FUJITSU

Fujitsu Software
PRIMECLUSTER

Installation and Administration Guide 4.7 Cloud Services

Linux

J2UL-2821-02ENZ0(00)
July 2023

Preface

This manual serves as your starting point for using PRIMECLUSTER in a cloud environment.

This manual explains the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system in a cloud environment.

Target Readers

This manual is intended for all users who use PRIMECLUSTER 4.7 and perform cluster system installation and operation management in a cloud environment. It is also intended for programmers who develop applications that operate on PRIMECLUSTER.

Configuration of This Documentation

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

Part 1 FJcloud-O Environment

Audience: System administrators who build the PRIMECLUSTER system in an FJcloud-O environment
Contents: This part describes the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system in an FJcloud-O environment.

Part 2 NIFCLOUD Environment

Audience: System administrators who build the PRIMECLUSTER system in a NIFCLOUD environment
Contents: This part describes the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system in a NIFCLOUD environment.

Part 3 FJcloud-Baremetal Environment

Audience: System administrators who build the PRIMECLUSTER system in an FJcloud-Baremetal environment
Contents: This part describes the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system in an FJcloud-Baremetal environment.

Part 4 AWS Environment

Audience: System administrators who build the PRIMECLUSTER system in an AWS environment
Contents: This part describes the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system in an AWS environment.

Part 5 Azure Environment

Audience: System administrators who build the PRIMECLUSTER system in an Azure environment
Contents: This part describes the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system in an Azure environment.

Appendix A Using the smart workload recovery feature

Audience: System administrator who creates the PRIMECLUSTER system that uses the smart workload recovery facility.
Contents: This part describes the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system using the smart workload recovery feature provided by PRIMECLUSTER Cloud Edition.

Appendix B Changes in Each Version

Audience: All users who use PRIMECLUSTER 4.6A20 or earlier
Contents: This appendix describes the changes made to the specifications of PRIMECLUSTER 4.7A00.

Appendix C Release Information

Audience: All users who use PRIMECLUSTER systems
Contents: This appendix lists the main changes in this manual.

Glossary

Audience: All users who use the PRIMECLUSTER system
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 Installation and Administration Guide
- PRIMECLUSTER Web-Based Admin View Operation Guide
- PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide
- PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide
- PRIMECLUSTER Global Disk Services Configuration and Administration Guide
- PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function
- PRIMECLUSTER Messages
- FJQSS (Information Collection Tool) User's Guide

Note

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

- PRIMECLUSTER Software Release Guide and Installation Guide

  These are the software release guide and the installation guide provided with each PRIMECLUSTER product. The data is stored on "DVD" of each package. For details on the file names, see the documentation.

Manual Printing

If you want to print a manual, use the PDF file found on the DVD 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, refer to 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, refer to "4.3.1 Assigning Users to Manage the Cluster" in "PRIMECLUSTER Installation and Administration Guide."

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).

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.
<Constant Width>	Variables that you must replace with an actual displayed value.
Bold	Items in a command line that you must type exactly as shown.
"Constant Width"	The title, documentation, screen, and etc of lookup destination.
[Constant Width]	Tool bar name, menu name, command name, button name, and icon names.

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.

See

Provides the names of manuals to be referenced.

Abbreviations

- Red Hat Enterprise Linux is abbreviated as RHEL.

- FUJITSU Hybrid IT Service FJcloud-O is abbreviated as FJcloud-O.
- FUJITSU Hybrid IT Service FJcloud-V is abbreviated as FJcloud-V.
- FUJITSU Hybrid IT Service FJcloud-Baremetal is abbreviated as FJcloud-Baremetal.
- FUJITSU Hybrid IT Service for Microsoft Azure is abbreviated as "for Azure."
- FUJITSU Hybrid IT Service for AWS is abbreviated as "for AWS."
- FJcloud-V sold by FUJITSU LIMITED and NIFCLOUD sold by FUJITSU CLOUD TECHNOLOGIES LIMITED are abbreviated as "NIFCLOUD" in this manual.
- "for Azure" sold by FUJITSU LIMITED and Microsoft Azure sold by Microsoft Corporation in the United States are abbreviated as "Azure" in this manual.
- "for AWS" sold by FUJITSU LIMITED and AWS (Amazon Web Services) sold by Amazon.com, Inc. are abbreviated as "AWS" in this manual.

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

Requests

- No part of this documentation may be reproduced or copied without permission of FUJITSU LIMITED.
- The contents of this documentation may be revised without prior notice.

Date of publication and edition

February 2023, First edition
July 2023, Second edition

Copyright notice

Contents

Part 1 FJcloud-O Environment

This part describes the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system in an FJcloud-O environment.

Chapter 1 Cluster System in an FJcloud-O Environment

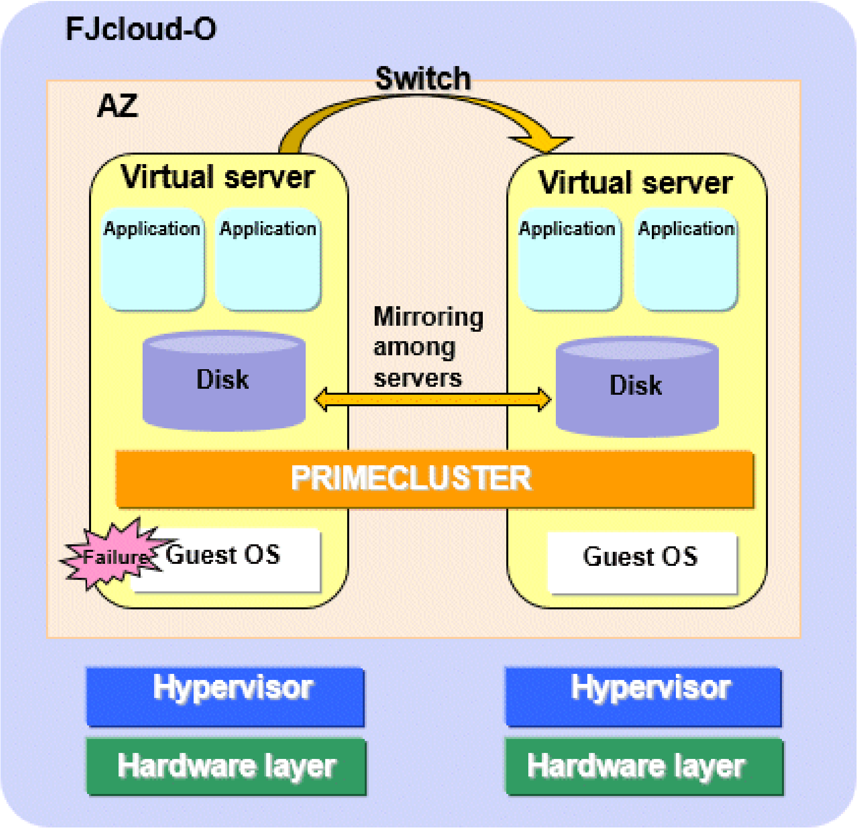
In an FJcloud-O environment, PRIMECLUSTER can be used on the virtual servers that are built within the same availability zone (hereinafter AZ). The mirroring among servers of Global Disk Services (hereinafter GDS) is used to take over the data between the virtual servers.

Compared to auto-scaling or automatic failover, PRIMECLUSTER provides the following advantages:

- The time to switch the standby virtual server from the operational virtual server is reduced because the standby VM is switched from the startup status of OS when a failure occurs in the virtual server.

- The standby virtual server can be switched with monitoring the operational application when a failure of the operational application occurs.

Figure 1.1 Cluster System in an FJcloud-O Environment



See

For details on FJcloud-O, refer to the documents provided by FJcloud-O.

1.1 Supported Range

This section explains the range of support of PRIMECLUSTER in an FJcloud-O environment.

Supported configurations

- Number of cluster nodes: 1 to 2 nodes

- Operation mode of the cluster system: 1:1 Standby operation, Mutual standby, Single-node cluster, Scalable operation (only when using Symfoware Server (Native) Hot Standby feature)

- Network configurations
    - The virtual servers and the management client in the cluster system must belong to the same availability zone and subnet.
    - For the cluster interconnect, its network must be independent from the network used with the administrative LAN, the public LAN, and the mirroring among the servers of GDS. This setting is not necessary in a single-node cluster.
    - The virtual servers in the cluster system must communicate with the API endpoints. This setting is not necessary in a single-node cluster.
- Security groups
    - One security group must be set among the virtual servers in the cluster system for a security reason.
    - Another security group must be set between the virtual servers and the management client in the cluster system.
    - The security group for the cluster interconnect must be set so that the node cannot communicate with any node outside of the cluster system. This setting is not necessary in a single-node cluster.

Supported monitoring functions

- Error of OS on the virtual server and the cluster interconnect

  The cyclic monitoring of the cluster interconnect (LAN) detects a hang-up of OS and the service is switched to the standby system.

- Error of the shared disk and the disk access path

  By combining the volume management function (GDS), a failure of the disk access and the disk access path can be detected (monitored by the Gds resource), and the service is switched to the standby system when the disk access is disabled or a failure of the whole system of the disk access path occurs.

- Error of the cluster application

  When a resource error of the cluster application occurs, the service is switched to the standby system.

Note

- PRIMECLUSTER cannot be used in the virtual server built in a separate availability zone.
- Set each virtual server so as to start it on each individual physical host.
- To take over the data from one virtual server to the other virtual server, setting the mirroring among the servers of GDS is required.
- To take over the IP address of the virtual server, setting the virtual NIC mode for Global Link Services (hereinafter GLS) is required.
- The snapshot to the virtual server can be acquired only when OS is stopped.
- The following functions for PRIMECLUSTER are not available:
    - Global File Services (hereinafter GFS)
    - GDS Snapshot
    - Root class and local class of GDS
    - Single volume of GDS, and disk groups of GDS other than the netmirror type (mirror, stripe, concatenation, and switch)
    - Cluster application using the takeover network resource
    - In addition to the above functions, the following functions for PRIMECLUSTER are not available in a single-node cluster.
        - GLS
        - GDS
- The auto-scaling function for FJcloud-O is not available.

- Can be used in conjunction with automatic failover function for FJcloud-O.

  [When the automatic failover function is enabled]

  If an error occurs on the physical host on which the operational virtual server is running, the operation is switched after the automatic failover operation (about 30 minutes to 60 minutes).

  [When the automatic failover function is disabled]

  If an error occurs on the physical host on which the operational virtual server is running, the operation does not switch automatically. Switching requires manual intervention. If an error occurs on a physical host, you must re-create the virtual server on the physical host on which the error occurred, regardless of whether it is on the operational or standby system. For this reason, take a snapshot of the virtual server in advance, and then use the snapshot to recover it when you re-create the virtual server.

  For cluster application continuity, we recommend that you enable the automatic failover function.

  For details on the automatic failover function, refer to the official FUJITSU Hybrid IT Service FJcloud documentation.

- Duplicate virtual server names cannot be used in the project on FJcloud-O.

- The console cannot be used in the virtual server in an FJcloud-O environment. Do not set the single user mode.

- In an FJcloud-O environment, when an OS panic occurs, the cluster node may be powered off while the memory dump is being output, and it may not be possible to collect a complete memory dump.

- Release the virtual server from the virtual server whose status is SHUTOFF (Stopped State).

Chapter 2 Design

You must prepare the items listed below before building the PRIMECLUSTER system in an FJcloud-O environment.

- Selecting the PRIMECLUSTER Product
- System Design
- Determining the Cluster System Operation Mode
- Determining the Web-Based Admin View Operation Mode
- Determining the Failover Timing of Cluster Application

Point

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

Information

For the flow to build the PRIMECLUSTER system, refer to "Chapter 1 Build Flow" in "PRIMECLUSTER Installation and Administration Guide."

2.1 Selecting the PRIMECLUSTER Product

Select a PRIMECLUSTER product.

In an FJcloud-O environment, you can select the following products.

For details on the PRIMECLUSTER products, refer to "2.1 PRIMECLUSTER Product Selection" in "PRIMECLUSTER Installation and Administration Guide."

- PRIMECLUSTER Enterprise Edition (EE)
- PRIMECLUSTER HA Server (HA)
- PRIMECLUSTER Clustering Base (CB)

2.2 System Design

Use PRIMECLUSTER Designsheets to design the system.

The installation operation of the PRIMECLUSTER system is performed based on the created PRIMECLUSTER Designsheets. Make sure to create the designsheets and confirm that all required items are described.

2.3 Determining the Cluster System Operation Mode

In the cluster system in an FJcloud-O environment, operation modes for 1:1 standby operation, mutual standby, and single-node cluster operation can be built.

For details on the operation mode of each cluster system, refer to "2.3 Determining the Cluster System Operation Mode" in "PRIMECLUSTER Installation and Administration Guide."

2.4 Determining the Web-Based Admin View Operation Mode

For details on the Web-Based Admin View operation mode, refer to "2.4 Determining the Web-Based Admin View Operation Mode" in "PRIMECLUSTER Installation and Administration Guide."

2.5 Determining the Failover Timing of Cluster Application

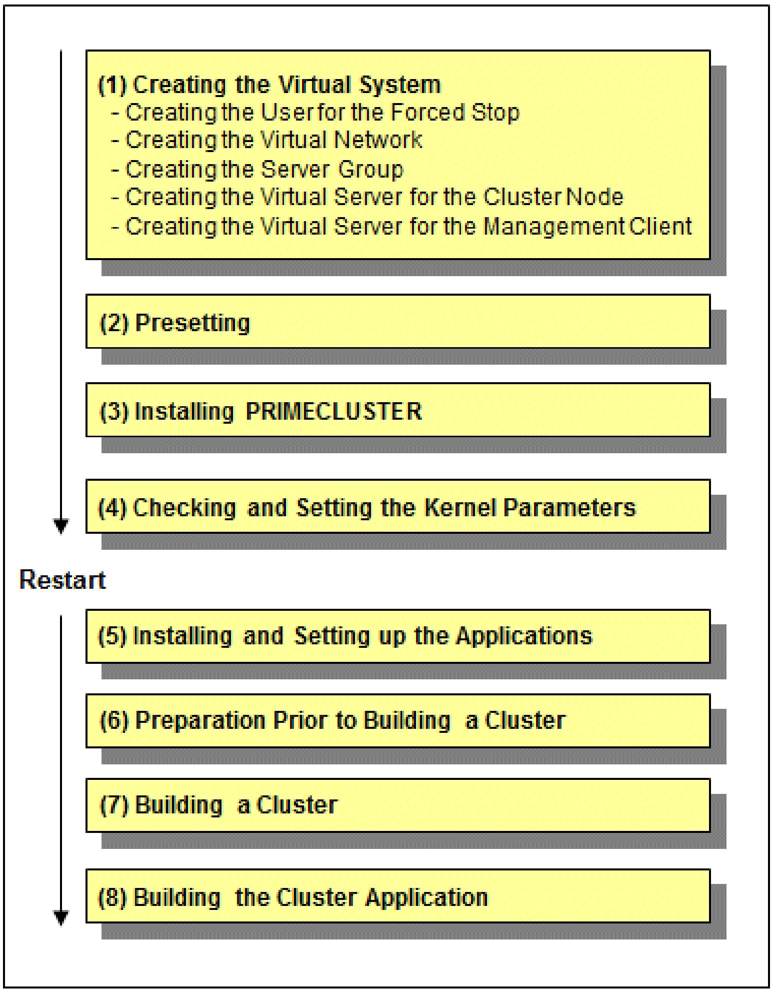
Determine the failover timing of cluster application.

For details, refer to "2.5 Determining the Failover Timing of Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

Chapter 3 Installation

This chapter explains the procedure to install PRIMECLUSTER in an FJcloud-O environment.

Perform the steps shown in the figure below.


(1) Creating the Virtual System
- Creating the User for the Forced Stop
- Creating the Virtual Network
- Creating the Server Group
- Creating the Virtual Server for the Cluster Node
- Creating the Virtual Server for the Management Client
(2) Presetting
(3) Installing PRIMECLUSTER
(4) Checking and Setting the Kernel Parameters
Restart
(5) Installing and Setting up the Applications
(6) Preparation Prior to Building a Cluster
(7) Building a Cluster
(8) Building the Cluster Application

Note

For how to check the operation environment, refer to "Chapter 2 Operation Environment" in the Installation Guide for PRIMECLUSTER.

3.1 Creating the Virtual System

This section explains how to create the virtual system for a cluster system in an FJcloud-O environment.

See

For details on FJcloud-O, refer to the documents provided by FJcloud-O.

Note

- To use the service provided by FJcloud-O IaaS with API, it is necessary to build an environment for using API. For details, refer to the documents provided by FJcloud-O.

- If the execution examples in this chapter are executed as they are described, an execution error may occur due to incompatibilities of API depending on a region of FJcloud-O. For details, refer to the documents provided by FJcloud-O.

3.1.1 Creating the User for the Forced Stop

On the FJcloud portal, create the user to forcibly stop the virtual server in the cluster system with the following values.

These settings are not necessary in a single-node cluster.

If the user is already created in an FJcloud-O environment, you can use that user without creating a new user.

Item name	Value
User name	Arbitrary user name
Role	Developer

Then, in the IaaS management of the FJcloud portal, add the created user to the project where the virtual server is to be created, and grant the following role.

Item name	Value
Role	Operator role

The user authentication method can be "Password Authentication" or "Certificate + Password Authentication".

On the FJcloud portal, set the authentication method to be used.

To use "Certificate + Password Authentication", issue and download the client certificate of the user from the FJcloud portal.

If you have already downloaded the client certificate, you do not need to do so.

Note

- The password for the user expires in 90 days and must be changed periodically. For the procedure on changing a password, refer to "6.1 Changing a Password Periodically."
- The client certificate for the user expires in 3 years and must be changed periodically. For the procedure on updating the client certificate, see "6.2 Updating the Certificate File".
- If FJcloud-O fails to authenticate a user consecutively, the user is locked out. While the user is locked out, the shutdown facility cannot operate normally and the cluster application will be suspended. We recommend that you create a dedicated user for the user for the forced stop. If the user is locked out, enter the correct password 30 minutes after locking or reissue the password to unlock it.

See

For details on the FJcloud portal, refer to the official FUJITSU Hybrid IT Service FJcloud documentation.

3.1.2 Creating the Virtual Network

Create subnets and security groups for the public LAN (used also for the administrative LAN) or the cluster interconnect.

Note

When creating multiple cluster systems, create the following security groups with every cluster system.

- For the public LAN (used also for the administrative LAN)
- For the cluster interconnect
- For Web-Based Admin View (on the cluster node side)

- For Web-Based Admin View (on the management client side)

3.1.2.1 Creating Subnets

Use the following values to create the subnets used for the public LAN (used also for the administrative LAN) and the cluster interconnect in the cluster system.

Item name	Value
Enable/Disable DHCP Auto Allocation	true (Default)
Pool for assigning IP address	A range of IP address assigned to each node (the takeover IP address is excluded from the range)

Create the virtual router so as to communicate to each endpoint from the virtual server in FJcloud-O and then connect to the subnet for the public LAN (used also for the administrative LAN).

3.1.2.2 Creating the Common Security Group

Create the security group used in common with the following values.

Communication direction	Communication destination	Protocol information	Starting port number	Ending port number
egress	169.254.169.254/32 (*1)	tcp	80	80
egress	IP address of DNS server	udp	53	53
egress	IP address of DNS server	tcp	53	53
egress	IP address of NTP server	udp	123	123

(*1) The IP address used by the virtual server to obtain information on the cloud side. (It has nothing to do with the cluster operation.)

3.1.2.3 Creating the Security Group for the Public LAN (Used also for the Administrative LAN)

Create the security group for the public LAN (used also for the administrative LAN) with the following values

Communication direction	Communication destination	Protocol information	Starting port number	Ending port number
egress (*1)	Not specified	tcp	443	443
ingress	Own security group	udp	9382	9382
egress	Own security group	udp	9382	9382
ingress	Own security group	udp	9796	9796
egress	Own security group	udp	9796	9796
ingress	Own security group	tcp	9797	9797
egress	Own security group	tcp	9797	9797
egress	IP address of the virtual gateway	icmp	Not specified	Not specified
ingress (*2)	Own security group	tcp	3260	3260
egress (*2)	Own security group	tcp	3260	3260
ingress	Own security group	icmp	Not specified	Not specified
egress	Own security group	icmp	Not specified	Not specified

(*1) This setting is not necessary in a single-node cluster.

(*2) This setting is not necessary when not using the mirroring among the servers of GDS.

3.1.2.4 Creating the Security Group for the Cluster Interconnect

Create the security group for the cluster interconnect with the following values.

These settings are not necessary in a single-node cluster.

Communication direction	Communication destination	Protocol information	Starting port number	Ending port number
egress	Own security group	123 (*1)	Not specified	Not specified
ingress	Own security group	123 (*1)	Not specified	Not specified

(*1) Use a protocol other than TCP/UDP/ICMP. Enter the above value for other protocols.

3.1.2.5 Creating the Security Groups for Web-Based Admin View

Create the security group for Web-Based Admin View (on the cluster node side) with the following values.

Communication direction	Communication destination	Protocol information	Starting port number	Ending port number
ingress	Own security group	tcp	8081	8081
ingress	Own security group	tcp	9798	9798
ingress	Own security group	tcp	9799	9799

Create the security group for Web-Based Admin View (on the management client side) with the following values.

Communication direction	Communication destination	Protocol information	Starting port number	Ending port number
egress	Own security group	tcp	8081	8081
egress	Own security group	tcp	9798	9798
egress	Own security group	tcp	9799	9799

3.1.2.6 Creating the Security Group for the Virtual Server Access

Create the security group for installing and maintaining the cluster node.

Create the security group for ssh connection to the cluster node with the following values.

Communication direction	Communication destination	Protocol information	Starting port number	Ending port number
ingress	Specified timely	tcp	22	22

Note

When using the yum command, create the security group with the following values.

Communication direction	Communication destination	Protocol information	Starting port number	Ending port number
egress	IP address of the repository server	tcp	80	80
egress	IP address of the repository server	tcp	443	443

Create the security group for installing and maintaining the management client.

Create the security group for the remote desktop connection to the management client with the following values.

Communication direction	Communication destination	Protocol information	Starting port number	Ending port number
ingress	Specified timely	tcp	3389	3389

3.1.2.7 Creating the Firewall Rule

When using the firewall service, add the following to the firewall rule.

Protocol	Source IP address	Destination IP address	Destination port number	Action
tcp (*1)	Subnet for the public LAN (used also for the administrative LAN)	Not specified	443	Allow
udp	Subnet for the public LAN (used also for the administrative LAN)	IP address of DNS server	53	Allow
tcp	Subnet for the public LAN (used also for the administrative LAN)	IP address of DNS server	53	Allow
udp	Subnet for the public LAN (used also for the administrative LAN)	IP address of NTP server	123	Allow

(*1) This setting is not necessary in a single-node cluster.

Note

- Add the settings to allow the connection via ssh or the remote desktop connection from the external network as necessary.
- When using the yum command, add the following settings. Add or delete these settings as necessary to enhance the security.

Communication direction	Communication destination	Protocol information	Starting port number	Action
egress	Subnet for the public LAN (used also for the administrative LAN)	IP address to the repository server	80	Allow
egress	Subnet for the public LAN (used also for the administrative LAN)	IP address to the repository server	443	Allow

3.1.3 Creating the Server Group

Create the server group so that the virtual servers in the cluster start on the different physical hosts. From the IaaS Service Portal of FJcloud-O, set the server group as the following on the availability zone where the virtual servers are assigned.

Table 3.1 Setting the server group

Item	Value
Server group name	Arbitrary server group name
Policy	anti-affinity

Note

When creating multiple cluster systems, create the server group with every cluster system.

3.1.4 Creating the Virtual Server for the Cluster Node

Create the virtual server for the cluster node.

Perform the following operations by the number of the nodes in the cluster system and create the virtual server for the cluster node.

- Creating the virtual server
- Creating the port for the public LAN (used also for the administrative LAN)
- Creating the port for the cluster interconnect
- Creating/Attaching the expanded storage
- Setting up the DNS client
- Application of the necessary OS patch
- Creating .curlrc

3.1.4.1 Creating the Virtual Server

From the IaaS Service Portal of FJcloud-O, for the server group created in "3.1.3 Creating the Server Group", set the virtual servers in the cluster system as "Table 3.2 Values of the virtual servers for the cluster node."

For security reasons, leave the virtual servers in the "SHUTOFF" state until "3.1.4.2 Creating the Port for the Public LAN (Used also for the Administrative LAN)" and "3.1.4.3 Creating the Port for the Cluster Interconnect" are completed. Also, delete the ports that are automatically added when the virtual servers are created. You can check the ports on the details screen of the virtual server.

Table 3.2 Values of the virtual servers for the cluster node

Item	Value
AZ (*1)	Availability zone to assign the virtual server
Server group (*2)	Server group created in "3.1.3 Creating the Server Group"
Virtual server name	Arbitrary virtual server name *Specify a virtual server name taking care that there are no virtual names in duplicate within the project.
Virtual server type	Arbitrary virtual server type according to the performance requirement (flavor)
Boot source of the virtual server	OS image used in the cluster node
Virtual network	Network created in "3.1.2.1 Creating Subnets"
Security group	Not specified
Keypair	Arbitrary keypair

(*1) Not displayed in East Japan third, West Japan third regions.

(*2) Not displayed in East Japan first/second, West Japan first/second regions.
You need to operate from the action of the server group created in "3.1.3 Creating the Server Group."

3.1.4.2 Creating the Port for the Public LAN (Used also for the Administrative LAN)

Set the port for the public LAN (used also for the administrative LAN) in the virtual server in the cluster system as follows.

1. Click the <+> button in the port list on [Details screen of the virtual server] of the virtual server created in "3.1.4.1 Creating the Virtual Server" to transit to the port creation screen.
2. After entering the following setup items on the port creation screen, click the <Create> button to create the port.

Table 3.3 Port to be created in the subnet of the public LAN and the administrative LAN

Item	Value
Port name	Enter an arbitrary port name.

Item	Value
Management state	Select "Up".
Network name	Select the network to which the port is connected.
Subnet name	Select the subnet for the public LAN (used also for the administrative LAN) created in "3.1.2.1 Creating Subnets."
Private IP address	Enter the IP address of the public LAN (used also for the administrative LAN).

3. Click the port name in the port list on the details screen of the virtual server to display the details screen of the port, and make sure that all settings are correct.

4. Select the action of the port on the details screen of the virtual server to set the following security groups.

   - ID of the security group created in "3.1.2.2 Creating the Common Security Group"

   - ID of the security group created in "3.1.2.3 Creating the Security Group for the Public LAN (Used also for the Administrative LAN)"

   - ID of the security group on the cluster node side created in "3.1.2.5 Creating the Security Groups for Web-Based Admin View"

   - ID of the security group for the installation and maintenance of the cluster node created in "3.1.2.6 Creating the Security Group for the Virtual Server Access"

   - If there are any necessary security groups for operations other than those above, add them.

5. When taking over the IP address between the virtual servers, execute the following API, and set the takeover IP address as "allowed_address_pairs".

```
# curl -X PUT https://networking.jp-east-3.cloud.global.fujitsu.com/v2.0/ports/
{id_of_created_port} -H 'X-Auth-Token:XXX' -H 'Content-Type:application/json' -H
'Accept:application/json' -d '{"port":{ "allowed_address_pairs" : [{"ip_address":"takeover IP
address"}]}}'
```

3.1.4.3 Creating the Port for the Cluster Interconnect

Set the port for the cluster interconnect as follows.

These settings are not necessary in a single-node cluster.

1. Click the <+> button in the port list on [Details screen of the virtual server] of the virtual server created in "3.1.4.1 Creating the Virtual Server" to transit to the port creation screen.

2. After entering the following setup items on the port creation screen, click the <Create> button to create the port.

Table 3.4 Port to be created in the subnet of the cluster interconnect

Item	Value
Port name	Enter an arbitrary port name.
Management state	Select "Up".
Network name	Select the network name of the interconnect created in "3.1.2.1 Creating Subnets."
Subnet name	Select the subnet name of the interconnect created in "3.1.2.1 Creating Subnets."
Private IP address	Enter the IP address of the cluster interconnect (*1).

(*1) This is the IP address used with CF over IP. For details on CF over IP, refer to "Chapter 9 CF over IP" in "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

3. Click the port name in the port list on the details screen of the virtual server to display the details screen of the port, and make sure that all settings are correct.

4. Select the action of the port on the details screen of the virtual server to set the following security group.

   - ID of the security group for the cluster interconnect created in "3.1.2.4 Creating the Security Group for the Cluster Interconnect"

3.1.4.4 Creating/Attaching the Expanded Storage

When using the mirroring among the servers of GDS, create the block storages used in the mirroring among the servers and attach to the virtual servers created in "3.1.4.1 Creating the Virtual Server" as the expanded storage.

Attach the same size of the block storage to the virtual server for each cluster node.

Note

Make sure to restart the virtual server after attaching the expanded storage.

3.1.4.5 Setting up the DNS Client

Note

- If this setting is done incorrectly by mistake, the system may not be accessible. Before setting the DNS client, acquire the snapshot to the system disk.
- This setting is not necessary in a single-node cluster.

Set the following in each node building the cluster.

The connection name of eth0 is "Wiredeth0" and the connection name of eth1 is "Wiredeth1". Read these names according to your environment.

1. Execute the following command, and make sure that ipv4.never-default and ipv4.ignore-auto-dns of eth0 are "no".

   ```
   # nmcli connection show Wiredeth0
   ```

   If they are not "no", execute the following commands and set them to "no".

   ```
   # nmcli connection modify Wiredeth0 ipv4.never-default no
   # nmcli connection modify Wiredeth0 ipv4.ignore-auto-dns no
   ```

2. Set the DNS server as follows.

   ```
   # nmcli connection modify Wiredeth0 ipv4.dns <IP address of the main DNS server>
   # nmcli connection modify Wiredeth0 +ipv4.dns <IP address of the sub DNS server>
   ```

3. For eth1, set ipv4.never-default and ipv4.ignore-auto-dns to yes.

   ```
   # nmcli connection modify Wiredeth1 ipv4.never-default yes
   # nmcli connection modify Wiredeth1 ipv4.ignore-auto-dns yes
   ```

3.1.4.6 Application of the Necessary OS Patch

Refer to the documents provided by FJcloud-O to set up Red Hat Update Infrastructure to the virtual server.

After setting it, execute the following command to apply the necessary OS patch.

```
# yum update curl
```

3.1.4.7 Creating .curlrc

Add the following line into the file, /root/.curlrc. If there is no file, create it and describe the following.

```
tlsv1.2
```

If you created the file, execute the following.

```
# chown root:root /root/.curlrc
# chmod 600 /root/.curlrc
```

3.1.5 Creating the Virtual Server for the Management Client

Create the virtual server for the management client.

Create the port for the public LAN (used also for the administrative LAN) for the management client to create the virtual server.

3.1.5.1 Creating the Virtual Server

From the IaaS Service Portal of FJcloud-O, create the virtual server of the management client.

Set the virtual server of the management client as "Table 3.5 Values of the virtual server for the management client."

For security reasons, leave the virtual server in the "SHUTOFF" state until "3.1.5.2 Creating the Port for the Public LAN (Used also for the Administrative LAN)" is completed. Also, delete the ports that are automatically added when the virtual servers are created. You can check the ports on the details screen of the virtual server.

Table 3.5 Values of the virtual server for the management client

Item	Value
AZ (*1)	Availability zone to assign the virtual server
Server group (*2)	Arbitrary server group
Virtual server name	Arbitrary virtual server name *Specify a virtual server name taking care that there are no virtual names in duplicate within the project.
Virtual server type	Arbitrary virtual server type according to the performance requirement (flavor)
Boot source of the virtual server	Image of Windows Server 2012 R2, image of Windows Server 2016, or image of Windows Server 2019
Virtual network	Network created in "3.1.2.1 Creating Subnets"
Security group	Not specified
Password creation	Manual
Password	Arbitrary password

(*1) Not displayed in East Japan third, West Japan third regions.

(*2) Only displayed in East Japan third, West Japan third regions.

3.1.5.2 Creating the Port for the Public LAN (Used also for the Administrative LAN)

Set the port for the public LAN (used also for the administrative LAN) in the virtual server for the management client as follows.

1. Click the <+> button in the port list on [Details screen of the virtual server] of the virtual server created in "3.1.5.1 Creating the Virtual Server" to transit to the port creation screen.
2. After entering the following setup items on the port creation screen, click the <Create> button to create the port.

Table 3.6 Port to be created in the subnet of the public LAN and the administrative LAN

Item	Value
Port name	Enter an arbitrary port name.
Management state	Select "Up".

Item	Value
Network name	Select the network to which the port is connected.
Subnet name	Select the subnet for the public LAN (used also for the administrative LAN) created in "3.1.2.1 Creating Subnets."
Private IP address	Enter the IP address of the public LAN (used also for the administrative LAN).

3. Click the port name in the port list on the details screen of the virtual server to display the details screen of the port, and make sure that all settings are correct.

4. Select the action of the port on the details screen of the virtual server to set the following security groups.

   - ID of the security group created in "3.1.2.2 Creating the Common Security Group"

   - ID of the security group on the management client side created in "3.1.2.5 Creating the Security Groups for Web-Based Admin View"

   - ID of the security group for the installation and maintenance of the management client created in "3.1.2.6 Creating the Security Group for the Virtual Server Access"

   - If there are any necessary security groups for operations other than those above, add them.

3.2 Presetting

1. Disable the firewall.

   Make sure that "firewalld" is disabled.

```
# systemctl is-enabled firewalld
```

   If it is enabled, disable it.

```
# systemctl stop firewalld
# systemctl disable firewalld
```

2. Set NTP.

   Make sure to set NTP when building the cluster to synchronize the time of each node in the cluster system.

   Set NTP before installing PRIMECLUSTER.

3.3 Installing PRIMECLUSTER

Use the installation script (CLI Installer) to install PRIMECLUSTER.

Install PRIMECLUSTER on each node in the system where Linux(R) software and Linux(R) related software are already installed. Use the same installation script when installing PRIMECLUSTER in the cluster management server.

Note

If OS has never been restarted since the virtual server was created, restart it and then install PRIMECLUSTER.

See

For details on the installation/uninstallation procedure, refer to the sections of the cloud environments described in the Installation Guide for PRIMECLUSTER.

3.4 Checking and Setting the Kernel Parameters

Change the kernel parameters depending on the environment.

Applicable nodes:

All nodes on which PRIMECLUSTER is to be installed

Depending on the products and components utilized, different kernel parameters are required.

Check PRIMECLUSTER Designsheets and if you need to modify the kernel parameters, set them again.

See

For details on kernel parameters, refer to "3.1.7 Checking and Setting the Kernel Parameters" in "PRIMECLUSTER Installation and Administration Guide."

Note

To activate the modified kernel parameters, restart OS.

3.5 Installing and Setting up the Applications

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

See

- For details on environment setup, refer to manuals for each application.
- For information on PRIMECLUSTER-related products that support FJcloud-O, refer to the documentation for each product.

3.6 Preparation Prior to Building a Cluster

Prior to building a cluster, perform presettings such as the initial GLS setup, setting the DNS client, creating the FJcloud-O environment information file, and staring the Web-Based Admin View screen.

3.6.1 Initial GLS Setup

When using GLS, execute the initial GLS setup to the network used for the public LAN (used also for the administrative LAN), according to the procedure below. For details on each setting, refer to "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function."

Note

If this setting is done incorrectly by mistake, the system may not be accessible. Before the initial GLS setup, acquire the snapshot to the system disk.

Set the following in each node building the cluster.

1. Set the system.

    1. In the /etc/hosts file, define the IP address and the host name to be used.

Example

```
172.16.0.10    node1    IP address of # node1
172.16.0.11    node2    IP address of # node2
172.16.0.100    takeover    # Takeover IP address
172.16.0.1    gw    # Gateway IP address
```

2. Modify the network settings.

   [RHEL8]

   Set BOOTPROTO=static, PEERDNS=no, and DEFROUTE=no in the/etc/sysconfig/network-scripts/ifcfg-eth0 file.

   - Contents of /etc/sysconfig/network-scripts/ifcfg-eth0

   ```
   DEVICE=eth0
   TYPE=Ethernet
   BOOTPROTO=static
   DEFROUTE=no
   UUID=<fixed value depending on an environment (change not required)>
   HOTPLUG=no
   ONBOOT=yes
   PEERDNS=no
   ```

   Note

   Describe all the parameters described in /etc/sysconfig/network-scripts/ifcfg-eth0.

   [RHEL9]

   The nmcli connection modify command sets the following parameters.

   ```
   ipv4.method "disabled"
   ipv4.ignore-auto-dns "yes"
   ipv4.never-default "yes"
   ```

   Example) To set ipv4.ignore-auto-dns to "yes"

   ```
   # nmcli connection modify eth0 ipv4.ignore-auto-dns yes
   ```

2. Create the virtual interface.

   Execute the following command, and create the virtual interface.

```
# /opt/FJSVhanet/usr/sbin/hanetconfig create -n sha0 -m v -t eth0
```

3. Set the virtual interface.

   [RHEL8]

   In the /etc/sysconfig/network-scripts/ifcfg-sha0 file, comment out IPADDR and PREFIX. Set BOOTPROTO=dhcp. Also add DEFROUTE=yes, PEERDNS=yes, and the settings of DNS1 and DNS2.

   - Contents of /etc/sysconfig/network-scripts/ifcfg-sha0

   ```
   DEVICE=sha0
   #IPADDR=
   #PREFIX=
   BOOTPROTO=dhcp
   DEFROUTE=yes
   ONBOOT=yes
   TYPE=Ethernet
   PEERDNS=yes
   ```

```
DNS1=<IP address of the main DNS server>
DNS2=<IP address of the sub DNS server>
```

Note

- Do not set ifcfg-sha0 to SHAMACADDR.
- Describe all the parameters described in the /etc/sysconfig/network-scripts/ifcfg-sha0 file.

[RHEL9]

In the nmcli connection modify command, set the following parameters for sha0.

```
ipv4.addresses ""
ipv4.method "auto"
ipv4.ignore-auto-dns "no"
ipv4.never-default "no"
```

Example) To set ipv4.ignore-auto-dns to "no".

```
# nmcli connection modify sha0 ipv4.ignore-auto-dns no
```

In the nmcli connection modify command, add the following parameter to sha0.

```
ipv4.dns <IP address of the primary DNS server>,<IP address of the secondary DNS server>
```

```
# nmcli connection modify sha0 +ipv4.dns <IP address of the primary DNS server> +ipv4.dns <IP
address of the secondary DNS server>
```

4. Set the network monitoring function.

   Set the virtual router to the monitoring destination. In consideration of a prolonged time stop in the virtual router, configure the settings to avoid the switchover of cluster when a failure of network route occurs.

Example

```
# /opt/FJSVhanet/usr/sbin/hanetpathmon target -n sha0 -p 172.16.0.1
# /opt/FJSVhanet/usr/sbin/hanetpathmon param -n sha0 -f no
```

5. Set the subnet mask of the takeover virtual interface.

Example

```
# /opt/FJSVhanet/usr/sbin/hanetmask create -i 172.16.0.0 -m 255.255.255.0
```

6. Create the takeover virtual interface.

Example

```
# /opt/FJSVhanet/usr/sbin/hanethvrsc create -n sha0 -i 172.16.0.100
```

7. Check the configuration.

   Make sure that the settings done from step 3 to step 6 are reflected.

Example

```
# /opt/FJSVhanet/usr/sbin/hanetconfig print
 [IPv4,Patrol / Virtual NIC]

 Name       Hostname      Mode Physical ipaddr    Interface List
+-----------+---------------+----+-----------------+---------------------------+
 sha0                          v                    eth0

[IPv6]

 Name        Hostname/prefix                  Mode Interface List
+-----------+---------------------------------+-----+----------------------------+
```

```
# /opt/FJSVhanet/usr/sbin/hanetpathmon target
[Target List]
Name   VID  Target
+-------+----+-----------------------------------------------------------+
sha0    -    172.16.0.1
```

```
# /opt/FJSVhanet/usr/sbin/hanetpathmon param
[Parameter List]
Name   Monitoring Parameter
+-------+----------------------------------------------------------+
sha0    auto_startup       =    yes
         interval           =      3 sec
         times              =      5 times
         repair_times       =      2 times
         idle               =     45 sec
         Auto fail-back     =     no
         FAILOVER Status    =     no
```

```
# /opt/FJSVhanet/usr/sbin/hanetmask print
network-address  netmask
+---------------+---------------+
172.16.0.0       255.255.255.0
```

```
# /opt/FJSVhanet/usr/sbin/hanethvrsc print
 ifname     takeover-ipv4     takeover-ipv6     vlan-id/logical ip address list
+----------+----------------+----------------+---------------------------------+
 sha0:65    172.16.0.100     -                -
```

8. Restart the system.

   Run the following command and restart the system.

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

3.6.2 Creating the FJcloud-O Environment Information File

To activate a cluster system in an FJcloud-O environment, create the FJcloud-O environment information file with the following procedure.

This setting is not necessary in a single-node cluster.

1. Create the /opt/SMAW/SMAWRrms/etc/k5_endpoint.cfg file on all nodes as shown below:

   ```
   DOMAIN_NAME=domainname
   PROJECT_NAME=projectname
   IDENTITY=identityurl
   COMPUTE=computeurl
   ```

```
domainname : Domain name of FJcloud-O (contractor number)
projectname: Project name building a cluster in FJcloud-O
identityurl: URL of the endpoint for the regional user management of the region used in
             FJcloud-O (*)
computeurl : URL of the endpoint for the compute (standard service) of the region used in
             FJcloud-O (*)

* For details on URL of the endpoint for the regional user management and the compute (standard
service), refer to the documents provided by FJcloud-O.
```

Example

```
DOMAIN_NAME=primecluster_domain
PROJECT_NAME=primecluster_project
IDENTITY=https://identity.jp-east-3.cloud.global.fujitsu.com
COMPUTE=https://compute.jp-east-3.cloud.global.fujitsu.com
```

2. Set the owner, group, and access rights as follows.

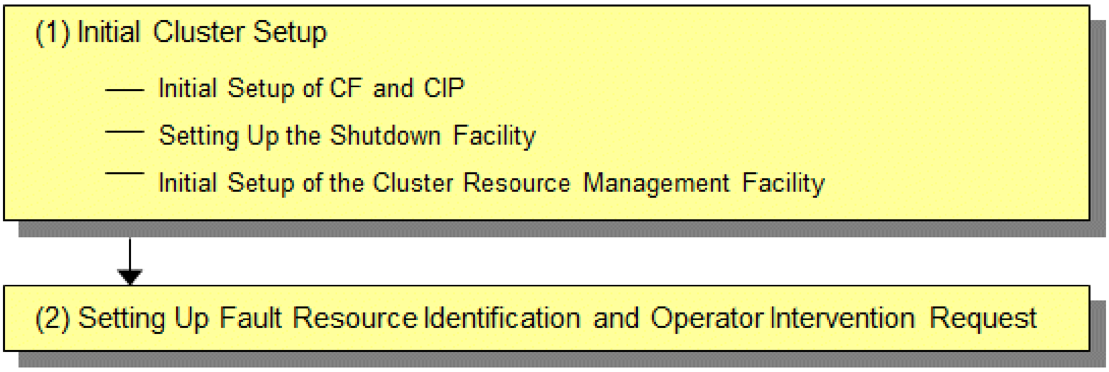
```
# chown root:root /opt/SMAW/SMAWRrms/etc/k5_endpoint.cfg
# chmod 600 /opt/SMAW/SMAWRrms/etc/k5_endpoint.cfg
```

3.6.3 Presettings for Building a Cluster

Refer to "Chapter 4 Preparation Prior to Building a Cluster" in "PRIMECLUSTER Installation and Administration Guide", and execute the initial setup for a cluster in the virtual server.

3.7 Building a Cluster

The procedure for building a PRIMECLUSTER cluster is shown below:


(1) Initial Cluster Setup
Initial Setup of 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

3.7.1 Initial Cluster Setup

This section explains the initial cluster setup for PRIMECLUSTER.

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

	Details	Manual reference location*
1	3.7.1.1 Initial Setup of CF and CIP (setting up cluster configuration information and IP addresses)	CF "1.1 CF, CIP, and CIM configuration"
2	3.7.1.2 Setting up the Shutdown Facility	CF "7 Shutdown Facility (SF)"

	Details	Manual reference location*
3	3.7.1.3 Initial Setup of the Cluster Resource Management Facility	CF "3.1 Resource Database configuration"

*The PRIMECLUSTER manual name is abbreviated as follows:

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

3.7.1.1 Initial Setup of CF and CIP

Refer to "5.1.1 Setting Up CF and CIP" in "PRIMECLUSTER Installation and Administration Guide" to set up CF and CIP.

Note

You need to use CF over IP in a cloud environment. In addition, you need to configure CIP because such as RMS uses it for inter-node communication in the cluster. CF over IP (IP interconnect) and CIP are different features. For more information, see "1.1.1 Differences between CIP and CF over IP" in the Cluster Foundation Configuration and Administration Guide.

3.7.1.2 Setting up the Shutdown Facility

In an FJcloud-O environment, only the SA_vmk5r shutdown agent is available for setup.

This section explains the method for setting up the SA_vmk5r shutdown agent as the shutdown facility.

For details on the survival priority, refer to "5.1.2.1 Survival Priority." in "PRIMECLUSTER Installation and Administration Guide."

Note

- This setting is not necessary in a single-node cluster.
- After setting up the shutdown agent, conduct a test for the forced stop of cluster nodes to make sure that the correct nodes can be forcibly stopped. For details of the test for the forced stop of cluster nodes, refer to "1.4 Test" in "PRIMECLUSTER Installation and Administration Guide."
- The contents of the SA_vmk5r.cfg file and the rcsd.cfg file of all nodes should be identical. If not, a malfunction will occur.

3.7.1.2.1 Distributing Client Certificates for the User for the Forced Stop

If the authentication method of the user for the forced stop of FJcloud-O created in "3.1.1 Creating the User for the Forced Stop" is Certificate + Password Authentication, perform the following steps. The following procedures are not required, if "Password Authentication".

Place the downloaded client certificate of the user for the forced stop on all nodes and convert it to a format that can be used by PRIMECLUSTER. Perform the following procedure on all nodes.

1. Distributing the client certificate for the user for the forced stop

   Place the client certificate for the user for the forced stop downloaded in

   "3.1.1 Creating the User for the Forced Stop" in any directory on all cluster nodes.

2. Creating a client certificate file placement directory for PRIMECLUSTER

   Create an any directory in which to place client certificate file for PRIMECLUSTER.

```
# mkdir -p <client certificate file placement directory>
# chown root:root <client certificate file placement directory>
# chmod 700 <client certificate file placement directory>
```

3. Creating the client certificate file for PRIMECLUSTER

   Create the client certificate file for PRIMECLUSTER by converting the client certificate for the user

   for the forced stop you placed in step 1. The file name of the client certificate for PRIMECLUSTER is optional.

```
# cd <client certificate file placement directory created in step 2>
# openssl pkcs12 -in <client certificate placed in step 1> -clcerts -nokeys -out <client
certificate file name for PRIMECLUSTER>
  Enter Import Password: <passphrase for the client certificate placed in step 1>
# chown root:root <client certificate file for PRIMECLUSTER>
# chmod 400 <client certificate file for PRIMECLUSTER>
```

4. Creating a directory for private key file

   Create an any directory in which to place the private key file.

```
# mkdir -p <private key file placement directory>
# chown root:root <private key file placement directory>
# chmod 700 <private key file placement directory>
```

Point

It is recommended that this directory be separate from the "client certificate file placement directory" created in step 2.

5. Creating the private key file

   Create the private key file from the client certificate for the user for the forced stop you placed in step 1. The private key file name is optional.

```
# cd <private key file placement directory created in step 4>
# openssl pkcs12 -in <client certificate placed in step 1> -nocerts -nodes -out <private key file
name>
  Enter Import Password: <passphrase for the client certificate placed in step 1>
# chown root:root <private key file>
# chmod 400 <private key file>
```

6. Deleting the client certificate for the user for the forced stop

   Delete the client certificate for the user for the forced stop you placed in step 1.

Information

- You can check the expiration date of the client certificate file for PRIMECLUSTER with the following command.

```
# openssl x509 -in <client certificate file for PRIMECLUSTER>
```

- You can verify that the client certificate file for PRIMECLUSTER and private key file combination

is correct by matching the modulus value in the output of the following command.

```
# openssl x509 -in <client certificate file for PRIMECLUSTER> -modulus -noout
# openssl rsa -in <private key file> -modulus -noout
```

3.7.1.2.2 Setting up the Shutdown Daemon

Create /etc/opt/SMAW/SMAWsf/rcsd.cfg with the following contents on all nodes in the cluster system.

```
CFNameX,weight=weight,admIP=myadmIP:agent=SA_vmk5r,timeout=125
CFNameX,weight=weight,admIP=myadmIP:agent=SA_vmk5r,timeout=125
```

```
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 used in the shutdown facility
                 of the cluster host.
                 Available IP addresses are IPv4.
                 When specifying a host name, make sure it is described in /etc/hosts.
```

Example) The following is a setup example.

```
# cat /etc/opt/SMAW/SMAWsf/rcsd.cfg
node1,weight=1,admIP=192.168.1.1:agent=SA_vmk5r,timeout=125
node2,weight=1,admIP=192.168.1.2:agent=SA_vmk5r,timeout=125
```

Create /etc/opt/SMAW/SMAWsf/rcsd.cfg and then set the owner, group, and access rights as follows.

```
# chown root:root /etc/opt/SMAW/SMAWsf/rcsd.cfg
# chmod 600 /etc/opt/SMAW/SMAWsf/rcsd.cfg
```

Information

When creating the /etc/opt/SMAW/SMAWsf/rcsd.cfg file, the /etc/opt/SMAW/SMAWsf/rcsd.cfg.template file can be used as a template.

3.7.1.2.3 Setting up SA_vmk5r Shutdown Agent

1. Encrypt the password.

   Execute the sfcipher command to encrypt a password of a user for forcibly stopping the virtual server of FJcloud-O. For details on how to use the sfcipher command, refer to the manual page of "sfcipher."

   ```
   # sfcipher -c
   ```

   Example) The following is a setup example.

   If a password is "k5admin$":

   ```
   # sfcipher -c
   Enter Password:    <- Enter k5admin$
   Re-Enter Password: <- Enter k5admin$
   O/gm+AYuWwE7ow3dgVG/Nw==
   ```

2. Set the shutdown agent.

   Create /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg with the following contents on all nodes in the cluster system.

   Delimit each item with a single space.

   ```
   CFNameX ServerName user passwd {cycle | leave-off} [ClientCertFile PrivateKeyFile]
   CFNameX ServerName user passwd {cycle | leave-off} [ClientCertFile PrivateKeyFile]
   ```

   ```
   CFNameX          : Specify the CF node name of the cluster host.
   ServerName       : Specify the virtual server name in FJcloud-O on which the cluster host
                      is running.
   user             : Specify a user name for forcibly stopping the virtual server in FJcloud-O.
   passwd           : Specify a password encrypted in step 2.
   cycle            : Restart the node after forcibly stopping the node.
   leave-off        : Power-off the node after forcibly stopping the node.
   ClientCertFile   : Specify the path to the client certificate file for PRIMECLUSTER that you
                      created in step 3 of
                     "3.7.1.2.1 Distributing Client Certificates for the User for the Forced Stop".
                      This item is required when the authentication method of the user for the
                      forced stop is "Certificate + Password Authentication".
                      This setting is not required for "Password Authentication"
   PrivateKeyFile   : Specify the path to the private key file that you created in step 5 of
                     "3.7.1.2.1 Distributing Client Certificates for the User for the Forced Stop".
                      This item is required when the authentication method of the user for the
                      forced stop is "Certificate + Password Authentication".
                      This setting is not required for "Password Authentication"
   ```

   Example) The following is a setup example.

   This example shows the following settings:

- The authentication method of the user for the forced stop is "Password Authentication"

- The CF node names of the cluster host are node1 and node2.

- The virtual server names are vm1 and vm2.

- The user name to forcibly stop the virtual servers is pcl.

- The node will be restarted when it is forcibly stopped.

```
# cat /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg
node1 vm1 pcl O/gm+AYuWwE7ow3dgVG/Nw== cycle
node2 vm2 pcl O/gm+AYuWwE7ow3dgVG/Nw== cycle
```

This example shows the following settings:

- The authentication method of the user for the forced stop is "Certificate + Password Authentication"

- The CF node names of the cluster host are node1 and node2.

- The virtual server names are vm1 and vm2.

- The user name to forcibly stop the virtual servers is pcl.

- The node will be restarted when it is forcibly stopped.

```
# cat /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg
node1 vm1 pcl O/gm+AYuWwE7ow3dgVG/Nw== cycle /root/cert/pcl_cert /root/key/pcl_key
node2 vm2 pcl O/gm+AYuWwE7ow3dgVG/Nw== cycle /root/cert/pcl_cert /root/key/pcl_key
```

Create /etc/opt/SMAW/SMAWsf/ SA_vmk5r.cfg and then set the owner, group, and access rights as follows.

```
# chown root:root /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg
# chmod 600 /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg
```

Note

- Make sure that the /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg file is set correctly. If the setting is incorrect, the shutdown facility cannot be performed normally.

- Make sure that the virtual server name (ServerName) corresponding to the CF node name (CFNameX) of the cluster host of the /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg file is set. If the setting is incorrect, an incorrect node will be forcibly stopped.

3.7.1.2.4 Start the shutdown facility

1. Start the shutdown facility.

   Check if the shutdown facility has been started on all nodes in the cluster system.

   ```
   # sdtool -s
   ```

   On a node where the shutdown facility has already been started, execute the following commands to restart the shutdown facility.

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

   On a node where the shutdown facility has not been started, execute the following command to start the shutdown facility.

   ```
   # sdtool -b
   ```

Information

You can check if the shutdown facility has already been started with the sdtool -s command. If "The RCSD is not running" is displayed, the shutdown facility is not started.

2. Check the status of the shutdown facility.

   Execute the following command with all nodes in the cluster system to check the status of the shutdown facility.

```
# sdtool -s
```

Note

- If "The RCSD is not running" is displayed, there is a failure in the shutdown daemon or shutdown agent settings. Perform the procedure from step 1 to 4 again.
- The password for the user created in "3.1.1 Creating the User for the Forced Stop " expires in 90 days and must be changed periodically. For the procedure on changing a password, refer to "6.1 Changing a Password Periodically".
- The client certificate for the user created in "3.1.1 Creating the User for the Forced Stop " expires in 3 years and must be changed periodically. For the procedure on updating the client certificate, see "6.2 Updating the Certificate File".
- If you changed the virtual server name created in "3.1.4 Creating the Virtual Server for the Cluster Node ", perform the procedure "3.7.1.2.3 Setting up SA_vmk5r Shutdown Agent " and "3.7.1.2.4 Start the shutdown facility " again.
- If you changed a user password created in "3.1.1 Creating the User for the Forced Stop ", perform the procedure "3.7.1.2.3 Setting up SA_vmk5r Shutdown Agent " and "3.7.1.2.4 Start the shutdown facility" again with a new password.

Information

Display results of the sdtool -s command

- If Unknown or Init-ing is displayed in Init State, wait for about one minute, and then check the status again.
- If Unknown is displayed in Shut State, it means that SF has not yet stopped the node. If Unknown is displayed in Init State, it means that SF has not yet initialized SA or tested the route. Unknown is displayed temporarily in Test State or Init State until the actual status can be confirmed.
- If TestFailed is displayed in Test State, 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 InitFailed is displayed in Init State, communication with the endpoint of the regional user management or the compute (standard service) in FJcloud-O is disabled or the setting might have a failure. Check the following and then set the following again.
  After the failure-causing problem is resolved and SF is restarted, the status display changes to InitWorked or TestWorked.

  a. Execute the following command and check if the virtual server on which the cluster host is running can communicate with the endpoint of the regional user management.

     ```
     # curl -k -s -X GET <URL of the endpoint of the regional user management>/v3/
     ```

     If an error occurs, check the following.

     - Application of the necessary OS patch
     If a version of curl displayed by executing rpm -q curl is 7.19.7-43 or earlier, the necessary OS patch is not applied. Perform "3.1.4.6 Application of the Necessary OS Patch."

     - .curlrc must be created.
     Refer to "3.1.4.7 Creating .curlrc" to make sure that .curlrc is created according to the procedure.

     - The security groups and the firewall service in FJcloud-O must be set properly.

     - The virtual router of FJcloud-O must be created.

     - The default router of the cluster host must be set in the virtual router.

     - URL of the endpoint of the regional user management must be correct.

     - The DNS server used in the cluster host must be set.

b. Execute the following command and check if the virtual server on which the cluster host is running can communicate with the endpoint of the compute (standard service).

```
# curl -k -s -X GET <URL of the endpoint of the compute (standard service)>/v2/
```

If the following 404 message was displayed, it is a normal operation.

```
{"nova_error":{"message":"{\"error\": {\"message\": \"Could not find token, .\"
, \"code\": 404, \"title\": \"Not Found\"}}","request_id":
"xxxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx"}}
```

If the following 401 message was displayed, check the following.

```
{"error": {"message": "The request you have made requires authentication.", "code": 401,
"title": "Unauthorized"}}
```

- A user name and a password for forcibly stopping the virtual server must be correct.

- If a password of a user for forcibly stopping the virtual server has not expired (90 days).

Refer to "6.1 Changing a Password Periodically" and change the password.

- An appropriate role must be set in a user for forcibly stopping the virtual server.

Refer to "3.1.1 Creating the User for the Forced Stop" and make sure that the role is set.

If a message other than the above 404 message or the above 401 message was displayed, check the following.

- The security groups and the firewall service in FJcloud-O must be set properly.

- The virtual router of FJcloud-O must be created.

- The default router of the cluster host must be set in the virtual router.

- URL of the endpoint of the compute (standard service) must be correct.

- The DNS server used in the cluster host must be set.

c. Make sure that the following settings are correct:

- The domain name, project name, URL of the endpoint for the regional user management, and URL of the endpoint for the compute (standard service) for the FJcloud-O environment information file (/opt/SMAW/SMAWRrms/etc/k5_endpoint.cfg)

- All of CF node name, virtual server name, user name, and encrypted password in the configuration file of the shutdown agent (/etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg)

3.7.1.3 Initial Setup of the Cluster Resource Management Facility

Refer to "5.1.3 Initial Setup of the Cluster Resource Management Facility" in "PRIMECLUSTER Installation and Administration Guide" to set up the resource database managed by the cluster resource management facility. In this setting, set the iSCSI device used in the mirroring among the servers of GDS and register it to the resource database.

3.7.2 Setting up Fault Resource Identification and Operator Intervention Request

Refer to "5.2 Setting up Fault Resource Identification and Operator Intervention Request" in "PRIMECLUSTER Installation and Administration Guide" to set up the fault resource identification and the operator intervention request.

3.8 Building the Cluster Application

For the detail on how to build the cluster application, refer to "Chapter 6 Building Cluster Applications" in "PRIMECLUSTER Installation and Administration Guide."

Set the mirroring among the servers of GDS (creating netmirror volume) in this setting.

It is not necessary to set "6.2 Initial GLS Setup" in "PRIMECLUSTER Installation and Administration Guide" because it has been already set in "3.6.1 Initial GLS Setup" above.

Note

- Change and set the values of the following tuning parameters configured with "Setting Tuning Parameters" in setting procedures of the iSCSI device for GDS ("Disk Setting for Performing Mirroring among Servers" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide").

Tuning parameter name	Value after change
ED_CMD_RETRY_COUNT	100
ED_DRV_RETRY_COUNT	100

Example:

```
ED_CMD_RETRY_COUNT=100
ED_DRV_RETRY_COUNT=100
```

To extend the timeout period (CLUSTER_TIMEOUT) of the CF heartbeat, change the above parameter values according to the following formula. Round up the values after the decimal point.

Calculation formula:

```
<Increased CLUSTER_TIMEOUT> / 3 + 100
```

Also, specify the IP address of the public LAN (used also for the administrative LAN) as the IP address for the mirroring among servers used in "Creating iSCSI Target," and "Establishing iSCSI Session" of the above manual.

- If the icmp communication between cluster nodes is not allowed in the security group configuration, the following message is displayed when the clchkcluster command is executed.

```
Admin IP <IP address> used by SF is not alive.
```

If this message is output, refer to "3.1.2.3 Creating the Security Group for the Public LAN (Used also for the Administrative LAN)", and set the icmp protocol rule to allow the icmp communication between cluster nodes. After that, execute the clchkcluster command again.

- When using the Symfoware Server (Native) Hot Standby feature, the standby and scalable userApplication configurations are required.

Set the standby userApplication attribute to the following value.

Attribute	Setup value
StandbyTransitions	ClearFaultRequest | SwitchRequest

Refer to "6) Scalable operation" in "6.7 Setting Up Cluster Applications" in the "PRIMECLUSTER Installation and Administration Guide" for detailed build instructions.

Chapter 4 Operations

For details on functions for managing PRIMECLUSTER system operations, refer to "Chapter 7 Operations" in "PRIMECLUSTER Installation and Administration Guide."

See

For detail on how to operate GDS, refer to "Operation and Maintenance" of "PRIMECLUSTER Global Disk Services Configuration and Administration Guide" and how to operate GLS, refer to "GLS operation on cluster systems" in "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function."

Note

In an FJcloud-O environment, a heartbeat may fail due to an error of the network node or an error of the storage controller, or scheduled maintenance for the infrastructure. This may switch the cluster application.

Chapter 5 Changing the Configurations

For details on changing the configuration information for the PRIMECLUSTER system, environment settings, the configuration of the cluster application, the operation attributes of the cluster system, refer to "Chapter 9 Changing the Cluster System Environment", "Chapter 10 Configuration Change of Cluster Applications", "Chapter 11 Changing the Operation Attributes of a Cluster System" in "PRIMECLUSTER Installation and Administration Guide." For details on changing the GDS configuration, refer to "Configuration Change" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

5.1 Changing the Authentication Method of the User for the Forced Stop

This section describes how to change the authentication method of the user for the forced stop.

1. Execute the following command on all nodes to stop the shutdown facility.

```
# sdtool -e
```

2. On the FJcloud portal, change the authentication method of the user for the forced stop.

   If you change from "Password Authentication" to "Certificate + Password Authentication", issue and download the client certificate of the user from the FJcloud portal.

   If you have already downloaded the client certificate, you do not need to do so.

   Also, refer to "3.7.1.2.1 Distributing Client Certificates for the User for the Forced Stop"

   to create the client certificate file for PRIMECLUSTER and private key file from the client certificate.

   See

   For details on the FJcloud portal, refer to the official FUJITSU Hybrid IT Service FJcloud documentation.

3. Modify the configuration definition file of the SA_vmk5r shutdown agent on all nodes.

   For the description of the configuration definition file, refer to step 2 in "3.7.1.2.3 Setting up SA_vmk5r Shutdown Agent"

4. Execute the following command on all nodes to start the shutdown facility.

```
# sdtool -b
```

5. Execute the following command on all nodes and make sure that the shutdown facility operates normally.

```
# sdtool -s
```

Chapter 6 Maintenance

When you maintain the PRIMECLUSTER system in an FJcloud-O environment, note the following points:

- For the procedure for applying/deleting urgent corrections in an FJcloud-O environment, refer to "6.3 Software Maintenance."
- For details on other items and procedures required for maintenance of the PRIMECLUSTER system, refer to "Chapter 12 Maintenance of the PRIMECLUSTER System" in "PRIMECLUSTER Installation and Administration Guide." For details on how to maintain GDS, refer to "Operation and Maintenance" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide." For details on how to maintain GLS, refer to "Maintenance" in "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function."

6.1 Changing a Password Periodically

A user created in "3.1.1 Creating the User for the Forced Stop" needs a periodical change of the password (every 90 days). If you do not change the password even after 90 days, the shutdown facility will not be operated.

To change a password, perform the following procedure.

1. Change a user password created in FJcloud-O.

2. Set the shutdown facility again with the changed password by following the instructions in "3.7.1.2.3 Setting up SA_vmk5r Shutdown Agent " and "3.7.1.2.4 Start the shutdown facility".

If you do not change the password even after 90 days or do not set the shutdown facility again even if you change the password, the following message is displayed in the file, /var/log/messages and TestState displayed with sdtool -s will be "TestFailed."

```
SF: The authentication request failed.
SMAWsf : SA SA_vmk5r to test host <CF node name>  failed
```

6.2 Updating the Certificate File

A user created in "3.1.1 Creating the User for the Forced Stop" needs a periodical change of the client certificate (every 3 years).If you do not update the client certificate even after 3 years, the shutdown facility will not be operated.

Use caution when you issue a new client certificate from the FJcloud portal, as the client certificate file for PRIMECLUTER and private key file is no longer valid.

1. Execute the following command on all nodes to stop the shutdown facility.

```
# sdtool -e
```

2. From the FJcloud portal, issue and download the new client certificate for the user for the forced stop.

   At this point, the previous client certificate file for PRIMECLUSTER and private key file is no longer valid.

   See

   For details on the FJcloud portal, refer to the official FUJITSU Hybrid IT Service FJcloud documentation.

3. Recreate the client certificate file for PRIMECLUSTER and private key file from the new client certificate,

   as described in "3.7.1.2.1 Distributing Client Certificates for the User for the Forced Stop".

   Re-create the client certificate file for PRIMECLUSTER and private key file in the same directory with the same file name as before.

4. Execute the following command on all nodes to start the shutdown facility.

```
# sdtool -b
```

5. Execute the following command on all nodes and make sure that the shutdown facility operates normally.

```
# sdtool -s
```

6.3 Software Maintenance

6.3.1 Notes on Applying Corrections to the PRIMECLUSTER System

For details on notes for applying an intensive correction to the cluster system, refer to "12.3.1 Notes on Applying Corrections to the PRIMECLUSTER System" in "PRIMECLUSTER Installation and Administration Guide."

Note

In an FJcloud-O environment, refer to "6.3.2 Overview of the Procedure for Applying/Deleting Corrections" to apply/delete the corrections in multi-user mode.

6.3.2 Overview of the Procedure for Applying/Deleting Corrections

Overview of the procedure is shown for applying each correction including an intensive correction to the cluster system in an FJcloud-O environment. In an environment that does not use GDS, the procedure related to GDS is not necessary.

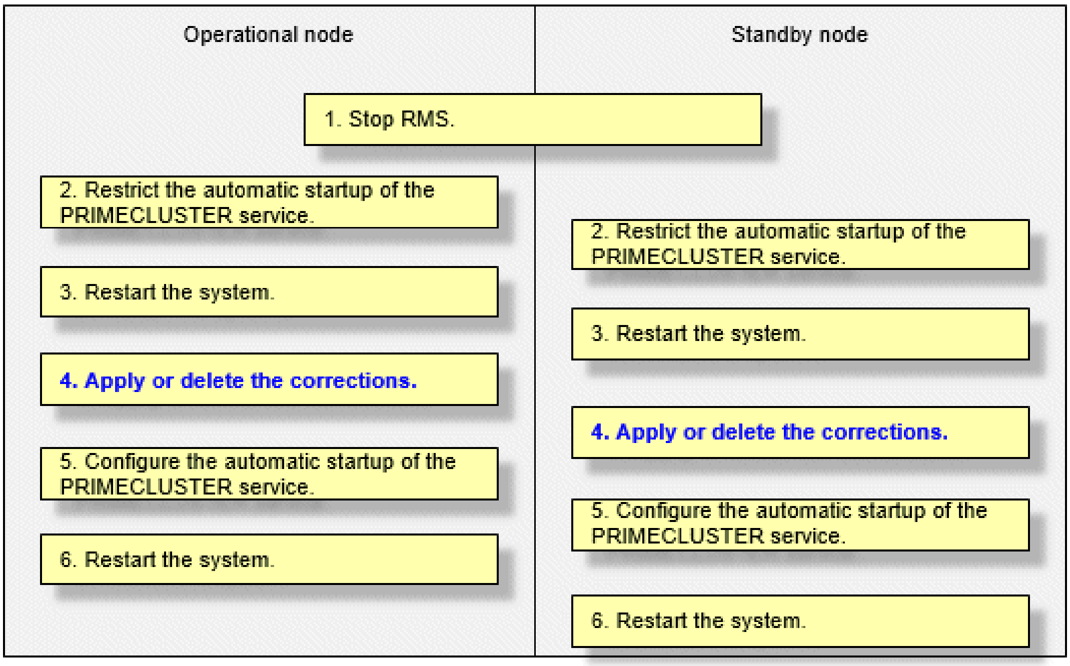
Note

Before applying/deleting corrections to PRIMECLUSTER, take a snapshot of the system storage.

6.3.2.1 Procedure for Applying/Deleting Corrections by Stopping the Entire System

This section explains the procedure for applying/deleting corrections by stopping the entire cluster system.

Flow of the operation


Operational node
Standby node
1. Stop RMS.
2. Restrict the automatic startup of the PRIMECLUSTER service.
3. Restart the system.
4. Apply or delete the corrections.
5. Configure the automatic startup of the PRIMECLUSTER service.
6. Restart the system.
2. Restrict the automatic startup of the PRIMECLUSTER service.
3. Restart the system.
4. Apply or delete the corrections.
5. Configure the automatic startup of the PRIMECLUSTER service.
6. Restart the system.

Operation procedure

Copy the corrections to be applied to each node to the local file system in advance.

1. Stop RMS.

   If RMS is running, execute the following command on any one node to stop RMS.

   ```
   # hvshut -a
   ```

Note

If RMS is stopped on all nodes during the synchronization copying of the GDS volume, the synchronization copying of the entire volume area is performed after the corrections are applied and all nodes are restarted.

If you do not want to perform the synchronization copying of the entire area of volume, stop RMS after the synchronization copying is completed.

To check the slice status of the GDS volume, execute the following command.

Execute the following command on any one node to check the value of the STATUS field of the command output.

The status of the copy destination slice is COPY during the synchronization copying, and after copying is complete, the status becomes ACTIVE or STOP.

```
# sdxinfo -S
```

2. Restrict the automatic startup of the PRIMECLUSTER service.

   Restrict the automatic startup of the PRIMECLUSTER service by executing the following command on all nodes.

   ```
   # /opt/FJSVpclinst/bin/pclservice off
   ```

3. Restart the system.

   Restart the system on all nodes.

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

4. Apply or delete the corrections.

   Apply the corrections that were copied to the local file system or delete the corrections.

   - Applying corrections

     Copy the corrections to the working directory and then execute the following commands.

     ```
     # cd <working directory>
     # /opt/FJSVfupde/bin/uam add -d ./ -i <correction number>
     ```

     If the following message is displayed, select "Y".

     ```
     It is required to update with single user mode. Do you want to apply the update now? (Y/N)Y
     ```

     After that, the following message is displayed. Select "N".

     ```
     Do you want to restart your computer immediately? (Y/N)N
     ```

   - Deleting corrections

     Execute the following command.

     ```
     # /opt/FJSVfupde/bin/uam remove -i <correction number>
     ```

     If the following message is displayed, select "Y".

     ```
     It is required to restore with single user mode. Do you want to restore the updated product
     to its pre-update state now? (Y/N)Y
     ```

     After that, the following message is displayed. Select "N".

     ```
     Do you want to restart your computer immediately? (Y/N)N
     ```

5. Configure the automatic startup of the PRIMECLUSTER service.

   Execute the following command on all nodes and change the PRIMECLUSTER service settings back to the state they were in before they were restricted in step 2.

```
# /opt/FJSVpclinst/bin/pclservice on
```

6. Restart the system.

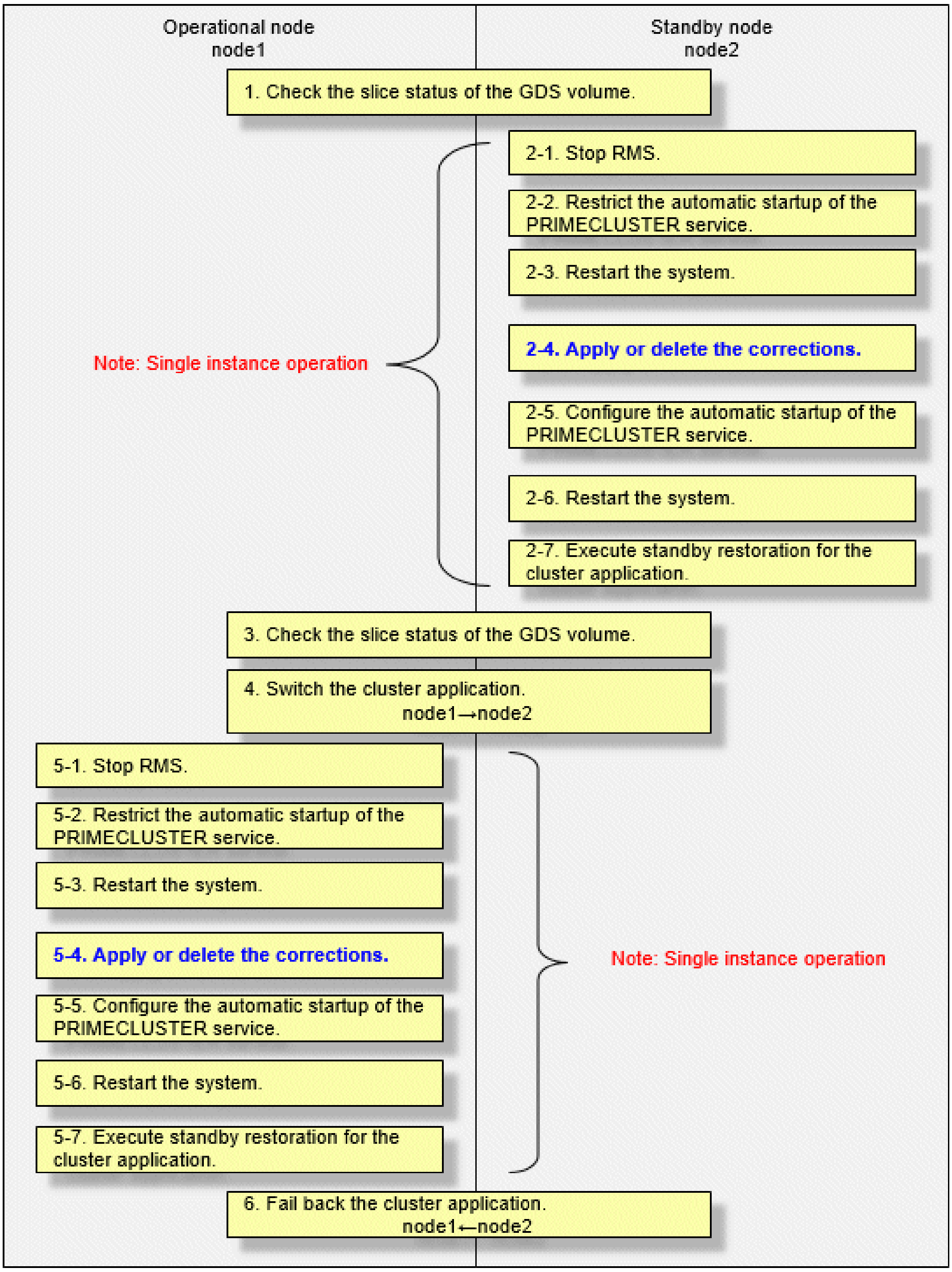
   Restart the system on all nodes.

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

6.3.2.2 Procedure for Applying/Deleting Corrections by Rolling Update

This section explains the procedure for applying corrections by rolling update.

Flow of the operation


Operational node
node1
Standby node
node2
1. Check the slice status of the GDS volume.
2-1. Stop RMS.
2-2. Restrict the automatic startup of the PRIMECLUSTER service.
2-3. Restart the system.
Note: Single instance operation
2-4. Apply or delete the corrections.
2-5. Configure the automatic startup of the PRIMECLUSTER service.
2-6. Restart the system.
2-7. Execute standby restoration for the cluster application.
3. Check the slice status of the GDS volume.
4. Switch the cluster application.
node1→node2
5-1. Stop RMS.
5-2. Restrict the automatic startup of the PRIMECLUSTER service.
5-3. Restart the system.
5-4. Apply or delete the corrections.
5-5. Configure the automatic startup of the PRIMECLUSTER service.
5-6. Restart the system.
5-7. Execute standby restoration for the cluster application.
Note: Single instance operation
6. Fail back the cluster application.
node1←node2

GDS: Global Disk Services

Operation procedure:

1. Check the slice status of the GDS volume.

   Execute the following command on any cluster node to check the value of the STATUS field of the command output.

   ```
   # sdxinfo -S
   ```

   If the COPY status slice exists in the netmirror volume, wait until the synchronization copying is complete.

   For problems caused by node operations during copying, refer to "Stopping or Restarting the Node" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

2. Execute the following operation with the standby node (node2).

    1. Stop RMS.

        Stop RMS to apply corrections to the standby node (node2). A cutoff state transition occurs according to the shutdown of RMS. In this case, make sure that the single instance operation continues until the standby restoration for the cluster application is executed.

        ```
        # hvshut -l
        ```

    2. Restrict the automatic startup of the PRIMECLUSTER service.

        Execute the following command to restrict the automatic startup of the PRIMECLUSTER service.

        ```
        # /opt/FJSVpclinst/bin/pclservice off
        ```

    3. Restart the system.

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

    4. Apply or delete the corrections.

        - Applying corrections

        Copy the corrections to the working directory and then execute the following commands.

        ```
        # cd <working directory>
        # /opt/FJSVfupde/bin/uam add -d ./ -i <correction number>
        ```

        If the following message is displayed, select "Y".

        ```
        It is required to update with single user mode. Do you want to apply the update now? (Y/N)Y
        ```

        After that, the following message is displayed. Select "N".

        ```
        Do you want to restart your computer immediately? (Y/N)N
        ```

        - Deleting corrections

        Execute the following command.

        ```
        # /opt/FJSVfupde/bin/uam remove -i <correction number>
        ```

        If the following message is displayed, select "Y".

        ```
        It is required to restore with single user mode. Do you want to restore the updated
        product to its pre-update state now? (Y/N)Y
        ```

        After that, the following message is displayed. Select "N".

        ```
        Do you want to restart your computer immediately? (Y/N)N
        ```

    5. Configure the automatic startup of the PRIMECLUSTER service.

        Execute the following command and change the PRIMECLUSTER service settings back to the state they were in before they were restricted in 2 of step 2.

        ```
        # /opt/FJSVpclinst/bin/pclservice on
        ```

    6. Restart the system.

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

    7. Execute standby restoration for the cluster application.

        If the node (node2) to which corrections have been applied is cut off from the cluster system, execute standby restoration for the node.

For details on how to execute cluster application standby restoration, refer to "7.2.2.1 Starting a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

3. Check the slice status of the GDS volume.

   After starting the standby node (node2), the synchronization copying of the netmirror volume is executed. Make sure that the synchronization copying is completely finished and all slices are either in ACTIVE or STOP status on any one node.

   To check the slice status of the netmirror volume, execute the following command:

   Execute the following command on any cluster node to check the value of the STATUS field of the command output.

   ```
   # sdxinfo -S
   ```

4. Switch the cluster application.

   To apply corrections to the operational node (node1), execute hvswitch and switch all cluster applications to the standby node (node2). For details on how to switch the cluster applications, refer to "7.2.2.3 Switching a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

5. Perform the following operation with the old operational node (node1).

   1. Stop RMS.

      Stop RMS to apply corrections to the operational node (node1). A cutoff state transition occurs according to the shutdown of RMS. In this case, make sure that the single instance operation continues until the standby restoration for the cluster application is executed.

      ```
      # hvshut -l
      ```

   2. Restrict the automatic startup of the PRIMECLUSTER service.

      Execute the following command to restrict the automatic startup of the PRIMECLUSTER service.

      ```
      # /opt/FJSVpclinst/bin/pclservice off
      ```

   3. Restart the system.

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

   4. Apply or delete the corrections.

      - Applying corrections

      Copy the corrections to the working directory and then execute the following commands.

      ```
      # cd <working directory>
      # /opt/FJSVfupde/bin/uam add -d ./ -i <correction number>
      ```

      If the following message is displayed, select "Y".

      ```
      It is required to update with single user mode. Do you want to apply the update now? (Y/N)Y
      ```

      After that, the following message is displayed. Select "N".

      ```
      Do you want to restart your computer immediately? (Y/N)N
      ```

      - Deleting corrections

      Execute the following command.

      ```
      # /opt/FJSVfupde/bin/uam remove -i <correction number>
      ```

      If the following message is displayed, select "Y".

      ```
      It is required to restore with single user mode. Do you want to restore the updated
      product to its pre-update state now? (Y/N)Y
      ```

      After that, the following message is displayed. Select "N".

```
Do you want to restart your computer immediately? (Y/N)N
```

    5. Configure the automatic startup of the PRIMECLUSTER service.

       Execute the following command and change the PRIMECLUSTER service settings back to the state they were in before they were restricted in 2 of step 5.

       ```
       # /opt/FJSVpclinst/bin/pclservice on
       ```

    6. Restart the system.

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

    7. Execute standby restoration for the cluster application.

       If the node (node1) to which corrections have been applied is cut off from the cluster system, execute standby restoration for the node. For details on how to execute cluster application standby restoration, refer to "7.2.2.1 Starting a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

6. Fail back the cluster application.

   Restore the state of the standby layout defined at installation by executing failback operation, as necessary. For details on failback, refer to "7.2.2.3 Switching a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

6.4 Procedure for Restoring OS with the Snapshot Function

If OS is restored with the snapshot function of FJcloud-O, perform the following procedure.

Note

If GLS is being used, when OS is restored from the snapshot according to the procedures described in this section, an interface configuration file of the physical network is rewritten to the initial state at the time when the virtual server is created. This duplicates the IP address of the physical NIC and the IP address of the GLS virtual interface, and functions or applications that are using the IP address of the virtual interface may not operate normally.

Therefore, if GLS is being used, acquire the snapshot with the following procedure.

[Procedure for acquiring the snapshot]

1. Stores the interface configuration of the physical network used by the GLS.

   [RHEL8]

   Back up the interface configuration file.

   Example:

   ```
   # cp /etc/sysconfig/network-script/ifcfg-eth1 /etc/sysconfig/network-script/ifcfg-eth1.bak
   ```

   [RHEL9]

   Check if the IP address configured with the nmcli connection show command.

2. Set the automatic startup of RMS to OFF.

   ```
   # hvsetenv HV_RCSTART 0
   # hvsetenv HV_RCSTART
   0 <- Check if "0" is output
   ```

3. Stop the virtual server from which the snapshot is acquired.

4. Acquire the snapshot.

5. After starting the virtual server, start RMS, and set the automatic startup to ON.

   Startup of RMS:

```
# hvcm -s <SysNode name of the virtual server from which the snapshot is acquired>
```

Automatic startup of RMS is set to ON:

```
# hvsetenv HV_RCSTART 1
# hvsetenv HV_RCSTART
1 <- Check if "1" is output
```

6.4.1 Procedure for Restoring One Node While the Operation is Working

1. Refer to "6.4.3 Restoring the Virtual Server from the Snapshot" to restore the virtual server.

   Note

   If you did not set the expanded volume used in the mirroring among the servers of GDS according to "6.4.3 Restoring the Virtual Server from the Snapshot", refer to "6.4.3 Restoring the Virtual Server from the Snapshot" to restore the virtual server again without attaching the target volume to the restored virtual server. If the target volume is attached to the restored virtual server, the process fails in the rest of the procedure.

2. When using the mirroring among the servers of GDS, check the slice status. If the status of the slice is INVALID, execute the following command to perform the synchronization copying of each volume after the node is started.

   ```
   # sdxcopy -B -c <class name> -v <volume name>
   ```

Note

If the virtual server name is changed in step 1, perform step 3 to 5 described in "3.7.1.2 Setting up the Shutdown Facility" on all nodes.

6.4.2 Procedure for Restoring Nodes While the Operation does not Work

1. If either or all of the nodes are started before restoring the nodes, stop RMS. Perform the restore procedure on either node that is started.

   ```
   # hvshut -a
   ```

2. Select the latest disk when all nodes are started before restoring nodes in an environment where the mirroring among the servers of GDS is used. For all classes of GDS, execute the following command on either node.

   ```
   # /etc/opt/FJSVsdx/bin/sdxnetdisk -S -c <class name>
   ```

3. Refer to "6.4.3 Restoring the Virtual Server from the Snapshot" to restore one node and then start the node.

   Note

   If you did not set the expanded volume used in the mirroring among the servers of GDS according to "6.4.3 Restoring the Virtual Server from the Snapshot", refer to "6.4.3 Restoring the Virtual Server from the Snapshot" to restore the virtual server again without attaching the target volume to the restored virtual server. If the target volume is attached to the restored virtual server, the process fails in the rest of the procedure.

4. When using the mirroring among the servers of GDS, start the node and then execute the following with the restored node.

   1. Delete the information of the iSCSI device.

      ```
      # rm -f /var/opt/FJSVsdx/log/.sdxnetmirror_disable.db
      # rm -f /var/opt/FJSVsdx/log/.sdxnetmirror_timestamp
      ```

    2. Stop RMS.

    ```
    # hvshut -l
    ```

5. When restoring the other node, refer to "6.4.3 Restoring the Virtual Server from the Snapshot" to restore the node.

    Note

    If the volume used in the mirroring among the servers of GDS is attached after creating a virtual server, the process fails in the rest of the procedure. Refer to "6.4.3 Restoring the Virtual Server from the Snapshot" to create the virtual server again.

6. When using the mirroring among the servers of GDS, delete the iSCSI device information with the restored node if the node is restored with step 5.

    ```
    # rm -f /var/opt/FJSVsdx/log/.sdxnetmirror_disable.db
    # rm -f /var/opt/FJSVsdx/log/.sdxnetmirror_timestamp
    ```

7. If all nodes are stopped before restoring nodes in an environment where the mirroring among the servers of GDS is used, check the status of the source slice for the synchronization copying. If the source slice for the synchronization copying is INVALID, restore the status of the slice. For the "-d" option of the "sdxfix " command, specify the source disk of the synchronization copying. Perform this procedure on either node.

    ```
    # sdxfix -V -c <class name> -v <volume name> -d <disk name> -x NoRdchk
    ```

8. If 2 of step 4 is performed, start RMS on the node where RMS is stopped.

    ```
    # hvcm
    ```

Note

If the virtual server name is changed in step 3 or step 5, perform step 3 to 5 in "3.7.1.2 Setting up the Shutdown Facility" on all nodes.

6.4.3 Restoring the Virtual Server from the Snapshot

Note

To use the service provided by FJcloud-O IaaS with API, it is necessary to build an environment for using API. For details, refer to the documents provided by FJcloud-O.

1. To restore the virtual server, refer to the documents provided by FJcloud-O.

    If GLS is being used, additionally perform the following procedure.

2. Modify the interface settings of the physical network used by the GLS.

    [RHEL8]

    Overwrite the interface configuration file of the physical network used in GLS with the backup file created in step 1 in "6.4 Procedure for Restoring OS with the Snapshot Function".

    Example:

    ```
    # cp /etc/sysconfig/network-script/ifcfg-eth1.bak /etc/sysconfig/network-script/ifcfg-eth1
    ```

    [RHEL9]

    Use the nmcli connection modify command to modify the following parameters.

    ```
    ipv4.method "manual"
    ipv4.address  <IP address that you checked in step 1 of "6.4 Procedure for Restoring OS with the
    ```

```
Snapshot Function">
ipv4.ignore-auto-dns "yes"
ipv4.never-default "yes"
```

3. Set the automatic startup of RMS to ON.

```
# hvsetenv HV_RCSTART 1
# hvsetenv HV_RCSTART
1 <- Check if "1" is output
```

4. Restart the virtual server.

Part 2 NIFCLOUD Environment

This part describes the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system in a NIFCLOUD environment.

Chapter 7 Cluster System in a NIFCLOUD Environment

PRIMECLUSTER provides clustering for the servers in a NIFCLOUD private LAN. This enables higher business availability than the availability provided by cloud services.

See

For details on NIFCLOUD, refer to the official NIFCLOUD documentation.

The following cluster systems are available in a NIFCLOUD environment:

- Cluster system in multiple zones
- Cluster system in a single zone

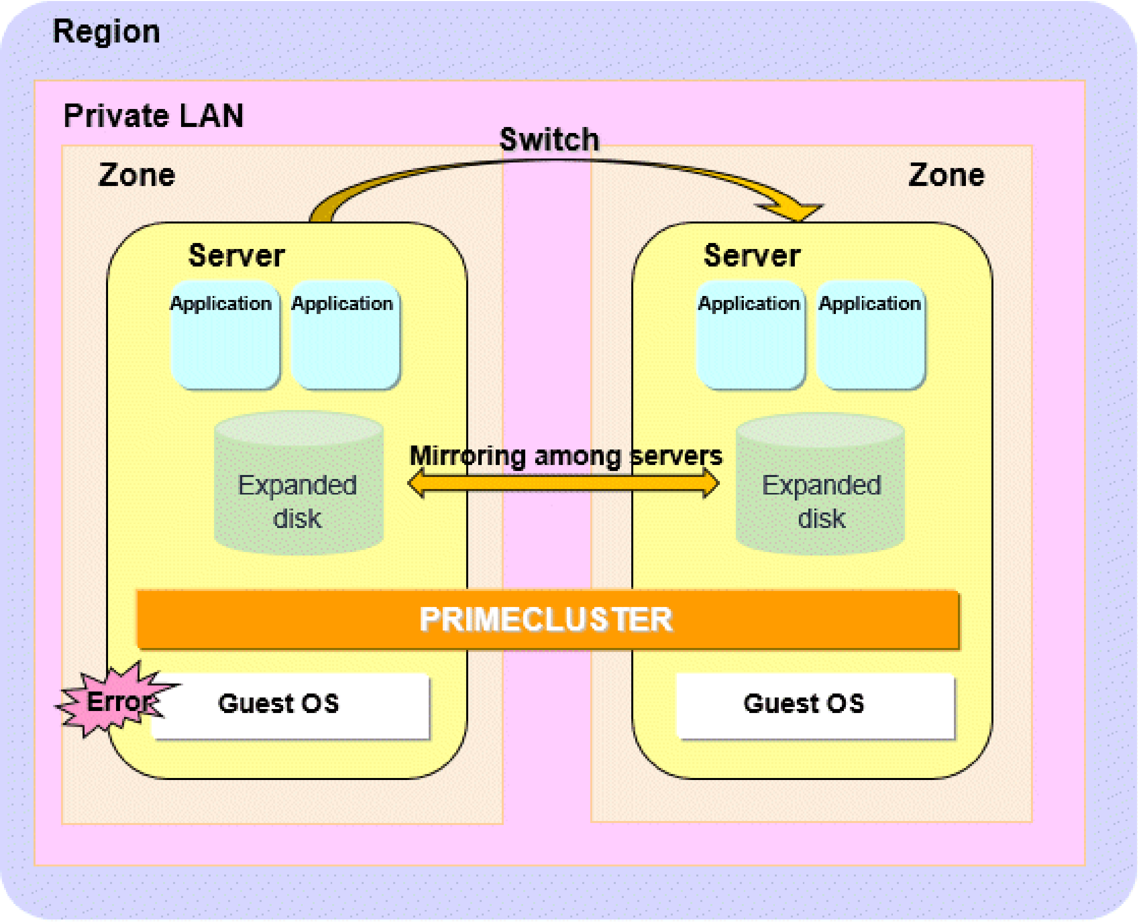
7.1 Cluster System in Multiple Zones

In this configuration, the cluster system can be operated on multiple zones.

By applying PRIMECLUSTER to the server, in the event of an error, an application can be switched from the operational server to the standby server in a short time to provide a highly reliable server environment.

Also, when a network failure occurs in the entire zone or when the zones become abnormal due to a large-scale disaster, business can be quickly restored by failing over to the standby system with fewer operations.

Figure 7.1 Cluster system in multiple zones



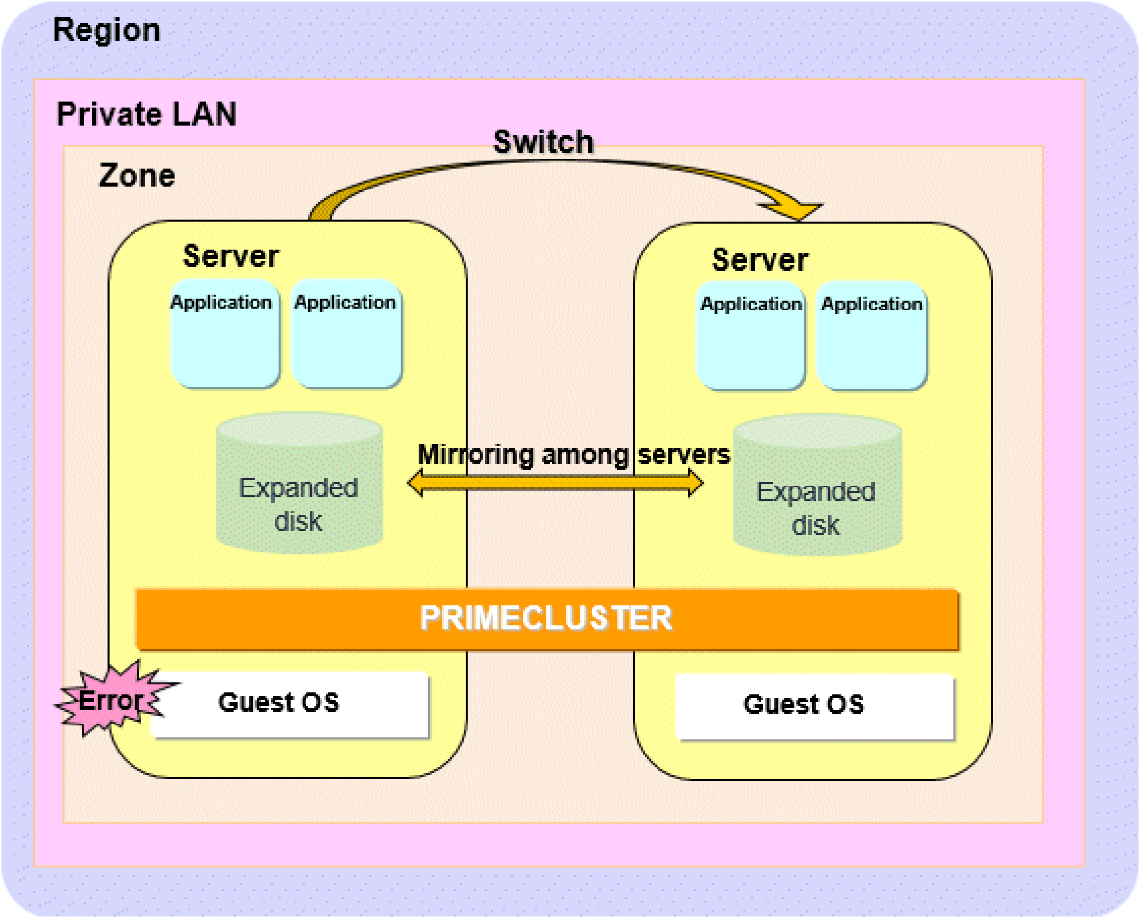
7.2 Cluster System in a Single Zone

In this configuration, the cluster system can be operated in one zone. By applying PRIMECLUSTER to the server, in the event of an error, an application can be switched from the operational server to the standby server in a short time to provide a highly reliable server environment.

Note

In the cluster system in a single zone, in the event of an error in the zone where the cluster is built, all cluster nodes stop and the business stops.

Figure 7.2 Cluster system in a single zone



7.3 Supported Range

This section describes the range of support of PRIMECLUSTER in a NIFCLOUD environment.

Supported configurations

- Number of cluster nodes: 1 to 2 nodes
- Operation mode of the cluster system: 1:1 Standby operation, Mutual standby, Single-node cluster, Scalable operation (only when using Symfoware Server (Native) Hot Standby feature)
- Network configurations

  The servers in the cluster system must communicate with the API endpoints. This setting is not necessary in a single-node cluster.
- Configurations to take over volume data

  Data takeover by the mirroring among servers of GDS using the expanded disk function

Supported monitoring functions

- Error of a zone (cluster system in multiple zones)

  The cyclic monitoring of the cluster interconnect detects an error of a zone, and the node becomes LEFTCLUSTER.

- Error of the public LAN

  The network monitoring using ICMP detects a route failure, and the service is automatically switched to the standby system.

- Error of the guest OS

  The cyclic monitoring of the cluster interconnect detects an error of the guest OS, and the service is automatically switched to the standby system.

- Error of the cluster interconnect

  The cyclic monitoring of the cluster interconnect detects an error of the cluster interconnect, and the service is switched or cut off.

- Error of the disk access

  GDS monitors I/O to a disk, and when an error of the disk access occurs, the disk is detached and the service continues.

  If an I/O error occurs in all slices in a mirror, the service is automatically switched to the standby system.

- Error of the cluster application

  When a resource error of the cluster application occurs, the service is automatically switched to the standby system.

Note

- Acquiring the backup to the server is recommended when OS is stopped.
- The following functions for PRIMECLUSTER are not available:
  - GFS
  - GLS
  - GDS Snapshot
  - Root class and local class of GDS
  - Single volume of GDS, and disk groups of GDS other than the netmirror type (mirror, stripe, concatenation, and switch)
- In addition to the above functions, the functions of GDS for PRIMECLUSTER are not available in a single-node cluster.
- Among the functions of NIFCLOUD. the following functions are not available.
  - Live migration
  - VM import
  - Auto scale
  - Internet VPN (H/W)
  - Red Hat Enterprise Linux 6 ELS (Extended Life-cycle Support)
  - NAS
  - Dedicated component
  - Private region
- If the additional NIC function is used, the server copy function cannot be used.
- In the cluster system in multiple zones, use the private bridge function for a cluster communication between zones.
- In the cluster system in a single zone, use the server separateness function of NIFCLOUD so that each server starts on a different physical host.
- In the cluster system in a single zone, it is recommended to use the unit separation function of an expanded disk.

- In a NIFCLOUD environment, when an OS panic occurs, the cluster node may be powered off while the memory dump is being output, and it may not be possible to collect a complete memory dump.

Chapter 8 Design

You must prepare the items listed below before building the PRIMECLUSTER system in a NIFCLOUD environment.

- Selecting the PRIMECLUSTER Product
- Selecting the Architectural Pattern
- Network Design
- System Design
- Determining the Cluster System Operation Mode
- Determining the Web-Based Admin View Operation Mode
- Determining the Failover Timing of Cluster Application

Point

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

Information

For the flow to build the PRIMECLUSTER system, refer to "Chapter 1 Build Flow" in "PRIMECLUSTER Installation and Administration Guide."

8.1 Selecting the PRIMECLUSTER Product

Select a PRIMECLUSTER product.

In a NIFCLOUD environment, you can select the following products.

For details on the PRIMECLUSTER products, refer to "2.1 PRIMECLUSTER Product Selection" in "PRIMECLUSTER Installation and Administration Guide."

- PRIMECLUSTER Enterprise Edition (EE)
- PRIMECLUSTER HA Server (HA)
- PRIMECLUSTER Clustering Base (CB)

8.2 Selecting the Architectural Pattern

In the PRIMECLUSTER system in a NIFCLOUD environment, select an architecture pattern for each item below.

- Ensuring Connectivity with the API Endpoint
- Ensuring Connectivity with Web-Based Admin View

8.2.1 Ensuring Connectivity with the API Endpoint

When using PRIMECLUSTER on public clouds, connectivity between the cluster node and the API endpoint must be ensured to deal with split-brain. Because the API endpoint is on the Internet, it must be connected to the Internet. Use the API endpoint to control the power of the cluster node.

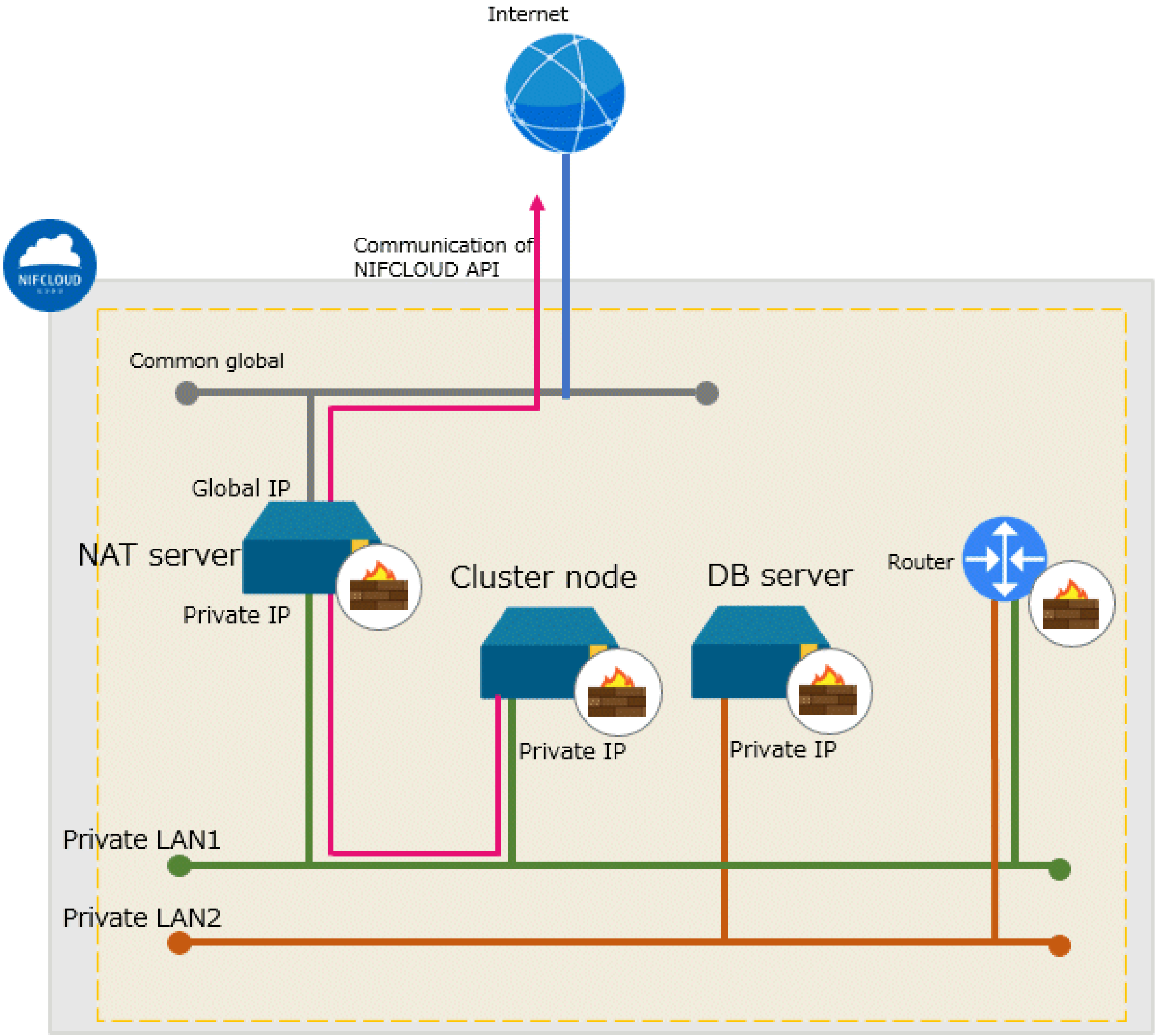
PRIMECLUSTER provides architectural patterns for ensuring the connectivity between the cluster node and the API endpoint. For smooth design of a cluster system, choose from these architectural patterns.

The following are the architectural patterns for ensuring the connectivity between the cluster node and the API endpoint and the appropriate scenarios for each pattern.

Table 8.1 Architectural patterns and appropriate scenarios for ensuring the connectivity between the cluster node and the API endpoint

Architectural pattern	Appropriate scenario	Note
Ensuring connectivity with the router	Ensuring low cost and secure connectivity of back-end servers	Operation management is not required by the user since the NAT function provided by the router of NIFCLOUD is used.
Ensuring connectivity with the NAT server	Ensuring secure connectivity of back-end servers without using the router function	Building and operation management by a user is required for the NAT server. For the following configurations that cannot be configured by the router function, secure the connectivity with the NAT server. - Using multiple global IPs without directly allocating them to the server, such as using different global IP between DNAT and SNAT - Connecting the Internet in an environment where the router is already used, such as in an environment in "Figure 8.1 Connecting only the private LAN1 to the Internet in an environment where the router is already used"
Ensuring connectivity with the global IP address	Ensuring low cost connectivity of front-end servers	To allow the access from the global network to the server, IP permission restrictions or the firewall rule must be firmly established. This architecture is simple because there are a low number of system components since the router is not required.

Figure 8.1 Connecting only the private LAN1 to the Internet in an environment where the router is already used



8.2.1.1 Ensuring Connectivity with the Router

An architecture pattern with the router can block access from the Internet, and it also ensures connectivity between the cluster node and the API endpoint.

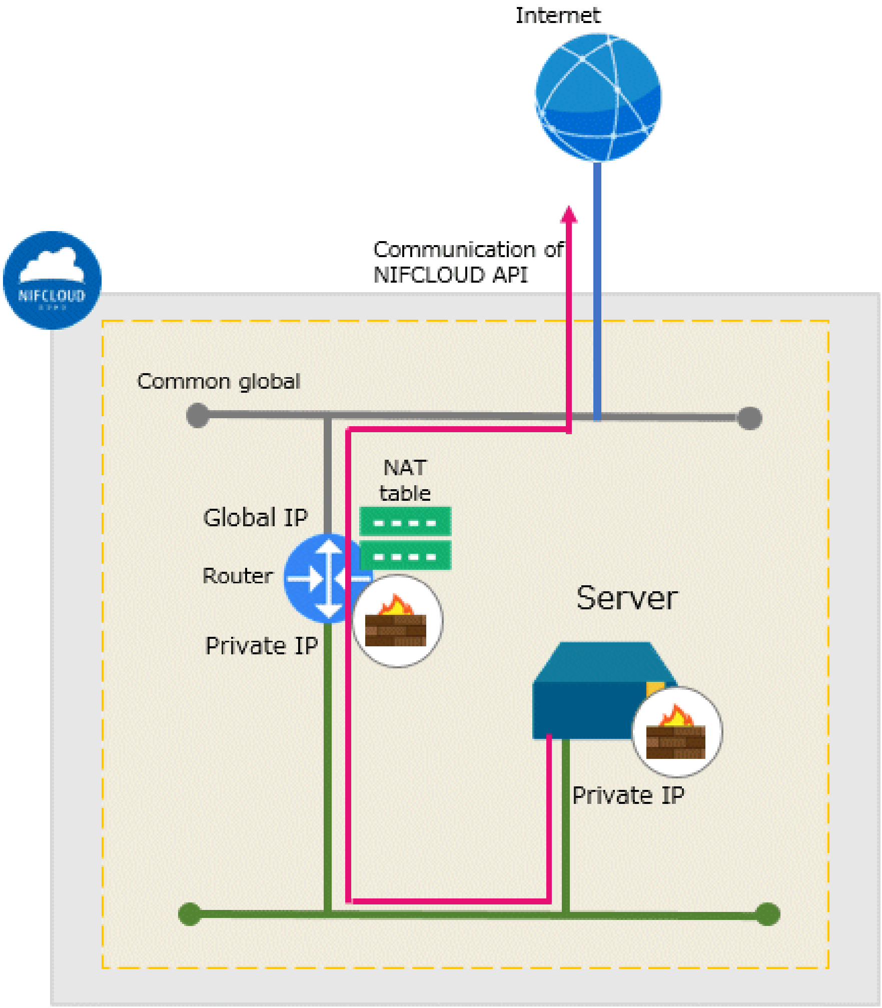
Place the cluster node of back-end servers in a private LAN. This server is not given a global IP address and has no direct connectivity with the Internet.

Use the router for connectivity with the endpoint since the API endpoint that forcibly stops PRIMECLUSTER exists over the Internet.

See

For details on the NAT function of the router, refer to the official NIFCLOUD documentation.

Figure 8.2 Ensuring connectivity with the router

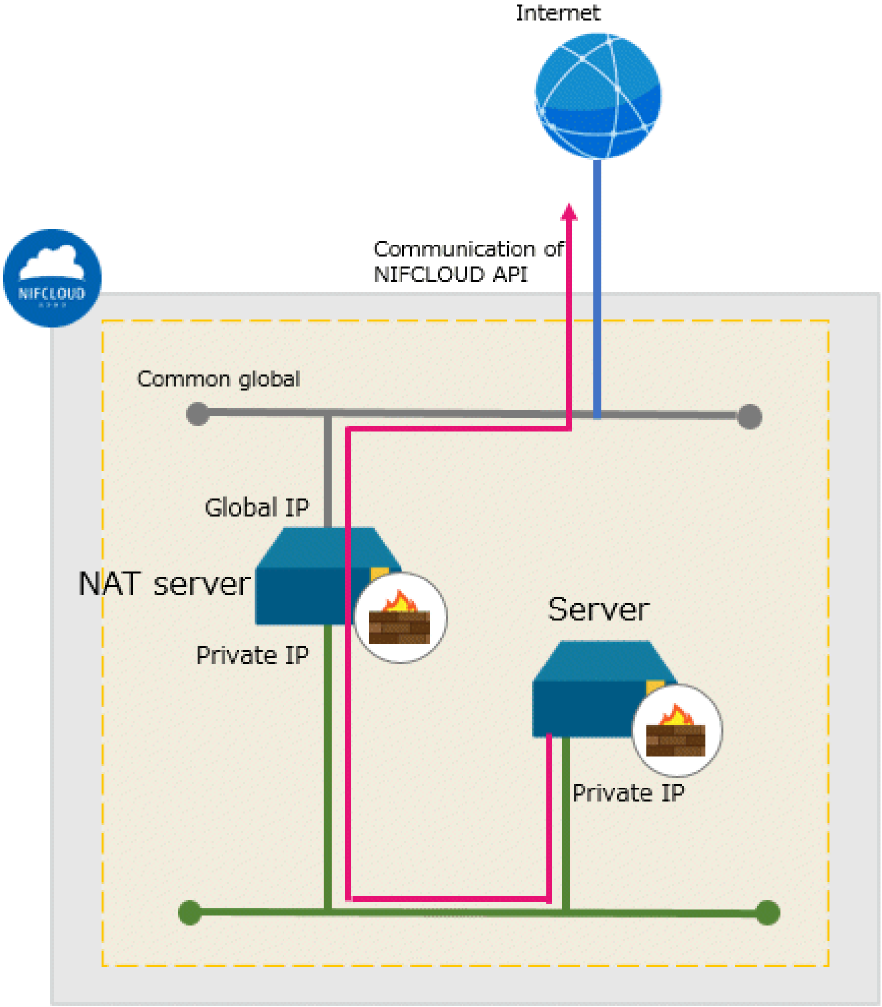


8.2.1.2 Ensuring Connectivity with the NAT Server

An architecture pattern with the NAT server can block access from the Internet, and it also ensures connectivity between the cluster node and the API endpoint.

This architectural pattern is identical to the architecture pattern with the router, except that the router is replaced by the NAT server in the placement of components.

Figure 8.3 Ensuring connectivity with the NAT server



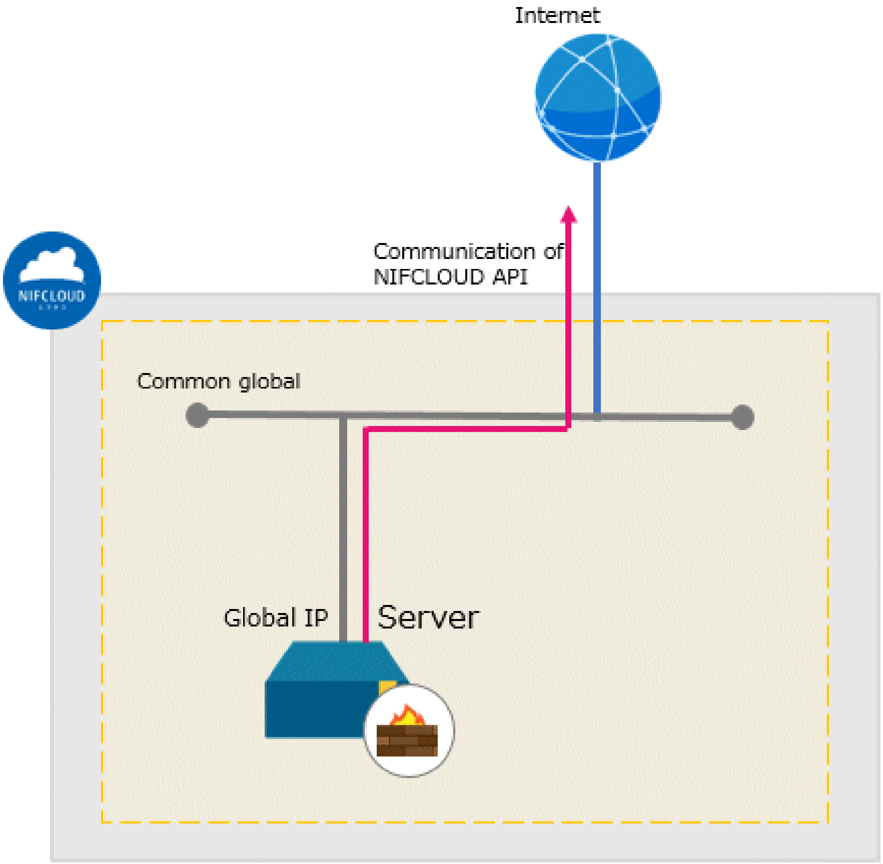
8.2.1.3 Ensuring Connectivity with the Global IP Address

This architectural pattern is simpler than architectural patterns with the router or the NAT server.

However, since the cluster node is accessible from the Internet, IP permission restrictions or the firewall rule must be firmly established.

When using this architectural pattern, you only need to give the global IP address to the cluster node.

Figure 8.4 Ensuring connectivity with the global IP address



8.2.2 Ensuring Connectivity with Web-Based Admin View

PRIMECLUSTER provides architectural patterns for ensuring the connectivity between a terminal directly operated by a user and the management view client.

For smooth designing of a cluster system, choose the appropriate one from these architectural patterns.

The following are the architectural patterns for the connectivity with the Web-Based Admin View and the appropriate scenarios for each pattern.

Table 8.2 Architectural patterns and appropriate scenarios for the connectivity with the Web-Based Admin View

Architectural pattern	Appropriate scenario	Note
Ensuring connectivity with a server for a client	Pattern using a server for a client	On NIFCLOUD, the server for the management view client that is connected with the private LAN for the administrative LAN must be deployed.
Ensuring connectivity using a VPN connection	Pattern using a VPN connection	An additional device is required for a VPN connection.

8.2.2.1 Ensuring Connectivity with a Server for a Client

In the pattern using a server for a client, to ensure the connectivity between a terminal directly operated by a user (remote control terminal of the management view client) and the management view client, prepare a server for the management view client deployed on NIFCLOUD.

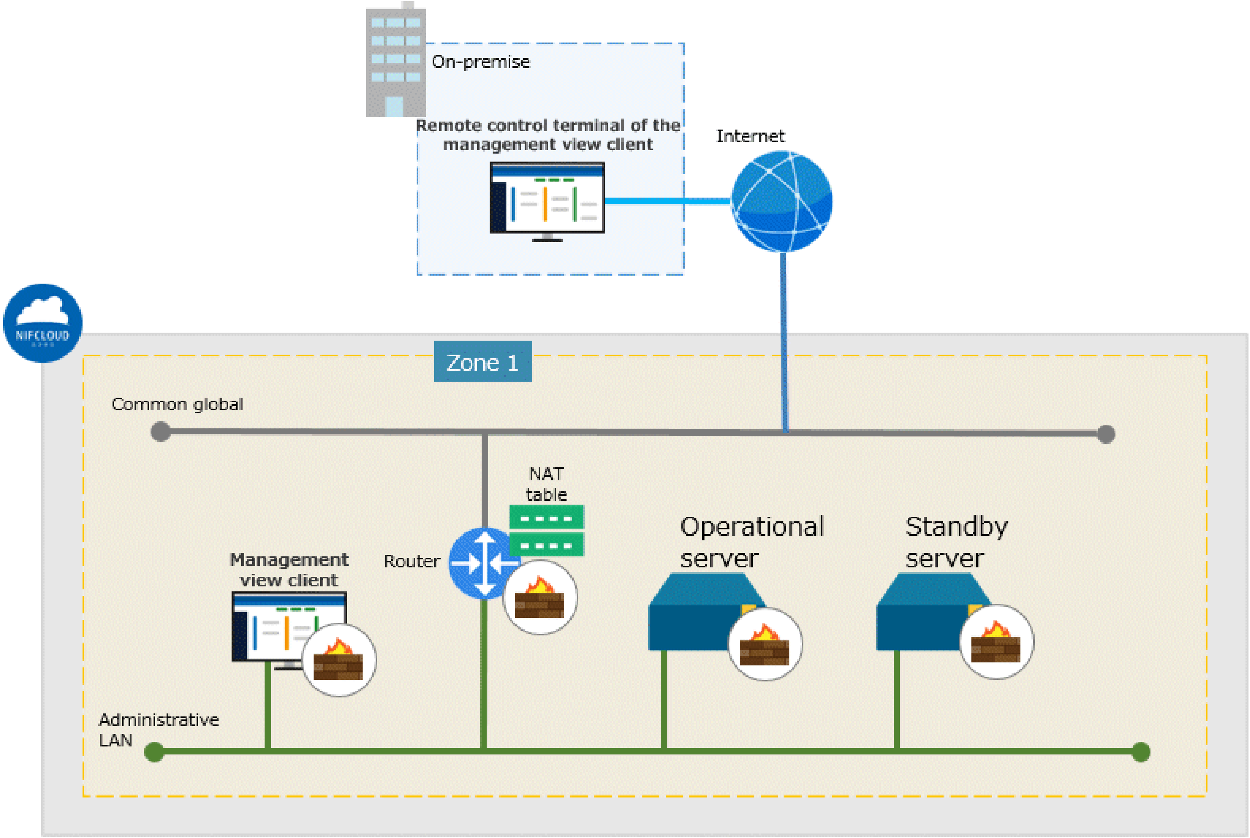
The system component for a VPN connection is not required and the configuration may be simple.

When selecting this architectural pattern, a user connects to the management view client via a remote desktop connection.

For a remote desktop connection from the Internet, it is necessary to connect a router to the Internet and use the DNAT function for a connection via the router, or to build the server connected to the Internet as a reverse proxy server on NIFCLOUD for a connection via the reverse proxy server.

Also, to allow the access to the server from the Internet, the firewall rule must be firmly established.

Figure 8.5 Ensuring connectivity with a server for a client



8.2.2.2 Ensuring Connectivity Using a VPN Connection

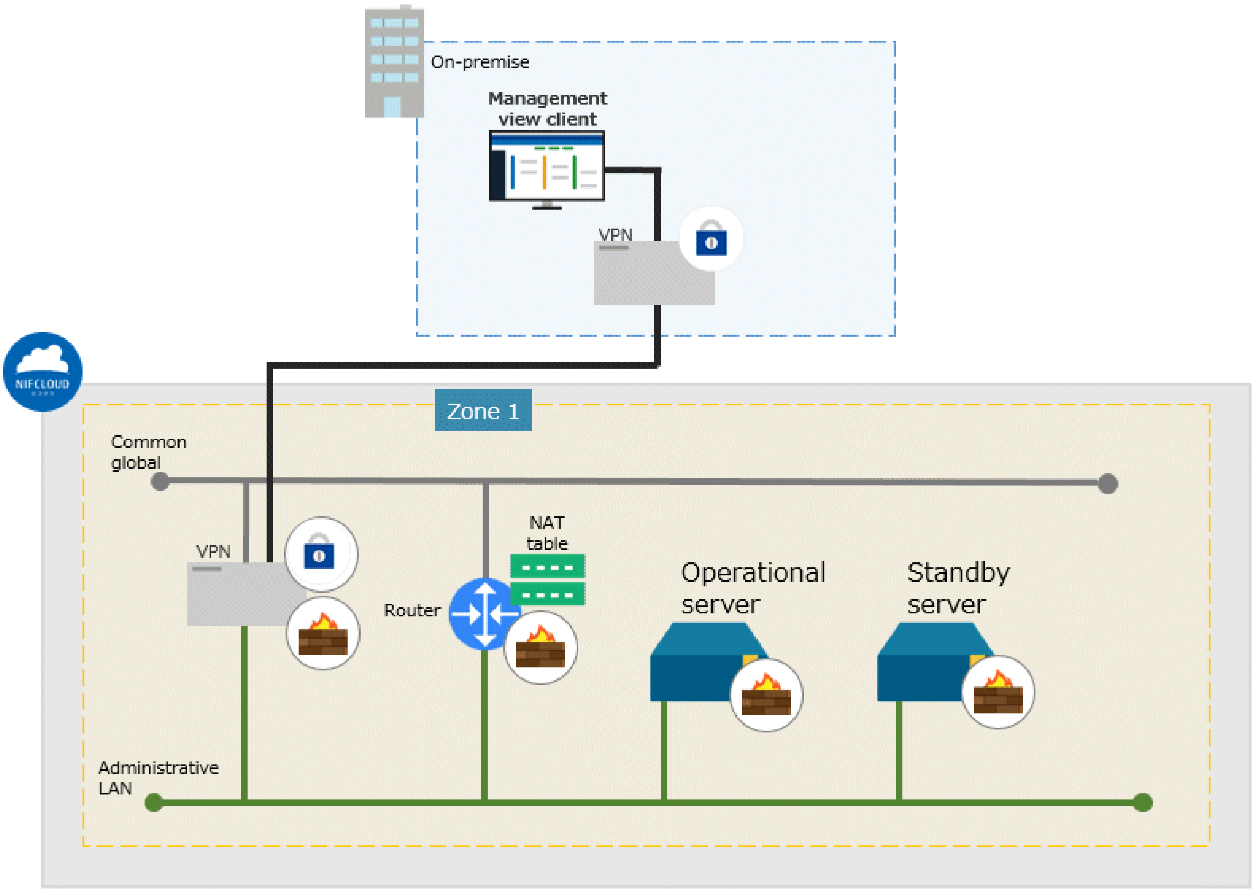
In the pattern using a VPN connection, to ensure the connectivity between the management view client and the operational server, or the standby server, a VPN connection is used.

No management view client is required on NIFCLOUD. It is also possible to securely connect to an environment on NIFCLOUD.

When selecting this architectural pattern, a terminal directly operated by a user is the management view client.

To provide a VPN connection, the VPN must be set or a device is required in the network where the terminal directly operated by a user is deployed.

Figure 8.6 Ensuring connectivity using a VPN connection



8.3 Network Design

In the cluster system in a NIFCLOUD environment, security rules to be applied to the server must be designed in advance.

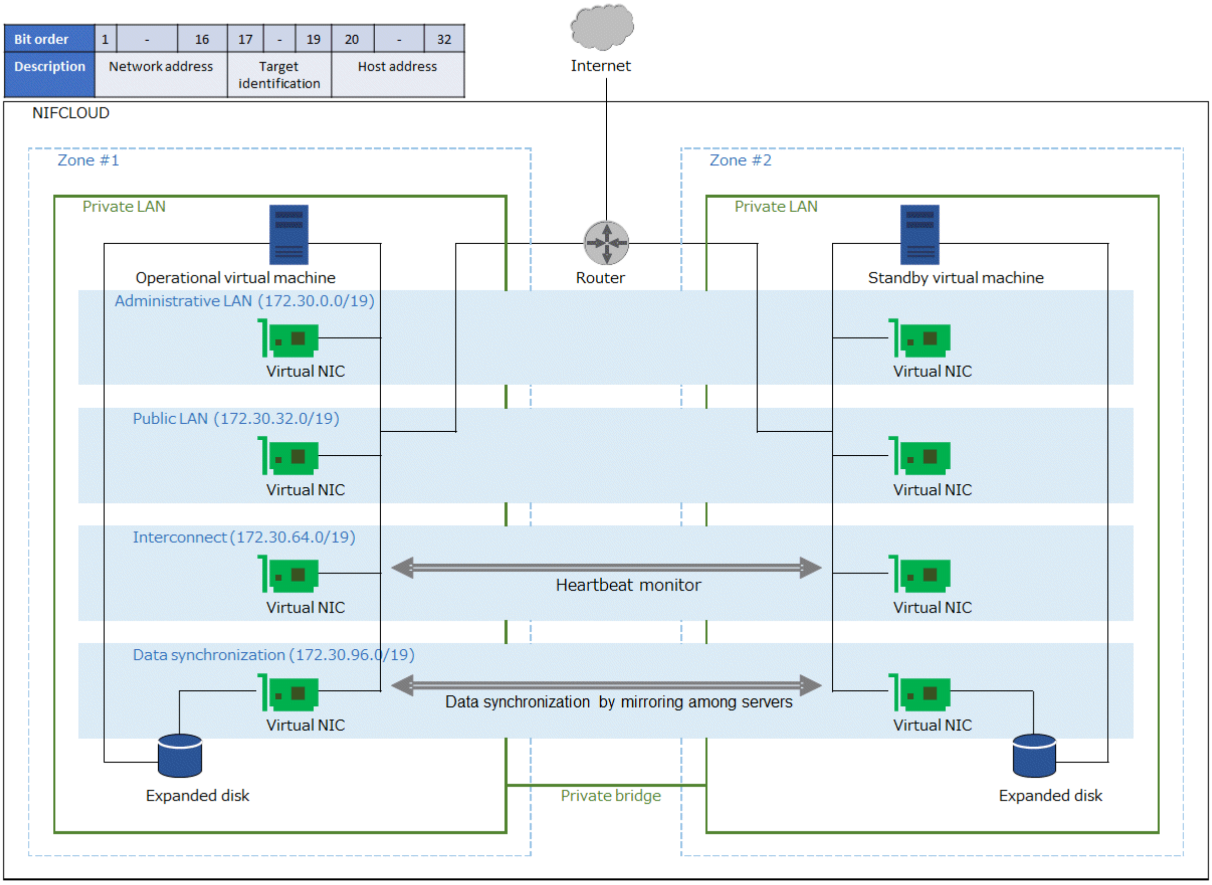
8.3.1 Subnet Design

8.3.1.1 Subnet Design of the Cluster System in Multiple Zones

For the cluster system in multiple zones, connect a private LAN prepared in each zone with a private bridge and assign a subnet for each purpose.

For easier access control, it is recommended that you use the upper bits of the subnet for role identification of the network.

Figure 8.7 Subnet design of the cluster system in multiple zones

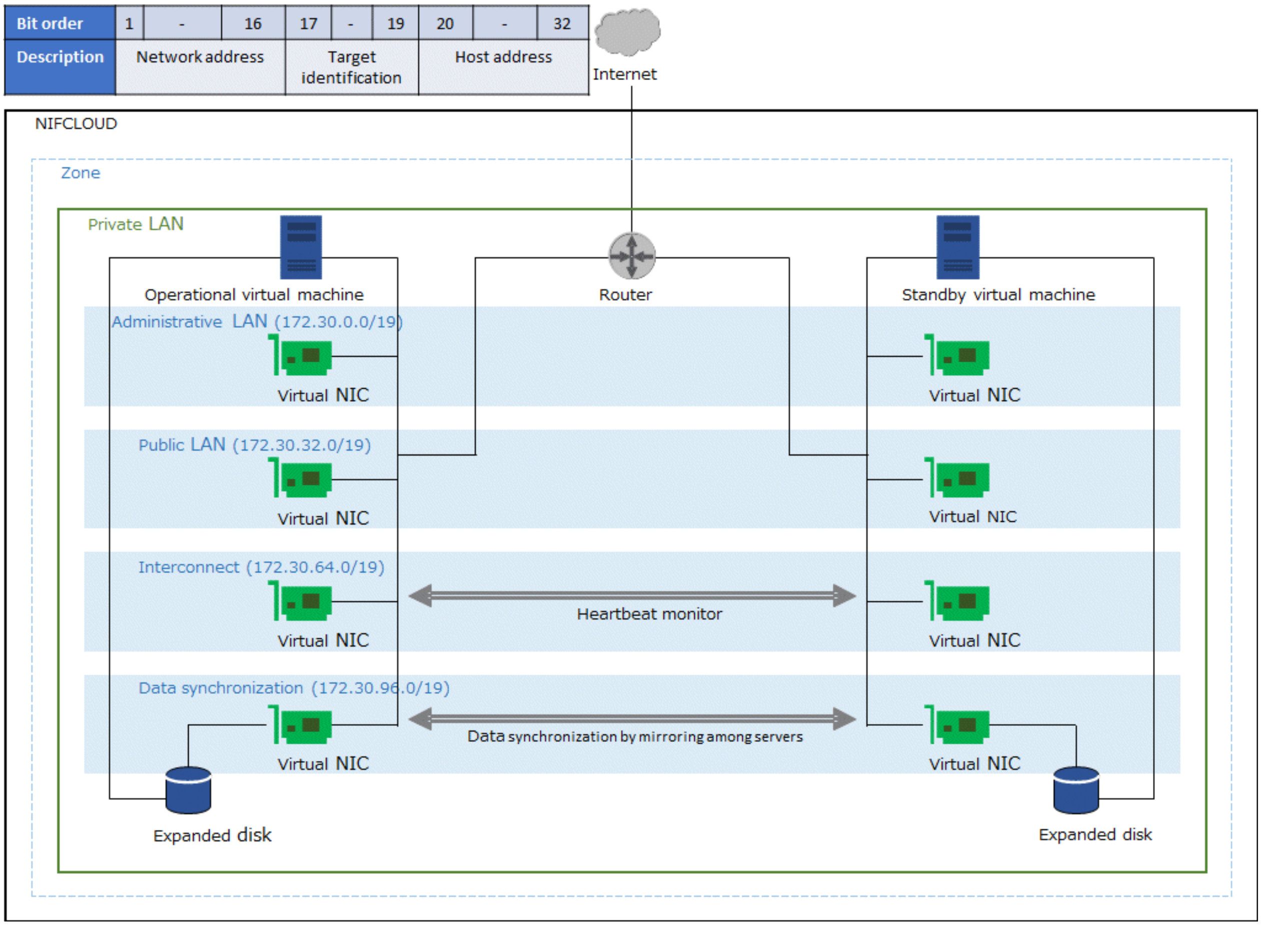


Bit order	1	-	16	17	-	19	20	-	32
Description	Network address			Target identification			Host address		

8.3.1.2 Subnet Design of the Cluster System in a Single Zone

For the cluster system in a single zone, prepare a private LAN and assign a subnet for each purpose.

Figure 8.8 Subnet design of the cluster system in a single zone



Bit order	1	-	16	17	-	19	20	-	32
Description	Network address			Target identification			Host address		

Internet
NIFCLOUD
Zone
Private LAN
Operational virtual machine
Router
Standby virtual machine
Administrative LAN (172.30.0.0/19)
Virtual NIC
Virtual NIC
Public LAN (172.30.32.0/19)
Virtual NIC
Virtual NIC
Interconnect (172.30.64.0/19)
Heartbeat monitor
Virtual NIC
Virtual NIC
Data synchronization (172.30.96.0/19)
Data synchronization by mirroring among servers
Virtual NIC
Virtual NIC
Expanded disk
Expanded disk

8.3.2 Firewall Design

This section describes the firewall group rule settings that are required to allow communication within the cluster.

PRIMECLUSTER uses several protocols/ports for communication within the cluster. When setting detailed rules, allow communication of protocols/ports for communication within the cluster.

In addition, you can add rules based on the security requirements of the customer to design firewall groups.

When adding rules according to requirements or adding rules required for the operation of other software, set the rules so that PRIMECLUSTER communication is not rejected.

The tables of the rules below describe the rules for the cluster system with a two-node configuration.

See

For details of the rules of the firewall groups, refer to the official NIFCLOUD documentation.

8.3.2.1 Rules Applied to the Administrative LAN

Design the rules of the firewall groups applied to the administrative LAN.

The IN rule setting is not necessary in a single-node cluster.

IN rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
UDP	9382	IP/CIDR	Administrative LAN NIC IP of remote cluster node	Used for the shutdown facility (SF)
UDP	9796	IP/CIDR	Administrative LAN NIC IP of remote cluster node	Used for the management view
TCP	9797	IP/CIDR	Administrative LAN NIC IP of remote cluster node	Used for the management view
ICMP	- (Specify all ports)	IP/CIDR	Administrative LAN NIC IP of remote cluster node	Used for clchkcluster

OUT rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
UDP (*1)	9382	IP/CIDR	Administrative LAN NIC IP of remote cluster node	Used for the shutdown facility (SF)
UDP (*1)	9796	IP/CIDR	Administrative LAN NIC IP of remote cluster node	Used for the management view
TCP (*1)	9797	IP/CIDR	Administrative LAN NIC IP of remote cluster node	Used for the management view
ICMP (*1)	- (Specify all ports)	IP/CIDR	Administrative LAN NIC IP of remote cluster node	Used for clchkcluster
TCP (*1)	53	IP/CIDR	CIDR of DNS	Used for the forced stop (Name resolution of the API endpoint)
UDP (*1)	53	IP/CIDR	CIDR of DNS	Used for the forced stop (Name resolution of the API endpoint)
TCP (*1)	443	IP/CIDR	0.0.0.0/0	Used for the forced stop (Communication with the API endpoint)
TCP	123	IP/CIDR	IP address of NTP server	Used for NTP server query
UDP	123	IP/CIDR	IP address of NTP server	Used for NTP server query

(*1) This setting is not necessary in a single-node cluster.

8.3.2.1.1 Rules Applied to Web-Based Admin View

Design the rules of the firewall groups applied to the Web-Based Admin View.

1) When ensuring the connectivity with a server for a client

Design the rules of the firewall groups applied to the Web-Based Admin View (cluster node side).

IN rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
TCP	8081	IP/CIDR	Server IP for the management view client	Used for the management view
TCP	9798	IP/CIDR	Server IP for the management view client	Used for the management view

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
TCP	9799	IP/CIDR	Server IP for the management view client	Used for the management view

Design the rules of the firewall groups applied to the Web-Based Admin View (management client side).

OUT rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
TCP	8081	IP/CIDR	Administrative LAN NIC IP of all cluster nodes	Used for the management view
TCP	9798	IP/CIDR	Administrative LAN NIC IP of all cluster nodes	Used for the management view
TCP	9799	IP/CIDR	Administrative LAN NIC IP of all cluster nodes	Used for the management view

Also, create an IN rule of the firewall group to allow a remote desktop connection from a remote control terminal of the management view client to a server for the management view client.

2) When ensuring the connectivity using a VPN connection

Design the rules of the firewall groups applied to the Web-Based Admin View (cluster node side).

IN rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
TCP	8081	IP/CIDR	CIDR of the management view client	Used for the management view
TCP	9798	IP/CIDR	CIDR of the management view client	Used for the management view
TCP	9799	IP/CIDR	CIDR of the management view client	Used for the management view

Design the rules of the firewall groups applied to the Web-Based Admin View (VPN gateway side).

OUT rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
TCP	8081	IP/CIDR	Administrative LAN NIC IP of all cluster nodes	Used for the management view
TCP	9798	IP/CIDR	Administrative LAN NIC IP of all cluster nodes	Used for the management view
TCP	9799	IP/CIDR	Administrative LAN NIC IP of all cluster nodes	Used for the management view

When using a VPN connection, refer to "8.3.2.5 Rules Applied to a VPN Gateway when Using VPN" and set the firewall groups applied to a VPN gateway as well.

8.3.2.1.2 Rules Applied to Server Access in Introduction and Maintenance

Design the rules of the firewall groups applied to server access in introduction and maintenance.

IN rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
TCP	22	IP/CIDR	CIDR of the access source	Used for the remote access by SSH

OUT rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
TCP	80	IP/CIDR	0.0.0.0/0	Used for installing dependent packages
TCP	443	IP/CIDR	0.0.0.0/0	Used for installing dependent packages

8.3.2.2 Rules Applied to the Cluster Interconnect

Design the rules of the firewall groups applied to the cluster interconnect.

This setting is not necessary in a single-node cluster.

IN rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
ANY	- (Specify all ports)	IP/CIDR	IP of NIC for cluster interconnect of remote cluster node	Used for the heartbeat

OUT rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
ANY	- (Specify all ports)	IP/CIDR	IP of NIC for cluster interconnect of remote cluster node	Used for the heartbeat

8.3.2.3 Rules Applied to the Public LAN

Design the rules of the firewall groups applied to the public LAN.

Add the rules that are required for application operations, and the following OUT rule.

OUT rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
ICMP	- (Specify all ports)	IP/CIDR	CIDR of the monitoring destination of the business network	Used for monitoring the business network

This rule is required when using the network monitoring function.

For details, refer to "6.7.3.6 Setting Up Takeover Network Resources" in "PRIMECLUSTER Installation and Administration Guide."

When using a router for an error monitoring of the public LAN, add the IN rule of the firewall group of the router to allow ICMP from the cluster.

8.3.2.4 Rules Applied to the Network for Data Synchronization

Design the rules of the firewall groups applied to the network for data synchronization.

This setting is not necessary in a single-node cluster.

IN rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
TCP	3260	IP/CIDR	IP of NIC for data synchronization of remote cluster node	Used for mirroring among servers

OUT rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
TCP	3260	IP/CIDR	IP of NIC for data synchronization of remote cluster node	Used for mirroring among servers

8.3.2.5 Rules Applied to a VPN Gateway when Using VPN

Design the rules of the firewall groups applied to a VPN gateway when using VPN.

IN rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
UDP	500	IP/CIDR	CIDR of the access source	Used for a VPN connection
UDP	4500	IP/CIDR	CIDR of the access source	Used for a VPN connection
ESP	- (Specify all ports)	IP/CIDR	CIDR of the access source	Used for a VPN connection

OUT rule

Protocol	Destination port	Connection source type	IP/CIDR, Group	Description
Protocol used for a communication via VPN	Port used for a communication via VPN	IP/CIDR	IP/CIDR of a cluster communicating via VPN	Used for a VPN connection

For the added OUT rule, add the IN rule of the firewall group of the server to allow the communication from the VPN connection source CIDR.

8.4 System Design

Use PRIMECLUSTER Designsheets to design the system.

The installation operation of the PRIMECLUSTER system is performed based on the created PRIMECLUSTER Designsheets. Make sure to create the designsheets and confirm that all required items are described.

8.5 Determining the Cluster System Operation Mode

In the cluster system in a NIFCLOUD environment, operation modes for 1:1 standby operation, mutual standby, and single-node cluster operation can be built.

For details on the operation mode of each cluster system, refer to "2.3 Determining the Cluster System Operation Mode" in "PRIMECLUSTER Installation and Administration Guide."

8.6 Determining the Web-Based Admin View Operation Mode

For details on the Web-Based Admin View operation mode, refer to "2.4 Determining the Web-Based Admin View Operation Mode" in "PRIMECLUSTER Installation and Administration Guide."

8.7 Determining the Failover Timing of Cluster Application

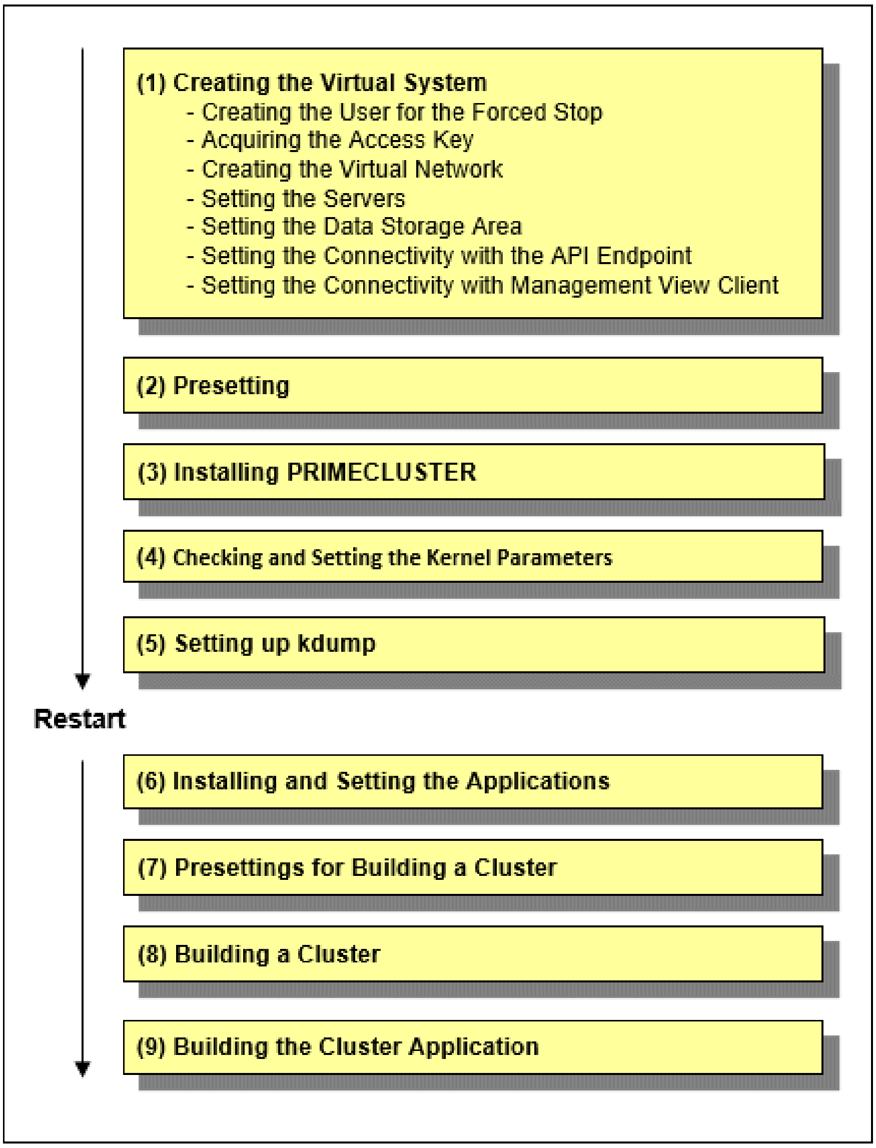
Determine the failover timing of cluster application.

For details, refer to "2.5 Determining the Failover Timing of Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

Chapter 9 Installation

This chapter describes the procedure to install PRIMECLUSTER in a NIFCLOUD environment.

Perform the steps shown in the figure below.


(1) Creating the Virtual System
- Creating the User for the Forced Stop
- Acquiring the Access Key
- Creating the Virtual Network
- Setting the Servers
- Setting the Data Storage Area
- Setting the Connectivity with the API Endpoint
- Setting the Connectivity with Management View Client
(2) Presetting
(3) Installing PRIMECLUSTER
(4) Checking and Setting the Kernel Parameters
(5) Setting up kdump
Restart
(6) Installing and Setting the Applications
(7) Presettings for Building a Cluster
(8) Building a Cluster
(9) Building the Cluster Application

Note

For how to check the operation environment, refer to "Chapter 2 Operation Environment" in the Installation Guide for PRIMECLUSTER.

9.1 Creating the Virtual System

This section describes how to create the virtual system for a cluster system in a NIFCLOUD environment.

By using the control panel, create the virtual system that was designed in "Chapter 8 Design." The creation procedure depends on the architectural pattern selected.

9.1.1 Creating the User for the Forced Stop

On the NIFCLOUD control panel, create the user to forcibly stop the server in the cluster system with the following values.

These settings are not necessary in a single-node cluster.

If the user with the following authorities already exists, you can use that user without creating a new user.

Item name	Value
Account name	Arbitrary account name
Authority	Administrator authority, operator authority, or viewer authority

The user for the forced stop must be allowed to stop the server.

When limiting the operations performed by the user for the forced stop, give only the viewer authority to the user, and add the policy that defines the following operation permission.

Service	Function	Operation
Computing	Server	Stop

See

For details on the NIFCLOUD control panel, refer to the official NIFCLOUD documentation.

9.1.2 Acquiring the Access Key

On the NIFCLOUD control panel, acquire the AccessKey and the SecretAccessKey for the user created in "9.1.1 Creating the User for the Forced Stop."

These keys are required by the shutdown facility for the authentication of the NIFCLOUD API.

When using the existing user without creating a new user for the forced stop, acquire the AccessKey and the SecretAccessKey for that user.

See

For details on the NIFCLOUD control panel, the AccessKey, and the SecretAccessKey, refer to the official NIFCLOUD documentation.

9.1.3 Creating the Virtual Network

According to the design in "8.3 Network Design", create the private LAN where the cluster system is to be deployed and the firewall to be applied.

9.1.4 Setting the Servers

Create the server that configures the cluster node and NIC.

When creating the server, do not use the DHCP function of the router, but set the IP address in the OS setting.

When creating NIC, use the additional NIC function of NIFCLOUD, and select the appropriate private LAN based on the network configurations designed in "8.3 Network Design."

When creating NIC and using NIC, for which a default gateway is not set, to communicate with NIC on different subnets, static routing must be set for that NIC. You do not need to set static routing when communicating only with NIC on the same subnet.

For how to configure static routing, refer to the example of static routing configuration in "21.1.3 Setting Instances."

9.1.5 Setting the Data Storage Area

When using the mirroring among servers, set the data storage area used by an application.

Create the virtual block device managed by the mirroring among servers, then attach it to the server.

If you are not using the mirroring among servers but configuring the data storage area by shared file system services or RDB services provided by NIFCLOUD, this setting is not necessary.

9.1.6 Setting the Connectivity with the API Endpoint

Create the components required for configuring the connectivity with the API endpoint. The required components depend on the architectural pattern.

When selecting a router to ensure connectivity

1. Create a router in a private LAN of the administrative LAN.
2. Create a NAT table and set it to the router created in step 1.

See

For details on the NAT function of the router, refer to the official NIFCLOUD documentation.

When selecting a NAT server to ensure connectivity

1. Build a server connected to the global network as a NAT server.
2. When connecting to the Internet from a cluster node, set the connection via the NAT server created in step 1.

When selecting a global IP address to ensure connectivity

Give a global IP address to a cluster node.

9.1.7 Setting the Connectivity with Management View Client

Set the components required for ensuring the connectivity between the management terminal and the management view client on the cluster node. The required components depend on the architectural pattern selected in "8.6 Determining the Web-Based Admin View Operation Mode."

When selecting a server for the client to ensure connectivity

On NIFCLOUD, deploy the server for the management view client connected with a private LAN of the administrative LAN.

When selecting a VPN connection to ensure connectivity

Create a VPN gateway, customer gateway, and VPN connection.

9.2 Presetting

Take the following procedure on all nodes.

1. Disable the firewall.

   Make sure that "firewalld" is disabled.

   ```
   # systemctl is-enabled firewalld
   ```

   If it is enabled, disable it.

   ```
   # systemctl stop firewalld
   # systemctl disable firewalld
   ```

2. Set NTP.

   Make sure to set NTP when building the cluster to synchronize the time of each node in the cluster system.

   Set NTP before installing PRIMECLUSTER.

9.3 Installing PRIMECLUSTER

Use the installation script (CLI Installer) to install PRIMECLUSTER.

Install PRIMECLUSTER on each node in the system where Linux(R) software and Linux(R) related software are already installed. Use the same installation script when installing PRIMECLUSTER on the cluster management server.

Note

If the OS has never been restarted since the server was created, restart it and then install PRIMECLUSTER.

See

For details on the installation and uninstallation procedures, refer to the descriptions of cloud environments described in the Installation Guide for PRIMECLUSTER.

9.4 Checking and Setting the Kernel Parameters

Change the kernel parameters depending on the environment.

Applicable nodes:

All nodes on which PRIMECLUSTER is to be installed

Depending on the products and components utilized, different kernel parameters are required.

Check PRIMECLUSTER Designsheets and if you need to modify the kernel parameters, set them again.

See

For details on kernel parameters, refer to "3.1.7 Checking and Setting the Kernel Parameters" in "PRIMECLUSTER Installation and Administration Guide."

Note

- To activate the modified kernel parameters, restart the OS.
- After uninstalling PRIMECLUSTER, change the kernel parameter settings back to the state before installing PRIMECLUSTER if necessary.

Set the following kernel parameters.

This setting is not necessary in a single-node cluster.

Parameter	Value	Remarks: meaning of (parameter)
kernel.panic	0	Seconds to wait until the kernel is restarted in case of a panic. If 0 is set, the kernel is not restarted.
kernel.sysrq	Other than 0	The SysRq key is enabled.

9.5 Setting up kdump

Set up the following on all nodes when using the kdump (recommended).

1. Set up the kdump parameter.

   In a two-node configuration, to power off the cluster node after the memory dump is output, set the following parameters in the /etc/kdump.conf.

   For details on the setup procedure of the kdump, refer to the OS manual.

Parameter	Value	Remarks: meaning of (parameter)
kdump_post	/opt/SMAW/SMAWsf/bin/poff.sh	The power-off script (poff.sh) is executed after the memory dump output ends.
default	poweroff	The power is turned off when the memory dump output fails.

2. Check the kdump.

   Check if the kdump server function is enabled. If not, enable the kdump.

   - Check the availability of the kdump with the systemctl(1) command.

     Example) The kdump is disabled if the state is as follows.

```
# /usr/bin/systemctl list-unit-files --type=service | grep  kdump.service
kdump.service                               disabled
```

   - If the kdump is enabled, restart the kdump with the systemctl(1) command.

```
# /usr/bin/systemctl restart kdump.service
```

   - If the kdump is disabled, enable it with the systemctl(1) command, and then start the kdump.

```
# /usr/bin/systemctl enable kdump.service
# /usr/bin/systemctl start kdump.service
```

Note

After uninstalling PRIMECLUSTER, make sure to change the kdump settings back to the state before installing PRIMECLUSTER.

9.6 Installing and Setting the Applications

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

See

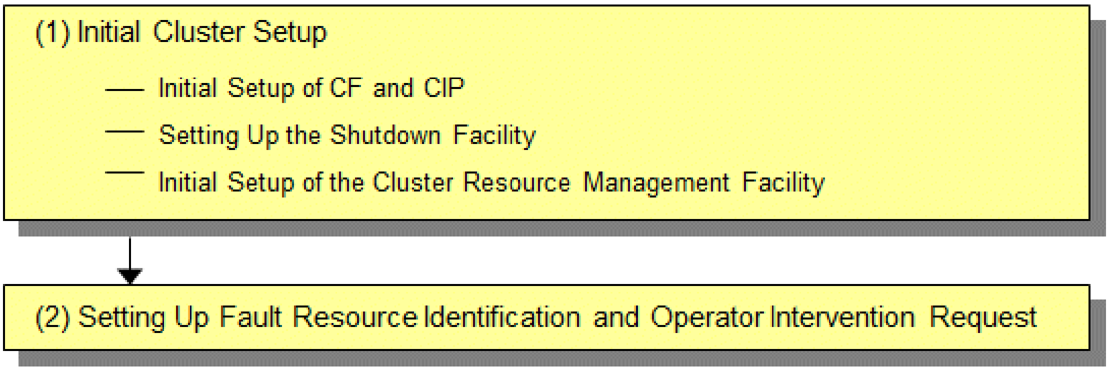
- For details on environment setup, refer to the manuals for each application.
- For information on PRIMECLUSTER-related products that support NIFCLOUD, refer to the documentation for each product.

9.7 Presettings for Building a Cluster

Prior to building a cluster, perform presettings such as staring the Web-Based Admin View screen. For details on presettings prior to building a cluster, refer to "Chapter 4 Preparation Prior to Building a Cluster" in "PRIMECLUSTER Installation and Administration Guide."

9.8 Building a Cluster

The procedure for building a PRIMECLUSTER cluster is shown below:


(1) Initial Cluster Setup
Initial Setup of 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

9.8.1 Initial Cluster Setup

This section describes the initial cluster setup for PRIMECLUSTER.

9.8.1.1 Initial Setup of CF and CIP

Refer to "5.1.1 Setting Up CF and CIP" in "PRIMECLUSTER Installation and Administration Guide" to set up CF and CIP.

Note

With the private bridge function of NIFCLOUD, a loss of communication may occur for about 15 seconds after an instantaneous loss of communication during maintenance and equipment failure accompanied by a loss of communication, and a heartbeat failure may occur.

To prevent a heartbeat failure during maintenance and equipment failure accompanied by a loss of communication, after setting CF, refer to "11.3.1 Changing Time to Detect CF Heartbeat Timeout" in "PRIMECLUSTER Installation and Administration Guide", and tune the cluster timeout value to 20 seconds or more.

You need to use CF over IP in a cloud environment. In addition, you need to configure CIP because such as RMS uses it for inter-node communication in the cluster. CF over IP (IP interconnect) and CIP are different features. For more information, see "1.1.1 Differences between CIP and CF over IP" in the Cluster Foundation Configuration and Administration Guide.

9.8.1.2 Setting up the Shutdown Facility

This section describes how to set up the shutdown facility in a NIFCLOUD environment.

The shutdown agent available in a NIFCLOUD environment is as follows.

- NIFCLOUD API (SA_vmnifclAsyncReset)

  The shutdown function of a node (server) using the NIFCLOUD API is provided.

  This setting is not necessary in a single-node cluster.

  The storage location of a log file is as follows.

```
/var/opt/SMAWsf/log/SA_vmnifclAsyncReset.log
```

For details on the survival priority, refer to "5.1.2.1 Survival Priority" in "PRIMECLUSTER Installation and Administration Guide."

Note

- After setting up the shutdown agent, conduct a test for the forced stop of cluster nodes to make sure that the correct nodes can be forcibly stopped. For details of the test for the forced stop of cluster nodes, refer to "1.4 Test" in "PRIMECLUSTER Installation and Administration Guide."
- The contents of the SA_vmnifclAsyncReset.cfg file and the rcsd.cfg file of all nodes should be identical. If not, a malfunction will occur.
- This setting is not necessary in a single-node cluster.

1. Set up the shutdown daemon.

   Create /etc/opt/SMAW/SMAWsf/rcsd.cfg with the following contents on all nodes in the cluster system.

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

```
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 used
                       in the shutdown facility of the cluster host.
                       Available IP addresses are IPv4.
                       When specifying a host name, make sure it is described in /etc/hosts.
SA_vmnifclAsyncReset: NIFCLOUD API shutdown agent.
timeout              : Specify the timeout duration (seconds) of the NIFCLOUD API shutdown agent.
                       Specify 30 seconds.
```

   Example) The following is a setup example.

   If the CF node names of the cluster host are node1 and node2, the weight of two nodes is 1, the IP address of the administrative LAN of node1 is 192.168.1.1, and the IP address of the administrative LAN of node2 is 192.168.1.2.

```
# cat /etc/opt/SMAW/SMAWsf/rcsd.cfg
node1,weight=1,admIP=192.168.1.1:agent=SA_vmnifclAsyncReset,timeout=30
node2,weight=1,admIP=192.168.1.2:agent=SA_vmnifclAsyncReset,timeout=30
```

   Create /etc/opt/SMAW/SMAWsf/rcsd.cfg and then set the owner, group, and access rights as follows.

```
# chown root:root /etc/opt/SMAW/SMAWsf/rcsd.cfg
# chmod 600 /etc/opt/SMAW/SMAWsf/rcsd.cfg
```

   ### Information

   When creating the /etc/opt/SMAW/SMAWsf/rcsd.cfg file, the /etc/opt/SMAW/SMAWsf/rcsd.cfg.template file can be used as a template.

2. Encrypt the SecretAccessKey.

   Execute the sfcipher command to encrypt a SecretAccessKey used for authenticating the NIFCLOUD API. For details on how to use the sfcipher command, refer to the manual page of "sfcipher."

```
# sfcipher -c
```

   Example) The following is a setup example.

   If the SecretAccessKey is "123456789ABCDEFGHIJKLMNOPQRSTUVWXYZ12345".

```
# sfcipher -c
Enter Password:      <- Enter 123456789ABCDEFGHIJKLMNOPQRSTUVWXYZ12345
Re-Enter Password:   <- Enter 123456789ABCDEFGHIJKLMNOPQRSTUVWXYZ12345
PIJgi9wm4sYJlfWShOxWP6fpz0Xb6t9KmRc4sPDsRBIGfm6gT5m0xH2H2xf1r38G
```

3. Set the shutdown agent.

   Create /etc/opt/SMAW/SMAWsf/SA_vmnifclAsyncReset.cfg with the following contents on all nodes in the cluster system.

   ### Information

   The template of the SA_vmnifclAsyncReset.cfg file can be found at the following location:

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

   Delimit each item with a single space.

```
CFNameX ServerName Region Accesskey SecretAccesskey
CFNameX ServerName Region Accesskey SecretAccesskey
```

```
CFNameX          : Specify the CF node name of the cluster host.
ServerName       : Specify the server name of NIFCLOUD on which the cluster host is operating.
Region           : Specify a region of the endpoint used in a NIFCLOUD environment.
                   Specify the "jp-east-X" format for the East Japan region
                   and the "jp-west-X" format for the West Japan region.
Accesskey        : Specify the AccessKey used for authenticating the NIFCLOUD API.
SecretAccesskey  : Specify the SecretAccessKey encrypted in step 2.
```

Example) The following is a setup example.

If the CF node names of the cluster host are node1 and node2, the server names are pclserver1 and pclserver2, and the region is east-1.

```
# cat /etc/opt/SMAW/SMAWsf/SA_vmnifclAsyncReset.cfg
node1 pclserver1 jp-east-1 123456789ABCDEFGHIJK
PIJgi9wm4sYJlfWShOxWP6fpz0Xb6t9KmRc4sPDsRBIGfm6gT5m0xH2H2xf1r38G
node2 pclserver2 jp-east-1 123456789ABCDEFGHIJK
PIJgi9wm4sYJlfWShOxWP6fpz0Xb6t9KmRc4sPDsRBIGfm6gT5m0xH2H2xf1r38G
```

Create /etc/opt/SMAW/SMAWsf/SA_vmnifclAsyncReset.cfg and then set the owner, group, and access rights as follows.

```
# chown root:root /etc/opt/SMAW/SMAWsf/SA_vmnifclAsyncReset.cfg
# chmod 600 /etc/opt/SMAW/SMAWsf/SA_vmnifclAsyncReset.cfg
```

Note

- Make sure that the /etc/opt/SMAW/SMAWsf/SA_vmnifclAsyncReset.cfg file is set correctly. If the setting is incorrect, the shutdown facility cannot be performed normally.

- Make sure that the server name (*ServerName*) and the region of the endpoint (*Region*) of NIFCLOUD corresponding to the CF node name (*CFNameX*) of the cluster host of the /etc/opt/SMAW/SMAWsf/SA_vmnifclAsyncReset.cfg file are set. If the setting is incorrect, an incorrect node will be forcibly stopped.

4. Start the shutdown facility.

   Check if the shutdown facility has been started on all nodes in the cluster system.

   ```
   # sdtool -s
   ```

   On a node where the shutdown facility has already been started, execute the following commands to restart the shutdown facility.

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

   On a node where the shutdown facility has not been started, execute the following command to start the shutdown facility.

   ```
   # sdtool -b
   ```

   Information

   You can check if the shutdown facility has already been started with the sdtool -s command. If "The RCSD is not running" is displayed, the shutdown facility is not started.

5. Check the status of the shutdown facility.

   Execute the following command on all nodes in the cluster system to check the status of the shutdown facility.

   ```
   # sdtool -s
   ```

Note

If "The RCSD is not running" is displayed, the setting of the shutdown daemon or the setting of the shutdown agent is not correct. Perform the procedure from step 1 to 4 again.

Information

Display results of the sdtool -s command

- If Unknown or Init-ing is displayed in Init State, wait for about one minute, and then check the status again.
- If Unknown is displayed in Shut State or Init State, it means that SF has not yet stopped the node, tested the route, or initialized the shutdown agent. Unknown is displayed temporarily in Test State or Init State until the actual state can be confirmed.
- If TestFailed is displayed in Test State, 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.

9.8.1.3 Initial Setup of the Cluster Resource Management Facility

Refer to "5.1.3 Initial Setup of the Cluster Resource Management Facility" in "PRIMECLUSTER Installation and Administration Guide" to set up the resource database managed by the cluster resource management facility. In this setting, set the iSCSI device used in the mirroring among servers of GDS and register it to the resource database.

9.8.2 Setting up the Fault Resource Identification and Operator Intervention Request

Refer to "5.2 Setting up Fault Resource Identification and Operator Intervention Request" in "PRIMECLUSTER Installation and Administration Guide" to set up the fault resource identification and the operator intervention request.

9.9 Building the Cluster Application

For details on how to build the cluster application, refer to "Chapter 6 Building Cluster Applications" in "PRIMECLUSTER Installation and Administration Guide."

When using the mirroring among servers, set the mirroring among servers of GDS (creating netmirror volume) while building the cluster application.

Note

- Note the following points when performing the settings in "Setting Tuning Parameters" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."
    - Among the tuning parameters to be set, set the following values for the following tuning parameters.

Tuning parameter name	Value after change
ED_CMD_RETRY_COUNT	100
ED_DRV_RETRY_COUNT	100

Example:

```
ED_CMD_RETRY_COUNT=100
ED_DRV_RETRY_COUNT=100
```

To extend the timeout period (CLUSTER_TIMEOUT) of the CF heartbeat, change the above parameter values according to the following formula. Round up the values after the decimal point.

    Calculation formula:

    ```
    <Increased CLUSTER_TIMEOUT> / 3 + 100
    ```

    Example: If the timeout period is extended from the default value of 10 seconds to 20 seconds.

    ```
    (20 - 10) / 3 + 100 = 104 seconds (Round up the value after the decimal point)
    ```

  - Set the following tuning parameter in /etc/opt/FJSVsdx/sdx.cf.

    ```
    SDX_NETMIRROR_IO_BLOCKADE=1
    ```

  - Comment out the following tuning parameter in /etc/opt/FJSVsdx/modules/sfdsk.conf and disable the slice preceding degenerated option.

    ```
    SDX_NETMIRROR_PRE_DETACH=1;
    ```

    Example:

    ```
    #SDX_NETMIRROR_PRE_DETACH=1;
    ```

- Note the following setting in "Checking and Setting Required Packages" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

  - Add "smawcf.service" to the value "After" in /etc/systemd/system/fjsvsdx.service.d/netmirror.conf.

    Example:

    ```
    After=target.service smawcf.service
    ```

- If the icmp communication between cluster nodes is not allowed in the security group configuration, the following message is displayed when the clchkcluster command is executed.

  ```
  Admin IP <IP address> used by SF is not alive.
  ```

  If this message is output, refer to "8.3.2.1 Rules Applied to the Administrative LAN", and set the icmp protocol rule to allow the icmp communication between cluster nodes. After that, execute the clchkcluster command again.

- When using the Symfoware Server (Native) Hot Standby feature, the standby and scalable userApplication configurations are required.

  Set the standby userApplication attribute to the following value.

Attribute	Setup value
StandbyTransitions	ClearFaultRequest | SwitchRequest

  Refer to "6) Scalable operation" in "6.7 Setting Up Cluster Applications" in the "PRIMECLUSTER Installation and Administration Guide" for detailed build instructions.

Chapter 10 Operations

For details on functions for managing PRIMECLUSTER system operations, refer to "Chapter 7 Operations" in "PRIMECLUSTER Installation and Administration Guide."

See

For details on how to operate GDS, refer to "Operation and Maintenance" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

Note

In a NIFCLOUD environment, a heartbeat may fail due to an error of the network node or an error of the storage controller, or scheduled maintenance for the infrastructure. This may switch the cluster application.

Chapter 11 Changing the Configurations

For details on changing the configuration information for the PRIMECLUSTER system, environment settings, the configuration of the cluster application, the operation attributes of the cluster system, refer to "Chapter 9 Changing the Cluster System Environment", "Chapter 10 Configuration Change of Cluster Applications", "Chapter 11 Changing the Operation Attributes of a Cluster System" in "PRIMECLUSTER Installation and Administration Guide." For details on changing the GDS configuration, refer to "Configuration Change" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

11.1 Changing the Configuration of the NIFCLOUD Environment

This section describes how to change the configuration of the NIFCLOUD environment.

11.1.1 Changing the Server Name of NIFCLOUD

This section describes how to change the server name of NIFCLOUD.

1. Execute the following command on all nodes to stop the shutdown facility.

```
# sdtool -e
```

2. On the NIFCLOUD control panel, change the name of the server created in "9.1.4 Setting the Servers."

See

For details on the NIFCLOUD control panel, refer to the official NIFCLOUD documentation.

3. Modify the configuration definition file of the NIFCLOUD API shutdown agent on all nodes.

   For the description of the configuration definition file, refer to step 3 to 5 in "9.8.1.2 Setting up the Shutdown Facility."

4. Execute the following command on all nodes to start the shutdown facility.

```
# sdtool -b
```

5. Execute the following command on all nodes and make sure that the shutdown facility operates normally.

```
# sdtool -s
```

11.1.2 Changing Credentials of the NIFCLOUD API

This section describes how to change the credentials of the NIFCLOUD API (AccessKey, SecretAccessKey).

1. Execute the following command on all nodes to stop the shutdown facility.

```
# sdtool -e
```

2. On the NIFCLOUD control panel, reissue the AccessKey or the SecretAccessKey.

See

For details on the NIFCLOUD control panel, refer to the official NIFCLOUD documentation.

3. Modify the configuration definition file of the NIFCLOUD API shutdown agent on all nodes.

   For the description of the configuration definition file, refer to step 2 to 5 in "9.8.1.2 Setting up the Shutdown Facility."

4. Execute the following command on all nodes to start the shutdown facility.

```
# sdtool -b
```

5. Execute the following command on all nodes and make sure that the shutdown facility operates normally.

```
# sdtool -s
```

Chapter 12 Maintenance

When you maintain the PRIMECLUSTER system in a NIFCLOUD environment, note the following points.

- For the procedure for applying/deleting urgent corrections in a NIFCLOUD environment, refer to "12.1 Software Maintenance."
- For details on other items and procedures required for maintenance of the PRIMECLUSTER system, refer to "Maintenance of the PRIMECLUSTER System" in "PRIMECLUSTER Installation and Administration Guide." For details on how to maintain GDS, refer to "Operation and Maintenance" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide." For details on how to maintain GLS, refer to "Maintenance" in "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function."

12.1 Software Maintenance

12.1.1 Notes on Applying Corrections to the PRIMECLUSTER System

For details on notes for applying an intensive correction to the cluster system, refer to "12.3.1 Notes on Applying Corrections to the PRIMECLUSTER System" in "PRIMECLUSTER Installation and Administration Guide."

Note

In a NIFCLOUD environment, refer to "9.5 Setting up kdump" to apply/delete the corrections in multi-user mode.

12.1.2 Overview of the Procedure for Applying/Deleting Corrections

Overview of the procedure is shown for applying each correction including an intensive correction to the cluster system in a NIFCLOUD environment. In an environment that does not use GDS, the procedure related to GDS is not necessary.

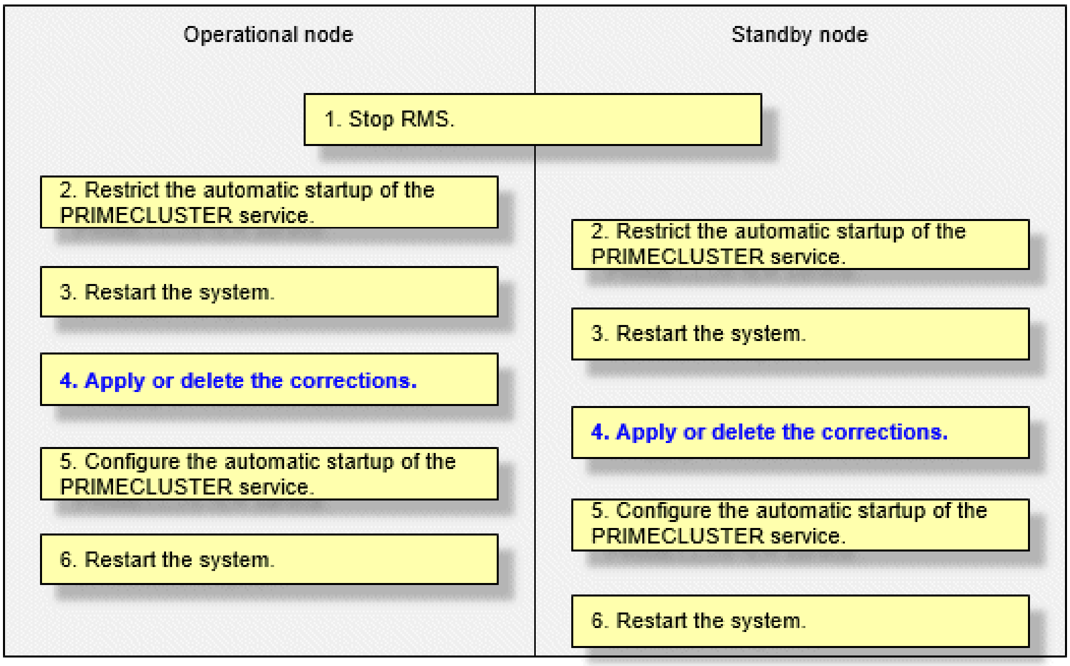
Note

Before applying/deleting corrections to PRIMECLUSTER, take a snapshot of the system storage.

12.1.2.1 Procedure for Applying or Deleting Corrections by Stopping the Entire System

This section describes the procedure for applying/deleting corrections by stopping the entire cluster system.

Flow of the operation


Operational node
Standby node
1. Stop RMS.
2. Restrict the automatic startup of the PRIMECLUSTER service.
3. Restart the system.
4. Apply or delete the corrections.
5. Configure the automatic startup of the PRIMECLUSTER service.
6. Restart the system.
2. Restrict the automatic startup of the PRIMECLUSTER service.
3. Restart the system.
4. Apply or delete the corrections.
5. Configure the automatic startup of the PRIMECLUSTER service.
6. Restart the system.

Operation procedure

Copy the corrections to be applied to each node to the local file system in advance.

1. Stop RMS.

   If RMS is running, execute the following command on any one node to stop RMS.

   ```
   # hvshut -a
   ```

   Note

   If RMS is stopped on all nodes during the synchronization copying of the GDS volume, the synchronization copying of the entire volume area is performed after the corrections are applied and all nodes are restarted.

   If you do not want to perform the synchronization copying of the entire area of volume, stop RMS after the synchronization copying is completed.

   To check the slice status of the GDS volume, execute the following command.

   Execute the following command on any one node to check the value of the STATUS field of the command output.

   The status of the copy destination slice is COPY during the synchronization copying, and after copying is complete, the status becomes ACTIVE or STOP.

   ```
   # sdxinfo -S
   ```

2. Restrict the automatic startup of the PRIMECLUSTER service.

   Restrict the automatic startup of the PRIMECLUSTER service by executing the following command on all nodes.

   ```
   # /opt/FJSVpclinst/bin/pclservice off
   ```

3. Restart the system.

   Restart the system on all nodes.

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

4. Apply or delete the corrections.

   Apply the corrections that were copied to the local file system, or delete the corrections.

- Applying corrections

  Copy the corrections to the working directory and then execute the following commands.

  ```
  # cd <working directory>
  # /opt/FJSVfupde/bin/uam add -d ./ -i <correction number>
  ```

  If the following message is displayed, select "Y".

  ```
  It is required to update with single user mode. Do you want to apply the update now? (Y/N)Y
  ```

  After that, the following message is displayed. Select "N".

  ```
  Do you want to restart your computer immediately? (Y/N)N
  ```

- Deleting corrections

  Execute the following command.

  ```
  # /opt/FJSVfupde/bin/uam remove -i <correction number>
  ```

  If the following message is displayed, select "Y".

  ```
  It is required to restore with single user mode. Do you want to restore the updated product
  to its pre-update state now? (Y/N)Y
  ```

  After that, the following message is displayed. Select "N".

  ```
  Do you want to restart your computer immediately? (Y/N)N
  ```

5. Configure the automatic startup of the PRIMECLUSTER service.

   Execute the following command on all nodes and change the PRIMECLUSTER service settings back to the state they were in before they were restricted in step 2.

   ```
   # /opt/FJSVpclinst/bin/pclservice on
   ```

6. Restart the system.

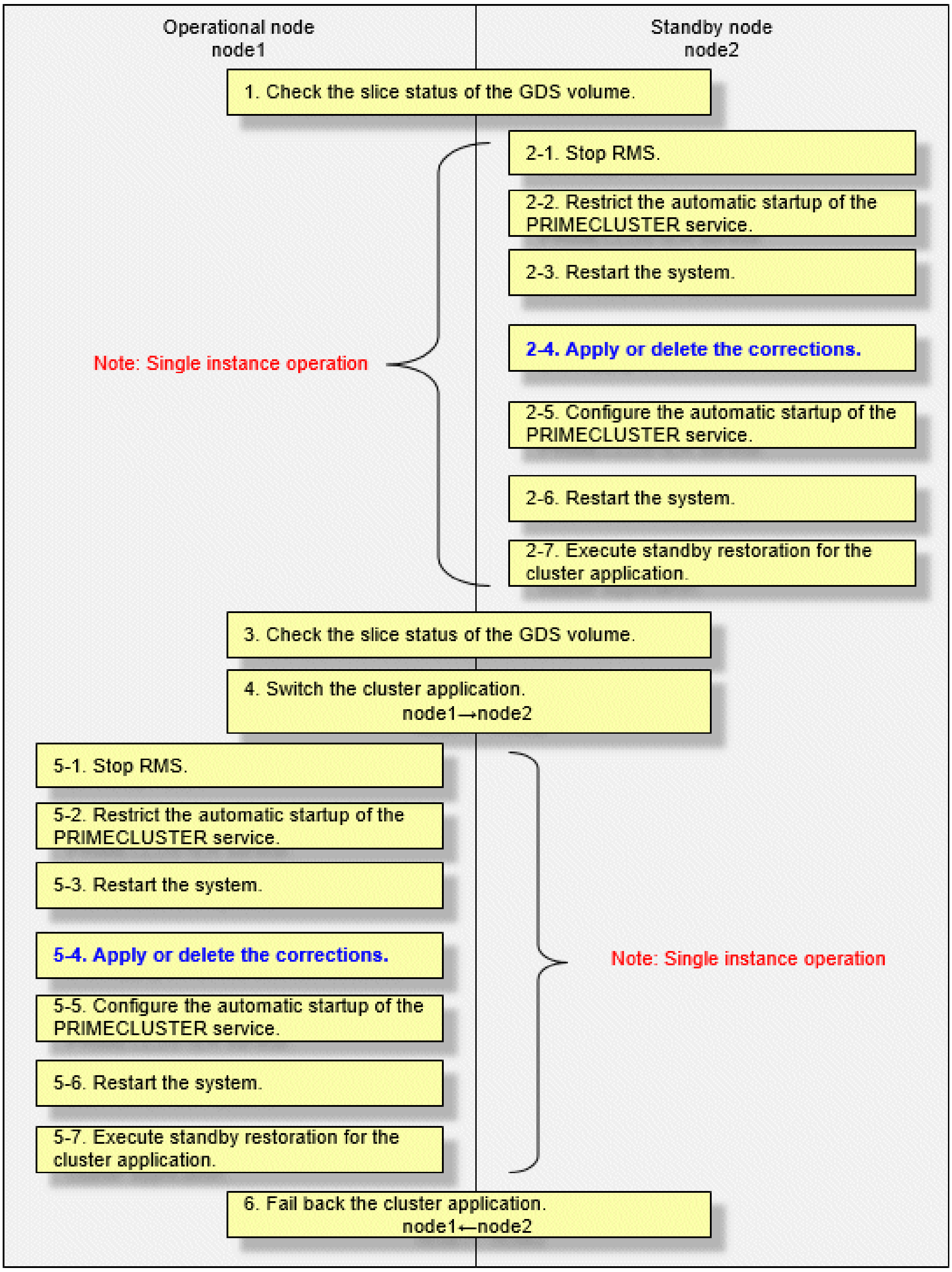
   Restart the system on all nodes.

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

12.1.2.2 Procedure for Applying or Deleting Corrections by Rolling Update

This section describes the procedure for applying corrections by rolling update.

Flow of the operation


Operational node
node1
Standby node
node2
1. Check the slice status of the GDS volume.
2-1. Stop RMS.
2-2. Restrict the automatic startup of the PRIMECLUSTER service.
2-3. Restart the system.
Note: Single instance operation
2-4. Apply or delete the corrections.
2-5. Configure the automatic startup of the PRIMECLUSTER service.
2-6. Restart the system.
2-7. Execute standby restoration for the cluster application.
3. Check the slice status of the GDS volume.
4. Switch the cluster application.
node1→node2
5-1. Stop RMS.
5-2. Restrict the automatic startup of the PRIMECLUSTER service.
5-3. Restart the system.
5-4. Apply or delete the corrections.
Note: Single instance operation
5-5. Configure the automatic startup of the PRIMECLUSTER service.
5-6. Restart the system.
5-7. Execute standby restoration for the cluster application.
6. Fail back the cluster application.
node1←node2

GDS: Global Disk Services

Operation procedure

1. Check the slice status of the GDS volume.

   Execute the following command on any cluster node to check the value of the STATUS field of the command output.

   ```
   # sdxinfo -S
   ```

   If the COPY status slice exists in the netmirror volume, wait until the synchronization copying is complete.

   For problems caused by node operations during copying, refer to "Stopping or Restarting the Node" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

2. Execute the following operation with the standby node (node2).

    1. Stop RMS.

        Stop RMS to apply corrections to the standby node (node2). A cutoff state transition occurs according to the shutdown of RMS. In this case, make sure that the single instance operation continues until the standby restoration for the cluster application is executed.

        ```
        # hvshut -l
        ```

    2. Restrict the automatic startup of the PRIMECLUSTER service.

        Execute the following command to restrict the automatic startup of the PRIMECLUSTER service.

        ```
        # /opt/FJSVpclinst/bin/pclservice off
        ```

    3. Restart the system.

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

    4. Apply or delete the corrections.

        - Applying corrections

        Copy the corrections to the working directory and then execute the following commands.

        ```
        # cd <working directory>
        # /opt/FJSVfupde/bin/uam add -d ./ -i <correction number>
        ```

        If the following message is displayed, select "Y".

        ```
        It is required to update with single user mode. Do you want to apply the update now? (Y/N)Y
        ```

        After that, the following message is displayed. Select "N".

        ```
        Do you want to restart your computer immediately? (Y/N)N
        ```

        - Deleting corrections

        Execute the following command.

        ```
        # /opt/FJSVfupde/bin/uam remove -i <correction number>
        ```

        If the following message is displayed, select "Y".

        ```
        It is required to restore with single user mode. Do you want to restore the updated
        product to its pre-update state now? (Y/N)Y
        ```

        After that, the following message is displayed. Select "N".

        ```
        Do you want to restart your computer immediately? (Y/N)N
        ```

    5. Configure the automatic startup of the PRIMECLUSTER service.

        Execute the following command and change the PRIMECLUSTER service settings back to the state they were in before they were restricted in 2 of step 2.

        ```
        # /opt/FJSVpclinst/bin/pclservice on
        ```

    6. Restart the system.

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

    7. Execute standby restoration for the cluster application.

        If the node (node2) to which corrections have been applied is cut off from the cluster system, execute standby restoration for the node.

    For details on how to execute cluster application standby restoration, refer to "7.2.2.1 Starting a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

3. Check the slice status of the GDS volume.

   After starting the standby node (node2), the synchronization copying of the netmirror volume is executed. Make sure that the synchronization copying is completely finished and all slices are either in ACTIVE or STOP status on any one node.

   To check the slice status of the netmirror volume, execute the following command:

   Execute the following command on any cluster node to check the value of the STATUS field of the command output.

   ```
   # sdxinfo -S
   ```

4. Switch the cluster application.

   To apply corrections to the operational node (node1), execute hvswitch and switch all cluster applications to the standby node (node2). For details on how to switch the cluster applications, refer to "7.2.2.3 Switching a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

5. Perform the following operation with the old operational node (node1).

   1. Stop RMS.

      Stop RMS to apply corrections to the operational node (node1). A cutoff state transition occurs according to the shutdown of RMS. In this case, make sure that the single instance operation continues until the standby restoration for the cluster application is executed.

      ```
      # hvshut -l
      ```

   2. Restrict the automatic startup of the PRIMECLUSTER service.

      Execute the following command to restrict the automatic startup of the PRIMECLUSTER service.

      ```
      # /opt/FJSVpclinst/bin/pclservice off
      ```

   3. Restart the system.

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

   4. Apply or delete the corrections.

      - Applying corrections

      Copy the corrections to the working directory and then execute the following commands.

      ```
      # cd <working directory>
      # /opt/FJSVfupde/bin/uam add -d ./ -i <correction number>
      ```

      If the following message is displayed, select "Y".

      ```
      It is required to update with single user mode. Do you want to apply the update now? (Y/N)Y
      ```

      After that, the following message is displayed. Select "N".

      ```
      Do you want to restart your computer immediately? (Y/N)N
      ```

      - Deleting corrections

      Execute the following command.

      ```
      # /opt/FJSVfupde/bin/uam remove -i <correction number>
      ```

      If the following message is displayed, select "Y".

      ```
      It is required to restore with single user mode. Do you want to restore the updated
      product to its pre-update state now? (Y/N)Y
      ```

      After that, the following message is displayed. Select "N".

```
Do you want to restart your computer immediately? (Y/N)N
```

   5. Configure the automatic startup of the PRIMECLUSTER service.

      Execute the following command and change the PRIMECLUSTER service settings back to the state they were in before they were restricted in 2 of step 5.

      ```
      # /opt/FJSVpclinst/bin/pclservice on
      ```

   6. Restart the system.

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

   7. Execute standby restoration for the cluster application.

      If the node (node1) to which corrections have been applied is cut off from the cluster system, execute standby restoration for the node. For details on how to execute cluster application standby restoration, refer to "7.2.2.1 Starting a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

6. Fail back the cluster application.

   Restore the state of the standby layout defined at installation by executing failback operation, as necessary. For details on failback, refer to "7.2.2.3 Switching a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

12.2 Procedure for Restoring with the Backup Related Services

When restoring servers with the backup related services (the One Day Snap Shot function, the Backup function, and the Customized image function), perform the following procedure.

Note

This is the procedure to restore the entire server including the expanded disks.

For backup and restore, configure the server with the same IP, firewall, and other settings as the server before the restore.

When using the One Day Snap Shot function, stop the server and take a snapshot.

For how to use the backup related services and notes, refer to the official NIFCLOUD documentation.

If the Backup function is used, the order of the NICs may change. In this case, refer to the official NIFCLOUD documentation.

If the additional NIC function is used, you cannot image the server with the customized image function. Use the One Day Snap Shot function or the Backup function.

12.2.1 Procedure for Restoring One Node

1. Restore the server that was backed up with the backup related services. For how to restore the server, refer to the official NIFCLOUD documentation.

   Note

   Do not start the node.

2. When not using the mirroring among servers of GDS, start the node restored in step 1.

   When using the mirroring among servers of GDS, perform the following procedure.

   1. Start the node restored in step 1 in single-user mode, and execute the following command from the NIFCLOUD console.

      ```
      # systemctl disable target.service
      ```

      After the setting, stop the node restored in step 1, and start the node in multi-user mode.

   2. After starting the node restored in step 1, execute the following command to detach the disk.

      Perform this step on the node that is not restored in step 1.

      ```
      # sdxswap -O -c <class name> -d <name of the disk on the node restored in step 1>
      ```

      If multiple netmirror groups exist, repeat this step for the number of groups.

   3. Execute the following commands on the node restored in step 1 to restart the node.

      ```
      # systemctl enable target.service
      # shutdown -r now
      ```

   4. After starting the node restored in step 1, execute the following command to restore the disk.

      Perform this step on the node that is not restored in step 1.

      If multiple netmirror groups exist, repeat this step for the number of groups.

      - When waiting for the synchronization copying

      Execute the following command.

      ```
      # sdxswap -I -c <class name> -d <name of the disk on the node restored in step 1>
      ```

      - When not waiting for the synchronization copying

      Execute the following command. After executing this command, use GUI or the sdxinfo command to check if the synchronization copying is completed.

      ```
      # sdxswap -I -c <class name> -d <name of the disk on the node restored in step 1> -e
      nowaitsync
      ```

3. If the server is restored with the backup related services other than the One Day Snap Shot function in step 1, perform step 3 to 5 in "9.8.1.2 Setting up the Shutdown Facility" on all nodes.

12.2.2 Procedure for Restoring Both Nodes

This section describes the procedure to restore both nodes.

The procedure depends on a period when the backup is acquired in each node in the cluster system.

If a period when the backup is acquired by using the backup related services depends on a node

1. Stop all nodes in the cluster system.

2. Restore one node and start it.

   For the node restoration, refer to the official NIFCLOUD documentation.

3. Stop RMS.

   ```
   # hvshut -l
   ```

4. Restore and start the other node.

5. When using the mirroring among servers of GDS, perform the following procedure.

   1. Check the status of the source slice for the synchronization copying.

      If the source slice for the synchronization copying is INVALID, restore the status of the slice.

      For the -d option of the sdxfix command, specify the source disk of the synchronization copying.

      Perform this procedure on either node.

      ```
      # sdxfix -V -c <class name> -v <volume name> -d <disk name> -x NoRdchk
      ```

      If multiple volumes exist, repeat this step for the number of volumes.

Note

If multiple volumes belong to the same netmirror group, specify the same <*disk name*> for the command.

2. Restore the disk.

   Perform this step on any one node.

   If multiple netmirror groups exist, execute the following commands for the number of groups.

   ```
   # sdxswap -O -c <class name> -d <copy destination disk name>
   # sdxswap -I -c <class name> -d <copy destination disk name> -e nowaitsync
   ```

6. Start RMS on the node restored in step 2.

   ```
   # hvcm -a
   ```

   When using the mirroring among servers of GDS, the entire synchronization copying is also performed.

   Use GUI or the sdxinfo command to check if the synchronization copying is completed.

7. If the server is restored with the backup related services other than the One Day Snap Shot function, perform step 3 to 5 in "9.8.1.2 Setting up the Shutdown Facility" on all nodes.

If the backup is acquired on both nodes at the same time by using the backup related services with the OS stopped on both nodes

1. Stop all nodes in the cluster system.

2. Restore both nodes and start them.

   For the node restoration, refer to the official NIFCLOUD documentation.

3. If the server is restored with the backup related services other than the One Day Snap Shot function, perform step 3 to 5 in "9.8.1.2 Setting up the Shutdown Facility" on all nodes.

Part 3 FJcloud-Baremetal Environment

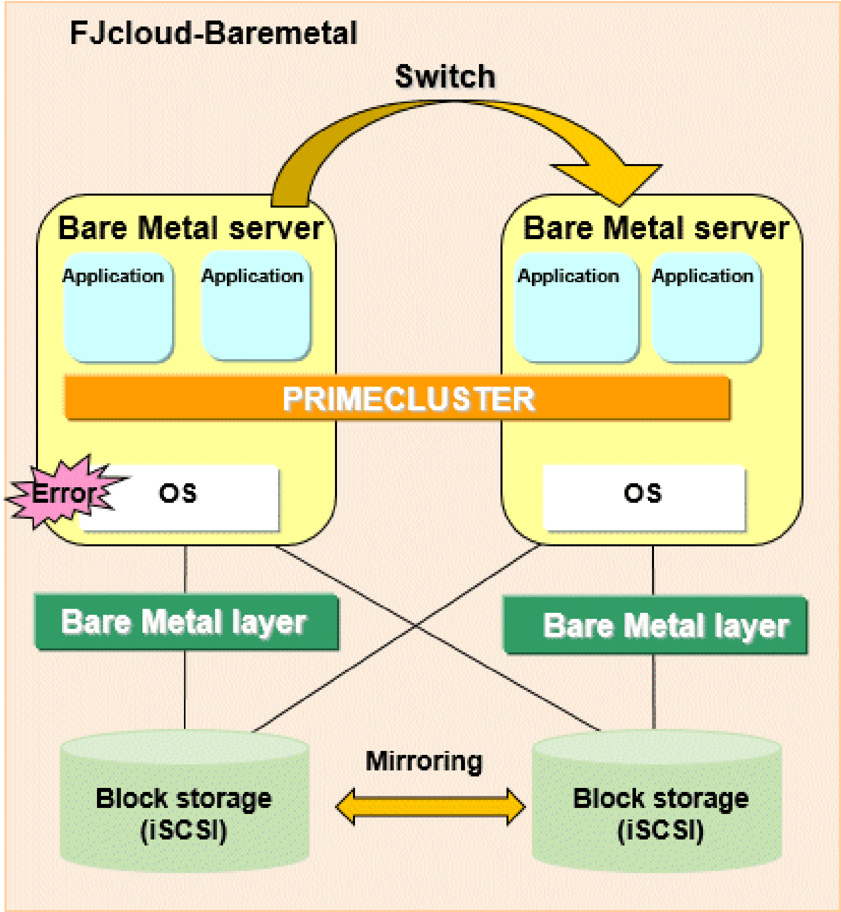
This part describes the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system in an FJcloud-Baremetal environment.

Chapter 13 Cluster System in an FJcloud-Baremetal Environment

In an FJcloud-Baremetal environment, PRIMECLUSTER can be used on the physical server built in the same segment. Even in a cloud, the business availability equal to the on-premise cluster system can be realized.

See

For details on FJcloud-Baremetal, refer to the official FJcloud-Baremetal documentation.



13.1 Supported Range

This section explains the range of support of PRIMECLUSTER in an FJcloud-Baremetal environment.

Supported configurations

- Number of cluster nodes: 1 to 2 nodes
- Operation mode of the cluster system: 1:1 Standby operation, Mutual standby, Single-node cluster
- Storage configurations: Block storage (iSCSI)
- Network configurations
  - For the cluster interconnect, its network must be independent from the network used with the administrative LAN and the public LAN. This setting is not necessary in a single-node cluster.

        - The servers in the cluster system must communicate with the API endpoints. This setting is not necessary in a single-node cluster.

    - Security groups

        - One security group must be set among the servers in the cluster system for a security reason.

        - Another security group must be set between the servers and the management client in the cluster system.

Supported monitoring functions

- Error of OS on the server and the cluster interconnect

  The cyclic monitoring of the cluster interconnect (LAN) detects a hang-up of OS and the service is switched to the standby system.

- Error of the shared disk and the disk access path

  By combining the volume management function (GDS), a failure of the disk access and the disk access path can be detected (monitored by the Gds resource), and the service is switched to the standby system when the disk access is disabled or a failure of the whole system of the disk access path occurs.

- Error of the cluster application

  When a resource error of the cluster application occurs, the service is switched to the standby system.

Note

- To take over the IP address of the server, setting the virtual NIC mode for GLS is required.

- To take over the data of the Bare Metal server, setting the shared disk mirroring of GDS is required.

- To take over the IP address of the Bare Metal server, setting the virtual NIC mode for GLS is required.

- The following functions for PRIMECLUSTER are not available:

    - GDS Snapshot

    - Root class of GDS, and disk groups of GDS other than the mirror type (netmirror, stripe, concatenation, and switch)

    - In addition to the above functions, the following functions for PRIMECLUSTER are not available in a single-node cluster.

        - GDS

        - GFS

- Duplicate virtual server names cannot be used in the project on FJcloud-Baremetal.

- For Bare Metal server names that use PRIMECLUSTER, the following ASCII characters can be used.
  Do not use other characters.

    - Uppercase letters

    - Lowercase letters

    - Numbers

    - "_" (Underscore)

    - "-" (Hyphen)

- In an FJcloud-Baremetal environment, when an OS panic occurs, the cluster node may be powered off while the memory dump is being output, and it may not be possible to collect a complete memory dump.

Chapter 14 Design

You must prepare the items listed below before building the PRIMECLUSTER system in an FJcloud-Baremetal environment.

- Selecting the PRIMECLUSTER Product
- Selecting the Architectural Pattern
- Network Design
- System Design
- Determining the Cluster System Operation Mode
- Determining the Web-Based Admin View Operation Mode
- Determining the Failover Timing of Cluster Application

Point

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

Information

For the flow to build the PRIMECLUSTER system, refer to "Chapter 1 Build Flow" in "PRIMECLUSTER Installation and Administration Guide."

14.1 Selecting the PRIMECLUSTER Product

Select a PRIMECLUSTER product.

In an FJcloud-Baremetal environment, you can select the following products.

For details on the PRIMECLUSTER products, refer to "2.1 PRIMECLUSTER Product Selection" in "PRIMECLUSTER Installation and Administration Guide."

- PRIMECLUSTER Enterprise Edition (EE)
- PRIMECLUSTER HA Server (HA)
- PRIMECLUSTER Clustering Base (CB)

14.2 Selecting the Architectural Pattern

In the PRIMECLUSTER system in an FJcloud-Baremetal environment, select an architecture pattern for the item below.

- Ensuring Connectivity with Web-Based Admin View

14.2.1 Ensuring Connectivity with Web-Based Admin View

PRIMECLUSTER provides architectural patterns for ensuring the connectivity between a terminal directly operated by a user and the management view.

For smooth designing of the cluster system, apply the following architectural pattern and design the cluster system.

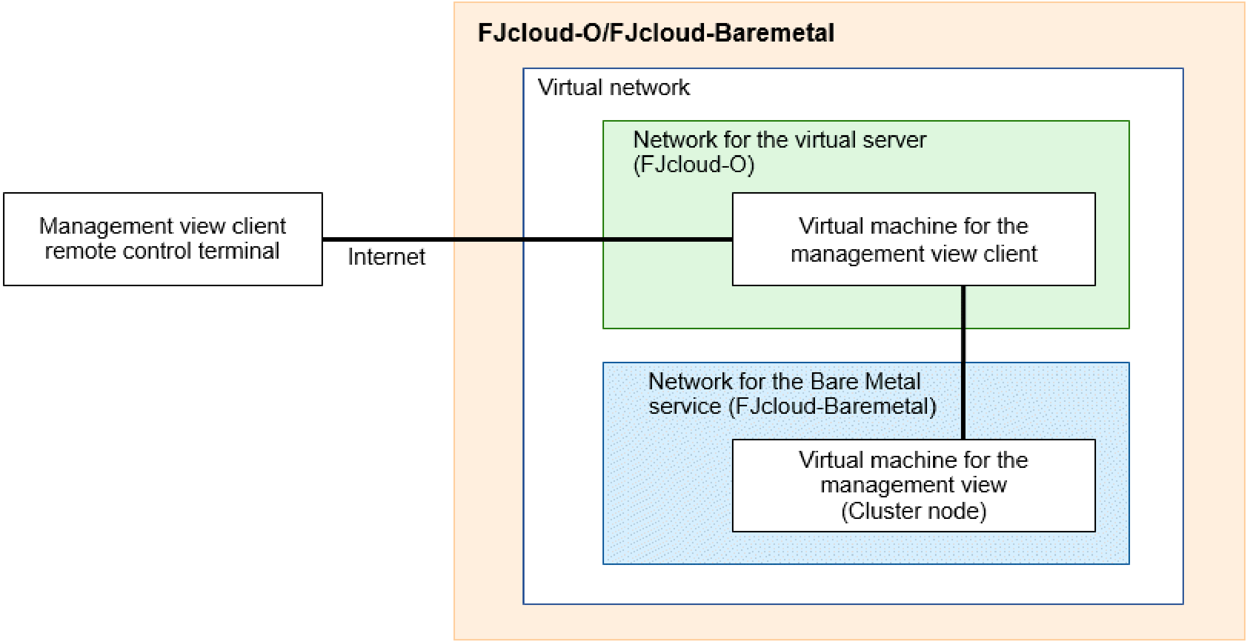
In the pattern using a virtual machine for a client, to ensure the connectivity between a terminal directly operated by a user and the management view, prepare a virtual machine for the management view client deployed in the network accessible from the Internet.

In this architectural pattern, a user connects to a client of the management view via a remote desktop connection.

The firewall service must be set in the virtual router in the network where the terminal directly operated by a user is deployed.

In addition, VLAN cannot be used in the network used by the Web-Based Admin View.

Figure 14.1 Ensuring connectivity with Web-Based Admin View


FJcloud-O/FJcloud-Baremetal
Virtual network
Network for the virtual server (FJcloud-O)
Virtual machine for the management view client
Management view client remote control terminal
Internet
Network for the Bare Metal service (FJcloud-Baremetal)
Virtual machine for the management view (Cluster node)

14.3 Network Design

In the cluster system in an FJcloud-Baremetal environment, securities must be designed in advance.

14.3.1 Subnet Design

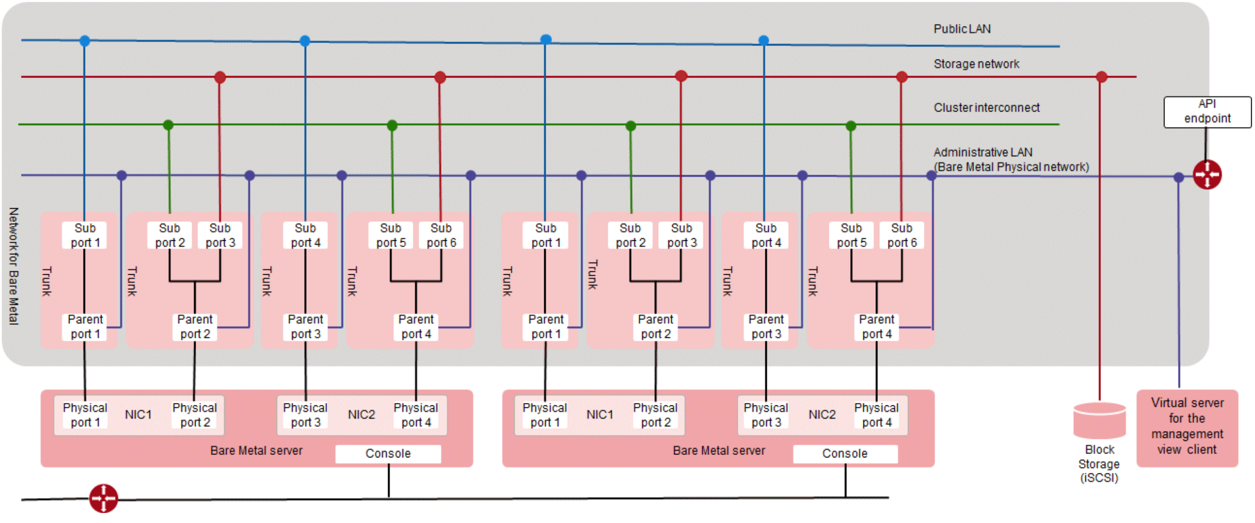
Create a subnet for each purpose.

To ensure the redundancy of NIC, build the subnet as follows.

Bundle the parent port 1 and the parent port 3 with GLS, and use it as the administrative LAN. Create the tagged VLAN interface for the bundled virtual NIC, and use it as the public LAN.

Also bundle the parent port 2 and the parent port 4 with GLS. Create two tagged VLAN interfaces for the bundled virtual NIC, and use each interface as the storage network and the cluster interconnect.

Figure 14.2 Subnet design of the cluster system in FJcloud-Baremetal



Note

- Due to a restriction in FJcloud-Baremetal, when connecting a virtual server to a network for Bare Metal services, two or more virtual servers must be connected on the same subnet.
- The parent port 2 and the parent port 4 are not used as the administrative LAN of PRIMECLUSTER.
- Storage network and Cluster interconnect are closed networks within the cluster and cannot be used to communicate with other environments.

14.3.2 Security Design

This section describes the security settings that are required to allow communication within the cluster.

PRIMECLUSTER uses several protocols/ports for communication within the cluster. Allow communication of protocols/ports for communication within the cluster as follows.

However, in FJcloud-Baremetal, if you allow the protocol/port used by PRIMECLUSTER, the maximum number of security rules will be exceeded. In this case, use the firewall on the OS to restrict communication.

14.3.2.1 Firewalls Applied to the Cluster Node

Refer to "Appendix J Using Firewall" in "PRIMECLUSTER Installation and Administration Guide", and set the firewall for each NIC.

In a Bare Metal environment, additional ports/protocols must be allowed.

In addition, add rules based on the security requirements of the customer.

- Firewall applied to the administrative LAN

  Allow sending to the following port numbers.

  These settings are not necessary in a single-node cluster.

Communication destination	Protocol	Port range	Description
IP address of DNS server	tcp	53	Used for the forced stop
IP address of DNS server	udp	53	Used for the forced stop
IP address of NTP server	udp	123	Used for NTP server query
Not specified	tcp	443	Used for the forced stop
IP address of the administrative LAN of a destination cluster node	icmp	-	Used for clchkcluster

  Allow receiving from the following port number.

  This setting is not necessary in a single-node cluster.

Communication destination	Protocol	Port range	Description
IP address of the administrative LAN of a destination cluster node	icmp	-	Used for clchkcluster

- Firewall applied to the cluster interconnect

  Allow sending and receiving for the following protocol.

  This setting is not necessary in a single-node cluster.

Communication destination	Protocol	Port range	Description
IP address of the interconnect of a destination cluster node	123	-	Used for the heartbeat

- Firewall applied to the storage network

  Allow sending to the following port number.

Communication destination	Protocol	Port range	Description
IP address of the block storage (iSCSI)	tcp	3260	Used for the connection with the block storage (iSCSI)

14.3.2.2 Security Group of the Management View Client

Create the security group for the Web-Based Admin View (on the management client side) with the following values.

Inbound rule

Communication source CIDR	Protocol	Port range	Description
Specified timely	tcp	3389	Used for the remote desktop connection

Outbound rule

Communication target CIDR	Protocol	Port range	Description
IP address of the cluster node	tcp	8081	Used for the management view
IP address of the cluster node	tcp	9798	Used for the management view
IP address of the cluster node	tcp	9799	Used for the management view

14.3.2.3 Firewall Rule

When using the firewall service, add the following to the firewall rule.

Protocol	Source IP address	Destination IP address	Destination port number	Action
tcp (*1)	Subnet for the administrative LAN	Not specified	443	Allow
udp	Subnet for the administrative LAN	IP address of DNS server	53	Allow
tcp	Subnet for the administrative LAN	IP address of DNS server	53	Allow
udp	Subnet for the administrative LAN	IP address of NTP server	123	Allow

(*1) This setting is not necessary in a single-node cluster.

Note

- Add the settings to allow the connection via ssh or the remote desktop connection from the external network as necessary.
- When using the yum command, add the following settings. Add or delete these settings as necessary to enhance the security.

Protocol	Source IP address	Destination IP address	Destination port number	Action
tcp	Subnet for the administrative LAN	IP address to the repository server	80	Allow
tcp	Subnet for the administrative LAN	IP address to the repository server	443	Allow

14.4 System Design

Use PRIMECLUSTER Designsheets to design the system.

The installation operation of the PRIMECLUSTER system is performed based on the created PRIMECLUSTER Designsheets. Make sure to create the designsheets and confirm that all required items are described.

14.5 Determining the Cluster System Operation Mode

In the cluster system in an FJcloud-Baremetal environment, the same operation mode as on-premise (physical environment) can be built.

For details on the operation mode of each cluster system, refer to "2.3 Determining the Cluster System Operation Mode" in "PRIMECLUSTER Installation and Administration Guide."

14.6 Determining the Web-Based Admin View Operation Mode

For details on the Web-Based Admin View operation mode, refer to "2.4 Determining the Web-Based Admin View Operation Mode" in "PRIMECLUSTER Installation and Administration Guide."

14.7 Determining the Failover Timing of Cluster Application

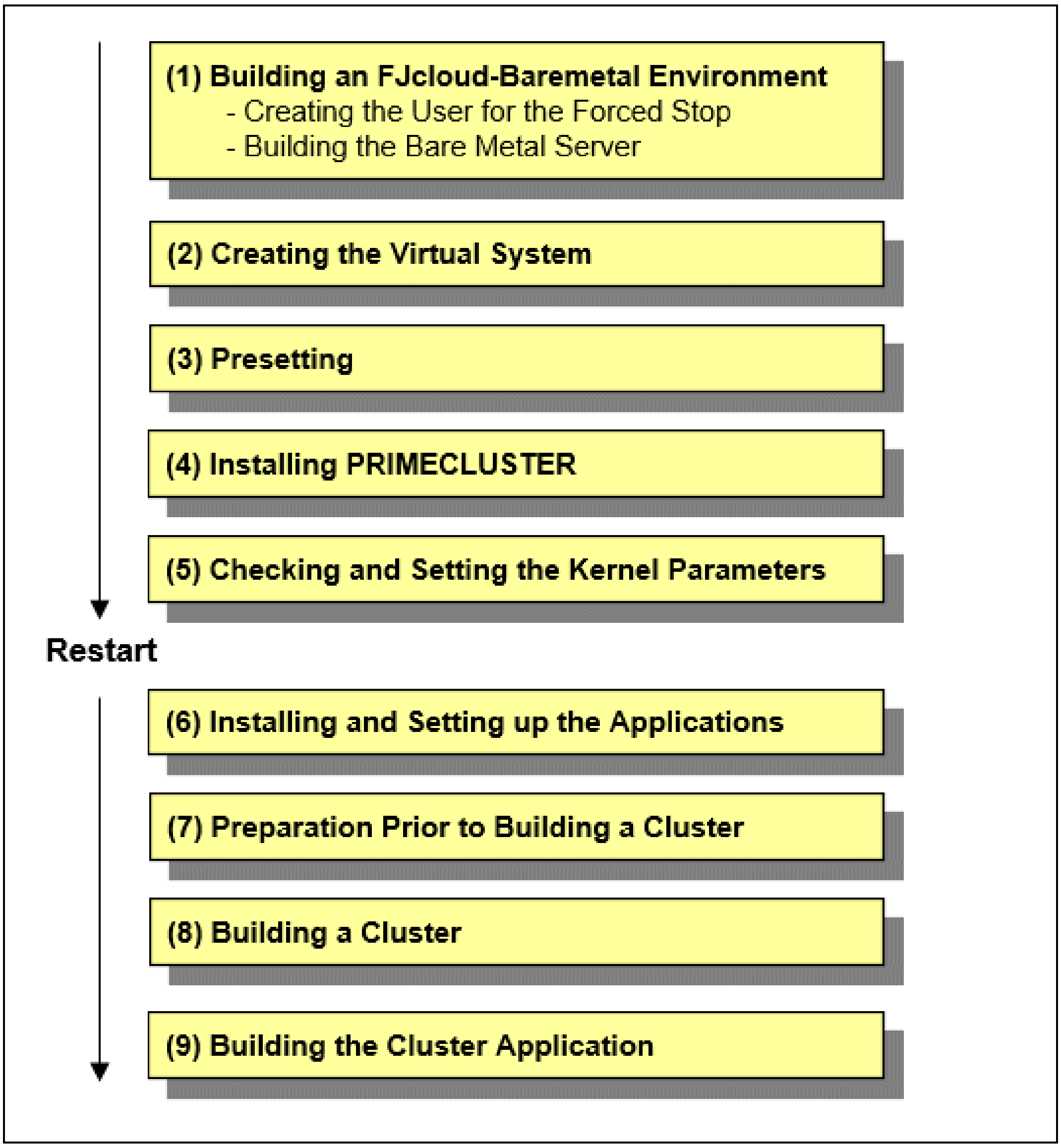
Determine the failover timing of cluster application.

For details, refer to "2.5 Determining the Failover Timing of Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

Chapter 15 Installation

This chapter describes the procedure to install PRIMECLUSTER in an FJcloud-Baremetal environment.

Perform the steps shown in the figure below.


(1) Building an FJcloud-Baremetal Environment
- Creating the User for the Forced Stop
- Building the Bare Metal Server
(2) Creating the Virtual System
(3) Presetting
(4) Installing PRIMECLUSTER
(5) Checking and Setting the Kernel Parameters
Restart
(6) Installing and Setting up the Applications
(7) Preparation Prior to Building a Cluster
(8) Building a Cluster
(9) Building the Cluster Application

Note

For how to check the operation environment, refer to "Chapter 2 Operation Environment" in the Installation Guide for PRIMECLUSTER.

15.1 Building an FJcloud-Baremetal Environment

This section describes how to build an FJcloud-Baremetal environment.

See

For how to set FJcloud-Baremetal, refer to the official FJcloud-Baremetal documentation.

Note

- To use the service provided by FJcloud-O IaaS/FJcloud-Baremetal with API, it is necessary to build an environment for using API. For details, refer to the documents provided by FJcloud-O.
- If the execution examples in this chapter are executed as they are described, an execution error may occur due to incompatibilities of API depending on a region of FJcloud-O.
  For details, refer to the documents provided by FJcloud-O/FJcloud-Baremetal.

15.1.1 Creating the User for the Forced Stop

On the FJcloud portal, create the user to forcibly stop the Bare Metal server in the cluster system with the following values.

These settings are not necessary in a single-node cluster.

If the user is already created in an FJcloud-Baremetal environment, you can use that user without creating a new user.

Item name	Value
User name	Arbitrary user name
Project	Project where the Bare Metal server is to be created
Role	Administrator role (cpf_admin), system owner role (cpf_systemowner), or operator role (cpf_operator)

Then, in the IaaS management of the FJcloud portal, add the created user to the project where the Bare Metal server is to be created, and grant the following role.

Item name	Value
Role	Operator role

Note

Do not change the authentication method of the created user from the password authentication.

See

For details on the FJcloud portal, refer to the official FUJITSU Hybrid IT Service FJcloud documentation.

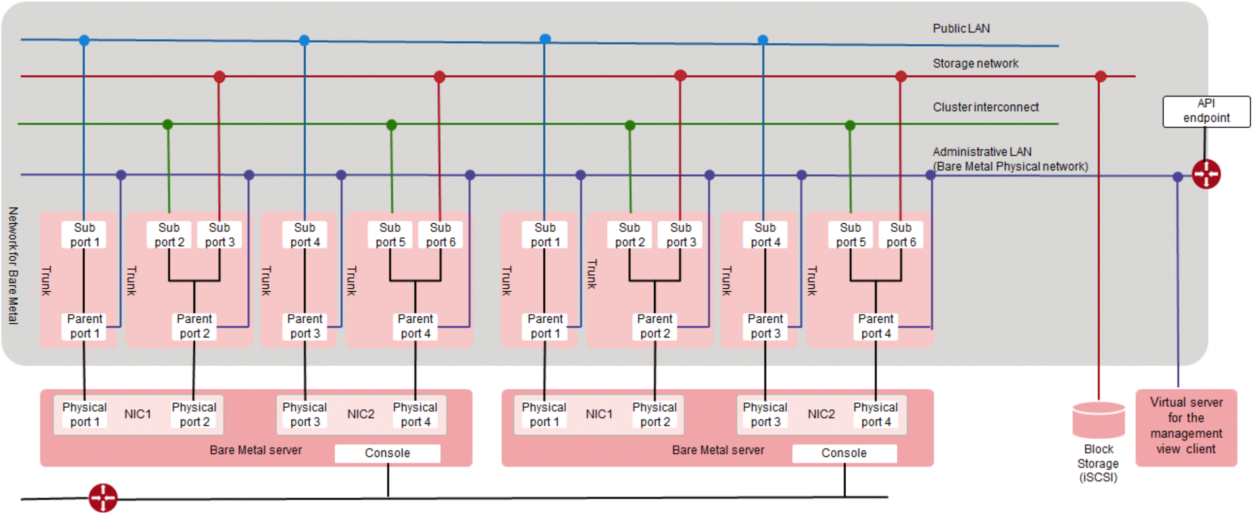
15.1.2 Building the Bare Metal Server

1. Refer to the documents provided by FJcloud-Baremetal and build the Bare Metal server and the network environment with the following network configurations.

   Setting the security group is not necessary. Specify an empty security group for each port.

Set the block storage (iSCSI) in "15.7.1 Connecting the Block Storage (iSCSI)."

Figure 15.1 Subnet design of the cluster system in FJcloud-Baremetal



Note

- When taking over the IP address between the servers, exclude the takeover IP address from the range of the pool for assigning the IP address of the subnet.
- Avoid duplicate server names within the project.
- For Bare Metal server names that use PRIMECLUSTER, the following ASCII characters can be used.
  Do not use other characters.
  - Uppercase letters
  - Lowercase letters
  - Numbers
  - "_" (Underscore)
  - "-" (Hyphen)

2. For each port where the IP address is taken over between the servers, execute the following API and set the takeover IP address as "allowed_address_pairs".

```
# curl -X PUT https://networking.jp-east-3.cloud.global.fujitsu.com/v2.0/ports/
{created_port_ID} -H 'X-Auth-Token:XXX' -H 'Content-Type:application/json' -H
'Accept:application/json' -d '{"port":{ "allowed_address_pairs" :
[{"ip_address":"takeover_IP_address"}]}}'
```

3. When using the firewall service, refer to "14.3.2.3 Firewall Rule" and set the firewall rule.

15.2 Creating the Virtual System

Refer to "3.1.5 Creating the Virtual Server for the Management Client" and create the management view client on FJcloud-O.

15.3 Presetting

1. Enable the firewall.

   Make sure that "firewalld" is enabled.

```
# systemctl is-enabled firewalld
```

If it is disabled, enable it.

```
# systemctl start firewalld
# systemctl enable firewalld
```

In addition, set the firewall according to "14.3.2.1 Firewalls Applied to the Cluster Node."

2. Set NTP.

   Make sure to set NTP when building the cluster to synchronize the time of each node in the cluster system.

   Set NTP before installing PRIMECLUSTER.

3. Disable the ACPI power management on the OS.

   1. Edit the following GRUB2 file on all nodes, and add the setting to disable the ACPI power management on the OS ("acpi=off") to the item GRUB_CMDLINE_LINUX.

      Example)

      ```
      # vi /etc/default/grub
      GRUB_TIMEOUT=5
      GRUB_DISTRIBUTOR="$(sed 's, release .*$,,g' /etc/system-release)"
      GRUB_DEFAULT=saved
      GRUB_DISABLE_SUBMENU=true
      GRUB_TERMINAL_OUTPUT="console"
      GRUB_CMDLINE_LINUX="crashkernel=auto spectre_v2=retpoline rd.lvm.lv=rhel/root
      rd.lvm.lv=rhel/swap rhgb quiet acpi=off"
      GRUB_DISABLE_RECOVERY="true"
      ```

   2. Rebuild the grub.cfg file on all nodes.

      - In a BIOS based environment

        ```
        # grub2-mkconfig -o /boot/grub2/grub.cfg
        ```

      - In a UEFI based environment

        ```
        # grub2-mkconfig -o /boot/efi/EFI/redhat/grub.cfg
        ```

   3. Restart all nodes.

   4. Make sure that "acpi=off" is added in the /proc/cmdline file on all nodes.

      Example)

      ```
      # cat /proc/cmdline
      BOOT_IMAGE=/vmlinuz-3.10.0-1127.el7.x86_64 root=/dev/mapper/rhel-root ro crashkernel=auto
      spectre_v2=retpoline rd.lvm.lv=rhel/root rd.lvm.lv=rhel/swap rhgb quiet LANG=ja_JP.UTF-8
      acpi=off
      ```

15.4 Installing PRIMECLUSTER

Use the installation script (CLI Installer) to install PRIMECLUSTER.

Install PRIMECLUSTER on each node in the system where Linux(R) software and Linux(R) related software are already installed. Use the same installation script when installing PRIMECLUSTER in the cluster management server.

Note

If OS has never been restarted since the Bare Metal server was built, restart it and then install PRIMECLUSTER.

See

For details on the installation/uninstallation procedure, refer to the sections of [Virtual environment, Red Hat OpenStack Platform environment] described in the Installation Guide for PRIMECLUSTER.

15.5 Checking and Setting the Kernel Parameters

Change the kernel parameters depending on the environment.

Applicable nodes:

All nodes on which PRIMECLUSTER is to be installed

Depending on the products and components utilized, different kernel parameters are required.

Check PRIMECLUSTER Designsheets and if you need to modify the kernel parameters, set them again.

See

For details on kernel parameters, refer to "3.1.7 Checking and Setting the Kernel Parameters" in "PRIMECLUSTER Installation and Administration Guide."

Note

To activate the modified kernel parameters, restart OS.

15.6 Installing and Setting up the Applications

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

See

- For details on environment setup, refer to manuals for each application.
- For information on PRIMECLUSTER-related products that support FJcloud-Baremetal, refer to the documentation for each product.

15.7 Preparation Prior to Building a Cluster

Prior to building a cluster, perform presettings such as the initial GLS setup, setting the DNS client, creating the FJcloud-Baremetal environment information file, and staring the Web-Based Admin View screen.

15.7.1 Connecting the Block Storage (iSCSI)

Refer to the documents provided by FJcloud-Baremetal and connect the block storage (iSCSI).

Also refer to "Setting Up DM-MP" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide" and set DM-MP.

15.7.2 Creating the FJcloud-Baremetal Environment Information File

To activate a cluster system in an FJcloud-Baremetal environment, create the FJcloud-Baremetal environment information file with the following procedure.

This setting is not necessary in a single-node cluster.

1. Create the /opt/SMAW/SMAWRrms/etc/k5_endpoint.cfg file on all nodes as shown below:

```
DOMAIN_NAME=domainname
PROJECT_NAME=projectname
IDENTITY=identityurl
COMPUTE=computeurl
```

```
domainname   : Domain name of FJcloud-Baremetal (contractor number)
projectname  : Project name building a cluster in FJcloud-Baremetal
identityurl  : URL of the endpoint for the regional user management of the region used in
               FJcloud-Baremetal (*)
computeurl   : URL of the endpoint for the compute (standard service) of the region used in
               FJcloud-Baremetal (*)

* For details on URL of the endpoint for the regional user management and the compute (standard
service), refer to the documents provided by FJcloud-Baremetal.
```

Example

```
DOMAIN_NAME=primecluster_domain
PROJECT_NAME=primecluster_project
IDENTITY=https://identity.jp-east-3.cloud.global.fujitsu.com
COMPUTE=https://compute.jp-east-3.cloud.global.fujitsu.com
```

2. Set the owner, group, and access rights as follows.

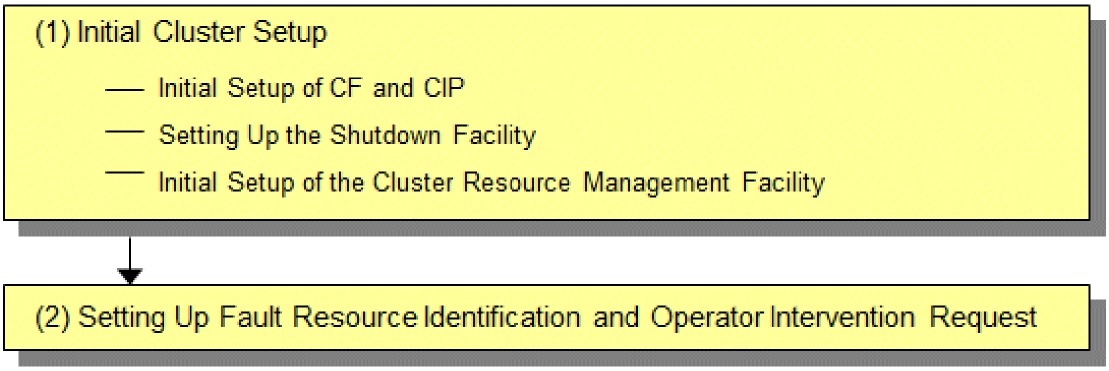
```
# chown root:root /opt/SMAW/SMAWRrms/etc/k5_endpoint.cfg
# chmod 600 /opt/SMAW/SMAWRrms/etc/k5_endpoint.cfg
```

15.7.3 Presettings for Building a Cluster

Refer to "Chapter 4 Preparation Prior to Building a Cluster" in "PRIMECLUSTER Installation and Administration Guide", and execute the initial setup for a cluster in the Bare Metal server.

15.8 Building a Cluster

The procedure for building a PRIMECLUSTER cluster is shown below:



(1) Initial Cluster Setup

- Initial Setup of 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

15.8.1 Initial Cluster Setup

This section describes the initial cluster setup for PRIMECLUSTER.

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

	Details	Manual reference location*
1	15.8.1.1 Initial Setup of CF and CIP (setting up cluster configuration information and IP addresses)	CF "1.1 CF, CIP, and CIM configuration"
2	15.8.1.2 Setting up the Shutdown Facility	CF "7 Shutdown Facility (SF)"
3	15.8.1.3 Initial Setup of the Cluster Resource Management Facility	CF "3.1 Resource Database configuration"

*The PRIMECLUSTER manual name is abbreviated as follows:

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

15.8.1.1 Initial Setup of CF and CIP

Refer to "5.1.1 Setting Up CF and CIP" in "PRIMECLUSTER Installation and Administration Guide" to set up CF and CIP.

However, in an FJcloud-Baremetal environment, CF cannot be built with Cluster Admin. For details on how to set up CF, refer to the configuration procedure in a cloud environment described in "1.1.6 Example of CF configuration by CLI" in "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

Note

You need to use CF over IP in a cloud environment. In addition, you need to configure CIP because such as RMS uses it for inter-node communication in the cluster. CF over IP (IP interconnect) and CIP are different features. For more information, see "1.1.1 Differences between CIP and CF over IP" in the Cluster Foundation Configuration and Administration Guide.

15.8.1.2 Setting up the Shutdown Facility

In an FJcloud-Baremetal environment, only the SA_vmk5r shutdown agent is available for setup.

This section describes the method for setting up the SA_vmk5r shutdown agent as the shutdown facility.

For details on the survival priority, refer to "5.1.2.1 Survival Priority." in "PRIMECLUSTER Installation and Administration Guide."

Note

- After setting up the shutdown agent, conduct a test for the forced stop of cluster nodes to make sure that the correct nodes can be forcibly stopped. For details of the test for the forced stop of cluster nodes, refer to "1.4 Test" in "PRIMECLUSTER Installation and Administration Guide."
  After the node was successfully forcibly stopped, make sure that the following messages are not output in /var/log/messages of the node.

  ```
  systemd-logind: Power key pressed.
  systemd-logind: Powering Off...
  systemd-logind: System is powering down.
  ```

- The contents of the SA_vmk5r.cfg file and the rcsd.cfg file of all nodes should be identical. If not, a malfunction will occur.
- If you changed a user password created in "15.1.1 Creating the User for the Forced Stop", perform this step again with a new password.
- Be sure to perform the following operations on all nodes.

1. Set the shutdown daemon.

   Create /etc/opt/SMAW/SMAWsf/rcsd.cfg with the following contents on all nodes in the cluster system.

   ```
   CFNameX,weight=weight,admIP=myadmIP:agent=SA_vmk5r,timeout=90
   CFNameX,weight=weight,admIP=myadmIP:agent=SA_vmk5r,timeout=90
   ```

   ```
   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 used in the shutdown facility
                 of the cluster host.
                 Available IP addresses are IPv4.
                 When specifying a host name, make sure it is described in /etc/hosts.
```

Example) The following is a setup example.

```
# cat /etc/opt/SMAW/SMAWsf/rcsd.cfg
node1,weight=1,admIP=192.168.1.1:agent=SA_vmk5r,timeout=90
node2,weight=1,admIP=192.168.1.2:agent=SA_vmk5r,timeout=90
```

Create /etc/opt/SMAW/SMAWsf/rcsd.cfg and then set the owner, group, and access rights as follows.

```
# chown root:root /etc/opt/SMAW/SMAWsf/rcsd.cfg
# chmod 600 /etc/opt/SMAW/SMAWsf/rcsd.cfg
```

Information

When creating the /etc/opt/SMAW/SMAWsf/rcsd.cfg file, the /etc/opt/SMAW/SMAWsf/rcsd.cfg.template file can be used as a template.

2. Encrypt the password.

   Execute the sfcipher command to encrypt a password of a user for forcibly stopping the Bare Metal server of FJcloud-Baremetal. For details on how to use the sfcipher command, refer to the manual page of "sfcipher."

   ```
   # sfcipher -c
   ```

   Example) The following is a setup example.

   If a password is "k5admin$":

   ```
   # sfcipher -c
   Enter Password:    <- Enter k5admin$
   Re-Enter Password: <- Enter k5admin$
   O/gm+AYuWwE7ow3dgVG/Nw==
   ```

3. Set the shutdown agent.

   Create /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg with the following contents on all nodes in the cluster system.

   Delimit each item with a single space.

   ```
   CFNameX ServerName user passwd {cycle | leave-off}
   CFNameX ServerName user passwd {cycle | leave-off}
   ```

   ```
   CFNameX           : Specify the CF node name of the cluster host.
   ServerName        : Specify the Bare Metal server name in FJcloud-Baremetal on which
                       the cluster host is running.
                       For Bare Metal server names that use PRIMECLUSTER,
                       the following ASCII characters can be used.
                       Do not use other characters.
                       - Uppercase letters
                       - Lowercase letters
                       - Numbers
                       - "_" (Underscore)
                       - "-" (Hyphen)
   user              : Specify a user name for forcibly stopping the Bare Metal server.
   passwd            : Specify a password encrypted in step 2.
   cycle             : Restart the node after forcibly stopping the node.
   leave-off         : Power-off the node after forcibly stopping the node.
   ```

   Example) The following is a setup example.

   This example shows the following settings:

- The CF node names of the cluster host are node1 and node2.

- The Bare Metal server names are vm1 and vm2.

- The user name to forcibly stop the Bare Metal server is pcl.

- The node will be restarted when it is forcibly stopped.

```
# cat /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg
node1 vm1 pcl O/gm+AYuWwE7ow3dgVG/Nw== cycle
node2 vm2 pcl O/gm+AYuWwE7ow3dgVG/Nw== cycle
```

Create /etc/opt/SMAW/SMAWsf/ SA_vmk5r.cfg and then set the owner, group, and access rights as follows.

```
# chown root:root /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg
# chmod 600 /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg
```

Note

- Make sure that the /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg file is set correctly. If the setting is incorrect, the shutdown facility cannot be performed normally.

- Make sure that the Bare Metal server name (ServerName) corresponding to the CF node name (CFNameX) of the cluster host of the /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg file is set. If the setting is incorrect, an incorrect node will be forcibly stopped.

4. Start the shutdown facility.

   Check if the shutdown facility has been started on all nodes in the cluster system.

```
# sdtool -s
```

   On a node where the shutdown facility has already been started, execute the following commands to restart the shutdown facility.

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

   On a node where the shutdown facility has not been started, execute the following command to start the shutdown facility.

```
# sdtool -b
```

Information

You can check if the shutdown facility has already been started with the sdtool -s command. If "The RCSD is not running" is displayed, the shutdown facility is not started.

5. Check the status of the shutdown facility.

   Execute the following command with all nodes in the cluster system to check the status of the shutdown facility.

```
# sdtool -s
```

Note

- If "The RCSD is not running" is displayed, there is a failure in the shutdown daemon or shutdown agent settings. Perform the procedure from step 1 to 4 again.

- A user created in "15.1.1 Creating the User for the Forced Stop" needs a periodical change of the password (every 90 days). For the procedure on changing a password, refer to "18.1 Changing a Password Periodically."

- If you changed the Bare Metal server name created in "15.1.2 Building the Bare Metal Server", perform the procedure from step 3 to 5 again.

Information

Display results of the sdtool -s command

- If Unknown or Init-ing is displayed in Init State, wait for about one minute, and then check the status again.
- If Unknown is displayed in Shut State, it means that SF has not yet stopped the node. If Unknown is displayed in Init State, it means that SF has not yet initialized SA or tested the route. Unknown is displayed temporarily in Test State or Init State until the actual status can be confirmed.
- If TestFailed is displayed in Test State, 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 InitFailed is displayed in Init State, communication with the endpoint of the regional user management or the compute (standard service) in FJcloud-Baremetal is disabled or the setting might have a failure. Check the following and then set the following again.
  After the failure-causing problem is resolved and SF is restarted, the status display changes to InitWorked or TestWorked.

  a. Execute the following command and check if the Bare Metal server on which the cluster host is running can communicate with the endpoint of the regional user management.

     ```
     # curl -k -s -X GET <URL of the endpoint of the regional user management>/v3/
     ```

     If an error occurs, check the following.

     - The security groups, the firewall service, and the firewall of OS in FJcloud-Baremetal must be set properly.

     - The virtual router of FJcloud-Baremetal must be created.

     - The default router of the cluster host must be set in the virtual router.

     - URL of the endpoint of the regional user management must be correct.

     - The DNS server used in the cluster host must be set.

  b. Execute the following command and check if the Bare Metal server on which the cluster host is running can communicate with the endpoint of the compute (standard service).

     ```
     # curl -k -s -X GET <URL of the endpoint of the compute (standard service)>/v2/
     ```

     If the following message is displayed, it is a normal operation.

     ```
     {"error": {"message": "The request you have made requires authentication.", "code": 401,
     "title": "Unauthorized"}}
     ```

     If a message other than the above message was displayed, check the following.

     - The security groups, the firewall service, and the firewall of OS in FJcloud-Baremetal must be set properly.

     - The virtual router of FJcloud-Baremetal must be created.

     - The default router of the cluster host must be set in the virtual router.

     - URL of the endpoint of the compute (standard service) must be correct.

     - The DNS server used in the cluster host must be set.

  c. Make sure that the following settings are correct:

     - The domain name (contractor number), project name, URL of the endpoint for the regional user management, and URL of the endpoint for the compute (standard service) for the FJcloud-Baremetal environment information file (/opt/SMAW/SMAWRrms/etc/k5_endpoint.cfg)

     - All of CF node name, Bare Metal server name, user name, and encrypted password in the configuration file of the shutdown agent (/etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg)

15.8.1.3 Initial Setup of the Cluster Resource Management Facility

Refer to "5.1.3 Initial Setup of the Cluster Resource Management Facility" in "PRIMECLUSTER Installation and Administration Guide" to set up the resource database managed by the cluster resource management facility.

15.8.2 Setting up Fault Resource Identification and Operator Intervention Request

Refer to "5.2 Setting up Fault Resource Identification and Operator Intervention Request" in "PRIMECLUSTER Installation and Administration Guide" to set up the fault resource identification and the operator intervention request.

15.9 Building the Cluster Application

For the detail on how to build the cluster application, refer to "Chapter 6 Building Cluster Applications" in "PRIMECLUSTER Installation and Administration Guide."

Set the shared disk of GDS in this setting.

Note

If the icmp communication between cluster nodes is not allowed in the firewall configuration, the following message is displayed when the clchkcluster command is executed.

```
Admin IP <IP address> used by SF is not alive.
```

If this message is output, refer to "14.3.2.1 Firewalls Applied to the Cluster Node", and set the icmp protocol rule to allow the icmp communication between cluster nodes. After that, execute the clchkcluster command again.

Chapter 16 Operations

For details on functions for managing PRIMECLUSTER system operations, refer to "Chapter 7 Operations" in "PRIMECLUSTER Installation and Administration Guide."

See

For detail on how to operate GDS, refer to "Operation and Maintenance" of "PRIMECLUSTER Global Disk Services Configuration and Administration Guide" and how to operate GLS, refer to "GLS operation on cluster systems" in "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function."

Chapter 17 Changing the Configurations

For details on changing the configuration information for the PRIMECLUSTER system, environment settings, the configuration of the cluster application, the operation attributes of the cluster system, refer to "Chapter 9 Changing the Cluster System Environment", "Chapter 10 Configuration Change of Cluster Applications", "Chapter 11 Changing the Operation Attributes of a Cluster System" in "PRIMECLUSTER Installation and Administration Guide." For details on changing the GDS configuration, refer to "Configuration Change" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

Chapter 18 Maintenance

When you maintain the PRIMECLUSTER system in an FJcloud-Baremetal environment, refer to "Chapter 12 Maintenance of the PRIMECLUSTER System" in "PRIMECLUSTER Installation and Administration Guide." For details on how to maintain GDS, refer to "Operation and Maintenance" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide." For details on how to maintain GLS, refer to "Maintenance" in "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function."

18.1 Changing a Password Periodically

A user created in "15.1.1 Creating the User for the Forced Stop" needs a periodical change of the password (every 90 days). If you do not change the password even after 90 days, the shutdown facility will not be operated.

To change a password, perform the following procedure.

1. Change a user password created in FJcloud-Baremetal.

2. Set the shutdown facility again with the changed password according to step 2 to 5 in "15.8.1.2 Setting up the Shutdown Facility."

If you do not change the password even after 90 days or do not set the shutdown facility again even if you change the password, the following message is displayed in the file, /var/log/messages and TestState displayed with sdtool -s will be "TestFailed."

```
SF: The authentication request failed.
SMAWsf : SA SA_vmk5r to test host <CF node name>  failed
```

Part 4 AWS Environment

This part describes the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system in an AWS environment.

Chapter 19 Cluster Systems in an AWS Environment

PRIMECLUSTER enables the clustering of an instance (a virtual server is called instance in AWS) in the virtual private cloud of AWS (hereinafter VPC). This enables higher business availability than the availability provided by cloud services.

Using PRIMECLUSTER enables the redundancy of back-end servers and front-end servers, the use of cloud-based cluster systems from on-premises systems, and building of the cluster system between Availability Zones.

If you are using PRIMECLUSTER Cloud Edition, see "Appendix A Using the Smart Workload Recovery Feature" in this guide.

See

For details on AWS, refer to the official AWS documentation.

The following cluster systems are available in an AWS environment:

- Cluster system in multiple Availability Zones (Multi-AZ)
- Cluster system in a single Availability Zone (Single-AZ)

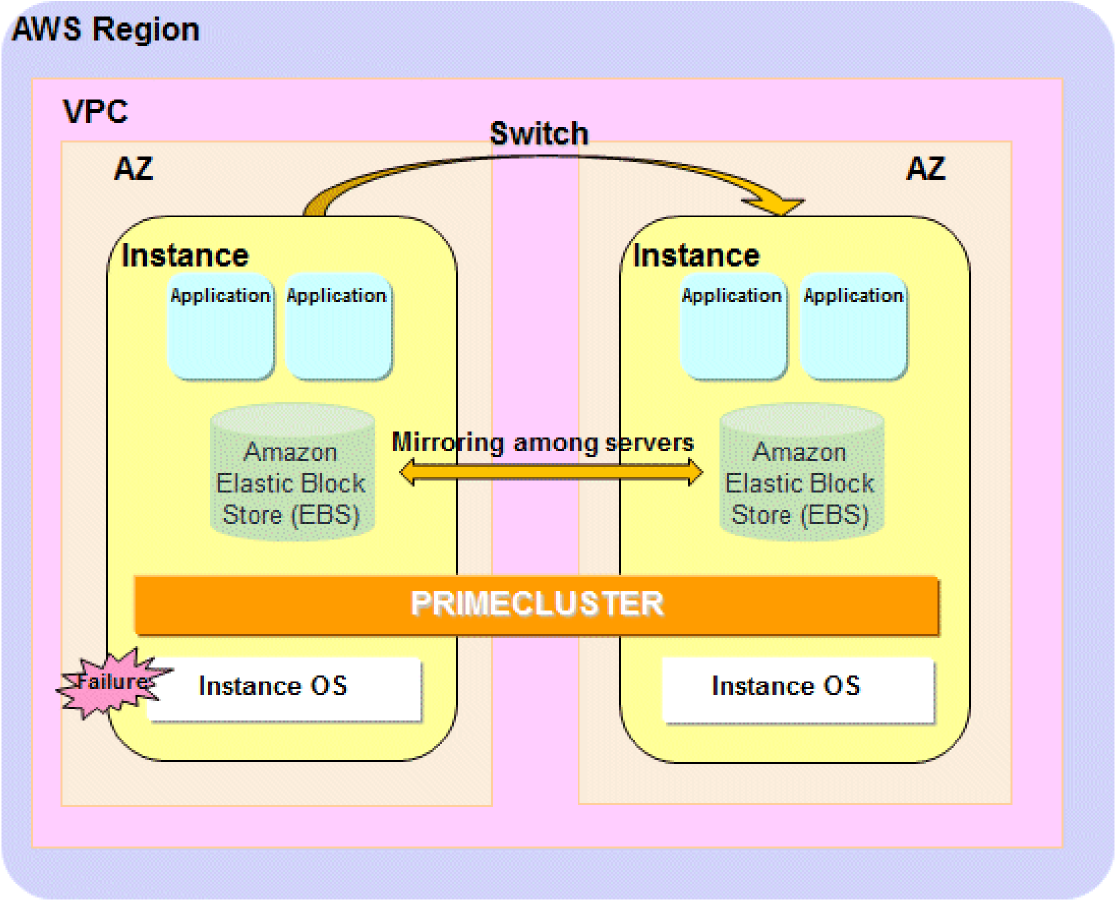
19.1 Cluster System in Multiple Availability Zones (Multi-AZ)

In this configuration, the cluster system can be operated on multiple Availability Zones.

By applying PRIMECLUSTER to the instance, in the event of an error, an application can be switched from the operational instance to the standby instance in a short time to provide a highly reliable instance environment.

Also when a network failure occurs in the entire Availability Zone or when Availability Zones become abnormal due to a large-scale disaster, business can be quickly restored by failing over to the standby system with fewer operations.

Figure 19.1 Cluster system in multiple Availability Zones (Multi-AZ)



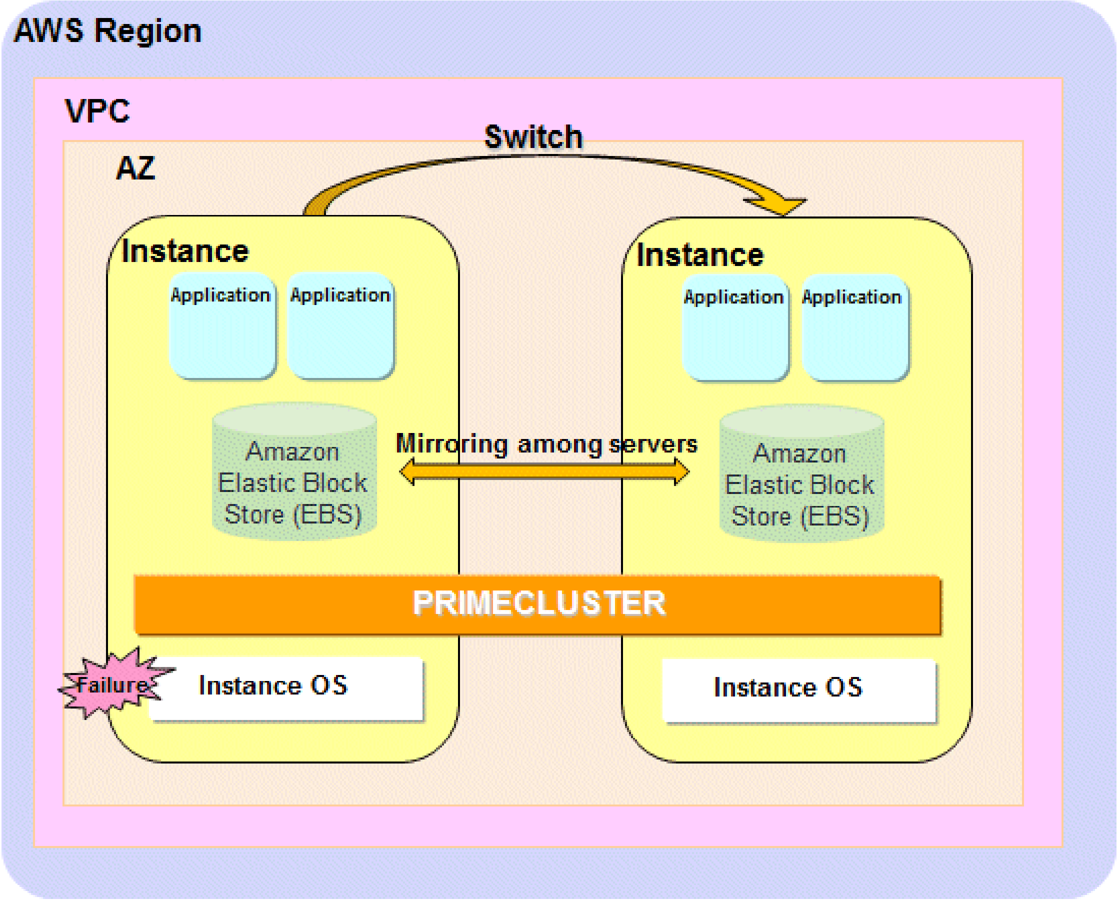
19.2 Cluster System in a Single Availability Zone (Single-AZ)

In this configuration, the cluster system can be operated in one Availability Zone. By applying PRIMECLUSTER to the instance, in the event of an error, an application can be switched from the operational instance to the standby instance in a short time to provide a highly reliable instance environment.

Note

In the cluster system in a single Availability Zone, in the event of an error in the AZ where the cluster is built, all cluster nodes stop and the business stops.

Figure 19.2 Cluster system in a single Availability Zone (Single-AZ)



19.3 Supported Range

This section describes the range of support of PRIMECLUSTER in an AWS environment.

Supported configurations

- Number of cluster nodes: 1 to 2 nodes
- Operation mode of the cluster system: 1:1 Standby operation, Mutual standby, Single-node cluster, Scalable operation (only when using Symfoware Server (Native) Hot Standby feature)
- Network configurations: Instances in the cluster system must communicate with the API endpoints.
- Configurations to take over volume data: Data takeover by the mirroring among servers of GDS using EBS

Supported monitoring functions

- Error of AZ (cluster system in multiple Availability Zones)

  The cyclic monitoring of the cluster interconnect detects a fault of the AZ and the service is automatically switched to the standby system.

- Error of the guest OS and the cluster interconnect

  The cyclic monitoring of the cluster interconnect detects a hang-up of the guest OS or an error of the cluster interconnect, and the service is automatically switched to the standby system.

- Error of the disk access

  By combining the volume management function of GDS, an error of the disk access can be detected (monitored by the Gds resource), and the service is automatically switched to the standby system when the disk access is disabled.

- Error of the cluster application

  When a resource error of the cluster application occurs, the service is automatically switched to the standby system.

Note

- The snapshot of the instance can be acquired only when the OS is stopped.
- The following functions for PRIMECLUSTER are not available:
  - GFS
  - GLS
  - GDS Snapshot
  - Root class and local class of GDS
  - Single volume of GDS, and disk groups of GDS other than the netmirror type (mirror, stripe, concatenation, and switch)
- The console cannot be used in the instance in an AWS environment. Do not set the single user mode.
- AWS features that involve node operations (Auto Scaling, Auto Recovery, etc.) are not available.
- In an AWS environment, when an OS panic occurs, the cluster node may be powered off while the memory dump is being output, and it may not be possible to collect a complete memory dump.
- When using the asynchronous forcible stop method, a shared file system service cannot be used.

Chapter 20 Design

You must prepare the items listed below before building the PRIMECLUSTER system in an AWS environment.

- Selecting the PRIMECLUSTER Product
- Selecting the Architectural Pattern
- Network Design
- System Design
- Determining the Cluster System Operation Mode
- Determining the Web-Based Admin View Operation Mode
- Forcible Stop Method
- Determining the Failover Timing of Cluster Application
- Policy Design

Point

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

Information

For the flow to build the PRIMECLUSTER system, refer to "Chapter 1 Build Flow" in "PRIMECLUSTER Installation and Administration Guide."

20.1 Selecting the PRIMECLUSTER Product

Select a PRIMECLUSTER product.

In an AWS environment, you can select the following products.

For details on the PRIMECLUSTER products, refer to "2.1 PRIMECLUSTER Product Selection" in "PRIMECLUSTER Installation and Administration Guide."

- PRIMECLUSTER Enterprise Edition (EE)
- PRIMECLUSTER HA Server (HA)
- PRIMECLUSTER Clustering Base (CB)

20.2 Selecting the Architectural Pattern

In the PRIMECLUSTER system in an AWS environment, select an architecture pattern for each item below.

- Network Takeover
- Ensuring Connectivity with API Endpoint
- Ensuring Connectivity with Web-Based Admin View

20.2.1 Network Takeover

PRIMECLUSTER provides architectural patterns for taking over the network in cluster systems on public clouds. For smooth design of a cluster system, choose from these architectural patterns.

The following are the architectural patterns for taking over the network and the appropriate scenarios for each pattern.

Table 20.1 Architectural patterns and appropriate scenarios for taking over the network

Architectural pattern	Appropriate scenario	Note
Network takeover by the virtual router	The cluster system of back-end servers used by the client in the VPC	- The cluster system is secured as an access from public sites is blocked. - The client must be deployed in the VPC.
Network takeover by replacing the Elastic IP address	The cluster system of front-end servers	- Accessible from public sites. - To assure the security, additional access control is required.
Network takeover by rewriting DNS records	The cluster system of back-end servers used by the client on-premises or in any other VPC	- Accessible via VPN from the client. - An additional device is required for a VPN connection.

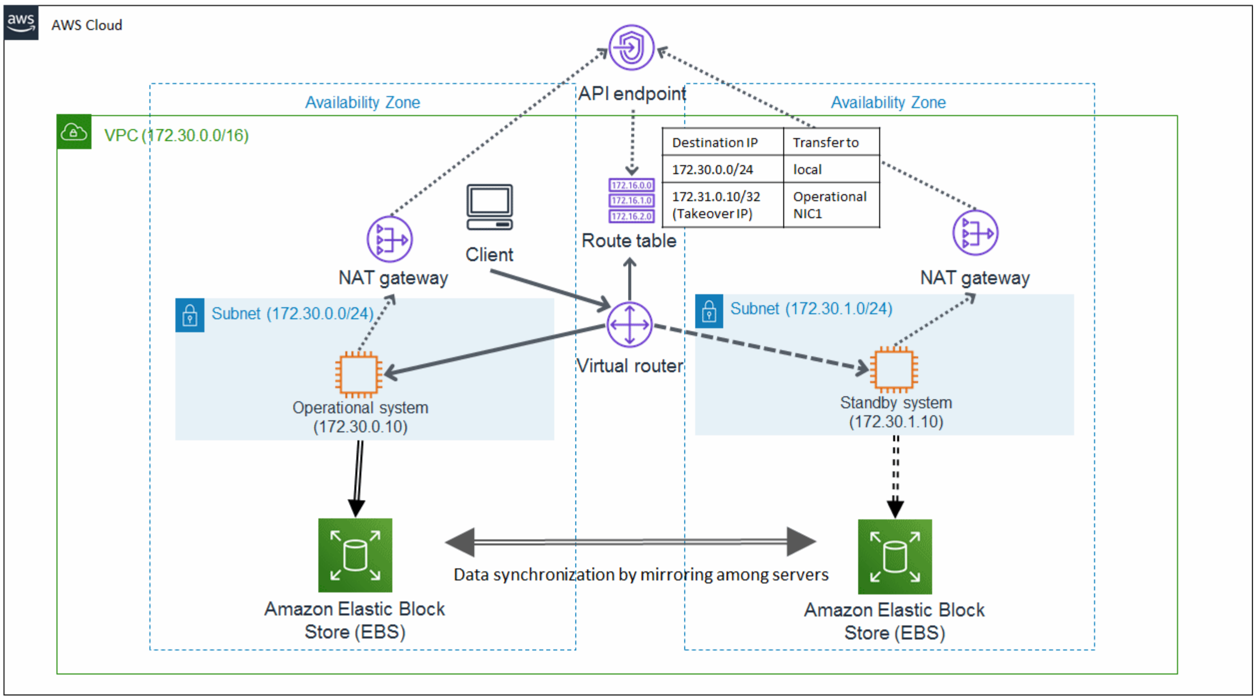
20.2.1.1 Network Takeover by the Virtual Router

This is an architectural pattern using the cluster system from the client in the VPC. Select this architecture when building the cluster system in back-end servers used by the client in the VPC.

In this architectural pattern, the network takeover is enabled by rewriting the route table of the virtual router provided by AWS. In the event of a cluster node error, PRIMECLUSTER automatically rewrites the route table and switches to the standby system to take over the IP.

An access from public sites (clients outside of the VPC) is blocked and the cluster system is secured.

Figure 20.1 Taking over the network by the virtual router



Destination IP	Transfer to
172.30.0.0/24	local
172.31.0.10/32 (Takeover IP)	Operational NIC1

20.2.1.2 Network Takeover by Replacing the Elastic IP Address

This is an architectural pattern that allows an access to the cluster node from public networks (the Internet). Select this architecture when building the cluster system in front-end servers.

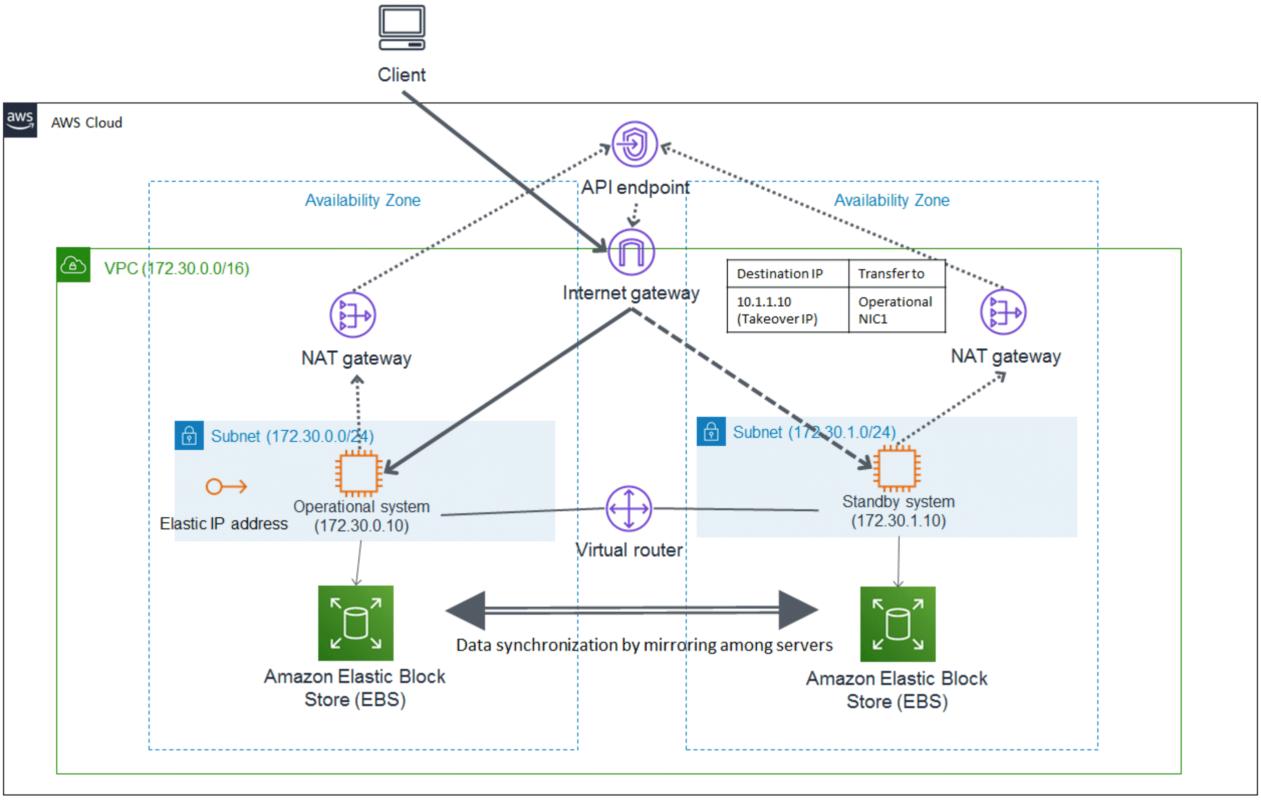
In this architectural pattern, the network takeover is enabled by associating and controlling the operational network interface as a transfer destination for the Elastic IP obtained as the takeover IP. In the event of a cluster node error, PRIMECLUSTER changes the transfer destination of the takeover IP to the network interface of the takeover destination of the application and switches to the standby system to take over the IP.

The cluster system provides the client with the Elastic IP address as the takeover IP. The client accesses this Elastic IP address.

Note

To prevent unauthorized accesses from public sites with the administrator authority, do not co-use the administrative LAN and the public LAN.

Figure 20.2 Taking over the network by replacing the Elastic IP address



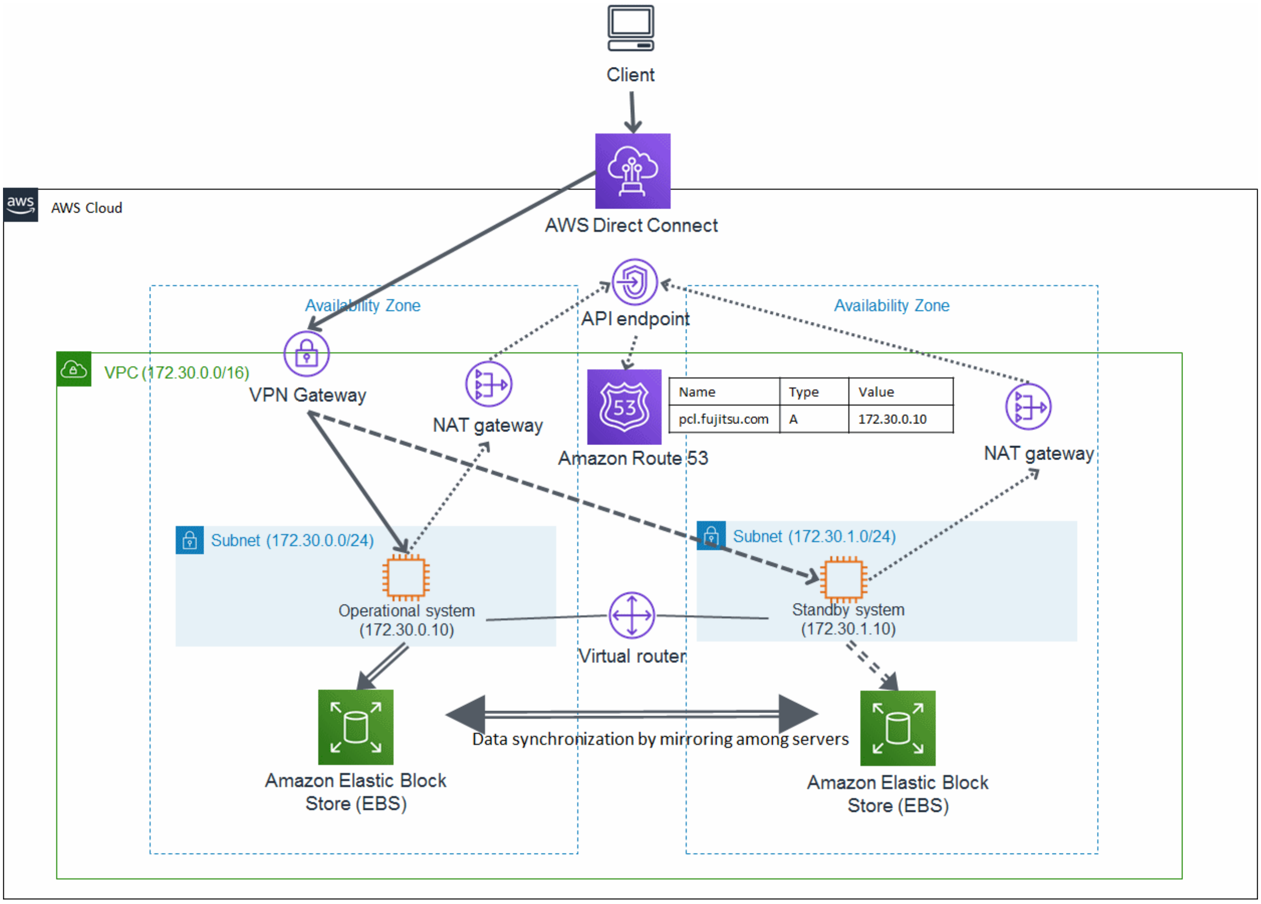
Destination IP	Transfer to
10.1.1.10 (Takeover IP)	Operational NIC1

20.2.1.3 Network Takeover by Rewriting DNS Records

This is an architectural pattern using a VPN service from the client system on-premises to allow a direct access to the cluster node. Hybrid configurations such as on-premises for the application layer and cloud for the database access layer are possible. Select this architecture if the network takeover by rewriting the route table is unavailable when using a VPN service.

In this architectural pattern, the network takeover is enabled by rewriting records of a DNS service provided by AWS. In the event of a cluster node error, PRIMECLUSTER takes over the IP by replacing the related records held by a DNS service with the standby private IP.

Figure 20.3 Taking over the network by rewriting DNS records



Name	Type	Value
pcl.fujitsu.com	A	172.30.0.10

20.2.2 Ensuring Connectivity with API Endpoint

When using PRIMECLUSTER on public clouds, connectivity between the cluster node and the API endpoint must be ensured for network takeover or to deal with split-brain. Use the API endpoint to control the power of instances or to change the settings of network components such as a virtual router.

PRIMECLUSTER provides architectural patterns for ensuring the connectivity between the cluster node and the API endpoint. For smooth design of a cluster system, choose from these architectural patterns.

The following are the architectural patterns for ensuring the connectivity between the cluster node and the API endpoint and the appropriate scenarios for each pattern.

Table 20.2 Architectural patterns and appropriate scenarios for ensuring the connectivity between the cluster node and the API endpoint

Architectural pattern	Appropriate scenario	Note
Ensuring connectivity with NAT gateway	Ensuring low cost and secure connectivity	Operation management is not required by the user since a NAT gateway is an AWS managed service.
Ensuring connectivity with NAT instance	Ensuring secure and flexible connectivity	Operation management is flexible for the user since a NAT instance is not an AWS managed service.
Ensuring connectivity with Elastic IP address	Ensuring connectivity of front-end servers with a simple architectural pattern	- To allow access from the public site to the administrative LAN that controls the server, the access control (ACL) or security rules of security groups must be firmly established.

Architectural pattern	Appropriate scenario	Note
		- This architecture is simple because there are a low number of components since NAT gateways and NAT instances are not required.
Ensuring connectivity with PrivateLink	Ensuring connectivity of back-end servers with a simple architectural pattern	- This pattern cannot be selected when selecting the network takeover pattern by rewriting DNS records. - This architecture is simple because there are a low number of components since NAT gateways and NAT instances are not required.

20.2.2.1 Ensuring Connectivity with NAT Gateway

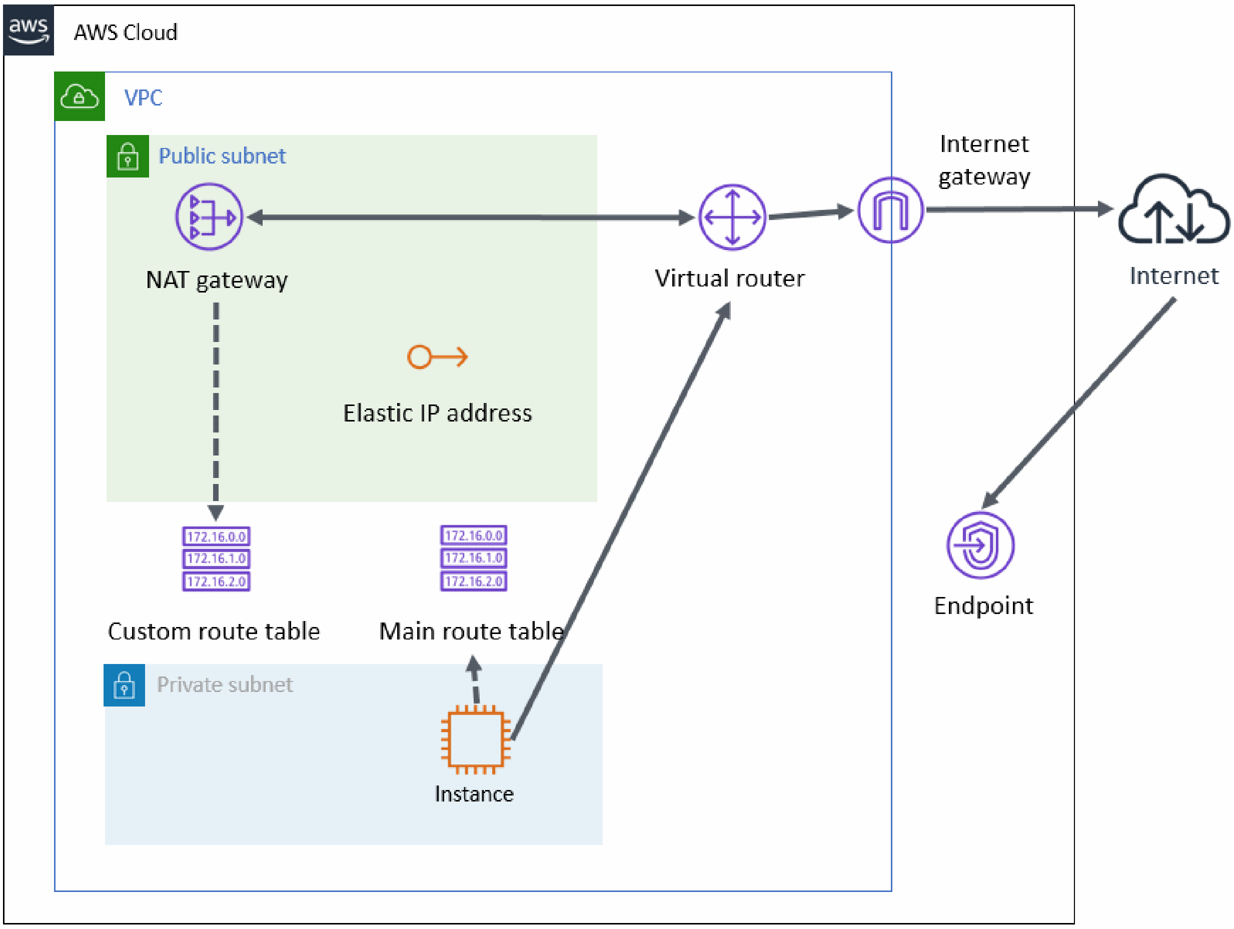
This is an architecture pattern with a NAT gateway that can block access from public sites outside the VPC, and ensure connectivity between the cluster node and the API endpoint.

Place one instance that configures the cluster system of back-end servers in a private subnet. This instance is not given a public IP and has no direct connectivity with the Internet. Use a NAT gateway for connectivity with the endpoint since the API endpoint that forcibly stops PRIMECLUSTER or enables the network takeover exists over the Internet. A NAT gateway is a service managed by AWS and the user does not have to operate it.

See

For details on NAT gateways, refer to the official AWS documentation.

Figure 20.4 Ensuring connectivity with NAT gateway



To control traffic forwarding when using this architectural pattern, set the following entries for the main route table and the custom route table.

Main route table (associated with the private subnet)

Destination CIDR	Forward to
CIDR of VPC	Network in the VPC
0.0.0.0/0	NAT gateway

Custom route table (associated with the public subnet)

Destination CIDR	Forward to
CIDR of VPC	Network in the VPC
0.0.0.0/0	Internet gateway

In this configuration, the traffic from the private subnet to the Internet is forwarded as follows and the NAT gateway blocks the traffic from public sites to the private subnet. This provides one-way connectivity from the private subnet to the Internet.

1. According to the entries of the main route table, the traffic from an instance in the private subnet to the Internet is forwarded to the NAT gateway by the virtual router.

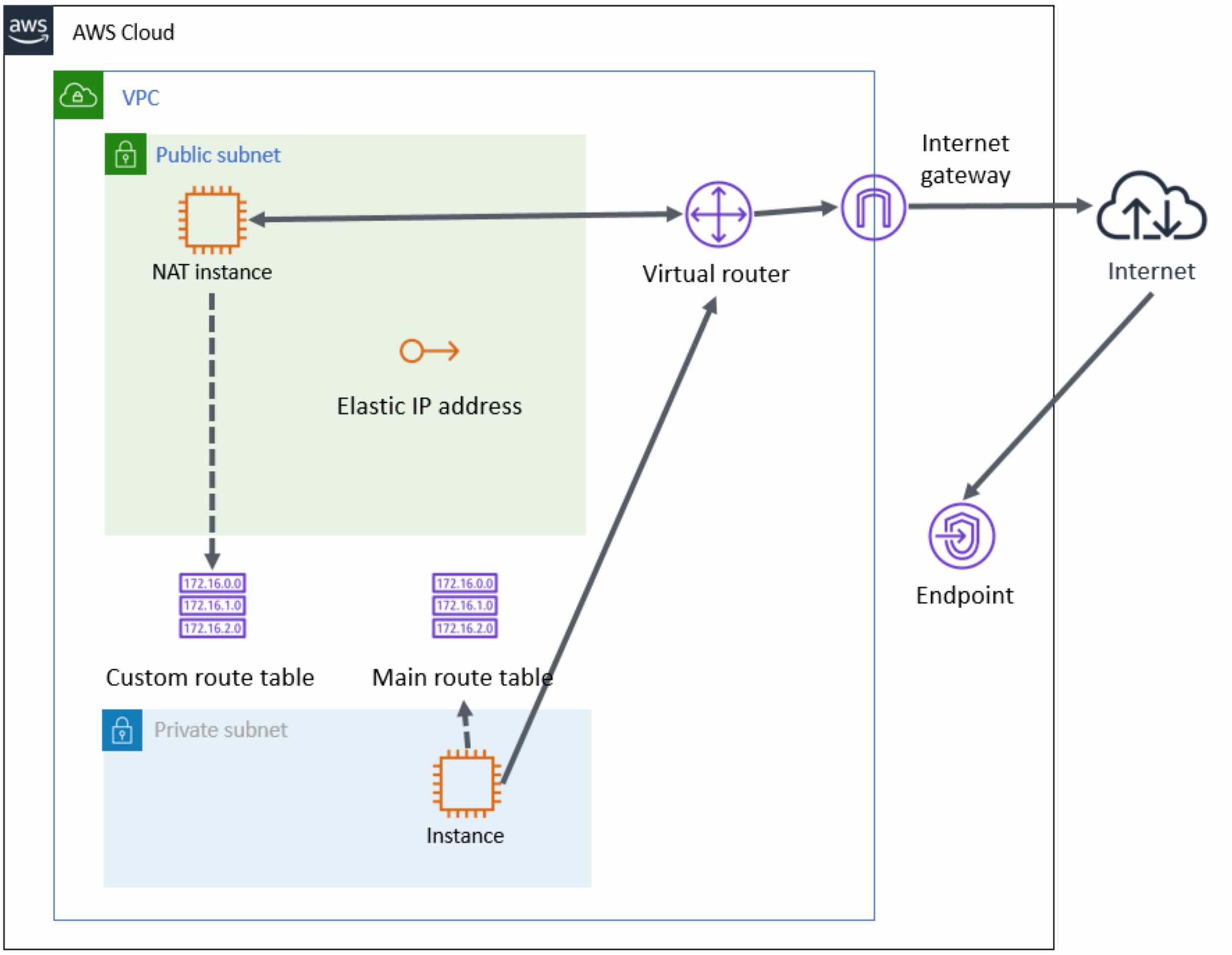
2. According to the entries of the custom route table, the traffic forwarded to the NAT instance is forwarded to the Internet gateway by the virtual router.

20.2.2.2 Ensuring Connectivity with NAT Instance

This is an architectural pattern with a NAT instance that can block access from public sites outside the VPC, and ensure connectivity between the cluster node and the API endpoint. This architectural pattern is identical to the NAT gateway architectural pattern, except that the NAT gateway is replaced by the NAT instance in the placement of components.

Unlike the NAT gateway, the NAT instance is not a service managed by AWS. The user can freely manage operations. When building the cluster system in the back-end servers and for flexible operations, select this architectural pattern.

Figure 20.5 Ensuring connectivity with NAT instance



To control traffic forwarding when using this architectural pattern, set the following entries for the main route table and the custom route table.

Main route table (associated with the private subnet)

Destination CIDR	Forward to
CIDR of VPC	Network in the VPC
0.0.0.0/0	NAT instance

Custom route table (associated with the public subnet)

Destination CIDR	Forward to
CIDR of VPC	Network in the VPC
0.0.0.0/0	Internet gateway

In this configuration, the traffic from the private subnet to the Internet is forwarded as follows and the NAT instance blocks the traffic from public sites to the private subnet. This provides one-way connectivity from the private subnet to the Internet.

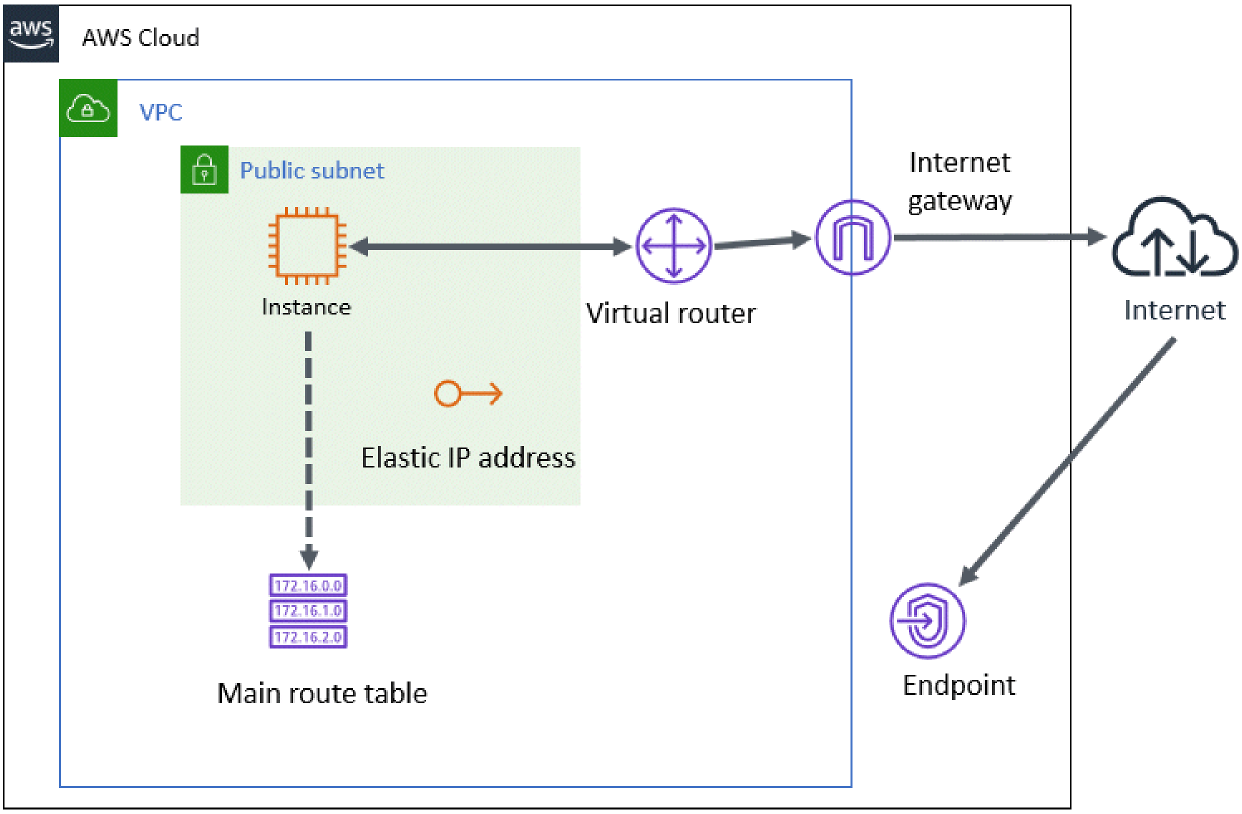
1. According to the entries of the main route table, the traffic from an instance in the private subnet to the Internet is forwarded to the NAT instance by the virtual router.
2. According to the entries of the custom route table, the traffic forwarded to the NAT instance is forwarded to the Internet gateway by the virtual router.

20.2.2.3 Ensuring Connectivity with Elastic IP Address

The architectural pattern with the Elastic IP address is accessible from public sites. It is suitable for the cluster system of front-end servers. This architectural pattern is simpler than architectural patterns with the NAT gateway or the NAT instance. However, since the cluster node is accessible from public sites, the access control (ACL) or security rules of security groups must be firmly established.

When using this architectural pattern, you only need to give the instance an Elastic IP address.

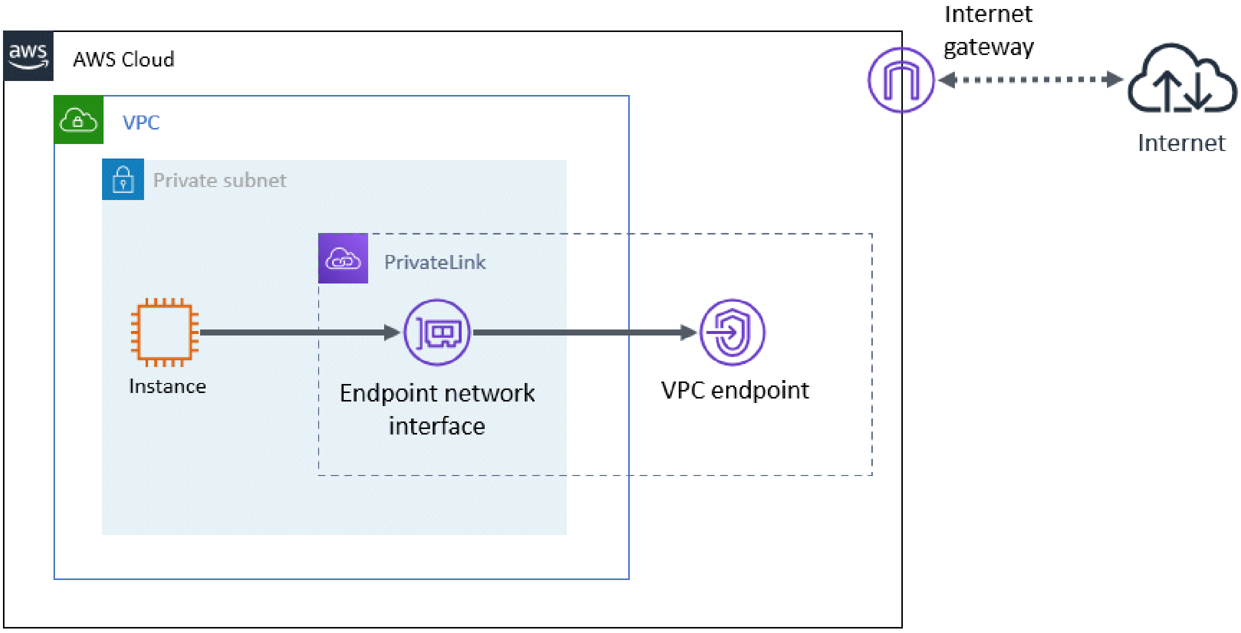
Figure 20.6 Ensuring connectivity with Elastic IP address



20.2.2.4 Ensuring connectivity with PrivateLink

The architectural pattern with the PrivateLink can ensure connectivity with the end point without preparing the public subnet by creating a VPC endpoint. This architectural pattern is simpler than the architectural patterns with the NAT gateway or the NAT instance. Also, unlike the pattern with the Elastic IP address, this pattern is securer since access from public sites can be blocked. However, the VPC endpoint of Amazon Route 53 (domain name system web service) cannot be created and if the architectural pattern of the network takeover by replacing the DNS records is selected, this architectural pattern cannot be selected.

Figure 20.7 Ensuring connectivity with PrivateLink



20.2.3 Ensuring Connectivity with Web-Based Admin View

PRIMECLUSTER provides architectural patterns for ensuring the connectivity between a terminal directly operated by a user and the management view.

For smooth designing of a cluster system, choose the appropriate one from these architectural patterns.

The following are the architectural patterns for the connectivity with the Web-Based Admin View and the appropriate scenarios for each pattern.

Table 20.3 Architectural patterns and appropriate scenarios for the connectivity with the Web-Based Admin View

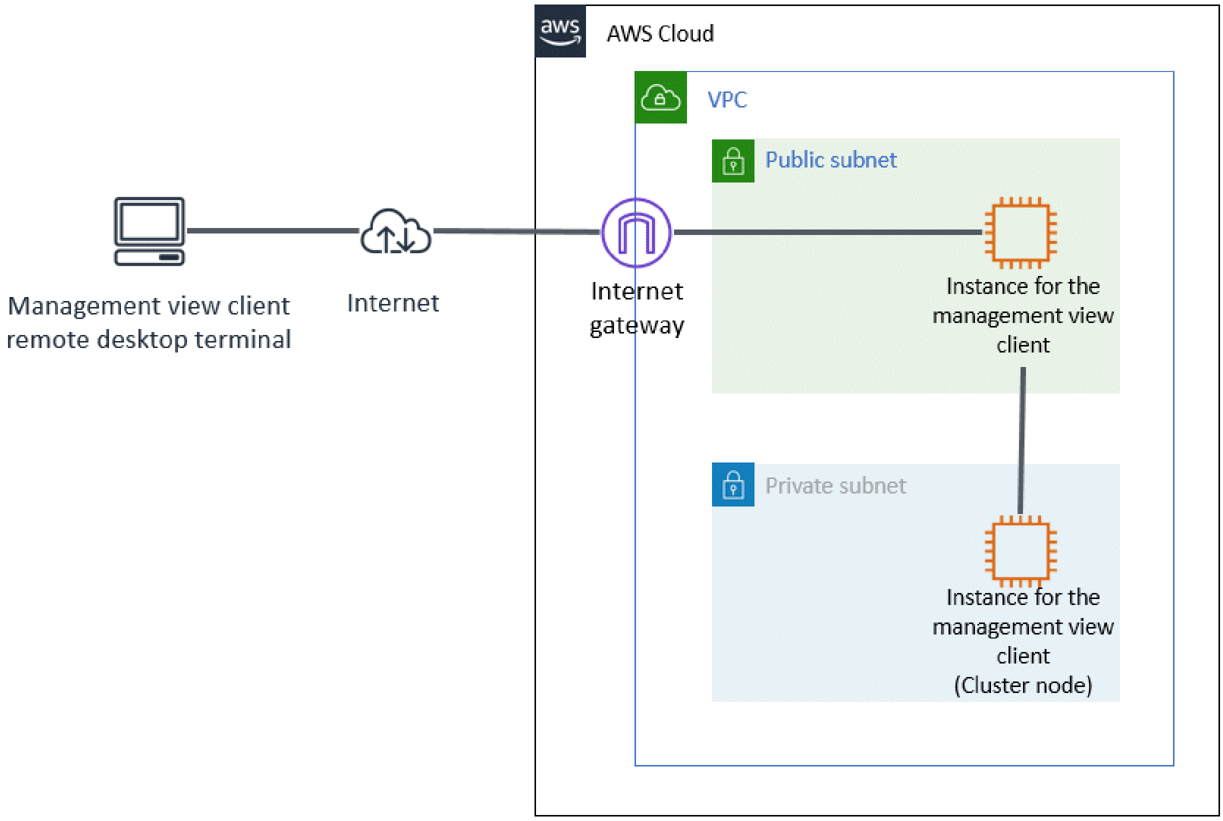
Architectural pattern	Appropriate scenario	Note
Ensuring connectivity with an instance for a client	Pattern using an instance for a client	An instance for the management view client must be deployed in the public subnet.
Ensuring connectivity using a VPN connection	Pattern using a VPN connection	An additional device is required for a VPN connection.

20.2.3.1 Ensuring Connectivity with Instance for Client

In the pattern using an instance for a client, to ensure the connectivity between a terminal directly operated by a user and the management view, prepare an instance for the management view client deployed in the public subnet. The system component for a VPN connection is not required and the configuration may be simple.

When selecting this architectural pattern, a user connects to a client of the management view via a remote desktop connection.

Figure 20.8 Ensuring connectivity with instance for client



AWS Cloud
VPC
Public subnet
Management view client remote desktop terminal
Internet
Internet gateway
Instance for the management view client
Private subnet
Instance for the management view client (Cluster node)

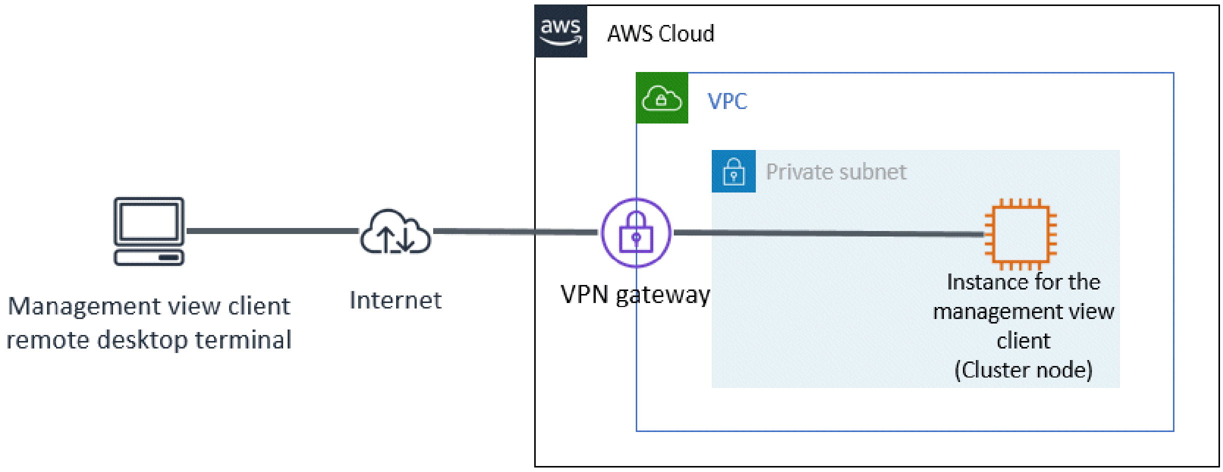
20.2.3.2 Ensuring Connectivity Using VPN Connection

In the pattern using a VPN connection, to ensure the connectivity between a terminal directly operated by a user and the management view, a VPN connection is used. No instances other than the management view are required in the VPC. This pattern is also secure because it does not use public subnets.

When selecting this architectural pattern, a terminal directly operated by a user is a client of the management view.

To provide a VPN connection, the VPN must be set or a device is required in the network where the terminal directly operated by a user is deployed.

Figure 20.9 Ensuring connectivity using VPN connection



AWS Cloud
VPC
Private subnet
Management view client remote desktop terminal
Internet
VPN gateway
Instance for the management view client (Cluster node)

20.3 Network Design

In the cluster system in an AWS environment, security rules to be applied to the subnet must be designed in advance.

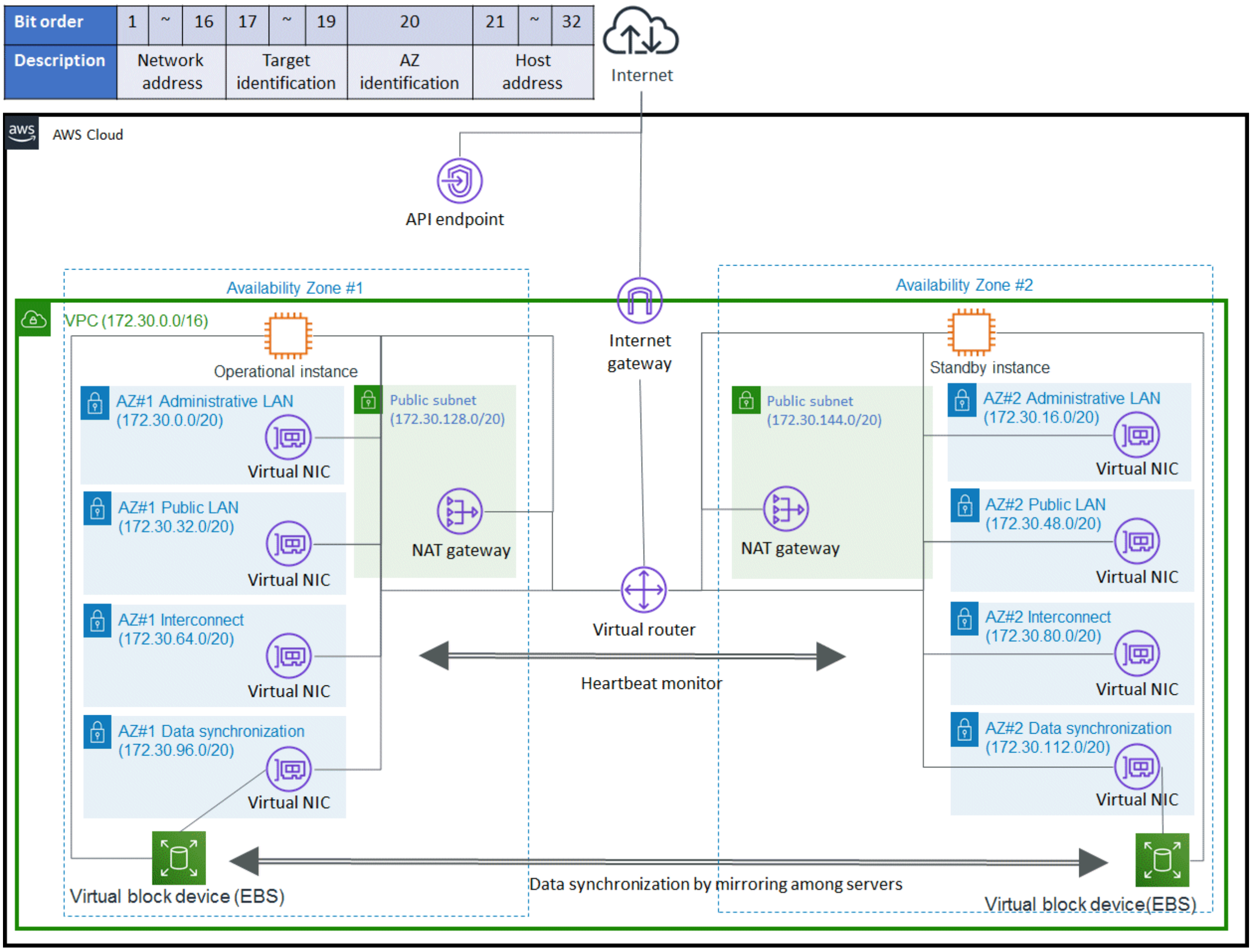
20.3.1 Subnet Design

20.3.1.1 Subnet Design of the Cluster System in Multiple Availability Zones (Multi-AZ)

For a cluster system in multiple Availability Zones (Multi-AZ), prepare the VPC and create a subnet for each purpose. Separate the subnet for each Availability Zone.

For easier access control, it is recommended that you use the upper bits of the subnet for role identification of the network and use the lower bits for identification of the Availability Zones.

Figure 20.10 Subnet design of the cluster system in multiple Availability Zones (Multi-AZ)



Bit order	1	~	16	17	~	19	20	21	~	32
Description	Network address			Target identification			AZ identification	Host address		

Internet
AWS Cloud
API endpoint
Availability Zone #1
Availability Zone #2
VPC (172.30.0.0/16)
Operational instance
Internet gateway
Standby instance
AZ#1 Administrative LAN (172.30.0.0/20)
Virtual NIC
Public subnet (172.30.128.0/20)
Public subnet (172.30.144.0/20)
AZ#2 Administrative LAN (172.30.16.0/20)
Virtual NIC
AZ#1 Public LAN (172.30.32.0/20)
Virtual NIC
NAT gateway
NAT gateway
AZ#2 Public LAN (172.30.48.0/20)
Virtual NIC
Virtual router
AZ#1 Interconnect (172.30.64.0/20)
Virtual NIC
AZ#2 Interconnect (172.30.80.0/20)
Virtual NIC
Heartbeat monitor
AZ#1 Data synchronization (172.30.96.0/20)
Virtual NIC
AZ#2 Data synchronization (172.30.112.0/20)
Virtual NIC
Virtual block device (EBS)
Data synchronization by mirroring among servers
Virtual block device(EBS)

The following is the procedure to create a private subnet in multiple Availability Zones (Multi-AZ).

1. Prepare the VPC to deploy the system.
2. Allocate CIDR to the VPC. (172.30.0.0/16 in the figure above)

   Select CIDR class according to the network size.
3. In the VPC, prepare the subnet for each purpose.

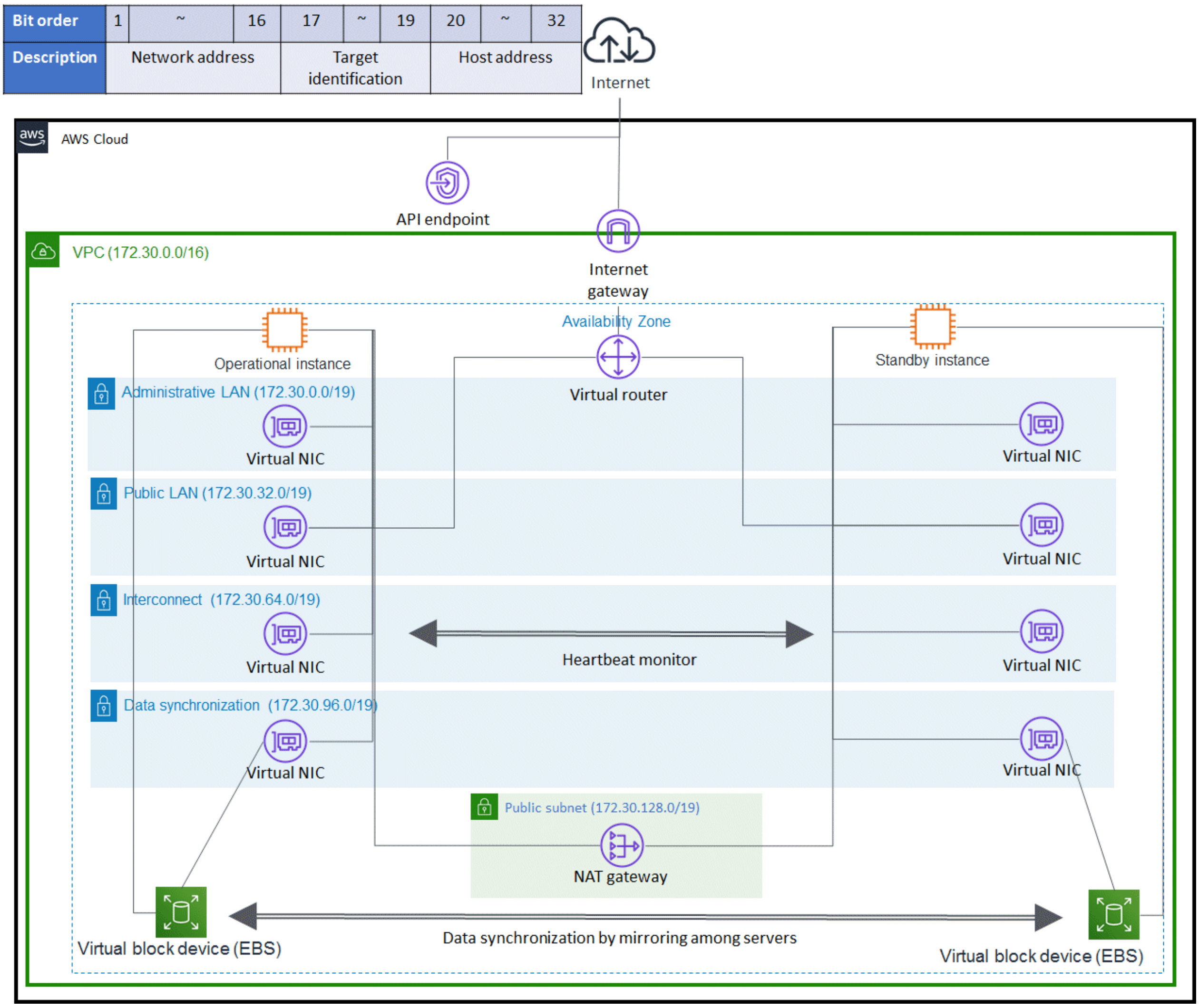
   Prepare the administrative LAN, the public LAN, the cluster interconnect, and the subnet for data synchronization for each Availability Zone. The network for data synchronization is required only when using mirroring among the servers.

   For the prefix length of the subnet, select the appropriate value according to the network size. (The prefix length is 20 in the figure above.)

20.3.1.2 Subnet Design of the Cluster System in a Single Availability Zone (Single-AZ)

For the single Availability Zone (Single-AZ), prepare the VPC and create a subnet for each purpose.

Figure 20.11 Subnet design of the cluster system in a single Availability Zone (Single-AZ)



Bit order	1	~	16	17	~	19	20	~	32
Description	Network address			Target identification			Host address		

Internet
AWS Cloud
API endpoint
VPC (172.30.0.0/16)
Internet gateway
Availability Zone
Operational instance
Standby instance
Virtual router
Administrative LAN (172.30.0.0/19)
Virtual NIC
Virtual NIC
Public LAN (172.30.32.0/19)
Virtual NIC
Virtual NIC
Interconnect (172.30.64.0/19)
Virtual NIC
Heartbeat monitor
Virtual NIC
Data synchronization (172.30.96.0/19)
Virtual NIC
Virtual NIC
Public subnet (172.30.128.0/19)
NAT gateway
Virtual block device (EBS)
Data synchronization by mirroring among servers
Virtual block device (EBS)

The following is the procedure to create a private subnet in a single Availability Zone (Single-AZ).

1. Prepare one VPC to deploy the system.
2. Allocate CIDR to the VPC. (172.30.0.0/16 in the figure above)

   Select CIDR class according to the network size.
3. In the VPC, prepare the subnet for each purpose.

   Prepare the administrative LAN, the public LAN, the cluster interconnect, and the subnet for data synchronization for each Availability Zone. The network for data synchronization is required only when using the mirroring among the servers.

   For the prefix length of the subnet, select the appropriate value according to the network size. (The prefix length is 19 in the figure above.)

   Depending on the selected architectural pattern, additional public subnets may be required.

20.3.2 Security Groups Design

This section describes the security group rule settings that are required to allow communication within the cluster.

PRIMECLUSTER uses several protocols/ports for communication within the cluster. By setting the rules described in this section, allow communication of protocols/ports for communication within the cluster.

In addition to the rules described in this section, you can add rules based on the security requirements of the customer to design security groups.

20.3.2.1 Rules Applied to the Administrative LAN

Design the security rules applied to the administrative LAN.

Inbound rule

Communication source CIDR	Protocol	Port range	Description
Own group	udp	9382	Used for the shutdown facility (SF)
Own group	udp	9796	Used for the management view
Own group	tcp	9797	Used for the management view
Own group	icmp	0-65535	Used for clchkcluster

Outbound rule

Communication target CIDR	Protocol	Port range	Description
0.0.0.0/0	tcp	443	Used for forced stop and network switching
0.0.0.0/0	tcp	53	Used for forced stop and network switching
Own group	icmp	0-65535	Used for clchkcluster
Own group	udp	9382	Used for the shutdown facility (SF)
Own group	udp	9796	Used for the management view
Own group	tcp	9797	Used for the management view

20.3.2.1.1 Creating Security Groups for Web-Based Admin View

Create the security groups for the Web-Based Admin View with the following setting values.

1) When ensuring the connectivity with an instance for a client

Create the security groups for the Web-Based Admin View (cluster node side) with the following setting values.

Inbound rule

Communication source CIDR	Protocol	Port range	Description
Own group	tcp	8081	Used for the management view
Own group	tcp	9798	Used for the management view
Own group	tcp	9799	Used for the management view

Create the security groups for the Web-Based Admin View (management client side) with the following setting values.

Outbound rule

Communication target CIDR	Protocol	Port range	Description
Own group	tcp	8081	Used for the management view
Own group	tcp	9798	Used for the management view
Own group	tcp	9799	Used for the management view

Also, create an inbound rule of the security group to allow a remote desktop connection from a remote control terminal of the management view client to an instance for the management view client.

2) When ensuring the connectivity using a VPN connection

Create the security groups for the Web-Based Admin View (cluster node side) with the following setting values.

Inbound rule

Communication source CIDR	Protocol	Port range	Description
CIDR of the management view client	tcp	8081	Used for the management view
CIDR of the management view client	tcp	9798	Used for the management view
CIDR of the management view client	tcp	9799	Used for the management view

Create the security groups for the Web-Based Admin View (management client side) with the following setting values.

Outbound rule

Communication target CIDR	Protocol	Port range	Description
CIDR of an instance for the management view (cluster node)	tcp	8081	Used for the management view
CIDR of an instance for the management view (cluster node)	tcp	9798	Used for the management view
CIDR of an instance for the management view (cluster node)	tcp	9799	Used for the management view

20.3.2.1.2 Rules Applied to Instance Access in Introduction and Maintenance

Design the security rules applied to instance access in introduction and maintenance.

Inbound rule

Communication source CIDR	Protocol	Port range	Description
CIDR of the access source	tcp	22	Used for the remote access by SSH

Outbound rule

Communication target CIDR	Protocol	Port range	Description
0.0.0.0/0	tcp	80	Used for installing dependent packages
0.0.0.0/0	tcp	443	Used for installing dependent packages

20.3.2.2 Rules Applied to the Cluster Interconnect

Design the security rules applied to the cluster interconnect.

This setting is not necessary in a single-node cluster.

Inbound rule

Communication source CIDR	Protocol	Port range	Description
Own group	123	0-65535	Used for the heartbeat

Outbound rule

Communication target CIDR	Protocol	Port range	Description
Own group	123	0-65535	Used for the heartbeat

20.3.2.3 Rules Applied to the Public LAN

Design the security rules applied to the public LAN.

Add the rules that are required for application operations and the following outbound rule.

Outbound rule

Communication target CIDR	Protocol	Port range	Description
CIDR of the monitoring destination of the business network	icmp	0-65535	Used for monitoring the business network

This rule is required when using the network monitoring function.

For details, refer to "6.7.3.6 Setting Up Takeover Network Resources" in "PRIMECLUSTER Installation and Administration Guide."

When monitoring a virtual router, the monitored virtual router and the public LAN of all cluster nodes must be on the same subnet.

20.3.2.4 Rules Applied to the Network for Data Synchronization

Set the security rules applied to the network for data synchronization.

Inbound rule

Communication source CIDR	Protocol	Port range	Description
Own group	tcp	3260	Used for mirroring among servers

Outbound rule

Communication target CIDR	Protocol	Port range	Description
Own group	tcp	3260	Used for mirroring among servers

20.4 System Design

Use PRIMECLUSTER Designsheets to design the system.

The installation operation of the PRIMECLUSTER system is performed based on the created PRIMECLUSTER Designsheets. Make sure to create the designsheets and confirm that all required items are described.

20.5 Determining the Cluster System Operation Mode

In the cluster system in an AWS environment, operation modes for 1:1 standby operation, mutual standby, and single-node cluster operation can be built.

For details on the operation mode of each cluster system, refer to "2.3 Determining the Cluster System Operation Mode" in "PRIMECLUSTER Installation and Administration Guide."

20.6 Determining the Web-Based Admin View Operation Mode

For details on the Web-Based Admin View operation mode, refer to "2.4 Determining the Web-Based Admin View Operation Mode" in "PRIMECLUSTER Installation and Administration Guide."

20.7 Forcible Stop Method

In the cluster system in an AWS environment, the asynchronous forcible stop method is used.

In this method, a node where an error occurs is forcibly stopped and the switching is started before the stop of a node is completed. This means that the switching time can be shorten and that a business does not stop in the LEFTCLUSTER state due to the incomplete stop of a node.

Also by using a function of the mirroring among servers, data synchronization communication is blocked at the time of the switching. This prevents simultaneous access to the business data from multiple nodes, and the data is secured.

When the cluster partition occurs, each node operates as follows according to the survival priority.

- Node with high survival priority
    1. Forcibly stopping a remote node by the panic instruction (in an AWS Nitro System environment) and the power-off instruction
    2. Operating a local node without waiting for the stop of a forcibly stopped node
- Node with low survival priority
    1. Waiting for the time until the forcible stop process of a node with high survival priority is completed
    2. Checking the state of the instance of a local node and panicking the local node

Note

- When an OS panic occurs, the cluster node may be powered off while the memory dump is being output, and it may not be possible to collect a complete memory dump.
- The slice preceding degenerated option of the mirroring among servers cannot be set.
- When an OS panic occurs, the OS must be restarted manually since the OS is not restarted automatically.
- A shared file system service cannot be used. The mirroring among servers must be used when taking over the volume data.

20.8 Determining the Failover Timing of Cluster Application

Determine the failover timing of cluster application.

For details, refer to "2.5 Determining the Failover Timing of Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

20.9 Policy Design

You can grant access permissions to the AWS CLI using an IAM role or an IAM user.

IAM roles grant access and operation permissions to AWS resources. Since access keys for IAM users are not saved in each server, access controls are secured.

When you cannot use IAM roles, you can attach policies to IAM users to grant access permissions to AWS resources.

See

For details on policies and IAM roles, refer to the official AWS documentation.

As a policy to attach to an IAM role or an IAM user, based on the architectural pattern selected in "20.2.1 Network Takeover", design a policy that grants access permissions to the following actions.

Network takeover by the virtual router

ec2:DescribeInstances
ec2:DescribeInstanceStatus
ec2:SendDiagnosticInterrupt
ec2:StopInstances
ec2:DescribeRouteTables
ec2:CreateRoute
ec2:ReplaceRoute
ec2:DescribeNetworkInterfaces

Network takeover by replacing the Elastic IP address

ec2:DescribeInstances
ec2:DescribeInstanceStatus
ec2:SendDiagnosticInterrupt
ec2:StopInstances
ec2:AssociateAddress
ec2:DescribeAddresses
ec2:DescribeNetworkInterfaces

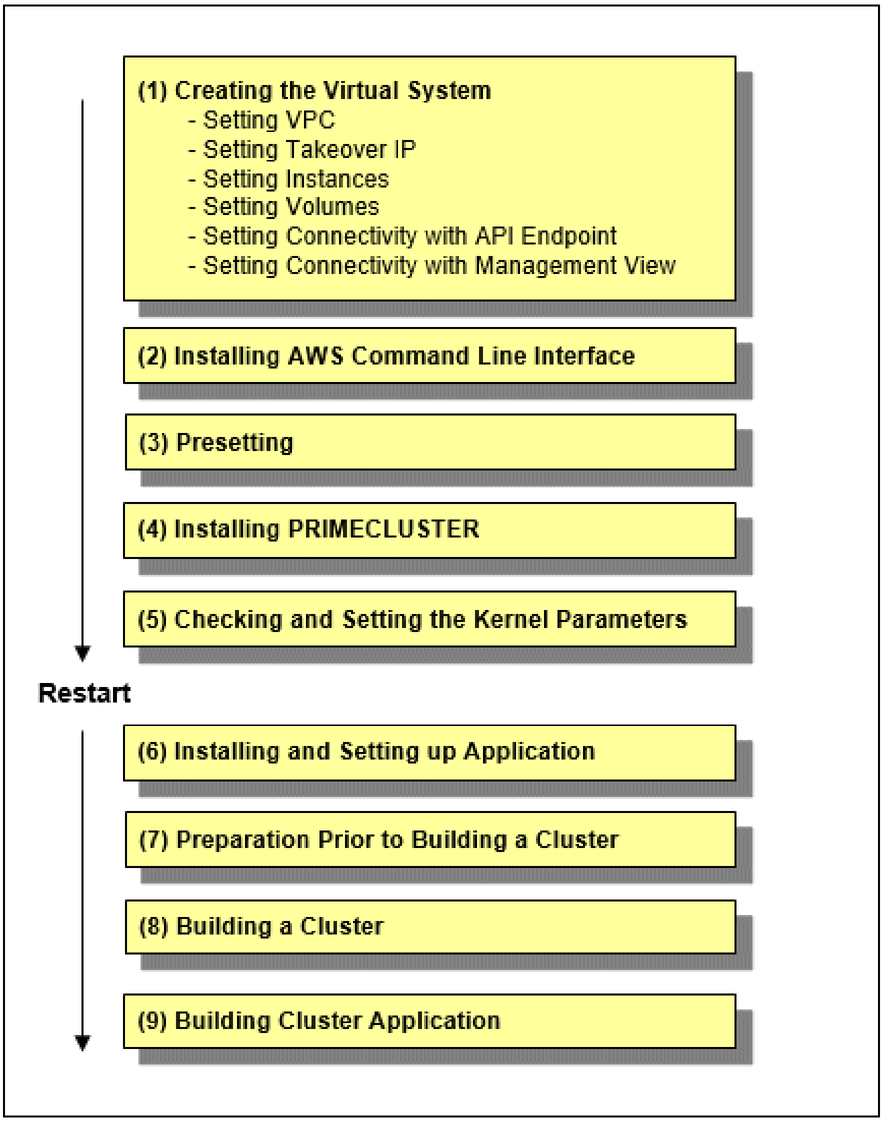
Network takeover by rewriting DNS records

ec2:DescribeInstances
ec2:DescribeInstanceStatus
ec2:SendDiagnosticInterrupt
ec2:StopInstances
route53:ChangeResourceRecordSets
route53:GetChange
route53:ListResourceRecordSets
route53:GetHostedZone

Chapter 21 Installation

This chapter describes the procedure to install PRIMECLUSTER in an AWS environment.

Perform the steps shown in the figure below.


(1) Creating the Virtual System
- Setting VPC
- Setting Takeover IP
- Setting Instances
- Setting Volumes
- Setting Connectivity with API Endpoint
- Setting Connectivity with Management View
(2) Installing AWS Command Line Interface
(3) Presetting
(4) Installing PRIMECLUSTER
(5) Checking and Setting the Kernel Parameters
Restart
(6) Installing and Setting up Application
(7) Preparation Prior to Building a Cluster
(8) Building a Cluster
(9) Building Cluster Application

Note

For how to check the operation environment, refer to "Operation Environment" in the Installation Guide for PRIMECLUSTER.

21.1 Creating the Virtual System

This section describes how to create the virtual system for a cluster system in an AWS environment.

By using the AWS Management Console, create the virtual system that was designed in "Chapter 20 Design." The creation procedure depends on the architectural pattern selected.

See

To create a virtual system or for details on each setting, refer to the official AWS documentation.

21.1.1 Setting VPC

Create the VPC where the cluster system is to be deployed. Then, create the subnet and set the security group for the created VPC according to the design created in "20.3 Network Design."

21.1.2 Setting Network Takeover

Create resources required for taking over the network.

The required resources depend on the architectural pattern selected for the network takeover in "20.2.1 Network Takeover."

When selecting network takeover by the virtual router

Create the route table.

In this case, you do not need to add a routing entry to take over the network.

When selecting network takeover by replacing the Elastic IP address

Allocate an Elastic IP address for takeover.

When selecting network takeover by rewriting DNS records

Create a VPN gateway, customer gateway, VPN connection, and private host zone.

21.1.3 Setting Instances

Create the instance and the virtual network interface that configure the cluster node.

When creating the virtual network interface, select the appropriate subnet based on the architectural pattern selected in "20.2.1 Network Takeover."

When selecting network takeover by the virtual router, uncheck the source/destination change.

When using an IAM role, refer to "20.9 Policy Design" and create a policy, and then attach the policy to the EC2 use case to create a role. After creating the role, associate the IAM role with the instances that use PRIMECLUSTER.

When using an IAM user, make sure that the policy described in "20.9 Policy Design" is attached to the IAM user.

If the policy is not attached, attach it.

When creating multiple virtual network interfaces and using a virtual network interface, for which a default gateway is not set, to communicate with interfaces on different subnets, static routing must be set for that virtual network interface. You do not need to set static routing when communicating only with interfaces on the same subnet.

Example

The following is an example of static routing configuration when using a network interface (eth1) different from the administrative LAN (eth0) as the cluster interconnect in the cluster system with a two-node configuration (CF node names are cluster node 1 and cluster node 2).

If the IP address for eth1 of the cluster node 1 is 172.30.64.4/20 and the IP address for eth1 of the cluster node 2 is 172.30.80.4/20, make the following settings on each node.

Cluster node 1

1. Check the CONNECTION of eth1.

```
# nmcli device status
DEVICE  TYPE      STATE      CONNECTION
eth0    ethernet  connected  System eth0
eth1    ethernet  connected  Wired connection 1
```

```
eth2    ethernet  connected  Wired connection 2
eth3    ethernet  connected  Wired connection 3
lo      loopback  unmanaged  --
```

2. Add a route to the CONNECTION confirmed in step 1.

```
# nmcli connection modify "Wired connection 1" +ipv4.routes "172.30.80.0/20 172.30.64.1
```

3. Make sure that the route has been successfully added to the CONNECTION.

```
# nmcli -g ipv4.routes connection show "Wired connection 1"
172.30.80.0/20 172.30.64.1
```

4. Restart the network connection.

```
# nmcli connection down "Wired connection 1"
# nmcli connection up "Wired connection 1"
```

5. Make sure that the route set in step 2 has been added.

```
# ip route show
172.30.80.0/20 via 172.30.64.1 dev eth1 proto static metric 113
```

6. Restart the OS.

Cluster node 2

1. Check the CONNECTION of eth1.

```
# nmcli device status
DEVICE  TYPE      STATE      CONNECTION
eth0    ethernet  connected  System eth0
eth1    ethernet  connected  Wired connection 1
eth2    ethernet  connected  Wired connection 2
eth3    ethernet  connected  Wired connection 3
lo      loopback  unmanaged  --
```

2. Add a route to the CONNECTION confirmed in step 1.

```
# nmcli connection modify "Wired connection 1" +ipv4.routes "172.30.64.0/20 172.30.80.1
```

3. Make sure that the route has been successfully added to the CONNECTION.

```
# nmcli -g ipv4.routes connection show "Wired connection 1"
172.30.64.0/20 172.30.80.1
```

4. Restart the network connection.

```
# nmcli connection down "Wired connection 1"
# nmcli connection up "Wired connection 1"
```

5. Make sure that the route set in step 2 has been added.

```
# ip route show
172.30.64.0/20 via 172.30.80.1 dev eth1 proto static metric 113
```

6. Restart the OS.

See

For details on how to configure the static routing, refer to the official OS documentation.

21.1.4 Setting Data Storage Area

When using the mirroring among servers, set the data storage area used by an application.

Create the virtual block device managed by the mirroring among servers, then attach it to the instance.

If you are not using the mirroring among servers but configuring the data storage area by shared file system services or RDB services provided by AWS, this setting is not necessary.

21.1.5 Setting Connectivity with API Endpoint

Create the components required for configuring the connectivity with the API endpoint. The required components depend on the architectural pattern.

When selecting a NAT gateway to ensure connectivity

1. Create the NAT gateway.
2. Set entries to the route tables in the public subnet and the private subnet.

   For information on setting route tables, refer to "20.2.2.1 Ensuring Connectivity with NAT Gateway."

When selecting a NAT instance to ensure connectivity

1. Create the NAT instance.
2. Set entries to the route tables in the public subnet and the private subnet.

   For information on setting route tables, refer to "20.2.2.2 Ensuring Connectivity with NAT Instance."

When selecting an Elastic IP address to ensure connectivity

Allocate an Elastic IP address for the administrative LAN.

When selecting PrivateLink to ensure connectivity

Create the VPC endpoint.

21.1.6 Setting Connectivity with Management View

Set the components required for ensuring the connectivity between the management terminal and the management view on the cluster node. The required components depend on the architectural pattern selected in "20.6 Determining the Web-Based Admin View Operation Mode."

When selecting an instance for the client to ensure connectivity

Create the public subnet where the instance for the client to be deployed, and create the instance for the client.

When selecting a VPN connection to ensure connectivity

Create a VPN gateway, customer gateway, and VPN connection.

21.2 Installing the AWS Command Line Interface

PRIMECLUSTER uses the AWS Command Line Interface to solve split brain or to take over the network.

On all nodes that configure the cluster, install the AWS Command Line Interface for the root user (/root/.local/bin), not in /usr/local/aws, /usr/local/bin. After the installation, verify the path to the AWS Command Line Interface.

Note

The AWS Command Line Interface must be version 1 (1.16 or later).

See

For details on the installation procedure, refer to the official AWS documentation.

21.3 Presetting

Take the following procedure on all nodes.

1. Disable the firewall.

   Make sure that "firewalld" is disabled.

   ```
   # systemctl is-enabled firewalld
   ```

   If it is enabled, disable it.

   ```
   # systemctl stop firewalld
   # systemctl disable firewalld
   ```

2. Set NTP.

   Make sure to set NTP when building the cluster to synchronize the time of each node in the cluster system.

   Set NTP before installing PRIMECLUSTER.

3. Set the credentials of the AWS Command Line Interface.

   Set the credentials with the root user with the "aws configure" command.

   If the named profile is specified when the settings are configured, make a note of the profile name that is required for setting the shutdown facility and setting the network takeover function.

   Example) The following are examples when setting the credentials for the AWS Command Line Interface.

   - When "userprofile1" is used for the profile name

     ```
     # aws configure --profile userprofile1
     ```

   - When the default profile is used

     ```
     # aws configure
     ```

   Specify "json" for the Default output format. Also, when using an IAM role for access permissions, leave the AWS Access Key ID and the AWS Secret Access Key blank (press the Enter key only).

4. Disable the nm-cloud-setup Service.

   Check the current auto-start configuration of the nm-cloud-setup service.

   ```
   # systemctl is-enabled nm-cloud-setup.timer
   # systemctl is-enabled nm-cloud-setup.service
   ```

   If the auto-start setting is enabled, disable it.

   ```
   # systemctl stop nm-cloud-setup.timer
   # systemctl disable nm-cloud-setup.timer
   # systemctl stop nm-cloud-setup.service
   # systemctl disable nm-cloud-setup.service
   ```

   See

   For details on setting the credentials of the AWS Command Line Interface, refer to the official AWS documentation.

21.4 Installing PRIMECLUSTER

Use the installation script (CLI Installer) to install PRIMECLUSTER.

Install PRIMECLUSTER on each node in the system where Linux(R) software and Linux(R) related software are already installed. Use the same installation script when installing PRIMECLUSTER on the cluster management server.

Note

If the OS has never been restarted since the instance was created, restart it and then install PRIMECLUSTER.

See

For details on the installation and uninstallation procedures, refer to the descriptions of cloud environments described in the Installation Guide for PRIMECLUSTER.

21.5 Checking and Setting the Kernel Parameters

Change the kernel parameters depending on the environment.

Applicable nodes:

All nodes on which PRIMECLUSTER is to be installed

Depending on the products and components utilized, different kernel parameters are required.

Check PRIMECLUSTER Designsheets and if you need to modify the kernel parameters, set them again.

See

For details on kernel parameters, refer to "3.1.7 Checking and Setting the Kernel Parameters" in "PRIMECLUSTER Installation and Administration Guide."

Note

- To activate the modified kernel parameters, restart the OS.
- After uninstalling PRIMECLUSTER, change the kernel parameter settings back to the state before installing PRIMECLUSTER if necessary.

Set the following kernel parameters.

This setting is not necessary in a single-node cluster.

Parameter	Value	Remarks: meaning of (parameter)
kernel.panic	0	Seconds to wait until the kernel is restarted in case of a panic. If 0 is set, the kernel is not restarted.
kernel.unknown_nmi_panic	1	A panic occurs in the NMI interrupt.
kernel.sysrq	Other than 0	The SysRq key is enabled.

21.6 Setting up kdump

Set up the following on all nodes when using the kdump (recommended).

1. Set up the kdump parameter.

   In a two-node configuration, to power off the cluster node after the memory dump is output, set the following parameters in the /etc/kdump.conf.

   For details on the setup procedure of the kdump, refer to the OS manual.

Parameter	Value	Remarks: meaning of (parameter)
kdump_post	/opt/SMAW/SMAWsf/bin/poff.sh	The power-off script (poff.sh) is executed after the memory dump output ends.
default	poweroff	The power is turned off when the memory dump output fails.

2. Check the kdump.

   Check if the kdump server function is enabled. If not, enable the kdump.

   - Check the availability of the kdump with the systemctl(1) command.

     Example) The kdump is disabled if the state is as follows.

```
# /usr/bin/systemctl list-unit-files --type=service | grep  kdump.service
kdump.service                               disabled
```

   - If the kdump is enabled, restart the kdump with the systemctl(1) command.

```
# /usr/bin/systemctl restart kdump.service
```

   - If the kdump is disabled, enable it with the systemctl(1) command, and then start the kdump.

```
# /usr/bin/systemctl enable kdump.service
# /usr/bin/systemctl start kdump.service
```

Note

After uninstalling PRIMECLUSTER, make sure to change the kdump settings back to the state before installing PRIMECLUSTER.

21.7 Installing and Setting up Application

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

See

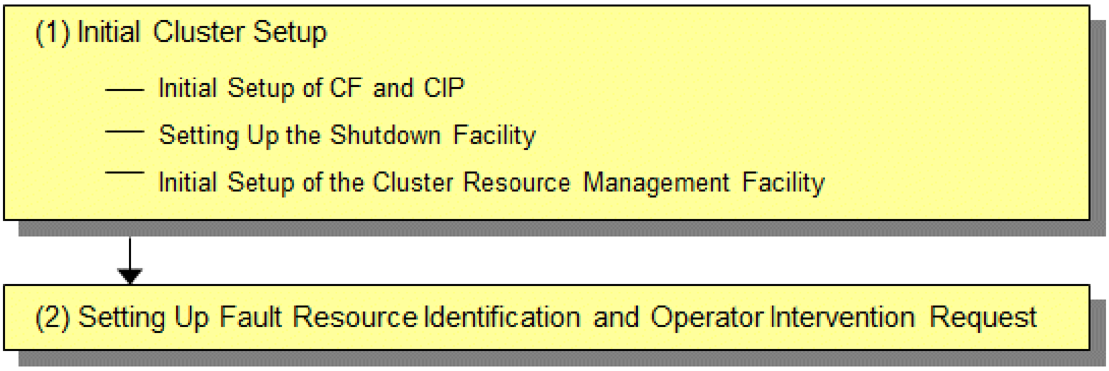
- For details on environment setup, refer to the manuals for each application.
- For information on PRIMECLUSTER-related products that support AWS, refer to the documentation for each product.

21.8 Presettings for Building a Cluster

Prior to building a cluster, perform presettings such as staring the Web-Based Admin View screen. For details on presettings prior to building a cluster, refer to "Chapter 4 Preparation Prior to Building a Cluster" in "PRIMECLUSTER Installation and Administration Guide."

21.9 Building a Cluster

The procedure for building a PRIMECLUSTER cluster is shown below:


(1) Initial Cluster Setup
Initial Setup of 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

21.9.1 Initial Cluster Setup

This section describes the initial cluster setup for PRIMECLUSTER.

21.9.1.1 Initial Setup of CF and CIP

Refer to "5.1.1 Setting Up CF and CIP" in "PRIMECLUSTER Installation and Administration Guide" to set up CF and CIP.

Note

You need to use CF over IP in a cloud environment. In addition, you need to configure CIP because such as RMS uses it for inter-node communication in the cluster. CF over IP (IP interconnect) and CIP are different features. For more information, see "1.1.1 Differences between CIP and CF over IP" in the Cluster Foundation Configuration and Administration Guide.

21.9.1.2 Setting up the Shutdown Facility

This section describes how to set up the shutdown facility in an AWS environment.

The shutdown agent available in an AWS environment is as follows.

- AWS CLI (SA_vmawsAsyncReset)

  The shutdown function of a node (instance) using the AWS Command Line Interface is provided.

  This setting is not necessary in a single-node cluster.

  The storage location of a log file is as follows.

```
/var/opt/SMAWsf/log/SA_vmawsAsyncReset.log
```

For details on the survival priority, refer to "5.1.2.1 Survival Priority" in "PRIMECLUSTER Installation and Administration Guide."

21.9.1.2.1 Setup Procedure of the Shutdown Facility for the Asynchronous Forcible Stop Method

This section describes the setup procedure of the shutdown facility for the asynchronous forcible stop method.

Perform the following procedure.

Note

- After setting up the shutdown agent, conduct a test for the forced stop of cluster nodes to make sure that the correct nodes can be forcibly stopped. For details of the test for the forced stop of cluster nodes, refer to "1.4 Test" in "PRIMECLUSTER Installation and Administration Guide."

- The contents of the SA_vmawsAsyncReset.cfg file and the rcsd.cfg file of all nodes should be identical. If not, a malfunction will occur.

1. Set up the shutdown daemon.

   Create SA_vmawsAsyncReset.cfg with the following contents on all nodes in the cluster system.

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

```
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 used
                       in the shutdown facility of the cluster host.
                       Available IP addresses are IPv4.
                       When specifying a host name, make sure it is described in /etc/hosts.
SA_vmawsAsyncReset   : AWS CLI shutdown agent.
timeout              : Specify the timeout duration (seconds) of the AWS CLI shutdown agent.
                       Specify 15 seconds.
```

   Example) The following is a setup example.

   If the CF node names of the cluster host are node1 and node2, the weight of two nodes is 1, the IP address of the administrative LAN of node1 is 192.168.250.1, and the IP address of the administrative LAN of node2 is 192.168.250.2.

```
# cat /etc/opt/SMAW/SMAWsf/rcsd.cfg
node1,weight=1,admIP=192.168.250.1:agent=SA_vmawsAsyncReset,timeout=15
node2,weight=1,admIP=192.168.250.2:agent=SA_vmawsAsyncReset,timeout=15
```

   Create /etc/opt/SMAW/SMAWsf/rcsd.cfg and then set the owner, group, and access rights as follows.

```
# chown root:root /etc/opt/SMAW/SMAWsf/rcsd.cfg
# chmod 600 /etc/opt/SMAW/SMAWsf/rcsd.cfg
```

   Information

   When creating the /etc/opt/SMAW/SMAWsf/rcsd.cfg file, the /etc/opt/SMAW/SMAWsf/rcsd.cfg.template file can be used as a template.

2. Set up the shutdown agent.

   Create /etc/opt/SMAW/SMAWsf/SA_vmawsAsyncReset.cfg with the following contents on all nodes in the cluster system.

   Information

   The template of the SA_vmawsAsyncReset.cfg file can be found at the following location:

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

   Delimit each item with a single space.

```
CFNameX InstanceID [ProfileName]
CFNameX InstanceID [ProfileName]
```

```
CFNameX            : Specify the CF node name of the cluster host.
InstanceID         : Specify the instance ID of AWS on which the cluster host is operating.
ProfileName        : Specify the profile name of the credentials used with the AWS Command Line
                     Interface configured in "21.3 Presetting."
                     When this item is omitted, the shutdown agent operates with the default
                     profile.
```

   Example) The following is a setup example.

   If the CF node names of the cluster host are node1 and node2, the instance IDs of AWS are i-abcdef0123456789a and i-abcdef0123456789b, and the profile name is userprofile1.

```
# cat /etc/opt/SMAW/SMAWsf/SA_vmawsAsyncReset.cfg
node1 i-abcdef0123456789a userprofile1
node2 i-abcdef0123456789b userprofile1
```

Create /etc/opt/SMAW/SMAWsf/SA_vmawsAsyncReset.cfg and then set the owner, group, and access rights as follows.

```
# chown root:root /etc/opt/SMAW/SMAWsf/SA_vmawsAsyncReset.cfg
# chmod 600 /etc/opt/SMAW/SMAWsf/SA_vmawsAsyncReset.cfg
```

Note

- Make sure that the /etc/opt/SMAW/SMAWsf/SA_vmawsAsyncReset.cfg file is set correctly. If the setting is incorrect, the shutdown facility cannot be performed normally.

- Make sure that the instance ID of AWS (*InstanceID*) and the profile name (*ProfileName*) corresponding to the CF node name (*CFNameX*) of the cluster host of the /etc/opt/SMAW/SMAWsf/SA_vmawsAsyncReset.cfg file are set. If the setting is incorrect, an incorrect node will be forcibly stopped.

3. Start the shutdown facility.

   Check if the shutdown facility has been started on all nodes in the cluster system.

   ```
   # sdtool -s
   ```

   On a node where the shutdown facility has already been started, execute the following commands to restart the shutdown facility.

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

   On a node where the shutdown facility has not been started, execute the following command to start the shutdown facility.

   ```
   # sdtool -b
   ```

Information

You can check if the shutdown facility has already been started with the sdtool -s command. If "The RCSD is not running" is displayed, the shutdown facility is not started.

4. Check the status of the shutdown facility.

   Execute the following command on all nodes in the cluster system to check the status of the shutdown facility.

   ```
   # sdtool -s
   ```

Note

If "The RCSD is not running" is displayed, the setting of the shutdown daemon or the setting of the shutdown agent is not correct. Perform the procedure from step 1 to 3 again.

Information

Display results of the sdtool -s command

- If Unknown or Init-ing is displayed in Init State, wait for about one minute, and then check the status again.

- If Unknown is displayed in Shut State, it means that SF has not yet stopped the node. If Unknown is displayed in Init State, it means that SF has not yet initialized SA or tested the route. Unknown is displayed temporarily in Test State or Init State until the actual state can be confirmed.

- If TestFailed is displayed in Test State, 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.

21.9.1.3 Initial Setup of the Cluster Resource Management Facility

Refer to "5.1.3 Initial Setup of the Cluster Resource Management Facility" in "PRIMECLUSTER Installation and Administration Guide" to set up the resource database managed by the cluster resource management facility. In this setting, set the iSCSI device used in the mirroring among the servers of GDS and register it to the resource database.

21.9.2 Setting up Fault Resource Identification and Operator Intervention Request

Refer to "5.2 Setting up Fault Resource Identification and Operator Intervention Request" in "PRIMECLUSTER Installation and Administration Guide" to set up the fault resource identification and the operator intervention request.

21.10 Building Cluster Application

For details on how to build the cluster application, refer to "Chapter 6 Building Cluster Applications" in "PRIMECLUSTER Installation and Administration Guide."

When using the mirroring among servers, set the mirroring among the servers of GDS (creating netmirror volume) while building the cluster application.

Also, depending on the architectural pattern selected for network takeover in "20.2.1 Network Takeover", network takeover needs to be registered in the cluster application. To build a cluster application for network takeover, refer to "21.10.1 Building Cluster Application for Network Takeover."

Note

- Note the following points when performing the settings in "Setting Tuning Parameters" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

  - Among the tuning parameters to be set, set the following values for the following tuning parameters.

Tuning parameter name	Value after change
ED_CMD_RETRY_COUNT	100
ED_DRV_RETRY_COUNT	100

Example:

```
ED_CMD_RETRY_COUNT=100
ED_DRV_RETRY_COUNT=100
```

To extend the timeout period (CLUSTER_TIMEOUT) of the CF heartbeat, change the above parameter values according to the following formula. Round up the values after the decimal point.

Calculation formula:

```
<Increased CLUSTER_TIMEOUT> / 3 + 100
```

  - When using the asynchronous forcible stop method, set the following tuning parameter in /etc/opt/FJSVsdx/sdx.cf.

```
SDX_NETMIRROR_IO_BLOCKADE=1
```

  - When using the asynchronous forcible stop method, comment out the following tuning parameter in the /etc/opt/FJSVsdx/modules/sfdsk.conf and disable the slice preceding degenerated option.

```
SDX_NETMIRROR_PRE_DETACH=1;
```

Example:

```
#SDX_NETMIRROR_PRE_DETACH=1;
```

- Note the following setting in "Checking and Setting Required Packages" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

  - When using the asynchronous forcible stop method, add "smawcf.service" to the value "After" in /etc/systemd/system/fjsvsdx.service.d/netmirror.conf.

    Example:

    ```
    After=target.service smawcf.service
    ```

- If the icmp communication between cluster nodes is not allowed in the security group configuration, the following message is displayed when the clchkcluster command is executed.

  ```
  Admin IP <IP address> used by SF is not alive.
  ```

  If this message is output, refer to "20.3.2.1 Rules Applied to the Administrative LAN", and set the icmp protocol rule to allow the icmp communication between cluster nodes. After that, execute the clchkcluster command again.

- When using the Symfoware Server (Native) Hot Standby feature, the standby and scalable userApplication configurations are required.

  Set the standby userApplication attribute to the following value.

Attribute	Setup value
StandbyTransitions	ClearFaultRequest | SwitchRequest

  Refer to "6) Scalable operation" in "6.7 Setting Up Cluster Applications" in the "PRIMECLUSTER Installation and Administration Guide" for detailed build instructions.

21.10.1 Building Cluster Application for Network Takeover

This section describes how to create the definition file and build the cluster application to use the architectural pattern selected in "20.2.1 Network Takeover."

21.10.1.1 Creating Definition File

For network takeover, create the following definition file on all nodes controlling the network devices of AWS.

```
/usr/opt/reliant/etc/hvawsconfig
```

Information

When you create the /usr/opt/reliant/etc/hvawsconfig file, you can use the /usr/opt/reliant/etc/hvawsconfig.template file as a template.

```
# cp -p /usr/opt/reliant/etc/hvawsconfig.template /usr/opt/reliant/etc/hvawsconfig
```

Note

- Create the hvawsconfig file with a root user and change the permission to 600.

- The hvawsconfig file must be the same contents on each cluster node.

The contents of the definition file depend on the selected architectural pattern.

Definition file for network takeover by the virtual router

To update the routes of the virtual router, the ID information of the device managed by AWS is required.

Check each ID in AWS described in the definition file (/usr/opt/reliant/etc/hvawsconfig) with the AWS Management Console.

Information in the definition file

```
KeyName Mode CFNameX InstanceID RouteTableID TakeoverIPaddress ENIID [ProfileName]
```

Item	Contents	Remarks
KeyName	Describe the ID up to 16 letters and numbers. *KeyName* is case-sensitive. This key name is defined in each script that is registered when setting the Cmdline resource. Specify the same *KeyName* for the paired operational system and standby system and define a *KeyName* for other systems to avoid duplicate names. This key name is specified as an argument to register a script when setting Cmdline resources explained later.	-
Mode	Specify the architectural pattern for network takeover. For network takeover by the virtual router, specify the string of ROUTE or route.	-
CFNameX	Specify the CF node name.	-
InstanceID	Specify the instance ID of AWS on which the cluster host is operating.	Check with the AWS Management Console
RouteTableID	Specify the route table ID.	Check with the AWS Management Console
TakeoverIPaddress	Takeover IP address (Specify by IPv4 address.)	-
ENIID	Specify ENIID of the network interface that takes over the IP.	Check with the AWS Management Console
ProfileName	Specify the profile name of the credentials used with the AWS Command Line Interface configured in "21.3 Presetting." When it is omitted, the device operates with the default profile.	-

Example

The takeover IP address is 172.31.0.10, and the profile name of the credentials used with the AWS Command Line Interface is userprofile1.

```
CmdR01 ROUTE node1 i-xxxxxxxx rtb-xxxxxxxx 172.31.0.10 eni-xxxxxxxx userprofile1
CmdR01 ROUTE node2 i-yyyyyyyy rtb-xxxxxxxx 172.31.0.10 eni-yyyyyyyy userprofile1
```

For multiple controls such as the mutual standby configuration, separate *KeyNames* and add the same contents.

The following example shows how to control two Takeover IP addresses, 172.31.0.10 and 172.32.0.20.

```
CmdR01 ROUTE node1 i-xxxxxxxx rtb-xxxxxxxx 172.31.0.10 eni-xxxxxxxx userprofile1
CmdR01 ROUTE node2 i-yyyyyyyy rtb-xxxxxxxx 172.31.0.10 eni-yyyyyyyy userprofile1
CmdR02 ROUTE node1 i-xxxxxxxx rtb-xxxxxxxx 172.32.0.20 eni-aaaaaaaa userprofile1
CmdR02 ROUTE node2 i-yyyyyyyy rtb-xxxxxxxx 172.32.0.20 eni-bbbbbbbb userprofile1
```

Note

- Use one line per node, separated by spaces.
- For *InstanceID* and *ENII*D, enter the ID information of the operational system and the standby system.
- The takeover IP address must be specified as an IPv4 address, not a host name.

- If the settings are not correct, the routing information cannot be updated correctly, resulting in a resource failure.

Definition file for network takeover by replacing the Elastic IP address

To associate an Elastic IP with ENIID, the ID information of the device managed by AWS is required.

Check each ID in AWS described in the definition file (/usr/opt/reliant/etc/hvawsconfig) with the AWS Management Console.

Information in the definition file

```
KeyName Mode CFNameX InstanceID AllocationID ENIID [ProfileName]
```

Item	Contents	Remarks
KeyName	Describe the ID up to 16 letters and numbers. *KeyName* is case-sensitive. This key name is defined in each script that is registered when setting the Cmdline resource. Specify the same *KeyName* for the paired operational system and standby system and define a *KeyName* for other systems to avoid duplicate names. This key name is specified as an argument to register a script when setting Cmdline resources explained later.	-
Mode	Specify the architectural pattern for network takeover. For network takeover by replacing the Elastic IP address, specify the string of ELASTIC or elastic.	-
CFNameX	Specify the CF node name.	-
InstanceID	Specify the instance ID of AWS on which the cluster host is operating.	Check with the AWS Management Console
AllocationID	Allocation ID of the Elastic IP address	Check with the AWS Management Console
ENIID	Specify ENIID of the network interface that is allocated to the Elastic IP address.	Check with the AWS Management Console
ProfileName	Specify the profile name of the credentials used with the AWS Command Line Interface configured in "21.3 Presetting." When it is omitted, the device operates with the default profile.	-

Example

The profile name of the credentials used with the AWS Command Line Interface is userprofile1.

```
CmdE01 ELASTIC node1 i-xxxxxxxx eipalloc-xxxxxxxx eni-xxxxxxxx userprofile1
CmdE01 ELASTIC node2 i-yyyyyyyy eipalloc-xxxxxxxx eni-yyyyyyyy userprofile1
```

For multiple controls such as the mutual standby configuration, separate *KeyNames* and add the same contents.

```
CmdE01 ELASTIC node1 i-xxxxxxxx eipalloc-xxxxxxxx eni-xxxxxxxx userprofile1
CmdE01 ELASTIC node2 i-yyyyyyyy eipalloc-xxxxxxxx eni-yyyyyyyy userprofile1
CmdE02 ELASTIC node1 i-xxxxxxxx eipalloc-xxxxxxxx eni-aaaaaaaa userprofile1
CmdE02 ELASTIC node2 i-yyyyyyyy eipalloc-xxxxxxxx eni-bbbbbbbb userprofile1
```

Note

- Use one line per node, separated by spaces.
- For *InstanceID* and *ENIID*, enter the ID information of the operational system and the standby system.

- If the settings are not correct, the Elastic IP cannot be associated correctly, resulting in a resource failure.

Definition file for network takeover by rewriting DNS records

To update DNS records, the ID information of the device managed by AWS and the information to update the DNS records are required.

When taking over the network by rewriting DNS records, create the definition file and the file for updating the DNS records (JSON format).

The definition file and the file for updating DNS records are explained in order.

Check each ID in AWS described in the definition file (/usr/opt/reliant/etc/hvawsconfig) with the AWS Management Console.

Information in the definition file

```
KeyName Mode CFNameX InstanceID HostZoneID change-batch [ProfileName]
```

Item	Contents	Remarks
KeyName	Describe the ID up to 16 letters and numbers. *KeyName* is case-sensitive. This key name is defined in each script that is registered when setting the Cmdline resource. Specify the same *KeyName* for the paired operational system and standby system and define a *KeyName* for other systems to avoid duplicate names. This key name is specified as an argument to register a script when setting Cmdline resources explained later.	-
Mode	Specify the architectural pattern for network takeover. For network takeover by rewriting the DNS records, specify the string of DNS or dns.	-
CFNameX	Specify the CF node name.	-
InstanceID	Specify the instance ID of AWS on which the cluster host is operating.	Check with the AWS Management Console
HostZoneID	Specify the host zone ID.	Check with the AWS Management Console
change-batch	Record information file to update DNS (specified by the absolute path.) To create the record information file, refer to "Record information to update DNS."	-
ProfileName	Specify the profile name of the credentials used with the AWS Command Line Interface configured in "21.3 Presetting." When it is omitted, the device operates with the default profile.	-

Example

The record information file to update DNS is /home/node1.json_sample, and the profile name of the credentials used with the AWS Command Line Interface is userprofile1.

```
CmdD01 DNS node1 i-xxxxxxxx xxxxxxxx /home/node1.json_sample userprofile1
CmdD01 DNS node2 i-yyyyyyyy xxxxxxxx /home/node2.json_sample userprofile1
```

For multiple controls such as the mutual standby configuration, separate *KeyNames* and add the same contents.

```
CmdD01 DNS node1 i-xxxxxxxx xxxxxxxx /home/node1.json_sample userprofile1
CmdD01 DNS node2 i-yyyyyyyy xxxxxxxx /home/node2.json_sample userprofile1
```

```
CmdD02 DNS node1 i-xxxxxxxx xxxxxxxx /home/node1.json_sample2 userprofile1
CmdD02 DNS node2 i-yyyyyyyy xxxxxxxx /home/node2.json_sample2 userprofile1
```

Note

- Use one line per node, separated by spaces.
- For *InstanceID*, enter the ID information of the operational system and the standby system.
- If the settings are not correct, the record of DNS information cannot be updated correctly, resulting in a resource failure.

Record information to update DNS

Use the following samples to create the file to update DNS records (JSON format) for the operational system node and the standby system node. Create the file with a root user and change the permission to 600. Create a JSON file in any location with any file name. For details on this file, refer to the documentation of Amazon Route 53 that is the official AWS documentation.

Example

Item	Value	Setting value	Description
Action	UPSERT	Fixed	To update resource records, specify UPSERT.
Name	Domain name	Specified	Describe the taken over domain name to be set for records. Total length of the domain name can be up to 255 letters. (AWS policy)
Type	A	Fixed	Specify A to associate the IPv4 address.
TTL	Time (second)	Specified	Specify the survival time (seconds) to cache the information of records. The initial value of TTL (Time to Leave) is 300 seconds.
Value	IP address of the operational/ standby system	Specified	In Online processing in the operational system and the standby system, specify the private IP address to associate with the domain in each file. Note that only this value is different for operational and standby systems.

Description example of the operational system: /home/node1.json_sample

```
{
 "Comment": "CREATE/DELETE/UPSERT a record ",
 "Changes": [{
 "Action": "UPSERT",
     "ResourceRecordSet": {
         "Name": "sub.fujitsu.com",
         "Type": "A",
         "TTL": 300,
         "ResourceRecords": [{ "Value": "172.30.10.10" }]
}}]
}
```

Description example of the standby system: /home/node2.json_sample

```
{
 "Comment": "CREATE/DELETE/UPSERT a record ",
 "Changes": [{
 "Action": "UPSERT",
     "ResourceRecordSet": {
         "Name": "sub.fujitsu.com",
         "Type": "A",
         "TTL": 300,
         "ResourceRecords": [{ "Value": "172.30.20.20" }]
}}]
}
```

21.10.1.2 Checking the Network Takeover Settings

Execute the following command on all nodes and make sure that the information described in the definition file (/usr/opt/reliant/etc/hvawsconfig) is correctly described to control the device of AWS.

```
# /opt/SMAW/bin/hvawschkconf
```

If there are no problems with the contents of the definition file, the display is as follows.

Example) The definition file is /usr/opt/reliant/etc/hvawsconfig.

```
# /opt/SMAW/bin/hvawschkconf
NOTICE: Check completed successfully. file=/usr/opt/reliant/etc/hvawsconfig
```

If there is a problem with the contents of the definition file, the setting value that must be checked is output. Follow the displayed message to take an action.

For details on the hvawschkconf (8) command messages, refer to "PRIMECLUSTER Messages."

21.10.1.3 Building Cluster Application

This section describes how to register the network takeover to the cluster application.

Architectural pattern	Registered resource	Usage
Network takeover by the virtual router	Cmdline resources Takeover network resources	In the operational system, the routes of the virtual router are updated. In the operational system, the takeover IP address is activated.
Network takeover by rewriting DNS records	Cmdline resources	In the operational system, the DNS record information is updated.
Network takeover by replacing the Elastic IP address	Cmdline resources	In the operational system, the Elastic IP address is associated with the network interface (ENIID).

21.10.1.3.1 Setting Cmdline Resources for Network Takeover

The procedure to register Cmdline resources is the same for all architectural patterns for network takeover.

For details on how to set Cmdline resources, refer to the procedure for setting up Cmdline resources described in "6.7.3 Setting Up Resources" in "PRIMECLUSTER Installation and Administration Guide", and set the setting values described in "Table 21.1 Creating Cmdline resources and setting Online/Offline/Check scripts."

Table 21.1 Creating Cmdline resources and setting Online/Offline/Check scripts

Parameter name	Setting value
StartCommands[0]	Set the following value. /opt/SMAW/bin/hvawsipalias -c *KeyName* *KeyName* Specify the *KeyName* that was predefined in the definition file (/usr/opt/reliant/etc/hvawsconfig).

Parameter name	Setting value
StopCommands[0]	Set the following. /opt/SMAW/bin/hvawsipalias -u *KeyName* *KeyName* Specify the argument equal to *KeyName* specified in StartCommands.
CheckCommands[0]	Set the following. /opt/SMAW/bin/hvawsipalias -m *KeyName* *KeyName* Specify the argument equal to *KeyName* specified in StartCommands.
CheckCommandTimeouts[0]	Specify the amount of time it takes until PRIMECLUSTER diagnoses an error when the command specified in CheckCommands[X] has hung up. Set the value of SCRIPTTIMEOUT (default value is 300 seconds).
Flags[0]	Settings of the AUTORECOVER attribute (Initial value is valid) and the TIMEOUT attribute are optional. All other attributes should be the default settings. Setting example: Flags[0]=XAT300 (AutoRecover valid) Flags[0]=XT300 (AutoRecover invalid)

Note

The AWS Command Line Interface is used in monitoring processing of Cmdline resources for network takeover.

If the path to the API endpoint is blocked by an error of the NAT instance or the NAT gateway, the AWS Command Line Interface ends with an error and the monitoring processing for network takeover fails. In this case, the cluster application is switched. If you wish to prevent the cluster application from switching because the path to the API endpoint is blocked, set the parameters in "Table 21.1 Creating Cmdline resources and setting Online/Offline/Check scripts" as follows and disable the monitoring of network takeover.

Parameter name	Setting value
StartCommands[0]	Set the same value as the value in "Table 21.1 Creating Cmdline resources and setting Online/Offline/Check scripts."
StopCommands[0]	Set the same value as the value in "Table 21.1 Creating Cmdline resources and setting Online/Offline/Check scripts."
CheckCommands[0]	Set "none".
CheckCommandTimeouts[0]	Set "none".
Flags[0]	Setting of the TIMEOUT attribute is optional. Setting example: Flags[0]=DT300

21.10.1.3.2 Setting Takeover Network Resources Used for the Network Takeover by the Virtual Router

Static physical IP addresses

Perform the following steps on all nodes to ensure that the physical IP addresses of all network interfaces for which you want to configure a Takeover IP address are persistent.

1. Determine the connection profile name (NAME) for the network interface (DEVICE) for which you want to set the Takeover IP address by executing the following command.

```
# nmcli connection show
```

Example: Result displayed when the command is executed

```
# nmcli connection show
NAME    UUID            TYPE    DEVICE
ens256    9f45e94f-1726-dd68-8a33-8022f72b550f    ethernet    ens256
ens192    03da7500-2101-c722-2438-d0d006c28c73    ethernet    ens192
cip0    c24ca981-b8d6-487f-89ba-e73c58da349d    ethernet    cip0
  :
```

2. Execute the following command to check the ipv4.method value of the connection profile name that you saw in 1.

```
# nmcli connection show <connection profile name>
```

If the ipv4.method value is not manual, record the following values and perform the steps from 3.

- IP4.ADDRESS,IP4.GATEWAY,IP4.DNS

If the ipv4.method value is manual, no further steps are required.

Example: The connection profile name is ens192

```
# nmcli connection show ens192
connection.id:                   ens192
connection.uuid:                 03da7500-2101-c722-2438-d0d006c28c73
  :
ipv4.method:                      auto
  :
GENERAL.MASTER-PATH:             --
IP4.ADDRESS[1]:                   172.31.0.10/24
IP4.GATEWAY:                       172.31.0.1
IP4.ROUTE[1]:                      dst = 172.31.0.0/24, nh = 0.0.0.0, mt = 100
IP4.DNS[1]:                         172.31.0.2
```

3. Execute the following command to set the value of IP4.ADDRESS as a fixed value from 2.

```
# nmcli connection modify <connection profile name> ipv4.addresses <IP4.ADDRESS>
# nmcli connection modify <connection profile name> ipv4.method manual
```

Example: The connection profile name is ens 192 and the IP4.ADDRESS value is 172.31.0.10/24

```
# nmcli connection modify ens192 ipv4.addresses 172.31.0.10/24
# nmcli connection modify ens192 ipv4.method manual
```

4. Execute the following command to set the IP4.GATEWAY, IP4.DNS values from 2 and activate the connection profile name.

```
# nmcli connection modify <connection profile name> ipv4.gateway <IP4.GATEWAY>
# nmcli connection modify <connection profile name> ipv4.dns <IP4.DNS>
# nmcli connection up <connection profile name>
```

Example: The connection profile name is ens192, the IP4.GATEWAY value is 172.31.0.1, and the IP4.DNS value is 172.31.0.2

```
# nmcli connection modify ens192 ipv4.gateway 172.31.0.1
# nmcli connection modify ens192 ipv4.dns 172.31.0.2
# nmcli connection up ens192
```

5. Execute the following command and verify that the values for ipv4.addresses, ipv4.method, ipv4.gateway, ipv4.dns appear as configured in 3.4.

```
# nmcli connection show <connection profile name>
```

Example: The connection profile name is ens192

```
# nmcli connection show ens192
connection.id:                         ens192
connection.uuid:                    03da7500-2101-c722-2438-d0d006c28c73
  :
ipv4.method:                          manual
ipv4.dns:                                   172.31.0.2
ipv4.dns-search:                            --
  :
ipv4.addresses:                             172.31.0.10/24
ipv4.gateway:                               172.31.0.1
```

Setting Takeover Network Resources Used for the Network Takeover by the Virtual Router

To set up the takeover network resources, refer to "6.7.3.6 Setting Up Takeover Network Resources" in "PRIMECLUSTER Installation and Administration Guide."

The cluster configuration file /usr/opt/reliant/etc/hvipalias for presetting this configuration must be described by the following rules:

```
CFNameX takeover interface netmask
```

```
CFNameX     : CF node name of the node which uses the takeover IP address
takeover    : Host name of the takeover IP address
interface   : Network interface name on which the takeover IP address will be activated
netmask     : Netmask for the takeover IP address(0xffffffff)
```

Note

The following steps are required to set up the takeover network resources when using network takeover by the virtual router.

- The IP address taken over by the network takeover by the virtual router must be the network different from the CIDR range of the VPC.
- For the netmask described in the cluster configuration file, specify 32-bit (specify 8 digits in hexadecimal) 0xffffffff.
- Set VIRTUAL (default value) for the BASE attribute and the VIRTUAL attribute of the takeover network resources.

About to define a target host (PingHosts)

- Setting up target hosts (PingHosts) is not required. Set the IP address of the target host only when there is an IP address that needs to be monitored for the operations.
  Note that in a Multi-AZ environment, you cannot specify a route table on a Public LAN as a target host.

Example

The CIDR range of VPC is 172.30.0.0/17.

- /etc/hosts

```
172.31.0.10 takeoverip # takeoverIP
```

- /usr/opt/reliant/etc/hvipalias

```
node1 takeoverip eth1 0xffffffff
node2 takeoverip eth1 0xffffffff
```

Chapter 22 Operations

For details on functions for managing PRIMECLUSTER system operations, refer to "Chapter 7 Operations" in "PRIMECLUSTER Installation and Administration Guide."

See

For details on how to operate GDS, refer to "Operation and Maintenance" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

Note

In an AWS environment, a heartbeat may fail due to an error of the network node or an error of the storage controller, or scheduled maintenance for the infrastructure. This may switch the cluster application.

Chapter 23 Changing the Configurations

For details on changing the configuration information for the PRIMECLUSTER system, environment settings, the configuration of the cluster application, the operation attributes of the cluster system, refer to "Chapter 9 Changing the Cluster System Environment", "Chapter 10 Configuration Change of Cluster Applications", "Chapter 11 Changing the Operation Attributes of a Cluster System" in "PRIMECLUSTER Installation and Administration Guide." For details on changing the GDS configuration, refer to "Configuration Change" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

23.1 Configuration Change of AWS Environment

This section describes how to change the configuration of the AWS environment.

23.1.1 Configuration Information of AWS Command Line Interface and Changing Credentials

This section describes how to change the configuration information of the AWS Command Line Interface (such as region name) and credentials (access key and secret key).

1. Execute the following command on any one of the nodes in the cluster system to stop RMS.

   ```
   # hvshut -a
   ```

2. Execute the following command on all nodes to stop the shutdown facility.

   ```
   # sdtool -e
   ```

3. Follow the procedures of the AWS Command Line Interface to change the configuration information and credentials.

4. Execute the following command on all nodes to start the shutdown facility.

   ```
   # sdtool -b
   ```

5. Execute the following command on all nodes and make sure that the shutdown facility operates normally.

   ```
   # sdtool -s
   ```

6. Execute the following command on any one of the nodes in the cluster system to start RMS.

   ```
   # hvcm -a
   ```

7. Execute the following command on any one of the nodes in the cluster system and make sure that RMS operates normally.

   If RMS is stopped, "RMS is not running" is output.

   ```
   # hvdisp -a
   ```

23.1.2 Changing Profile of AWS Command Line Interface

This section describes how to change the profile of the AWS Command Line Interface.

1. Execute the following command on any one of the nodes in the cluster system to stop RMS.

   ```
   # hvshut -a
   ```

2. Execute the following command on all nodes to stop the shutdown facility.

   ```
   # sdtool -e
   ```

3. Follow the procedures of the AWS Command Line Interface to set the profile on all nodes.

4. Modify the configuration definition file of the AWS shutdown agent on all nodes.

   For the description of the configuration definition file, refer to step 2 in "21.9.1.2.1 Setup Procedure of the Shutdown Facility for the Asynchronous Forcible Stop Method."

5. Modify the definition file for controlling network devices of AWS on all nodes.

   For the description of the definition file, refer to "21.10.1.1 Creating Definition File."

6. Execute the following command on all nodes to start the shutdown facility.

   ```
   # sdtool -b
   ```

7. Execute the following command on all nodes and make sure that the shutdown facility operates normally.

   ```
   # sdtool -s
   ```

8. Execute the following command on any one of the nodes in the cluster system to start RMS.

   ```
   # hvcm -a
   ```

9. Execute the following command on any one of the nodes in the cluster system and make sure that RMS operates normally.

   If RMS is stopped, "RMS is not running" is output.

   ```
   # hvdisp -a
   ```

Chapter 24 Maintenance

When you maintain the PRIMECLUSTER system in an AWS environment, note the following points:

- For the procedure on applying/deleting urgent corrections in an AWS environment, refer to "24.1 Software Maintenance."
- For details on other items and procedures required for maintenance of the PRIMECLUSTER system, refer to "Chapter 12 Maintenance of the PRIMECLUSTER System" in "PRIMECLUSTER Installation and Administration Guide." For details on how to maintain GDS, refer to "Operation and Maintenance" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

24.1 Software Maintenance

24.1.1 Notes on Applying Corrections to the PRIMECLUSTER System

For details on notes for applying an intensive correction to the cluster system, refer to "12.3.1 Notes on Applying Corrections to the PRIMECLUSTER System" in "PRIMECLUSTER Installation and Administration Guide."

Note

In an AWS environment, refer to "24.1.2 Overview of the Procedure for Applying/Deleting Corrections" to apply/delete the corrections in multi-user mode.

24.1.2 Overview of the Procedure for Applying/Deleting Corrections

Overview of the procedure is shown for applying each correction including an intensive correction to the cluster system in an AWS environment. In an environment that does not use GDS, the procedure related to GDS is not necessary.

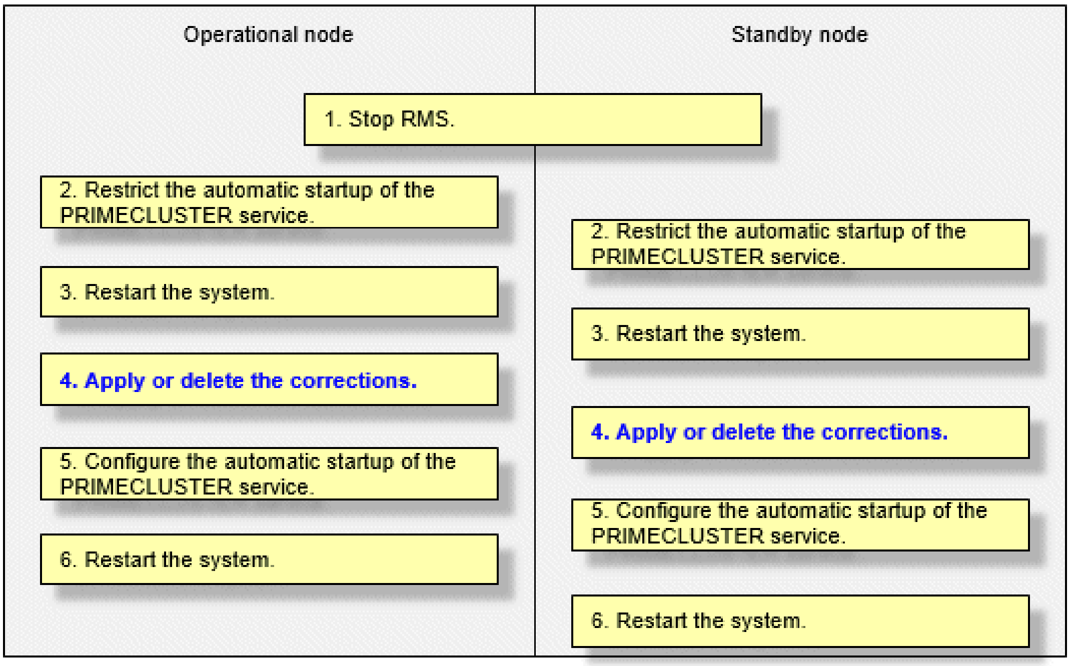
Note

Before applying/deleting corrections to PRIMECLUSTER, take a snapshot of the system storage.

24.1.2.1 Procedure for Applying/Deleting Corrections by Stopping the Entire System

This section describes the procedure for applying/deleting corrections by stopping the entire cluster system.

Flow of the operation


Operational node
Standby node
1. Stop RMS.
2. Restrict the automatic startup of the PRIMECLUSTER service.
3. Restart the system.
4. Apply or delete the corrections.
5. Configure the automatic startup of the PRIMECLUSTER service.
6. Restart the system.
2. Restrict the automatic startup of the PRIMECLUSTER service.
3. Restart the system.
4. Apply or delete the corrections.
5. Configure the automatic startup of the PRIMECLUSTER service.
6. Restart the system.

Operation procedure

Copy the corrections to be applied to each node to the local file system in advance.

1. Stop RMS.

   If RMS is running, execute the following command on any one node in the cluster system to stop RMS.

   ```
   # hvshut -a
   ```

   Note

   If RMS is stopped on all nodes during the synchronization copying of the GDS volume, the synchronization copying of the entire volume area is performed after the corrections are applied and all nodes are restarted.

   If you do not want to perform the synchronization copying of the entire area of volume, stop RMS after the synchronization copying is completed.

   To check the slice status of the GDS volume, execute the following command.

   Execute the following command on any one node in the cluster system to check the value of the STATUS field of the command output.

   The status of the copy destination slice is COPY during the synchronization copying, and after copying is complete, the status becomes ACTIVE or STOP.

   ```
   # sdxinfo -S
   ```

2. Restrict the automatic startup of the PRIMECLUSTER service.

   Restrict the automatic startup of the PRIMECLUSTER service by executing the following command on all nodes.

   ```
   # /opt/FJSVpclinst/bin/pclservice off
   ```

3. Restart the system.

   Restart the system on all nodes.

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

4. Apply or delete the corrections.

   Apply the corrections that were copied to the local file system or delete the corrections.

   - Applying corrections

     Copy the corrections to the working directory and then execute the following commands.

     ```
     # cd <working directory>
     # /opt/FJSVfupde/bin/uam add -d ./ -i <correction number>
     ```

     If the following message is displayed, select "Y".

     ```
     It is required to update with single user mode. Do you want to apply the update now? (Y/N)Y
     ```

     After that, the following message is displayed. Select "N".

     ```
     Do you want to restart your computer immediately? (Y/N)N
     ```

   - Deleting corrections

     Execute the following command.

     ```
     # /opt/FJSVfupde/bin/uam remove -i <correction number>
     ```

     If the following message is displayed, select "Y".

     ```
     It is required to restore with single user mode. Do you want to restore the updated product
     to its pre-update state now? (Y/N)Y
     ```

     After that, the following message is displayed. Select "N".

     ```
     Do you want to restart your computer immediately? (Y/N)N
     ```

5. Configure the automatic startup of the PRIMECLUSTER service.

   Execute the following command on all nodes and change the PRIMECLUSTER service settings back to the state they were in before they were restricted in step 2.

   ```
   # /opt/FJSVpclinst/bin/pclservice on
   ```

6. Restart the system.

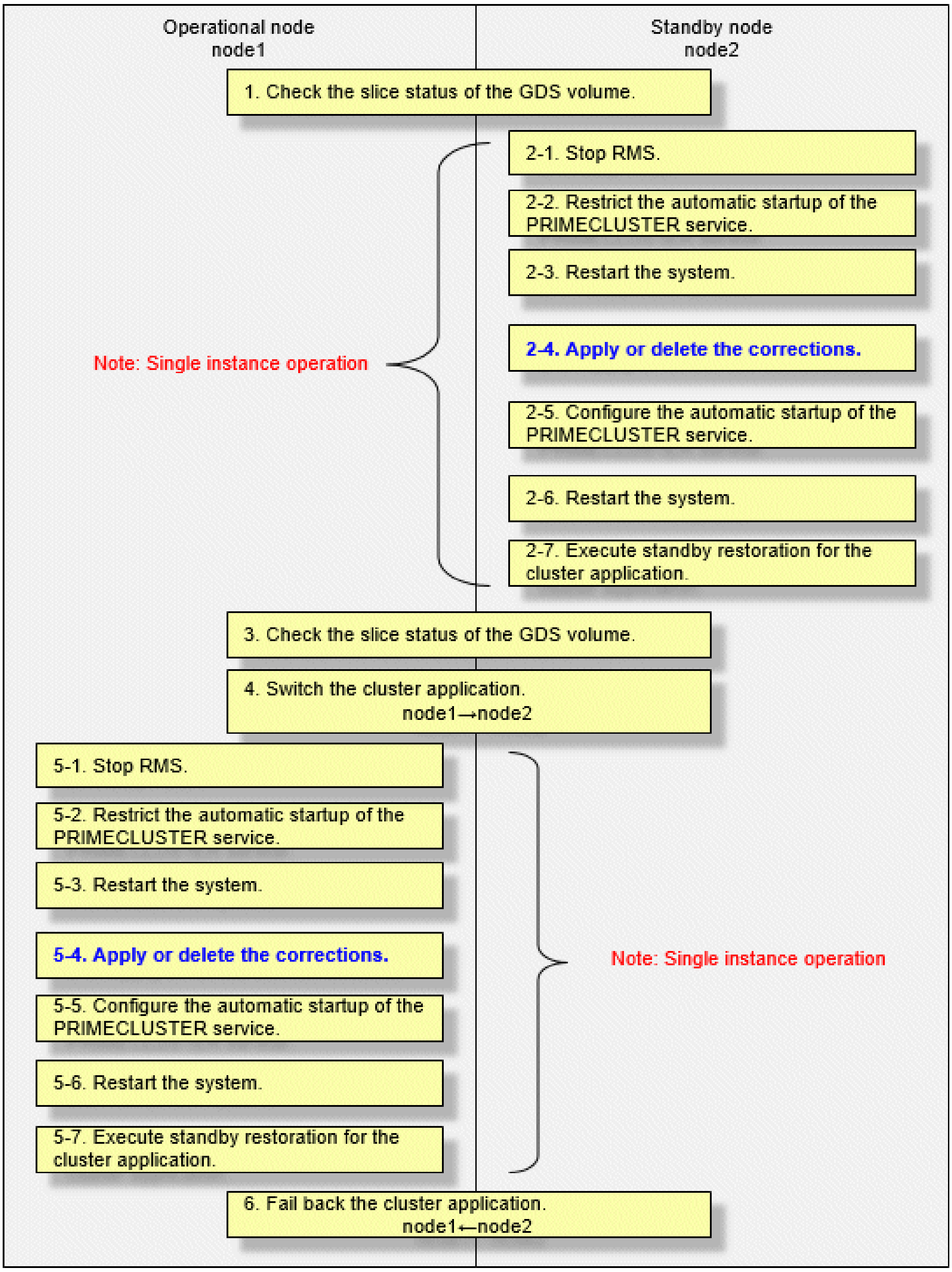
   Restart the system on all nodes.

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

24.1.2.2 Procedure for Applying/Deleting Corrections by Rolling Update

This section describes the procedure for applying corrections by rolling update.

Flow of the operation



Operational node
node1
Standby node
node2
1. Check the slice status of the GDS volume.
2-1. Stop RMS.
2-2. Restrict the automatic startup of the PRIMECLUSTER service.
2-3. Restart the system.
Note: Single instance operation
2-4. Apply or delete the corrections.
2-5. Configure the automatic startup of the PRIMECLUSTER service.
2-6. Restart the system.
2-7. Execute standby restoration for the cluster application.
3. Check the slice status of the GDS volume.
4. Switch the cluster application.
node1→node2
5-1. Stop RMS.
5-2. Restrict the automatic startup of the PRIMECLUSTER service.
5-3. Restart the system.
5-4. Apply or delete the corrections.
Note: Single instance operation
5-5. Configure the automatic startup of the PRIMECLUSTER service.
5-6. Restart the system.
5-7. Execute standby restoration for the cluster application.
6. Fail back the cluster application.
node1←node2

GDS: Global Disk Services

Operation procedure:

1. Check the slice status of the GDS volume.

   Execute the following command on any cluster node to check the value of the STATUS field of the command output.

   ```
   # sdxinfo -S
   ```

   If the COPY status slice exists in the netmirror volume, wait until the synchronization copying is complete.

   For problems caused by node operations during copying, refer to "Stopping or Restarting the Node" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

2. Execute the following operation with the standby node (node2).

    1. Stop RMS.

       Stop RMS to apply corrections to the standby node (node2). A cutoff state transition occurs according to the shutdown of RMS. In this case, make sure that the single instance operation continues until the standby restoration for the cluster application is executed.

       ```
       # hvshut -l
       ```

    2. Restrict the automatic startup of the PRIMECLUSTER service.

       Execute the following command to restrict the automatic startup of the PRIMECLUSTER service.

       ```
       # /opt/FJSVpclinst/bin/pclservice off
       ```

    3. Restart the system.

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

    4. Apply or delete the corrections.

       - Applying corrections

       Copy the corrections to the working directory and then execute the following commands.

       ```
       # cd <working directory>
       # /opt/FJSVfupde/bin/uam add -d ./ -i <correction number>
       ```

       If the following message is displayed, select "Y".

       ```
       It is required to update with single user mode. Do you want to apply the update now? (Y/N)Y
       ```

       After that, the following message is displayed. Select "N".

       ```
       Do you want to restart your computer immediately? (Y/N)N
       ```

       - Deleting corrections

       Execute the following command.

       ```
       # /opt/FJSVfupde/bin/uam remove -i <correction number>
       ```

       If the following message is displayed, select "Y".

       ```
       It is required to restore with single user mode. Do you want to restore the updated
       product to its pre-update state now? (Y/N)Y
       ```

       After that, the following message is displayed. Select "N".

       ```
       Do you want to restart your computer immediately? (Y/N)N
       ```

    5. Configure the automatic startup of the PRIMECLUSTER service.

       Execute the following command and change the PRIMECLUSTER service settings back to the state they were in before they were restricted in 2 of step 2.

       ```
       # /opt/FJSVpclinst/bin/pclservice on
       ```

    6. Restart the system.

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

    7. Execute standby restoration for the cluster application.

       If the node (node1) to which corrections have been applied is cut off from the cluster system, execute standby restoration for the node.

For details on how to execute cluster application standby restoration, refer to "7.2.2.1 Starting a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

3. Check the slice status of the GDS volume.

   After starting the standby node (node2), the synchronization copying of the netmirror volume is executed. Make sure that the synchronization copying is completely finished and all slices are either in ACTIVE or STOP status on any one node.

   To check the slice status of the netmirror volume, execute the following command:

   Execute the following command on any cluster node to check the value of the STATUS field of the command output.

   ```
   # sdxinfo -S
   ```

4. Switch the cluster application.

   To apply corrections to the operational node (node1), execute hvswitch and switch all cluster applications to the standby node (node2). For details on how to switch the cluster applications, refer to "7.2.2.3 Switching a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

5. Perform the following operation with the operational node (node1).

   1. Stop RMS.

      Stop RMS to apply corrections to the operational node (node1). A cutoff state transition occurs according to the shutdown of RMS. In this case, make sure that the single instance operation continues until the standby restoration for the cluster application is executed.

      ```
      # hvshut -l
      ```

   2. Restrict the automatic startup of the PRIMECLUSTER service.

      Execute the following command to restrict the automatic startup of the PRIMECLUSTER service.

      ```
      # /opt/FJSVpclinst/bin/pclservice off
      ```

   3. Restart the system.

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

   4. Apply or delete the corrections.

      - Applying corrections

      Copy the corrections to the working directory and then execute the following commands.

      ```
      # cd <working directory>
      # /opt/FJSVfupde/bin/uam add -d ./ -i <correction number>
      ```

      If the following message is displayed, select "Y".

      ```
      It is required to update with single user mode. Do you want to apply the update now? (Y/N)Y
      ```

      After that, the following message is displayed. Select "N".

      ```
      Do you want to restart your computer immediately? (Y/N)N
      ```

      - Deleting corrections

      Execute the following command.

      ```
      # /opt/FJSVfupde/bin/uam remove -i <correction number>
      ```

      If the following message is displayed, select "Y".

      ```
      It is required to restore with single user mode. Do you want to restore the updated
      product to its pre-update state now? (Y/N)Y
      ```

      After that, the following message is displayed. Select "N".

```
Do you want to restart your computer immediately? (Y/N)N
```

    5. Configure the automatic startup of the PRIMECLUSTER service.

       Execute the following command and change the PRIMECLUSTER service settings back to the state they were in before they were restricted in 2 of step 5.

       ```
       # /opt/FJSVpclinst/bin/pclservice on
       ```

    6. Restart the system.

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

    7. Execute standby restoration for the cluster application.

       If the node (node1) to which corrections have been applied is cut off from the cluster system, execute standby restoration for the node. For details on how to execute cluster application standby restoration, refer to "7.2.2.1 Starting a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

6. Fail back the cluster application.

   Restore the state of the standby layout defined at installation by executing failback operation, as necessary. For details on failback, refer to "7.2.2.3 Switching a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

24.2 Procedure for Restoring OS with the Snapshot Function

If the OS is restored with the snapshot function of AWS, perform the following procedure.

24.2.1 Procedure for Restoring One Node While the Operation is Working

1. Restore the instance. To restore the snapshot or attach the EBS volumes, refer to the official AWS documentation.

   Note

   When using the mirroring among the servers of GDS, replace the system volume of the restoration target instance with the restored volume, not newly create the instance.

2. When using the mirroring among the servers of GDS, check the slice status. If the status of the slice is INVALID, execute the following command to perform the synchronization copying of each volume after the node is started. Perform this procedure on either node.

   ```
   # sdxcopy -B -c <class name> -v <volume name>
   ```

3. If the instance name is changed in step 1, perform step 2 to 4 described in "21.9.1.2 Setting up the Shutdown Facility" on all nodes, and modify the InstanceID of the /usr/opt/reliant/etc/hvawsconfig file on all nodes created in "21.10.1.1 Creating Definition File."

24.2.2 Procedure for Restoring Nodes While the Operation does not Work

1. If either or all of the nodes are started before restoring the nodes, stop RMS. Perform the restoration procedure on either node that is started.

   ```
   # hvshut -a
   ```

2. Select the latest disk when all nodes are started before restoring nodes in an environment where the mirroring among the servers of GDS is used. For all classes of GDS, execute the following command on either node.

   ```
   # /etc/opt/FJSVsdx/bin/sdxnetdisk -S -c <class name>
   ```

3. Restore the instance. To restore the snapshot or attach the EBS volumes, refer to the official AWS documentation.

Note

When using the mirroring among the servers of GDS, replace the system volume of the restoration target instance with the restored volume, not newly create the instance.

4. When using the mirroring among the servers of GDS, start the node and then execute the following with the restored node.

    1. Delete the information of the iSCSI device.

        ```
        # rm -f /var/opt/FJSVsdx/log/.sdxnetmirror_disable.db
        # rm -f /var/opt/FJSVsdx/log/.sdxnetmirror_timestamp
        ```

    2. Stop RMS.

        ```
        # hvshut -l
        ```

5. Restore the other node if necessary.

Note

When using the mirroring among the servers of GDS, replace the system volume of the restoration target instance with the restored volume, not newly create the instance.

6. When using the mirroring among the servers of GDS, delete the iSCSI device information with the restored node if the node is restored with step 5.

    ```
    # rm -f /var/opt/FJSVsdx/log/.sdxnetmirror_disable.db
    # rm -f /var/opt/FJSVsdx/log/.sdxnetmirror_timestamp
    ```

7. If all nodes are stopped before restoring nodes in an environment where the mirroring among the servers of GDS is used, check the status of the source slice for the synchronization copying. If the source slice for the synchronization copying is INVALID, restore the status of the slice. For the "-d" option of the "sdxfix " command, specify the source disk of the synchronization copying. Perform this procedure on either node.

    ```
    # sdxfix -V -c <class name> -v <volume name> -d <disk name> -x NoRdchk
    ```

8. Perform step 2 to 4 described in "21.9.1.2.1 Setup Procedure of the Shutdown Facility for the Asynchronous Forcible Stop Method" on all nodes, and modify the InstanceID of the /usr/opt/reliant/etc/hvawsconfig file on all nodes created in "21.10.1.1 Creating Definition File."

9. If 2 of step 4 is performed, start RMS on the node where RMS is stopped.

    ```
    # hvcm
    ```

Part 5 Azure Environment

This part describes the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system in an Azure environment.

Chapter 25 Cluster Systems in an Azure Environment

PRIMECLUSTER provides clustering for the virtual machines in an Azure virtual network (hereinafter VNet). This enables higher business availability than the availability provided by cloud services.

See

For details on Azure, refer to the official Azure documentation.

The following cluster systems are available in an Azure environment:

- Cluster system in multiple Availability Zones
- Cluster system in a single Availability Zone

25.1 Cluster System in Multiple Availability Zones

In this configuration, the cluster system can be operated on multiple Availability Zones.

By applying PRIMECLUSTER to the virtual machine, in the event of an error, an application can be switched from the operational virtual machine to the standby virtual machine in a short time to provide a highly reliable virtual machine environment.

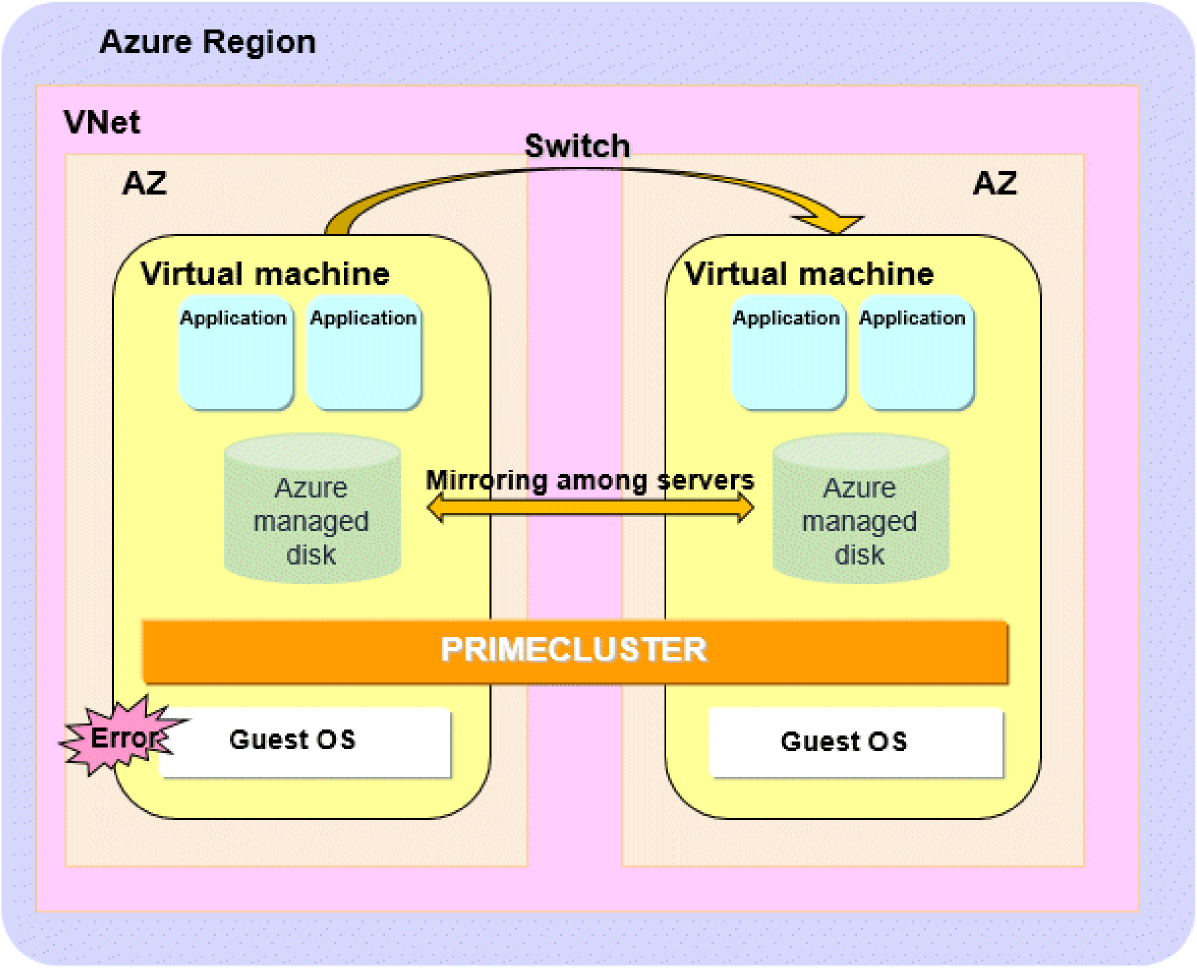
Also, when a network failure occurs in the entire Availability Zone or when Availability Zones become abnormal due to a large-scale disaster, business can be quickly restored by failing over to the standby system with fewer operations.

Note

In the cluster system of multiple Availability Zones, if an error of an Availability Zone occurs, automatic switch to the standby node is not performed, and the node becomes LEFTCLUSTER. To restore the operation, it is necessary to recover from the LEFTCLUSTER state.

For how to recover from the LEFTCLUSTER state, refer to "PRIMECLUSTER Cluster Foundation (CF) Configuration and Administration Guide."

Figure 25.1 Cluster system in multiple Availability Zones



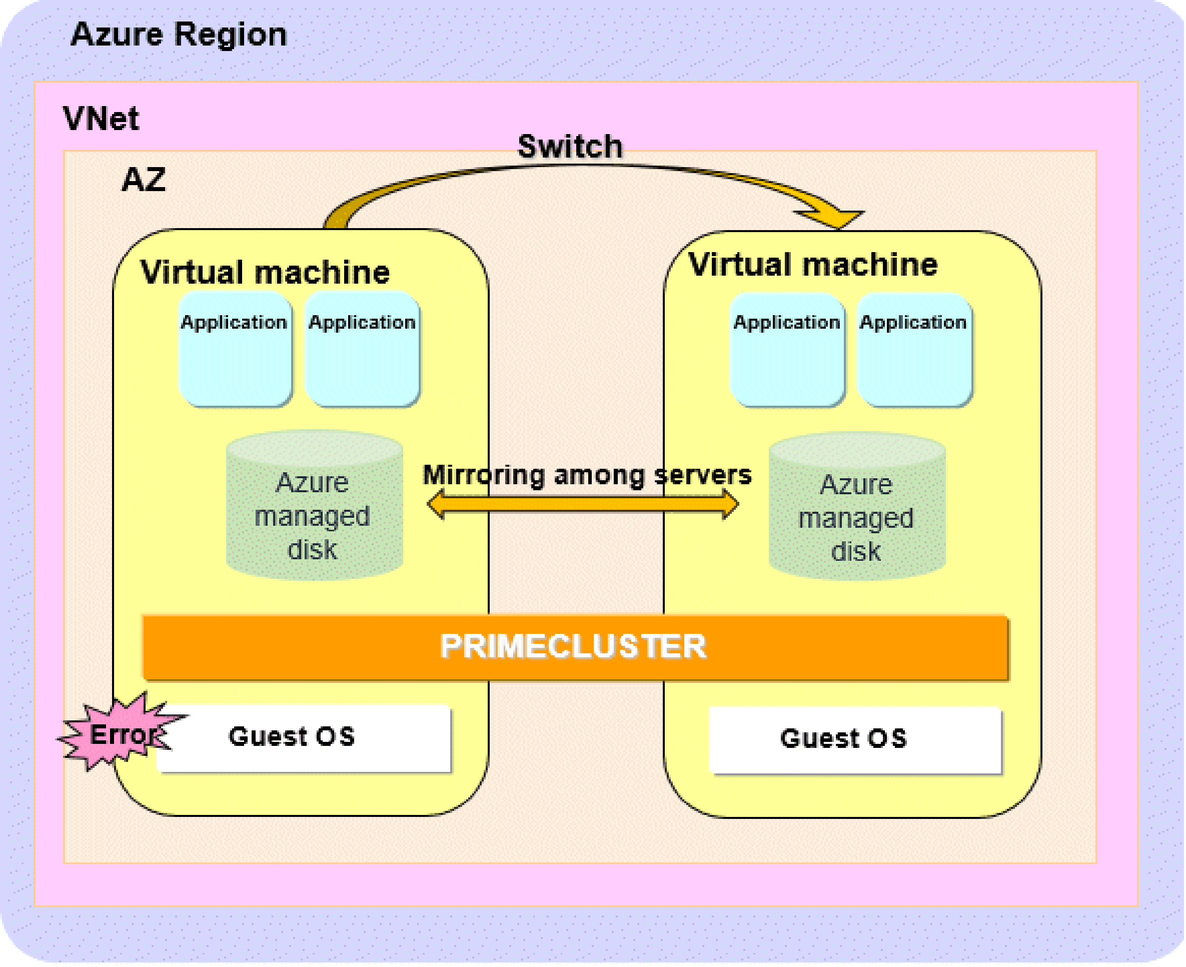
25.2 Cluster System in a Single Availability Zone

In this configuration, the cluster system can be operated in one Availability Zone. By applying PRIMECLUSTER to the virtual machine, in the event of an error, an application can be switched from the operational virtual machine to the standby virtual machine in a short time to provide a highly reliable virtual machine environment.

Note

In the cluster system in a single Availability Zone, in the event of an error in the AZ where the cluster is built, all cluster nodes stop and the business stops.

Figure 25.2 Cluster system in a single Availability Zone



25.3 Supported Range

This section describes the range of support of PRIMECLUSTER in an Azure environment.

Supported configurations

- Number of cluster nodes: 1 to 2 nodes
- Operation mode of the cluster system: 1:1 Standby operation, Mutual standby, Single-node cluster, Scalable operation (only when using Symfoware Server (Native) Hot Standby feature)
- Network configurations: The virtual machines in the cluster system must communicate with the API endpoints.
- Configurations to take over volume data
  - Data takeover by the mirroring among servers of GDS using Azure managed disks

Supported monitoring functions

- Error of AZ (cluster system in multiple Availability Zones)

  The cyclic monitoring of the cluster interconnect detects an error of the AZ, and the node becomes LEFTCLUSTER.

- Error of the guest OS and the cluster interconnect

  The cyclic monitoring of the cluster interconnect detects a hang-up of the guest OS, and the service is automatically switched to the standby system.

- Error of the disk access

  By combining the volume management function of GDS, an error of the disk access can be detected (monitored by the Gds resource), and the service is automatically switched to the standby system when the disk access is disabled.

- Error of the cluster application

  When a resource error of the cluster application occurs, the service is automatically switched to the standby system.

Conditions for the operation

- Operating cloud: Microsoft Azure
- Operating region: East Japan, West Japan
- Operating OS: Red Hat Enterprise Linux 8.2 (for Intel64) or later, Red Hat Enterprise Linux 9.0 (for Intel64) or later

Note

- The snapshot of the virtual machine can be acquired only when the OS is stopped.
- The following functions for PRIMECLUSTER are not available:
  - GFS
  - GLS
  - GDS Snapshot
  - Root class and local class of GDS
  - Single volume of GDS, and disk groups of GDS other than the netmirror type (mirror, stripe, concatenation, and switch)
- Azure features that involve node operations (Auto Scaling, etc.) cannot be used.
- In an Azure environment, when an OS panic occurs, the cluster node may be powered off while the memory dump is being output, and it may not be possible to collect a complete memory dump.

Chapter 26 Design

You must prepare the items listed below before building the PRIMECLUSTER system in an Azure environment.

- Selecting the PRIMECLUSTER Product
- Selecting the Architectural Pattern
- Network Design
- System Design
- Determining the Cluster System Operation Mode
- Determining the Web-Based Admin View Operation Mode
- Determining the Failover Timing of Cluster Application

Point

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

Information

For the flow to build the PRIMECLUSTER system, refer to "Chapter 1 Build Flow" in "PRIMECLUSTER Installation and Administration Guide."

26.1 Selecting the PRIMECLUSTER Product

Select a PRIMECLUSTER product.

In an Azure environment, you can select the following products.

For details on the PRIMECLUSTER products, refer to "2.1 PRIMECLUSTER Product Selection" in "PRIMECLUSTER Installation and Administration Guide."

- PRIMECLUSTER Enterprise Edition (EE)
- PRIMECLUSTER HA Server (HA)
- PRIMECLUSTER Clustering Base (CB)

26.2 Selecting the Architectural Pattern

In the PRIMECLUSTER system in an Azure environment, select an architecture pattern for the item below.

- Ensuring Connectivity with Web-Based Admin View

26.2.1 Ensuring Connectivity with Web-Based Admin View

PRIMECLUSTER provides architectural patterns for ensuring the connectivity between a terminal directly operated by a user and the management view.

For smooth designing of the cluster system, choose the appropriate one from these architectural patterns.

The following are the architectural patterns for the connectivity with the Web-Based Admin View and the appropriate scenarios for each pattern.

Table 26.1 Architectural patterns and appropriate scenarios for the connectivity with the Web-Based Admin View

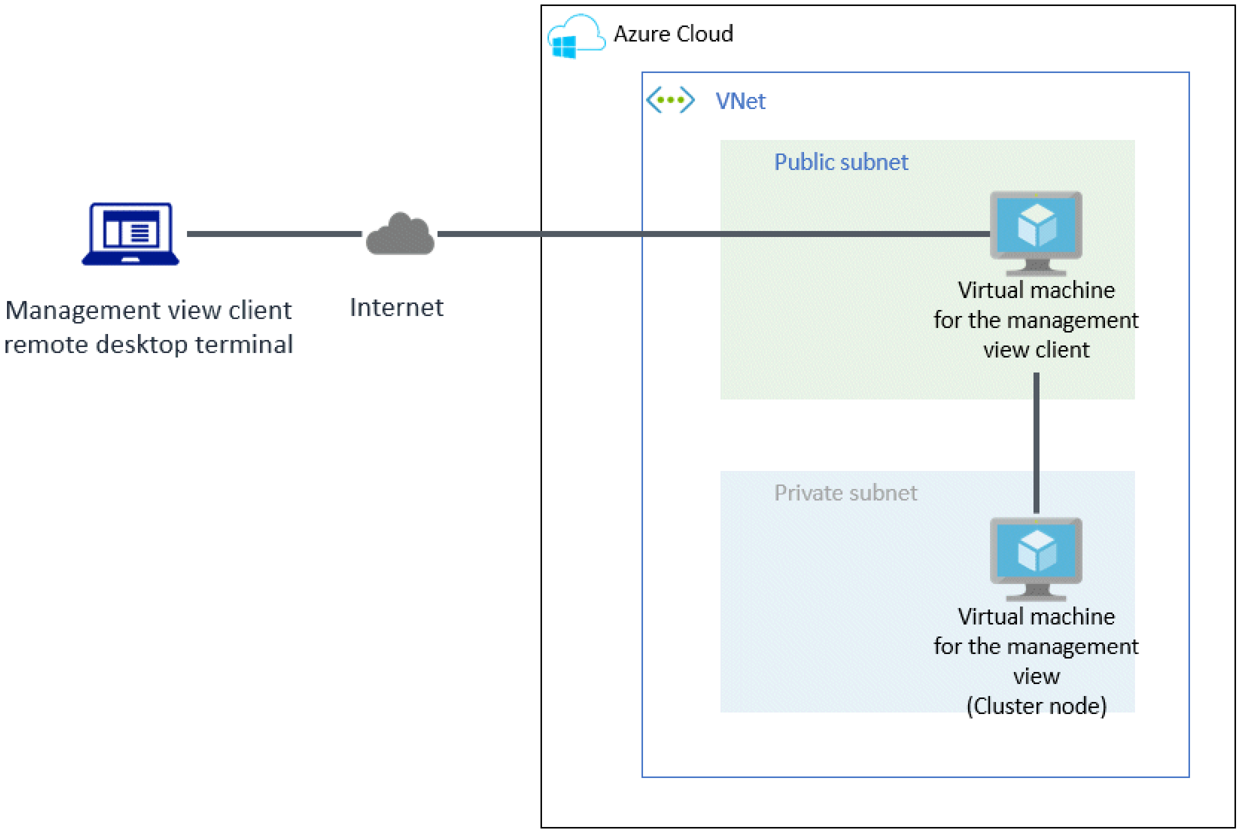
Architectural pattern	Appropriate scenario	Note
Ensuring connectivity with a virtual machine for a client	Pattern using a virtual machine for a client	A virtual machine for the management view client must be deployed in the public subnet.
Ensuring connectivity using a VPN connection	Pattern using a VPN connection	An additional device is required for a VPN connection.

26.2.1.1 Ensuring Connectivity with a Virtual Machine for a Client

In the pattern using a virtual machine for a client, to ensure the connectivity between a terminal directly operated by a user and the management view, prepare a virtual machine for the management view client deployed in the public subnet. The system component for a VPN connection is not required and the configuration may be simple.

When selecting this architectural pattern, a user connects to a client of the management view via a remote desktop connection.

Figure 26.1 Ensuring connectivity with a virtual machine for a client



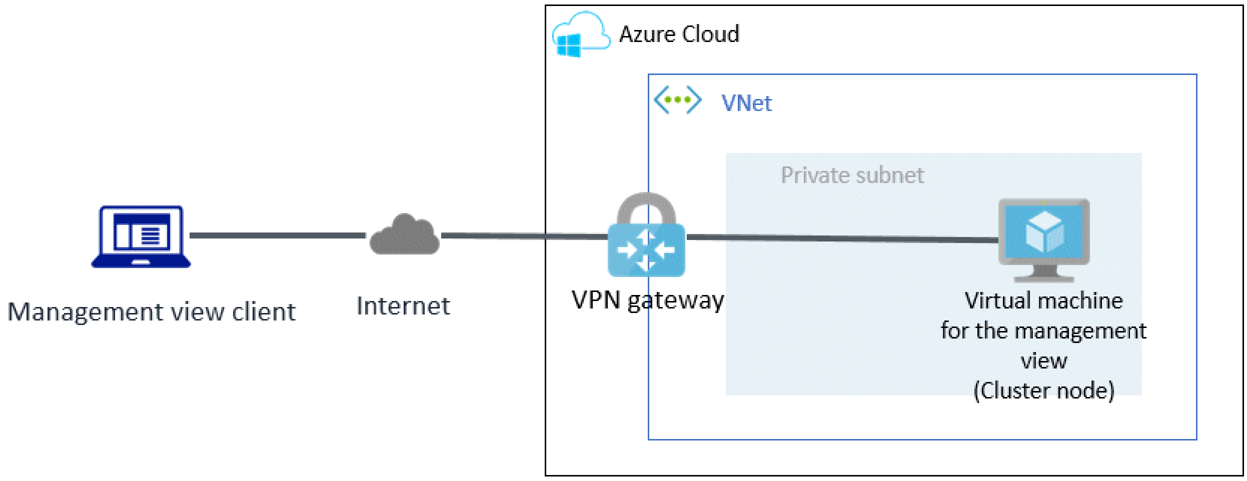
26.2.1.2 Ensuring Connectivity Using a VPN Connection

In the pattern using a VPN connection, to ensure the connectivity between a terminal directly operated by a user and the management view, a VPN connection is used. No virtual machines other than the management view are required in the VNet. This pattern is also secure because it does not use public subnets.

When selecting this architectural pattern, a terminal directly operated by a user is a client of the management view.

To provide a VPN connection, the VPN must be set or a device is required in the network where the terminal directly operated by a user is deployed.

Figure 26.2 Ensuring connectivity using a VPN connection



26.3 Network Design

In the cluster system in an Azure environment, security rules to be applied to the subnet must be designed in advance.

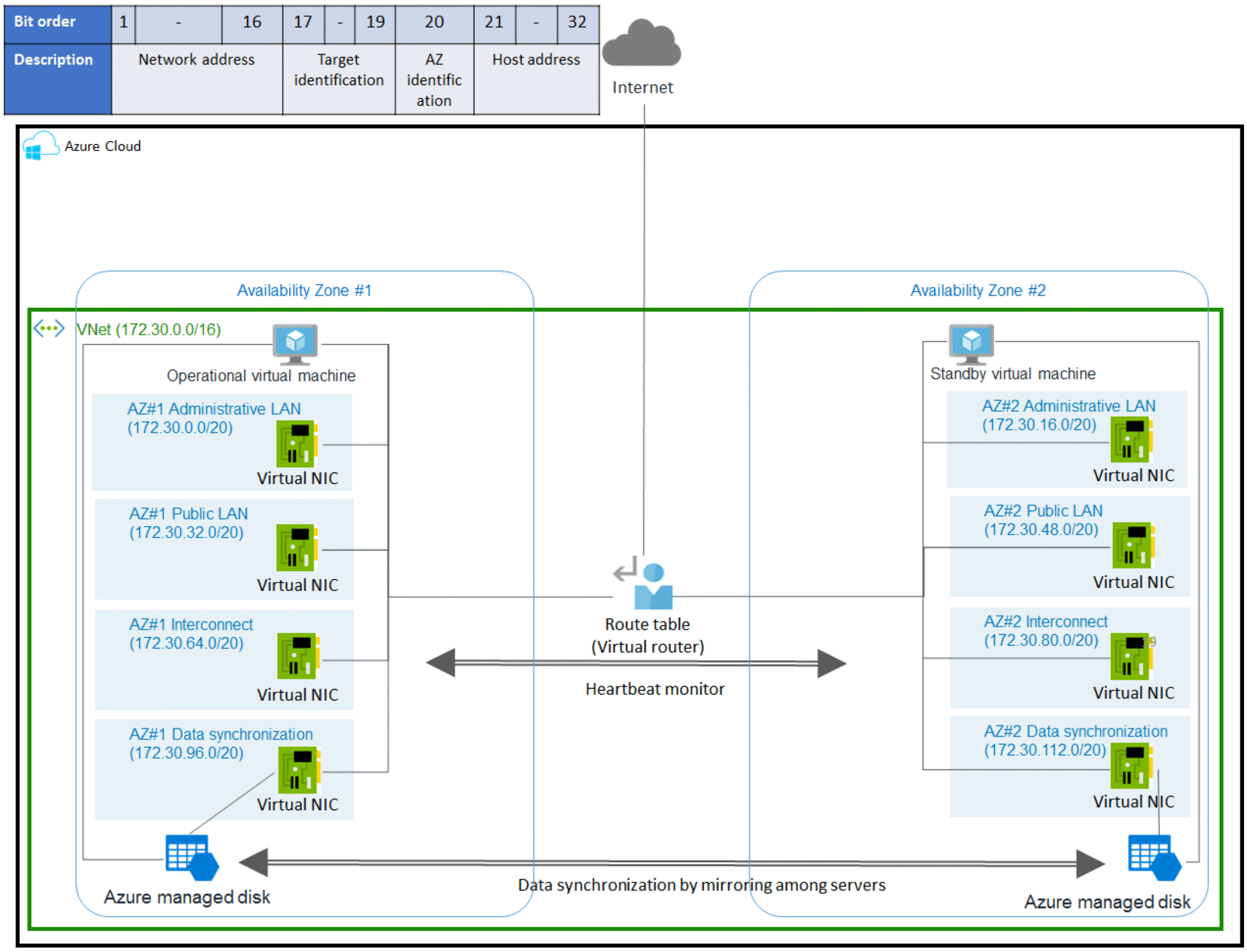
26.3.1 Subnet Design

26.3.1.1 Subnet Design of the Cluster System in Multiple Availability Zones

For the cluster system in multiple Availability Zones, prepare a virtual network and create a subnet for each purpose.

For easier access control, it is recommended that you use the upper bits of the subnet for role identification of the network and use the lower bits for identification of the Availability Zones.

Figure 26.3 Subnet design of the cluster system in multiple Availability Zones



Bit order	1	-	16	17	-	19	20	21	-	32
Description	Network address			Target identification			AZ identification	Host address		

The following is the procedure to create a private subnet in multiple Availability Zones.

1. Prepare a virtual network (VNet) to deploy the system.
2. Define the address space in the VNet. (172.30.0.0/16 in the figure above)

   Select the address range according to the network size.
3. In the VNet, prepare the subnet for each purpose.

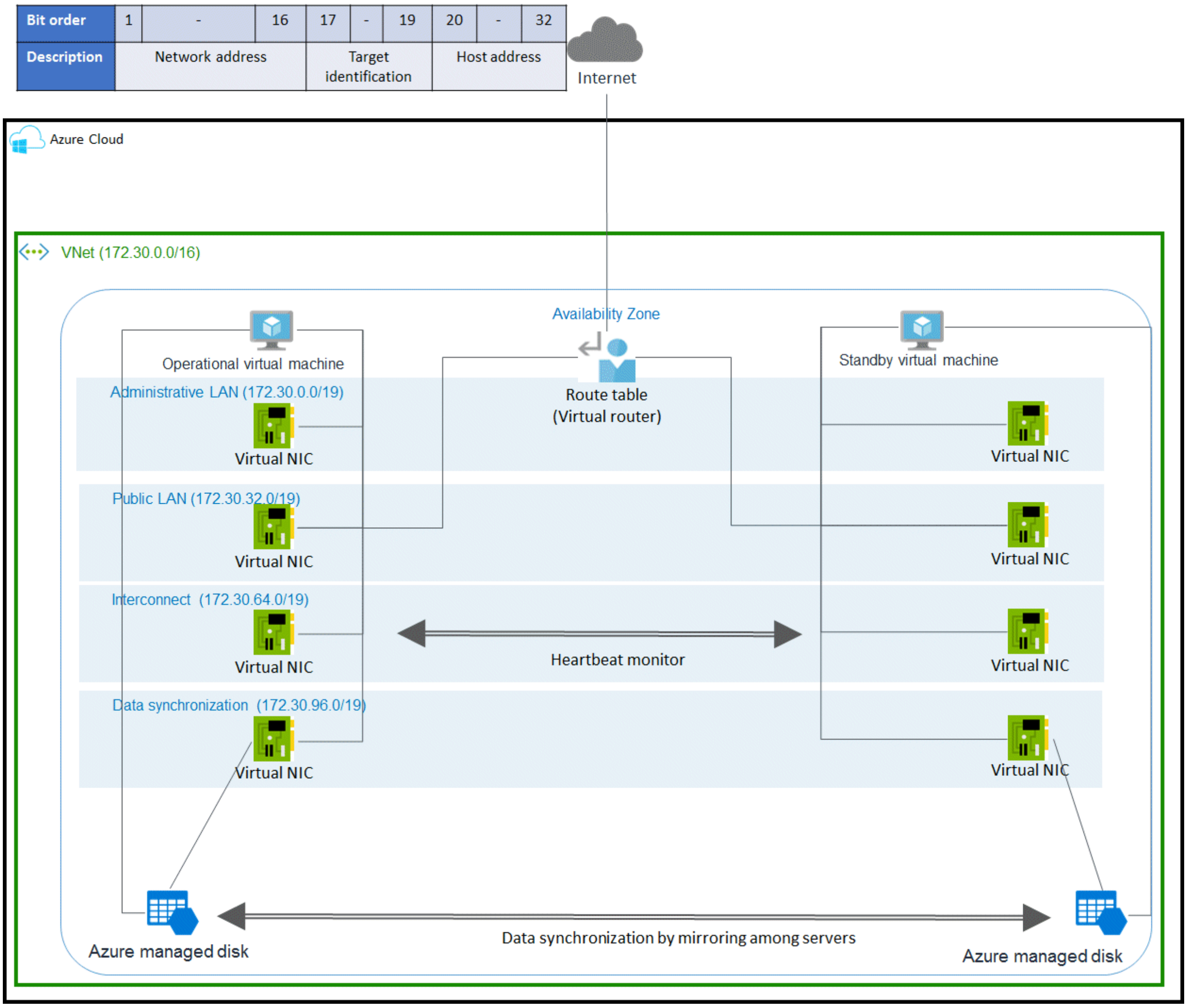
   Prepare the administrative LAN, the public LAN, the cluster interconnect, and the subnet for data synchronization.

   For the prefix length of the subnet, select the appropriate value according to the network size. (The prefix length is 20 in the figure above.)

26.3.1.2 Subnet Design of the Cluster System in a Single Availability Zone

For a single Availability Zone, prepare a virtual network and create a subnet for each purpose.

Figure 26.4 Subnet design of the cluster system in a single Availability Zone



Bit order	1	-	16	17	-	19	20	-	32
Description	Network address			Target identification			Host address		

Internet
Azure Cloud
VNet (172.30.0.0/16)
Availability Zone
Operational virtual machine
Standby virtual machine
Route table
(Virtual router)
Administrative LAN (172.30.0.0/19)
Public LAN (172.30.32.0/19)
Interconnect (172.30.64.0/19)
Data synchronization (172.30.96.0/19)
Virtual NIC
Heartbeat monitor
Azure managed disk
Data synchronization by mirroring among servers
Azure managed disk

The following is the procedure to create a private subnet in a single Availability Zone.

1. Prepare one virtual network (VNet) to deploy the system.
2. Define the address space in the VNet. (172.30.0.0/16 in the figure above)

   Select the address range according to the network size.
3. In the VNet, prepare the subnet for each purpose.

   Prepare the administrative LAN, the public LAN, the cluster interconnect, and the subnet for data synchronization.

   For the prefix length of the subnet, select the appropriate value according to the network size. (The prefix length is 19 in the figure above.)

   Depending on the selected architectural pattern, additional public subnets may be required.

26.3.2 Network Security Group Design

This section describes the security rule settings for the network security group that are required to allow communication within the cluster.

PRIMECLUSTER uses several protocols/ports for communication within the cluster.

When creating a new network security group, all communication within the Azure Virtual Network is allowed according to the default security rules of Azure. When setting detailed security rules, create the security rules described in this section, and allow communication of protocols/ports for communication within the cluster.

In addition to the security rules described in this section, you can add security rules based on the security requirements of the customer to design network security groups.

When adding security rules according to requirements or adding security rules required for the operation of other software, set the priority so that PRIMECLUSTER communication is not rejected.

The tables below describe the security rules for the cluster system with a two-node configuration (CF node names are cluster node 1 and cluster node 2).

See

For details and priorities of the security rules, refer to the official Azure documentation.

26.3.2.1 Security Rules Applied to the Administrative LAN

Design the security rules applied to the administrative LAN.

Inbound security rules

Source	Source port range	Destination	Destination port range	Protocol	Action	Description
Administrative LAN IP of cluster node 1	*(Specify all ports)	Administrative LAN IP of cluster node 2	9382	UDP	Allow	Used for the shutdown facility (SF)
Administrative LAN IP of cluster node 2		Administrative LAN IP of cluster node 1				
Administrative LAN IP of cluster node 1	*(Specify all ports)	Administrative LAN IP of cluster node 2	9796	UDP	Allow	Used for the management view
Administrative LAN IP of cluster node 2		Administrative LAN IP of cluster node 1				
Administrative LAN IP of cluster node 1	*(Specify all ports)	Administrative LAN IP of cluster node 2	9797	TCP	Allow	Used for the management view
Administrative LAN IP of cluster node 2		Administrative LAN IP of cluster node 1				
Administrative LAN IP of cluster node 1	*(Specify all ports)	Administrative LAN IP of cluster node 2	*(Specify all ports)	ICMP	Allow	Used for clchkcluster
Administrative LAN IP of cluster node 2		Administrative LAN IP of cluster node 1				

Replace sources and destinations with VNet/Application security groups according to your requirements.

Outbound security rules

Source	Source port range	Destination	Destination port range	Protocol	Action	Description
Administrative LAN IP of cluster node 1	*(Specify all ports)	Administrative LAN IP of cluster node 2	9382	UDP	Allow	Used for the shutdown facility (SF)
Administrative LAN IP of cluster node 2		Administrative LAN IP of cluster node 1				
Administrative LAN IP of cluster node 1	*(Specify all ports)	Administrative LAN IP of cluster node 2	9796	UDP	Allow	Used for the management view
Administrative LAN IP of cluster node 2		Administrative LAN IP of cluster node 1				
Administrative LAN IP of cluster node 1	*(Specify all ports)	Administrative LAN IP of cluster node 2	9797	TCP	Allow	Used for the management view

Source	Source port range	Destination	Destination port range	Protocol	Action	Description
Administrative LAN IP of cluster node 2		Administrative LAN IP of cluster node 1				
Administrative LAN IP of cluster node 1	*(Specify all ports)	Administrative LAN IP of cluster node 2	*(Specify all ports)	ICMP	Allow	Used for clchkcluster
Administrative LAN IP of cluster node 2		Administrative LAN IP of cluster node 1				
Administrative LAN IP of cluster node 1	*(Specify all ports)	168.63.129.16 (DNS server)	53	TCP	Allow	Used for forced stop and network switching
Administrative LAN IP of cluster node 2						
Administrative LAN IP of cluster node 1	*(Specify all ports)	168.63.129.16 (DNS server)	53	UDP	Allow	Used for forced stop and network switching
Administrative LAN IP of cluster node 2						

Source	Source port range	Destination	Destination service tag	Destination port range	Protocol	Action	Description
Administrative LAN IP of cluster node 1	*(Specify all ports)	Service Tag	AzureCloud	443	TCP	Allow	Used for forced stop and network switching (communication with API endpoint)
Administrative LAN IP of cluster node 2							
Administrative LAN IP of cluster node 1	*(Specify all ports)	Service Tag	Internet	123	TCP	Allow	Used for NTP server query
Administrative LAN IP of cluster node 2							
Administrative LAN IP of cluster node 1	*(Specify all ports)	Service Tag	Internet	123	UDP	Allow	Used for NTP server query
Administrative LAN IP of cluster node 2							

Replace sources and destinations with VNet/Application security groups according to your requirements.

26.3.2.1.1 Security Rules Applied to Web-Based Admin View

Design the security rules applied to the Web-Based Admin View.

1) When ensuring the connectivity with a virtual machine for a client

Design the security rules applied to the Web-Based Admin View (cluster node side).

Inbound security rules

Source	Source port range	Destination	Destination port range	Protocol	Action	Description
Virtual machine IP for the management view client.	*(Specify all ports)	Administrative LAN IP of cluster node 1	8081	TCP	Allow	Used for the management view
		Administrative LAN IP of cluster node 2				
Virtual machine IP for the management view client.	*(Specify all ports)	Administrative LAN IP of cluster node 1	9798	TCP	Allow	Used for the management view
		Administrative LAN IP of cluster node 2				
Virtual machine IP for the management view client.	*(Specify all ports)	Administrative LAN IP of cluster node 1	9799	TCP	Allow	Used for the management view
		Administrative LAN IP of cluster node 2				

Replace sources and destinations with VNet/Application security groups according to your requirements.

Design the security rules applied to the Web-Based Admin View (management client side).

Outbound security rules

Source	Source port range	Destination	Destination port range	Protocol	Action	Description
Virtual machine IP for the management view client.	*(Specify all ports)	Administrative LAN IP of cluster node 1	8081	TCP	Allow	Used for the management view
		Administrative LAN IP of cluster node 2				
Virtual machine IP for the management view client.	*(Specify all ports)	Administrative LAN IP of cluster node 1	9798	TCP	Allow	Used for the management view
		Administrative LAN IP of cluster node 2				
Virtual machine IP for the management view client.	*(Specify all ports)	Administrative LAN IP of cluster node 1	9799	TCP	Allow	Used for the management view
		Administrative LAN IP of cluster node 2				

Replace sources and destinations with VNet/Application security groups according to your requirements.

Also, create an inbound security rule for the network security group to allow a remote desktop connection from a remote control terminal of the management view client to a virtual machine for the management view client.

2) When ensuring the connectivity using a VPN connection

Design the security rules applied to the Web-Based Admin View (cluster node side).

Inbound security rules

Source	Source port range	Destination	Destination port range	Protocol	Action	Description
CIDR of the management view client	*(Specify all ports)	Administrative LAN IP of cluster node 1	8081	TCP	Allow	Used for the management view
		Administrative LAN IP of cluster node 2				

Source	Source port range	Destination	Destination port range	Protocol	Action	Description
CIDR of the management view client	*(Specify all ports)	Administrative LAN IP of cluster node 1	9798	TCP	Allow	Used for the management view
		Administrative LAN IP of cluster node 2				
CIDR of the management view client	*(Specify all ports)	Administrative LAN IP of cluster node 1	9799	TCP	Allow	Used for the management view
		Administrative LAN IP of cluster node 2				

Replace sources and destinations with VNet/Application security groups according to your requirements.

Design the security rules applied to the Web-Based Admin View (management client side).

Outbound security rules

Source	Source port range	Destination	Destination port range	Protocol	Action	Description
CIDR of the management view client	*(Specify all ports)	Administrative LAN IP of cluster node 1	8081	TCP	Allow	Used for the management view
		Administrative LAN IP of cluster node 2				
CIDR of the management view client	*(Specify all ports)	Administrative LAN IP of cluster node 1	9798	TCP	Allow	Used for the management view
		Administrative LAN IP of cluster node 2				
CIDR of the management view client	*(Specify all ports)	Administrative LAN IP of cluster node 1	9799	TCP	Allow	Used for the management view
		Administrative LAN IP of cluster node 2				

Replace sources and destinations with VNet/Application security groups according to your requirements.

26.3.2.1.2 Security Rules Applied to the Virtual Machine Access During Installation and Maintenance

Design the security rules applied to the virtual machine access during installation and maintenance.

Inbound security rules

Source	Source port range	Destination	Destination port range	Protocol	Action	Description
CIDR of the access source	*(Specify all ports)	Administrative LAN IP of cluster node 1	22	TCP	Allow	Used for the SSH remote access
		Administrative LAN IP of cluster node 2				

Replace sources and destinations with VNet/Application security groups according to your requirements.

Outbound security rules

Source	Source port range	Destination	Destination service tag	Destinatio n port range	Protocol	Action	Description
Administrative LAN IP of cluster node 1 Administrative LAN IP of cluster node 2	*(Specify all ports)	Service Tag	Internet	80	TCP	Allow	Used for installing dependent packages
Administrative LAN IP of cluster node 1 Administrative LAN IP of cluster node 2	*(Specify all ports)	Service Tag	Internet	443	TCP	Allow	Used for installing dependent packages

Replace sources and destinations with VNet/Application security groups according to your requirements.

26.3.2.2 Security Rules Applied to the Cluster Interconnect

Design the security rules applied to the cluster interconnect.

This setting is not necessary in a single-node cluster.

Inbound security rules

Source	Source port range	Destination	Destination port range	Protocol	Action	Description
IP of NIC for cluster interconnect of cluster node 1	*(Specify all ports)	IP of NIC for cluster interconnect of cluster node 2	*(Specify all ports)	Any	Allow	Used for the heartbeat
IP of NIC for cluster interconnect of cluster node 2		IP of NIC for cluster interconnect of cluster node 1				

Replace sources and destinations with VNet/Application security groups according to your requirements.

Outbound security rules

Source	Source port range	Destination	Destination port range	Protocol	Action	Description
IP of NIC for cluster interconnect of cluster node 1	*(Specify all ports)	IP of NIC for cluster interconnect of cluster node 2	*(Specify all ports)	Any	Allow	Used for the heartbeat
IP of NIC for cluster interconnect of cluster node 2		IP of NIC for cluster interconnect of cluster node 1				

Replace sources and destinations with VNet/Application security groups according to your requirements.

26.3.2.3 Security Rules Applied to the Public LAN

Design the security rules applied to the public LAN.

Add the security rules that are required for application operations.

26.3.2.4 Security Rules Applied to the Network for Data Synchronization

Set the security rules applied to the network for data synchronization.

Inbound security rules

Source	Source port range	Destination	Destination port range	Protocol	Action	Description
IP of NIC for data synchronization of cluster node 1	*(Specify all ports)	IP of NIC for data synchronization of cluster node 2	3260	TCP	Allow	Used for mirroring among servers
IP of NIC for data synchronization of cluster node 2		IP of NIC for data synchronization of cluster node 1				

Replace sources and destinations with VNet/Application security groups according to your requirements.

Outbound security rules

Source	Source port range	Destination	Destination port range	Protocol	Action	Description
IP of NIC for data synchronization of cluster node 1	*(Specify all ports)	IP of NIC for data synchronization of cluster node 2	3260	TCP	Allow	Used for mirroring among servers
IP of NIC for data synchronization of cluster node 2		IP of NIC for data synchronization of cluster node 1				

Replace sources and destinations with VNet/Application security groups according to your requirements.

26.4 System Design

Use PRIMECLUSTER Designsheets to design the system.

The installation operation of the PRIMECLUSTER system is performed based on the created PRIMECLUSTER Designsheets. Make sure to create the designsheets and confirm that all required items are described.

26.5 Determining the Cluster System Operation Mode

In the cluster system in an Azure environment, operation modes for 1:1 standby operation, mutual standby, and single-node cluster operation can be built.

For details on the operation mode of each cluster system, refer to "2.3 Determining the Cluster System Operation Mode" in "PRIMECLUSTER Installation and Administration Guide."

26.6 Determining the Web-Based Admin View Operation Mode

For details on the Web-Based Admin View operation mode, refer to "2.4 Determining the Web-Based Admin View Operation Mode" in "PRIMECLUSTER Installation and Administration Guide."

26.7 Determining the Failover Timing of Cluster Application

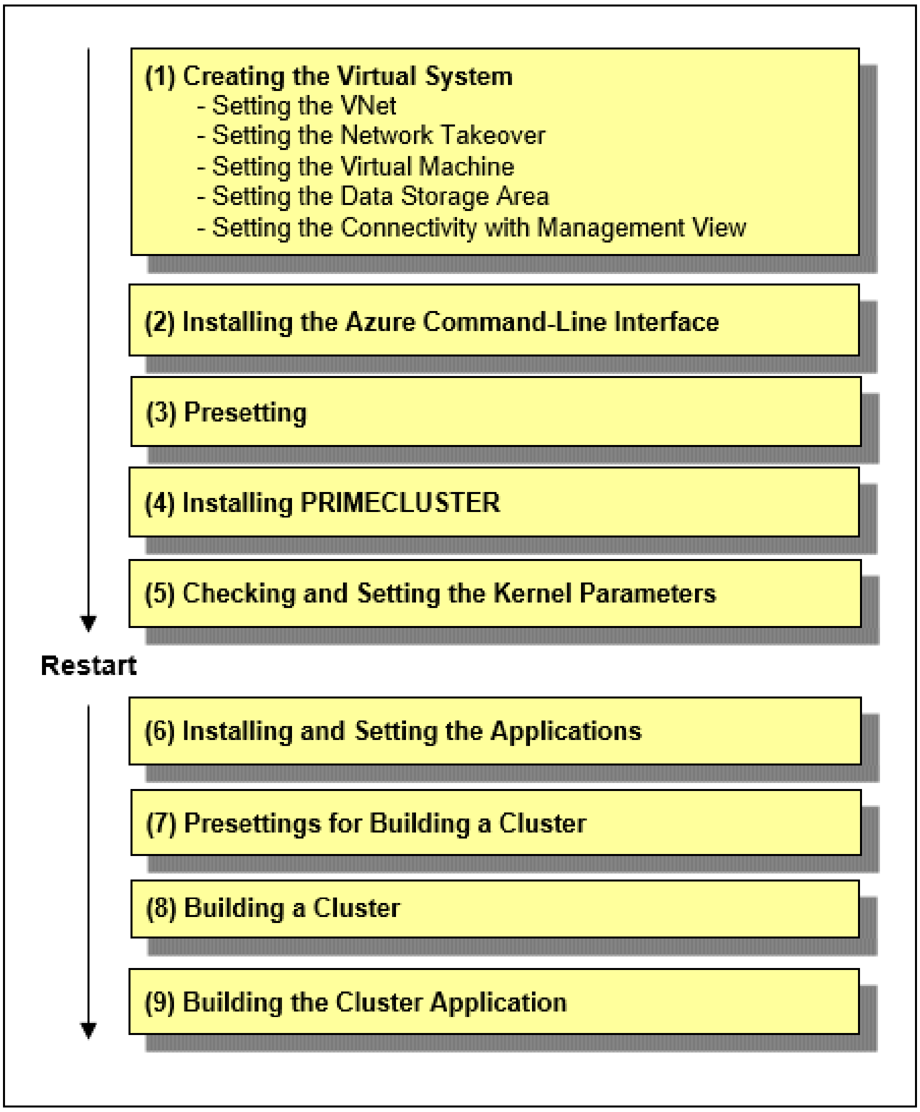
Determine the failover timing of cluster application.

For details, refer to "2.5 Determining the Failover Timing of Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

Chapter 27 Installation

This chapter describes the procedure to install PRIMECLUSTER in an Azure environment.

Perform the steps shown in the figure below.


(1) Creating the Virtual System
- Setting the VNet
- Setting the Network Takeover
- Setting the Virtual Machine
- Setting the Data Storage Area
- Setting the Connectivity with Management View
(2) Installing the Azure Command-Line Interface
(3) Presetting
(4) Installing PRIMECLUSTER
(5) Checking and Setting the Kernel Parameters
Restart
(6) Installing and Setting the Applications
(7) Presettings for Building a Cluster
(8) Building a Cluster
(9) Building the Cluster Application

Note

For how to check the operation environment, refer to "Operation Environment" in the Installation Guide for PRIMECLUSTER.

27.1 Creating the Virtual System

This section describes how to create the virtual system for the cluster system in an Azure environment.

Use the Azure portal to create a virtual system designed in "Chapter 26 Design." The creation procedure depends on the architectural pattern selected.

See

To create a virtual system or for details on each setting, refer to the official Azure documentation.

27.1.1 Setting the VNet

Create the virtual network where the cluster system is to be deployed. Then, create the subnet and set the network security group for the created virtual network according to the design created in "26.3 Network Design."

27.1.2 Setting the Network Takeover

PRIMECLUSTER provides a network takeover method using a secondary IP address change for network takeover in cluster systems on public clouds.

The network takeover method using a secondary IP address change achieves network takeover by Using the Azure CLI to change secondary IP address. In the event of a cluster node error, PRIMECLUSTER automatically secondary IP address change and switches to the standby system to take over the IP address.

Clients can be deployed inside or outside the VNet.

If the client is deployed inside the VNet, access from public sites (clients outside the VNet) will be blocked, and the cluster system will be secure.

Figure 27.1 Network takeover by the secondary IP address change (for clients inside the VNet)

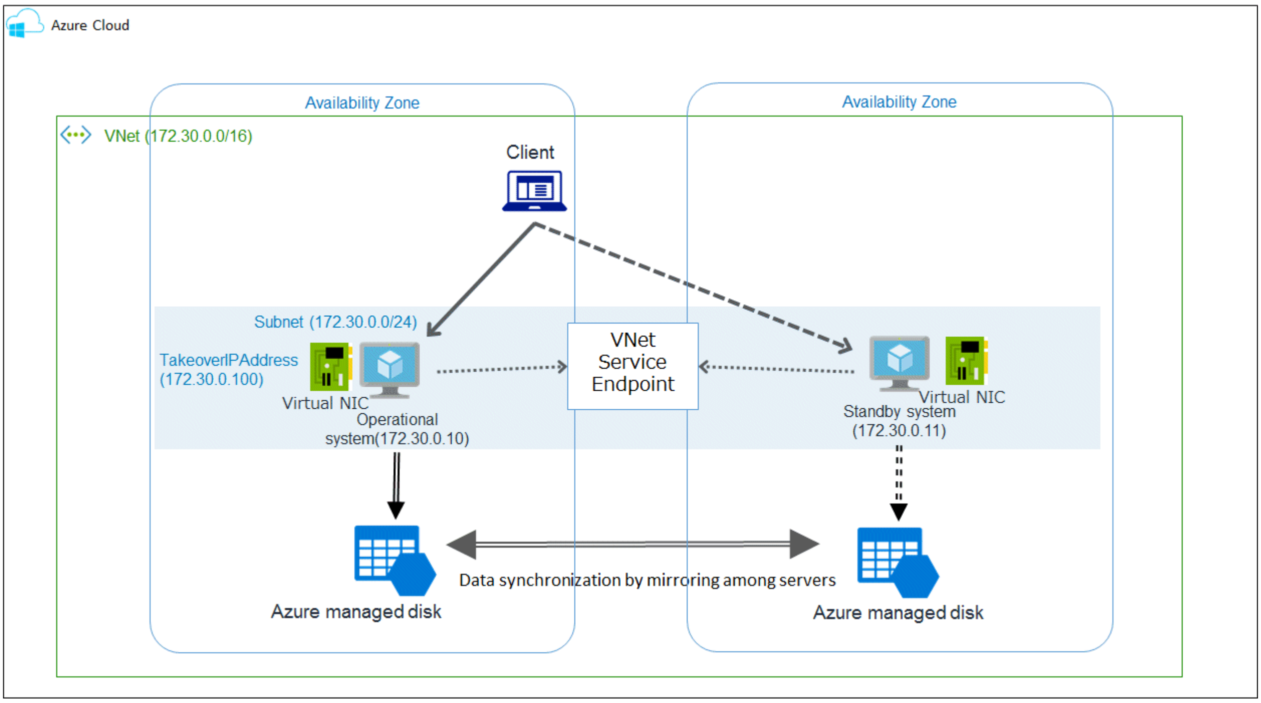
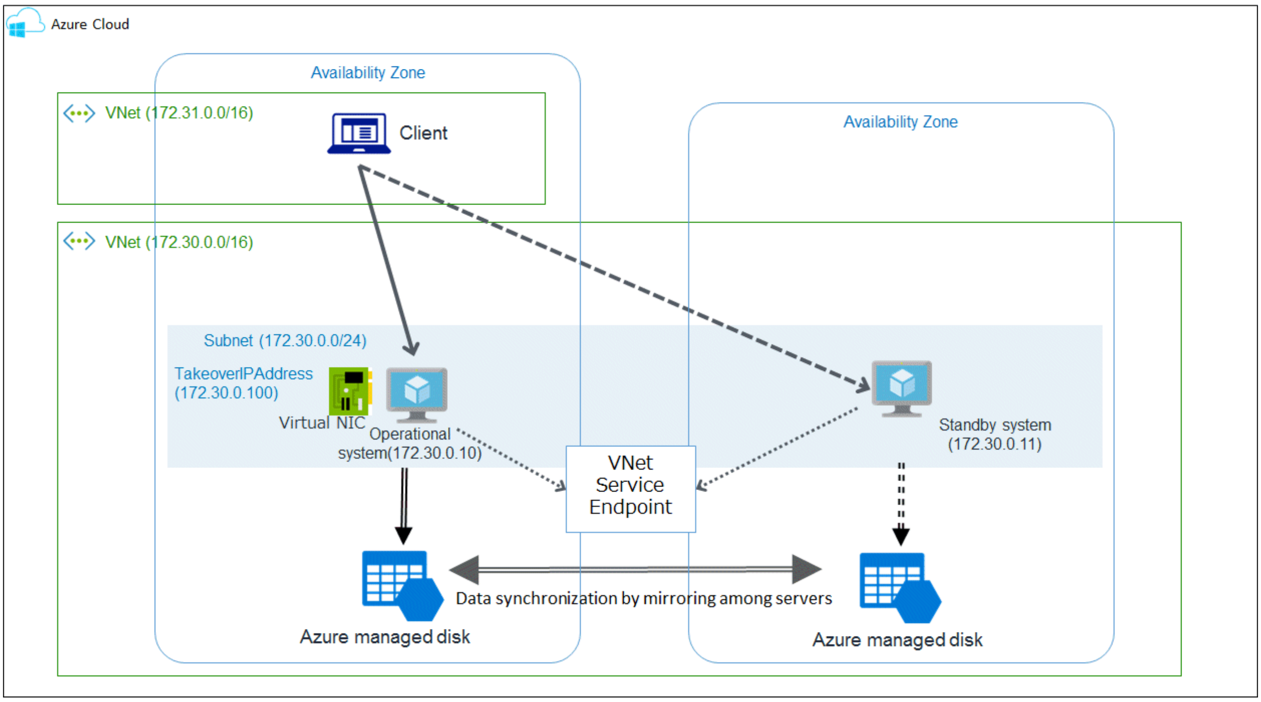


Figure 27.2 Network takeover by the secondary IP address change (for clients outside the VNet)



If you want to access using VPN, create a VPN gateway in addition to the above.

When accessing from another Vnet, configure peering to enable communication between Vnets.

27.1.3 Setting the Virtual Machine

Create the virtual machine and the virtual network interface that configure the cluster node.

When creating a virtual network interface, select an appropriate subnet based on the network configuration designed in "26.3 Network Design."

For the network interface of the public LAN used as the takeover IP address, enable the IP forwarding on Azure. For how to enable the IP forwarding, refer to the official Azure documentation.

When creating multiple virtual network interfaces and using a virtual network interface, for which a default gateway is not set, to communicate with interfaces on different subnets, static routing must be set for that virtual network interface. You do not need to set static routing when communicating only with interfaces on the same subnet. For how to configure static routing, refer to the example of static routing configuration in "21.1.3 Setting Instances."

27.1.4 Setting the Data Storage Area

Set the data storage area used by the applications.

Create the virtual block device managed by the mirroring among servers, then attach it to the virtual machine.

27.1.5 Setting the Connectivity with Management View

Set the components required for ensuring the connectivity between the management terminal and the management view on the cluster node. The required components depend on the architectural pattern selected in "26.6 Determining the Web-Based Admin View Operation Mode."

When selecting a virtual machine for the client to ensure connectivity

Create the public subnet where the virtual machine for the client to be deployed, and then create the virtual machine for the client.

When selecting a VPN connection to ensure connectivity

Create a VPN gateway, customer gateway, and VPN connection.

27.2 Installing the Azure Command-Line Interface

PRIMECLUSTER uses the Azure command-line interface (hereinafter Azure CLI) for split brain resolution and network takeover.

Install it on all nodes that configure the cluster so that it can be executed by the root user. After the installation, verify the path to the Azure CLI.

Note

Use Azure CLI version 2.0.72 or later.

See

For details on the installation procedure, refer to the official Azure documentation.

27.3 Presetting

Note

It is necessary to log in to all nodes using the az login command in the Azure CLI installation procedure.

If not logged in, perform it before the following procedure.

1. Disable the firewall.

   Make sure that "firewalld" is disabled on all nodes.

   ```
   # systemctl is-enabled firewalld
   ```

   If it is enabled, disable it.

   ```
   # systemctl stop firewalld
   # systemctl disable firewalld
   ```

2. Set NTP.

   Make sure to set NTP when building the cluster to synchronize the time of each node in the cluster system.

   Set NTP before installing PRIMECLUSTER.

3. Register service principals and set certificates.

   The service principal is a dedicated authentication method used to execute programs on Azure.

   Set it according to the following procedure.

   3-1) Log in with an Azure account on any one node in the cluster system.

   ```
   # az login -u account
   ```

   Register a service principal and create a certificate.

   ```
   # az ad sp create-for-rbac --create-cert
   {
     "appId": "d5b7dac1-718f-448b-8e11-4a8cca6d9004",
     "displayName": "azure-cli-2019-09-13-02-57-50",
     "fileWithCertAndPrivateKey": "/root/tmprjbQbI.pem",
     "name": "http://azure-cli-2019-09-13-02-57-50",
     "password": null,
   ```

```
  "tenant": "8ff7ddfd-fbcb-4700-ae52-6d071ac8d1b4"
}
```

The application ID (appId) and the tenant ID (tenant) displayed here are used for "27.8.1.2 Setting up the Shutdown Facility" and "27.9.1.1 Creating the Definition File." Record them down.

Log out from Azure.

```
# az logout
```

3-2) Store the certificate created in fileWithCertAndPrivateKey in any same location on all cluster nodes and set the permission to 600.
Example) When the certificate path is "/root/examplecert.pem"

```
# cp /root/tmprjbQbI.pem /root/examplecert.pem
# chmod 600 /root/examplecert.pem
```

Delete the certificate created in fileWithCertAndPrivateKey using the rm command.

The path for storing the certificate is used for "27.8.1.2 Setting up the Shutdown Facility" and "27.9.1.1 Creating the Definition File." Record it down.

3-3) Make sure that you can log in with the registered service principal on all cluster nodes.

```
# az login --service-principal --username appID --tenant tenant --password the path in 3-2) where
the certificate is stored
```

3-4) Log out from Azure on all cluster nodes.

```
# az logout
```

Note

When using Azure CLI version 2.31.0 or earlier, service principal roles are created by Contributor.
When using Azure CLI version 2.32.0 or later, roles are not created for service principals.
Therefore, when changing to another role, refer to "29.1.3 Changing the Role of the Service Principal."

The service principal certificate has an expiration date.
Before the certificate expires, renew the certificate in accordance with the certificate renewal for use by the "29.1.1 Updating the Certificate Used by the Service Principal". If the update is not performed, a forced stop cannot be performed and the cluster application may be stopped.

See

For details on setting up a service principal for the Azure CLI, refer to the official Azure documentation.

4. Set Device Name Deviation Prevention

   In an Azure environment, the network device name immediately after virtual machine deployment is ethX.

   If you continue to use this device name, device name deviation may occur.

   If there is more than one network device, refer to the Linux User's Manual for the operating system and take steps to prevent device name deviation.

27.4 Installing PRIMECLUSTER

Use the installation script (CLI Installer) to install PRIMECLUSTER.

Install PRIMECLUSTER on each node in the system where Linux(R) software and Linux(R) related software are already installed. Use the same installation script when installing PRIMECLUSTER on the cluster management server.

Note

If the OS has never been restarted since the virtual machine was created, restart it and then install PRIMECLUSTER.

See

For details on the installation and uninstallation procedures, refer to the descriptions of cloud environments described in the Installation Guide for PRIMECLUSTER.

27.5 Checking and Setting the Kernel Parameters

Change the kernel parameters depending on the environment.

Applicable nodes:

All nodes on which PRIMECLUSTER is to be installed

Depending on the products and components utilized, different kernel parameters are required.

Check PRIMECLUSTER Designsheets and if you need to modify the kernel parameters, set them again.

See

For details on kernel parameters, refer to "3.1.7 Checking and Setting the Kernel Parameters" in "PRIMECLUSTER Installation and Administration Guide."

Note

- To activate the modified kernel parameters, restart the OS.
- After uninstalling PRIMECLUSTER, change the kernel parameter settings back to the state before installing PRIMECLUSTER if necessary.

27.6 Installing and Setting the Applications

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

See

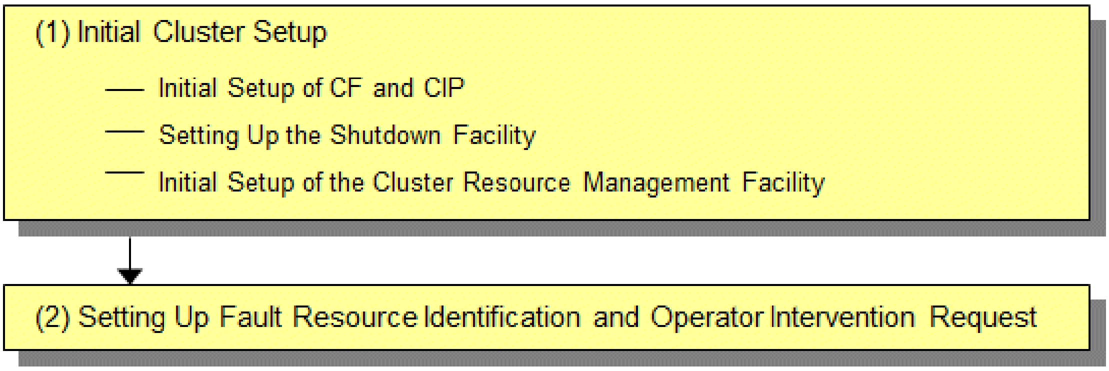
- For details on environment setup, refer to the manuals for each application.
- For information on PRIMECLUSTER-related products that support Azure, refer to the documentation for each product.

27.7 Presettings for Building a Cluster

Prior to building a cluster, perform presettings such as staring the Web-Based Admin View screen. For details on presettings prior to building a cluster, refer to "Chapter 4 Preparation Prior to Building a Cluster" in "PRIMECLUSTER Installation and Administration Guide."

27.8 Building a Cluster

The procedure for building a PRIMECLUSTER cluster is shown below:


(1) Initial Cluster Setup
Initial Setup of 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

27.8.1 Initial Cluster Setup

This section describes the initial cluster setup for PRIMECLUSTER.

27.8.1.1 Initial Setup of CF and CIP

Refer to "5.1.1 Setting Up CF and CIP" in "PRIMECLUSTER Installation and Administration Guide" to set up CF and CIP.

Note

In an Azure environment, the virtual machine may pause for up to 30 seconds without a notification for maintenance, and a heartbeat failure may occur. To prevent the heartbeat failure during the maintenance without a notification, after setting CF, refer to "11.3.1 Changing Time to Detect CF Heartbeat Timeout" in "PRIMECLUSTER Installation and Administration Guide", and tune the cluster timeout value to 30 seconds or more.

You need to use CF over IP in a cloud environment. In addition, you need to configure CIP because such as RMS uses it for inter-node communication in the cluster. CF over IP (IP interconnect) and CIP are different features. For more information, see "1.1.1 Differences between CIP and CF over IP" in the Cluster Foundation Configuration and Administration Guide.

27.8.1.2 Setting up the Shutdown Facility

This section describes how to set up the shutdown facility in an Azure environment.

The shutdown agent available in an Azure environment is as follows.

- Azure CLI

  The Azure CLI shutdown agent SA_vmazureReset provides the function of shutting down nodes (virtual machines) using the Azure CLI in an Azure environment.

  The storage location of a log file is as follows.

```
/var/opt/SMAWsf/log/SA_vmazureReset.log
```

For details on the survival priority, refer to "5.1.2.1 Survival Priority" in "PRIMECLUSTER Installation and Administration Guide."

Note

- After setting up the shutdown agent, conduct a test for the forced stop of cluster nodes to make sure that the correct nodes can be forcibly stopped. For details of the test for the forced stop of cluster nodes, refer to "1.4 Test" in "PRIMECLUSTER Installation and Administration Guide."
- The contents of the SA_vmazureReset.cfg file and the rcsd.cfg file on all nodes should be identical. If not, a malfunction will occur.
- This setting is not necessary in a single-node cluster.

1. Set up the shutdown daemon.

   Create /etc/opt/SMAW/SMAWsf/rcsd.cfg with the following contents on all nodes in the cluster system.

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

```
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 used in the shutdown facility
                 of the cluster host.
                 Available IP addresses are IPv4.
                 When specifying a host name, make sure it is described in /etc/hosts.
SA_vmazureReset: Azure shutdown agent.
timeout        : Specify the timeout duration (seconds) of the Azure shutdown agent.
                 Specify 45 seconds.
```

   Example) The following is a setup example.

   If the CF node names of the cluster host are node1 and node2, the weight of two nodes is 1, the IP address of the administrative LAN of node1 is 192.168.250.1, and the IP address of the administrative LAN of node2 is 192.168.250.2.

```
# cat /etc/opt/SMAW/SMAWsf/rcsd.cfg
node1,weight=1,admIP=192.168.250.1:agent=SA_vmazureReset,timeout=45
node2,weight=1,admIP=192.168.250.2:agent=SA_vmazureReset,timeout=45
```

   Create /etc/opt/SMAW/SMAWsf/rcsd.cfg and then set the owner, group, and access rights as follows.

```
# chown root:root /etc/opt/SMAW/SMAWsf/rcsd.cfg
# chmod 600 /etc/opt/SMAW/SMAWsf/rcsd.cfg
```

   Information

   When creating the /etc/opt/SMAW/SMAWsf/rcsd.cfg file, the /etc/opt/SMAW/SMAWsf/rcsd.cfg.template file can be used as a template.

2. Set up the shutdown agent.

   Create /etc/opt/SMAW/SMAWsf/SA_vmazureReset.cfg with the following contents on all nodes in the cluster system.

   Information

   The template of the SA_vmazureReset.cfg file can be found at the following location:

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

   Delimit each item with a single space.

```
CFNameX ResourceID AppID TenantID CertPath {cycle | leave-off}
CFNameX ResourceID AppID TenantID CertPath {cycle | leave-off}
```

```
CFNameX    : Specify the CF node name of the cluster host.
ResourceID : Specify the resource ID of the Azure virtual machine on which
             the cluster host is running.
AppID      : Specify the application ID when registering the service principal
             (the value recorded in step 3 of "27.3 Presetting").
TenantID   : Specify the tenant ID when registering the service principal
             (the value recorded in step 3 of "27.3 Presetting").
CertPath   : Specify the path of the certificate
             (the value recorded in step 3 of "27.3 Presetting").
```

```
cycle      : Restart the node after forcibly stopping the node.
leave-off  : The node is not restarted after it is forcibly stopped.
```

Example) This is a setting example when the node1 and the node2 are as follows, and the node is forcibly stopped and then restarted.

CFNameX	ResourceID	AppID	TenantID	CertPath
node1	/subscriptions/ 1e234d12-39b8-49db-881a-35aa03b402b9/ resourceGroups/pcl/providers/ Microsoft.Compute/virtualMachines/node1	d5b7dac1-718f-448b-8e11-4a8cca6d9004	8ff7ddfd-fbcb-4700-ae52-6d071ac8d1b4	/root/ examplecert.pem
node2	/subscriptions/ 1e234d12-39b8-49db-881a-35aa03b402b9/ resourceGroups/pcl/providers/ Microsoft.Compute/virtualMachines/node2			

```
# cat /etc/opt/SMAW/SMAWsf/SA_vmazureReset.cfg
node1 /subscriptions/1e234d12-39b8-49db-881a-35aa03b402b9/resourceGroups/pcl/providers/
Microsoft.Compute/virtualMachines/node1 d5b7dac1-718f-448b-8e11-4a8cca6d9004 8ff7ddfd-fbcb-4700-
ae52-6d071ac8d1b4 /root/examplecert.pem cycle
node2 /subscriptions/1e234d12-39b8-49db-881a-35aa03b402b9/resourceGroups/pcl/providers/
Microsoft.Compute/virtualMachines/node2 d5b7dac1-718f-448b-8e11-4a8cca6d9004 8ff7ddfd-fbcb-4700-
ae52-6d071ac8d1b4 /root/examplecert.pem cycle
```

Create /etc/opt/SMAW/SMAWsf/SA_vmazureReset.cfg and then set the owner, group, and access rights as follows.

```
# chown root:root /etc/opt/SMAW/SMAWsf/SA_vmazureReset.cfg
# chmod 600 /etc/opt/SMAW/SMAWsf/SA_vmazureReset.cfg
```

Note

- Make sure that the /etc/opt/SMAW/SMAWsf/SA_vmazureReset.cfg file is set correctly. If the setting is incorrect, the shutdown facility cannot be performed normally.
- Make sure that the resource ID (*ResourceID*) of the virtual machine corresponding to the CF node name (*CFNameX*) of the cluster host in the /etc/opt/SMAW/SMAWsf/SA_vmazureReset.cfg file is set. Also, make sure that the application ID (*AppID*), the tenant ID (*TenantID*), and the certificate path (*CertPath*) that were generated when registering the service principal and creating the certificate are set correctly. If the setting is incorrect, an incorrect node will be forcibly stopped.

3. Start the shutdown facility.

Check if the shutdown facility has been started on all nodes in the cluster system.

```
# sdtool -s
```

On a node where the shutdown facility has already been started, execute the following commands to restart the shutdown facility.

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

On a node where the shutdown facility has not been started, execute the following command to start the shutdown facility.

```
# sdtool -b
```

Information

You can check if the shutdown facility has already been started with the sdtool -s command. If "The RCSD is not running" is displayed, the shutdown facility is not started.

4. Check the status of the shutdown facility.

   Execute the following command on all nodes in the cluster system to check the status of the shutdown facility.

```
# sdtool -s
```

Note

If "The RCSD is not running" is displayed, the setting of the shutdown daemon or the setting of the shutdown agent is not correct. Perform the procedure from step 1 to 3 again.

Information

Display results of the sdtool -s command

- If Unknown or Init-ing is displayed in Init State, wait for about one minute, and then check the status again.
- If Unknown is displayed in Shut State, it means that SF has not yet stopped the node. If Unknown is displayed in Init State, it means that SF has not yet initialized SA or tested the route. Unknown is displayed temporarily in Test State or Init State until the actual state can be confirmed.
- If TestFailed is displayed in Test State, it means that a problem occurred while the agent was testing whether 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.

27.8.1.3 Initial Setup of the Cluster Resource Management Facility

Refer to "5.1.3 Initial Setup of the Cluster Resource Management Facility" in "PRIMECLUSTER Installation and Administration Guide" to set up the resource database managed by the cluster resource management facility. In this setting, set the iSCSI device used in the mirroring among the servers of GDS and register it to the resource database.

27.8.2 Setting up the Fault Resource Identification and Operator Intervention Request

Refer to "5.2 Setting up Fault Resource Identification and Operator Intervention Request" in "PRIMECLUSTER Installation and Administration Guide" to set up the fault resource identification and the operator intervention request.

27.9 Building the Cluster Application

For details on how to build the cluster application, refer to "Chapter 6 Building Cluster Applications" in "PRIMECLUSTER Installation and Administration Guide."

During this configuration, set the mirroring among servers of GDS (creating netmirror volume).

Also, network takeover must be registered to the cluster application. To build a cluster application for network takeover, refer to "27.9.1 Building the Cluster Application for Network Takeover."

Note

- Among the tuning parameters to be set in "Disk Setting for Performing Mirroring among Servers" in the "PRIMECLUSTER Global Disk Services Configuration and Administration Guide," set the following values for the following tuning parameters.

Tuning parameter name	Value after change
ED_CMD_RETRY_COUNT	100
ED_DRV_RETRY_COUNT	100

Example:

```
ED_CMD_RETRY_COUNT=100
ED_DRV_RETRY_COUNT=100
```

To extend the timeout period (CLUSTER_TIMEOUT) of the CF heartbeat, change the above parameter values according to the following formula. Round up the values after the decimal point.

Calculation formula:

```
<Increased CLUSTER_TIMEOUT> / 3 + 100
```

- If the icmp communication between cluster nodes is not allowed in the security group configuration, the following message is displayed when the clchkcluster command is executed.

```
Admin IP <IP address> used by SF is not alive.
```

If this message is output, refer to "26.3.2.1 Security Rules Applied to the Administrative LAN", and set the icmp protocol rule to allow the icmp communication between cluster nodes. After that, execute the clchkcluster command again.

- When executing the clchkcluster command, the following warning message is displayed. Ignore it.

Example) When LC_ALL is C

```
[warning]
The Shutdown Agent configuration is not redundant.
[action]
Add another agent so that a failed node can be eliminated even if the first agent fails to
eliminate the node.

SA configuration: Passed
```

- When using the Symfoware Server (Native) Hot Standby feature, the standby and scalable userApplication configurations are required.

Set the standby userApplication attribute to the following value.

Attribute	Setup value
StandbyTransitions	ClearFaultRequest | SwitchRequest

Refer to "6) Scalable operation" in "6.7 Setting Up Cluster Applications" in the "PRIMECLUSTER Installation and Administration Guide" for detailed build instructions.

27.9.1 Building the Cluster Application for Network Takeover

This section describes how to create the definition file and build the cluster application for network takeover.

27.9.1.1 Creating the Definition File

For network takeover, create the following definition file on all nodes controlling the network devices of Azure.

```
/usr/opt/reliant/etc/hvazurenicconfig
```

Information

When you create the /usr/opt/reliant/etc/hvazurenicconfigfile, you can use the /usr/opt/reliant/etc/hvazurenicconfig.template file as a template.

```
# cp -p /usr/opt/reliant/etc/hvazurenicconfig.template /usr/opt/reliant/etc/hvazurenicconfig
```

Note

Create the hvazurenicconfigfile with a root user and change the permission to 600.

Information in the definition file

```
TakeoverIPAddress CFName1 NICName1 CFName2 NICName2 ResourceGroup AppID TenantID CertPath
```

Item	Contents	Remarks
TakeoverIPAddress	Takeover IP address (Specify the IPv4 address.)	-
CFName1	Specify the CF node name.	-
NICName1	Specify the network interface name of the virtual machine in CFName1.	Check it in the Azure portal
CFName2	Specify the CF node name.	-
NICName2	Specify the network interface name of the virtual machine in CFName2.	Check it in the Azure portal
ResourceGroupName	Specify the resource group name.	Check it in the Azure portal
AppID	Specify the application ID of the service principal.	The value recorded in step 3 of "27.3 Presetting"
TenantID	Specify the tenant ID of the service principal.	The value recorded in step 3 of "27.3 Presetting"
CertPath	Specify the path of the service principal certificate.	The value recorded in step 3 of "27.3 Presetting"

Example

On each node, enter information on the operational system and the standby system in the created /usr/opt/reliant/etc/hvazurenicconfig file.

The following is an example when both 172.31.0.10 and 172.31.0.11 are controlled as the takeover IP address.

TakeoverIPAddress	CFName1	NICName1	CFName2	NICName2	ResourceGroupName	AppID	TenantID	CertPath
172.31.0.10	node0	Node0-eth1	node1	Node1-eth1	MyRG	d5b7dac1-718f-448b-8e11-4a8cca6d9004	8ff7ddfd-fbcb-4700-ae52-6d071ac8d1b4	/root/examplecert.pem
172.31.0.11	node0	Node0-eth1	node1	Node1-eth1	MyRG	d5b7dac1-718f-448b-8e11-4a8cca6d9004	8ff7ddfd-fbcb-4700-ae52-6d071ac8d1b4	/root/examplecert.pem

```
172.31.0.10 node0 Node0-eth1 node1 Node1-eth1 MyRG d5b7dac1-718f-448b-8e11-4a8cca6d9004 8ff7ddfd-
fbcb-4700-ae52-6d071ac8d1b4 /root/examplecert.pem
172.31.0.11 node0 Node0-eth1 node1 Node1-eth1 MyRG d5b7dac1-718f-448b-8e11-4a8cca6d9004 8ff7ddfd-
fbcb-4700-ae52-6d071ac8d1b4 /root/examplecert.pem
```

Note

- Separated by a space.

- For *CFName1* and *CFName2*, enter each CF node name of the operational system and the standby system. There are no restrictions on the order in which they are listed.

- Specify *TakeoverIPAddress* with IPv4 addresses, not host names.
- The contents of the definition file of the operational system and the standby system should be identical.
- If the definition file is not correct, or if the specified IP address is already in use on the Azure, UserApplication fails to start and an error message is output to the application log. For details on the messages, refer to "PRIMECLUSTER Messages."

27.9.1.2 Building the Cluster Application

This section describes how to register the network takeover to the cluster application.

The resources to be registered and their uses are as follows.

Registered target	Usage
PreOnline script of userApplication	Add the secondary IP address to the network interface of the virtual machine on the operational system.
Takeover network resources	In the operational system, the takeover IP address is activated.

Note

In an Azure environment, a virtual machine may pause for up to 30 seconds for maintenance without notice, causing a resource failure.

To prevent a resource failure during maintenance without notification, set the following parameters by adding 30 seconds to the calculated timeout period.

- ScriptTimeout for userApplication
- CheckCommandTimeouts for the Cmdline resource
- Timeout on takeover network resources

27.9.1.2.1 Setting PreOnline scripts for Network Takeover with Secondary IP address change

Register the following script in the PreOnlineScript of userApplication that registers the Takeover Network Resources.

```
/opt/SMAW/bin/hvazurenic
```

ScriptTimeout of userApplication should be greater than the following formula.

```
ScriptTimeout > Number of takeover IP address x ( 105 + retry interval (2seconds) ) (seconds)
```

Consider processing delays and perform verification. Then set the ScriptTimeout of userApplication high enough.

Note

If you change ScriptTimeout for userApplication, the following scripts change their timeout values to the same value.

- PreOnlineScript
- PostOnlineScript
- PreOfflineScript
- PostOfflineScript
- OfflineDoneScript
- FaultScript

27.9.1.2.2 Setting up the Takeover Network Resources

Static physical IP addresses

Perform the following steps on all nodes to ensure that the physical IP addresses of all network interfaces for which you want to configure a Takeover IP address are persistent.

1. Determine the connection profile name (NAME) for the network interface (DEVICE) for which you want to set the Takeover IP address by executing the following command.

```
# nmcli connection show
```

Example: Result displayed when the command is executed

```
# nmcli connection show
NAME    UUID            TYPE    DEVICE
ens256    9f45e94f-1726-dd68-8a33-8022f72b550f    ethernet    ens256
ens192    03da7500-2101-c722-2438-d0d006c28c73    ethernet    ens192
cip0    c24ca981-b8d6-487f-89ba-e73c58da349d    ethernet    cip0
  :
```

2. Execute the following command to check the ipv4.method value of the connection profile name that you saw in 1.

```
# nmcli connection show <connection profile name>
```

If the ipv4.method value is not manual, record the following values and perform the steps from 3.

- IP4.ADDRESS,IP4.GATEWAY,IP4.DNS

If the ipv4.method value is manual, no further steps are required.

Example: The connection profile name is ens192

```
# nmcli connection show ens192
connection.id:                   ens192
connection.uuid:                 03da7500-2101-c722-2438-d0d006c28c73
  :
ipv4.method:                     auto
  :
GENERAL.MASTER-PATH:             --
IP4.ADDRESS[1]:                  172.31.0.10/24
IP4.GATEWAY:                     172.31.0.1
IP4.ROUTE[1]:                    dst = 172.31.0.0/24, nh = 0.0.0.0, mt = 100
IP4.DNS[1]:                      172.31.0.2
```

3. Execute the following command to set the value of IP4.ADDRESS as a fixed value from 2.

```
# nmcli connection modify <connection profile name> ipv4.addresses <IP4.ADDRESS>
# nmcli connection modify <connection profile name> ipv4.method manual
```

Example: The connection profile name is ens 192 and the IP4.ADDRESS value is 172.31.0.10/24

```
# nmcli connection modify ens192 ipv4.addresses 172.31.0.10/24
# nmcli connection modify ens192 ipv4.method manual
```

4. Execute the following command to set the IP4.GATEWAY, IP4.DNS values from 2 and activate the connection profile name.

```
# nmcli connection modify <connection profile name> ipv4.gateway <IP4.GATEWAY>
# nmcli connection modify <connection profile name> ipv4.dns <IP4.DNS>
# nmcli connection up <connection profile name>
```

Example: The connection profile name is ens192, the IP4.GATEWAY value is 172.31.0.1, and the IP4.DNS value is 172.31.0.2

```
# nmcli connection modify ens192 ipv4.gateway 172.31.0.1
# nmcli connection modify ens192 ipv4.dns 172.31.0.2
# nmcli connection up ens192
```

5. Execute the following command and verify that the values for ipv4.addresses, ipv4.method, ipv4.gateway, ipv4.dns appear as configured in 3.4.

   ```
   # nmcli connection show <connection profile name>
   ```

   Example: The connection profile name is ens192

   ```
   # nmcli connection show ens192
   connection.id:                    ens192
   connection.uuid:                  03da7500-2101-c722-2438-d0d006c28c73
     :
   ipv4.method:                      manual
   ipv4.dns:                         172.31.0.2
   ipv4.dns-search:                  --
     :
   ipv4.addresses:                   172.31.0.10/24
   ipv4.gateway:                     172.31.0.1
   ```

Setting Takeover Network Resources Used for the Network Takeover by the Virtual Router

To set up the takeover network resources, refer to "6.7.3.6 Setting Up Takeover Network Resources" in "PRIMECLUSTER Installation and Administration Guide."

Note

The following steps are required to set up the takeover network resources.

- Set VIRTUAL (default value) for the BASE attribute and the VIRTUAL attribute of the takeover network resources.

About to define a target host(PingHosts)

- In an Azure Environment, virtual network gateways do not respond to pings. Therefore, you cannot define target hosts (PingHosts) and use ICMP to monitor network connectivity.

27.9.1.3 Checking the Network Takeover Settings

Execute the following command on all nodes and make sure that the information described in the definition file (/usr/opt/reliant/etc/hvazurenicconfig) is correctly described to control the device of Azure.

```
# /opt/SMAW/bin/hvazurechkconf
```

If there are no problems with the contents of the definition file, the display is as follows.

Example) The definition file is hvazurenicconfig

```
# /opt/SMAW/bin/hvazurechkconf
NOTICE: Check completed successfully. file=/usr/opt/reliant/etc/hvazurenicconfig
```

If there is a problem with the contents of the definition file, the setting value that must be checked is output. Follow the displayed message to take an action.

For details on the hvazurechkconf (8) command messages, refer to "PRIMECLUSTER Messages."

Chapter 28 Operations

For details on functions for managing PRIMECLUSTER system operations, refer to "Chapter 7 Operations" in "PRIMECLUSTER Installation and Administration Guide."

See

For details on how to operate GDS, refer to "Operation and Maintenance" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

Note

In an Azure environment, a heartbeat may fail due to an error of the network node or an error of the storage controller, or scheduled maintenance for the infrastructure. This may switch the cluster application.

Chapter 29 Changing the Configurations

For details on changing the configuration information for the PRIMECLUSTER system, environment settings, the configuration of the cluster application, the operation attributes of the cluster system, refer to "Chapter 9 Changing the Cluster System Environment", "Chapter 10 Configuration Change of Cluster Applications", "Chapter 11 Changing the Operation Attributes of a Cluster System" in "PRIMECLUSTER Installation and Administration Guide." For details on changing the GDS configuration, refer to "Configuration Change" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

29.1 Changing the Configuration of the Azure Environment

This section describes how to change the configuration of the Azure environment.

29.1.1 Updating the Certificate Used by the Service Principal

This section describes how to update the certificate used by the service principal.

1. Execute the following command on any one of the nodes in the cluster system to stop RMS.

```
# hvshut -a
```

2. Execute the following command on all nodes to stop the shutdown facility.

```
# sdtool -e
```

3. Update the certificate used by the service principal.

   3-1) Log in with an Azure account on any one node in the cluster.

```
# az login -u account
```

   Create a certificate to be used by the service principal.
   Example) When appID is "d5b7dac1-718f-448b-8e11-4a8cca6d9004"

```
# az ad sp credential reset --name d5b7dac1-718f-448b-8e11-4a8cca6d9004 --create-cert
{
  "appId": "d5b7dac1-718f-448b-8e11-4a8cca6d9004",
  "fileWithCertAndPrivateKey": "/root/tmphzHer5.pem",
  "name": "d5b7dac1-718f-448b-8e11-4a8cca6d9004",
  "password": null,
  "tenant": "8ff7ddfd-fbcb-4700-ae52-6d071ac8d1b4"
}
```

   For the --name option, you can also specify the name when registering the service principal.
   For details on registering service principals, refer to step 3 of "27.3 Presetting."

   Log out from Azure.

```
# az logout
```

   3-2) Store the certificate created in fileWithCertAndPrivateKey in the location specified by CertPath in step 2 of "27.8.1.2 Setting up the Shutdown Facility" on all cluster nodes, and set the permissions to 600.
   Example) When CertPath is "/root/examplecert.pem"

```
# cp /root/tmprjbQbI.pem /root/examplecert.pem
# chmod 600 /root/examplecert.pem
```

   Delete the certificate created in fileWithCertAndPrivateKey using the rm command.

   3-3) Make sure that you can log in with the service principal using the updated certificate on all cluster nodes.

```
# az login --service-principal --username appID --tenant tenant --password CertPath confirmed in
3-2)
```

3-4) Log out from Azure on all cluster nodes.

```
# az logout
```

4. Execute the following command on all nodes to start the shutdown facility.

```
# sdtool -b
```

5. Execute the following command on all nodes and make sure that the shutdown facility operates normally.

```
# sdtool -s
```

6. Execute the following command on any one of the nodes in the cluster system to start RMS.

```
# hvcm -a
```

7. Execute the following command on any one of the nodes in the cluster system and make sure that RMS operates normally.

   If RMS is stopped, "RMS is not running" is output.

```
# hvdisp -a
```

29.1.2 Changing the Azure CLI Resources

This section describes how to change the Azure CLI resources (virtual machine resource ID, application ID, tenant ID, certificate path).

1. Execute the following command on any one of the nodes in the cluster system to stop RMS.

```
# hvshut -a
```

2. Execute the following command on all nodes to stop the shutdown facility.

```
# sdtool -e
```

3. Modify the configuration definition file of the Azure shutdown agent on all nodes.

   For the description of the configuration definition file, refer to step 2 in "27.8.1.2 Setting up the Shutdown Facility."

4. Modify the definition file for controlling network devices of Azure on all nodes.

   For the description of the definition file, refer to "27.9.1.1 Creating the Definition File."

5. Execute the following command on all nodes to start the shutdown facility.

```
# sdtool -b
```

6. Execute the following command on all nodes and make sure that the shutdown facility operates normally.

```
# sdtool -s
```

7. Execute the following command on any one of the nodes in the cluster system to start RMS.

```
# hvcm -a
```

8. Execute the following command on any one of the nodes in the cluster system and make sure that RMS operates normally.

   If RMS is stopped, "RMS is not running" is output.

```
# hvdisp -a
```

29.1.3 Changing the Role of the Service Principal

When changing the role of a service principal, note the following.

For the Azure resources (ResourceID and RouteTableID) specified in SA_vmazureReset.cfg and hvazurenicconfig, set a role that has privileges for the following operations.

- Microsoft.Compute/virtualMachines/read
- Microsoft.Compute/virtualMachines/start/action
- Microsoft.Compute/virtualMachines/powerOff/action
- Microsoft.Network/routeTables/read
- Microsoft.Network/routeTables/routes/read
- Microsoft.Network/routeTables/routes/write
- Microsoft.Network/networkInterfaces/read
- Microsoft.Network/networkInterfaces/write
- Microsoft.Network/virtualNetworks/subnets/join/action

Note that depending on your environment, a role may require additional privileges.

Set permissions on the role if:.

If you assign a public IP to the network interface:

Microsoft.Network/publicIPAddresses/join/action

If you have associated a network security group with a network interface:

Microsoft.Network/networkSecurityGroups/join/action

Chapter 30 Maintenance

When you maintain the PRIMECLUSTER system in an Azure environment, note the following points.

- For the procedure on applying/deleting urgent corrections in an Azure environment, refer to "30.1 Software Maintenance."

- For details on other items and procedures required for maintenance of the PRIMECLUSTER system, refer to "Chapter 12 Maintenance of the PRIMECLUSTER System" in "PRIMECLUSTER Installation and Administration Guide." For details on how to maintain GDS, refer to "Operation and Maintenance" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

30.1 Software Maintenance

30.1.1 Notes on Applying Corrections to the PRIMECLUSTER System

For details on notes for applying an intensive correction to the cluster system, refer to "12.3.1 Notes on Applying Corrections to the PRIMECLUSTER System" in "PRIMECLUSTER Installation and Administration Guide."

Note

In an Azure environment, refer to "30.1.2 Overview of the Procedure for Applying/Deleting Corrections" to apply/delete the corrections in multi-user mode.

30.1.2 Overview of the Procedure for Applying/Deleting Corrections

Overview of the procedure is shown for applying each correction including an intensive correction to the cluster system in an Azure environment. In an environment that does not use GDS, the procedure related to GDS is not necessary.

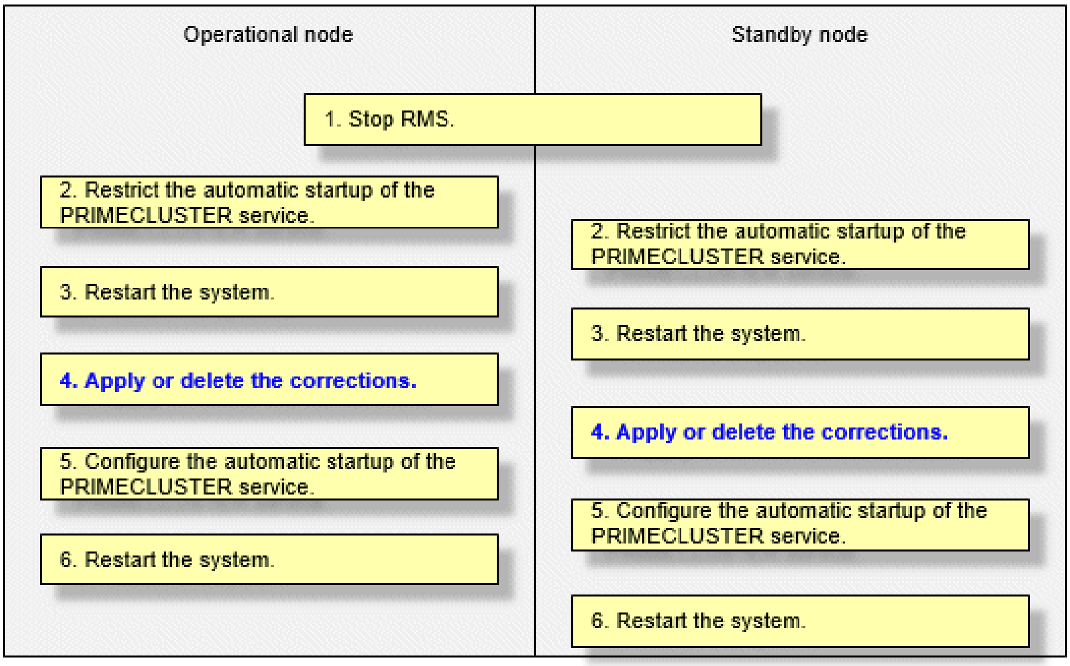
Note

Before applying or deleting corrections to PRIMECLUSTER, take a snapshot of the system storage.

30.1.2.1 Procedure for Applying or Deleting Corrections by Stopping the Entire System

This section describes the procedure for applying/deleting corrections by stopping the entire cluster system.

Flow of the operation


Operational node
Standby node
1. Stop RMS.
2. Restrict the automatic startup of the PRIMECLUSTER service.
3. Restart the system.
4. Apply or delete the corrections.
5. Configure the automatic startup of the PRIMECLUSTER service.
6. Restart the system.
2. Restrict the automatic startup of the PRIMECLUSTER service.
3. Restart the system.
4. Apply or delete the corrections.
5. Configure the automatic startup of the PRIMECLUSTER service.
6. Restart the system.

Operation procedure

Copy the corrections to be applied to each node to the local file system in advance.

1. Stop RMS.

   If RMS is running, execute the following command on any one node in the cluster system to stop RMS.

   ```
   # hvshut -a
   ```

   Note

   If RMS is stopped on all nodes during the synchronization copying of the GDS volume, the synchronization copying of the entire volume area is performed after the corrections are applied and all nodes are restarted.

   If you do not want to perform the synchronization copying of the entire area of volume, stop RMS after the synchronization copying is completed.

   To check the slice status of the GDS volume, execute the following command.

   Execute the following command on any one node in the cluster system to check the value of the STATUS field of the command output.

   The status of the copy destination slice is COPY during the synchronization copying, and after copying is complete, the status becomes ACTIVE or STOP.

   ```
   # sdxinfo -S
   ```

2. Restrict the automatic startup of the PRIMECLUSTER service.

   Restrict the automatic startup of the PRIMECLUSTER service by executing the following command on all nodes.

   ```
   # /opt/FJSVpclinst/bin/pclservice off
   ```

3. Restart the system.

   Restart the system on all nodes.

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

4. Apply or delete the corrections.

   Apply the corrections that were copied to the local file system on all nodes, or delete the corrections.

   - Applying corrections

     Copy the corrections to the working directory and then execute the following commands.

     ```
     # cd <working directory>
     # /opt/FJSVfupde/bin/uam add -d ./ -i <correction number>
     ```

     If the following message is displayed, select "Y".

     ```
     It is required to update with single user mode. Do you want to apply the update now? (Y/N)Y
     ```

     After that, the following message is displayed. Select "N".

     ```
     Do you want to restart your computer immediately? (Y/N)N
     ```

   - Deleting corrections

     Execute the following command.

     ```
     # /opt/FJSVfupde/bin/uam remove -i <correction number>
     ```

     If the following message is displayed, select "Y".

     ```
     It is required to restore with single user mode. Do you want to restore the updated product
     to its pre-update state now? (Y/N)Y
     ```

     After that, the following message is displayed. Select "N".

     ```
     Do you want to restart your computer immediately? (Y/N)N
     ```

5. Configure the automatic startup of the PRIMECLUSTER service.

   Execute the following command on all nodes and change the PRIMECLUSTER service settings back to the state they were in before they were restricted in step 2.

   ```
   # /opt/FJSVpclinst/bin/pclservice on
   ```

6. Restart the system.

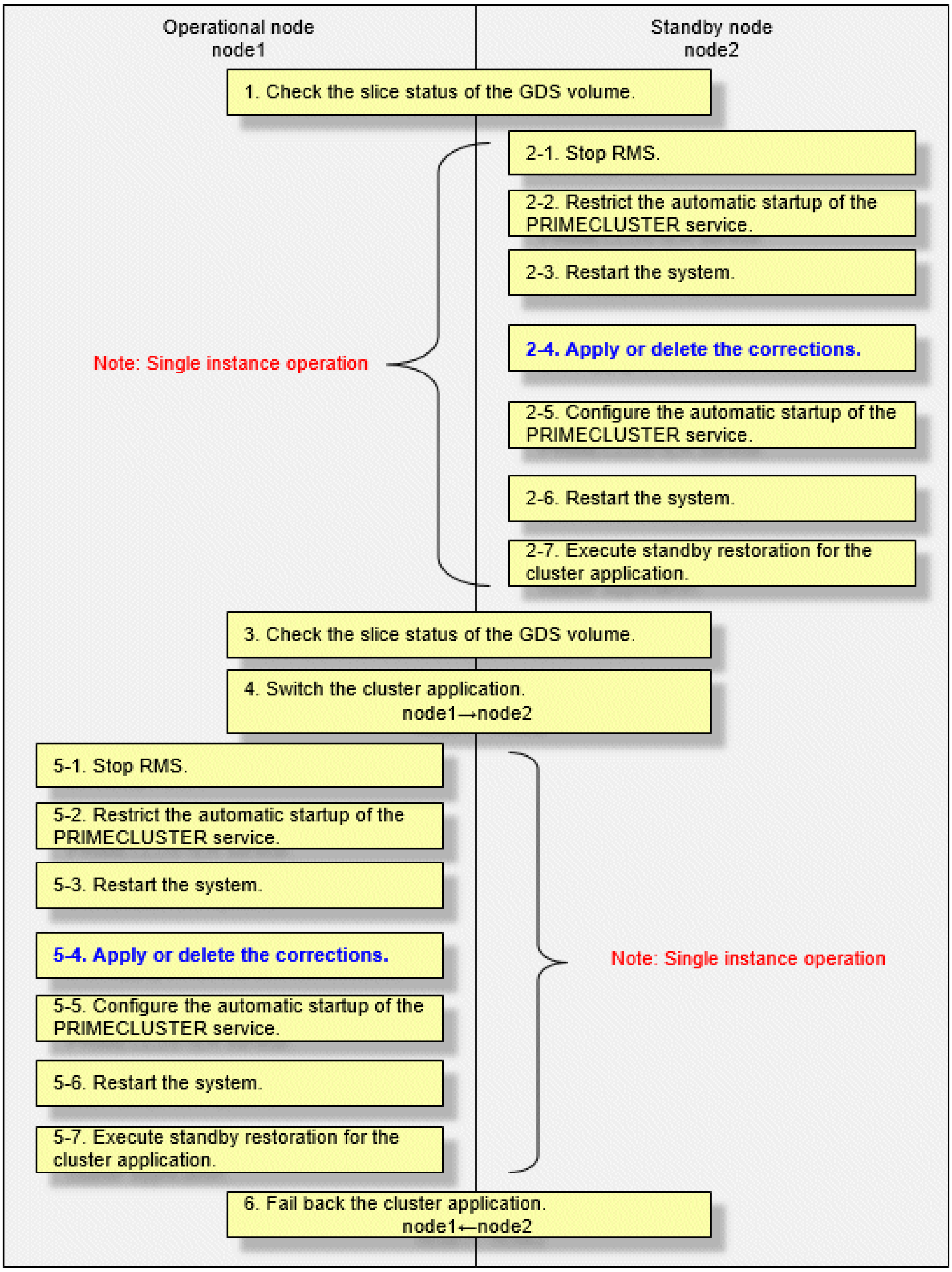
   Restart the system on all nodes.

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

30.1.2.2 Procedure for Applying or Deleting Corrections by Rolling Update

This section describes the procedure for applying corrections by rolling update.

Flow of the operation


Operational node
node1
Standby node
node2
1. Check the slice status of the GDS volume.
2-1. Stop RMS.
2-2. Restrict the automatic startup of the PRIMECLUSTER service.
2-3. Restart the system.
2-4. Apply or delete the corrections.
Note: Single instance operation
2-5. Configure the automatic startup of the PRIMECLUSTER service.
2-6. Restart the system.
2-7. Execute standby restoration for the cluster application.
3. Check the slice status of the GDS volume.
4. Switch the cluster application.
node1→node2
5-1. Stop RMS.
5-2. Restrict the automatic startup of the PRIMECLUSTER service.
5-3. Restart the system.
5-4. Apply or delete the corrections.
Note: Single instance operation
5-5. Configure the automatic startup of the PRIMECLUSTER service.
5-6. Restart the system.
5-7. Execute standby restoration for the cluster application.
6. Fail back the cluster application.
node1←node2

GDS: Global Disk Services

Operation procedure

1. Check the slice status of the GDS volume.

   Execute the following command on any cluster node to check the value of the STATUS field of the command output.

   ```
   # sdxinfo -S
   ```

   If the COPY status slice exists in the netmirror volume, wait until the synchronization copying is complete.

   For problems caused by node operations during copying, refer to "Stopping or Restarting the Node" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide."

2. Execute the following operation with the standby node (node2).

    1. Stop RMS.

        Stop RMS to apply corrections to the standby node (node2). A cutoff state transition occurs according to the shutdown of RMS. In this case, make sure that the single instance operation continues until the standby restoration for the cluster application is executed.

        ```
        # hvshut -l
        ```

    2. Restrict the automatic startup of the PRIMECLUSTER service.

        Execute the following command to restrict the automatic startup of the PRIMECLUSTER service.

        ```
        # /opt/FJSVpclinst/bin/pclservice off
        ```

    3. Restart the system.

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

    4. Apply or delete the corrections.

        - Applying corrections

        Copy the corrections to the working directory and then execute the following commands.

        ```
        # cd <working directory>
        # /opt/FJSVfupde/bin/uam add -d ./ -i <correction number>
        ```

        If the following message is displayed, select "Y".

        ```
        It is required to update with single user mode. Do you want to apply the update now? (Y/N)Y
        ```

        After that, the following message is displayed. Select "N".

        ```
        Do you want to restart your computer immediately? (Y/N)N
        ```

        - Deleting corrections

        Execute the following command.

        ```
        # /opt/FJSVfupde/bin/uam remove -i <correction number>
        ```

        If the following message is displayed, select "Y".

        ```
        It is required to restore with single user mode. Do you want to restore the updated
        product to its pre-update state now? (Y/N)Y
        ```

        After that, the following message is displayed. Select "N".

        ```
        Do you want to restart your computer immediately? (Y/N)N
        ```

    5. Configure the automatic startup of the PRIMECLUSTER service.

        Execute the following command and change the PRIMECLUSTER service settings back to the state they were in before they were restricted in 2 of step 2.

        ```
        # /opt/FJSVpclinst/bin/pclservice on
        ```

    6. Restart the system.

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

    7. Execute standby restoration for the cluster application.

        If the node (node1) to which corrections have been applied is cut off from the cluster system, execute standby restoration for the node.

For details on how to execute cluster application standby restoration, refer to "7.2.2.1 Starting a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

3. Check the slice status of the GDS volume.

   After starting the standby node (node2), the synchronization copying of the netmirror volume is executed. Make sure that the synchronization copying is completely finished and all slices are either in ACTIVE or STOP status on any one node.

   To check the slice status of the netmirror volume, execute the following command:

   Execute the following command on any cluster node to check the value of the STATUS field of the command output.

   ```
   # sdxinfo -S
   ```

4. Switch the cluster application.

   To apply corrections to the operational node (node1), execute hvswitch and switch all cluster applications to the standby node (node2). For details on how to switch the cluster applications, refer to "7.2.2.3 Switching a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

5. Perform the following operation with the operational node (node1).

   1. Stop RMS.

      Stop RMS to apply corrections to the operational node (node1). A cutoff state transition occurs according to the shutdown of RMS. In this case, make sure that the single instance operation continues until the standby restoration for the cluster application is executed.

      ```
      # hvshut -l
      ```

   2. Restrict the automatic startup of the PRIMECLUSTER service.

      Execute the following command to restrict the automatic startup of the PRIMECLUSTER service.

      ```
      # /opt/FJSVpclinst/bin/pclservice off
      ```

   3. Restart the system.

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

   4. Apply or delete the corrections.

      - Applying corrections

      Copy the corrections to the working directory and then execute the following commands.

      ```
      # cd <working directory>
      # /opt/FJSVfupde/bin/uam add -d ./ -i <correction number>
      ```

      If the following message is displayed, select "Y".

      ```
      It is required to update with single user mode. Do you want to apply the update now? (Y/N)Y
      ```

      After that, the following message is displayed. Select "N".

      ```
      Do you want to restart your computer immediately? (Y/N)N
      ```

      - Deleting corrections

      Execute the following command.

      ```
      # /opt/FJSVfupde/bin/uam remove -i <correction number>
      ```

      If the following message is displayed, select "Y".

      ```
      It is required to restore with single user mode. Do you want to restore the updated
      product to its pre-update state now? (Y/N)Y
      ```

      After that, the following message is displayed. Select "N".

```
Do you want to restart your computer immediately? (Y/N)N
```

    5. Configure the automatic startup of the PRIMECLUSTER service.

       Execute the following command and change the PRIMECLUSTER service settings back to the state they were in before they were restricted in 2 of step 5.

       ```
       # /opt/FJSVpclinst/bin/pclservice on
       ```

    6. Restart the system.

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

    7. Execute standby restoration for the cluster application.

       If the node (node1) to which corrections have been applied is cut off from the cluster system, execute standby restoration for the node. For details on how to execute cluster application standby restoration, refer to "7.2.2.1 Starting a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

6. Fail back the cluster application.

   Restore the state of the standby layout defined at installation by executing failback operation, as necessary. For details on failback, refer to "7.2.2.3 Switching a Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

30.2 Procedure for Restoring OS with the Snapshot Function

If the OS is restored with the snapshot function of Azure, perform the following procedure.

30.2.1 Procedure for Restoring One Node While the Operation is Working

1. Restore the virtual machine. To restore the snapshot or attach the Azure managed disk volumes, refer to the official Azure documentation.

   Note

   Instead of creating a new virtual machine, replace the system volume of the restoration target virtual machine with the restored volume.

2. Check the slice status of GDS. If the status of the slice is INVALID, execute the following command to perform the synchronization copying of each volume after the node is started. Perform this procedure on either node.

   ```
   # sdxcopy -B -c <class name> -v <volume name>
   ```

3. If the virtual machine name is changed in step 1, perform steps 2 to 4 described in "27.8.1.2 Setting up the Shutdown Facility" on all nodes, and modify the ResourceID of the hvazurenicconfig file created in "27.9.1.1 Creating the Definition File."

30.2.2 Procedure for Restoring Nodes While the Operation does not Work

1. If either or all the nodes are started before restoring the nodes, stop RMS. Perform the restoration procedure on either node that is started.

   ```
   # hvshut -a
   ```

2. If all nodes are running before the restoration, select the latest disk. For all classes of GDS, execute the following command on either node.

   ```
   # /etc/opt/FJSVsdx/bin/sdxnetdisk -S -c <class name>
   ```

3. Restore the virtual machine. To restore the snapshot or attach the Azure managed disk volumes, refer to the official Azure documentation.

Note

Instead of creating a new virtual machine, replace the system volume of the restoration target virtual machine with the restored volume.

4. After starting the node, execute the following on the restored node.

   1. Delete the information of the iSCSI device.

      ```
      # rm -f /var/opt/FJSVsdx/log/.sdxnetmirror_disable.db
      # rm -f /var/opt/FJSVsdx/log/.sdxnetmirror_timestamp
      ```

   2. Stop RMS.

      ```
      # hvshut -l
      ```

5. Restore the other node if necessary.

Note

Instead of creating a new virtual machine, replace the system volume of the restoration target virtual machine with the restored volume.

6. When restoring the node in step 5, delete the iSCSI device information on the restored node.

   ```
   # rm -f /var/opt/FJSVsdx/log/.sdxnetmirror_disable.db
   # rm -f /var/opt/FJSVsdx/log/.sdxnetmirror_timestamp
   ```

7. If all nodes have stopped before the restoration, check the status of the slice that is the copy source of the synchronization copy. If the source slice for the synchronization copying is INVALID, restore the status of the slice. For the "-d" option of the "sdxfix " command, specify the source disk of the synchronization copying. Perform this procedure on either node.

   ```
   # sdxfix -V -c <class name> -v <volume name> -d <disk name> -x NoRdchk
   ```

8. If the virtual machine name is changed in step 3 or step 5, perform step 2 to 4 described in "27.8.1.2 Setting up the Shutdown Facility" on all nodes, and modify the ResourceID of the hvazurenicconfig file created in "27.9.1.1 Creating the Definition File."

9. If 2 of step 4 is performed, start RMS on the node where RMS is stopped.

   ```
   # hvcm
   ```

Appendix A Using the Smart Workload Recovery Feature

Note

The smart workload recovery feature is not available in this version.

This part describes the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system using the smart workload recovery feature provided by PRIMECLUSTER Cloud Edition.

The smart workload recovery feature is a standby-less switching feature available in AWS environments.

By linking the HA manager (RMS) of PRIMECLUSTER with the services such as serverless infrastructure and resource monitor provided in the AWS environment, you can improve the availability of business operations by switching instances not only in the case of an abnormality of an instance that can be switched by the standard HA function of the AWS environment but also in the case of a cluster application abnormality or Availability Zone (AZ) failure.

For an overview of the smart workload recovery feature, see the "PRIMECLUSTER Concept Guide".

Supported Configurations

- Network Configuration
  - Taking over the IP address between the instances using ELB
- Configurations to takeover data
  - Data takeover by launching a new instance using Amazon Machine Image (hereinafter AMI)
  - Shared data takeover by using Amazon Elastic File System (EFS)

Supported monitoring features

- Failure of the cluster application

  When a failure or a hang-up of the cluster application occurs, launch the instance in a different AZ.

- Failure of the instance

  When an abnormal outages, hangs, or panics of the instance occurs, launch the instance in a different AZ.

Prerequisites

- The instance can communicate with the API endpoint.

Conditions for the operation

- Supported cloud platform: AWS
- Supported regions: ap-northeast-1, ap-northeast-3

Cloud services required for smart workload recovery feature

The smart workload recovery feature uses the following cloud services.

- Amazon EC2
- AWS IAM
- Amazon Elastic Block Store
- Amazon CloudWatch
- Amazon CloudWatch Logs
- Amazon EventBridge

- AWS CloudTrail
- AWS Lambda
- Elastic Load Balancing (excluding CLB)
- Amazon EFS
- Amazon DynamoDB
- AWS PrivateLink

Note

- The console cannot be used in the instance in an AWS environment. Do not set the single user mode.
- For the most up-to-date information on requirements, go to our company Web site.
- The smart workload recovery feature uses AWS APIs to provide switching. For this reason, using AWS APIs to successfully interact with the cloud service is a prerequisite for the switch to work correctly.

A.1 Design

You must prepare the items listed below before building the PRIMECLUSTER system that uses the smart workload recovery feature.

- Selecting the PRIMECLUSTER Product
- System Configuration
- Network Design
- Disk Design
- Policy Design
- System Design

Point

For an overview of the smart workload recovery feature, see the "PRIMECLUSTER Concept Guide".

A.1.1 Selecting the PRIMECLUSTER Product

Select a PRIMECLUSTER products.

When using the smart workload recovery feature, you can select the following products.

For details on the PRIMECLUSTER products, refer to "2.1 PRIMECLUSTER Product Selection" in the "PRIMECLUSTER Installation and Administration Guide".

- PRIMECLUSTER Cloud Edition

A.1.2 System Configuration

This subsection describes the system configuration of the smart workload recovery feature. Design and deploy your AWS environment based on your system configuration. For deployment instructions, see "A.2 Installation".

Point

For an overview of smart workload recovery feature, see "1.9 Smart workload recovery" in the "PRIMECLUSTER Concept Guide".

For more information about AWS resources and services, see the official AWS documentation.

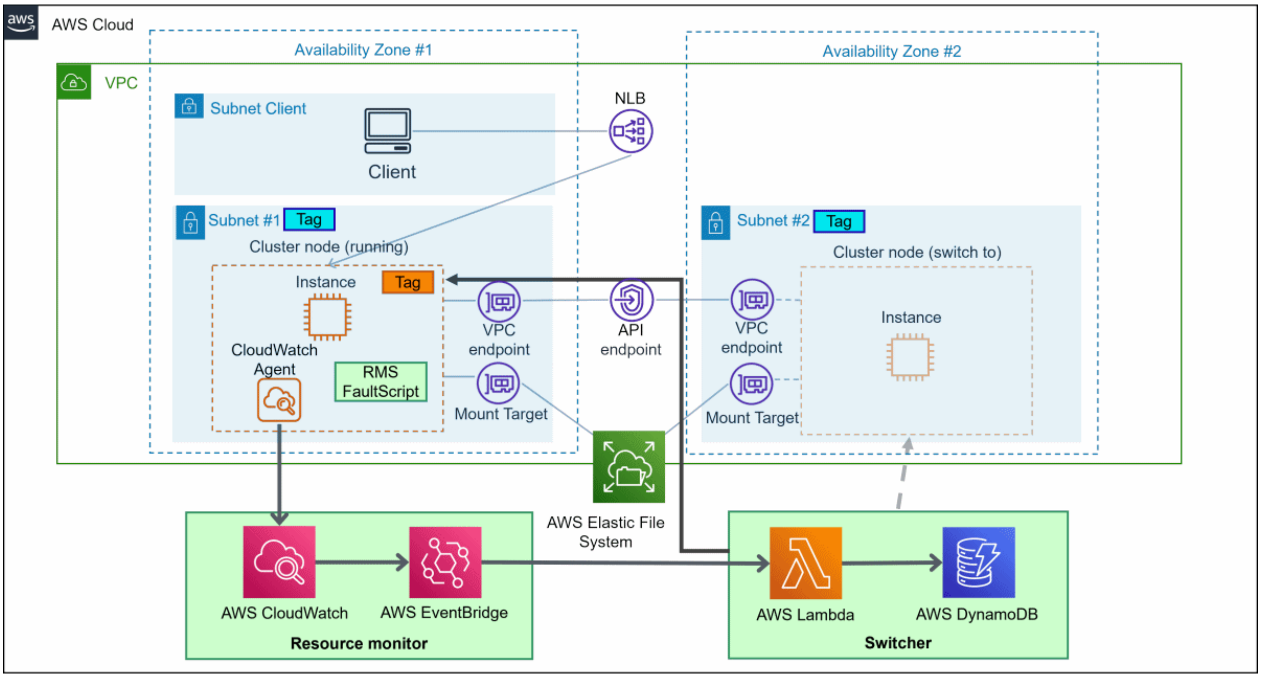
System configuration

Smart workload recovery is a single-node cluster system that is operational only and does not have a standby system. Remove the instance from the AZ in which the cluster node instance is running in the event of failure and launch a new instance in another AZ.

Before you can switch to an AZ, you must create a network environment, such as a subnet.

Resource monitor and switcher work with services such as AWS CloudWatch and AWS Lambda to provide instance switching. The tag identifies the cluster node instance and the subnet to which it switches. Therefore, you set the specified tags on your instances and subnets.

Figure A.1 System configuration



Component Description

A description of each component is provided below. It also provides an overview of the tasks required to design and deploy an AWS environment. The tasks described below that are required at the time of implementation are performed in "A.2 Installation" and later.

- Network

  Network configuration involves deploying an Elastic Load Balancing (ELB) Network Load Balancer (NLB) or an Application Load Balancer (ALB) for network handover by creating a subnet for each AZ that you want to switch over in a single Amazon Virtual Private Cloud (VPC). It also ensures connectivity with API endpoints so that CloudWatch Agent can access Amazon CloudWatch services.

  Tag the subnet so that switcher recognizes it as the subnet on which you want to redeploy the instance. Subnet tags are described in "A. 2.1.1 Creating VPC and Subnet".

  Note

  You can select either NLB or ALB for network takeover. The design and implementation diagrams are illustrated using NLB as an example.

- Security groups

  Design your security groups according to your security requirements.

  You also have a blackhole security group as a security group that blocks all traffic. Switcher removes an instance from the AZ in which the cluster node instance is running in the event of failure and launches a new instance in a different AZ. Quarantine a cluster node by configuring a blackhole security group on the instance that you want to delete.

- Disk

  You place the data that you want to share on EFS. Create a mount target for EFS for each AZ that you want to switch.

  The deployment creates an EFS file system to store RMS logs. If you want to store data that users share, create a separate file system for EFS.

- Instance

  Prepare an AMI for the instance that will be your cluster node. Create an instance of the AMI, install RMS of PRIMECLUSTER on the instance to detect error in the application, and create and register FaultScript. You also install a CloudWatch Agent on the instance to pass metrics and logs for the instance to Amazon CloudWatch, resource monitor. After you install the required software, you create an image of your instance.

  Tag the instances so that switcher can recognize them.

- Resource monitor

  Resource monitor is collectively referred to as Amazon CloudWatch and Amazon EventBridge. It works with Amazon CloudWatch to collect metrics and logs from your instances. It also works with Amazon EventBridge to notify switcher of instance error as events.

  The deployment involves adding CloudWatch Alarm and Amazon EventBridge rules.

- Switcher

  Switcher detects failure in the target instance with an event in Amazon EventBridge. After the event is detected, AWS Lambda redeploys the instance to the subnet in the unhealthy AZ. You store the data held by switcher in Amazon DynamoDB.

  The deployment involves registering your AWS Lambda function and your Amazon DynamoDB table.

Point

About Creating Multiple Resources

When you design a smart workload recovery feature, you might have multiple resources of the same type. The situation is described below. Here's how to design a smart workload recovery feature in that situation.

- Region

  If you want to use multiple regions, design them as follows.

  - Amazon DynamoDB as a Switcher

    You must have one Amazon DynamoDB table for each region.

  - Tags to set for instances on cluster nodes ([fujitsu.pclswr.id] key)

    The [fujitsu.pclswr.id] key in the tag must be a unique integer value for each instance. Unique integer values must be unique for each region.

- VPC

  If you use multiple VPCs, design them as follows.

  - Blackhole Security Group

    You must have one Blackhole security group for each VPC.

  - Switcher for AWS Lambda

    You must have one AWS Lambda function for each VPC.

  - Switcher for Amazon EventBridge

    You must have one set of Amazon EventBridge events (two events) for each VPC.

- Cluster node instances

  When using instances of multiple cluster nodes, design as follows.

  - ELB

    Prepare a NLB of ELB or ALB for an instance.

- EFS

  Prepare an EFS to store the data that you want to share for an instance.

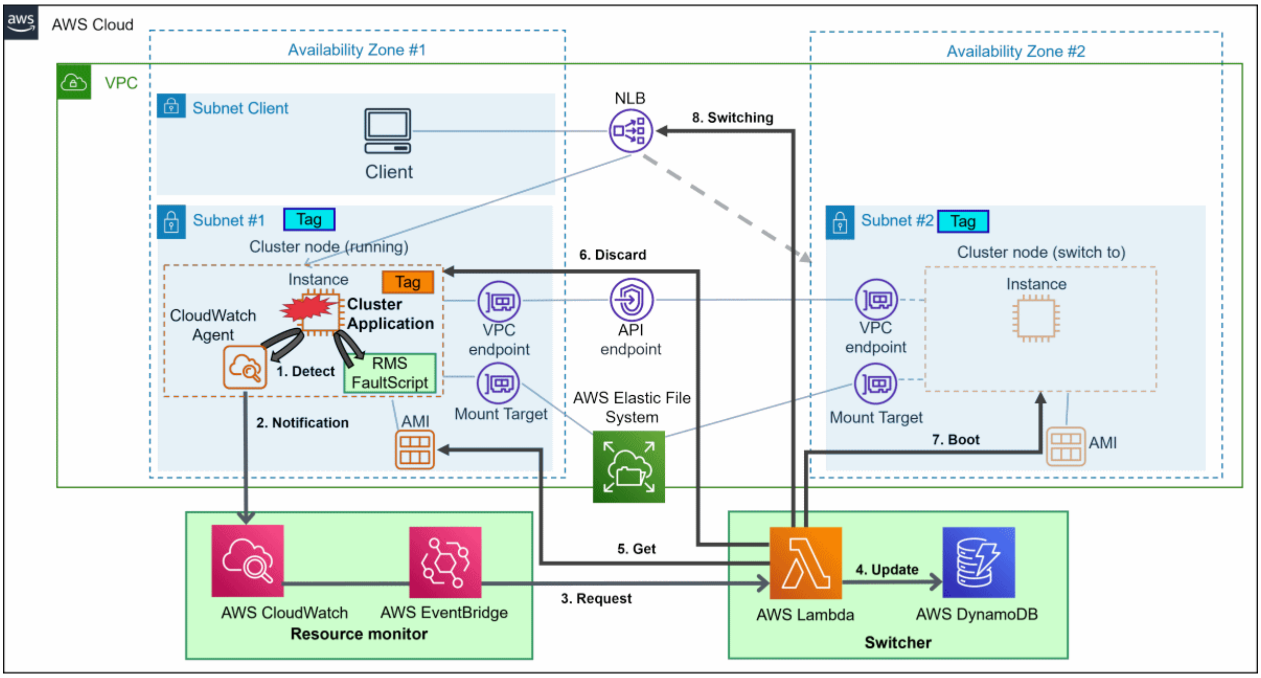
- Amazon CloudWatch for Resource monitor

  You must have one set of Amazon CloudWatch Alarm (two alarms) for each instance.

Operation

The following describes the behavior of each component when a cluster application error/RMS error occurs in a system using the smart workload recovery feature in an AWS environment.

Figure A.2 How smart workload recovery works in AWS environment



When error occurs in a cluster application

1. RMS detects cluster application error
2. RMS notifies resource monitor by running FaultScript and shutting down the instance
3. When resource monitor receives an error, it requests switcher to switch.
4. Switcher Updates tables in Amazon DynamoDB
5. Switcher retrieves the AMI of the instance from which it was switched
6. Switcher destroys the switching instance
7. Switcher launches instances in any AZ that is different from the source
8. Switcher switches instances of load balancer target groups

When RMS error occurs

1. CloudWatch Agent detects RMS error
2. CloudWatch Agent notifies resource monitor
3. When resource monitor receives an error, it requests switcher to switch
4. Switcher updates tables in Amazon DynamoDB
5. Switcher retrieves the AMI of the instance from which it was switched

6. Switcher destroys the switching instance

7. Switcher launches instances in any AZ that is different from the source

8. Switcher switches instances of load balancer target groups

The instance inherits the following when you switch.

- AMI ID
- Instance type
- Key pair name
- Security group ID
- Tags
- IAM Role
- CloudWatch Alarm
- RMS log
- User's shared data

Note

When you switch instances, only CloudWatch Alarm required by resource monitor are guaranteed to be carried over to the switched instance.

A.1.3 Network Design

The following describes the network configuration required to operate smart workload recovery. Design the network based on the network configuration.

- VPC and Subnet Design
- Ensuring Connectivity with API Endpoints
- Identifying Mount Targets for EFS
- Network takeover
- Designing Security Groups

Note

Because the smart workload recovery feature creates an instance when you switch, the switch can be interrupted if your AWS environment runs out of AZ resources.

Because AZ resource exhaustion cannot be predicted in advance, we recommend that you design the following items with all availability zones in the region in mind, rather than limiting the switched to only one availability zone.

- Subnet Design
- Ensuring Connectivity with API Endpoints
- Identifying Mount Targets for EFS
- Network Inheritance

In the event of AZ resource exhaustion, it is necessary to take measures to add a new switching destination AZ in accordance with "A.3.2.3 Changing the settings of the switch destination AZ".

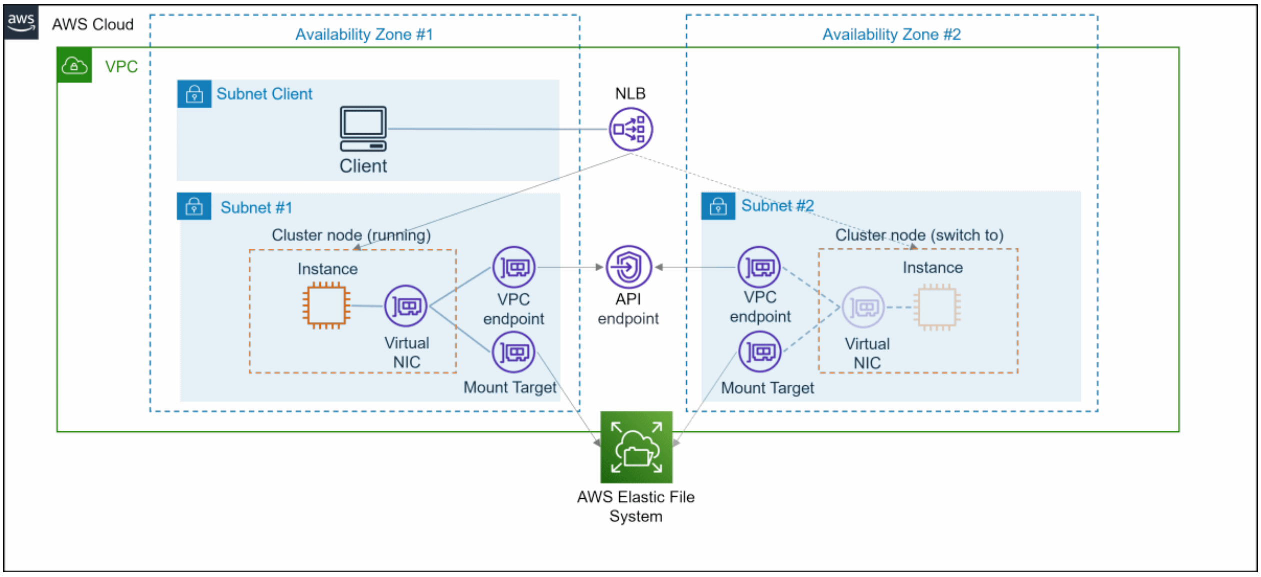
A.1.3.1 VPC and Subnet Design

Prepare a VPC and create a subnet for each AZ that you want to switch.

Configure your VPC to use DNS. When you create a VPC, you specify the following options. For more information, see the official AWS documentation.

Setting item	Value
DNS hostname	Enabled
DNS resolution	Enabled

Figure A.3 Creating Subnets per AZ



A.1.3.2 Ensuring Connectivity with API Endpoints

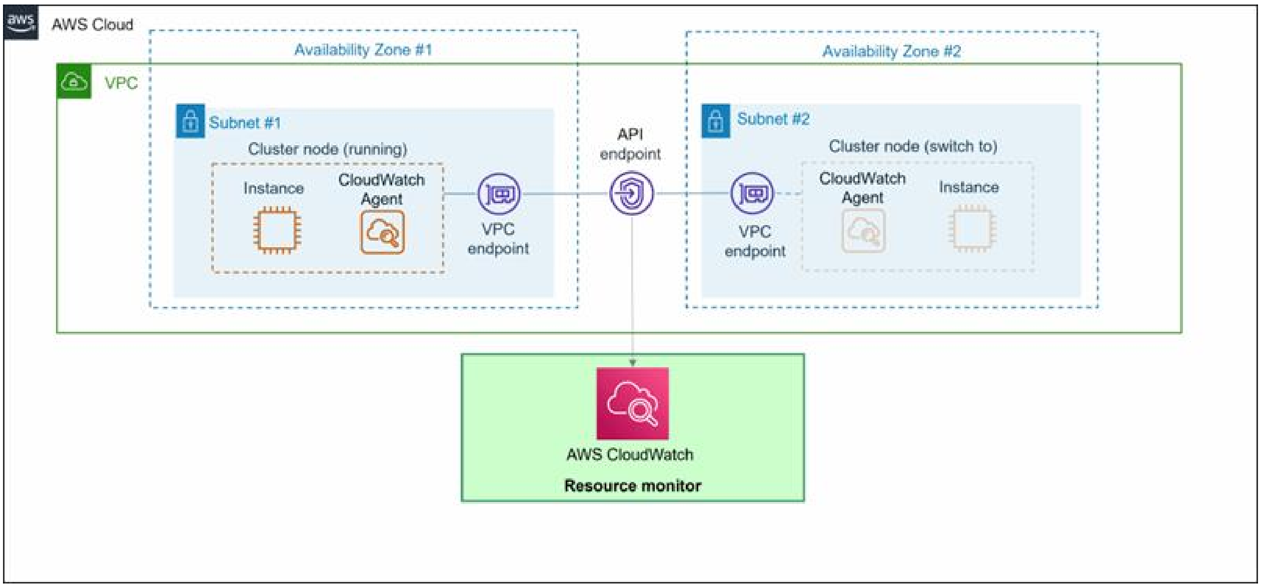
To access the API endpoint, deploy the AWS PrivateLink VPC endpoint in all AZ where you deploy cluster nodes.

CloudWatch Agent installed on your instance forwards metrics and logs to the Amazon CloudWatch API endpoint.

If you are creating a VPC endpoint for PrivateLink, specify the following options. For more information, see the official AWS documentation.

Setting item	Value	Description
Service name	monitoring ec2 logs	Specifies the service that the CloudWatch agent uses.
Private DNS names enabled	Enable	Configure DNS for use.

Figure A.4 Ensuring Connectivity with AWS PrivateLink



A.1.3.3 Verifying EFS Mount Targets

To use EFS from an instance on a cluster node, you must have an EFS mount target on the subnet of all AZ where you deploy the cluster node.

Ensure that you have an EFS mount target for each EFS file system that is required to operate smart workload recovery. For information about EFS file systems required to operate smart workload recovery, see "A.1.4 Disk Design".

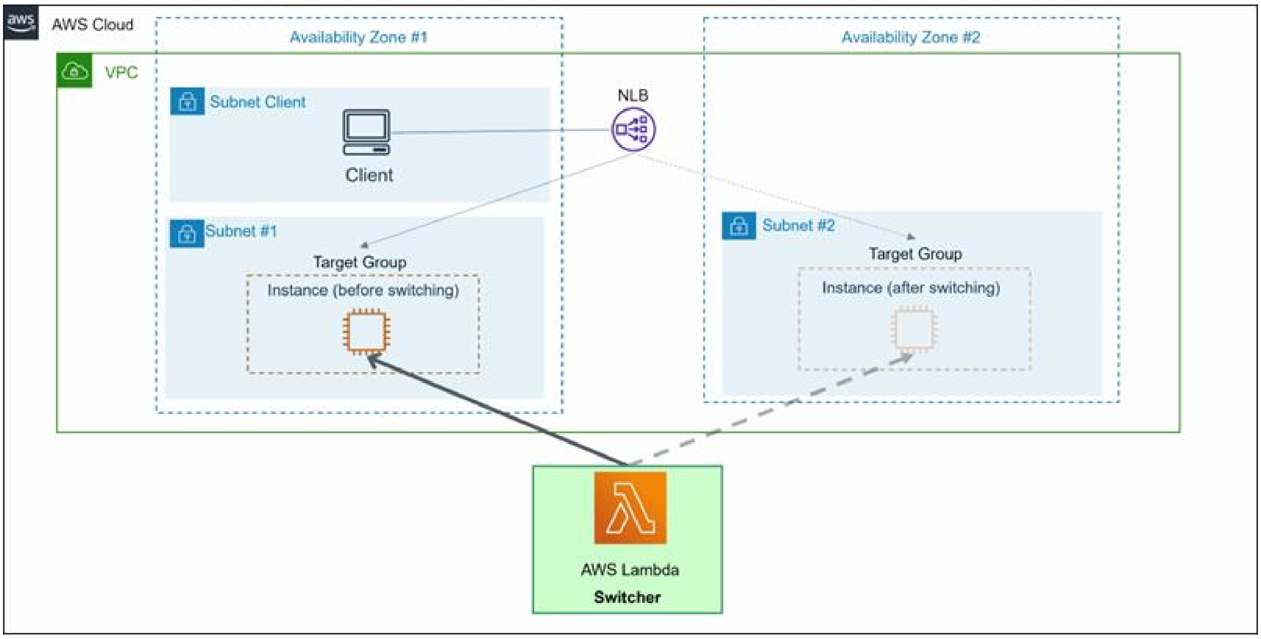
A.1.3.4 Network Takeover

Use ELB NLB or ALB for network takeover.

Switcher terminates the instance and launches the instance in a different AZ when failure occurs. At that time, we removed the instances in the target group and added the newly launched instances.

Client access NLB and ALB based on AWS specifications. In the event of a cluster node failure, the instance to which NLB or ALB is forwarded is automatically switched for continued access. For more information on how to access it, see the official AWS documentation.

Figure A.5 ELB Network takeover



An example of NLB is shown above, but it can also be configured with ALB. When configuring with ALB, configure the network, security groups, etc. appropriately to meet your security requirements.

The target group for an ELB or ALB must be an instance. When you create a target group, you specify the following options.

Setting item	Value
Target type	Instance

Note

When handling multiple instances for switching, separate ELBs (NLB or ALB) for each instance and set up network takeover for each instance.

A.1.3.5 Security Groups Design

Describes the security group rule settings.

In addition to the rules listed here, design your security group by adding rules according to your security requirements.

A.1.3.5.1 Rules for Instances

Design the security rules that you want to apply to the cluster nodes.

Add the rules required for the application to work, as well as the following rules.

Inbound rule

Allow access from NLB or ALB. For more information, see the official AWS documentation.

Outbound rule

Destination	Protocol	Port Range	Description
Security groups for VPC endpoints	tcp	443	Used by the CloudWatch Agent to access API endpoints. Specify the "A.1.3.5.2 Rules That Apply to VPC Endpoints" security group as the destination.
Reserved IP address for Route 53 Resolver	tcp	53	Used to access DNS servers for AWS resources.
EFS mount target security group	tcp	2049	Use to access an EFS mount target.

A.1.3.5.2 Rules That Apply to VPC Endpoints

Design the security rules that you want to apply to your VPC endpoint.

Add the following rules to your security group.

Inbound rule

Destination	Protocol	Port Range	Description
Security groups that you configure for instances on cluster nodes	tcp	443	Allow access from instances on cluster nodes. Specify the "A.1.3.5.1 Rules for Instances" security group as the destination.

Outbound rule

None

A.1.3.5.3 Rules for EFS Mount Targets

Design the security rules to apply to mount targets for EFS.

Add the following rules to your security group.

Inbound rule

Destination	Protocol	Port Range	Description
Security groups that you configure for instances on cluster nodes	tcp	2049	Allow access from instances on cluster nodes. Specify the "A.1.3.5.1 Rules for Instances" security group as the destination.

Outbound rule

None

A.1.3.5.4 Rules for Blackhole Security Group

Blackhole security group is a security group that blocks all traffic. Blackhole security groups are used by switcher.

Create blackhole security group with the following rules.

Inbound rule

None

Outbound rule

None

A.1.4 Disk Design

Describes the disk configuration required to operate a smart workload recovery feature. The design should be based on the disk configuration.

Place shared data in EFS. If you are creating an EFS file system, select "Regional" for the availability and durability options. Redundant data is stored in multiple AZ. Design the mount target for EFS to be on subnet of AZ for which you want to switch cluster nodes.

EFS file systems are provided for the following purposes.

RMS logs

Provide a file system for EFS to store RMS logs. When you create an EFS file system, specify the following options.For more information, see the official AWS documentation.

Setting item	Value
Availability and durability	Regional
Performance mode	General Purpose
Throughput mode	Bursting

Data shared by users

If you want to store data that users share, have an EFS file system that is different from RMS log. When you create an EFS file system, specify the following options.For more information, see the official AWS documentation.

Setting item	Value
Availability and durability	Regional

A.1.5 Policy Design

To grant access to an AWS service or resource, you must attach an IAM role to the service or resource. The policies that you assign to an IAM role are described below.

See

For more information about policies and IAM roles, see the official AWS documentation.

Cluster node instances

Attach an IAM role to the instance (EC2) to install the following on the instance of the cluster node.

- CloudWatch Agents required for operation of resource monitor
- Amazon EFS Client (amazon-efs-utils)

  When you create an IAM role, add the following policy.

- CloudWatchAgentServerPolicy

Note

If an instance on a cluster node does not have an IAM role attached, the instance fails to mount EFS, and the resource monitoring mechanism fails to collect metrics and logs from the instance.

AWS Lambda Switcher

Attach an IAM role to AWS Lambda running switcher. When you create the IAM role, add the following policy.

- AWSLambdaBasicExecutionRole
- New Policy to Create

  The policy defines the scope to which access is allowed so that Lambda function of the switcher can work with AWS resources. Create a new policy based on JSON. There are strings in JSON that need to be rewritten. The string is shown below.

String	Value to be rewritten	Description
${Region}	AWS Region If you do not specify a region, replace ${Region} with "*".	Specify the AWS region that you want to allow access to. If you specify a region, you must provide policies and IAM roles for each region. If you do not specify a region, you can use both policies and IAM roles.
${Account}	AWS Account	Specify the AWS account that you want to grant access to.

The JSON is shown below.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "VisualEditor0",
      "Effect": "Allow",
      "Action": "ec2:CreateTags",
      "Resource": [
        "arn:aws:ec2:${Region}:${Account}:volume/*",
        "arn:aws:ec2:${Region}:${Account}:instance/*",
        "arn:aws:ec2:${Region}:${Account}:network-interface/*"
      ],
      "Condition": {
        "StringEquals": {
          "ec2:CreateAction": "RunInstances"
        }
      }
    },
    {
      "Sid": "VisualEditor1",
      "Effect": "Allow",
      "Action": [
        "cloudwatch:PutMetricAlarm",
        "cloudwatch:DescribeAlarms",
        "ec2:TerminateInstances",
        "ec2:ModifyNetworkInterfaceAttribute",
        "ec2:RunInstances",
        "elasticloadbalancing:RegisterTargets",
        "elasticloadbalancing:DeregisterTargets",
        "dynamodb:ConditionCheckItem",
        "dynamodb:PutItem",
        "dynamodb:DeleteItem",
        "dynamodb:GetItem",
        "dynamodb:UpdateItem"
      ],
      "Resource": [
        "arn:aws:cloudwatch:${Region}:${Account}:alarm:*",
        "arn:aws:ec2:${Region}:${Account}:volume/*",
        "arn:aws:ec2:${Region}:${Account}:subnet/*",
        "arn:aws:ec2:${Region}:${Account}:security-group/*",
        "arn:aws:ec2:${Region}:${Account}:instance/*",
        "arn:aws:ec2:${Region}:${Account}:network-interface/*",
        "arn:aws:ec2:${Region}:${Account}:key-pair/*",
        "arn:aws:ec2:*::image/*",
        "arn:aws:elasticloadbalancing:${Region}:${Account}:targetgroup/*/*",
        "arn:aws:dynamodb:${Region}:${Account}:table/*"
      ]
    },
    {
      "Sid": "VisualEditor2",
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeImages",
        "ec2:DescribeInstances",
        "ec2:DescribeSubnets",
        "ec2:DescribeSecurityGroups",
        "elasticloadbalancing:DescribeTargetHealth",
        "elasticloadbalancing:DescribeTargetGroups"
      ],
      "Resource": "*"
    },
    {
      "Sid": "VisualEditor3",
      "Effect": "Allow",
      "Action": "iam:PassRole",
      "Resource": "arn:aws:iam::${Account}:role/*",
      "Condition": {
        "StringEquals": {
          "iam:PassedToService": "ec2.amazonaws.com"
        }
      }
    },
    {
      "Sid": "VisualEditor4",
      "Effect": "Allow",
      "Action": [
        "elasticloadbalancing:RegisterTargets",
        "elasticloadbalancing:DeregisterTargets"
      ],
      "Resource": "arn:aws:elasticloadbalancing:${Region}:${Account}:targetgroup/*/*"
    }
  ]
}
```

A.1.6 System Design

Use PRIMECLUSTER Designsheets to design the system.

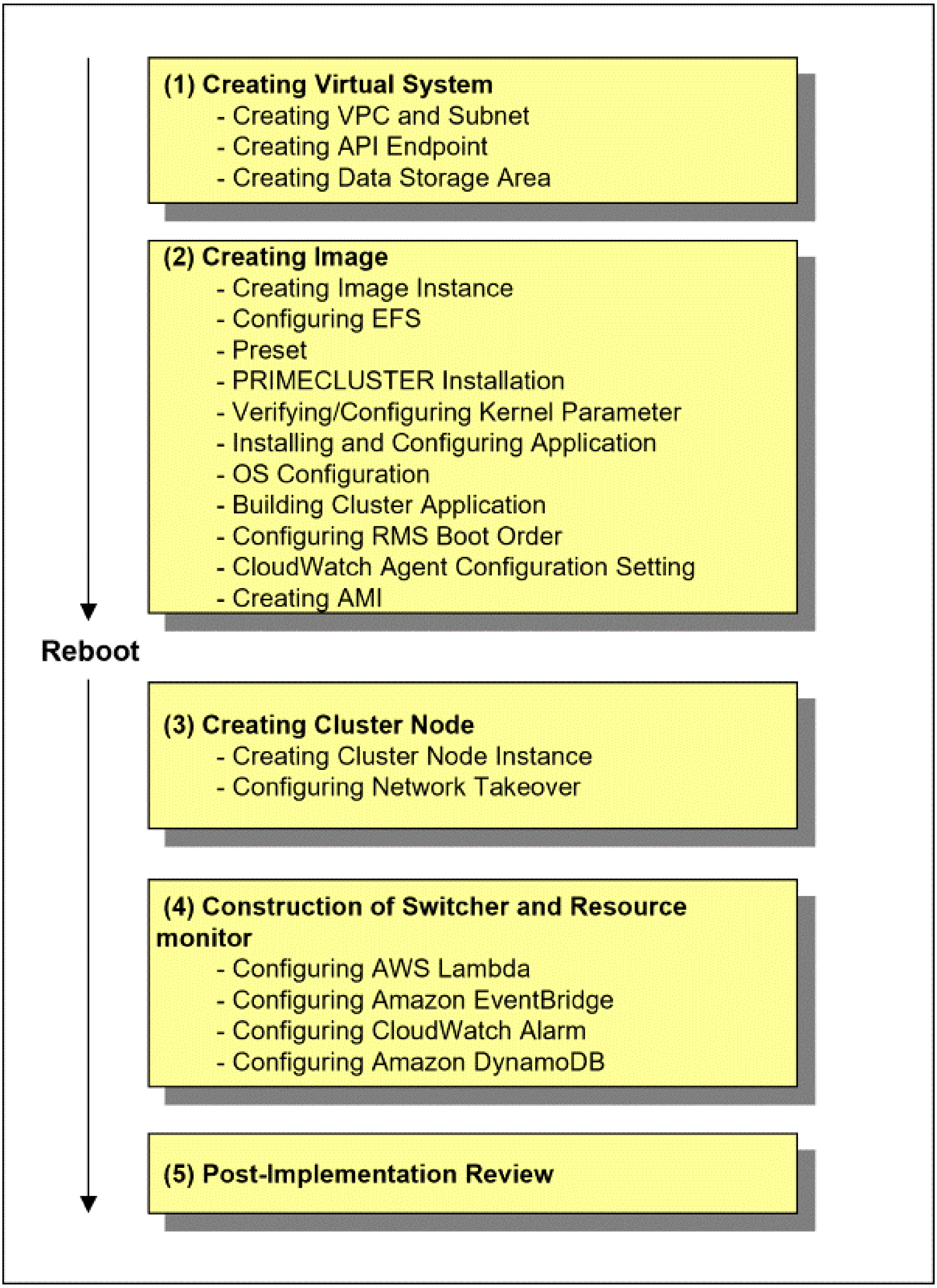
The installation operation of the PRIMECLUSTER system using the smart workload recovery feature is performed based on the created PRIMECLUSTER Designsheets. Make sure to create the designsheets and confirm that all required items are described.

A.2 Installation

Provides step-by-step instructions on how to deploy a cluster system that uses the smart workload recovery feature.

Perform the operations shown in the following figure.

Figure A.6 Instructions for deploying a system using the smart workload recovery feature


(1) Creating Virtual System
- Creating VPC and Subnet
- Creating API Endpoint
- Creating Data Storage Area
(2) Creating Image
- Creating Image Instance
- Configuring EFS
- Preset
- PRIMECLUSTER Installation
- Verifying/Configuring Kernel Parameter
- Installing and Configuring Application
- OS Configuration
- Building Cluster Application
- Configuring RMS Boot Order
- CloudWatch Agent Configuration Setting
- Creating AMI
Reboot
(3) Creating Cluster Node
- Creating Cluster Node Instance
- Configuring Network Takeover
(4) Construction of Switcher and Resource monitor
- Configuring AWS Lambda
- Configuring Amazon EventBridge
- Configuring CloudWatch Alarm
- Configuring Amazon DynamoDB
(5) Post-Implementation Review

Note

How to check the operation environment, refer to "Operation Environment" in the "PRIMECLUSTER Cloud Edition Installation Guide".

A.2.1 Creating Virtual System

Describes how to create a virtual system for a cluster system in an AWS environment.

Use the AWS Management Console to create the virtual system that you designed in "A.1 Design".

See

For detailed instructions on creating virtual system and configuring each, see the official AWS documentation.

A.2.1.1 Creating VPC and Subnet

Create a VPC and subnet to deploy your cluster system. Create a VPC and subnet as designed in "A.1.3.1 VPC and Subnet Design".

Required tags for switched subnets

Set the following tags for the subnet you want to switch to. The switcher recognizes in the tag the subnet on which to switch instances of a cluster node.

Key	Value
fujitsu.pclswr.idlist	Specify an integer value(1 to 10^{38}-1) that identifies the cluster node. Separate multiple cluster nodes into subnets with comma(,).
fujitsu.pclswr.is_recovery_target	Specify whether it is the subnet to switch to. true: applicable false: not applicable

Point

Value of the [fujitsu.pclswr.idlist] key

The value of the [fujitsu.pclswr.idlist] key must be an integer equal to the value of 1 to 10^{38}-1 as the [fujitsu.pclswr.id] key specified in "A.2.3.1 Creating Cluster Node Instance".

[Example]

The following is an example of an integer value that identifies the cluster node specified in the instance and subnet tags.

- Switching an instance to two subnets

Resources	Tag Key	Tag Value
Instance1	fujitsu.pclswr.id	1
Subnet1	fujitsu.pclswr.idlist	1
Subnet2	fujitsu.pclswr.idlist	1

- Switching two instances to two subnets

Resources	Tag Key	Tag Value
Instance1	fujitsu.pclswr.id	1
Instance2	fujitsu.pclswr.id	2
Subnet1	fujitsu.pclswr.idlist	1,2
Subnet2	fujitsu.pclswr.idlist	1,2

A.2.1.2 Creating API Endpoint

To configure connectivity with the API endpoint, create a VPC endpoint as designed in "A.1.3.2 Ensuring Connectivity with API Endpoints". In addition, set the security group that was designed in "A.1.3.5.2 Rules That Apply to VPC Endpoints".

A.2.1.3 Creating Data Storage Area

To create a data storage area, create a file system for EFS as designed in "A.1.4 Disk Design". Also, make sure that you have a mount target for EFS, as described in "A.1.3.3 Verifying EFS Mount Targets".Mount targets for EFS must be set to the security group designed in "A.1.3.5.3 Rules for EFS Mount Targets".

A.2.2 Creating Image

Provides step-by-step instructions on how to create an AMI for an instance on a cluster node.

A.2.2.1 Creating Image Instance

Create an instance to be a cluster node and do the following.

1. Configure the security group that you designed in "A.1.3.5.1 Rules for Instances".

2. Create IAM role and attach it to the instances that will be cluster nodes, as described in "A.1.5 Policy Design".

A.2.2.2 Configuring EFS

Because you are creating a file system for EFS in "A.1.4 Disk Design", the following file systems must be mounted.

- RMS logs

- Shared data (if created by the user)

To mount EFS file system on a cluster node.

1. Install amazon-efs-utils as described in "Manually installing the Amazon EFS client" in the Amazon Elastic File System User Guide.

2. Create mount points before installing PRIMECLUSTER.

```
# mkdir -p /var/opt/SMAWRrms/log
# chmod -R 755 /var/opt/SMAWRrms
# mkdir -p <Any directory>
```

3. Set to mount automatically at startup.

   Fixing fstab

   fstab settings (for IAM authenticated mounts)

```
# When the file-system-id of the RMS log is added to /etc/fstab:
file-system-id:/ /var/opt/SMAWRrms/log efs _netdev,noresvport,tls,iam 0 0
# When the file-system-id of the user-shared data is added to /etc/fstab:
file-system-id:/ <Any directory> efs _netdev,noresvport,tls,iam 0 0
```

   Note

   - Note that forgetting "_netdev" option causes the instance to become unresponsive when automounting fails.

   - RHEL 8 and later have SELinux enabled by default. When SELinux is enabled, users are restricted from accessing shared data, so you may need to specify a context in the mount options.

4. After you configure automount, reboot to ensure that RMS log area and shared data are successfully mounted at startup.

A.2.2.3 Preset

Perform the following steps on the cluster nodes to disable the Firewall. Before installing PRIMECLUSTER, disable the firewall on the cluster node. The procedure is as follows.

Make sure firewalld is disabled.

```
# systemctl is-enabled firewalld
```

If enabled, disable it.

```
# systemctl stop firewalld
# systemctl disable firewalld
```

A.2.2.4 PRIMECLUSTER Installation

To install PRIMECLUSTER, use the installation script (CLI Installer).

Install PRIMECLUSTER using the installation script on an instance where Linux (R) software and related software are already installed.

Note

If the OS has not been restarted since the instance was created, restart it before installing PRIMECLUSTER.

See

For more information on how to install/uninstall, see the "PRIMECLUSTER Cloud Edition Installation Guide".

A.2.2.5 Verifying/Configuring Kernel Parameter

It is necessary to change the kernel parameters required by RMS for each environment of the cluster node.

Check the PRIMECLUSTER design sheet and reconfigure kernel parameters if necessary.

See

For more information about kernel parameters, see "RMS Configuration" in "3.1.7 Checking and Setting the Kernel Parameters" in the "PRIMECLUSTER Installation and Administration Guide".

Note

- Restart the OS to enable the changed kernel parameters.
- After uninstalling PRIMECLUSTER, return the kernel parameter settings to the state before installing PRIMECLUSTER, if necessary.

A.2.2.6 Installing and Configuring Application

Install and configure applications to run on a cluster system, as necessary.

See

- Refer to the documentation for your application for instructions on how to set up your preferences.
- For information about how AWS handles PRIMECLUSTER-related products, see the documentation for each product.

A.2.2.7 OS Configuration

Set the name of /etc/hosts file on the cluster node by appending string "RMS" to the hostname (the node name output by the "hostname" command).

Add the name you configure to only one of the loopback addresses or the addresses assigned to the node you are building.

Example of adding to a loopback address

Result of executing the hostname command

```
# /bin/hostname
hostname
```

[Before adding]

```
# /bin/cat /etc/hosts
127.0.0.1   localhost localhost.localdomain localhost4 localhost4.localdomain4
```

[After adding]

```
# /bin/cat /etc/hosts
127.0.0.1   localhost localhost.localdomain localhost4 localhost4.localdomain4 hostnameRMS
```

Additional example when the output result of the hostname command is "node1"

```
# /bin/cat /etc/hosts
127.0.0.1   localhost localhost.localdomain localhost4 localhost4.localdomain4 node1RMS
```

A.2.2.8 Building Cluster Application

For more information on how to build a cluster application on a cluster node, refer to "Chapter 6 Building Cluster Applications" in the "PRIMECLUSTER Installation and Administration Guide" to create a 1:1 standby cluster application and configure it as you would in a single-node cluster operation.

Also, you need to register FaultScript in "6.7.2 Setting Up userApplication" of the "PRIMECLUSTER Installation and Administration Guide". For information about creating a script to register, see "A.2.2.8.1 Creating FaultScript". "6.2 Initial GLS Setup", "6.3 GDS Configuration Setup" and "6.4 Initial GFS Setup" of the "PRIMECLUSTER Installation and Administration Guide" do not need to be set.

Also, because the clchkcluster command is not available, you do not need to perform the "6.9 Checking the Cluster Environment" in the "PRIMECLUSTER Installation and Administration Guide".

Note

- Do not select "Does not start up automatically." for "6.1 Initial RMS Setup" in the "PRIMECLUSTER Installation and Administration Guide".

- Set the AutoStartUp attribute to "yes" for the userApplication attribute that is set in "6.7.2 Setting Up userApplication" in the "PRIMECLUSTER Installation and Administration Guide".

  If you do not set it to "yes", RMS does not start automatically when an instance switch occurs due to an error.

- You do not need to configure the nodes that make up userApplication in step 4 of "6.7.2.1 Creating Standby Cluster Applications" in the "PRIMECLUSTER Installation and Administration Guide".

Register "*hostname*RMS" configured in "A.2.2.7 OS Configuration" by performing the following steps before "6.7.2 Setting Up userApplication" in the "PRIMECLUSTER Installation and Administration Guide".

1. Run the hvw command.

   To start RMS Wizard as the configuration file name (testconf)

   ```
   # /opt/SMAW/SMAWRrms/bin/hvw -n testconf
   ```

2. Select "RMS-CreateMachine" from "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: 16
```

3. Select "FREECHOICE" and enter "*hostname*RMS".

   "*hostname*RMS" is "*hostname*RMS" set in "A.2.2.7 OS Configuration".

   The following is an example of setting "node1RMS".

```
Creation: Add hosts to a cluster:
Current set:
1) HELP
2) QUIT
3) RETURN
4) FREECHOICE
Choose the host to add: 4
        >> node1RMS
```

4. Select "QUIT" to finish the setting.

```
Creation: Add hosts to a cluster:
Current set: node1RMS
1) HELP
2) QUIT
3) RETURN
4) FREECHOICE
Choose the host to add: 2
```

Note

- In "Chapter 6 Building Cluster Applications" of the "PRIMECLUSTER Installation and Administration Guide" start RMS Wizard by specifying the configuration file used to register "*hostname*RMS".
- To configure userApplication for your application, see "6.6 Setting Up Online/Offline Scripts" in the "PRIMECLUSTER Installation and Administration Guide".

A.2.2.8.1 Creating FaultScript

Create FaultScript that registers with userApplication. This script enables you to stop an instance and switch jobs when an application fails.

1. Create a file for the script. Specify any name for the script file name.

```
# touch script
# chmod 700 script
```

2. Add the following to the script.

```
#!/bin/sh
LOG_TRANSFER_SECONDS=10

NODE=${HV_NODENAME:-unknown}
```

```
STATE=${HV_SCRIPT_TYPE:-unknown}

function Msg {
    echo ${NODE}: ${STATE}: $(date "+%Y-%m-%d %H:%M:%S"): $*
}

Msg NOTICE: Waiting ${LOG_TRANSFER_SECONDS} seconds for log transfer.
/usr/bin/sleep ${LOG_TRANSFER_SECONDS}

/usr/bin/systemctl -f poweroff
Ret=$?
if [ ${Ret} != 0 ]; then
    # systemctl Message when the command fails
    Msg ERROR: systemctl failed. Return code ${Ret}.
    exit 1
fi

exit 0
```

The script uses the following variables to specify values to adjust for your environment.

Parameter	Description
LOG_TRANSFER_SECONDS	In CloudWatch Agent settings, set a value that is twice the force_flush_interval in the Logs section. The unit is seconds. For details about the force_flush_interval check method, default values, and other specifications, see the official AWS documentation.

Note

- OfflineScript does not run when userApplication is notified of a fault. If you have something to do before stopping the instance, do the following.

    - Add an action to the third line of the script to be registered.

    - Specify multiple commands during FaultScript registration procedure.
      Note that the scripts for FaultScript must be listed at the end.
      (In the example below,/var/tmp/FaultScript.sh is the script for FaultScript.)

```
/var/tmp/command ; /var/tmp/FaultScript.sh
```

- An extension of the systemctl command unmounts the file system. If unmounting takes a long time (for example, due to a file system error), the switch may take a long time. Also, if the systemctl command fails, use the AWS Management Console to stop the machine.

A.2.2.9 Configuring RMS Boot Order

Configure RMS to start after the log area of RMS is mounted.

1. Setting the Startup Order

   Create a drop-in unit file for RMS service to configure the boot order.

   Create /etc/systemd/system/smawrs.service.d directory.

```
# mkdir /etc/systemd/system/smawrrms.service.d
# chmod 755 /etc/systemd/system/smawrrms.service.d
```

   Create a configuration file in the directory you created.

```
# touch /etc/systemd/system/smawrrms.service.d/pclswr-efs.conf
# chmod 644 /etc/systemd/system/smawrrms.service.d/pclswr-efs.conf
```

Unit file settings

```
[Unit]
Wants=var-opt-SMAWRrms-log.mount
```

After creating the unit file, execute the following command.

```
# systemctl daemon-reload
```

2. Checking the Startup Order

   Verify that RMS service dependency includes "var-opt-SMAWRrms-log.mount", which is the unit file associated with the mount, by running the following command.

```
# systemctl list-dependencies smawrrms.service
```

Sample Output (RMS Service Dependencies)

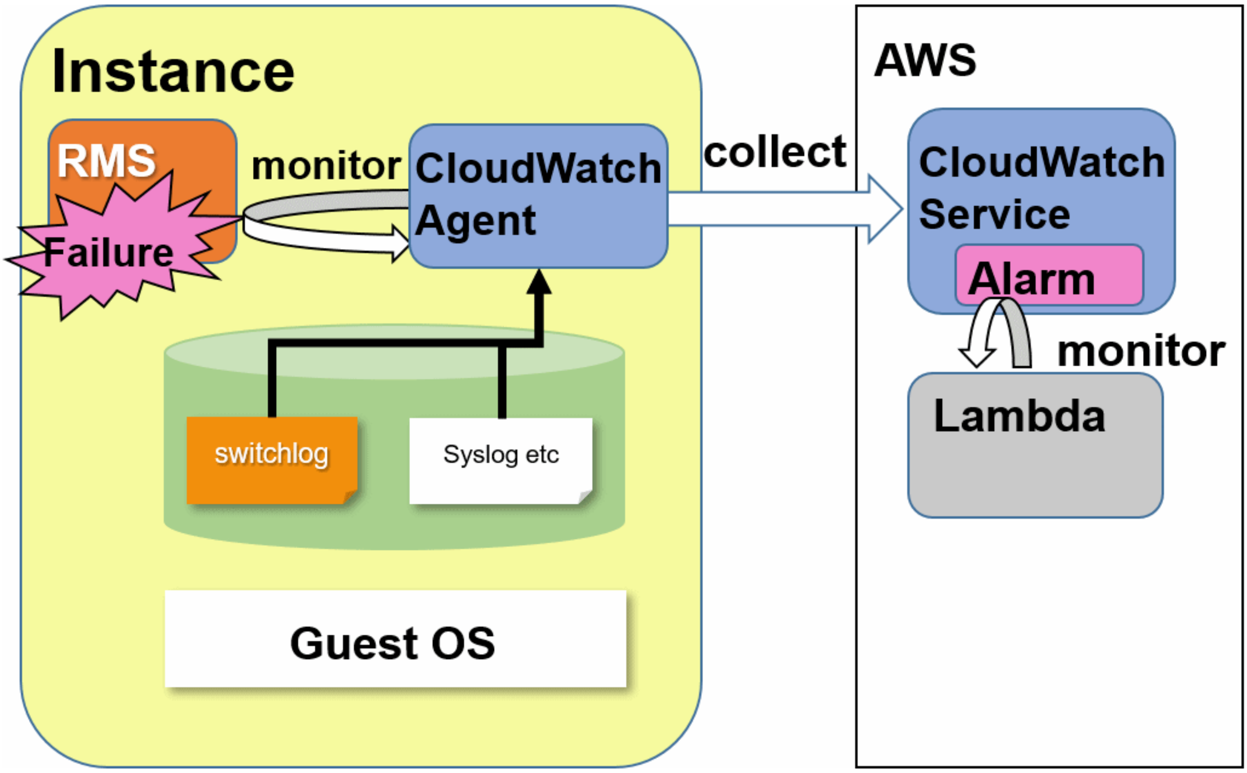
```
# systemctl list-dependencies smawrrms.service
smawrrms.service
* |-system.slice
* |-var-opt-SMAWRrms-log.mount
* `-sysinit.target
*   |-dev-hugepages.mount
(omit)
```

A.2.2.10 CloudWatch Agent Configuration Setting

You can collect CloudWatch Agent information by installing it on your instances and configuring the information that you want to collect. You can use the collected information to monitor RMS service in CloudWatch services and switch instances when failure is detected.

The Smart Workload Recovery feature uses CloudWatch Agent to monitor and log RMS process survival.

Figure A.7 Monitoring and Logging RMS with CloudWatch Agent



A.2.2.10.1 Installing CloudWatch Agent

Install CloudWatch Agent by using the command line, AWS Systems Manager, or AWS CloudFormation methods. For more information, see "Installing CloudWatch Agent" in the "Amazon CloudWatch User Guide".

A.2.2.10.2 Configuring CloudWatch Agent

Provides step-by-step instructions on how to configure CloudWatch Agent.

1. Configuring CloudWatch Agent

   Create a CloudWatch Agent configuration file. Specify the configuration file that you created to launch CloudWatch Agent. For more information, see "Create CloudWatch Agent configuration file" in the "Amazon CloudWatch User Guide". Include the following in the configuration file.

   Configuration Contents

```
{
    "agent": {
        "metrics_collection_interval": 60,
        "logfile": "/opt/aws/amazon-cloudwatch-agent/logs/amazon-cloudwatch-agent.log",
        "debug": false
    },
    "metrics": {
        "namespace": "RMS",
        "append_dimensions": {
            "InstanceId":"${aws:InstanceId}",
            "InstanceType":"${aws:InstanceType}"
        },
        "metrics_collected": {
            "procstat": [
                {
                    "pattern": "/opt/SMAW/SMAWRrms/bin/bm",
                    "measurement": [
                        "pid_count"
                    ],
                    "metrics_collection_interval": 60
                }
            ]
        },
        "force_flush_interval": 60
    },
    "logs": {
        "logs_collected": {
            "files": {
                "collect_list": [
                    {
                        "file_path": "/var/log/messages*",
                        "log_group_name": "system1_messages",
                        "log_stream_name": "{instance_id}"
                    },
                    {
                       "file_path": "/opt/aws/amazon-cloudwatch-agent/logs/amazon-cloudwatch-
agent.log",
                        "log_group_name": "system1_cwa_messages",
                        "log_stream_name": "{instance_id}"
                    }
                ]
            }
        },
        "force_flush_interval":5
    }
}
```

When you choose all metrics from the CloudWatch menu, the custom namespace displays the "RMS" that you specified for "namespace" in the metrics section. After you choose a log group from the CloudWatch menu, the logs are displayed by instance ID under the "system1_messages" that you specify for "log_group_name" in the logs section.

Information

You can change the values of the following parameters that set how often metrics and logs are collected.

- metrics_collection_interval
- force_flush_interval

To add or delete a monitoring log, add or delete the "collect_list" entry in the logs section.

Note

- The "RMS" specified in "namespace" in the metrics section should only be used in this configuration.
- After CloudWatch Agent starts, CloudWatch Agent configuration file is renamed and moved to the following directory.

  /opt/aws/amazon-cloudwatch-agent/etc/amazon-cloudwatch-agent.d

  Example) CloudWatch Agent configuration files are stored in the

  /opt/aws/amazon-cloudwatch-agent/etc/amazon-cloudwatch-agent.json

  CloudWatch Agent configuration file is:

  /opt/aws/amazon-cloudwatch-agent/etc/amazon-cloudwatch-agent.d/file_amazon-cloudwatch-agent.json

2. Configuring CloudWatch services startup order

   Configure CloudWatch services to start after RMS startup.

Note

If the /etc/systemd/system/smawrs.service.d directory does not exist, you may not have performed a "A.2.2.9 Configuring RMS Boot Order". Please check if it is implemented.

1. Sets the start/stop order for CloudWatch Agent service.

   Create a configuration file in the /etc/systemd/system/smawrs.service.d directory.

   ```
   # touch /etc/systemd/system/smawrrms.service.d/pclswr-cwa.conf
   # chmod 644 /etc/systemd/system/smawrrms.service.d/pclswr-cwa.conf
   ```

   Include the following in the configuration file.

   ```
   [Unit]
   Before=amazon-cloudwatch-agent.service
   ```

   Reflects the start/stop order setting for CloudWatch Agent service.

   ```
   # systemctl daemon-reload
   ```

2. Check the start/stop order settings for CloudWatch Agent service.

   Ensure that amazon-cloudwatch-agent.service is included.

   ```
   # systemctl show smawrrms.service | grep "Before="
   Before=shutdown.target amazon-cloudwatch-agent.service multi-user.target
   ```

A.2.2.11 Creating AMI

Use the AMI when switching instances.

Use the following procedure to create an AMI from the AWS Management Console.

1. Stop the instance that will be the cluster node
2. Create an AMI for the instance that will be the cluster node.
3. Terminate the instance that will be the cluster node.

See

For more information about AMIs and instructions for creating AMIs, see the official AWS documentation.

Note

- Be sure to create an AMI. If you don't, PRIMECLUSTER launches an instance from an AMI for which "A.2 Installation" has not been implemented when you switch instances. In this case, the instance is not guaranteed to work.
- If you make changes to the OS settings (such as network settings) while the instance is running, recreate the AMI.

  For instructions about creating an AMI in the operating instance, see Create an AMI "A.5.8 Procedure for Getting AMI in Operation".
- After you launch a new instance, the AMI, snapshot, and CloudWatch logs associated with the previous instance remain, so delete them if you don't need them to prevent charges. For instructions, see the official AWS documentation.

A.2.3 Creating Cluster Node

Provides step-by-step instructions on how to create a cluster node.

A.2.3.1 Creating Cluster Node Instance

To tie the AMI you created in "A.2.2.11 Creating AMI" to an instance, create an instance based on the AMI you created, and then do the following.

- Configure the security group that you designed in "A.1.3.5.1 Rules for Instances".
- Attach the IAM role for the instances on the cluster nodes.

Required Tags for Instances on Cluster Nodes

Configure the following tags so that switcher identifies instances on cluster nodes.

Key	Value
fujitsu.pclswr.id	Specifies an integer value(1 to 10^{38}-1) that identifies the cluster node. The integer value identifying the cluster node must be unique for each region. The integer value is specified referring to "About Creating Multiple Resources" in "A.1.2 System Configuration".
fujitsu.pclswr.is_recovery_target	Specifies whether the node is a cluster node. true: a cluster node false: not a cluster node

A.2.3.2 Configuring Network Takeover

Create the resources required to set up network takeover.

1. Create an NLB or ALB with the target group as designed in "A.1.3.4 Network Takeover".

2. Add the instance you created in "A.2.3.1 Creating Cluster Node Instance" to the target group.

A.2.4 Construction of Switcher and Resource monitor

Provides step-by-step instructions on how to build switcher and resource monitor.

A.2.4.1 Configuring AWS Lambda

Describes how to configure AWS Lambda for switcher.

Creating Blackhole Security Groups

Create a blackhole security group as described in "A.1.3.5.4 Rules for Blackhole Security Group".

Creating Execution Role

Create the IAM role that you specify for the AWS Lambda execution role. Create the IAM role that you designed in "A.1.5 Policy Design".

Creating an AWS Lambda Function

Select the "Create From Scratch" option and create a Lambda function with the following settings.

Setting item	Value
Function name	Fujitsu-Pclswr-Function-Switcher-"Any string" "Any string" must be unique for the VPC. Please specify "Any string" referring to "About Creating Multiple Resources" in "A.1.2 System Configuration". Function names must conform to AWS rules.
Runtime	Python 3.7
Default Execution Role	Specify the IAM role that you created in "Creating Execution Role".
Enable Network	uncheck

After you create the Lambda function, continue to modify settings on the following Lambda function tabs.

Code

Upload the Lambda function zip file on the software DVD.

The Lambda function zip file is located at.

Tool/pclswr/FJSVpcl-swr-switcher.zip

Set the following values for "Handler" in "Runtime Settings".

Setting item	Value
Handler	pclswr.lambda_handler

Configuration

Under general configuration, set the timeout value to.

Setting item	Value
Timeout	15 min 0 sec

Display Environment Variables to add the following environment variables.

Key	Value
PCLSWR_SYSTEM_LIST	Specify the list of cluster nodes to be handled by switcher. Separate multiple cluster nodes with spaces.
PCLSWR_BLACKHOLE	Specify the blackhole security group identifier.

A.2.4.2 Configuring Amazon EventBridge

This subsection describes how to set the switcher and the resource monitor.

Create an Amazon EventBridge rule with the following settings.

Instance State Change Event

Classification	Setting item	Value
Name and description	Name	Fujitsu-Pclswr-Rule-InstanceStopped-"Any string" "Any string" must be unique for the VPC. Please specify "Any string" referring to "About Creating Multiple Resources" in "A.1.2 System Configuration". Function names must conform to AWS rules.
Define pattern	Pattern	Event pattern
	Event matching pattern	Pre-defined pattern by service
	Service provider	AWS
	Service name	EC2
	Event type	EC2 Instance State-change Notification
	State	Specific state(s)
		stopped
	Instance	Any instance
Select targets	Target	Lambda function
	Function	Specify the Lambda function that you created in "A.2.4.1 Configuring AWS Lambda".

CloudWatch Alarm State Change Events

Classification	Setting item	Value
Name and description	Name	Fujitsu-Pclswr-Rule-CloudWatchStatusCheck-"Any string" "Any string" must be unique for the VPC. Please specify "Any string" referring to "About Creating Multiple Resources" in "A.1.2 System Configuration". Function names must conform to AWS rules.
Define pattern	Pattern	Event pattern
	Event matching pattern	Pre-defined pattern by service
	Service provider	AWS
	Service name	CloudWatch
	Event type	CloudWatch Alarm State Change
Select targets	Target	Lambda function
	Function	Specify the Lambda function that you created in "A.2.4.1 Configuring AWS Lambda".

A.2.4.3 Configuring CloudWatch Alarm

This subsection describes how to set up the resource monitor.

If you created instances in "A.2.3.1 Creating Cluster Node Instance", create a CloudWatch Alarm for each instance.

CloudWatch Alarm create two alarms: a switcher and an alarm for CloudWatch Agent. To create a CloudWatch Alarm, specify the following settings.

Switcher alarms

Category	Setting Item	Value
Select Metrics	AWS Namespace	EC2
	Metric	Per-Instance Metrics
	Instance name (InstanceId)	Select an instance of a cluster node
	Metric name	StatusCheckFailed Select StatusCheckFailed for the instance on the screen.
Metric	Statistic	Maximum
	Period	1 minute
Conditions	Threshold types	Static
	Alarm condition	Greater/Equal
	Threshold	0.99
	Datapoints to alarm	2/2
	Missing data treatment	Treat missing data as missing
Action settings	Notification	None Click the [Delete] button to delete the notification. Turn off notification.
	Auto Scaling action	None
	EC2 action	None
	Systems Manager action	None
Name and description	Alarm name	Fujitsu-Pclswr-Alarm-Instance-StatusCheckFailed-"Integer value identifying cluster node" The integer value that identifies the cluster node must be unique for each region. For the integer value, refer to "About Creating Multiple Resources" in "A.1.2 System Configuration".

CloudWatch Agent Alarms

Category	Setting Item	Value
Select Metrics	Custom Namespace	RMS Displays the value you set for namespace in CloudWatch Agent configuration file.
	Dimension	InstanceId, InstanceType, pattern, pid_finder
	Instance name (InstanceId)	Select an instance of a cluster node.
	Metric name	procstat_lookup_pid_count Select procstat_lookup_pid_count for the instance on the screen.
Metric	Statistic	Average
	Period	1 minute
Conditions	Threshold types	Static
	Alarm condition	Lower/Equal
	Threshold	0

Category	Setting Item	Value
	Datapoints to alarm	1/1
	Missing data treatment	Treat missing data as missing
Action settings	Notification	None Click [Delete] button to delete the notification. Turn off notification.
	Auto Scaling action	None
	EC2 action	None
	Systems Manager action	None
Name and description	Alarm name	Fujitsu-Pclswr-Alarm-Instance-RMS-"Integer value identifying cluster node" The integer value that identifies the cluster node must be unique for each region. For the integer value, refer to "About Creating Multiple Resources" in "A.1.2 System Configuration".

Point

- Create an alarm for CloudWatch Agent after you start collecting metrics on CloudWatch Agent.
- CloudWatch Agent alarms are configured for survival monitoring to detect RMS outages.
- RMS shutdown is determined by whether "/opt/SMAW/SMAWRrms/bin/bm" process has stopped. The judgment condition is when the number of "/opt/SMAW/SMAWRrms/bin/bm" processes (pid_count) becomes 0 or lower. The evaluation is once a minute. If the condition is satisfied in the evaluation, the alarm transitions from OK to ALARM.
- Alarms can be in the following states. When the status changes from OK to ALARM, the process of switching instances is performed.
  - OK (within thresholds for which metrics are defined)
  - ALARM (the metric is above a defined threshold)
  - INSUFFICIENT_DATA (CloudWatch Agent stopped, metrics unavailable, and so on)
- To create an alarm, see "Create a CloudWatch Alarm based on a static threshold" in the "Amazon CloudWatch User Guide".

A.2.4.4 Configuring Amazon DynamoDB

Describes how to configure Amazon DynamoDB, which is required for switcher.

Switcher relies on Amazon DynamoDB to manage information for the cluster nodes. Create a table in Amazon DynamoDB. Create as many items as there are cluster nodes. If you add a cluster node later, add an item.

Creating Tables

Create an Amazon DynamoDB table with the following settings.

Setting item	Value
Table name	Fujitsu-Pclswr-DB-Switcher
Partition key	SystemID / Number
Sort Key	InstanceID / String
Settings	Default settings

Creating Item

Items store information about cluster nodes. For each instance of the cluster node, create an item as follows.

Add new attribute	Attribute name	Value
- (Added)	SystemID (Partition key)	Specifies the same integer value as the [fujitsu.pclswr.id] key specified in "A.2.3.1 Creating Cluster Node Instance".
- (Added)	InstanceID (Sort key)	Instance ID of the cluster node
String	State	NOT_SWITCHED

Point

The attribute name state contains a value indicating whether the cluster node instance is switched. Its values are described below.

- NOT_SWITCHED

  Indicates that the cluster node instance is not switched. Specify NOT_SWITCHED if no initial value or switch has been performed.

- SWITCHING

  Indicates that the cluster node instance is switched.

A.2.5 Post-Implementation Review

After the deployment is complete, do the following to verify that the deployment is successful.

1. Check the security group and verify that the Blackhole security group is not misconfigured. If this setting is incorrect, data corruption may occur.

2. Stop the instance on the cluster node and verify that the switch occurs.

3. Ensure that the instance created by the switch in Step 2 inherits the following information from the source instance.

   - AMI ID
   - Instance type
   - Key pair name
   - Security group ID
   - Tags
   - IAM Role
   - CloudWatch Alarm
   - RMS logs
   - User shared data

A.3 Operations

Logs for systems that use the smart workload recovery feature are output to Amazon CloudWatch Logs. To detect when switcher fails, a system that uses the smart workload recovery feature must monitor the messages in the logs output to Amazon CloudWatch Logs. Check the message to see if a switching event occurred during operation.

For detailed operational procedure using Amazon CloudWatch that are required to operate a system using the smart workload recovery feature, see "A.3.1 Using Amazon CloudWatch for Operations".

For information about what to do if a switchover interruption occurs in a system that uses the smart workload recovery feature, see "A.3.2 Actions to take when switching is interrupted".

For information about viewing logs in Amazon CloudWatch Logs, see the official AWS documentation.

Note

On systems that use the smart workload recovery feature, CloudWatch Agent monitors RMS. If RMS goes down abnormally, a switch occurs regardless of the state of the cluster application.

A.3.1 Using Amazon CloudWatch for Operations

Describes how to use Amazon CloudWatch to operate your system with the smart workload recovery feature.

A.3.1.1 Changing the monitoring target settings

If you want to change the frequency at which metrics and logs are collected, or if you want to add or delete logs to monitor, modify CloudWatch Agent configuration file that you created in "A.2.2.10.2 Configuring CloudWatch Agent". Then start CloudWatch Agent with the configuration file. For more information about CloudWatch Agent configuration files, see "Installing the CloudWatch agent" in the "Amazon CloudWatch User Guide".

A.3.1.2 Operation when RMS startup is stopped / RMS monitoring is disabled

Perform the following procedure when you want to operate (start or stop) RMS that is monitored, for example, when applying RMS corrections.

Preparation

You must first create two files as configuration files for the CloudWatch agent.

a. RMS/log monitoring file

   Configuration files with metrics and logs sections

   Refer to "A.2.2.10.2 Configuring CloudWatch Agent" for the creation method.

b. Log monitoring file

   Configuration files in the logs section only

RMS Operation

- When to Stop RMS

  1. Load the configuration file in (b) above and change the setting to log monitoring only.

  2. Stop RMS.

     To stop RMS, see "7.1.3 Stopping RMS" in the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide".

- When to Start RMS

  1. Start RMS.

     To start RMS, see "7.1.1 Starting RMS" in the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide".

  2. Load the configuration file in (a) above and change RMS monitoring and log monitoring settings.

Note

- Loading a configuration file is the same as starting CloudWatch Agent.

  For more information about launching CloudWatch Agent, see "Installing the CloudWatch agent" in the "Amazon CloudWatch User Guide".

- If you load CloudWatch Agent configuration file in the following incorrect order, an instance switch occurs.

  - When you stop RMS with RMS/log monitoring file (a) loaded

- When you stop RMS with log monitoring file (b) loaded and then load (a) with it loaded

A.3.1.3 Troubleshooting CloudWatch Agent

For issues with installing and operating the CloudWatch Agent, see "Troubleshooting the CloudWatch agent" in the "Amazon CloudWatch User Guide".

A.3.2 Actions to take when switching is interrupted

Switching to a system using the smart workload recovery function may be interrupted due to an API or service error. This subsection describes what to do if switching is interrupted.

Note

If the problem persists, collect troubleshooting information and contact your our company service representative (SE).

If the switch is aborted, see "A.3.2.1 Reviewing Error Messages" for an error message.

Take the appropriate action as described in the message, then perform "A.3.2.2 Recovering According to Instance State" or "A.3.2.3 Changing the settings of the switch destination AZ".

Note that you can use the AWS Management Console to handle the switch interruption. For instructions about using the AWS Management Console, see the official AWS documentation.

A.3.2.1 Reviewing Error Messages

If the switch is interrupted while the system is operating with the smart workload recovery feature, check the Lambda function log for error messages about the switch interruption. To check for error messages, use the AWS Management Console and do the following:.

1. From the Region Name drop-down list, select the region in which your system is running.

2. Open the Amazon CloudWatch service administration screen and choose "Log Group" from the drop-down list that displays "Log" from the sidebar.

3. From the list of log groups displayed in the Log Groups screen, choose a log group name in the following format that includes the Lambda function name you are using on your production system.

```
/aws/lambda/< Lambda function name >
```

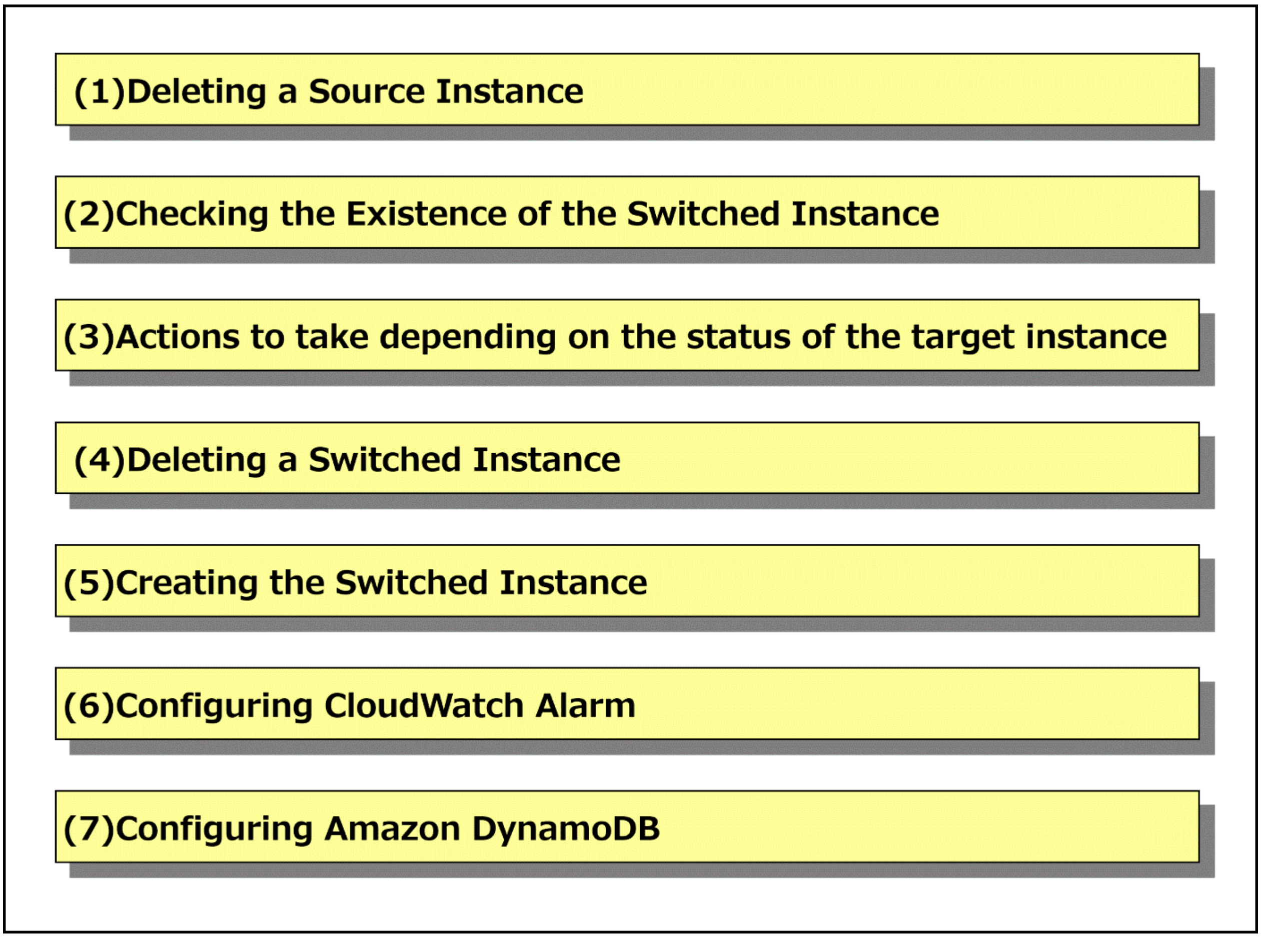
4. From the list of log streams displayed in the "Log Stream" block, select an item that contains a log that is close to the time when the switch was interrupted in the active system. Check the time in the "Last event time" column to see if this is the log stream that contains the log.

5. Check the list of messages displayed on the screen for error messages.

6. Refer to "8.1.4 Error (ERROR) Message" in the "PRIMECLUSTER Messages" and take action according to the error message identified in step 5.

A.3.2.2 Recovering According to Instance State

If you need to recover the instance as per the error message identified in "A.3.2.1 Reviewing Error Messages", recover the instance from which the switch was interrupted. To recover your instance, use the AWS Management Console and perform the following steps:.



(1)Deleting a Source Instance

(2)Checking the Existence of the Switched Instance

(3)Actions to take depending on the status of the target instance

(4)Deleting a Switched Instance

(5)Creating the Switched Instance

(6)Configuring CloudWatch Alarm

(7)Configuring Amazon DynamoDB

A.3.2.2.1 Deleting a Source Instance

To delete the source instance, use the AWS Management Console and do the following.

This procedure is not necessary if the source instance does not exist in the displayed list of instances when you open the Manage Amazon EC2 Service screen, or if the state of the source instance is terminated.

1. Sets the tag "fujitsu.pclswr.is_recovery_target" value of the source instance to false.
2. Stop the source instance.
   Not required if the instance state is stopped.
3. Note the integer value that identifies the cluster node for the tag "fujitsu.pclswr.id" of the source instance that you stopped in Step 2. The saved integer value is used by "A.3.2.2.5 Creating the target instance" when the error message does not identify the cluster node of the source instance. Do not use if the error message confirms it. Terminate the instance that you stopped in step 2.

A.3.2.2.2 Checking the Existence of the target instance

To verify the existence of the switched instance, use the AWS Management Console to check if the error message contains "new instance_id = <*new_instance_id*>", and then do one of the following:.

- If the Error Message Contains "new instance_id = <*new_instance_id*>"
- If the Error Message Does Not Contain "new_instance_id = <*new_instance_id*>"

If the Error Message Contains "new instance_id = <*new_instance_id*>"

Depending on the instance ID value of the switched instance described in <*new_instance_id*>, do the following:.

- <*new_instance_id*> is not None

  Verify that an instance with the same instance ID exists.

1. Open the Manage Amazon EC2 Service screen and choose Instances from the sidebar.

2. From the list of displayed instances, verify that the instance ID in the Instance ID column is the same as the instance ID of the switched instance. If the same instance ID is found, the switched instance exists.

3. Proceed according to the following table depending on whether you have switched to an instance or not.

Whether to switch to the instance	Next Steps
There is a switched to instance.	Proceed to "A.3.2.2.3 Actions to take depending on the status of the target instance".
Destination Instance Does Not Exist	Proceed to "A.3.2.2.5 Creating the target instance".

- The value of <*new_instance_id*> is None.

  Verify that there is an instance with an integer tag that identifies the same cluster node.

  1. Open the Manage Amazon EC2 Service screen and choose Tags from the sidebar.

  2. Select the [Manage Tags]button.

  3. Ensure that Instance is selected in the filter drop-down list.

  4. From the list of displayed instances, verify that the value in the "fujitsu.pclswr.id" column is the same as the value set in "A.2.3.1 Creating Cluster Node Instance" for the following tag, and note the value in the "Instance ID" column.

Key	Value
fujitsu.pclswr.id	Same as the system_id value in the message

  5. From the sidebar, select "Instance" and from the list of displayed instances, verify that the instance ID in the "Instance ID" column is the same as the instance ID you noted in step 4. If the same instance ID is found, the switched instance exists.

  6. Proceed to the procedure depending on the existence of the switched instance according to the table below.

Whether to switch to the instance	Next Steps
There is a switched to instance.	Proceed to "A.3.2.2.3 Actions to take depending on the status of the target instance".
Destination Instance Does Not Exist	Proceed to "A.3.2.2.5 Creating the target instance".

If the Error Message Does Not Contain "new_instance_id = <*new_instance_id*>"

The destination instance does not exist.

Proceed to "A.3.2.2.5 Creating the target instance".

A.3.2.2.3 Actions to take depending on the status of the target instance

Check the status of the instance that you checked in "A.3.2.2.2 Checking the Existence of the target instance", and take appropriate action. The procedure is as follows:.

1. Check the status of the destination instance.

   From the AWS Management Console, open the Manage Amazon EC2 Service screen and check the Instance State column from the displayed list of instances.

2. Take action according to the status of the switched instance.

   Follow the table below to take action depending on the state of the switched instance.

Instance State	treatment content
terminated	Because the destination instance does not exist, skip to step 1 in "A.3.2.2.5 Creating the target instance".

Instance State	treatment content
Running	Because the destination instance was created successfully, skip to step 2 in "A.3.2.2.5 Creating the target instance". Note the instance ID of the switched instance from < *new_instance_id* > in the error message.
pending	Proceed to "A.3.2.2.4 Deleting a target instance" because the destination instance was not created successfully.
Other than the above	Because an unexpected error might have occurred while creating the instance, collect the survey information and contact your our company service representative (SE). If you are in a hurry to switch the system, collect the survey information and proceed to "A.3.2.2.4 Deleting a target instance".

A.3.2.2.4 Deleting a target instance

To delete the switched instance, use the AWS Management Console and do the following:.

1. Open the "Amazon EC2 Service" administration screen and select "Instances" from the sidebar.
2. From the list of displayed instances, select the check box for the instance.
3. From the Instance State drop-down list, select Terminate Instance.

A.3.2.2.5 Creating the target instance

To create a switched instance, use the AWS Management Console and do the following:.

1. Create a destination instance from the AMI.

   After creating the switched instance as described in "A.2.3.1 Creating Cluster Node Instance", note the instance ID. The instance ID of this switched to instance is used in Step 2, "A.3.2.2.6 Configuring CloudWatch Alarm", and Step 2 in "A.3.2.2.7 Configuring Amazon DynamoDB".

2. Configure the target group for the ELB.

   This step is not required if the target instance has already been added to the ELB target group.

   Add the instance ID of the switched instance to the ELB target group.

   For information about configuring target groups, see "A.2.3.2 Configuring Network Takeover". However, do not perform Step 1 of "A.2.3.2 Configuring Network Takeover" and add to an existing target group.

   If you provided the instance ID of the switched instance in step 2 of "A.3.2.2.3 Actions to take depending on the status of the target instance", add it to the target group.

A.3.2.2.6 Configuring CloudWatch Alarm

Configure CloudWatch alarms.

If the target instance is already configured for CloudWatch alarms, this step is not required.

Set the instance ID of the switched instance to the instance name (InstanceID) for the CloudWatch alarm.

For information about configuring CloudWatch alarms, see "A.2.4.3 Configuring CloudWatch Alarm".

If you recorded the instance ID of the switched instance in step 2 of "A.3.2.2.3 Actions to take depending on the status of the target instance", set the recorded instance ID to a CloudWatch alarm.

A.3.2.2.7 Configuring Amazon DynamoDB

1. Review the tables in Amazon DynamoDB.

   Verify that the table "Fujitsu-Pclswr-DB-Switcher" has the following entries:.

Attribute name	Value
SystemID	Same value as the tag for the instance on the cluster node "fujitsu.pclswr.id"

Attribute name	Value
InstanceID	Same as the <*instance id*> of the source instance in the error message

2. Update the table in Amazon DynamoDB according to the results of step 1.

   - If the item checked in step 1 exists:

     Update the table Fujitsu-Pclswr-DB-Switcher from step 1.

     Update the value of the attribute "InstanceID" to the instance ID of the switched instance.

     If you recorded the instance ID of the switched instance in step "A.3.2.2.3 Actions to take depending on the status of the target instance", set that instance ID to the attribute "InstanceID".

     If the value of attribute "State" is SWITCHING, update the value of attribute "State" to NOT_SWITCHED.

Attribute name	Value
SystemID	Same value as the tag for the instance on the cluster node "fujitsu.pclswr.id"
InstanceID	Instance ID of the switched instance
State	NOT_SWITCHED

   - If the item checked in step 1 does not exist:

     Add the following to the table "Fujitsu-Pclswr-DB-Switcher":. If you recorded the instance ID of the switched instance in step "A.3.2.2.3 Actions to take depending on the status of the target instance", set that instance ID to the attribute "InstanceID".

Attribute name	Value
SystemID	Same value as the tag for the instance on the cluster node "fujitsu.pclswr.id"
InstanceID	Instance ID of the switched instance
State	NOT_SWITCHED

A.3.2.3 Changing the settings of the switch destination AZ

Follow the actions for the error message identified in "A.3.2.1 Reviewing Error Messages", and take action in case the switch is interrupted due to AZ resource exhaustion. Because AZ resources are depleted, remove the subnet of the AZ whose resources are depleted from the switched AZ. Also, add a subnet for the AZ that has not experienced resource exhaustion to the switched AZ. Use the AWS Management Console to do the following:.

1. Create an instance according to "A.3.2.2.5 Creating the target instance" in the switch-source AZ.

2. Change the value of the tag "fujitsu.pclswr.is_recovery_target" of the subnet to which you want to switch AZ that is experiencing resource exhaustion to false.

3. Note the integer value that identifies the cluster node from the tag "fujitsu.pclswr.idlist" of the subnet to which AZ is switched when resource exhaustion occurs. The integer value that identifies the cluster node may have multiple values. This integer value is used in step 4.

4. Configure the subnet to set AZ as the switch destination.

   To add a new switched subnet

   a. Create a virtual system that contains the subnet to switch to, as described in "A.2.1 Creating Virtual System". Make sure that the subnet you create can use API endpoints and mount targets for EFS.

   b. Configure network takeover, as described in "A.2.3.2 Configuring Network Takeover". Make sure that NLB and ALB are available on the subnet you created.

   c. Add the integer value that identifies the cluster node you recorded in Step 3 to the subnet tag "fujitsu.pclswr.idlist". This step is not required if it has already been added.

   d. Change the subnet tag "fujitsu.pclswr.is_recovery_target" to true. This step is not required if it has already been changed.

   You already have a subnet to switch to.

a. Based on the value of the subnet tag "fujitsu.pclswr.idlist" and the integer value you recorded in Step 3 that identifies the cluster node, consider the integer value that identifies the cluster node you want to set as a tag for the source and destination subnets so that the switch destination is set to multiple AZ.

b. Change the subnet tag "fujitsu.pclswr.is_recovery_target" to true.

Example

Here's how to set the subnet tag "fujitsu.pclswr.idlist":.

The following is an example of setting change when resource exhaustion occurs in AZ of SubnetB.

Note that SubnetA, SubnetB, and SubnetC are different AZ subnets.

- Example 1) Only one configured subnet

  Change the value of the SubnetB tag "fujitsu.pclswr.is_recovery_target" to false and the SubnetC tag "fujitsu.pclswr.is_recovery_target" to true.

  The following shows the settings before and after the change.

Table A.1 Previous setting (only one subnet has been set)

Subnet Name	Tag "fujitsu.pclswr.idlist" Value	Tag "fujitsu.pclswr.is_recovery_target" Value
SubnetA	1,2,3,4	true
SubnetB	1,2,3,4	true
SubnetC	tag "fujitsu.pclswr.idlist" value	false

Table A.2 The new setting, if there is only one subnet.

Subnet Name	Tag "fujitsu.pclswr.idlist" Value	Tag "fujitsu.pclswr.is_recovery_target" Value
SubnetA	1,2,3,4	true
SubnetB	-	false
SubnetC	1,2,3,4	true

- Example 2) Two or more configured subnets

  Change the value of the SubnetB tag "fujitsu.pclswr.is_recovery_target" to false.

  The following shows the settings before and after the change.

Table A.3 Previous setting (if there are two or more configured subnets)

Subnet Name	Tag "fujitsu.pclswr.idlist" Value	Tag "fujitsu.pclswr.is_recovery_target" Value
SubnetA	1,3,4	true
SubnetB	1,2,4	true
SubnetC	2,3,4	true

Table A.4 The new setting, if there are two or more configured subnets

Subnet Name	Tag "fujitsu.pclswr.idlist" Value	Tag "fujitsu.pclswr.is_recovery_target" Value
SubnetA	1,2,3,4	true
SubnetB	-	false
SubnetC	1,2,3,4	true

A.4 Changing the Configuration

System configuration changes using the smart workload recovery feature include.

- Cluster application configuration changes
    - Add/Remove/Modify Cluster Applications
    - Add/Remove/Modify Resources to Cluster Applications
- Add/Remove Cluster Nodes
- Change the configuration of an instance on a cluster node
- Change settings of switching mechanism and resource monitoring mechanism

For instructions on modifying the cluster application configuration, see "Chapter 10 Configuration Change of Cluster Applications" in the "PRIMECLUSTER Installation and Administration Guide". However, you must do the following before stopping and after starting RMS.

Before stopping RMS

1. Disables tags on the instance.

   Use the AWS Management Console to set the value of the tag "fujitsu.pclswr.is_recovery_target" for the instance to false.

2. Disables CloudWatch Agent.

   Disables monitoring of RMS by the CloudWatch agent and changes the setting to log monitoring only. For more information, see "A.3.1.2 Operation when RMS startup is stopped / RMS monitoring is disabled".

After starting RMS

1. Enables CloudWatch Agent.

   Enable RMS monitoring in CloudWatch Agent configuration file. For more information, see "A.3.1.2 Operation when RMS startup is stopped / RMS monitoring is disabled".

2. Enables tags on the instance.

   Use the AWS Management Console to set the value of the tag "fujitsu.pclswr.is_recovery_target" for the instance to true.

If you want to add or remove cluster nodes, see "A.5.4 Procedure for Adding Cluster Node" and "A.5.5 Procedure for Removing Cluster Node".

For instructions on changing the configuration of instances on a cluster node, see the official AWS documentation.

For instructions on how to change the configuration of switcher and resource monitor, see the official AWS documentation. Note that operation is not guaranteed if the setting is changed to a value other than that described in "A.2.4 Construction of Switcher and Resource monitor".

Note

- If you made configuration changes to your instance (such as changing the instance type), you must recreate the AMI. For instructions about creating an AMI in the operating instance, see Create an AMI "A.5.8 Procedure for Getting AMI in Operation".
- Always perform a switch test after a configuration change to detect misconfiguration settings.

A.5 Maintenance

This subsection describes maintenance on a system using the smart workload recovery feature.

A.5.1 Software Maintenance

A.5.1.1 Notes on applying software maintenance fixes

When applying fixes to a system that uses the smart workload recovery feature, note the following.

- Refer to "A.5.1.2 Summary of Application/Removal Procedure" and apply/delete in multi-user mode.

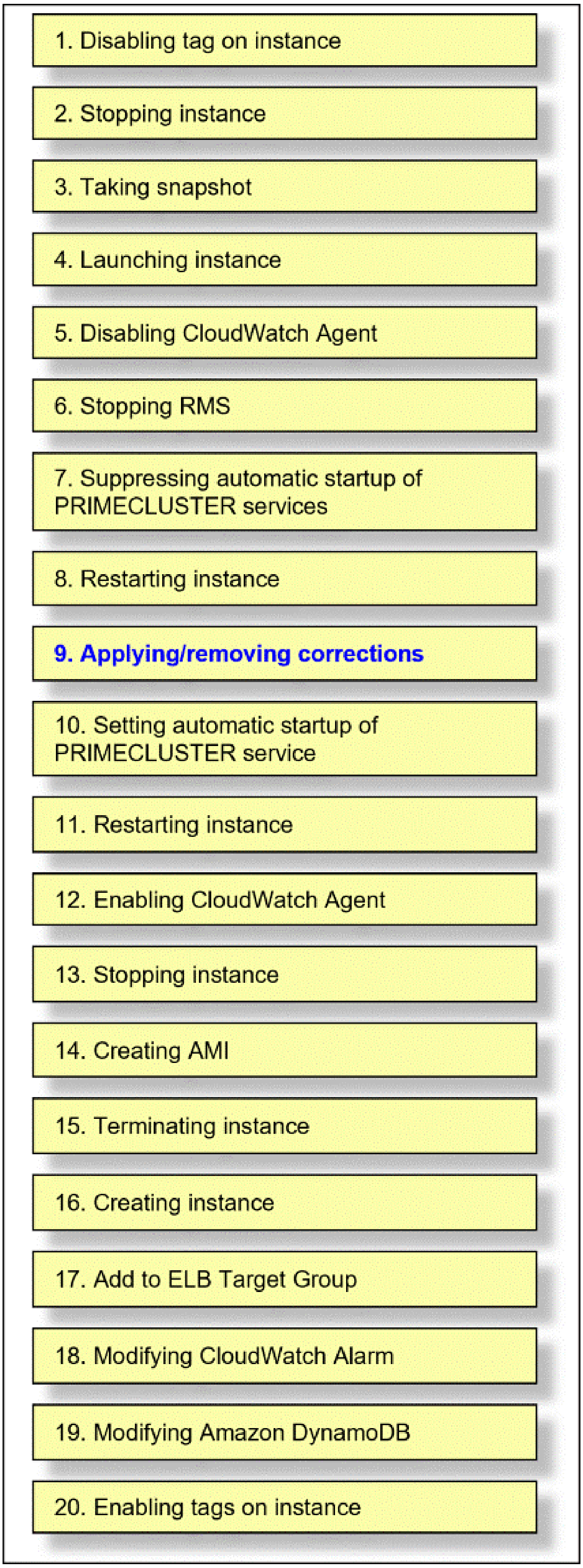
A.5.1.2 Summary of Application/Removal Procedure

Provides an overview of the procedure for applying various corrections to a system that uses the smart workload recovery feature.

Note

- Be sure to take a snapshot and back up the storage for the instance before applying/removing the corrections.
- Snapshots can only be taken when the OS is stopped.

Figure A.8 Flow of Operations


1. Disabling tag on instance
2. Stopping instance
3. Taking snapshot
4. Launching instance
5. Disabling CloudWatch Agent
6. Stopping RMS
7. Suppressing automatic startup of PRIMECLUSTER services
8. Restarting instance
9. Applying/removing corrections
10. Setting automatic startup of PRIMECLUSTER service
11. Restarting instance
12. Enabling CloudWatch Agent
13. Stopping instance
14. Creating AMI
15. Terminating instance
16. Creating instance
17. Add to ELB Target Group
18. Modifying CloudWatch Alarm
19. Modifying Amazon DynamoDB
20. Enabling tags on instance

Procedure

Copy the corrections that should be applied to the instance to the instance's local file system in advance.

1. Disabling tag on instance

   Using the AWS Management Console, set the value of the tag "fujitsu.pclswr.is_recovery_target" for the instance of the cluster node to false.

2. Stopping instance

   Using the AWS Management Console, stop the instance of the cluster node.

3. Taking snapshot

   Using the AWS Management Console, take a snapshot of the instance of the cluster node.

4. Launching instance

   Using the AWS Management Console, launch an instance of the cluster node.

5. Disabling CloudWatch Agent

   Disables monitoring of RMS by the CloudWatch agent and changes the setting to log monitoring only. For more information, see "A.3.1.2 Operation when RMS startup is stopped / RMS monitoring is disabled".

6. Stopping RMS

   Stop RMS by running the following command on the instance of the cluster node.

   ```
   # hvshut -a
   ```

7. Suppressing automatic startup of PRIMECLUSTER services

   Execute the following command on the instance of the cluster node to suppress the automatic startup of the PRIMECLUSTER service.

   ```
   # /opt/FJSVpclinst/bin/pclservice off
   ```

8. Restarting instance

   Restart the instance by running the following command on the instance of the cluster node .

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

9. Applying/removing corrections

   Apply the corrections previously copied to the local file system on the instance of the cluster node.

   Or delete the correction.

   - Applying corrections

     After copying the corrections to the working directory, run the following command.

     ```
     # cd <working directory>
     # /opt/FJSVfupde/bin/uam add -d ./ -i <correction number>
     ```

     If the following message is output at this time, select "Y".

     When displayed in Japanese

     ```
     シングルユーザモードで適用が必要な修正です。マルチユーザモードで修正を適用しますか？(Y/N)Y
     ```

     When displayed in English

     ```
     It is required to update with single user mode. Do you want to apply the update now? (Y/N)Y
     ```

     When the following message is output, select "N".

     When displayed in Japanese

```
すぐに再起動しますか？(Y/N)N
```

When displayed in English

```
Do you want to restart your computer immediately? (Y/N)N
```

- Deleting corrections

  Run the following command on the instance.

```
# /opt/FJSVfupde/bin/uam remove -i <correction number>
```

  If the following message is output at this time, select "Y".

  When displayed in Japanese

```
シングルユーザモードで復元が必要な修正です。マルチユーザモードで修正を復元しますか？(Y/N)Y
```

  When displayed in English

```
It is required to restore with single user mode. Do you want to restore the updated product to
its pre-update state now? (Y/N)Y
```

  When the following message is output, select "N".

  When displayed in Japanese

```
すぐに再起動しますか？(Y/N)N
```

  When displayed in English

```
Do you want to restart your computer immediately? (Y/N)N
```

10. Setting automatic startup of PRIMECLUSTER service

    Execute the following command on the instance of the cluster node to restore the PRIMECLUSTER service settings that you suppressed in step 7.

```
# /opt/FJSVpclinst/bin/pclservice on
```

11. Restarting instance

    Restart the instance by running the following command on the instance of the cluster node .

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

12. Enabling CloudWatch Agent

    Enables monitoring of RMS with the CloudWatch agent. For more information, see "A.3.1.2 Operation when RMS startup is stopped / RMS monitoring is disabled".

13. Stopping instance

    Using the AWS Management Console, stop an instance on a cluster node. Note the integer value and instance ID that identifies the cluster node, which is the value of the cluster node instance tag "fujitsu.pclswr.id".

14. Creating AMI

    Using the AWS Management Console, create an AMI for the instances on your cluster nodes.

15. Terminating instance

    Terminate the instance that you stopped in step 13.

16. Create instance

    Create an instance from the AMI created in Step 14, as described in "A.2.3.1 Creating Cluster Node Instance". Set the value of the cluster node instance tag "fujitsu.pclswr.id" to an integer value that identifies the cluster node in step 13. Also note the instance ID of the instance you created.

17. Add to ELB Target Group

    Add the instance you created to the ELB's target group.

18. Modifying CloudWatch Alarm

    Change the instance ID for the CloudWatch alarm from the instance ID in Step 13 to the instance ID in Step 16. Determine what to change by the integer value that identifies the cluster node at the end of the alarm name. For more information about alarm names, see "A.2.4.3 Configuring CloudWatch Alarm".

19. Modifying Amazon DynamoDB

    Modifies the items of a eligible instance registered to the table "Fujitsu-Pclswr-DB-Switcher" in Amazon DynamoDB. The items to be modified are checked from the items "SystemID" and "InstanceID". Change the value of the item "InstanceID" from the instance ID in step 13 to the instance ID in step 16. For more information about the items, see "A.2.4.4 Configuring Amazon DynamoDB".

20. Enabling tags on instance

    Using the AWS Management Console, verify that the value of the tag "fujitsu.pclswr.is_recovery_target" for the instance on the cluster node is true.

Note

- If you want to apply a fix to the switcher Lambda function, use the AWS Management Console to reconfigure the Lambda function that contains the fix. See "A.2.4 Construction of Switcher and Resource monitor" for instructions.

See

For instructions about using the AWS Management Console, see the official AWS documentation.

A.5.2 Procedure for Starting Cluster Node

To start a cluster node, use the AWS Management Console and perform the following steps.

1. Verify that the tag "fujitsu.pclswr.is_recovery_target" value for the instance of the cluster node is false. Set to false if not false.
2. Start the instance of the cluster node.
3. Sets the value of the tag "fujitsu.pclswr.is_recovery_target" of the instance of the cluster node to true.

A.5.3 Procedure for Stopping Cluster Node

To stop a cluster node, use the AWS Management Console and perform the following steps.

1. Sets the tag "fujitsu.pclswr.is_recovery_target" value of the instance of the cluster node to false.
2. Stop the instance of the cluster node.

A.5.4 Procedure for Adding Cluster Node

To add a new cluster node, use the AWS Management Console and perform the following steps.

1. Configure a security group on the VPC endpoint as described in "A.2.1.2 Creating API Endpoint".
2. Create EFS according to "A.2.1.3 Creating Data Storage Area".
3. Create an AMI as described in "A.2.2 Creating Image".
4. Create an instance as described in "A.2.3 Creating Cluster Node".
5. Configure CloudWatch alarms as described in "A.2.4.3 Configuring CloudWatch Alarm".
6. Add items to the table as described in Creating Items in "A.2.4.4 Configuring Amazon DynamoDB".

7. Add an integer value that identifies the cluster node to the subnet tag "fujitsu.pclswr.idlist". For more information about tags, see "A.2.1.1 Creating VPC and Subnet".

8. Add an integer value to the AWS Lambda environment variable "PCLSWR_SYSTEM_LIST" that identifies the cluster node. For more information about environment variables, see "A.2.4.1 Configuring AWS Lambda".

9. Verify that the deployment was successful according to the "A.2.5 Post-Implementation Review".

Note

The options you choose when you create the instance (Instance type, set instance details, add storage, etc.) depend on the requirements of the cluster node that you are adding.

A.5.5 Procedure for Removing Cluster Node

To remove an existing cluster node, use the AWS Management Console and perform the following steps.

1. Set the value of the tag "fujitsu.pclswr.is_recovery_target" of the instance of the cluster node to false.

2. Stop the instance of the cluster node.

3. Note the integer value and instance ID that identifies the cluster node with the tag "fujitsu.pclswr.id" for the instance that you stopped in step 2.

4. Terminate the instance that you stopped in step 2.

5. Deregister the AMI for the instance that you stopped in step 2.

6. If you created a snapshot on a volume attached to the instance you stopped in step 2, delete the snapshot.

7. Delete the load balancer for the ELB used by the instance in step 2.

8. Delete the ELB target group used by the instance in step 2.

9. Delete the mounted EFS file system on the instance in step 2.

10. Delete the CloudWatch alarm that you configured for the instance in step 2. Determine what to delete by the integer value that identifies the cluster node at the end of the alarm name. For more information about alarm names, see "A.2.4.3 Configuring CloudWatch Alarm".

11. Remove the security group that allows the instances in step 2 that you are configuring on your VPC endpoint.

12. Delete the integer value that identifies the cluster node that you want to delete from the AWS Lambda environment variable "PCLSWR_SYSTEM_LIST". For more information about environment variables, see "A.2.4.1 Configuring AWS Lambda".

13. Remove the integer value that identifies the cluster node to remove from the subnet tag "fujitsu.pclswr.idlist". For more information about tags, see "A.2.1.1 Creating VPC and Subnet".

14. Delete the item for the deleted instance that is registered to the table "Fujitsu-Pclswr-DB-Switcher" in Amazon DynamoDB. The items for the instance to be deleted are checked from the items "SystemID" and "InstanceID". For more information about the items, see "A.2.4.4 Configuring Amazon DynamoDB".

A.5.6 Procedure for Restoring OS with the Snapshot Function

If the OS is restored with the snapshot function of AWS, perform the following procedure.

1. Set the value of the tag "fujitsu.pclswr.is_recovery_target" of the instance of the cluster node to false.

2. Restore a snapshot of a volume (EBS) attached to an instance.

   For information about restoring snapshots and attaching EBS volumes, see the official AWS documentation.

3. Stop the instance of the cluster node. Note the integer value and instance ID that identifies the cluster node, which is the value of the cluster node instance tag "fujitsu.pclswr.id".

4. Create an AMI for the instance of the cluster node.

5. Terminate the instance that you stopped in step 3.

6. Create an instance according to "A.2.3.1 Creating Cluster Node Instance" from the AMI created in step 4. At this time, set the value of the cluster node instance tag "fujitsu.pclswr.id" to an integer value that identifies the cluster node in step 3. Also note the instance ID of the instance you created.

7. Add the instance you created to the ELB's target group.

8. Change the instance ID of the CloudWatch alarm from the instance ID in Step 3 to the instance ID in Step 6.Determine what to change by the integer value that identifies the cluster node at the end of the alarm name. For more information about alarm names, see "A.2.4.3 Configuring CloudWatch Alarm".

9. Modifies the items of a eligible instance registered to the table "Fujitsu-Pclswr-DB-Switcher" in Amazon DynamoDB. The items to be modified are checked from the items "SystemID" and "InstanceID". Change the value of the item "InstanceID" from the instance ID in step 3 to the instance ID in step 6. For more information about the items, see "A.2.4.4 Configuring Amazon DynamoDB".

10. Verify that the tag "fujitsu.pclswr.is_recovery_target" value of the instance you created in step 6 is true.

See

For instructions about using the AWS Management Console, see the official AWS documentation.

Note

- In a cluster system that uses the smart workload recovery feature of PRIMECLUSTER, you cannot restore using the snapshot function while the business is still in operation.

- Synchronize with EFS if restoring with snapshot.

A.5.7 Procedure for changing the operating AZ

To change the running AZ while the instance is running, use the AWS Management Console and perform the following steps.

1. Set the value of the tag "fujitsu.pclswr.is_recovery_target" of the instance of the cluster node to false.

2. Stop the instance of the cluster node. Note the integer value and instance ID that identifies the cluster node, which is the value of the cluster node instance tag "fujitsu.pclswr.id".

3. Terminate the instance that you stopped in step 2.

4. Create an instance in the AZ to be modified according to "A.2.3.1 Creating Cluster Node Instance" from the AMI tied to the instance of the cluster node. At this time, set the value of the cluster node instance tag "fujitsu.pclswr.id" to an integer value that identifies the cluster node in Step 2. Also note the instance ID of the instance you created.

5. Add the instance you created to the ELB's target group.

6. Change the instance ID of the CloudWatch alarm from the instance ID in Step 2 to the instance ID in Step 4. Determine what to change by the integer value that identifies the cluster node at the end of the alarm name. For more information about alarm names, see "A.2.4.3 Configuring CloudWatch Alarm".

7. Modify the items of a eligible instance registered to the table "Fujitsu-Pclswr-DB-Switcher" in Amazon DynamoDB. The items to be modified are checked from the items "SystemID" and "InstanceID". Change the value of the item "InstanceID" from the instance ID in step 2 to the instance ID in step 4. For more information about the items, see "A.2.4.4 Configuring Amazon DynamoDB".

8. Verify that the tag "fujitsu.pclswr.is_recovery_target" for the instance you created in step 4 has a value of true.

See

For instructions about using the AWS Management Console, see the official AWS documentation.

A.5.8 Procedure for Getting AMI in Operation

To get an AMI while your instance is running, use the AWS Management Console and perform the following steps.

1. Set the value of the tag "fujitsu.pclswr.is_recovery_target" of the instance of the cluster node to false.
2. Stop the instance of the cluster node. Note the integer value and instance ID that identifies the cluster node, which is the value of the cluster node instance tag "fujitsu.pclswr.id".
3. Create an AMI for the instance of the cluster node.
4. Terminate the instance that you stopped in step 2.
5. Create an instance from the AMI created in step 3 by following the instructions "A.2.3.1 Creating Cluster Node Instance". At this time, set the value of the cluster node instance tag "fujitsu.pclswr.id" to an integer value that identifies the cluster node in step 2. Also note the instance ID of the instance you created.
6. Add the instance you created to the ELB's target group.
7. Change the instance ID for the CloudWatch alarm from the instance ID in step 2 to the instance ID in step 5. Determine what to change by the integer value that identifies the cluster node at the end of the alarm name. For more information about alarm names, see "A.2.4.3 Configuring CloudWatch Alarm".
8. Modify the items of a target instance registered to the table "Fujitsu-Pclswr-DB-Switcher" in Amazon DynamoDB. Check the items "SystemID" and "InstanceID" for the items to be changed. Change the value of the item "InstanceID" from the instance ID in step 2 to the instance ID in step 5. For more information about the items, see "A.2.4.4 Configuring Amazon DynamoDB".
9. Verify that the tag "fujitsu.pclswr.is_recovery_target" value of the instance you created in step 5 is true.

See

For instructions about using the AWS Management Console, see the official AWS documentation.

A.5.9 Procedure for deleting cloud resources

To finish using the smart workload recovery feature, use the AWS Management Console and follow these steps to delete your cloud resources.

Note

This procedure also deletes the user's application and various data. Back up important data before performing this procedure.

1. Set the value of the tag "fujitsu.pclswr.is_recovery_target" of the instance of the cluster node to false.
2. Stop the instance of the cluster node.
3. Terminate the instance of the cluster node.
4. Deregister the AMI associated with the instance in step 3.
5. If you created a snapshot on a volume attached to the instance in step 3, delete the snapshot.
6. If the IAM role attached to the instance in step 3 is not used by another instance, delete the IAM role.

   If the IAM role is being used by other instances, and you do not want CloudWatchAgentServerPolicy policy on other instances, remove it from the IAM role.
7. Delete Lambda function in the target system.
8. Delete IAM role that was used for the Lambda function in Step 7.
9. Delete policy that you created for use with the Lambda function in Step 7.
10. Delete the Amazon EventBridge rules in the system you want to delete.

11. Delete CloudWatch alarms used on the instance in step 3.
12. Delete load balancer for the ELB used by the instance in step 3.
13. Delete target group for the ELB that was used in the instance in Step 3.
14. Delete mounted EFS file system on the instance in step 3.
15. Delete Amazon DynamoDB table in the target system.
16. Delete security group in the target system.
17. Delete log group for CloudWatch Logs in the deleted system.
18. If it is not being used by another system, remove the following.
    - Security groups configured for the PrivateLink VPC and VPC endpoint in the deleted system
    - Subnets in the target system
    - VPC in the system to be deleted

See

For instructions about using the AWS Management Console, see the official AWS documentation.

Appendix B Changes in Each Version

For details on the changes made to the specifications of PRIMECLUSTER 4.7A00, refer to "Appendix M Changes in Each Version" in "PRIMECLUSTER Installation and Administration Guide."

Appendix C Release Information

This appendix lists the main changes in this manual.

No	Edition	Location	Description
1	Second edition	1.1 Supported Range	Added the note about the range of support for FJcloud-O environments.
2	Second edition	3.6.1 Initial GLS Setup	Changed values set for parameters of the nmcli connection modify command.
3	Second edition	19.3 Supported Range	Added the note about the range of support for AWS environments.
4	Second edition	21.10.1.3.2 Setting Takeover Network Resources Used for the Network Takeover by the Virtual Router	Added the note about takeover network resources in AWS environment.
5	Second edition	25.3 Supported Range	Added the note about the range of support for Azure environments.
6	Second edition	27.3 Presetting 3.Register service principals and set certificates.	Changed incompatibility notes for Azure CLI versions.
7	Second edition	27.9.1.2.1 Setting PreOnline scripts for Network Takeover with Secondary IP address change	Changed ScriptTimeout value calculation when used in network takeover with secondary IP address replacement in Azure environment.
8	Second edition	27.9.1.2.2 Setting up the Takeover Network Resources	Added the note about takeover network resources for Azure Environments.
9	Second edition	29.1.3 Changing the Role of the Service Principal	Added the notes about incompatibilities in the Azure CLI version.

Glossary

API (application program interface)

Refer to Application Program Interface.

API endpoint

The destination of a communication when a client accesses the API.

Application Program Interface

A shared boundary between a service provider and the application that uses that service.

Availability Zone (AZ)

A unit for sharing physical facilities, such as data center facilities and service provision facilities.

CF (Cluster Foundation)

Refer to Cluster Foundation.

CIM

Cluster Integrity Monitor

CIP

Cluster Interconnect Protocol

CLI

command-line interface

Cluster Foundation

The set of PRIMECLUSTER modules that provides basic clustering communication services.

cluster interconnect

The set of private network connections used exclusively for PRIMECLUSTER communications.

Global Disk Services

This optional product provides volume management that improves the availability and manageability of information stored on the disk unit.

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.

guest OS (Guest OS)

An OS running on a virtual machine.

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.

node

A host which is a member of a cluster. A computer node is a computer.

public LAN

The local area network (LAN) by which normal users access a machine.

Reliant Monitor Services (RMS)

The package that maintains high availability of user-specified resources by providing monitoring and switchover capabilities.

RMS (Reliant Monitor Services)

Refer to Reliant Monitor Services.

SA

Shutdown Agent. SA forcibly stops the target node by receiving instructions from the Shutdown Facility.

SF

Shutdown Facility

Shutdown Facility

A facility that forcibly stops a node in which a failure has occurred. When PRIMECLUSTER decides that system has reached a state in which the quorum is not maintained, it uses the Shutdown Facility (SF) to return the cluster system to the quorum state.

Virtual Machine

A logical computer separated from a physical computer by virtualization technology.

VPC

Virtual Private Cloud. A virtual network dedicated to your AWS account.

Web-Based Admin View

This is a common base enabling use of the Graphic User Interface of PRIMECLUSTER. This interface is in Java.

Index