

Corrective action:

Confirm that Netdump is normally set up.

WARNING: /etc/panicinfo.conf file already exists.
(I)nitialize, (C)opy or (Q)uit (I/C/Q) ?

Content:

The panicinfo configuration file already exists.

Corrective action:

Enter "I" to initialize the panicinfo configuration file, "C" to distribute the panicinfo configuration file on the command execution node, and "Q" to terminate the command.

ERROR: failed to change mode of index file on *<node>*

Content:

Changing of the index file mode on *<node>* node failed.

Corrective action:

Confirm that the "chmod" command can be executed normally on <*node*> node for the "/etc/opt/FJSVcllkcd/etc/SA_lkcd.idx" file, and then re-execute the command.

ERROR: failed to patch rcsd.cfg on <*node*>

Content:

Updating of rcsd.cfg failed on the <*node*> node.

Corrective action:

Confirm that the "/etc/opt/SMAW/SMAWsf/rcsd.cfg" file is included in the <*node*> node and then, re-execute the command.

ERROR: failed to change owner of index and rcsd.cfg file on <*node*>

Content:

Changing of the owner on the <*node*> node failed.

Corrective action:

Confirm that the "chown" command can be executed for the "/etc/opt/FJSVcllkcd/etc/SA_lkcd.idx" file and for the "/etc/opt/SMAW/ SMAWsf/rcsd.cfg" file on the <*node*> node, and then re-execute the command.

ERROR: failed to change group of index and rcsd.cfg file on <*node*>

Content:

Changing of the owner group on the <*node*> node failed.

Corrective action:

Confirm that the "chgrp" command can be executed for the "/etc/opt/FJSVcllkcd/etc/SA_lkcd.idx" file and for the "/etc/opt/SMAW/ SMAWsf/rcsd.cfg" file on the <*node*> node, and then re-execute the command.

ERROR: failed to change mode of /etc/panicinfo.conf file on <*node*>

Content:

Changing of the panicinfo configuration file mode on <*node*> node failed.

Corrective action:

Confirm that the "chmod" command can be executed normally on <*node*> node for the "/etc/panicinfo.conf" file, and then re-execute the command.

ERROR: failed to change owner of /etc/panicinfo.conf file on <*node*>

Content:

Changing of the owner on the <*node*> node failed.

Corrective action:

Confirm that the "chmod" command can be executed normally for the "/etc/panicinfo.conf" file on the <*node*> node, and then re-execute the command.

ERROR: failed to change group of /etc/panicinfo.conf file on <*node*>

Content:

Changing of the owner group on the <*node*> node failed.

Corrective action:

Confirm that the "chgrp" command can be executed normally for the "/etc/panicinfo.conf" file on the <*node*> node, and then re-execute the command.

ERROR: internal error, ...

Content:

Internal error.

Corrective action:

Copy this message, and then contact your Fujitsu systems engineer (SE).

ERROR: The Shutdown Facility configuration cannot be found.

Content:

The Shutdown Facility configuration file is not found. You might not have completed configuring the Shutdown Facility.

Corrective action:

Confirm that Shutdown Facility configuration is completed then execute the command again.

ERROR: Reading the Shutdown Facility configuration failed.

Content:

Reading Shutdown Facility configuration failed. You might have set the Shutdown Facility incorrectly.

Corrective action:

Confirm that Shutdown Facility configuration is correct then execute the command again.

ERROR: Reading the Shutdown Agent configuration failed.

Content:

Reading Shutdown Agent configuration failed. You might have set the Shutdown Agent incorrectly.

Corrective action:

Confirm that Shutdown Agent configuration is correct then execute the command again.

ERROR: *<File name>* generation failed.

Content:

Creation of the *<file name>* file failed.

Corrective action:

Confirm that the *<file name>* file can be created then execute the command again.

ERROR: The Blade Shutdown Agent configuration cannot be found.

Content:

The Blade Shutdown Agent configuration file is not found. You might not have completed configuration the Blade Shutdown Agent.

Corrective action:

Confirm that Blade Shutdown Agent configuration is completed then execute the command again.

ERROR: The IPMI Shutdown Agent configuration cannot be found.

Content:

The IPMI Shutdown Agent configuration file is not found. You might not have completed configuration the IPMI Shutdown Agent.

Corrective action:

Confirm that IPMI Shutdown Agent configuration is completed then execute the command again.

Appendix E Registering, Changing, and Deleting State Transition Procedure Resources for PRIMECLUSTER Compatibility

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.

E.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.

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" command to register the state transition procedure.

   ### See
   ....
   For details on the "clsetproc" 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" command to register the procedure resource.

   ### See
   ....
   For details on the "claddprocrsc" command, see the manual page for claddprocrsc .
   ....

   ### 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
```

E.2 Changing a Procedure Resource

This section explains how to change a procedure resource.

E.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" command to retrieve the state transition procedure.

    ### See

    For details on the "clgetproc" command, see the manual page.

    ### 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

    Register the state transition procedure by using the "clsetproc" command.

    For registration, specify the "-o" option to overwrite the state transition procedure.

    ### 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 -o /tmp/program
    ```

E.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 a procedure resource again.

For more details, see "8.1 Changing the Cluster 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
```
....

E.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" command to change the registration information of the procedure resource.

See
....

For details on the "clsetprocrsc" command, see the manual page.
....

Example
....

When the procedure resource with the following conditions receives a state transition request of START RUN BEFORE in addition to START RUN AFTER and STOP RUN BEFORE;

- The resource key of the procedure resource is SDISK,

- This procedure resource uses the state transition procedure registered in the BasicApplication class, and

- The procedure resource is registered in the node (NODE1)

```
# /etc/opt/FJSVcluster/bin/clsetprocrsc -n SDISK -c BasicApplication -s NODE1 -K BEFORE,AFTER
-S BEFORE
```
....

E.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" command to delete the procedure resource.

   📇 **See**
   ....
   For details on the "cldelprocrsc" 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" command to delete the state transition procedure.

   📇 **See**
   ....
   For details on the "cldelproc" 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 F Registering PRIMECLUSTER-compatible Resources

This section explains how to register a resource of a PRIMECLUSTER-compatible product to a cluster service instance and how to delete a resource.

F.1 Associating Resources to a Cluster Service Instance

The PRIMECLUSTER-compatible cluster service instance needs to be associated with the following product:

- Interstage Traffic Director

![Note icon] **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 PRIMECLUSTER-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 icon] **Note**

   Perform Step 2 only if userApplication was set up with the CUI and IP address takeover was set during userApplication setup. If IP address takeover was not set during userApplication setup, Step 2 is unnecessary.

2. Register the IP address takeover resources.

   - Execute the "clgettree" 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" 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...(*)
takeover-IP-address(**)
```

* For "*resource-ID-of-PublicLAN-resource...*," specify the resource IDs of the PublicLAN resources of all nodes that were confirmed with the previous "clgettree" command.
If takeover of the same IP address is set to multiple PublicLAN resources in one node, specify the resource ID of one of the PublicLAN resources being used.

** For "*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 (*)

     * If IP address takeover resources were registered in Step 2, associated them to cluster service instances.

   - "Interstage Traffic Director" resources

     1. Execute the "clgettree" 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 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 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" 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(*) -t  resource-ID-of-cluster-service-instance(**)
```

* Specify the resource IDs of the IP address takeover and Interstage Traffic Director resources that were obtained with the "clgettree" command in each node.

** Execute the "clgettree" 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
```

F.2 Deleting Takeover Network Resources

This section describes the procedure to be performed in the following cases:

- If userApplication is deleted

- If takeover network resources are deleted

Delete entries which are configured for the takeover network resources in the following environment files:

- /usr/opt/reliant/etc/hvipalias

- /etc/inet/hosts

Appendix G Configuration Update Service for SA

This appendix explains Configuration Update Service for SA.

G.1 Feature Description

This function is enabled when building a cluster system in combination with ServerView Resource Orchestrator Virtual Edition.

This configuration provides active, standby and spare nodes. Normally, cluster applications are started on the active node. In the event of a failure on the active node, the cluster applications fail over to the standby node.

After this, the spare node is started up using data from the storage, thus keeping the cluster configuration alive and ensuring even higher availability.

Configuration Update Service for SA is a function that automatically retrieves the BMC or iRMC IP address of the spare node and then updates the configurations of shutdown agents on the local node and other cluster nodes when starting the operating system.
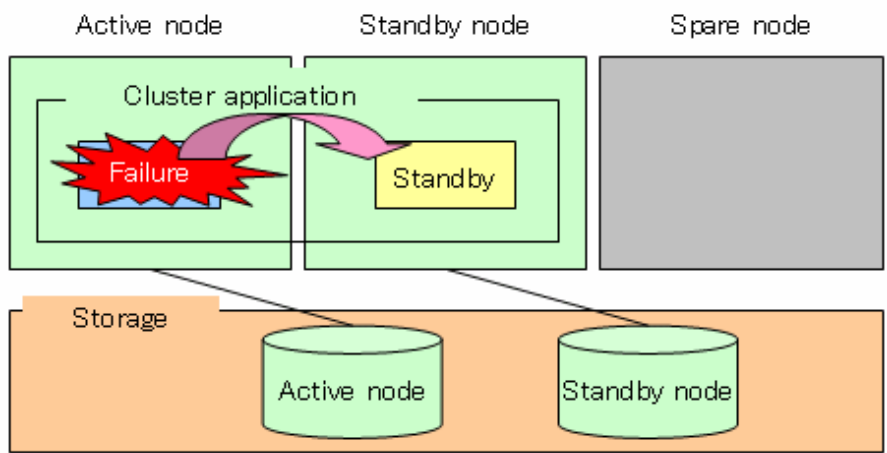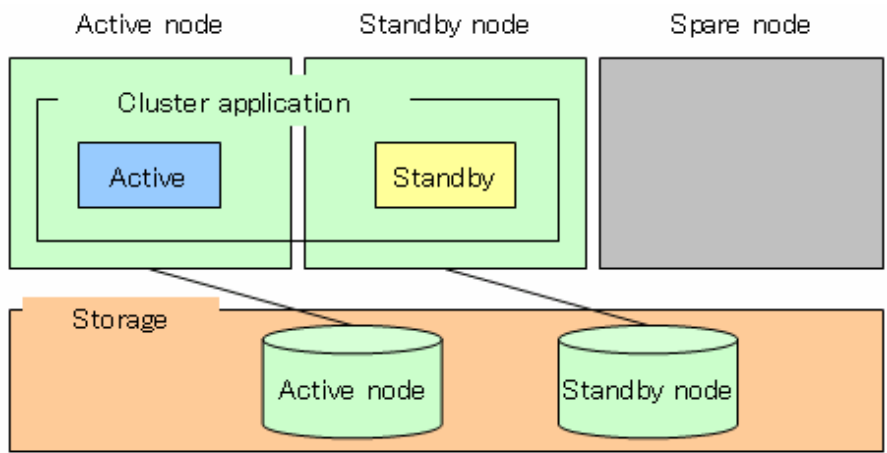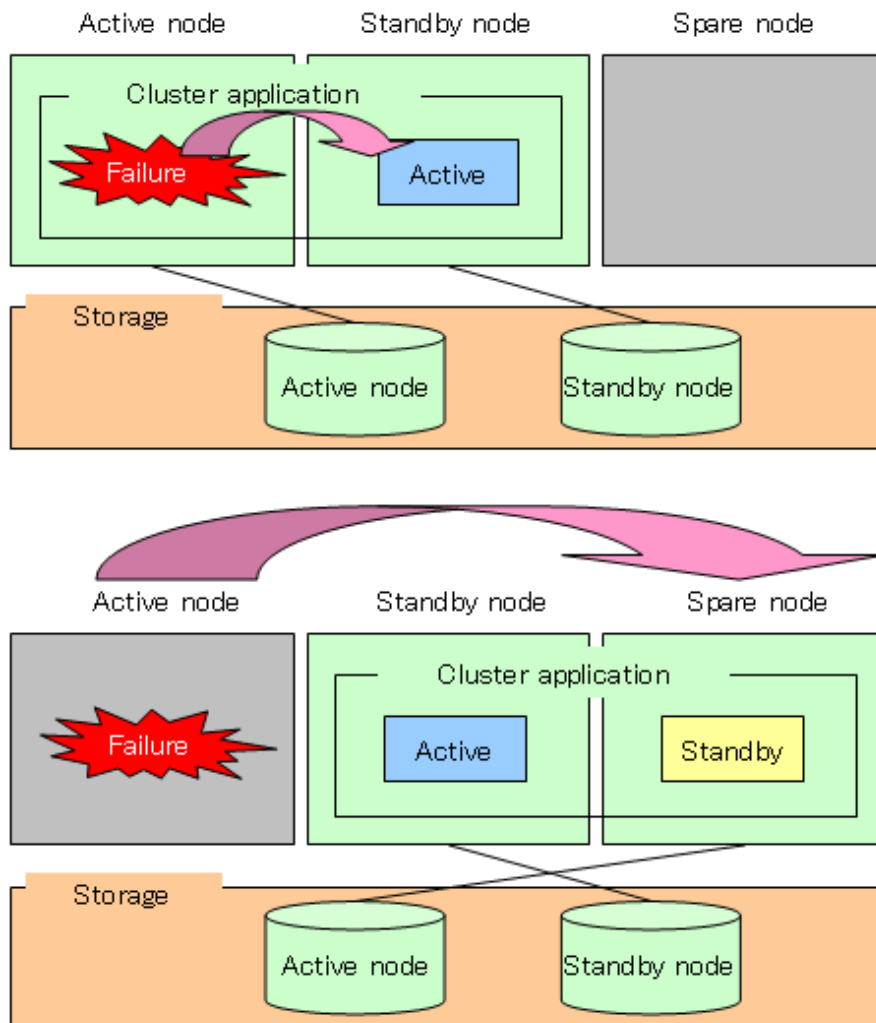
![Note icon] **Note**

- - Use the same user name and password for BMC or iRMC on every blade.

- - If the PersistentFault attribute of RMS is set to "1," the Fault information is kept even if RMS is started on a normal spare blade. (The default value of the PersistentFault attribute is "0.")

- - When you update the configuration file for the shutdown agent, the updated configuration file is distributed to nodes in which the communication is available. The file is not distributed to nodes in which operation is stopped or the network communication is not available.

  In addition, when you start multiple nodes simultaneously, the configuration file for the shutdown agent is updated and distributed on multiple nodes at the same time. In this case, inconsistencies may occur in the information of the configuration file for the shutdown agent stored in each node.

  To check that correct information is distributed to all nodes, execute the following command on any node when all nodes are activated.

  ```
  # /opt/SMAW/SMAWsf/bin/sfsacfgupdate -s
  ```

  As a result of check, if inconsistencies occur in the information, restore the service according to the procedure in "G.6 Restoration."

G.2 Operation Environment

You need the following environment to use Configuration Update Service for SA:

- Server model

    Models using the IPMI shutdown agent (SA_ipmi)

See

For details on models using the IPMI shutdown agent, see "5.1.2 Setting Up the Shutdown Facility."

- Operating system

    The following operating systems are supported:

    - Red Hat Enterprise Linux 5 (for x86)

    - Red Hat Enterprise Linux 5 (for Intel64)

    - Red Hat Enterprise Linux 6 (for x86)

    - Red Hat Enterprise Linux 6 (for Intel64)

Note

This service is not available in a virtual machine environment.

- Required package

    - Red Hat Enterprise Linux 5

        - OpenIPMI

        - OpenIPMI-tools

    - Red Hat Enterprise Linux 6

        - OpenIPMI

        - ipmitool

    Check that the packages described above are installed by executing the rpm command. Install packages if they are not installed.

    Packages are included in the installation media for the operating system.

G.3 Configuration

This section describes how to set up this service.

G.3.1 Startup Configuration for the IPMI Service

Configuration Update Service for SA performs startup configuration of the IPMI service to use the IPMI service when obtaining the BMC or iRMC IP address of the server.

1. Starting the IPMI service

    Execute the following command on all nodes to check the startup status of the IPMI service.

    ```
    # /sbin/service ipmi status
    ```

    Execute the following command on all nodes in which the IPMI service is not activated to start the IPMI service.

    ```
    # /sbin/service ipmi start
    Starting ipmi drivers:                                    [  OK  ]
    ```

2. Setting the run level of the IPMI service

    Check the current run level of the IPMI service on all nodes.

```
# /sbin/chkconfig --list ipmi
ipmi            0:off   1:off   2:off   3:on    4:off   5:off   6:off
```

Note

....
Copy the value when the run level is set to "on." The above example shows that the run level is 3. This value is required for canceling this service and restoring the environment.
....

Execute the following command on all nodes to read the IPMI service on startup.

```
# /sbin/chkconfig ipmi on
```

Information

....
You can set "on" to run levels only that you want to activate this service. In this case, specify run levels in the range from 2 to 5.
....

G.3.2 Activating Configuration Update Service for SA

Execute the sfsacfgupdate command to activate this service.

See

....
For details on the sfsacfgupdate command, see "G.7 sfsacfgupdate."
....

G.3.2.1 Setting Run Levels of Configuration Update Service for SA

Execute the following command on all nodes to activate this service.

```
# /opt/SMAW/SMAWsf/bin/sfsacfgupdate -e
```

Information

....
If you set "on" to run levels only that you want to activate this service in Step 2 in "G.3.1 Startup Configuration for the IPMI Service," specify the values of the run levels.

- Example: When activating this service in run levels 3 and 5

```
# /opt/SMAW/SMAWsf/bin/sfsacfgupdate -e 35
```

When run levels are omitted, this service is activated in run levels 2 to 5.
....

G.3.2.2 Checking the Configuration

Check the setup status on all nodes.

```
# /opt/SMAW/SMAWsf/bin/sfsacfgupdate -c
Configuration file exists.            [ OK ]
ipmitool command exists.              [ OK ]
ipmi service has been started.        [ OK ]
ipmi service's run level :
0:off   1:off   2:on    3:on    4:on    5:on    6:off
Configuration Update Service's run level :
0:off   1:off   2:on    3:on    4:on    5:on    6:off
```

The following describes items to be displayed and contents to be checked.

Configuration file exists.

The existence of the configuration file for the shutdown agent is displayed.

Check that the status is "OK."

If the status is "NG," the configuration file for the shutdown agent does not exist. Review the configuration of the shutdown agent.

ipmitool command exists.

The existence of the ipmitool command is displayed.

Check that the status is "OK."

If the status is "NG," the ipmitool command does not exist. Check that the ipmitool command is correctly installed.

ipmi service has been started.

The startup status of the IPMI service is displayed.

Check that the status is "OK."

If the status is "NG," the IPMI service is not activated. Perform Step 1 in "G.3.1 Startup Configuration for the IPMI Service" again.

ipmi service's run level :

The statuses of run levels for the IPMI service are displayed.

Check that the status of each run level you set in Step 2 in "G.3.1 Startup Configuration for the IPMI Service" is "on."

If the status of each run level is different, perform Step 2 in "G.3.1 Startup Configuration for the IPMI Service" again.

Configuration Update Service's run level :

The status of the run level for Configuration Update Service for SA is displayed.

Check that the status is "on" in all or any levels from 2 to 5.

When the following error message is output, perform "G.3.2.1 Setting Run Levels of Configuration Update Service for SA" again.

```
Configuration Update Service's run level :
sfsacfgupdate: ERROR: "sfsacfgupdate -e [<levels>]" is not executed.
```

Note

Check that run levels, which are "on" in "Configuration Update Service's run level" are also "on" in "ipmi service's run level."

If the status of each run level is not identical, any setting may be incorrect. Review the processes of Step 2 in "G.3.1 Startup Configuration for the IPMI Service" and "G.3.2.1 Setting Run Levels of Configuration Update Service for SA."

G.3.2.3 Checking the BMC or iRMC IP Address and the Configuration Information of the Shutdown Agent

To check the BMC or iRMC IP address and the configuration information of the shutdown agent, execute the sfsacfgupdate command on any node.

Check that the following information is consistent with the displayed contents.

- BMC or iRMC IP address of each node ("BMC IP Address :")

- BMC or iRMC IP address of each node stored in the configuration file for the shutdown agent in each node. ("Configuration file :")

Example) Three-node cluster with nodeA, nodeB, and nodeC

The BMC IP address of each node is as follows:

```
nodeA: 10.20.30.41
nodeB: 10.20.30.42
nodeC: 10.20.30.43
```

```
# /opt/SMAW/SMAWsf/bin/sfsacfgupdate -s
Node : nodeA
 Node status : UP
 Configuration Update Service status : ENABLE
 BMC IP Address :
  10.20.30.41
 Configuration file :
  nodeA 10.20.30.41
  nodeB 10.20.30.42
  nodeC 10.20.30.43

Node : nodeB
 Node status : UP
 Configuration Update Service status : ENABLE
 BMC IP Address :
  10.20.30.42
 Configuration file :
  nodeA 10.20.30.41
  nodeB 10.20.30.42
  nodeC 10.20.30.43

Node : nodeC
 Node status : UP
 Configuration Update Service status : ENABLE
 BMC IP Address :
  10.20.30.43
 Configuration file :
  nodeA 10.20.30.41
  nodeB 10.20.30.42
  nodeC 10.20.30.43
```

The following describes displayed items.

Node :

The node name is displayed.

Node status :

The startup status of the node is displayed.

When the node is running, the status is "UP." For other than "UP," the subsequent information is not displayed.

Configuration Update Service status :

The setup status of Configuration Update Service for SA is displayed.

If no problem is found in "G.3.2.2 Checking the Configuration," the status is "ENABLE." For other than "ENABLE," the subsequent information is not displayed.

BMC IP Address :

The current BMC or iRMC IP address is displayed.

Configuration file :

The BMC or iRMC IP address of each node stored in the current configuration file for the shutdown agent is displayed.

G.4 Operation Check

This section describes how to check the operation of this service.

G.4.1 Operation Check by Restarting the System

This service operates on operating system startup. Therefore, you need to restart the system to check the operation.

The following describes how to check the operation by temporarily and manually updating the configuration file for the shutdown agent and restarting the system.

1. Backing up the configuration file for the shutdown agent

   Back up the SA_ipmi.cfg file on a node, hereafter referred to as nodeA, for which you check the operation.

   ```
   # cp -p /etc/opt/SMAW/SMAWsf/SA_ipmi.cfg /etc/opt/SMAW/SMAWsf/SA_ipmi.cfg.bk
   ```

2. Updating the configuration file for the shutdown agent

   Change the BMC or iRMC IP address of nodeA in the configuration file for the shutdown agent in nodeA to an unused IP address.

   ```
   # vi /etc/opt/SMAW/SMAWsf/SA_ipmi.cfg
   ~~~
   nodeA 10.20.30.41:user:pass cycle
      The new address is as follows:
   nodeA 255.255.255.255:user:pass cycle <- Change to an unused IP address
   ```

   ## 🖙 Note

   . . . .

   When you change the IP address, the following message may be displayed on Syslog. As a result of execution of sdtool -s, the state of SA_ipmi may be "TestFailed," however, there is no problem.

   ```
   SMAWsf : SA SA_ipmi to test host <node> failed
   ```

   . . . .

3. Restarting the system

   Restart nodeA.

   ```
   # shutdown -r now
   ```

4. Checking the configuration file for the shutdown agent

   Check that the BMC or iRMC IP address of nodeA is updated in the configuration file for the shutdown agent in nodeA.

   ```
   # vi /etc/opt/SMAW/SMAWsf/SA_ipmi.cfg
   ~~~
   nodeA 10.20.30.41:user:pass cycle
   ```

5. Deleting the backup file

   Delete the backed up configuration file for the shutdown agent on nodeA.

   ```
   # rm -f  /etc/opt/SMAW/SMAWsf/SA_ipmi.cfg.bk
   ```

G.5 Cancellation

The following describes how to cancel this service.

G.5.1 Deactivating Configuration Update Service for SA

Execute the following command on all nodes to deactivate the configuration of this service.

```
# /opt/SMAW/SMAWsf/bin/sfsacfgupdate -d
```

G.5.2 Restoring the Startup Configuration of the IPMI Service

When you do not need the IPMI service, return the run level of the IPMI service to the status before Step 2 in "G.3.1 Startup Configuration for the IPMI Service" is performed.

G.6 Restoration

This section describes restoration methods if correct information is not distributed to all nodes when this service operates.

G.6.1 Restoration Method When Correct Information is not Distributed to All Nodes

If communication is not available in other nodes when Configuration Update Service for SA operates, the configuration file for the shutdown agent is not distributed to other nodes.

In this case, consistency of the information stored in each node is not ensured and the shutdown agent cannot operate normally.

The following example shows the restoration methods in such a case.

If distribution of the configuration file for the shutdown agent to nodeB fails because the BMC IP address of nodeA is changed and nodeB is stopped:

nodeA: 10.20.30.41 -> Changed to 10.20.30.51
nodeB: 10.20.30.42 (Stopped)
nodeC: 10.20.30.43

1. Checking the message

   On nodeA, for which the BMC IP address is changed, check that any of the following messages are output on Syslog.

   ```
   sfsacfgupdate: ERROR: Failed to copy the backup of <file> on node <node>.
   ```

   ```
   sfsacfgupdate: ERROR: Failed to distribute <file> to node <node>.
   ```

   ```
   sfsacfgupdate: ERROR: Failed to change the access permission of <file> on node <node>.
   ```

   ```
   sfsacfgupdate: ERROR: Failed to change the group of <file> on node <node>.
   ```

   ```
   sfsacfgupdate: ERROR: Failed to change the owner of <file> on node <node>.
   ```

   If any of the above messages are output, the process for *<node>* has failed.

2. Checking the stopped node

   If all the other nodes stop while Configuration Update Service for SA is operating, messages in Step 1 are not output. Check if any nodes stop.

3. Restoring the stopped node

   Check that the status of the stopped node and restore it.

4. Checking the current status

   Execute the following command on any node to check the current status.

   ```
   # /opt/SMAW/SMAWsf/bin/sfsacfgupdate -s
   Node : nodeA
    Node status : UP
    Configuration Update Service status : ENABLE
    BMC IP Address :
     10.20.30.51 <- Changed from 10.20.30.41
    Configuration file :
     nodeA 10.20.30.51  <- Updated with the changed information on nodeA
     nodeB 10.20.30.42
     nodeC 10.20.30.43

   Node : nodeB
    Node status : UP
    Configuration Update Service status : ENABLE
    BMC IP Address :
     10.20.30.42
   ```

```
 Configuration file :
  nodeA 10.20.30.41 <- Not updated with the changed information on nodeB
  nodeB 10.20.30.42
  nodeC 10.20.30.43

Node : nodeC
 Node status : UP
 Configuration Update Service status : ENABLE
 BMC IP Address :
  11.22.33.46
 Configuration file :
  nodeA 10.20.30.51 <- Updated with the changed information on nodeC
  nodeB 10.20.30.42
  nodeC 10.20.30.43
```

In the above example, you can see the BMC IP address of nodeA is not updated with the changed information in the configuration file for the shutdown agent stored in nodeB.

5. Restoring the information

Execute the following command on a node storing the correct information to restore the status.

In this case, execute the command on nodeA or nodeC because the information stored in nodeA and nodeC is correct.

```
# /opt/SMAW/SMAWsf/bin/sfsacfgupdate -r
```

6. Checking the status after restoration

Execute the following command on any node to check the current status.

In the following example, you can see the BMC IP address of nodeA is updated with the changed information in the configuration file for the shutdown agent stored in nodeB.

```
# /opt/SMAW/SMAWsf/bin/sfsacfgupdate -s
Node : nodeA
 Node status : UP
 Configuration Update Service status : ENABLE
 BMC IP Address :
  10.20.30.51
 Configuration file :
  nodeA 10.20.30.51
  nodeB 10.20.30.42
  nodeC 10.20.30.43

Node : nodeB
 Node status : UP
 Configuration Update Service status : ENABLE
 BMC IP Address :
  10.20.30.42
 Configuration file :
  nodeA 10.20.30.51 <- Updated with the changed information on nodeB
  nodeB 10.20.30.42
  nodeC 10.20.30.43

Node : nodeC
 Node status : UP
 Configuration Update Service status : ENABLE
 BMC IP Address :
  11.22.33.46
 Configuration file :
  nodeA 10.20.30.51
  nodeB 10.20.30.42
  nodeC 10.20.30.43
```

G.7　sfsacfgupdate

The following describes how to use the sfsacfgupdate command.

Name

sfsacfgupdate(8) - Management of Configuration Update Service for SA

Synopsis

/opt/SMAW/SMAWsf/bin/sfsacfgupdate {-e [*<levels>*]|-d|-c|-s|-r}

Feature description

This command manages Configuration Update Service for SA.

When Configuration Update Service for SA is activated, the configuration information of the shutdown agent is automatically updated on operating system startup. Execute this command with the system administrator authority.

Options

-e

    Activates Configuration Update Service for SA.

    Specify the value of run levels 2 to 5 which you want to activate for *<levels>*. You can specify several run levels.

    For example, when you specify "-e 35," run levels 3 and 5 will be activated.

    When you omit the value, all run levels from 2 to 5 will be activated.

-d

    Deactivates Configuration Update Service for SA.

-c

    Checks the setup status of Configuration Update Service for SA.

-s

    Displays the configuration information of the shutdown agent stored in all nodes.

-r

    Restores the configuration information of the shutdown agent.

Example

```
# /opt/SMAW/SMAWsf/bin/sfsacfgupdate -c [Return]
Configuration file exists.          [ OK ]
ipmitool command exists.            [ OK ]
ipmi service has been started.      [ OK ]
ipmi service's run level :
0:off   1:off   2:on    3:on    4:on    5:on    6:off
Configuration Update Service's run level :
0:off   1:off   2:on    3:on    4:on    5:on    6:off
#
```

Exit status

0 : Normal exit

Other than 0 : Abnormal exit

G.8 Output Message (Syslog)

The following shows the messages output to Syslog.

---

sfsacfgupdate: ERROR: *<command>* command failed. return_value=*<value>*.

Content:

<command> abnormally ended with the return value *<value>*.

Corrective action:

Copy this message, and then contact your Fujitsu systems engineer (SE).

---

sfsacfgupdate: ERROR: Could not find *<file>*.

Content:

<file> does not exist.

Corrective action:

Create *<file>*.

---

sfsacfgupdate: ERROR: Could not find ipmitool command.

Content:

The ipmitool command does not exit.

Corrective action:

Install the ipmitool command.

---

sfsacfgupdate: ERROR: ipmi service doesn't start.

Content:

The ipmi service does not start.

Corrective action:

Start the ipmi service.

---

sfsacfgupdate: ERROR: *<file>* is invalid.

Content:

Contents described in *<file>* include incorrect information.

Corrective action:

Check the contents in *<file>* and enter the correct information.

---

sfsacfgupdate: ERROR: Reading the Shutdown Agent configuration failed.

Content:

Reading the configuration file for the shutdown agent failed.

Corrective action:

Review the contents of the configuration file for the shutdown agent, and check if the correct information is entered.

---

sfsacfgupdate: ERROR: Failed to copy the backup of *<file>* on node *<node>*.

Content:

Copying the backup of *<file>* failed on *<node>*.

Corrective action:

Check that the communication with *<node>* is available. After restoring the state of *<node>*, execute this command with the -r option and restore the configuration information of the shutdown agent.

sfsacfgupdate: ERROR: *<file>* generation failed.

Content:

Creating a file failed.

Corrective action:

Copy this message, and then contact your Fujitsu systems engineer (SE).

sfsacfgupdate: ERROR: Failed to distribute *<file>* to node *<node>*.

Content:

Distributing *<file>* to *<node>* failed.

Corrective action:

Check that the communication with *<node>* is available. After restoring the state of *<node>*, execute this command with the -r option and restore the configuration information of the shutdown agent.

sfsacfgupdate: ERROR: Failed to change the access permission of *<file>* on node *<node>*.

Content:

Changing the mode of *<file>* failed on *<node>*.

Corrective action:

Check that the communication with *<node>* is available. After restoring the state of *<node>*, execute this command with the -r option and restore the configuration information of the shutdown agent.

sfsacfgupdate: ERROR: Failed to change the group of *<file>* on node *<node>*.

Content:

Changing the group of *<file>* failed on *<node>*.

Corrective action:

Check that the communication with *<node>* is available. After restoring the state of *<node>*, execute this command with the -r option and restore the configuration information of the shutdown agent.

sfsacfgupdate: ERROR: Failed to change the owner of *<file>* on node *<node>*.

Content:

Changing the owner of *<file>* failed on *<node>*.

Corrective action:

Check that the communication with *<node>* is available. After restoring the state of *<node>*, execute this command with the -r option and restore the configuration information of the shutdown agent.

Appendix H Using PRIMECLUSTER in a VMware Environment

This appendix explains how to use PRIMECLUSTER in a VMware environment.

H.1 Cluster Systems in a VMware Environment

When using PRIMECLUSTER in a VMware environment, clustering (virtual machine function) between guest OSes on multiple ESX hosts and clustering (virtual machine function) between guest OSes on a single ESX host are available.

When an error occurs on an ESX host or a guest OS within a VMware environment, applications on that guest OS will no longer work. In clustering between guest OSes on multiple ESX hosts, with PRIMECLUSTER applied to guest OSes, when an error occurs, applications will failover from the active guest OS to a standby guest OS in the event of a failure, which creates a highly reliable guest OS environment.

When using shared disks and a failover occurs with guest OSes not completely stopped (e.g. when the OS is hanging), I/O fencing can be used for putting the unstopped guest OSes into the panic state and thereby securely stop guest OSes. Also, it is possible to prevent access from both guest OSes to the shared disk in order to effect a safe and secure failover.

In clustering between guest OSes on a single ESX host, an automatic switchover occurs only for an application error on guest OSes.

Figure H.1 Cluster Systems in a VMware Environment
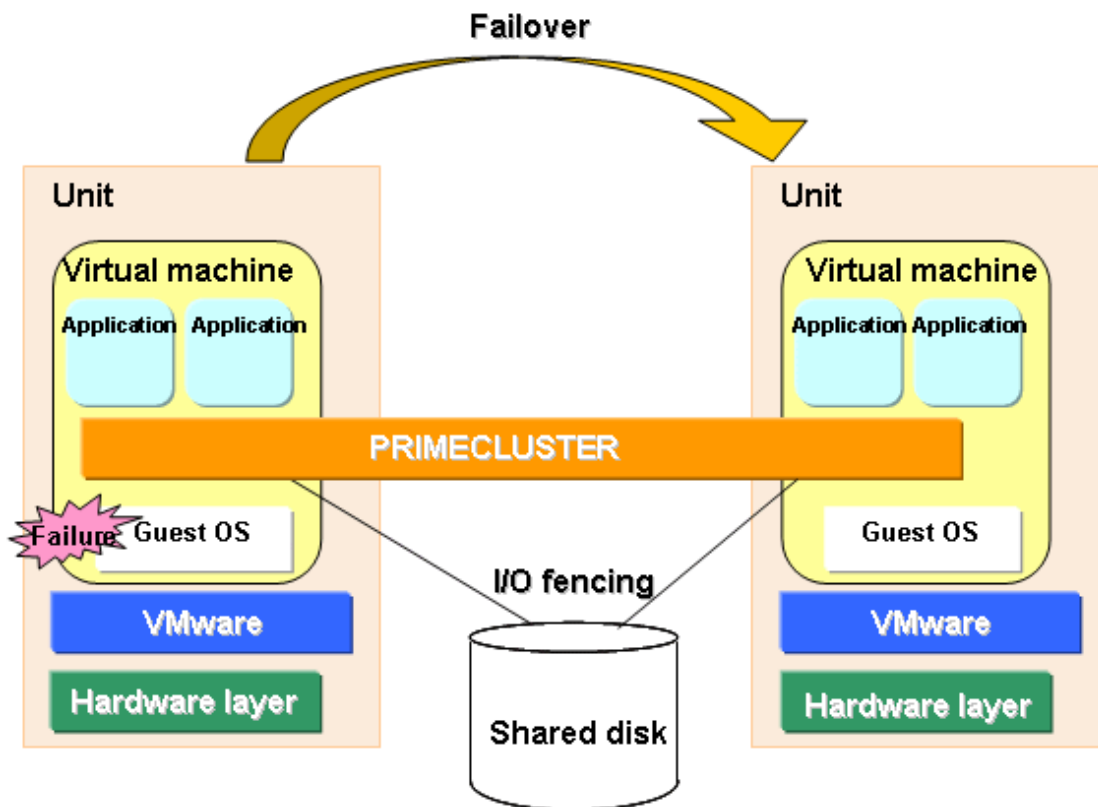
🗂 **See**
. . . .
For details on VMware, see the documentation for VMware.
. . . .

🛈 **Note**
. . . .

- Up to two nodes can be added to one cluster system.

- Since an automatic switchover does not occur when a guest OS error occurs in clustering between guest OSes on a single ESX host, manually switch the OS.

- Since cluster systems between guest OSes on multiple ESX hosts in a VMware environment check guest OS statuses via network paths (administrative LAN or interconnect) before effecting a failover, it may happen that guest OSes on which an error occurred are not completely stopped (e.g. when the OS is hanging). Therefore, when using shared disks, you should make sure to set up I/O fencing.

- I/O fencing must be set up at the earlier stage of configuring the cluster application.

- When using I/O fencing, the shared disk device should be managed by GDS.

- To use a switchover file system on the shared disk, you cannot specify the device of the file system to be mounted by defining names such as the label name or the udev functionality in the /etc/fstab.pcl file. Use the device name which can be specified by the entry beginning with /dev/sfdsk.

- If a failure in the cluster interconnect occurs, a forced shutdown of a node depending on the survival priority, and an automatic switchover are not made and each node enters the LEFTCLUSTER state. In this case, the system administrator needs to recover the LEFTCLUSTER state manually. Perform the procedure in "6.2 Recovering from LEFTCLUSTER" in the "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide.*"

- Remove a virtual machine that installs PRIMECLUSTER from targets for cluster functions (such as VMware HA, VMware FT, VMware DRS, and VMware DPM) of VMware.

- The following functions are not available in a virtual machine in which PRIMECLUSTER is to be installed.

    - VMware vMotion

    - VMware Storage vMotion

    - Migration with VMware vCenter Converter

    - Snapshot of VMware

    - Hot clone

    - Backup by Data Recovery

    - Backup by VCB

- Set the path policy for the Native Multipathing (NMP) as follows:

    - In the case of using VMware vSphere 4.x
      Set to "Most Recently Used".

    - In the case of using VMware vSphere 5.0 Update1 or later, or VMware vSphere 5.1 or later
      Set to "Most Recently Used" or "Round Robin".
      However, if you set to "Round Robin", the following patches must be applied in the Intel64 environment:

        - T005785LP-03 or later (RHEL5)

        - T005873LP-03 or later (RHEL6)

    Settings other than above are not supported.

- Support for multipath software from third parties, contact your Fujitsu systems engineer (SE).

- The following environments and functions are not supported:

    - iSCSI and FCoE

    - ESX hosts with different versions

    - N-Port ID Virtualization (NPIV)

- It is necessary to use the hardware version 7 for VMware vSphere 4.x of the virtual machine.

- When you cannot refer to a system volume on a disk device in a SAN boot configuration, the PRIMECLUSTER failure detection function cannot be operated depending on the status of the system. In this case, perform a manual switchover by an operator.

- Use a shared disk responding to SCSI-3 Persistent Reservation.

- When using the file system that is created on the shared disk as Fsystem resources, you need to register all the file systems that are created on the same disk (LUN) or on the same disk class to the same userApplication. Due to the restriction of I/O fencing, you cannot create multiple file systems on one disk (LUN) or on one disk class and register each file system to the different userApplications to monitor and control them.

....

H.2 Installation (between guest OSes on multiple ESX hosts)

This section describes procedures for installing PRIMECLUSTER between guest OSes on multiple ESX hosts in a VMware environment.

H.2.1 Software Installation

Install the software required for PRIMECLUSTER on each node.

The explanation is divided into the following topics:

- Installation and configuration of related software

- Installation and environment configuration of applications

H.2.1.1 Installation and Configuration of Related Software

After installing the software related to PRIMECLUSTER, you need to take it into operation and make various settings for the OS and the hardware.

Perform the following steps as necessary.
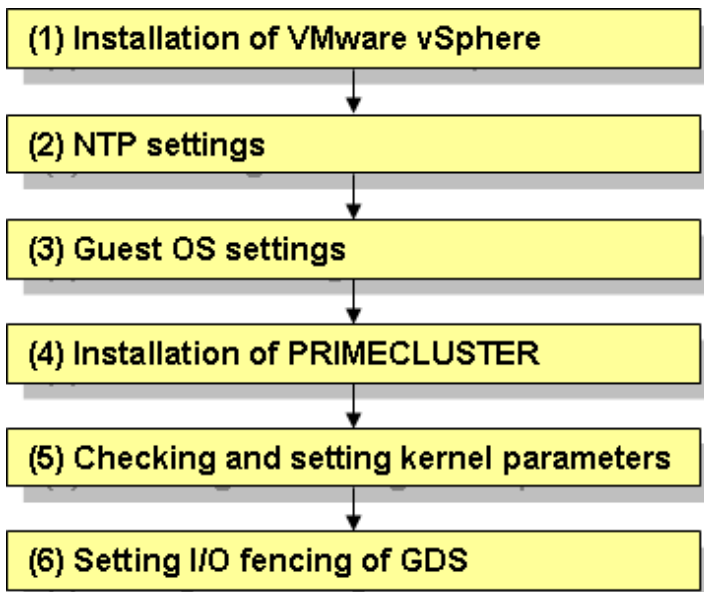
1. Installation of VMware vSphere

   Refer to the documentation for VMware vSphere to implement all setup work from installing VMware vSphere to installing the guest OS and configuring the environment.

   When creating a virtual machine, note the following points:

   - When you create a new virtual machine using the wizard executed from vSphere Client version 5, proceed to "Create Disk" and select Eager Zeroed.

   - When you create a new virtual machine using the wizard executed from vSphere Client version 4, proceed to "Create Disk." Select "Support clustering features such as Fault Tolerance" and create disks in eagerzeroedthick format.

   - For types of SCSI controllers, set to "LSI Logic Parallel" for the RHEL5 environment, and set to "LSI Logic Parallel" or "VMware Paravirtual" for the RHL6 environment.

- Setting up disks and related devices (shared disks)

    When creating the first virtual machine, create shared disks taken over in the cluster system with Raw DeviceMapping (RDM), and set the compatibility mode to "Physical."

    For virtual machines other than the first one, select "Use an existing virtual disk" and specify the first RDM disk you created. Select "Physical" as the compatibility mode as well.

    Create the virtual machine disk (.vmdk) file in the data store in the shared disk.

    For SCSI controllers, select a new virtual device node such as SCSI(1:0) and use the same one on each node in the cluster system.

    For types of SCSI controllers, set the same type as the system disk on a guest OS. For sharing SCSI buses, set to "Physical."

- Setting up the virtual network

    When creating the virtual machine, create at least two network systems for the cluster interconnect and connect them to different physical adapters.

    Also, for sharing a physical network adapter by multiple clusters, allocate a port group to each cluster system for a vSwitch.

    For VLAN ID, use different values for each port group.

2. NTP settings (Guest OS)

    These settings serve to synchronize the time of each node in the cluster system configuration. Be sure to make these settings when you configure the cluster.

    Make these settings on the guest OS before you install PRIMECLUSTER.

3. Guest OS settings (Guest OS)

    In the guest OS in the cluster system, it is necessary to make network settings such as IP addresses for the public LAN and the administrative LAN.

    Implement these settings on the guest OS that you are going to run as a cluster.

4. Installation of PRIMECLUSTER (Guest OS)

    For installing PRIMECLUSTER, an installation script (CLI Installer) is available.

    This script method installs PRIMECLUSTER node by node on systems that already have Linux(R) and related software installed. It is also utilized for installation on cluster management servers.

See

For details on the installation procedure, see the Installation Guide for PRIMECLUSTER.

5. Checking and setting the kernel parameters

    Depending on the environment, the kernel parameters must be modified.

    Applicable nodes:

        All nodes on which PRIMECLUSTER is to be installed

    Depending on the utilized products and components, different kernel parameters are required.

    Check the Kernel Parameter Worksheet and modify the settings as necessary.

See

For details on the kernel parameters, see "A.6 Kernel Parameter Worksheet."

6. Setting I/O fencing of GDS

    When a shared disk is registered to a GDS class, set up I/O fencing of GDS.

    Add the following line into the /etc/opt/FJSVsdx/sdx.cf file:

```
SDX_VM_IO_FENCE=on
```

Applicable nodes:

   All nodes on which PRIMECLUSTER is to be installed

7. Setting up the /etc/hostid file

   To run I/O fencing properly, you may need to set up the /etc/hostid file depending on the environment.

   According to the following steps, check whether setting up the /etc/hostid file is required, and then, set it up if needed.

   How to check
   Execute the hostid command and check the output.

   When the output is other than "00000000," setting up the /etc/hostid file is not necessary.

```
# /usr/bin/hostid
a8c00101
```

   When the output is "00000000," follow the setting procedure below so that outputs of the hostid will be different on all nodes that configure the cluster.

   Setting procedure

   1. Create the /etc/hostid file.

```
# touch /etc/hostid
```

   2. Create the following python script file.
      [Contents of the file to be created]

```
#!/usr/bin/python
from struct import pack
filename = "/etc/hostid"
hostid = pack("I",int("0x<hhhhhhhh>",16))
open(filename, "wb").write(hostid)
```

      (hhhhhhhh: Describe the intended host identifier in base 16, 8 digit numbers)

   3. Set the execute permissions to the created script file and then, execute it.

```
# chmod +x <created script file name>
# ./<created script file name>
```

   4. Execute the hostid command to check if the specified host identifier is obtained.

```
# hostid
hhhhhhhh
```

      (hhhhhhhh: host identifier that is specified in the script file)

Note

To activate the modified kernel parameters and I/O fencing of GDS, restart the guest OS after installation settings for related software is complete.

H.2.1.2 Installation and Environment Configuration of Applications

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

See

- For details on environment setup, see manuals for each application.

- For information on PRIMECLUSTER-related products supporting VMware, see the documentation for each product.

H.2.2 Preparation Prior to Building a Cluster

Refer to "Chapter 4 Preparation Prior to Building a Cluster" to make the initial cluster setup on the guest OS.

H.2.3 Building a Cluster

This section describes procedures for setting up a cluster with PRIMECLUSTER in a VMware environment.

(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 Fault Resource Identification and Operator Intervention Request

H.2.3.1 Initial Setup of CF and CIP

Refer to "5.1.1 Setting Up CF and CIP" to set up CF and CIP on the guest OS.

H.2.3.2 Setting Up the Shutdown Facility

For details on survival priority, see "5.1.2.2 Survival priority."

In VMware environments, only the SA_icmp shutdown agent is available for setup.

This section explains the method for setting up the SA_icmp shutdown agent as the shutdown facility.

📝 Note

Be sure to perform the following operations on all guest OSes (nodes).

1. Setting up the shutdown facility

   Specify the shutdown agent.

   Create /etc/opt/SMAW/SMAWsf/SA_icmp.cfg with the following contents on all guest OSes (nodes) of the cluster:

   ```
   TIME_OUT=value
   cfname:ip-address-of-node:NIC-name1,NIC-name2
   ```

   value

   Specify the interval (in seconds) for checking whether the node is alive. The recommended value is "5" (s).

   cfname

   Specify the name of the CF node.

   ip-address-of-node

   Specify one of the IP addresses below for cfname.

   - Administrative LAN (Required)

   - Public LAN

- Cluster interconnect (IP address of CIP)

Enter at least one for every node in the cluster system. Add a line each for specifying multiple LAN paths. To check LAN paths, we recommend that you use multiple ones.

NIC-nameX

Specify the network interface of the local guest OS (node) utilized for checking whether the node defined by ip-address-of-node is alive. If there is more than one, delimit them with commas (",").

Note

Registering network interfaces

- For duplicating by GLS, define all redundant network interfaces. (Example: eth0,eth1)

- If you are bonding NICs, define the bonding device behind the IP address. (Example: bond0)

- For registering the cluster interconnect, define all network interfaces that are used on all paths of the cluster interconnect. (Example: eth2,eth3)

- Do not use the takeover IP address (takeover virtual Interface).

Example

In a cluster between guest OSes on multiple ESX hosts, the following example shows settings for the public LAN (duplicated by GLS (eth0,eth1)) and the administrative LAN (eth2) in a two-node configuration.

```
TIME_OUT=5
node1:10.20.30.100:eth0,eth1
node1:10.20.40.200:eth2
node2:10.20.30.101:eth0,eth1
node2:10.20.40.201:eth2
```

2. Setting up the shutdown daemon

Create /etc/opt/SMAW/SMAWsf/rcsd.cfg with the following contents on all guest OSes (nodes) of the cluster:

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

```
CFNameX         : CF node name of the cluster host.
weight          : Weight of the SF node.
myadmIP         : Specify the IP address of the administrative LAN for CFNameX
timeout         : Set the value determined by below calculation.
                  A) From the LAN path information provided in SA_icmp.cfg, extract lines
                     from each CF node.
                  B) Determine the times required for checking LAN path for each CF node.
                     ((TIME_OUT value in SA_icmp.cfg) + 2 [s]) x (number of lines extracted
                        in step A))
                  C) Perform steps A) and B) for all CF nodes described in SA_icmp.cfg,
                     and set the largest value as the timeout value.
                     However, if the value determined in step C) is smaller than 20 seconds,
                     set the timeout value to "20."
```

Note

The rcsd.cfg file must be the same on all guest OSes (nodes). Otherwise, operation errors might occur.

3. Starting the shutdown facility

Check that the shutdown facility has started.

```
# sdtool -s
```

If the shutdown facility has already started, execute the following command to restart the shutdown facility.

```
# sdtool -r
```

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

```
# sdtool -b
```

4. Checking the status of the shutdown facility

    Check that the status of the shutdown facility is either "InitWorked" or "TestWorked." If the displayed status is "TestFailed" or "InitFailed," check the shutdown daemon settings for any mistakes.

```
# sdtool -s
```

H.2.3.3 Initial Setup of the Cluster Resource Management Facility

Refer to "5.1.3 Initial Setup of the Cluster Resource Management Facility" to set up the resource database managed by the cluster resource management facility (hereafter referred to as "CRM") on the guest OS.

H.2.3.4 Setting Up Fault Resource Identification and Operator Intervention Request

Refer to "5.2 Setting Up Fault Resource Identification and Operator Intervention Request" to make the settings for identifying fault resources and for requesting operator intervention.

H.2.4 Building Cluster Applications

For details on how to configure cluster applications, see "Chapter 6 Building Cluster Applications."

When configuring cluster applications on guest OSes, you must set up I/O fencing for all cluster applications to which Gds resources or Fsystem resources corresponding to the shared disk are registered.

H.2.4.1 Setting Up I/O Fencing

This section explains how to set up I/O fencing for all cluster applications to which Gds resources or Fsystem resources corresponding to the shared disk are registered.

Perform the following to set up I/O fencing:

- Creating Cmdline resources and setting up Online and Offline scripts

- Setting up userApplication scripts (PreOnline, OfflineDone, and Fault scripts)

1. Creating Cmdline resources and setting up Online and Offline scripts

    1. In the Cmdline resource settings, add the Start script, the Stop script, and the Check script in the following format:

       <Start script>

       /opt/SMAW/bin/hvsgpr -c

       <Stop script>

       /opt/SMAW/bin/hvsgpr -u

       <Check script>

       /opt/SMAW/bin/hvsgpr -m

       To create Cmdline resources, see, "6.7.3.1 Setting Up Cmdline Resources."

    2. In the attribute settings of the Cmdline resources, set the AutoRecover attribute to disabled ("0"). Do not change the default settings for other attributes.

2. Setting up userApplication scripts

1. In the cluster application settings, add the PreOnline and OfflineDone scripts in the following format.

<PreOnline script>

/opt/SMAW/bin/hvsgpr -r

<OfflineDone script>

/opt/SMAW/bin/hvsgpr -o

1. Select "(PreOnlineScript=)" of "Machines+Basics."

```
Consistency check ...

Machines+Basics (app1:consistent)
 1) HELP
 2) -
 3) SAVE+EXIT
 4) REMOVE+EXIT
 5) AdditionalMachine
 6) AdditionalConsole
 7) Machines[0]=vm21RMS
 8) Machines[1]=vm22RMS
 9) (PreCheckScript=)
10) (PreOnlineScript=)
11) (PostOnlineScript=)
12) (PreOfflineScript=)
13) (OfflineDoneScript=)
14) (FaultScript=)
15) (AutoStartUp=yes)
16) (AutoSwitchOver=HostFailure|ResourceFailure|ShutDown)
17) (PreserveState=no)
18) (PersistentFault=0)
19) (ShutdownPriority=)
20) (OnlinePriority=)
21) (StandbyTransitions=ClearFaultRequest|StartUp|SwitchRequest)
22) (LicenseToKill=no)
23) (AutoBreak=yes)
24) (AutoBreakMaintMode=no)
25) (HaltFlag=yes)
26) (PartialCluster=0)
27) (ScriptTimeout=)
Choose the setting to process:10
```

2. Select "FREECHOICE" and enter the full path of the PreOnline script.

```
1) HELP
2) RETURN
3) NONE
4) FREECHOICE
Enter the command line to start prior to the application becoming ONLINE:4
       >> /opt/SMAW/bin/hvsgpr -r
```

3. Select "(OfflineDoneScript=)" of "Machines+Basics."

```
Consistency check ...

Machines+Basics (app1:consistent)
 1) HELP
 2) -
 3) SAVE+EXIT
 4) REMOVE+EXIT
 5) AdditionalMachine
 6) AdditionalConsole
```

```
 7) Machines[0]=vm21RMS
 8) Machines[1]=vm22RMS
 9) (PreCheckScript=)
10) (PreOnlineScript='/opt/SMAW/bin/hvsgpr~-r')
11) (PostOnlineScript=)
12) (PreOfflineScript=)
13) (OfflineDoneScript=)
14) (FaultScript=)
15) (AutoStartUp=yes)
16) (AutoSwitchOver=HostFailure|ResourceFailure|ShutDown)
17) (PreserveState=no)
18) (PersistentFault=0)
19) (ShutdownPriority=)
20) (OnlinePriority=)
21) (StandbyTransitions=ClearFaultRequest|StartUp|SwitchRequest)
22) (LicenseToKill=no)
23) (AutoBreak=yes)
24) (AutoBreakMaintMode=no)
25) (HaltFlag=yes)
26) (PartialCluster=0)
27) (ScriptTimeout=)
Choose the setting to process:13
```

4. Select "FREECHOICE" and enter the full path of the OfflineDone script.

```
1) HELP
2) RETURN
3) NONE
4) FREECHOICE
Enter the command line to start prior to the application becoming ONLINE:4
        >> /opt/SMAW/bin/hvsgpr -o
```

2. In the attribute settings of the cluster application, if the HaltFlag attribute is set to enabled ("1"), add the Fault script in the following format.

<Fault script>

/opt/SMAW/bin/hvsgpr -f

1. Select "(FaultScript=)" of "Machines+Basics."

```
Consistency check ...

Machines+Basics (app1:consistent)
 1) HELP
 2) -
 3) SAVE+EXIT
 4) REMOVE+EXIT
 5) AdditionalMachine
 6) AdditionalConsole
 7) Machines[0]=vm21RMS
 8) Machines[1]=vm22RMS
 9) (PreCheckScript=)
10) (PreOnlineScript='/opt/SMAW/bin/hvsgpr~-r')
11) (PostOnlineScript=)
12) (PreOfflineScript=)
13) (OfflineDoneScript='/opt/SMAW/bin/hvsgpr~-o')
14) (FaultScript=)
15) (AutoStartUp=yes)
16) (AutoSwitchOver=HostFailure|ResourceFailure|ShutDown)
17) (PreserveState=no)
18) (PersistentFault=0)
19) (ShutdownPriority=)
20) (OnlinePriority=)
```

```
21) (StandbyTransitions=ClearFaultRequest|StartUp|SwitchRequest)
22) (LicenseToKill=no)
23) (AutoBreak=yes)
24) (AutoBreakMaintMode=no)
25) (HaltFlag=yes)
26) (PartialCluster=0)
27) (ScriptTimeout=)
Choose the setting to process:14
```

2. Select "FREECHOICE" and enter the full path of the Fault script.

```
1) HELP
2) RETURN
3) NONE
4) FREECHOICE
Enter the command line to start prior to the application becoming ONLINE:4
        >> /opt/SMAW/bin/hvsgpr -f
```

H.3 Installation (between guest OSes on a single ESX host)

This section describes procedures for installing PRIMECLUSTER between guest OSes on a single ESX host in a VMware environment.

H.3.1 Software Installation

Install the software required for PRIMECLUSTER on each node.

The explanation is divided into the following topics:

- Installation and configuration of related software

- Installation and environment configuration of applications

H.3.1.1 Installation and Configuration of Related Software

After installing the software related to PRIMECLUSTER, you need to take it into operation and make various settings for the OS and the hardware.

Perform the following steps as necessary
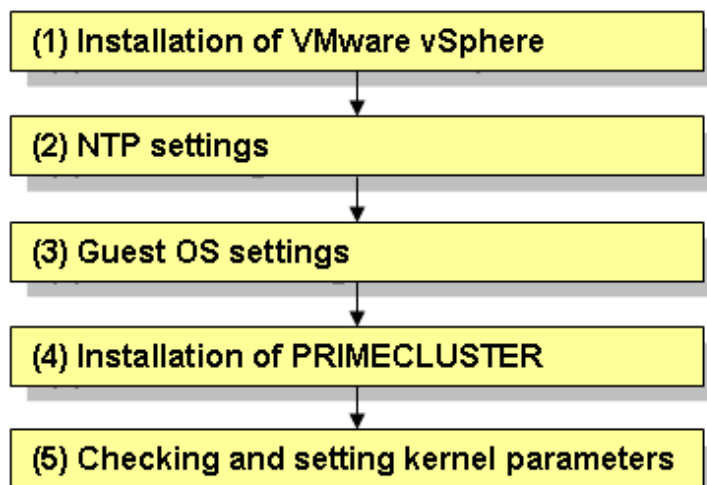
1. Installation of VMware vSphere

   Refer to the documentation for VMware vSphere to implement all setup work from installing VMware vSphere to installing the guest OS and configuring the environment.

- Setting up disks and related devices

    - When creating the virtual machine, create shared disks taken over in the cluster system with Raw DeviceMapping (RDM), and set the compatibility mode to "Physical."

    - For types of SCSI controllers, set the same type as the system disk on a guest OS.

    - When you create a new virtual machine using the wizard executed from vSphere Client version 5, proceed to "Create Disk" and select Eager Zeroed.

    - When you create a new virtual machine using the wizard executed from vSphere Client version 4, proceed to "Create Disk." Select "Support clustering features such as Fault Tolerance" and create disks in eagerzeroedthick format.

- Setting up the SCSI controller

    - Set to "Virtual" for sharing of the SCSI bus on the property dialog box of the virtual machine.

    - For types of SCSI controllers, set to "LSI Logic Parallel" for the RHEL5 environment, and set to "LSI Logic Parallel" or "VMware Paravirtual" for the RHL6 environment.

- Setting up the virtual network

    When creating the virtual machine, create at least two network systems for the cluster interconnect.

2. NTP settings (Guest OS)

    These settings serve to synchronize the time of each node in the cluster system configuration. Be sure to make these settings when you configure the cluster.

    Make these settings on the guest OS before you install PRIMECLUSTER.

3. Guest OS settings (Guest OS)

    In the guest OS in the cluster system, it is necessary to make network settings such as IP addresses for the public LAN and the administrative LAN.

    Implement these settings on the guest OS that you are going to run as a cluster.

4. Installation of PRIMECLUSTER (Guest OS)

    For installing PRIMECLUSTER, an installation script (CLI Installer) is available.

    This script method installs PRIMECLUSTER node by node on systems that already have Linux(R) and related software installed. It is also utilized for installation on cluster management servers.

    ### See

    ....

    For details on the installation procedure, see the Installation Guide for PRIMECLUSTER.

    ....

5. Checking and setting the kernel parameters

    Depending on the environment, the kernel parameters must be modified.

    Applicable nodes:

    All nodes on which PRIMECLUSTER is to be installed

    Depending on the utilized products and components, different kernel parameters are required.

    Check the Kernel Parameter Worksheet and modify the settings as necessary.

    ### See

    ....

    For details on the kernel parameters, see "A.6 Kernel Parameter Worksheet."

    ....

Note

To activate the modified kernel parameters, restart the guest OS after installation and settings for related software are complete.

H.3.1.2 Installation and Environment Configuration of Applications

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

See

- For details on environment setup, see manuals for each application.

- For information on PRIMECLUSTER-related products supporting VMware, see the documentation for each product.

H.3.2 Preparation Prior to Building a Cluster

Refer to "Chapter 4 Preparation Prior to Building a Cluster" to make the initial cluster setup on the guest OS.

H.3.3 Building a Cluster

For procedures for setting up a cluster with PRIMECLUSTER between guest OSes on a single ESX host, see " Chapter 5 Building a Cluster."

Note

Do not configure the shutdown facility between guest OSes on a single ESX host."

H.3.4 Building Cluster Applications

For details on how to configure cluster applications between guest OSes on a single ESX host, see "Chapter 6 Building Cluster Applications."

H.4 Operations

For details on functions for managing PRIMECLUSTER system operations, see "Chapter 7 Operations."

Note

When you configure the cluster system between guest OSes on a single ESX host, an automatic switchover occurs only for an application error on guest OSes.

Since an automatic switchover does not occur when a guest OS error occurs, stop the guest OS in which an error occurs and manually switch the cluster application by using hvswitch.

H.5 Changing the Configuration

For details on additions or changes of configuration nodes in the PRIMECLUSTER system and cluster applications, see "Chapter 8 Changing the Operation Configuration."

H.6 Maintenance

For details on items and procedures required for maintenance of the PRIMECLUSTER system, see "Chapter 9 Maintenance of the PRIMECLUSTER System."

H.7   Sample Configurations

H.7.1   Configuring a Cluster System Between Guest OSes on Two ESX Hosts

Figure H.2 Sample configuration of a cluster between guest OSes on two ESX hosts
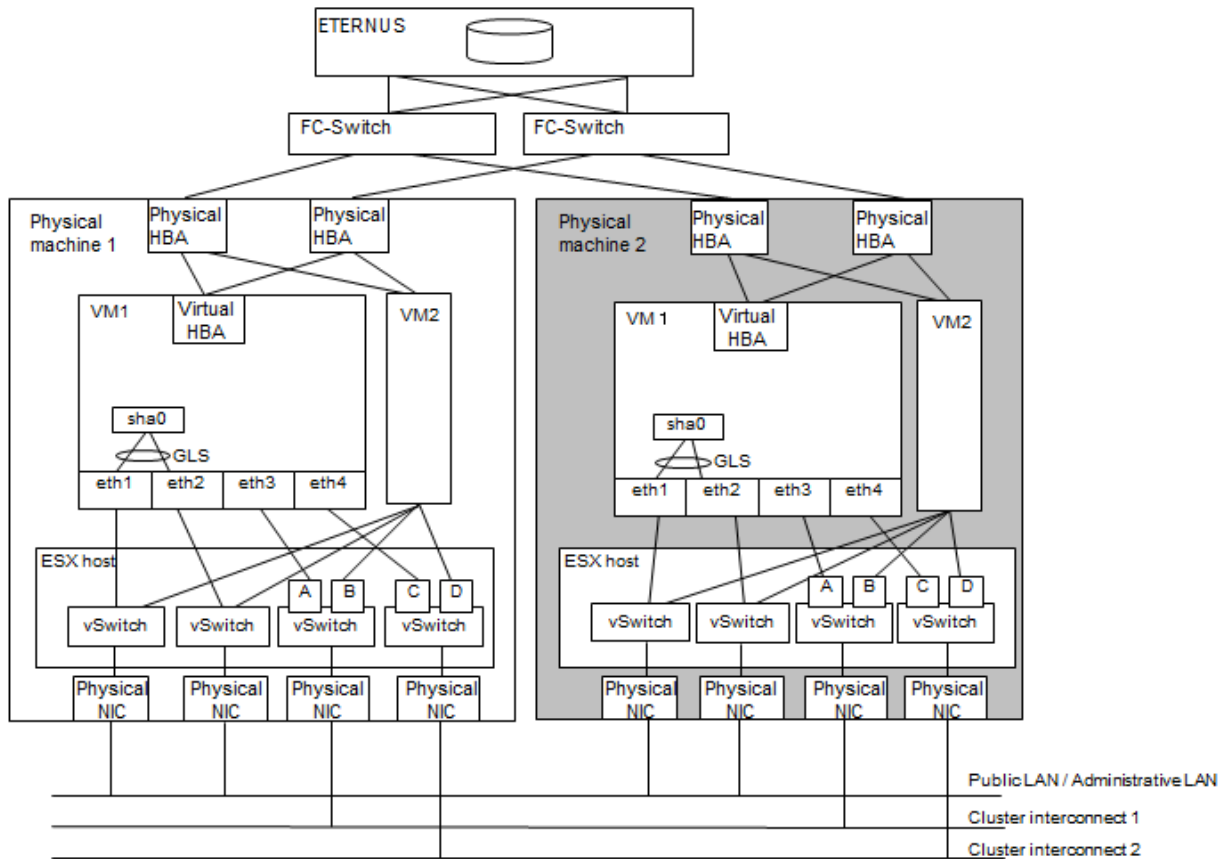
Note

- Perform time synchronization with the guest OS set as an NTP client.

- For sharing a vSwitch by multiple clusters, allocate a port group to each cluster system via VLAN.

Used hardware

- RX300 S6 x 2

- FC-Switch x 2

- LAN card x 8

- Fibre channel card x 4

- Disk array unit x 1

Used software

- Red Hat Enterprise Linux

- PRIMECLUSTER 4.3

- VMware vSphere

- Microsoft Internet Explorer 8 (installed on the operation management personal computer)

- J2SE (TM) Runtime Environment 6 (installed on the operation management personal computer)

H.7.1.1 Cluster Configuration Worksheet

Shown below is an example of the Cluster Configuration Worksheet.

Item				Setting			Remarks
Product type				<Y>PRIMECLUSTER Enterprise Edition			
				<>PRIMECLUSTER HA Server			
Shared disk unit: Type 1				ETERNUS3000			
	Interface			Fibrechannel			
	Path			mutilpath			
	File system selection			ext3			
Shared disk unit: Type 2				Expansion file unit			
	Interface			Fibrechannel			
	Path			mutilpath			
	File system selection			ext3			
NTP	Operation mode						
	Protocol (only when broadcast is specified)						
	NTP Network configuration	External NTP servers	Host name	host01,host02			
			IP address	10.20.40.50,10.20.40.51			
		NTP server settings in cluster	Primary NTP server	Host name			
				IP address			
			Secondary NTP server	Host name			
				IP address			
Operation management topology	Web-Based Admin View operation mode			2		layers	
	Public LAN						
	Primary management server IP			10.20.40.40			
	Secondary management server IP			10.20.40.41			
User group				Group name		User name	
				<Y>wvroot		root	
				<Y>clroot		root	
				<>cladmin			
				<>clmon			
Cluster name				PRIMECLUSTER1			
Node information	CIP	Number of subnets		1			
		Subnet IP		192.168.1.0			
		Net mask		255.255.255.0			
	Node 1	Node name (uname -n)		primecl01			
		CF node name		primecl01			
		Public LAN	Network interface name	eth1,eth2			
			IP address	10.20.30.40,10.20.30.44			
			Net mask	255.255.255.0			
			Gateway	-			
		Administrative LAN	Network interface name	-			
			IP address	-			
			Net mask	-			
			Gateway	-			
		Cluster interconnect	Path 0 NIC	eth3			
			Path 1 NIC	eth4			
		weight		1			Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area		/dev/sdd1 /var/crash			
		Node in CF quorum set		<Y>[Yes] <>No			Do not change the default setting.
	Node 2	Node name (uname -n)		primecl02			
		CF node name		primecl02			
		Public LAN	Network interface name	eth1,eth2			
			IP address	10.20.30.41,10.20.30.45			
			Net mask	255.255.255.0			
			Gateway	-			
		Administrative LAN	Network interface name	-			
			IP address	-			
			Net mask	-			
			Gateway	-			
		Cluster interconnect	Path 0 NIC	eth3			
			Path 1 NIC	eth4			
		weight		2			Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area		/dev/sdd1 /var/crash			
		Node in CF quorum set		<Y>[Yes] <>No			Do not change the default setting.
	CF remote service usage			[ ]cfcp (CF file copy)			
				[ ]cfsh (CF remote command execution)			
	IP interconnect setup			<Y>[Not used]			
				<>Used (1P interconnect count)			
	CIP			<Y>[Used by RMS]			
				<>Not used by RMS			
				Suffix:			
Cluster resource management facility	Hardware device to be registered in the resource database			[ ]Network device *1			*1) Required when public LAN takeover is used
				[ ]Disk units *2			*2) Required when GDS and GFS are used
Shutdown facility (SF)	Means of stopping a node forcibly			Means			
				[ ]icmp	value (Timeout)	5	Virtual machine function used in a VMware environment

H.7.1.2 GLS Setup Worksheet

Shown below is an example of the GLS Setup Worksheet.

Worksheet for NIC switching mode (virtual IP takeover)

Item				Setting
GLS settings	Switching mode			\<Y\>NIC switching mode (virtual IP takeover) \< \>NIC switching mode (physical IP takeover I) \< \>NIC switching mode (physical IP takeover II) \< \>Fast switching mode
	Takeover virtual interface name			sha1:65
		Takeover virtual IP address (or host name)		10.20.30.46
		Subnet mask		255.255.255.0
	Node name (1)			primecl01
		Configuration information	Virtual interface name	sha1
			Primary physical interface name	eth3
			Secondary physical interface name	eth4
			Physical IP address (or host name)	10.20.30.40
			Logical IP address (or host name)	10.20.30.46
		Monitoring destination	IP address (or host name) of primary	10.20.30.61
			IP address (or host name) of secondary	10.20.30.62
		Optional function	Standby NIC patrol	\<\>Disable \<Y\>Enable
	Node name (2)			primecl02
		Configuration information	Virtual interface name	sha1
			Primary physical interface name	eth3
			Secondary physical interface name	eth4
			Physical IP address (or host name)	10.20.30.41
			Logical IP address (or host name)	10.20.30.46
		Monitoring destination	IP address (or host name) of primary	10.20.30.61
			IP address (or host name) of secondary	10.20.30.62
		Optional function	Standby NIC patrol	\<\>Disable \<Y\>Enable

H.7.1.3 GDS Configuration Worksheet

Shown below is an example of the GDS Configuration Worksheet.

	Item			Setting
Class 1	Class name			class0005
	Class scope (node name)	Node 1		primecl01
		Node 2		primecl02
	Disk 1 connected to group	SDX disk name		Disk1
		Physical disk name in node 1		sdc
		Physical disk name in node 2		sdc
	Disk 2 connected to group	SDX disk name		Disk2
		Physical disk name in node 1		sdd
		Physical disk name in node 2		sdd
	Highest-order group 1	Group name		group0005
		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	volume0005
			Size	1024 blocks
Class 2	Class name			class0006
	Class scope (node name)	Node 1		primecl11
		Node 2		primecl12
	Disk 1 connected to group	SDX disk name		Disk3
		Physical disk name in node 1		sde
		Physical disk name in node 2		sde
	Disk 2 connected to group	SDX disk name		Disk4
		Physical disk name in node 1		sdf
		Physical disk name in node 2		sdf
	Highest-order group 1	Group name		group0006
		Group type		mirror
		Stripe width		-
		Disk/low-order group name	Disk/low-order group 1	Disk3
			Disk/low-order group 2	Disk4
		Volume 1	Volume name	volume0006
			Size	1024 blocks

H.7.2 Configuring a Cluster System Between Guest OSes on One ESX Host

Figure H.3 Sample configuration of a cluster between guest OSes on one ESX host
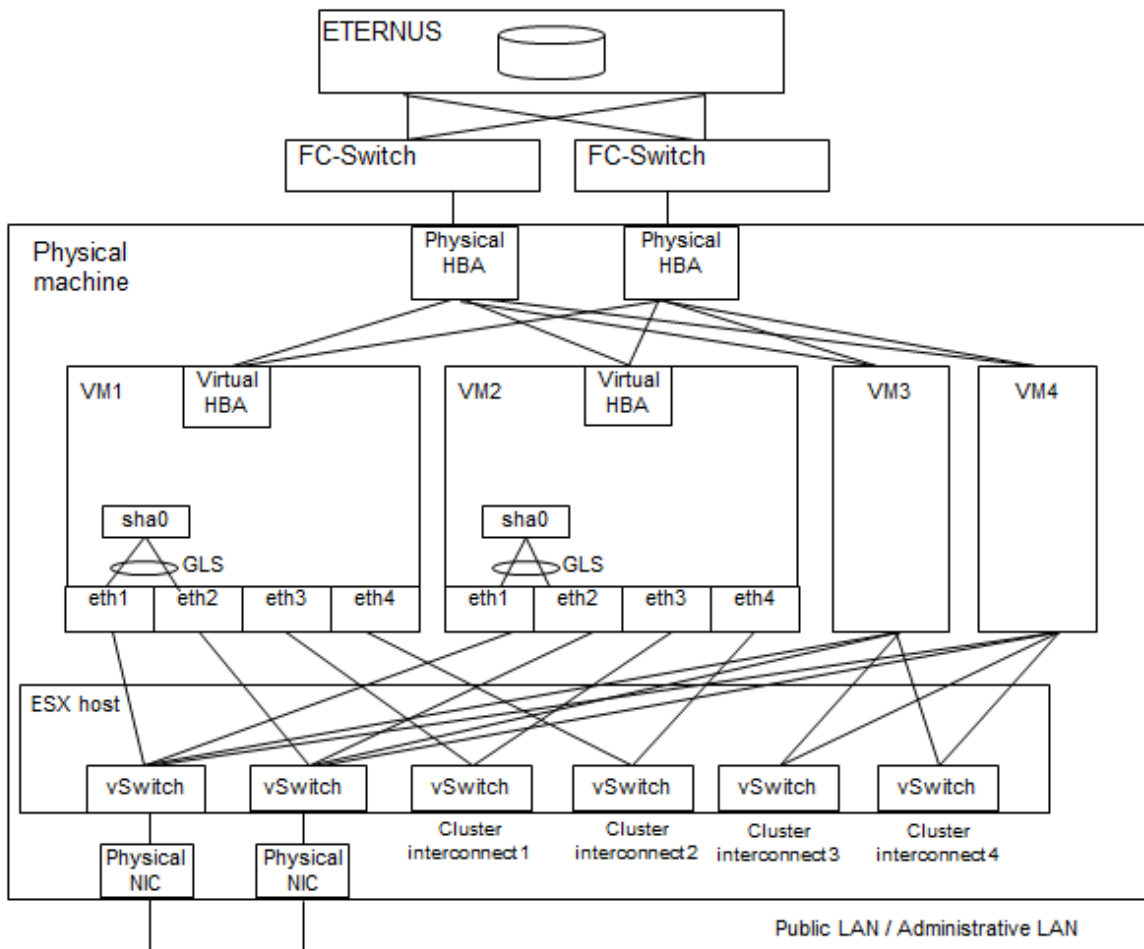
> ![Note]
> ....
>
> - Perform time synchronization with the guest OS set as an NTP client.
>
> - Create at least two network systems for the cluster interconnect.
> ....

Used hardware

- RX300 S6 x 1

- FC-Switch x 2

- LAN card x 2

- Fibre channel card x 2

- Disk array unit x 1

Used software

- Red Hat Enterprise Linux

- PRIMECLUSTER 4.3

- VMware vSphere

- Microsoft Internet Explorer 8 (installed on the operation management personal computer)

- J2SE(TM) Runtime Environment 6 (installed on the operation management personal computer)

H.7.2.1 Cluster Configuration Worksheet

Shown below is an example of the Cluster Configuration Worksheet.

Item				Setting			Remarks
Product type				<Y>PRIMECLUSTER Enterprise Edition			
				<>PRIMECLUSTER HA Server			
Shared disk unit: Type 1				ETERNUS3000			
	Interface			Fibrechannel			
	Path			mutilpath			
	File system selection			ext3			
Shared disk unit: Type 2				Expansion file unit			
	Interface			Fibrechannel			
	Path			mutilpath			
	File system selection			ext3			
NTP	Operation mode						
	Protocol (only when broadcast is specified)						
	NTP Network configuration	External NTP servers	Host name	host01,host02			
			IP address	10.20.40.50,10.20.40.51			
		NTP server settings in cluster	Primary NTP server	Host name			
				IP address			
			Secondary NTP server	Host name			
				IP address			
Operation management topology	Web-Based Admin View operation mode			2		layers	
	Public LAN						
	Primary management server IP			10.20.40.40			
	Secondary management server IP			10.20.40.41			
User group				Group name		User name	
				<Y>wvroot		root	
				<Y>clroot		root	
				<>cladmin			
				<>clmon			
Cluster name				PRIMECLUSTER1			
Node information	CIP	Number of subnets		1			
		Subnet IP		192.168.1.0			
		Net mask		255.255.255.0			
	Node 1	Node name (uname -n)		primecl01			
		CF node name		primecl01			
		Public LAN	Network interface name	eth1,eth2			
			IP address	10.20.30.40,10.20.30.44			
			Net mask	255.255.255.0			
			Gateway	-			
		Administrative LAN	Network interface name	-			
			IP address	-			
			Net mask	-			
			Gateway	-			
		Cluster interconnect	Path 0 NIC	eth3			
			Path 1 NIC	eth4			
		weight		1			Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area		/dev/sdd1 /var/crash			
		Node in CF quorum set		<Y>[Yes]　　<>No			Do not change the default setting.
	Node 2	Node name (uname -n)		primecl02			
		CF node name		primecl02			
		Public LAN	Network interface name	eth1,eth2			
			IP address	10.20.30.41,10.20.30.45			
			Net mask	255.255.255.0			
			Gateway	-			
		Administrative LAN	Network interface name	-			
			IP address	-			
			Net mask	-			
			Gateway	-			
		Cluster interconnect	Path 0 NIC	eth3			
			Path 1 NIC	eth4			
		weight		2			Node weight (priority if cluster partition occurs) used by SF
		Dump device for kdump and dump save area		/dev/sdd1 /var/crash			
		Node in CF quorum set		<Y>[Yes]　　<>No			Do not change the default setting.
	CF remote service usage			[ ]cfcp (CF file copy)			
				[ ]cfsh (CF remote command execution)			
	IP interconnect setup			<Y>[Not used]			
				<>Used　　　　　　　(1P interconnect count)			
	CIP			<Y>[Used by RMS]			
				<>Not used by RMS			
				Suffix:			
Cluster resource management facility	Hardware device to be registered in the resource database			[ ]Network device　*1			*1) Required when public LAN takeover is used
				[ ]Disk units　*2			*2) Required when GDS and GFS are used
Shutdown facility (SF)	Means of stopping a node forcibly			Means			
				[ ]icmp	value (Timeout)	5	Virtual machine function used in a VMware environment

H.7.2.2 GLS Setup Worksheet

Shown below is an example of the GLS Setup Worksheet.

Worksheet for NIC switching mode (virtual IP takeover)

Item			Setting
GLS settings	Switching mode		<Y>NIC switching mode (virtual IP takeover) < >NIC switching mode (physical IP takeover I) < >NIC switching mode (physical IP takeover II) < >Fast switching mode
	Takeover virtual interface name		sha1:65
		Takeover virtual IP address (or host name)	10.20.30.46
		Subnet mask	255.255.255.0
	Node name (1)		primecl01
	Configuration information	Virtual interface name	sha1
		Primary physical interface name	eth3
		Secondary physical interface name	eth4
		Physical IP address (or host name)	10.20.30.40
		Logical IP address (or host name)	10.20.30.46
	Monitoring destination	IP address (or host name) of primary	10.20.30.61
		IP address (or host name) of secondary	10.20.30.62
	Optional function	Standby NIC patrol	<>Disable <Y>Enable
	Node name (2)		primecl02
	Configuration information	Virtual interface name	sha1
		Primary physical interface name	eth3
		Secondary physical interface name	eth4
		Physical IP address (or host name)	10.20.30.41
		Logical IP address (or host name)	10.20.30.46
	Monitoring destination	IP address (or host name) of primary	10.20.30.61
		IP address (or host name) of secondary	10.20.30.62
	Optional function	Standby NIC patrol	<>Disable <Y>Enable

H.7.2.3 GDS Configuration Worksheet

Shown below is an example of the GDS Configuration Worksheet.

	Item			Setting
Class 1	Class name			class0005
	Class scope (node name)	Node 1		primecl01
		Node 2		primecl02
	Disk 1 connected to group	SDX disk name		Disk1
		Physical disk name in node 1		sdc
		Physical disk name in node 2		sdc
	Disk 2 connected to group	SDX disk name		Disk2
		Physical disk name in node 1		sdd
		Physical disk name in node 2		sdd
	Highest-order group 1	Group name		group0005
		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	volume0005
			Size	1024 blocks
Class 2	Class name			class0006
	Class scope (node name)	Node 1		primecl11
		Node 2		primecl12
	Disk 1 connected to group	SDX disk name		Disk3
		Physical disk name in node 1		sde
		Physical disk name in node 2		sde
	Disk 2 connected to group	SDX disk name		Disk4
		Physical disk name in node 1		sdf
		Physical disk name in node 2		sdf
	Highest-order group 1	Group name		group0006
		Group type		mirror
		Stripe width		-
		Disk/low-order group name	Disk/low-order group 1	Disk3
			Disk/low-order group 2	Disk4
		Volume 1	Volume name	volume0006
			Size	1024 blocks

Appendix I Startup Scripts and Startup Daemons, and Port Numbers in PRIMECLUSTER

This appendix provides explanations on scripts and daemons that are started by PRIMECLUSTER, and the port numbers being used.

I.1   Explanation Formats

Scripts and daemons, and port numbers are explained with the following formats:

Startup script

Name of startup script.

Function

> Function of startup script and daemon.

Effect if stopped

> Effect if startup script and daemon are stopped.

Startup daemon

> Daemon started by startup script.

Utilized port

Port

> Port number.

Protocol

> Protocol - TCP or UDP.

Send/Receive

> "s" if port sends data, "r" if it receives data, "s, r" for both.

Network

> Utilized network - any of Cluster interconnect, administrative LAN, or public LAN.

Target

> Node that uses the port.

Communication target

Port

> Port number of communication target.

Target

> Node or device that uses the port of the communication target.

Remarks

> Remarks

I.2   Startup Script List

init

init (inittab)

Function

> Basic part of GDS.

Effect if stopped

GDS functions cannot be used.

Startup daemon

/usr/sbin/sdxmond

Utilized port

None.

Remarks

None.

init (inittab)

Function

Monitoring of shutdown facility.

Effect if stopped

If shutdown facility terminates abnormally, it will not be restarted.

Startup daemon

/opt/SMAW/SMAWsf/bin/rcsd_monitor

Utilized port

None.

Remarks

None.

/etc/rc3.d

S06clonltrc

Function

Loading of the online trace driver.

Effect if stopped

Investigation information for CRM cannot be collected.

Startup daemon

None.

Utilized port

None.

Remarks

None.

S07clapi

Function

Beginning of online tracing of the Cluster Resource Management facility (1).

Effect if stopped

The cluster cannot be started.

Startup daemon

None.

Utilized port

    None.

Remarks

    None.

S08clrms

Function

    Beginning of online tracing of the Cluster Resource Management facility (2).

Effect if stopped

    The cluster cannot be started.

Startup daemon

    None.

Utilized port

    None.

Remarks

    None.

S12cf

Function

    Loading of CF and CIP drivers.

Effect if stopped

    The cluster cannot be started.

Startup daemon

    /opt/SMAW/SMAWcf/bin/cfregd

Utilized port

    None.

Remarks

    None.

S12zcldevmon

Function

    Startup of MMB asynchronous monitoring.

Effect if stopped

    MMB asynchronous monitoring cannot be used.

Startup daemon

    /etc/opt/FJSVcluster/sys/devmmbd
    /etc/opt/FJSVcluster/sys/devmmbmond
    /etc/opt/FJSVcluster/sys/devmmbmonitord
    /etc/opt/FJSVcluster/sys/devmalogd

Utilized port

Port	Protocol	Send/ Receive	Network	Target	Communication target	
					Port	Target
7000-7100	UDP	s, r	Administrative LAN	Cluster node	7000-7100	Remote cluster node

Port	Protocol	Send/Receive	Network	Target	Communication target	
					Port	Target
162	UDP	r	Administrative LAN	Local cluster node	ANY	Remote cluster node

Remarks

None.

S13SMAWsf

Function

Startup of Shutdown Facility.

Effect if stopped

Shutdown Facility cannot be used.

Startup daemon

/opt/SMAW/SMAWsf/bin/rcsd

Utilized port

Port	Protocol	Send/Receive	Network	Target	Communication target	
					Port	Target
2316	UDP	s,r	Administrative LAN	Cluster node	ANY	Remote cluster node
623	UDP	s,r	Administrative LAN	BMC/iRMC	ANY	Remote cluster node
161	UDP	s,r	Administrative LAN	Management blade	ANY	Remote cluster node

Remarks

The 2316 port is used when remote nodes are forcibly stopped in split-brain processing.

The 623 port is used in split-brain processing of S2 and S3 models.

S11hanet

Function

Startup of daemons and activation of virtual interfaces.

Effect if stopped

Creation of LAN redundancy using the Redundant Line Control function is not available.

Startup daemon

/opt/FJSVhanet/etc/sbin/hanetctld
/opt/FJSVhanet/etc/sbin/hanetselect (*1)
/opt/FJSVhanet/etc/sbin/hanetpathmd (*2)

Utilized port

Port	Protocol	Send/Receive	Network	Target	Communication target	
					Port	Target
1807	UDP	s, r	Public LAN	Cluster node	1807	Remote node (GS or SURE)

Remarks

(*1) This daemon is started only when GS linkage mode is used.

(*2) Availability of startup and the number of processes rely on the configuration. Also, this may be suspended according to the monitoring status.

S24hanet2

Function

Startup of monitoring daemon and self check daemon.

Effect if stopped

The line monitoring function and the self-checking function cannot work.

Startup daemon

/opt/FJSVhanet/etc/sbin/hanetmond (*1)
/opt/FJSVhanet/etc/sbin/hanetselect (*2)

Utilized port

None.

Remarks

(*1) This daemon is started only when the self-checking function is used.
(*2) Availability of startup and the number of processes rely on the configuration. Also, this may be suspended according to the monitoring status.

S51cldbm

Function

Startup of cluster configuration management facility.

Effect if stopped

The cluster cannot be started.

Startup daemon

/etc/opt/FJSVcluster/FJSVcldbm/daemons/dcmmond
/etc/opt/FJSVcluster/FJSVcldbm/daemons/dcmmstd
/etc/opt/FJSVcluster/FJSVcldbm/daemons/dcmevmd
/etc/opt/FJSVcluster/FJSVcldbm/daemons/dcmfcpd
/etc/opt/FJSVcluster/FJSVcldbm/daemons/dcmsynd
/etc/opt/FJSVcluster/FJSVcldbm/daemons/dcmprcd
/etc/opt/FJSVcluster/FJSVcldbm/daemons/dcmcfmd
/etc/opt/FJSVcluster/FJSVcldbm/daemons/dcmdbud
/etc/opt/FJSVcluster/FJSVcldbm/daemons/dcmcomd
/etc/opt/FJSVcluster/FJSVcldbm/daemons/dcmdbcd
/etc/opt/FJSVcluster/FJSVcldbm/daemons/dcmlckd
/etc/opt/FJSVcluster/FJSVclrms/daemons/clwatchlogd

Utilized port

Port	Protocol	Send/ Receive	Network	Target	Communication target	
					Port	Target
9331	TCP	s, r	Interconnect	Cluster node	ANY	Local cluster node
9379	TCP	s, r	Interconnect	Cluster node	ANY	Local and remote cluster nodes
9378	TCP	s, r	Interconnect	Cluster node	ANY	Local cluster node
9377	TCP	s, r	Interconnect	Cluster node	ANY	Local and remote cluster nodes
9376	TCP	s, r	Interconnect	Cluster node	ANY	Local cluster node
9375	TCP	s, r	Interconnect	Cluster node	ANY	Local cluster node
9383	TCP	s, r	Interconnect	Cluster node	ANY	Remote cluster node
9384	TCP	s, r	Interconnect	Cluster node	ANY	Remote cluster node

Remarks

None.

S51clrmgr

Function

Startup of Cluster Resource Management facility.

Effect if stopped

The cluster cannot be started.

Startup daemon

/etc/opt/FJSVcluster/FJSVcldbm/daemons/clrmd

Utilized port

None

Remarks

None.

S51clrwz

Function

Setting of cluster applications.

Effect if stopped

Cluster applications cannot be configured correctly, or will not work correctly.

Startup daemon

None.

Utilized port

None.

Remarks

None.

S52sfdsk

Function

Basic part of GDS.

Effect if stopped

GDS functions cannot be used.

Startup daemon

/usr/sbin/sdxlogd
/usr/sbin/sdxservd
/usr/sbin/sdxexd

Utilized port

None.

Remarks

None.

S53clctrl

Function

Waiting for completion of startup of Cluster Resource Management facility.

Effect if stopped

The cluster cannot be started.

Startup daemon

/usr/sbin/sdxclc
/usr/sbin/sdxcle
/usr/sbin/sdxcld

Utilized port

None.

Remarks

None.

S53sfdsk2

Function

Basic part of GDS.

Effect if stopped

GDS functions cannot be used.

Startup daemon

None.

Utilized port

None.

Remarks

None.

S57sfcfsrm

Function

Startup control for monitoring facility of GFS shared file system, mount control for GFS shared file system.

Effect if stopped

Functions of GFS shared file system cannot be used.

Startup daemon

/usr/lib/fs/sfcfs/sfcpncd
/usr/lib/fs/sfcfs/sfcprmd
/usr/lib/fs/sfcfs/sfchnsd
/usr/lib/fs/sfcfs/sfcfrmd
/usr/lib/fs/sfcfs/sfcfsd
/usr/lib/fs/sfcfs/sfcfsmg

Utilized port

Port	Protocol	Send/ Receive	Network	Target	Communication target	
					Port	Target
9300 (*1)	TCP	s, r	Interconnect	Cluster node	ANY	Remote cluster node
9200-9263 (*2)	TCP	s, r	Interconnect Administrative LAN	Cluster node	ANY	Local and remote cluster nodes

Remarks

(*1) No. 9300 is set to support the service name "sfcfsrm."

(*2) From No. 9200 to No. 9263 are set to support the service names from sfcfs-1 to sfcfs-64.

S76clprmd

Function

Startup of process monitoring facility.

Effect if stopped

Applications using the process monitoring functions will not work.

Startup daemon

/etc/opt/FJSVcluster/FJSVclapm/daemons/prmd

Utilized port

None.

Remarks

Exclusive for PRIMECLUSTER products.

S99SMAWRrms

Function

Startup of RMS.

Effect if stopped

Even if HV_RCSTART=1 is set, RMS will not start automatically at node startup.

Startup daemon

/opt/SMAW/SMAWRrms/bin/bm
/opt/SMAW/SMAWRrms/bin/hvdet_xxxx
(Detectors and applications used in cluster applications will start.)

Utilized port

Port	Protocol	Send/ Receive	Network	Target	Communication target	
					Port	Target
9786	TCP	s, r	Interconnect	Cluster node	9786	Remote cluster node
8000	UDP	s, r	Interconnect	Cluster node	8000	Remote cluster node
11111	UDP	s, r	Interconnect	Cluster node	11111	Remote cluster node

Remarks

None.

S99fjsvwvbs

Function

Startup of daemons on Web-Based Admin View management server or monitoring nodes.

Effect if stopped

Settings and monitoring via the GUI provided by Web-Based Admin View will not be available.

Startup daemon

[For nodes working as primary or secondary management servers]

/opt/SMAW/SMAWcj2re/jre/bin/java
wvAgent /opt/FJSVwvbs (2 processes)
/etc/opt/FJSVwvfrm/sbin/wvClEventd (0 to 2 processes)
/etc/opt/FJSVwvfrm/sbin/wvFaultEventd (0 to 2 processes)

[For nodes other than those described above]

wvAgent /opt/FJSVwvbs (2 processes)
/etc/opt/FJSVwvfrm/sbin/wvClEventd (0 to 2 processes)
/etc/opt/FJSVwvfrm/sbin/wvFaultEventd (0 to 2 processes)

Utilized port

Port	Protocol	Send/ Receive	Network	Target	Communication target	
					Port	Target
9799	TCP	s, r	Administrative LAN	Administrative server (*1)	ANY	WebView client (*2)
9798	TCP	s, r	Administrative LAN	Administrative server (*1)	ANY	WebView client (*2)
9797	TCP	s, r	Administrative LAN	Administrative server (*1)	ANY	Local and remote nodes
9796	UDP	s, r	Administrative LAN	Administrative server (*1)	ANY	Local and remote nodes

Remarks

(*1) Including concurrent use with cluster nodes.

(*2) PC or Linux(R) workstation.

S99fjsvwvcnf

Function

WWW server for sending Java applets, Java classes, and HTML contents to clients.

Effect if stopped

Settings and monitoring via the GUI provided by Web-Based Admin View will not be available.

Startup daemon

/opt/FJSVwvcnf/bin/wvcnfd

Utilized port

Port	Protocol	Send/ Receive	Network	Target	Communication target	
					Port	Target
8081	TCP	s, r	Administrative LAN	Administrative server (*1)	ANY	Browser (*2)

Remarks

(*1) Including concurrent use with cluster nodes.

(*2) PC or Linux(R) workstation.

For wvcnfd of the Web-Based Admin View process, there is an additional child process of the same name while processing a request from a client. This process, however, terminates immediately after processing the request.

I.3   Necessary Daemons for PRIMECLUSTER to Operate

Necessary daemons other than PRIMECLUSTER for PRIMECLUSTER to operate are as follows:

- crond

- ntpd

- radvd (*1)

- syslog (syslogd, klogd) / rsyslog (rsyslogd) (*2)

- snmptrapd (*3)

- libvirtd (*4)

(*1) The radvd daemon is necessary only when both Fast switching mode and IPv6 communication are used.

(*2) The syslogd and klogd daemons are necessary for RHEL5. The rsyslogd daemon is necessary for RHEL6.

(*3) The snmptrapd daemon is necessary only when MMB asynchronous monitoring is used.

(*4) The libvirtd daemon is used only when PRIMECLUSTER 4.3A10 or later is used in a KVM environment. It is also necessary when the virtual NIC mode of PRIMECLUSTER 4.3A10 or later is connected to a virtual bridge in virtual machine or KVM hypervisor environments.

Appendix J Using Firewall

When using Firewall, perform either of the following procedures because the cluster may not operate normally.

- Allow the communication interface used by PRIMECLUSTER.

- Allow the port number used by PRIMECLUSTER.

This chapter provides an example to configure Firewall by using iptables.

Configure the settings in accordance with the security policy.

🔖 See
....

- For details on iptables, see the man manual or other related documentation for the iptables(8) command.
....

When allowing the communication interface used by PRIMECLUSTER

With PRIMECLUSTER, communication interfaces are used on the administrative LAN and the cluster interconnects. Configure the settings to allow both communication interfaces.

The following is an example to allow sending and receiving the communication interface "cip0" of the cluster interconnect.

```
Format:  -A INPUT -i <input-interface> -j ACCEPT
         -A OUTPUT -o <output-interface> -j ACCEPT


Example: -A INPUT -i cip0 -j ACCEPT
         -A OUTPUT -o cip0 -j ACCEPT
```

When allowing the port number used by PRIMECLUSTER

See "Appendix I Startup Scripts and Startup Daemons, and Port Numbers in PRIMECLUSTER" and allow communication of all port numbers used by PRIMECLUSTER.

The following is an example to allow communications of some port numbers used by the cluster resource management facility between the local node and other nodes.

```
Format:  -A <INPUT/OUTPUT> -p <tcp/udp> -m <tcp/udp> --dport <destination-port-number> -j ACCEPT
         -A <INPUT/OUTPUT> -p <tcp/udp> -m <tcp/udp> --sport <destination-port-number> -j ACCEPT

Example: -A INPUT -p tcp -m tcp --dport 9383 -j ACCEPT
         -A OUTPUT -p tcp -m tcp --sport 9383 -j ACCEPT
         -A INPUT -p tcp -m tcp --sport 9383 -j ACCEPT
         -A OUTPUT -p tcp -m tcp --dport 9383 -j ACCEPT
```

📝 Note
....

- If you change the configuration of iptables, do not restart the iptables service. Instead, execute either of the following actions:

    - Restarting the cluster node

    - Applying the configuration change by using iptables-restore

- When using the state module in iptables, configure settings to allow communications of PRIMECLUSTER before the state module settings.
  In the following example, communications of cluster interconnects are allowed before the state module settings.

```
    Example: -A INPUT -i cip0 -j ACCEPT
             -A OUTPUT -o cip0 -j ACCEPT
             -A INPUT -m state --state RELATED,ESTABLISHED -j ACCEPT
             -A INPUT -p tcp -m tcp ! --tcp-flags SYN,RST,ACK SYN -m state --state NEW -j DROP
```
....

Appendix K Changes in Each Version

This chapter explains the changes made to the specifications of PRIMECLUSTER 4.3A10.

The changes are listed in the following table.

Table K.1 List of changes

Category	Item	Version
Incompatible commands	clgettree(1) command	(Before change) PRIMECLUSTER 4.1A20 or earlier (After change) PRIMECLUSTER 4.3A10
	hvdisp command	(Before change) PRIMECLUSTER 4.1A20 or earlier (After change) PRIMECLUSTER 4.3A10
	ciptool command	(Before change) PRIMECLUSTER 4.1A30 or earlier (After change) PRIMECLUSTER 4.3A10
Incompatible functions	Posting Notification of a Resource Failure or Recovery	(Before change) PRIMECLUSTER 4.0A20 (After change) PRIMECLUSTER 4.3A10
	Operator Intervention Function	(Before change) PRIMECLUSTER 4.0A20 (After change) PRIMECLUSTER 4.3A10
	Shutdown Facility	(Before change) PRIMECLUSTER 4.0A20 (After change) PRIMECLUSTER 4.3A10
	Setting Up Cluster Applications	(Before change) PRIMECLUSTER 4.0A20 (After change) PRIMECLUSTER 4.3A10
	Changing the Cluster Application	(Before change) PRIMECLUSTER 4.0A20 (After change) PRIMECLUSTER 4.3A10
	State of the Main CF Table	(Before change) PRIMECLUSTER 4.0A20 (After change) PRIMECLUSTER 4.3A10
	Node state	(Before change) PRIMECLUSTER 4.1A20 or earlier (After change) PRIMECLUSTER 4.3A10
	Operation Procedures and Displayed Items for Cluster Application Setup and Modification	(Before change) PRIMECLUSTER 4.1A30 or earlier (After change) PRIMECLUSTER 4.3A10
	Registration and Deletion of Resources for PRIMECLUSTER-compatible Products	(Before change) PRIMECLUSTER 4.1A30 or earlier (After change) PRIMECLUSTER 4.3A10
	Setting Up Fsystem Resources	(Before change) PRIMECLUSTER 4.2A00 or earlier (After change) PRIMECLUSTER 4.3A10
	Client Environment for Web-Based Admin View	(Before change) PRIMECLUSTER 4.3A00 or earlier (After change) PRIMECLUSTER 4.3A10

Category	Item	Version
	Changes of the Behavior of CF Startup	(Before change) PRIMECLUSTER 4.3A00 or earlier (After change) PRIMECLUSTER 4.3A10
	HV_CONNECT_TIMEOUT	(Before change) PRIMECLUSTER 4.3A00 or earlier (After change) PRIMECLUSTER 4.3A10
Incompatible message	Changes of the RMS message	(Before change) PRIMECLUSTER 4.3A00 or earlier (After change) PRIMECLUSTER 4.3A10
	Changes of the importance of the message in the RMS wizard	(Before change) PRIMECLUSTER 4.3A00 or earlier (After change) PRIMECLUSTER 4.3A10

K.1   Changes in PRIMECLUSTER 4.3A10 from 4.0A20

Incompatible commands

The following commands of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.0A20.

- K.1.1 clgettree(1) command

- K.1.2 hvdisp command

- K.1.3 ciptool command

Incompatible functions

The following functions of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.0A20.

- K.1.4 Posting Notification of a Resource Failure or Recovery

- K.1.5 Operator Intervention Function

- K.1.6 Shutdown Facility

- K.1.7 Setting Up Cluster Applications

- K.1.8 Changing the Cluster Application

- K.1.9 State of the Main CF Table

- K.1.10 Node state

- K.1.11 Operation Procedures and Displayed Items for Cluster Application Setup and Modification

- K.1.12 Registration and Deletion of Resources for PRIMECLUSTER-compatible Products

- K.1.13 Setting Up Fsystem Resources

- K.1.14 Client Environment for Web-Based Admin View

- K.1.15 Changes of the Behavior of CF Startup

- K.1.16 HV_CONNECT_TIMEOUT

Incompatible messages

The following messages of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.0A20.

- K.1.17 Changes of the RMS message

- K.1.18 Changes of the importance of the message in the RMS wizard

K.1.1 clgettree(1) command

Details on incompatibilities

Cluster class resource names, which are output with the "clgettree(1)" command, are changed.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

A fixed string "cluster" is displayed when the resource management facility is configured.

After upgrading [PRIMECLUSTER 4.3A10]

The cluster class uses the same name as the CF cluster when the resource management facility is configured.

Note

None.

K.1.2 hvdisp command

Details on incompatibilities

The number of characters in "Type" in the resources that are displayed with "hvdisp -a" or "hvdisp -u" has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

The entire contents of "Type" in the resource are displayed.

After upgrading [PRIMECLUSTER 4.3A10]

The contents of "Type" in the resource may not be completely displayed. Use "hvdip <*resource-name*>" to check the whole contents of "Resource Type."

Note

None.

Details on incompatibilities 2

StateDetails information is displayed.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

StateDetails information is not displayed.

After upgrading [PRIMECLUSTER 4.3A10]

StateDetails information is displayed.

Note

None

K.1.3 ciptool command

Details on incompatibilities

Symbolic links of the ciptool commands were not made for /usr/bin.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

Symbolic links of the ciptool commands are made for /usr/bin.

After upgrading [PRIMECLUSTER 4.3A10]

Symbolic links of the ciptool commands are not made for /usr/bin.

Note

Specify /opt/SMAW/SMAWcf/bin/ciptool by full passing when you use the ciptool command.

K.1.4   Posting Notification of a Resource Failure or Recovery

Details on incompatibilities

If a failure occurs in the resource or if the resource recovers from a failure, the failure or recovery of the resource can be posted by sending the message shown below to syslogd. The default setting at installation is that notification of a resource failure or recovery is not posted with PRIMECLUSTER 4.3A10. For details, see "5.2 Setting Up Fault Resource Identification and Operator Intervention Request."

Message No	Message overview
2700	Recovering from a resource failure
2701	Recovering from a node failure
6750	Resource failure
6751	Node failure

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

Notification of a resource failure or recovery will be posted.

After upgrading **[PRIMECLUSTER 4.3A10]**

A resource failure or recovery will be posted only when the AppWatch parameter is set to ON with clsetparam. The default value of AppWatch at cluster installation is OFF and notification of the resource failure or recovery will not be posted.

Note

After you have changed the AppWatch parameter with clsetparam, you have to reboot all the nodes to validate the setting.

K.1.5   Operator Intervention Function

Details on incompatibilities

To enable the operator intervention function, you have to set the AppWatch parameter to ON with clsetparam. With the default settings made when the cluster was installed, the operator intervention function is disabled. For details, see "5.2 Setting Up Fault Resource Identification and Operator Intervention Request."

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

The operator intervention function is always enabled.

After upgrading [PRIMECLUSTER 4.3A10]

The operator intervention function is enabled only when the AppWatch parameter is set to ON with clsetparam. The default value of AppWatch set when the cluster was installed is set to OFF, and the operator intervention function will not work with this default value.

Note

After you have changed the AppWatch parameter with clsetparam, you have to reboot all the nodes to validate the setting.

K.1.6 Shutdown Facility

Details on incompatibilities

"SF node weight" is added as a new setting item for the shutdown facility of PRIMECLUSTER 4.3A10. For the same survival priority as PRIMECLUSTER 4.0A20, specify 1 in "SF node weight" for all the nodes. For details, see "5.1.2 Setting Up the Shutdown Facility."

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

Survival priority is determined by "ShutdownPriority of userApplication."

After upgrading [PRIMECLUSTER 4.3A10]

Survival priority is determined by "ShutdownPriority of userApplication" and "SF node weight."

Note

None.

K.1.7 Setting Up Cluster Applications

Details on incompatibilities

If a cluster application for a standby operation is used with PRIMECLUSTER 4.3A10, set up "ClearFaultRequest|StartUp|SwitchRequest" to the StandbyTransition attribute. With this setting, a standby state transition can be performed in the same way as PRIMECLUSTER 4.0A20 for the following cases:

- When a cluster application in the Faulted state is set to operating state

- When a cluster application is started up

- When a cluster application is switched

For details on the settings of the cluster application, see "6.7 Setting Up Cluster Applications."

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

If a resource for support hot-standby operation is available, a standby state transition will be made.

After upgrading [PRIMECLUSTER 4.3A10]

A standby state transition will be performed only when a resource to support hot-standby operation is available and "ClearFaultRequest|StartUp|SwitchRequest" is set to the StandbyTransition attribute.

Note

None.

K.1.8 Changing the Cluster Application

Details on incompatibilities

If the "hvw" command is executed while RMS is running, the following menus, that could be displayed with PRIMECLUSTER 4.0A20, cannot be displayed.

- Application-Create

- Application-Edit

- Application-Remove

- Configuration-Activate

To change the cluster application, you have to stop RMS before you execute the "hvw" command. This is the same with PRIMECLUSTER 4.0A20.

If you execute the "hvw" command after stopping RMS, the same menus as PRIMECLUSTER 4.0A20 will be displayed.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

If you execute the "hvw" command while RMS is running, the following menus will be displayed.

- Application-Create

- Application-Edit

- Application-Remove

- Configuration-Activate

After upgrading [PRIMECLUSTER 4.3A10]

If you execute the "hvw" command while RMS is running, the following menus will not be displayed.

- Application-Create

- Application-Edit

- Application-Remove

- Configuration-Activate

Note

To change the cluster application, you need to stop RMS before you execute the "hvw" command. This is the same with PRIMECLUSTER 4.0A20.

For details on changing the cluster application, see "8.1 Changing the Cluster Configuration."

K.1.9 State of the Main CF Table

Details on incompatibilities

"INVALID" is added to PRIMECLUSTER 4.3A10 as a new state of the local node in the main CF table.

For details on the main CF table, see "5 GUI Management" in "*PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide.*"

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

If the configuration of the node is invalid, "UNCONFIGURED" or "UNKNOWN" will be displayed as the state of the local node on the main CF table.

After upgrading [PRIMECLUSTER 4.3A10]

If the node configuration is invalid, "INVALID" will be displayed as the state of the local node on the main CF table.

Note

None.

K.1.10 Node state

Details on incompatibilities

When any one of all the stopped nodes is started, the state of nodes that are separated from the cluster is changed when it is displayed by hvdisp (1M) command or output on the RMS main window.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

The nodes that are separated from the cluster show Offline state.

After upgrading [PRIMECLUSTER 4.3A10]

The nodes that are separated from the cluster show Faulted state.
See details on the state of the nodes (the value of the StateDetails attribute) that are displayed by hvdisp (1M) command or output on the RMS main window to identify if Faulted state is due to the nodes being separated from the cluster or due to an abnormal shutdown of the nodes by a panic or other errors.

	State	StateDetails
When a node is not joining the cluster	Faulted	Shutdown
Abnormal shutdown by panic. etc	Faulted	Killed

Note

None.

K.1.11 Operation Procedures and Displayed Items for Cluster Application Setup and Modification

Details on incompatibilities

The procedures and displayed items for creating userApplications and for creating and changing procedure resources have been changed.

Changes

1. Creating a userApplication for standby operation

Before upgrading [PRIMECLUSTER 4.0A20]

To create a userApplication for standby operation, select "CRM" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP
 2) QUIT
 3) RETURN
 4) OPTIONS
 5) CRM
 6) GENERIC
 7) SYMFOWARE
Application Type: 5
```

↓

```
Settings of turnkey wizard "CRM"
 1) HELP                4) REMOVE+EXIT            7) Machines+Basics(-)
 2) NO-SAVE+EXIT        5) ApplicationName=APP1
 3) SAVE+EXIT           6) BeingControlled=no
Choose the setting to process: 7
```

After upgrading [PRIMECLUSTER 4.3A10]

"CRM" is not displayed in the "Application type selection menu."

To create a userApplication for standby operation, select "STANDBY" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP                          7) SCALABLE
 2) QUIT                          8) STANDBY
 3) RETURN
 4) OPTIONS
 5) DEMO
 6) GENERIC
Application Type: 8
```

⬇

```
Consistency check ...
Yet to do: process the basic settings using Machines+Basics
Yet to do: process at least one of the non-basic settings

Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                          4) REMOVE+EXIT
 2) NO-SAVE+EXIT                  5) ApplicationName=APP1
 3) SAVE+EXIT                     6) Machines+Basics(-)
Choose the setting to process: 6
```

2. Creating a userApplication for scalable operation

Before upgrading [PRIMECLUSTER 4.0A20]

To create a userApplication for scalable operation, select "Controller" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP                  9) Controller        17) Rawdisk
 2) QUIT                 10) Crm               18) SYMFOWARE
 3) RETURN               11) Fsystem           19) Symfoware
 4) OPTIONS              12) GENERIC           20) SystemState2
 5) Application          13) Gds               21) SystemState3
 6) BasicApplication     14) Gls               22) foreign-code
 7) CRM                  15) Ipaddress         23) generic
 8) Cmdline              16) Lvm
Application Type: 9
```

⬇

```
Settings of application type "Controller" (not yet consistent)
 1) HELP                          5) AdditionalAppToControl
 2) NO-SAVE+EXIT                  6) (InParallel=)
 3) ApplicationName=Controller   7) (FaultScript=)
 4) ControlPolicy=FOLLOW
Choose the setting to process: 4
```

After upgrading [PRIMECLUSTER 4.3A10]

"Controller" is not displayed in the "Application type selection menu."

To create a userApplication for scalable operation, select "SCALABLE" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP                          7) SCALABLE
 2) QUIT                          8) STANDBY
 3) RETURN
 4) OPTIONS
 5) DEMO
 6) GENERIC
Application Type: 7
```

↓

```
Consistency check ...
Yet to do: process the basic settings using Machines+Basics
Yet to do: process at least one of the non-basic settings

Settings of turnkey wizard "SCALABLE" (APP2:not yet consistent)
 1) HELP                          4) REMOVE+EXIT
 2) NO-SAVE+EXIT                  5) ApplicationName=APP2
 3) SAVE+EXIT                     6) Machines+Basics(-)
Choose the setting to process: 6
```

3. Creating a procedure resource

   Before upgrading [PRIMECLUSTER 4.0A20]

   To create or change a procedure resource, select "CRM" from the "turnkey wizard "CRM"" menu, and then select the resource class name.

   Note: Shown below is an example of registering a procedure resource of the BasicApplication class to a userApplication.

```
Settings of turnkey wizard "CRM"
 1) HELP                          10) LocalFileSystems(-)
 2) -                             11) RemoteFileSystems(-)
 3) SAVE+EXIT                     12) IpAddresses(-)
 4) -                             13) Gds:Global-Disk-Services(-)
 5) ApplicationName=APP1          14) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)
 7) CommandLines(-)
 8) Controllers(-)
 9) CRM(-)
Choose the setting to process: 9
```

↓

```
Crm (CRM_APP1:not yet consistent)
 1) HELP
 2) -
 3) SAVE+EXIT
 4) REMOVE+EXIT
 5) SystemState2(-)
 6) SystemState3(-)
 7) BasicApplication(-)
 8) Application(-)
Choose the setting to process: 7
```

After upgrading [PRIMECLUSTER 4.3A10]

The "turnkey wizard "CRM"" menu is not displayed.

To create or change a procedure resource, select "Procedure:resource-class-name" from the "turnkey wizard STANDBY" menu.

Note: Shown below is an example of registering a procedure resource of the BasicApplication class to a userApplication.

```
Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                            9) Procedure:BasicApplication(-)
 2) -                              10) Symfoware(-)
 3) SAVE+EXIT                      11) Procedure:SystemState3(-)
 4) -                              12) Procedure:SystemState2(-)
 5) ApplicationName=APP1           13) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)          14) IpAddresses(-)
 7) CommandLines(-)                15) LocalFileSystems(-)
 8) Procedure:Application(-)       16) Gds:Global-Disk-Services(-)
Choose the setting to process:9
```

4. Changing the timeout value of a procedure resource

Before upgrading [PRIMECLUSTER 4.0A20]

To change the timeout value, select "Flags[0]=OTtimeout-time" and then select "SCRIPTTIMEOUT."

```
Settings of "Application"
 1) HELP               4) REMOVE+RETURN         7) Flags[0]=OT1800
 2) NO-SAVE+RETURN     5) AdditionalResource    8) Priority[0]=65535
 3) SAVE+RETURN        6) ResourceName[0]=rsc1
Choose the setting to process: 7
```

↓

```
Set Flags[0]:Currently set: STANDBYCAPABLE,SCRIPTTIMEOUT (OT1800)
 1) HELP                5) NULLDETECTOR(D)        9) MONITORONLY(M)
 2) -                   6) LIEOFFLINE(L)         10) NOT:STANDBYCAPABLE(O)
 3) SAVE+RETURN         7) CLUSTEREXCLUSIVE(X)   11) SCRIPTTIMEOUT(T)
 4) DEFAULT             8) AUTORECOVER(A)
Choose one of the flags: 11
```

↓

```
 1) HELP
 2) RETURN
 3) FREECHOICE
 4) 1800
Set an appropriate script timeout:
```

After upgrading [PRIMECLUSTER 4.3A10]

"Flags[0]=OTtimeout-time" is not displayed.

To change the timeout value, select "TIMEOUT."

```
Set flags for Procedure resource :
Currently set: TIMEOUT (T1800), PRIORITY (P65535)
 1) HELP
 2) -
 3) SAVE+RETURN
 4) TIMEOUT
 5) PRIORITY
Choose one of the flags: 4
```

↓

```
 1) HELP
 2) RETURN
 3) FREECHOICE
 4) 1800
Set an appropriate timeout:
```

5. Changing the priority in a resource class of a procedure resource.

   Before upgrading [PRIMECLUSTER 4.0A20]

   To change the priority in a resource class, select "Priority[0]=*priority*."

```
Settings of "Application"
 1) HELP               4) REMOVE+RETURN       7) Flags[0]=OT1800
 2) NO-SAVE+RETURN     5) AdditionalResource  8) Priority[0]=65535
 3) SAVE+RETURN        6) ResourceName[0]=rscl
Choose the setting to process: 8
```

↓

```
 1) HELP
 2) RETURN
 3) FREECHOICE
Set a priority:
```

   After upgrading [PRIMECLUSTER 4.3A10]

   "Priority[0]=*priority*" is not displayed.

   To change the priority in a resource class, select "PRIORITY."

```
Set flags for Procedure resource :
Currently set: TIMEOUT (T1800), PRIORITY (P65535)
 1) HELP
 2) -
 3) SAVE+RETURN
 4) TIMEOUT
 5) PRIORITY
Choose one of the flags: 5
```

↓

```
 1) HELP
 2) RETURN
 3) FREECHOICE
 4) 65535
Set an appropriate priority:
```

Note

None.

K.1.12 Registration and Deletion of Resources for PRIMECLUSTER-compatible Products

Details on incompatibilities

The procedures for registering and deleting resources for PRIMECLUSTER-compatible products have been changed.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

1. Registering resources for PRIMECLUSTER-compatible products

   Execute multiple commands to register the following resources for PRIMECLUSTER-compatible products:

   - Resources of cluster services

   - Resources of cluster service instances

2. Deleting resource for PRIMECLUSTER-compatible products

   Execute multiple commands to delete the following resources for PRIMECLUSTER-compatible products:

   - Resources of cluster services

   - Resources of cluster service instances

3. Reconfiguring a cluster application

   Delete the cluster service, and then register the cluster service again.

After upgrading [PRIMECLUSTER 4.3A10]

1. Registering resources for PRIMECLUSTER-compatible products

   Execute the "clrwzconfig" command to register the following resources for PRIMECLUSTER-compatible products:

   - Resources of cluster services

   - Resources of cluster service instances

2. Deleting resources for PRIMECLUSTER-compatible products

   Execute the "clrwzconfig" command to delete the following resources for PRIMECLUSTER-compatible products:

   - Resources of cluster services

   - Resources of cluster service instances

3. Reconfiguring a cluster application

   The cluster service must be registered when a cluster application is reconfigured.

   Use the "clrwzconfig" command to check whether the cluster service needs to be set up again. If re-setup of the cluster service is necessary, execute the "clrwzconfig" command.

Note

The following operations have not been changed:

- Registering a resource for IP address takeover and associating a resource to a cluster service instance

- Deleting a resource for IP address takeover

K.1.13 Setting Up Fsystem Resources

Details on incompatibilities

When you use shared disk device, you need to reserve the dedicated monitoring disk area, and register it to the userApplication as an Fsystem resource.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

   The monitoring disk area was not needed.

After upgrading [PRIMECLUSTER 4.3A10]

   The dedicated monitoring disk area must be registered to the userApplication as an Fsystem resource.

   For details, see "6.7.3.2 Setting Up Fsystem Resources."

Note

None.

Details on incompatibilities

The file which defines the mount point of the filesystem used as Fsystem resource has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

   The mount point was defined in /etc/fstab.

After upgrading [PRIMECLUSTER 4.3A10]

   It is necessary to define the mount point in /etc/fstab.pcl.

   For details, see "6.7.3.2 Setting Up Fsystem Resources."

Note

None

K.1.14 Client Environment for Web-Based Admin View

Details on incompatibilities

Linux(R) is not supported as a client environment for Web-Based Admin View by PRIMECLUSTER 4.3A10.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

Linux(R) is supported as a client environment for Web-Based Admin View.

After upgrading [PRIMECLUSTER 4.3A10]

Linux(R) is not supported as a client environment for Web-Based Admin View.

Note

None.

K.1.15 Changes of the Behavior of CF Startup

Details on incompatibilities

CF starts even if some of the network interfaces for the cluster interconnects are not recognized.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

CF does not start unless all of the network interfaces for the cluster interconnects are recognized.

After upgrading [PRIMECLUSTER 4.3A10]

CF starts if at least one of the network interfaces for the cluster interconnects is recognized.

Note

If there are any network interfaces that are not recognized on CF startup, the following message appears:

CF: <NIC>: device not found.

<NIC> will be the name of the network interface (e.g. eth0).

This message is also available in 4.0A20.

K.1.16 HV_CONNECT_TIMEOUT

Details on incompatibilities

The default value of the RMS local environment variables HV_CONNECT_TIMEOUT is changed.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

The default value of HV_CONNECT_TIMEOUT is 5 (seconds).

After upgrading [PRIMECLUSTER 4.3A10]

The default value of HV_CONNECT_TIMEOUT is 30 (seconds).

Note

For details on HV_CONNECT_TIMEOUT, see "12.3 Local environment variables" in the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide".

K.1.17 Changes of the RMS message

Details on incompatibilities

The RMS message (SYS, 8) logged in the syslog have been changed.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung.

After upgrading [PRIMECLUSTER 4.3A10]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung. An operator intervention is required.

Note

None.

K.1.18 Changes of the importance of the message in the RMS wizard

Details on incompatibilities

The importance of the following message in the RMS wizard has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.0A20]

WARNING: cannot grab mount lock for dostat() check_getbdev(), returning previous state

After upgrading [PRIMECLUSTER 4.3A10]

NOTICE: cannot grab mount lock for dostat() check_getbdev(), returning previous state

Note

None.

K.2 Changes in PRIMECLUSTER 4.3A10 from 4.1A20

Incompatible commands

The following commands of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.1A20.

- K.2.1 clgettree(1) command

- K.2.2 hvdisp command

- K.2.3 ciptool command

Incompatible functions

The following functions of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.1A20.

- K.2.4 Node state

- K.2.5 Operation Procedures and Displayed Items for Cluster Application Setup and Modification

- K.2.6 Registration and Deletion of Resources for PRIMECLUSTER-compatible Products

- K.2.7 Setting Up Fsystem Resources

- K.2.8 Client Environment for Web-Based Admin View

- K.2.9 Changes of the Behavior of CF Startup

Incompatible messages

The following messages of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.1A20.

K.2.1 clgettree(1) command

Details on incompatibilities

Cluster class resource names, which are output with the "clgettree(1)" command, are changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A20]

A fixed string "cluster" is displayed when the resource management facility is configured.

After upgrading [PRIMECLUSTER 4.3A10]

The cluster class uses the same name as the CF cluster when the resource management facility is configured.

Note

None.

K.2.2 hvdisp command

Details on incompatibilities

The number of characters in "Type" in the resources that are displayed with "hvdisp -a" or "hvdisp -u" has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A20]

The entire contents of "Type" in the resource are displayed.

After upgrading [PRIMECLUSTER 4.3A10]

The contents of "Type" in the resource may not be completely displayed. Use "hvdip <*resource-name*>" to check the whole contents of "Resource Type."

Note

None.

Details on incompatibilities 2

StateDetails information is displayed.

Changes

Before upgrading [PRIMECLUSTER 4.1A20]

StateDetails information is not displayed.

After upgrading [PRIMECLUSTER 4.3A10]

StateDetails information is displayed.

None.

K.2.3 ciptool command

Details on incompatibilities

Symbolic links of the ciptool commands were not made for /usr/bin.

Changes

Before upgrading [PRIMECLUSTER 4.1A20]

Symbolic links of the ciptool commands are made for /usr/bin.

After upgrading [PRIMECLUSTER 4.3A10]

Symbolic links of the ciptool commands are not made for /usr/bin.

Note

Specify /opt/SMAW/SMAWcf/bin/ciptool by full passing when you use the ciptool command.

K.2.4 Node state

Details on incompatibilities

When any one of all the stopped nodes is started, the state of nodes that are separated from the cluster is changed when it is displayed by hvdisp (1M) command or output on the RMS main window.

Changes

Before upgrading [PRIMECLUSTER 4.1A20]

The nodes that are separated from the cluster show Offline state.

After upgrading [PRIMECLUSTER 4.3A10]

The nodes that are separated from the cluster show Faulted state.
See details on the state of the nodes (the value of the StateDetails attribute) that are displayed by hvdisp (1M) command or output on the RMS main window to identify if Faulted state is due to the nodes being separated from the cluster or due to an abnormal shutdown of the nodes by a panic or other errors.

	State	StateDetails
When a node is not joining the cluster	Faulted	Shutdown
Abnormal shutdown by panic. etc	Faulted	Killed

Note

None.

K.2.5 Operation Procedures and Displayed Items for Cluster Application Setup and Modification

Details on incompatibilities

The procedures and displayed items for creating userApplications and for creating and changing procedure resources have been changed.

Changes

1. Creating a userApplication for standby operation

   Before upgrading [PRIMECLUSTER 4.1A20]

   To create a userApplication for standby operation, select "CRM" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP
 2) QUIT
 3) RETURN
 4) OPTIONS
 5) CRM
 6) GENERIC
 7) SYMFOWARE
Application Type: 5
```

```
Settings of turnkey wizard "CRM"
 1) HELP                 4) REMOVE+EXIT          7) Machines+Basics(-)
 2) NO-SAVE+EXIT         5) ApplicationName=APP1
 3) SAVE+EXIT            6) BeingControlled=no
Choose the setting to process: 7
```

   After upgrading [PRIMECLUSTER 4.3A10]

   "CRM" is not displayed in the "Application type selection menu."

   To create a userApplication for standby operation, select "STANDBY" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP                          7) SCALABLE
 2) QUIT                          8) STANDBY
 3) RETURN
 4) OPTIONS
 5) DEMO
 6) GENERIC
Application Type: 8
```

```
Consistency check ...
Yet to do: process the basic settings using Machines+Basics
Yet to do: process at least one of the non-basic settings

Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                          4) REMOVE+EXIT
 2) NO-SAVE+EXIT                  5) ApplicationName=APP1
 3) SAVE+EXIT                     6) Machines+Basics(-)
Choose the setting to process: 6
```

2. Creating a userApplication for scalable operation

   Before upgrading [PRIMECLUSTER 4.1A20]

   To create a userApplication for scalable operation, select "Controller" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP                    9) Controller          17) Rawdisk
 2) QUIT                   10) Crm                 18) SYMFOWARE
 3) RETURN                 11) Fsystem             19) Symfoware
 4) OPTIONS                12) GENERIC             20) SystemState2
 5) Application            13) Gds                 21) SystemState3
 6) BasicApplication       14) Gls                 22) foreign-code
 7) CRM                    15) Ipaddress           23) generic
 8) Cmdline                16) Lvm
Application Type: 9
```

```
Settings of application type "Controller" (not yet consistent)
 1) HELP                             5) AdditionalAppToControl
 2) NO-SAVE+EXIT                     6) (InParallel=)
 3) ApplicationName=Controller      7) (FaultScript=)
 4) ControlPolicy=FOLLOW
Choose the setting to process: 4
```

After upgrading [PRIMECLUSTER 4.3A10]

"Controller" is not displayed in the "Application type selection menu."

To create a userApplication for scalable operation, select "SCALABLE" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP                            7) SCALABLE
 2) QUIT                            8) STANDBY
 3) RETURN
 4) OPTIONS
 5) DEMO
 6) GENERIC
Application Type: 7
```

```
Consistency check ...
Yet to do: process the basic settings using Machines+Basics
Yet to do: process at least one of the non-basic settings

Settings of turnkey wizard "SCALABLE" (APP2:not yet consistent)
 1) HELP                     4) REMOVE+EXIT
 2) NO-SAVE+EXIT             5) ApplicationName=APP2
 3) SAVE+EXIT               6) Machines+Basics(-)
Choose the setting to process: 6
```

3. Creating a procedure resource

   Before upgrading [PRIMECLUSTER 4.1A20]

   To create or change a procedure resource, select "CRM" from the "turnkey wizard "CRM"" menu, and then select the resource class name.

   Note: Shown below is an example of registering a procedure resource of the BasicApplication class to a userApplication.

```
Settings of turnkey wizard "CRM"
 1) HELP                              10) LocalFileSystems(-)
 2) -                                 11) RemoteFileSystems(-)
 3) SAVE+EXIT                         12) IpAddresses(-)
 4) -                                 13) Gds:Global-Disk-Services(-)
 5) ApplicationName=APP1              14) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)
 7) CommandLines(-)
 8) Controllers(-)
 9) CRM(-)
Choose the setting to process: 9
```

```
Crm (CRM_APP1:not yet consistent)
 1) HELP
 2) -
 3) SAVE+EXIT
 4) REMOVE+EXIT
 5) SystemState2(-)
 6) SystemState3(-)
 7) BasicApplication(-)
 8) Application(-)
Choose the setting to process: 7
```

After upgrading [PRIMECLUSTER 4.3A10]

The "turnkey wizard "CRM"" menu is not displayed.

To create or change a procedure resource, select "Procedure:resource-class-name" from the "turnkey wizard STANDBY" menu.

Note: Shown below is an example of registering a procedure resource of the BasicApplication class to a userApplication.

```
Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                              9) Procedure:BasicApplication(-)
 2) -                                10) Symfoware(-)
 3) SAVE+EXIT                        11) Procedure:SystemState3(-)
 4) -                               12) Procedure:SystemState2(-)
 5) ApplicationName=APP1            13) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)           14) IpAddresses(-)
 7) CommandLines(-)                 15) LocalFileSystems(-)
 8) Procedure:Application(-)        16) Gds:Global-Disk-Services(-)
Choose the setting to process:9
```

4. Changing the timeout value of a procedure resource

Before upgrading [PRIMECLUSTER 4.1A20]

To change the timeout value, select "Flags[0]=OTtimeout-time" and then select "SCRIPTTIMEOUT."

```
Settings of "Application"
 1) HELP                4) REMOVE+RETURN          7) Flags[0]=OT1800
 2) NO-SAVE+RETURN      5) AdditionalResource     8) Priority[0]=65535
 3) SAVE+RETURN         6) ResourceName[0]=rsc1
Choose the setting to process: 7
```

↓

```
Set Flags[0]:Currently set: STANDBYCAPABLE,SCRIPTTIMEOUT (OT1800)
 1) HELP               5) NULLDETECTOR(D)        9) MONITORONLY(M)
 2) -                  6) LIEOFFLINE(L)         10) NOT:STANDBYCAPABLE(O)
 3) SAVE+RETURN        7) CLUSTEREXCLUSIVE(X)   11) SCRIPTTIMEOUT(T)
 4) DEFAULT            8) AUTORECOVER(A)
Choose one of the flags: 11
```

↓

```
 1) HELP
 2) RETURN
 3) FREECHOICE
 4) 1800
Set an appropriate script timeout:
```

After upgrading [PRIMECLUSTER 4.3A10]

"Flags[0]=OTtimeout-time" is not displayed.

To change the timeout value, select "TIMEOUT."

```
Set flags for Procedure resource :
Currently set: TIMEOUT (T1800), PRIORITY (P65535)
 1) HELP
 2) -
 3) SAVE+RETURN
 4) TIMEOUT
 5) PRIORITY
Choose one of the flags: 4
```

↓

```
 1) HELP
 2) RETURN
 3) FREECHOICE
 4) 1800
Set an appropriate timeout:
```

5. Changing the priority in a resource class of a procedure resource.

Before upgrading [PRIMECLUSTER 4.1A20]

To change the priority in a resource class, select "Priority[0]=*priority*."

```
Settings of "Application"
 1) HELP               4) REMOVE+RETURN        7) Flags[0]=OT1800
 2) NO-SAVE+RETURN     5) AdditionalResource   8) Priority[0]=65535
 3) SAVE+RETURN        6) ResourceName[0]=rsc1
Choose the setting to process: 8
```

↓

```
 1) HELP
 2) RETURN
 3) FREECHOICE
Set a priority:
```

After upgrading [PRIMECLUSTER 4.3A10]

"Priority[0]=*priority*" is not displayed.

To change the priority in a resource class, select "PRIORITY."

```
Set flags for Procedure resource :
Currently set: TIMEOUT (T1800), PRIORITY (P65535)
 1) HELP
 2) -
 3) SAVE+RETURN
 4) TIMEOUT
 5) PRIORITY
Choose one of the flags: 5
```

↓

```
 1) HELP
 2) RETURN
 3) FREECHOICE
 4) 65535
Set an appropriate priority:
```

Note

None.

K.2.6 Registration and Deletion of Resources for PRIMECLUSTER-compatible Products

Details on incompatibilities

The procedures for registering and deleting resources for PRIMECLUSTER-compatible products have been changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A20]

1. Registering resources for PRIMECLUSTER-compatible products

   Execute multiple commands to register the following resources for PRIMECLUSTER-compatible products:

   - Resources of cluster services

- Resources of cluster service instances

2. Deleting resource for PRIMECLUSTER-compatible products

Execute multiple commands to delete the following resources for PRIMECLUSTER-compatible products:

- Resources of cluster services

- Resources of cluster service instances

3. Reconfiguring a cluster application

Delete the cluster service, and then register the cluster service again.

After upgrading [PRIMECLUSTER 4.3A10]

1. Registering resources for PRIMECLUSTER-compatible products

Execute the "clrwzconfig" command to register the following resources for PRIMECLUSTER-compatible products:

- Resources of cluster services

- Resources of cluster service instances

2. Deleting resources for PRIMECLUSTER-compatible products

Execute the "clrwzconfig" command to delete the following resources for PRIMECLUSTER-compatible products:

- Resources of cluster services

- Resources of cluster service instances

3. Reconfiguring a cluster application

The cluster service must be registered when a cluster application is reconfigured.

Use the "clrwzconfig" command to check whether the cluster service needs to be set up again. If re-setup of the cluster service is necessary, execute the "clrwzconfig" command.

Note

The following operations have not been changed:

- Registering a resource for IP address takeover and associating a resource to a cluster service instance

- Deleting a resource for IP address takeover

K.2.7 Setting Up Fsystem Resources

Details on incompatibilities

When you use shared disk device, you need to reserve the dedicated monitoring disk area, and register it to the userApplication as an Fsystem resource.

Changes

Before upgrading [PRIMECLUSTER 4.1A20]

The monitoring disk area was not needed.

After upgrading [PRIMECLUSTER 4.3A10]

The dedicated monitoring disk area must be registered to the userApplication as an Fsystem resource.

For details, see "6.7.3.2 Setting Up Fsystem Resources."

Note

None.

Details on incompatibilities

The file which defines the mount point of the filesystem used as Fsystem resource has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A20]

The mount point was defined in /etc/fstab.

After upgrading PRIMECLUSTER 4.3A10]

It is necessary to define the mount point in /etc/fstab.pcl.

For details, see "6.7.3.2 Setting Up Fsystem Resources."

Note

None.

K.2.8   Client Environment for Web-Based Admin View

Details on incompatibilities

Linux(R) is not supported as a client environment for Web-Based Admin View by PRIMECLUSTER 4.3A10.

Changes

Before upgrading [PRIMECLUSTER 4.1A20]

Linux(R) is supported as a client environment for Web-Based Admin View.

After upgrading [PRIMECLUSTER 4.3A10]

Linux(R) is not supported as a client environment for Web-Based Admin View.

Note

None.

K.2.9   Changes of the Behavior of CF Startup

Details on incompatibilities

CF starts even if some of the network interfaces for the cluster interconnects are not recognized.

Changes

Before upgrading [PRIMECLUSTER 4.1A20]

CF does not start unless all of the network interfaces for the cluster interconnects are recognized.

After upgrading [PRIMECLUSTER 4.3A10]

CF starts if at least one of the network interfaces for the cluster interconnects is recognized.

Note

If there are any network interfaces that are not recognized on CF startup, the following message appears:

CF: <NIC>: device not found.

<NIC> will be the name of the network interface (e.g. eth0).

This message is also available in 4.1A20.

K.2.10   HV_CONNECT_TIMEOUT

Details on incompatibilities

The default value of the RMS local environment variables HV_CONNECT_TIMEOUT is changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A20]

The default value of HV_CONNECT_TIMEOUT is 5 (seconds).

After upgrading [PRIMECLUSTER 4.3A10]

The default value of HV_CONNECT_TIMEOUT is 30 (seconds).

Note

For details on HV_CONNECT_TIMEOUT, see "12.3 Local environment variables" in the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide".

K.2.11 Changes of the RMS message

Details on incompatibilities

The RMS message (SYS, 8) logged in the syslog have been changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A20]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung.

After upgrading [PRIMECLUSTER 4.3A10]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung. An operator intervention is required.

Note

None.

K.2.12 Changes of the importance of the message in the RMS wizard

Details on incompatibilities

The importance of the following message in the RMS wizard has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A20]

WARNING: cannot grab mount lock for dostat() check_getbdev(), returning previous state

After upgrading [PRIMECLUSTER 4.3A10]

NOTICE: cannot grab mount lock for dostat() check_getbdev(), returning previous state

Note

None.

K.3 Changes in PRIMECLUSTER 4.3A10 from 4.1A30

Incompatible commands

The following commands of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.1A30.

- K.3.1 ciptool command

Incompatible functions

The following functions of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.1A30.

- K.3.2 Operation Procedures and Displayed Items for Cluster Application Setup and Modification
- K.3.3 Registration and Deletion of Resources for PRIMECLUSTER-compatible Products
- K.3.4 Setting Up Fsystem Resources
- K.3.5 Client Environment for Web-Based Admin View
- K.3.6 Changes of the Behavior of CF Startup
- K.3.7 HV_CONNECT_TIMEOUT

Incompatible messages

The following messages of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.1A30.

- K.3.8 Changes of the RMS message
- K.3.9 Changes of the importance of the message in the RMS wizard

K.3.1 ciptool command

Details on incompatibilities

Symbolic links of the ciptool commands were not made for /usr/bin.

Changes

Before upgrading [PRIMECLUSTER 4.1A30]

Symbolic links of the ciptool commands are made for /usr/bin.

After upgrading [PRIMECLUSTER 4.3A10]

Symbolic links of the ciptool commands are not made for /usr/bin.

Note

Specify /opt/SMAW/SMAWcf/bin/ciptool by full passing when you use the ciptool command.

K.3.2 Operation Procedures and Displayed Items for Cluster Application Setup and Modification

Details on incompatibilities

The procedures and displayed items for creating userApplications and for creating and changing procedure resources have been changed.

Changes

1. Creating a userApplication for standby operation

    Before upgrading [PRIMECLUSTER 4.1A30]

    To create a userApplication for standby operation, select "CRM" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP
 2) QUIT
 3) RETURN
 4) OPTIONS
 5) CRM
 6) GENERIC
 7) SYMFOWARE
Application Type: 5
```

↓

```
Settings of turnkey wizard "CRM"
 1) HELP                 4) REMOVE+EXIT              7) Machines+Basics(-)
 2) NO-SAVE+EXIT         5) ApplicationName=APP1
 3) SAVE+EXIT            6) BeingControlled=no
Choose the setting to process: 7
```

After upgrading [PRIMECLUSTER 4.3A10]

"CRM" is not displayed in the "Application type selection menu."

To create a userApplication for standby operation, select "STANDBY" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP                          7) SCALABLE
 2) QUIT                          8) STANDBY
 3) RETURN
 4) OPTIONS
 5) DEMO
 6) GENERIC
Application Type: 8
```

↓

```
Consistency check ...
Yet to do: process the basic settings using Machines+Basics
Yet to do: process at least one of the non-basic settings

Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                          4) REMOVE+EXIT
 2) NO-SAVE+EXIT                  5) ApplicationName=APP1
 3) SAVE+EXIT                     6) Machines+Basics(-)
Choose the setting to process: 6
```

2. Creating a userApplication for scalable operation

   Before upgrading [PRIMECLUSTER 4.1A30]

      To create a userApplication for scalable operation, select "Controller" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP                    9) Controller           17) Rawdisk
 2) QUIT                   10) Crm                   18) SYMFOWARE
 3) RETURN                 11) Fsystem               19) Symfoware
 4) OPTIONS                12) GENERIC               20) SystemState2
 5) Application            13) Gds                   21) SystemState3
 6) BasicApplication       14) Gls                   22) foreign-code
 7) CRM                    15) Ipaddress             23) generic
 8) Cmdline                16) Lvm
Application Type: 9
```

↓

```
Settings of application type "Controller" (not yet consistent)
 1) HELP                           5) AdditionalAppToControl
 2) NO-SAVE+EXIT                   6) (InParallel=)
 3) ApplicationName=Controller     7) (FaultScript=)
 4) ControlPolicy=FOLLOW
Choose the setting to process: 4
```

After upgrading [PRIMECLUSTER 4.3A10]

"Controller" is not displayed in the "Application type selection menu."

To create a userApplication for scalable operation, select "SCALABLE" from the "Application type selection menu."

```
Creation: Application type selection menu:
 1) HELP                           7) SCALABLE
 2) QUIT                           8) STANDBY
 3) RETURN
 4) OPTIONS
 5) DEMO
 6) GENERIC
Application Type: 7
```

↓

```
Consistency check ...
Yet to do: process the basic settings using Machines+Basics
Yet to do: process at least one of the non-basic settings

Settings of turnkey wizard "SCALABLE" (APP2:not yet consistent)
 1) HELP                           4) REMOVE+EXIT
 2) NO-SAVE+EXIT                   5) ApplicationName=APP2
 3) SAVE+EXIT                      6) Machines+Basics(-)
Choose the setting to process: 6
```

3. Creating a procedure resource

Before upgrading [PRIMECLUSTER 4.1A30]

To create or change a procedure resource, select "CRM" from the "turnkey wizard "CRM"" menu, and then select the resource class name.

Note: Shown below is an example of registering a procedure resource of the BasicApplication class to a userApplication.

```
Settings of turnkey wizard "CRM"
 1) HELP                            10) LocalFileSystems(-)
 2) -                               11) RemoteFileSystems(-)
 3) SAVE+EXIT                       12) IpAddresses(-)
 4) -                               13) Gds:Global-Disk-Services(-)
 5) ApplicationName=APP1            14) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)
 7) CommandLines(-)
 8) Controllers(-)
 9) CRM(-)
Choose the setting to process: 9
```

↓

```
Crm (CRM_APP1:not yet consistent)
 1) HELP
 2) -
 3) SAVE+EXIT
 4) REMOVE+EXIT
 5) SystemState2(-)
 6) SystemState3(-)
 7) BasicApplication(-)
 8) Application(-)
Choose the setting to process: 7
```

After upgrading [PRIMECLUSTER 4.3A10]

The "turnkey wizard "CRM"" menu is not displayed.

To create or change a procedure resource, select "Procedure:resource-class-name" from the "turnkey wizard STANDBY" menu.

Note: Shown below is an example of registering a procedure resource of the BasicApplication class to a userApplication.

```
Settings of turnkey wizard "STANDBY" (APP1:not yet consistent)
 1) HELP                             9) Procedure:BasicApplication(-)
 2) -                               10) Symfoware(-)
 3) SAVE+EXIT                       11) Procedure:SystemState3(-)
 4) -                               12) Procedure:SystemState2(-)
 5) ApplicationName=APP1            13) Gls:Global-Link-Services(-)
 6) Machines+Basics(app1)           14) IpAddresses(-)
 7) CommandLines(-)                 15) LocalFileSystems(-)
 8) Procedure:Application(-)        16) Gds:Global-Disk-Services(-)
Choose the setting to process:9
```

4. Changing the timeout value of a procedure resource

  Before upgrading [PRIMECLUSTER 4.1A30]

    To change the timeout value, select "Flags[0]=OTtimeout-time" and then select "SCRIPTTIMEOUT."

```
Settings of "Application"
 1) HELP               4) REMOVE+RETURN        7) Flags[0]=OT1800
 2) NO-SAVE+RETURN     5) AdditionalResource   8) Priority[0]=65535
 3) SAVE+RETURN        6) ResourceName[0]=rscl
Choose the setting to process: 7
```

↓

```
Set Flags[0]:Currently set: STANDBYCAPABLE,SCRIPTTIMEOUT (OT1800)
 1) HELP               5) NULLDETECTOR(D)        9) MONITORONLY(M)
 2) -                  6) LIEOFFLINE(L)         10) NOT:STANDBYCAPABLE(O)
 3) SAVE+RETURN        7) CLUSTEREXCLUSIVE(X)   11) SCRIPTTIMEOUT(T)
 4) DEFAULT            8) AUTORECOVER(A)
Choose one of the flags: 11
```

↓

```
 1) HELP
 2) RETURN
 3) FREECHOICE
 4) 1800
Set an appropriate script timeout:
```

After upgrading [PRIMECLUSTER 4.3A10]

"Flags[0]=OTtimeout-time" is not displayed.

To change the timeout value, select "TIMEOUT."

```
Set flags for Procedure resource :
Currently set: TIMEOUT (T1800), PRIORITY (P65535)
 1) HELP
 2) -
 3) SAVE+RETURN
 4) TIMEOUT
 5) PRIORITY
Choose one of the flags: 4
```

↓

```
 1) HELP
 2) RETURN
 3) FREECHOICE
 4) 1800
Set an appropriate timeout:
```

5. Changing the priority in a resource class of a procedure resource.

Before upgrading [PRIMECLUSTER 4.1A30]

To change the priority in a resource class, select "Priority[0]=*priority*."

```
Settings of "Application"
 1) HELP              4) REMOVE+RETURN       7) Flags[0]=OT1800
 2) NO-SAVE+RETURN    5) AdditionalResource  8) Priority[0]=65535
 3) SAVE+RETURN       6) ResourceName[0]=rsc1
Choose the setting to process: 8
```

↓

```
 1) HELP
 2) RETURN
 3) FREECHOICE
Set a priority:
```

After upgrading [PRIMECLUSTER 4.3A10]

"Priority[0]=*priority*" is not displayed.

To change the priority in a resource class, select "PRIORITY."

```
Set flags for Procedure resource :
Currently set: TIMEOUT (T1800), PRIORITY (P65535)
 1) HELP
 2) -
 3) SAVE+RETURN
 4) TIMEOUT
 5) PRIORITY
Choose one of the flags: 5
```

↓

```
 1) HELP
 2) RETURN
 3) FREECHOICE
 4) 65535
Set an appropriate priority:
```

Note

None.

K.3.3 Registration and Deletion of Resources for PRIMECLUSTER-compatible Products

Details on incompatibilities

The procedures for registering and deleting resources for PRIMECLUSTER-compatible products have been changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A30]

1. Registering resources for PRIMECLUSTER-compatible products

   Execute multiple commands to register the following resources for PRIMECLUSTER-compatible products:

   - Resources of cluster services

- Resources of cluster service instances

2. Deleting resource for PRIMECLUSTER-compatible products

    Execute multiple commands to delete the following resources for PRIMECLUSTER-compatible products:

    - Resources of cluster services

    - Resources of cluster service instances

3. Reconfiguring a cluster application

    Delete the cluster service, and then register the cluster service again.

After upgrading [PRIMECLUSTER 4.3A10]

1. Registering resources for PRIMECLUSTER-compatible products

    Execute the "clrwzconfig" command to register the following resources for PRIMECLUSTER-compatible products:

    - Resources of cluster services

    - Resources of cluster service instances

2. Deleting resources for PRIMECLUSTER-compatible products

    Execute the "clrwzconfig" command to delete the following resources for PRIMECLUSTER-compatible products:

    - Resources of cluster services

    - Resources of cluster service instances

3. Reconfiguring a cluster application

    The cluster service must be registered when a cluster application is reconfigured.

    Use the "clrwzconfig" command to check whether the cluster service needs to be set up again. If re-setup of the cluster service is necessary, execute the "clrwzconfig" command.

Note

The following operations have not been changed:

- Registering a resource for IP address takeover and associating a resource to a cluster service instance

- Deleting a resource for IP address takeover

K.3.4 Setting Up Fsystem Resources

Details on incompatibilities

The file which defines the mount point of the filesystem used as Fsystem resource has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A30]

   The mount point was defined in /etc/fstab.

After upgrading [PRIMECLUSTER 4.3A10]

   It is necessary to define the mount point in /etc/fstab.pcl.

   For details, see "6.7.3.2 Setting Up Fsystem Resources."

Note

None.

K.3.5 Client Environment for Web-Based Admin View

Details on incompatibilities

Linux(R) is not supported as a client environment for Web-Based Admin View by PRIMECLUSTER 4.3A10.

Changes

Before upgrading [PRIMECLUSTER 4.1A30]

Linux(R) is supported as a client environment for Web-Based Admin View.

After upgrading [PRIMECLUSTER 4.3A10]

Linux(R) is not supported as a client environment for Web-Based Admin View.

Note

None.

K.3.6 Changes of the Behavior of CF Startup

Details on incompatibilities

CF starts even if some of the network interfaces for the cluster interconnects are not recognized.

Changes

Before upgrading [PRIMECLUSTER 4.1A30]

CF does not start unless all of the network interfaces for the cluster interconnects are recognized.

After upgrading [PRIMECLUSTER 4.3A10]

CF starts if at least one of the network interfaces for the cluster interconnects is recognized.

Note

If there are any network interfaces that are not recognized on CF startup, the following message appears:

CF: <NIC>: device not found.

<NIC> will be the name of the network interface (e.g. eth0).

This message is also available in 4.1A30.

K.3.7 HV_CONNECT_TIMEOUT

Details on incompatibilities

The default value of the RMS local environment variables HV_CONNECT_TIMEOUT is changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A30]

The default value of HV_CONNECT_TIMEOUT is 5 (seconds).

After upgrading [PRIMECLUSTER 4.3A10]

The default value of HV_CONNECT_TIMEOUT is 30 (seconds).

Note

For details on HV_CONNECT_TIMEOUT, see "12.3 Local environment variables" in the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide".

K.3.8 Changes of the RMS message

Details on incompatibilities

The RMS message (SYS, 8) logged in the syslog have been changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A30]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung.

After upgrading [PRIMECLUSTER 4.3A10]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung. An operator intervention is required.

Note

None.

K.3.9 Changes of the importance of the message in the RMS wizard

Details on incompatibilities

The importance of the following message in the RMS wizard has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A30]

WARNING: cannot grab mount lock for dostat() check_getbdev(), returning previous state

After upgrading [PRIMECLUSTER 4.3A10]

NOTICE: cannot grab mount lock for dostat() check_getbdev(), returning previous state

Note

None.

K.4 Changes in PRIMECLUSTER 4.3A10 from 4.1A40

Incompatible functions

The following functions of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.1A40.

- K.4.1 Setting Up Fsystem Resources
- K.4.2 Client Environment for Web-Based Admin View
- K.4.3 Changes of the Behavior of CF Startup
- K.4.4 HV_CONNECT_TIMEOUT

Incompatible messages

The following messages of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.1A40.

- K.4.5 Changes of the RMS message
- K.4.6 Changes of the importance of the message in the RMS wizard

K.4.1 Setting Up Fsystem Resources

Details on incompatibilities

The file which defines the mount point of the filesystem used as Fsystem resource has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A40]

The mount point was defined in /etc/fstab.

After upgrading [PRIMECLUSTER 4.3A10]

It is necessary to define the mount point in /etc/fstab.pcl.

For details, see "6.7.3.2 Setting Up Fsystem Resources."

Note

None.

K.4.2 Client Environment for Web-Based Admin View

Details on incompatibilities

Linux(R) is not supported as a client environment for Web-Based Admin View by PRIMECLUSTER 4.3A10.

Changes

Before upgrading [PRIMECLUSTER 4.1A40]

Linux(R) is supported as a client environment for Web-Based Admin View.

After upgrading [PRIMECLUSTER 4.3A10]

Linux(R) is not supported as a client environment for Web-Based Admin View.

Note

None.

K.4.3 Changes of the Behavior of CF Startup

Details on incompatibilities

CF starts even if some of the network interfaces for the cluster interconnects are not recognized.

Changes

Before upgrading [PRIMECLUSTER 4.1A40]

CF does not start unless all of the network interfaces for the cluster interconnects are recognized.

After upgrading [PRIMECLUSTER 4.3A10]

CF starts if at least one of the network interfaces for the cluster interconnects is recognized.

Note

If there are any network interfaces that are not recognized on CF startup, the following message appears:

CF: <NIC>: device not found.

<NIC> will be the name of the network interface (e.g. eth0).

This message is also available in 4.1A40.

K.4.4 HV_CONNECT_TIMEOUT

Details on incompatibilities

The default value of the RMS local environment variables HV_CONNECT_TIMEOUT is changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A40]

The default value of HV_CONNECT_TIMEOUT is 5 (seconds).

After upgrading [PRIMECLUSTER 4.3A10]

The default value of HV_CONNECT_TIMEOUT is 30 (seconds).

Note

For details on HV_CONNECT_TIMEOUT, see "12.3 Local environment variables" in the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide".

K.4.5 Changes of the RMS message

Details on incompatibilities

The RMS message (SYS, 8) logged in the syslog have been changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A40]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung.

After upgrading [PRIMECLUSTER 4.3A10]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung. An operator intervention is required.

Note

None.

K.4.6 Changes of the importance of the message in the RMS wizard

Details on incompatibilities

The importance of the following message in the RMS wizard has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.1A40]

WARNING: cannot grab mount lock for dostat() check_getbdev(), returning previous state

After upgrading [PRIMECLUSTER 4.3A10]

NOTICE: cannot grab mount lock for dostat() check_getbdev(), returning previous state

Note

None.

K.5 Changes in PRIMECLUSTER 4.3A10 from 4.2A00

Incompatible functions

The following functions of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.2A00.

- K.5.1 Setting Up Fsystem Resources

- K.5.2 Client Environment for Web-Based Admin View

- K.5.3 Changes of the Behavior of CF Startup

- K.5.4 HV_CONNECT_TIMEOUT

Incompatible messages

The following messages of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.2A00.

- K.5.5 Changes of the RMS message

- K.5.6 Changes of the importance of the message in the RMS wizard

K.5.1   Setting Up Fsystem Resources

Details on incompatibilities

The file which defines the mount point of the filesystem used as Fsystem resource has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.2A00]

The mount point was defined in /etc/fstab.

After upgrading [PRIMECLUSTER 4.3A10]

It is necessary to define the mount point in /etc/fstab.pcl.

For details, see "6.7.3.2 Setting Up Fsystem Resources."

Note

None.

K.5.2   Client Environment for Web-Based Admin View

Details on incompatibilities

Linux(R) is not supported as a client environment for Web-Based Admin View by PRIMECLUSTER 4.3A10.

Changes

Before upgrading [PRIMECLUSTER 4.2A00]

Linux(R) is supported as a client environment for Web-Based Admin View.

After upgrading [PRIMECLUSTER 4.3A10]

Linux(R) is not supported as a client environment for Web-Based Admin View.

Note

None.

K.5.3   Changes of the Behavior of CF Startup

Details on incompatibilities

CF starts even if some of the network interfaces for the cluster interconnects are not recognized.

Changes

Before upgrading [PRIMECLUSTER 4.2A00]

CF does not start unless all of the network interfaces for the cluster interconnects are recognized.

After upgrading [PRIMECLUSTER 4.3A10]

CF starts if at least one of the network interfaces for the cluster interconnects is recognized.

Note

If there are any network interfaces that are not recognized on CF startup, the following message appears:

CF: <NIC>: device not found.

<NIC> will be the name of the network interface (e.g. eth0).

This message is also available in 4.2A00.

K.5.4 HV_CONNECT_TIMEOUT

Details on incompatibilities

The default value of the RMS local environment variables HV_CONNECT_TIMEOUT is changed.

Changes

Before upgrading [PRIMECLUSTER 4.2A00]

The default value of HV_CONNECT_TIMEOUT is 5 (seconds).

After upgrading [PRIMECLUSTER 4.3A10]

The default value of HV_CONNECT_TIMEOUT is 30 (seconds).

Note

For details on HV_CONNECT_TIMEOUT, see "12.3 Local environment variables" in the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide".

K.5.5 Changes of the RMS message

Details on incompatibilities

The RMS message (SYS, 8) logged in the syslog have been changed.

Changes

Before upgrading [PRIMECLUSTER 4.2A00]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung.

After upgrading [PRIMECLUSTER 4.3A10]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung. An operator intervention is required.

Note

None.

K.5.6 Changes of the importance of the message in the RMS wizard

Details on incompatibilities

The importance of the following message in the RMS wizard has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.2A00]

WARNING: cannot grab mount lock for dostat() check_getbdev(), returning previous state

After upgrading [PRIMECLUSTER 4.3A10]

NOTICE: cannot grab mount lock for dostat() check_getbdev(), returning previous state

Note

None.

K.6   Changes in PRIMECLUSTER 4.3A10 from 4.2A30

Incompatible functions

The following functions of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.2A30.

- K.6.1 Client Environment for Web-Based Admin View

- K.6.2 Changes of the Behavior of CF Startup

- K.6.3 HV_CONNECT_TIMEOUT

Incompatible messages

The following messages of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.2A30.

- K.6.4 Changes of the RMS message

- K.6.5 Changes of the importance of the message in the RMS wizard

K.6.1   Client Environment for Web-Based Admin View

Details on incompatibilities

Linux(R) is not supported as a client environment for Web-Based Admin View by PRIMECLUSTER 4.3A10.

Changes

Before upgrading [PRIMECLUSTER 4.2A30]

Linux(R) is supported as a client environment for Web-Based Admin View.

After upgrading [PRIMECLUSTER 4.3A10]

Linux(R) is not supported as a client environment for Web-Based Admin View.

Note

None.

K.6.2   Changes of the Behavior of CF Startup

Details on incompatibilities

CF starts even if some of the network interfaces for the cluster interconnects are not recognized.

Changes

Before upgrading [PRIMECLUSTER 4.2A30]

CF does not start unless all of the network interfaces for the cluster interconnects are recognized.

After upgrading [PRIMECLUSTER 4.3A10]

CF starts if at least one of the network interfaces for the cluster interconnects is recognized.

Note

If there are any network interfaces that are not recognized on CF startup, the following message appears:

CF: <NIC>: device not found.

<NIC> will be the name of the network interface (e.g. eth0).

This message is also available in 4.2A30.

K.6.3 HV_CONNECT_TIMEOUT

Details on incompatibilities

The default value of the RMS local environment variables HV_CONNECT_TIMEOUT is changed.

Changes

Before upgrading [PRIMECLUSTER 4.2A30]

The default value of HV_CONNECT_TIMEOUT is 5 (seconds) in the RHEL-AS environment, and 30 (seconds) in the RHEL5 environment.

After upgrading [PRIMECLUSTER 4.3A10]

The default value of HV_CONNECT_TIMEOUT is 30 (seconds).

Note

There are no incompatibilities when upgrading PRIMECLUSTER from 4.2A30 for RHEL5 to 4.3A10.

For details on HV_CONNECT_TIMEOUT, see "12.3 Local environment variables" in the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide".

K.6.4 Changes of the RMS message

Details on incompatibilities

The RMS message (SYS, 8) logged in the syslog have been changed.

Changes

Before upgrading [PRIMECLUSTER 4.2A30]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung.

After upgrading [PRIMECLUSTER 4.3A10]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung. An operator intervention is required.

Note

None.

K.6.5 Changes of the importance of the message in the RMS wizard

Details on incompatibilities

The importance of the following message in the RMS wizard has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.2A30]

WARNING: cannot grab mount lock for dostat() check_getbdev(), returning previous state

After upgrading [PRIMECLUSTER 4.3A10]

NOTICE: cannot grab mount lock for dostat() check_getbdev(), returning previous state

Note

None.

K.7 Changes in PRIMECLUSTER 4.3A10 from 4.3A00

Incompatible functions

The following functions of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.3A00.

- K.7.1 Client Environment for Web-Based Admin View

- K.7.2 Changes of the Behavior of CF Startup

- K.7.3 HV_CONNECT_TIMEOUT

Incompatible messages

The following messages of PRIMECLUSTER 4.3A10 are incompatible with PRIMECLUSTER 4.3A00.

- K.7.4 Changes of the RMS message

- K.7.5 Changes of the importance of the message in the RMS wizard

K.7.1 Client Environment for Web-Based Admin View

Details on incompatibilities

Linux(R) is not supported as a client environment for Web-Based Admin View by PRIMECLUSTER 4.3A10.

Changes

Before upgrading [PRIMECLUSTER 4.3A00]

Linux(R) is supported as a client environment for Web-Based Admin View.

After upgrading [PRIMECLUSTER 4.3A10]

Linux(R) is not supported as a client environment for Web-Based Admin View.

Note

None.

K.7.2 Changes of the Behavior of CF Startup

Details on incompatibilities

CF starts even if some of the network interfaces for the cluster interconnects are not recognized.

Changes

Before upgrading [PRIMECLUSTER 4.3A00]

CF does not start unless all of the network interfaces for the cluster interconnects are recognized.

After upgrading [PRIMECLUSTER 4.3A10]

CF starts if at least one of the network interfaces for the cluster interconnects is recognized.

Note

If there are any network interfaces that are not recognized on CF startup, the following message appears:

CF: <NIC>: device not found.

<NIC> will be the name of the network interface (e.g. eth0).

This message is also available in 4.3A00.

K.7.3 HV_CONNECT_TIMEOUT

Details on incompatibilities

The default value of the RMS local environment variables HV_CONNECT_TIMEOUT is changed.

Changes

Before upgrading [PRIMECLUSTER 4.3A00]

The default value of HV_CONNECT_TIMEOUT is 5 (seconds) in the RHEL-AS environment, and 30 (seconds) in the RHEL5 environment.

After upgrading [PRIMECLUSTER 4.3A10]

The default value of HV_CONNECT_TIMEOUT is 30 (seconds).

Note

There are no incompatibilities when upgrading PRIMECLUSTER from 4.3A00 for RHEL5 to 4.3A10.

For details on HV_CONNECT_TIMEOUT, see "12.3 Local environment variables" in the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide".

K.7.4 Changes of the RMS message

Details on incompatibilities

The RMS message (SYS, 8) logged in the syslog have been changed.

Changes

Before upgrading [PRIMECLUSTER 4.3A00]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung.

After upgrading [PRIMECLUSTER 4.3A10]

(SYS, 8): ERROR: RMS failed to shut down the host <host> via a Shutdown Facility, no further kill functionality is available.

The cluster is now hung. An operator intervention is required.

Note

None.

K.7.5 Changes of the importance of the message in the RMS wizard

Details on incompatibilities

The importance of the following message in the RMS wizard has been changed.

Changes

Before upgrading [PRIMECLUSTER 4.3A00]

WARNING: cannot grab mount lock for dostat() check_getbdev(), returning previous state

After upgrading [PRIMECLUSTER 4.3A10]

NOTICE: cannot grab mount lock for dostat() check_getbdev(), returning previous state

Note

None.

Appendix L   Release Information

This chapter explains the outline, details, and the part changed with PRIMECLUSTER 4.3.

No	Version / Level	Outline	Details	Change Location
1	4.3	Change of supported hard device	The supported hard device was changed as follows. PRIMERGY RX300 S3, TX200 S3, RX600 S3, RX300 S4, TX300 S4, BX620 S4	- 2.1.1 Product Selection - 3.1.4 Setting Up the Cluster High-Speed Failover Function - 5.1.2 Setting Up the Shutdown Facility - A.5 Cluster Installation Environment Worksheet - A.11 System Design Examples (for the Virtual Machine Function not used) - A.11.1.1 Cluster Configuration Worksheet - B.7 Shutdown Facility (SF)
2	4.3	Change of supported OS	Supported OS was changed to RHEL4.6 and RHEL5.1.	- 3.1.4 Setting Up the Cluster High-Speed Failover Function - 6.2.1 GLS Setup - A.5 Cluster Installation Environment Worksheet - A.11 System Design Examples (for the Virtual Machine Function not used)
3	4.3	Addition of exit code	The exit code of Check script was added.	- 6.6 Setting Up Online/Offline Scripts
4	4.3	Correction of setup value of RMS automatic startup	The setup value of RMS automatic startup (HV_RCSTART) was changed from "Does not start up automatically" to "Start up automatically".	- 6.1 Initial RMS Setup
5	4.3	Addition of setting of GLS in RHEL5.1 environment	The procedure which nullified the setting of the Hot Plug in the RHEL5.1 environment was added to the operational procedure.	- 6.2.1 GLS Setup
6	4.3	Addition of attention by configuration setting of GDS	Attention by the configuration setting of GDS was added.	- Chapter 5 Building a Cluster - 6.3 GDS Configuration Setup
7	4.3 Second edition	Addition of supported hard device	RX600 S4 was added.	- 5.1.2 Setting Up the Shutdown Facility
8	4.3 Second edition	Addition of supported OS	RHEL5.2 has been added.	- 3.1.4 Setting Up the Cluster High-Speed Failover Function - A.5 Cluster Installation Environment Worksheet

No	Version / Level	Outline	Details	Change Location
9	4.3 Third edition	Change of company name	Changed company name to "Fujitsu Technology Solutions"	- Throughout document
10	4.3 Fourth edition	Addition of supported OS	RHEL AS 4.7 and RHEL5.3 have been added.	- 3.1.4 Setting Up the Cluster High-Speed Failover Function - 5.1.2 Setting Up the Shutdown Facility - A.5 Cluster Installation Environment Worksheet
11	4.3 Fourth edition	Addition of supported hard device	The supported hard device was added as follows. PRIMERGY BX920 S1, RX200 S5, TX300 S5, BX620 S5	- 5.1.2 Setting Up the Shutdown Facility
12	4.3 Fifth edition	Addition of description for maintenance mode	Descriptions of "7.2.2.6 Entering maintenance mode for Cluster Application" were added.	- 7.2.2.6 Entering maintenance mode for Cluster Application
13	4.3 Fifth edition	Addition of supported OS	The supported OS was added as follows. RHEL-AS 4.8, RHEL5.4	- 3.1.4 Setting Up the Cluster High-Speed Failover Function - 5.1.2 Setting Up the Shutdown Facility - A.5 Cluster Installation Environment Worksheet
14	4.3 Fifth edition	Addition of supported hard device	RX300 S5 was added.	- 5.1.2 Setting Up the Shutdown Facility
15	4.3 Fifth edition	Addition of the virtual machine function (cluster between guest operating systems)	The function to operate guest operating systems in the virtual machine function as a cluster system was added.	- 1.1 Planning - 2.2.1 Virtual Machine Function - 3.2 Installation and Setup of Related Software (for the Virtual Machine Function used) - Chapter 5 Building a Cluster - 5.1.2 Setting Up the Shutdown Facility - 5.1.2.1 Checking the Shutdown Agent Information - VMGuest (virtual machine) - 8.9 Changing Virtual Machine Settings - A.7 Cluster Configuration Worksheet - A.12 System Design Examples (for the Virtual Machine Function used in Xen Environment) - B.7 Shutdown Facility (SF) - C.1.2 Crash Dump - D.1 Searching for a Message

No	Version / Level	Outline	Details	Change Location
16	4.3 Sixth edition	Addition of supported hard device	BX960 was added.	- 5.1.2 Setting Up the Shutdown Facility - 5.1.2.3.4 kdump
17	4.3 Sixth edition	Change of supported OS	Supported OS was changed to RHEL5.3 or later, and RHEL6.X.	- 6.2.1 GLS Setup - A.5 Cluster Installation Environment Worksheet
18	4.3 Sixth edition	Addition of description along with addition of supported OS	Descriptions about RHEL6 were added.	- 5.2 Setting Up Fault Resource Identification and Operator Intervention Request
19	4.3 Sixth edition	Addition and change of description for the virtual machine function	Descriptions when the cluster is operated between guest OSes on multiple host OSes with the virtual machine function were added and changed.	- 1.1 Planning - 2.2 System Design - 2.2.1 Virtual Machine Function - 3.2.3 When building a cluster system between guest OSes on multiple host OSes using Host OS failover function - 5.1.2 Setting Up the Shutdown Facility - 5.1.2.5.1 vmSP (Virtual Machine Service Provider) - 7.2 Operating the PRIMECLUSTER System - 8.9.1 Xen environment - A.12 System Design Examples (for the Virtual Machine Function used in Xen Environment)
20	4.3 Sixth edition	Change of the path name of the GFS shared file system	The file name of the GFS shared file system was changed.	- 6.4 Initial GFS Setup
21	4.3 Sixth edition	Addition of "physical machine"	The explanation of "physical machine" was added.	- Glossary
22	4.3 Sixth edition	Addition of Configuration Update Service for SA	Descriptions of Configuration Update Service for SA were added.	- Appendix G Configuration Update Service for SA
23	4.3 Sixth edition	Addition and change of description for using VMware environment	Descriptions for using the VMware environment with the virtual machine function were added.	- 2.2 System Design - 2.2.1 Virtual Machine Function - 3.2 Installation and Setup of Related Software (for the Virtual Machine Function used) - A.12 System Design Examples (for the Virtual Machine Function used in Xen Environment) - Appendix H Using PRIMECLUSTER in a VMware Environment
24	4.3 Sixth edition	Change of the settings for the CF node name	Settings for the CF node name were changed.	- 5.1.1 Setting Up CF and CIP

No	Version / Level	Outline	Details	Change Location
25	4.3 Sixth edition	Addition of note	Notes were added.	- 5.1.2.3.1 Setting up the shutdown daemon - 5.1.2.3.4 kdump - 6.6 Setting Up Online/Offline Scripts - 6.7.3.2 Setting Up Fsystem Resources - 7.2.2.1 Starting a Cluster Application - 7.2.2.6 Entering maintenance mode for Cluster Application - 7.4.1.1 Failure Detection and Cause Identification if a Failure Occurs -
26	4.3 Sixth edition	Addition of procedure to change the MMB IP address	Procedures to change the MMB IP address were added.	- 8.8.1 Changing the MMB IP Address
27	4.3 Sixth edition	Addition of procedure to change IPMI settings	Procedures to change IPMI settings were added.	- 8.10 Changing IPMI Settings
28	4.3 Sixth edition	Addition of procedure to change Blade settings	Procedures to change Blade settings were added.	- 8.11 Changing Blade Settings
29	4.3 Sixth edition	Addition of procedure to add, delete, and change hardware	Procedures to add, delete, and change hardware were added.	- 8.12 Adding, Deleting, and Changing Hardware
30	4.3 Sixth edition	Addition of procedure to change the startup priority of a state transition procedure	Procedures to change the startup priority of a state transition procedure were added.	- E.2 Changing a Procedure Resource
31	4.3 Sixth edition	Addition of test	The test for forced shutdown of cluster nodes was added.	- 1.4 Test
32	4.3 Sixth edition	Deletion of Diskdump	Descriptions of Diskdump were deleted.	- 3.1.4.1 PRIMERGY - 5.1.2.3.2 IPMI - 5.1.2.3.3 Blade - A.3 Overall Design Worksheet - A.7 Cluster Configuration Worksheet - A.11.1.1 Cluster Configuration Worksheet - C.1.2 Crash Dump - Glossary
33	4.3 Sixth edition	Addition of check for the cluster environment	Descriptions of check for the cluster environment were added.	- Chapter 6 Building Cluster Applications - 6.9 Checking the Cluster Environment
34	4.3 Sixth edition	Change of the name of Installation Guide	The name of Installation Guide was changed to Software Release Guide.	- 3.1.2 Setting Up Disk Units - 3.2.1.1 Host OS setup - 3.2.2.1 Host OS setup

No	Version / Level	Outline	Details	Change Location
				- 3.4 Installation and Environment Setup of Applications
35	4.3 Sixth edition	Change of procedures	Procedures to register a state transition procedure were added.	- E.2 Changing a Procedure Resource
36	4.3 Sixth edition	Deletion of PRIMEQUEST 500/400 series	Descriptions about PRIMEQUEST 500/400 series were deleted.	- 2.2.1 Virtual Machine Function - 3.1.4.2 PRIMEQUEST - 5.1.2 Setting Up the Shutdown Facility - 5.1.2.1 Checking the Shutdown Agent Information - VMGuest (virtual machine) - PRIMEQUEST 500/400 series - C.1.3 PSA Information
37	4.3 Sixth edition	Deletion of description for NetScape	Descriptions about NetScape were deleted.	- 4.4 Starting the Web-Based Admin View Screen - 4.6 Exiting the Web-Based Admin View Screen - A.11.2 PRIMEQUEST - A.12.1 When building a cluster system between guest OSes on one host OS
38	4.3 Sixth edition	Addition of incompatibilities	Incompatibilities for the client environment for Web-Based Admin View were added.	- K.7 Changes in PRIMECLUSTER 4.3A10 from 4.3A00
39	4.3 Sixth edition	Change of the number of supported nodes	Changed the number of supported nodes.	- 1.1 Planning
40	4.3 Sixth edition	Addition and change of description for using KVM environment	Descriptions for using a KVM environment with the virtual machine function were added and chanted.	- 2.2 System Design - 2.2.1 Virtual Machine Function - 3.2 Installation and Setup of Related Software (for the Virtual Machine Function used) - 5.1.2 Setting Up the Shutdown Facility - 6.3.1 Setting Up System Disk Mirroring - 8.9 Changing Virtual Machine Settings - A.7 Cluster Configuration Worksheet - A.13 System Design Examples (for the Virtual Machine Function used in KVM Environment) - C.1 Collecting Troubleshooting Information
41	4.3 Sixth edition	Change of PSA to PSA/ SVmco	Changed PSA to PSA/ SVmco.	- 3.1.4.2 PRIMEQUEST

No	Version / Level	Outline	Details	Change Location
42	4.3 Sixth edition	Addition of recovery procedure for a failure of the cluster interconnect	Added the description to recover the cluster interconnect when an error occurs.	- 7.4.1.3 Recovery of Failed Cluster Interconnect
43	4.3 Sixth edition	Addition of server models	Added the supported server models.	- C.1.3 PSA Information
44	4.3 Sixth edition	Addition of SVmco information	Added the description on how to collect the SVmco information.	- C.1.4 SVmco Information
45	4.3 Sixth edition	Addition of incompatible messages	Added incompatible messages.	- Appendix K Changes in Each Version
46	4.3 Sixth edition	Addition of incompatible functions	Added incompatible functions for the Changes of the Behavior of CF Startup.	- Appendix K Changes in Each Version

Glossary

AC (Access Client)

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*.

API (application program interface)

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 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.

linuxBMC (Baseboard Management Controller)

A dedicated processor for monitoring and diagnosis of environmental factors (e.g. temperature, voltage) and parts and units.

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 Monito

CIP

Cluster Interconnect Protocol

class (GDS)

See *disk class*.

CLI

command-line interface

cluster

A set of computers that work together as a single computing source. Specifically, a cluster performs a distributed form of parallel computing.

See also *RMS configuration*.

Cluster Foundation

The set of PRIMECLUSTER modules that provides basic clustering communication services.

See also *base cluster foundation*.

cluster interconnect (CF)

The set of private network connections used exclusively for PRIMECLUSTER communications.

cluster partition

The state in which communication with some of the nodes that constitute the cluster has been stopped.

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 only.)

See also *simple virtual disk, striped 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.

Crash dump collection facility

The facility that collects crash dumps if an OS error (panic, etc.) occurs. The crash dump collection facility includes kdump.

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

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 (Event Notification Services)(CF)

See *Event Notification Services*.

environment variable (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 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 Linux(R) 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*.

See *disk group*.

GDS (Global Disk Services)

See *Global Disk Services*.

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 (Global File Services)

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 Linux(R) systems that are connected to the shared disk device, 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 (Global Link Service)

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*.

GUI (graphical user interfac)

See *graphical user interface*.

HA (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

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.

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*.

I/F

Interface

I/O

input/output

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 networ)

See *public LAN*.

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*.

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*.

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 (Meta Data Server)

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

mirror group (GDS)

A disk group of the mirror type. This is a collection of mutually mirrored disks or low-order groups.

mirror volume (GDS)

A volume that is created in a mirror group. Data redundancy is created by mirroring.

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.

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.

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

LAN duplexed mode that is provided by GLS. The duplexed NIC is used exclusively, and LAN monitoring between the Linux(R) 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 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 (Oracle Parallel Server)

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.

OSD (operating system dependent) (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).

physical machine

A server configured with actual hardware. This is used in contrast with a virtual machine, and is also referred to as a physical server.

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 LEFCLUSTER node).

RAO

RMS-Add on

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 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 as displayed in a system graph.

resource state (RMS)

Current state of a resource.

RMS (Reliant Monitor Services)

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*.

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.

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.

SA

Shutdown Agent. SA forcibly stops the target node by receiving instructions from the Shutdown Facility.

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.

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, a 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 devices 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 reach 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*.

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 not belong to a group. There is no data redundancy.

SIS (Scalable Internet Services)

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*, *simple 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.

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 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

A disk accessible from a virtual machine.

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

Index