FUJITSU

FUJITSU Software PRIMECLUSTER

Installation and Administration Guide 4.4 FUJITSU Cloud Service K5

Linux

J2UL-2112-01ENZ0(01)
June 2017

Preface

This manual serves as your starting point for using PRIMECLUSTER on FUJITSU Cloud Service K5.

This manual explains the workflow of the series of operations from installation to operation management of the PRIMECLUSTER system on FUJITSU Cloud Service K5.

Target Readers

This manual is intended for all users who use PRIMECLUSTER 4.4 and perform cluster system installation and operation management on FUJITSU Cloud Service K5. It is also intended for programmers who develop applications that operate on PRIMECLUSTER.

Related Documentation

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

- PRIMECLUSTER Installation and Administration Guide
- PRIMECLUSTER Concepts 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 Web-Based Admin View Operation Guide
- PRIMECLUSTER Messages
- FJQSS (Information Collection Tool) User's Guide
- PRIMECLUSTER Software Release Guide and Installation Guide
- FUJITSU Cloud Service K5 IaaS Service Portal User Guide
- FUJITSU Cloud Service K5 IaaS API User Guide
- FUJITSU Cloud Service K5 IaaS API Reference
- FUJITSU Cloud Service K5 IaaS Features Handbook

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 "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 Cloud Service K5 is abbreviated as K5.

Export Controls

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

Trademarks

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

Linux is a registered trademark of Linus Torvalds.

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

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

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

Requests

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

Date of publication and edition

April 2017, First edition
June 2017, 1.1 edition

Copyright notice

Revision History

Revision	Location	Edition
Changed the section title and the description of "2.2 Setting Up NTP."	Chapter 2 Installation	1.1
Added the note about GDS settings.	2.8 Building Cluster Application	
Changed the maintenance procedure.	5.2.2 Overview of the Procedure for Applying/ Deleting Corrections	

Contents

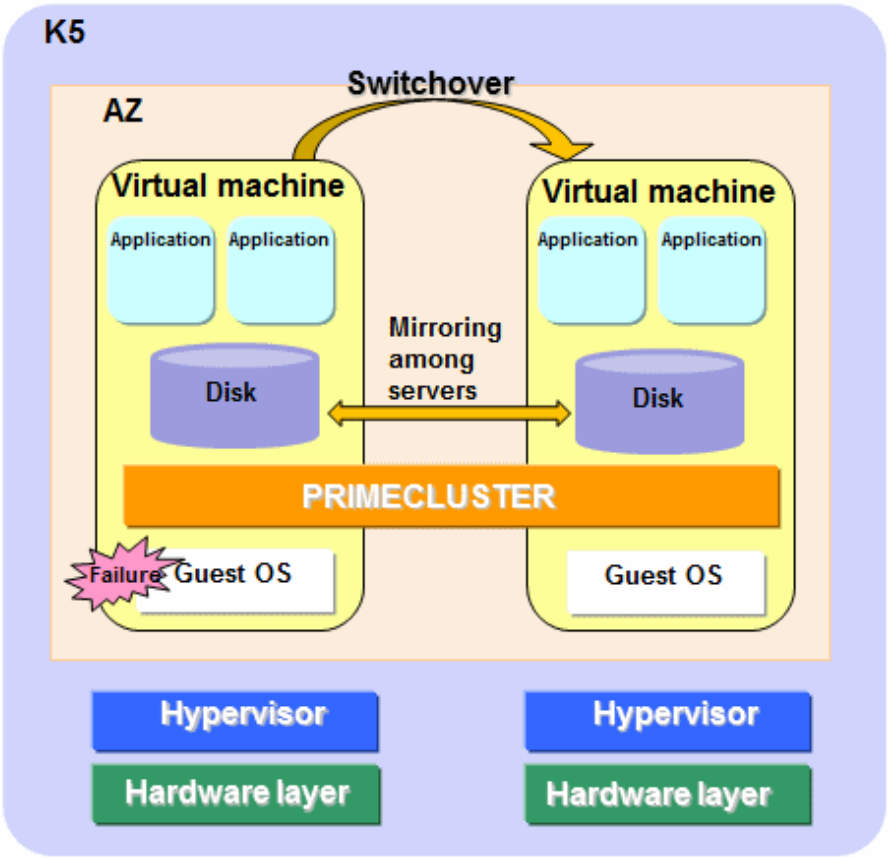
Chapter 1 Cluster System on K5

On K5, PRIMECLUSTER can be used on the virtual servers that are built within the same availability zone (hereinafter referred to as "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 on K5



See

For details on K5, see "FUJITSU Cloud Service K5 IaaS Service Portal User Guide", "FUJITSU Cloud Service K5 IaaS API User Guide" and "FUJITSU Cloud Service K5 IaaS API Reference."

1.1 Supported Range

This section explains the range of support of PRIMECLUSTER on K5.

Supported configurations

- Number of cluster nodes : Two nodes
- Operation mode of the cluster system:1:1 Standby operation, Mutual standby

- 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.
  - The virtual servers in the cluster system must communicate with the API endpoints.
- 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.

Supported monitoring functions

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

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

- Failure of the shared disk and disk access path

  By combining the volume management function (GDS), a failure of the disk access and disk access path can be detected (monitored by the Gds resource), and the system 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.

- Cluster application failure

  The system is switched to the standby system when a resource failure of the cluster application occurs.

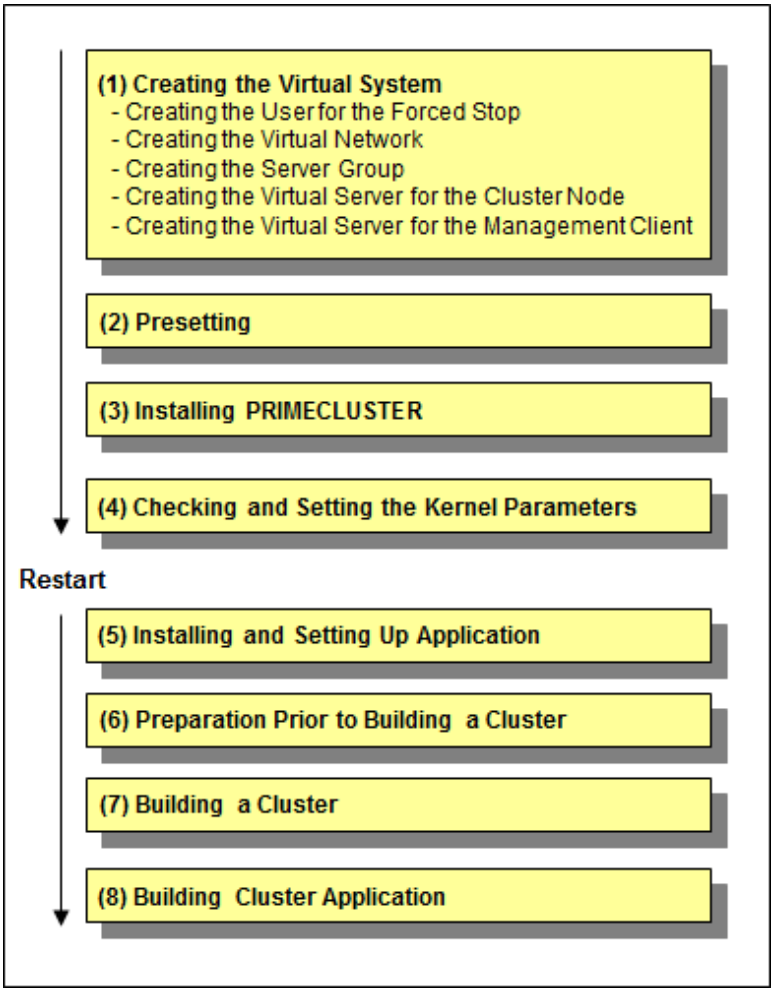
Note

- 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 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
  - GDS Snapshot
  - Root class and local class of GDS
  - Single volume, mirroring, concatenation, and striping of GDS
  - Scalable operation
  - Cluster application using the takeover network resource
- The following functions for K5 are not available:
  - Auto-scaling
  - Automatic failover
- Duplicate virtual server names cannot be used in the project on K5.
- The console cannot be used in the virtual server on K5. Do not set the single user mode.

Chapter 2 Installation

This chapter describes the procedure to install PRIMECLUSTER on K5.

Perform the steps shown in the figure below.



2.1 Creating the Virtual System

This section explains how to create the virtual system for a cluster system on K5.

See

For details on the setup methods for K5, see "FUJITSU Cloud Service K5 IaaS API User Guide" and "FUJITSU Cloud Service K5 IaaS API Reference."

Note

To use the service provided by FUJITSU Cloud Service K5 IaaS with API, it is necessary to build an environment for using API. Refer to "FUJITSU Cloud Service K5 IaaS API User Guide" to build an environment for using API.

2.1.1 Creating the User for the Forced Stop

Create the user on K5 to forcibly stop the virtual server in the cluster system with the following values.

Item name	Value
User name	Arbitrary user name
Default project ID	Project ID in the project creating the virtual server
Role	Operator role

2.1.2 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)

2.1.2.1 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 on K5 and then connect to the subnet for the public LAN (used also for the administrative LAN).

2.1.2.2 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	tcp	80	80
egress	IP address of DNS server	udp	53	53
egress	IP address of NTP server	udp	123	123

2.1.2.3 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	Not specified	tcp	443	443
ingress	Own security group	udp	9382	9382

Communication direction	Communication destination	Protocol information	Starting port number	Ending port number
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	Own security group	tcp	3260	3260
egress	Own security group	tcp	3260	3260

2.1.2.4 Creating the Security Group for the Cluster Interconnect

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

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

2.1.2.5 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

2.1.2.6 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

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

2.1.2.7 Creating the Firewall Rule

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

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

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	Ending port number
egress	Subnet for the public LAN (used also for the administrative LAN)	IP address to the repository server	80	Allow

2.1.3 Creating the Server Group

Set the server group as the following so that the virtual servers in the cluster start on the different physical hosts.

Item	Value
Server group name	Arbitrary server group name
Policies of server group	anti-affinity
Availability zone	Availability zone to assign the virtual server

Execution API (Example)

```
# SERVER_GROUP_NAME=<server group name>
# AZ=<availability zone>
# curl --tlsv1.2 -s $COMPUTE/v2/$PROJECT_ID/os-server-groups -X POST -H "X-Auth-Token:
$OS_AUTH_TOKEN" -H "Content-Type: application/json" -d '{"server_group": { "name":
"'$SERVER_GROUP_NAME'", "policies": [ "anti-affinity" ] }, "availability_zone": "'$AZ'"}' | jq .
```

Note

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

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

2.1.4.1 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 configuring the cluster system as follows.

Table 2.1 Port to be created in the subnet of the public LAN and the administrative LAN

Item	Value
Port name	Arbitrary port name
Network ID	Network ID
Subnet ID	Network ID of the subnet for the public LAN (used also for the administrative LAN) created in "2.1.2.1 Creating Subnets"
Private IP address	IP address of the public LAN (used also for the administrative LAN)
ID list of the security group	- ID of the security group created in "2.1.2.2 Creating the Common Security Group" - ID of the security group created in "2.1.2.3 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 "2.1.2.5 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 "2.1.2.6 Creating the Security Group for the Virtual Server Access" - If there are any necessary security groups for operations other than those above, add them.

Item	Value
IP address list of Allowed address pairs	Takeover IP address
Availability zone	Availability zone to assign the virtual server

Execution API (Example)

```
# PORT_NAME=<port name>
# NETWORK_ID=<network ID>
# SUBNET_ID=<subnet ID>
# FIXED_IP_ADDRESS=<private IP address>
# SG_ID1=<ID of the common security group>
# SG_ID2=<ID of the security group for the public LAN (used also for the administrative LAN)>
# SG_ID3=<ID of the security group for Web-Based Admin View>
# SG_ID4=<ID of the security group for the virtual server access>
# ALLOWED_ADDRESS_PAIRS=<takeover IP address>
# AZ=<availability zone>
# curl --tlsv1.2 -s $NETWORK/v2.0/ports -X POST -H "X-Auth-Token: $OS_AUTH_TOKEN" -H "Content-Type:
application/json" -d '{"port":{"network_id": "'$NETWORK_ID'", "name": "'$PORT_NAME'",
"availability_zone" : "'$AZ'", "allowed_address_pairs": [{"ip_address":
"'$ALLOWED_ADDRESS_PAIRS'"}], "fixed_ips": [{"subnet_id": "'$SUBNET_ID'", "ip_address":
"'$FIXED_IP_ADDRESS'"}], "security_groups": ["'$SG_ID1'", "'$SG_ID2'", "'$SG_ID3'",
"'$SG_ID4'"] } }' | jq .
```

Note

If you want to enhance the security, the security groups created in "2.1.2.5 Creating the Security Groups for Web-Based Admin View" and "2.1.2.6 Creating the Security Group for the Virtual Server Access" can be added or deleted from the setting for the port as necessary after the installation.

2.1.4.2 Creating the Port for the Cluster Interconnect

Set the port for the cluster interconnect in the virtual servers in the cluster system as follows.

Table 2.2 Port to be created in the subnet of the cluster interconnect

Item	Value
Port name	Arbitrary port name
Network ID	Network ID
Subnet ID	Network ID of the subnet for the public LAN (used also for the administrative LAN) created in "2.1.2.1 Creating Subnets"
Private IP address	IP address for the cluster interconnect
ID list of the security group	Security group for the cluster interconnect created in "2.1.2.4 Creating the Security Group for the Cluster Interconnect"
Availability zone	Availability zone to assign the virtual server

Execution API (Example)

```
# PORT_NAME=<port name>
# NETWORK_ID=<network ID>
# SUBNET_ID=<subnet ID>
# FIXED_IP_ADDRESS=<private IP address>
# SG_ID=<ID of the security group for the cluster interconnect>
# AZ=<availability zone>
# curl --tlsv1.2 -s $NETWORK/v2.0/ports -X POST -H "X-Auth-Token: $OS_AUTH_TOKEN" -H "Content-Type:
application/json" -d '{"port":{"network_id": "'$NETWORK_ID'", "name": "'$PORT_NAME'",
```

```
"availability_zone" : "'$AZ'", "fixed_ips": [{"subnet_id": "'$SUBNET_ID'", "ip_address":
"'$FIXED_IP_ADDRESS'"}], "security_groups": ["'$SG_ID'"] } }' | jq .
```

2.1.4.3 Creating the Virtual Server

Set the virtual servers in the cluster system as follows.

Item	Value
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	Flavor ID of the arbitrary virtual server type as the performance requirement *Refer to "FUJITSU Cloud Service K5 IaaS API User Guide" to acquire the flavor ID.
OS image	RedHat Enterprise Linux 6.x 64bit (English)
Connection port	Port (eth0) created in "2.1.5.1 Creating the Port for the Public LAN (Used also for the Administrative LAN)" Port (eth1) created in "2.1.4.2 Creating the Port for the Cluster Interconnect"
Security group	Not specified (Specified in the port)
Automatic failover	Disabled
Server group ID	Server group ID created in "2.1.3 Creating the Server Group"
Minimum number of servers	1
Maximum number of servers	1
Availability zone	Availability zone to assign the virtual server

Execution API (Example)

```
# VM_NAME=<virtual server name>
# FLAVOR_REF=<ID of the virtual server type>
# OS_IMAGE_ID=<ID on the OS image>
# VOL_SIZE=<volume size>
# IS_DELETE=<0:Do not delete the block storages in deleting the virtual server.1: Delete the block
storages in deleting the virtual server>
# KEYNAME=<Keypair name>
# PORT_ID1=<port ID of eth0>
# PORT_ID2=<port ID of eth1>
# SERVER_GROUP_ID=<server group ID>
# AZ=<availability zone>
# curl --tlsv1.2 -i $COMPUTE/v2/$PROJECT_ID/servers -X POST -H "X-Auth-Token: $OS_AUTH_TOKEN" -H
"Content-Type: application/json" -d '{"server": {"name": "'$VM_NAME'", "availability_zone":
"'$AZ'", "imageRef": "", "flavorRef": "'$FLAVOR_REF'", "block_device_mapping_v2":[ {"boot_index":
"0", "uuid":"'$OS_IMAGE_ID'", "volume_size": "'$VOL_SIZE'", "device_name": "/dev/vda",
"source_type": "image", "destination_type": "volume", "delete_on_termination": "'$IS_DELETE'"} ] ,
"key_name": "'$KEYNAME'", "max_count": "1", "min_count": "1", "networks": [{"port":
"'$PORT_ID1'"}, {"port": "'$PORT_ID2'"}] }, "os:scheduler_hints": {"group":
"'$SERVER_GROUP_ID'"} }'
```

2.1.4.4 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 "2.1.4.3 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.

2.1.4.5 Setting Up 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.
- If /etc/sysconfig/network-scripts/ifcfg-eth1 does not exist, set the following.

  1. Create /etc/sysconfig/network-scripts/ifcfg-eth1 as follows.

     ```
     DEVICE=eth1
     TYPE=Ethernet
     ONBOOT=yes
     NM_CONTROLLED=no
     BOOTPROTO=dhcp
     ```

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

     ```
     # chown root:root /etc/sysconfig/network-scripts/ifcfg-eth1
     # chmod 644 /etc/sysconfig/network-scripts/ifcfg-eth1
     ```

Set the following in each node building the cluster.

1. Add the following lines to the file, /etc/sysconfig/network-scripts/ifcfg-eth0.

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

2. Add the following into the file, /etc/sysconfig/network-scripts/ifcfg-eth1.

   ```
   PEERDNS=no
   ```

3. Add the following into the file, /etc/sysconfig/network.

   ```
   GATEWAYDEV=eth0
   ```

4. Restart the network service.

   ```
   # service network restart
   ```

2.1.4.6 Application of the Necessary OS Patch

Refer to "FUJITSU Cloud Service K5 IaaS Features Handbook" 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
```

2.1.4.7 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
```

2.1.5 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.

2.1.5.1 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.

Table 2.3 Port to be created in the subnet of the public LAN and the administrative LAN

Item	Value
Port name	Arbitrary port name
Network ID	Network ID
Subnet ID	Network ID of the subnet for the public LAN (used also for the administrative LAN) created in "2.1.2.1 Creating Subnets"
Private IP address	IP address of the management client
ID list of the security group	- ID of the security group created in "2.1.2.2 Creating the Common Security Group" - ID of the security group on the management client side created in "2.1.2.5 Creating the Security Groups for Web-Based Admin View" - ID of the security group for installing and maintaining the management client created in "2.1.2.6 Creating the Security Group for the Virtual Server Access" - If there are any necessary security groups for operations, add them.
Availability zone	Availability zone to assign the virtual server

Execution API (Example)

```
# PORT_NAME=<port name>
# NETWORK_ID=<network ID>
# SUBNET_ID=<subnet ID>
# FIXED_IP_ADDRESS=<private IP address>
# SG_ID1=<ID of the common security group>
# SG_ID2=<ID of the security group for Web-Based Admin View>
# SG_ID3=<ID of the security group for the virtual server access>
# AZ=<availability zone>
# curl --tlsv1.2 -s $NETWORK/v2.0/ports -X POST -H "X-Auth-Token: $OS_AUTH_TOKEN" -H "Content-Type:
application/json" -d '{"port":{"network_id": "'$NETWORK_ID'", "name": "'$PORT_NAME'",
"availability_zone" : "'$AZ'", "fixed_ips": [{"subnet_id": "'$SUBNET_ID'", "ip_address":
"'$FIXED_IP_ADDRESS'"}], "security_groups": ["'$SG_ID1'", "'$SG_ID2'", "'$SG_ID3'"] } }' | jq .
```

Note

If you want to enhance the security, the security groups created in "2.1.2.5 Creating the Security Groups for Web-Based Admin View" and "2.1.2.6 Creating the Security Group for the Virtual Server Access" can be added or deleted from the setting for the port as necessary after the installation.

2.1.5.2 Creating the Virtual Server

Create the virtual server of the management client.

Set the virtual server of the management client as follows.

Item	Value
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	Flavor ID of the arbitrary virtual server type as the performance requirement *Refer to "FUJITSU Cloud Service K5 IaaS API User Guide" to acquire the flavor ID.
OS image	Any one of the following OS images: Windows Server 2012 R2 64bit Windows Server 2012 64bit Windows Server 2008 R2 SP1 64bit
Connection port	Port (eth0) created in "2.1.5.1 Creating the Port for the Public LAN (Used also for the Administrative LAN)"
Security group	Not specified (Specified in the port)
Availability zone	Availability zone to assign the virtual server

Execution API (Example)

```
# VM_NAME=<virtual server name>
# FLAVOR_REF=<ID of the virtual server type>
# OS_IMAGE_ID=<ID on the OS image>
# VOL_SIZE=<volume size>
# IS_DELETE=<0:Do not delete the block storages in deleting the virtual server.1: Delete the block storages in deleting the virtual server>
# PORT_ID=<port ID of eth0>
# ADMIN_PASS=<password >
# AZ=<availability zone>
# curl --tlsv1.2 -i $COMPUTE/v2/$PROJECT_ID/servers -X POST -H "X-Auth-Token: $OS_AUTH_TOKEN" -H "Content-Type: application/json" -d '{"server": {"name": "'$VM_NAME'", "availability_zone": "'$AZ'", "imageRef": "", "flavorRef": "'$FLAVOR_REF'", "block_device_mapping_v2":[ {"boot_index": "0", "uuid":"'$OS_IMAGE_ID'", "volume_size": "'$VOL_SIZE'", "device_name": "/dev/vda", "source_type": "image", "destination_type": "volume", "delete_on_termination": "'$IS_DELETE'"} ] , "networks": [{"port": "'$PORT_ID'"}], "metadata": {"admin_pass": "'$ADMIN_PASS'" } } }'
```

2.2 Presetting

1. Disabling firewall

   Make sure that iptables and ip6tables are disabled.

```
# chkconfig --list iptables
# chkconfig --list ip6tables
```

   If they are enabled, disable them.

```
# service iptables stop
# chkconfig iptables off
# service ip6tables stop
# chkconfig ip6tables off
```

2. Setting up 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.

2.3 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.

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

For details on the installation procedure, see Installation Guide for PRIMECLUSTER.

2.4 Checking and Setting the Kernel Parameters

Change the kernel parameters depending on the environment.

Applicable nodes:

All the nodes on which PRIMECLUSTER is to be installed

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

Check PRIMECLUSTER Designsheets and if modifying the kernel parameters is needed, set them again.

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

To activate the modified kernel parameters, restart OS.

2.5 Installing and Setting Up Application

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

- For details on environment setup, see manuals for each application.
- For information on PRIMECLUSTER-related products supporting K5, see the documentation for each product.

2.6 Preparation Prior to Building a Cluster

Prior to building a cluster, perform presettings such as the initial GLS setup, setting the DNS client, creating information files on K5, and staring the Web-Based Admin View screen.

2.6.1 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 the procedure below. For details on each setting, see "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. System setup

   1. Define the IP address and the host name in the file, /etc/hosts.

      ### 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. In the /etc/sysconfig/network-scripts/ifcfg-eth0 file, comment out TYPE. Set BOOTPROTO=static and PEERDNS=no. Add HOTPLUG=no and DEVICETYPE=hanet.

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

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

   3. Set GATEWAYDEV of the file, /etc/sysconfig/network to sha0.

      ```
      GATEWAYDEV=sha0
      ```

2. Creating the virtual interface

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

3. Setting up the virtual interface

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

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

   ```
   DEVICE=sha0
   #IPADDR=
   #NETMASK=
   ```

```
BOOTPROTO=dhcp
ONBOOT=yes
DEVICETYPE=sha
HOTPLUG=no
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.

4. Setting up 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. Setting up 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. Creating the takeover virtual interface

Example

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

7. Checking 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. Restarting the system

   Run the following command and restart the system.

/sbin/shutdown -r now

2.6.2 Creating Information Files on K5

To activate a cluster system on K5, create the information file on K5 with the following procedure.

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

```
DOMAIN_NAME=K5DomainName
PROJECT_NAME=K5ProjectName
IDENTITY=IdentityURL
COMPUTE=ComputeURL
```

K5DomainName : Domain name of FUJITSU Cloud Service K5

K5ProjectName: Project name building a cluster in FUJITSU Cloud Service K5

IdentityURL : URL of the endpoint for the identity service of the region used in FUJITSU Cloud Service K5(*)

ComputeURL : URL of the endpoint for the computer service of the region used in FUJITSU Cloud Service K5(*)

*For details on URL of the endpoint for the identity service and the compute service, see init.sh created in the preparation for an environment setup in "FUJITSU Cloud Service K5 IaaS API User Guide."

Example

```
DOMAIN_NAME=primecluster_domain
PROJECT_NAME=primecluster_project
IDENTITY=https://identity.cloud.global.fujitsu.com
COMPUTE=https://compute.jp-east-1.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
```

2.6.3 Presettings for Building a Cluster

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

Note

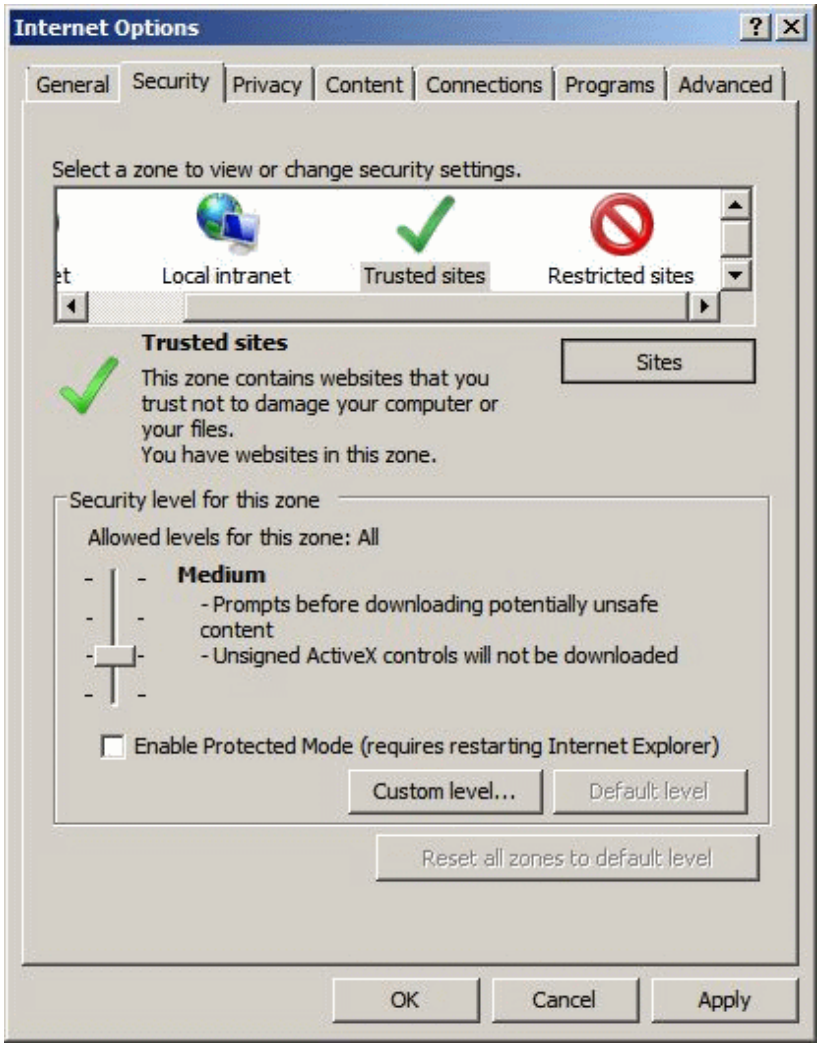
- For checking the operation environment, see "Chapter 2 Operation Environment " in the Installation Guide for PRIMECLUSTER.
- If you cannot connect to the Java SE download page in Oracle Corp from the management client, transfer the Java by copy & paste to the destination management client from the connection source of the remote desktop.
- When Web-Based Admin View is connected after performing the preparations for starting the Web-Based Admin View screen, the connection is not available with the following pop up on the display.

Perform the following procedure to allow the connection to the Web-Based Admin View.

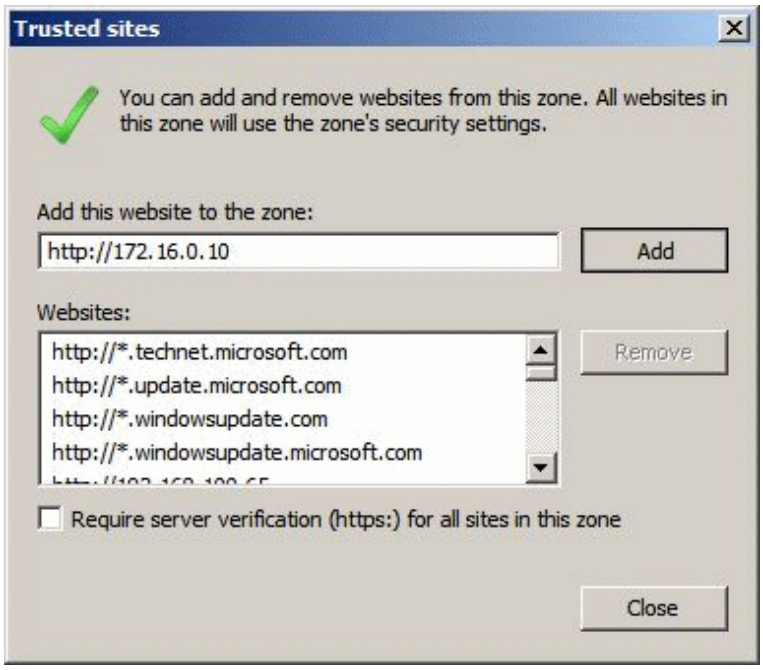
1. Start InternetExplorer.

2. Select [Internet options] in [Tools] menu.

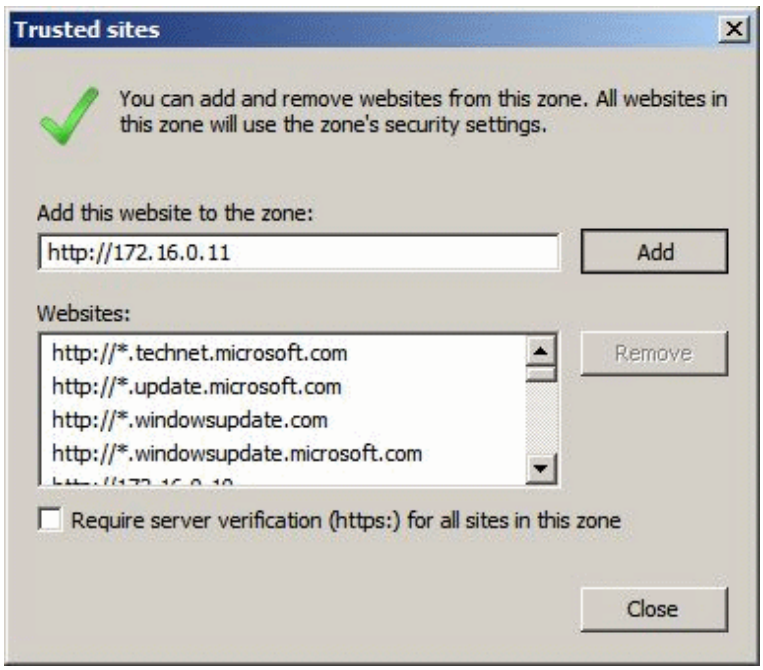
3. Select [Trusted sites] in [Security] tab and click [Sites].



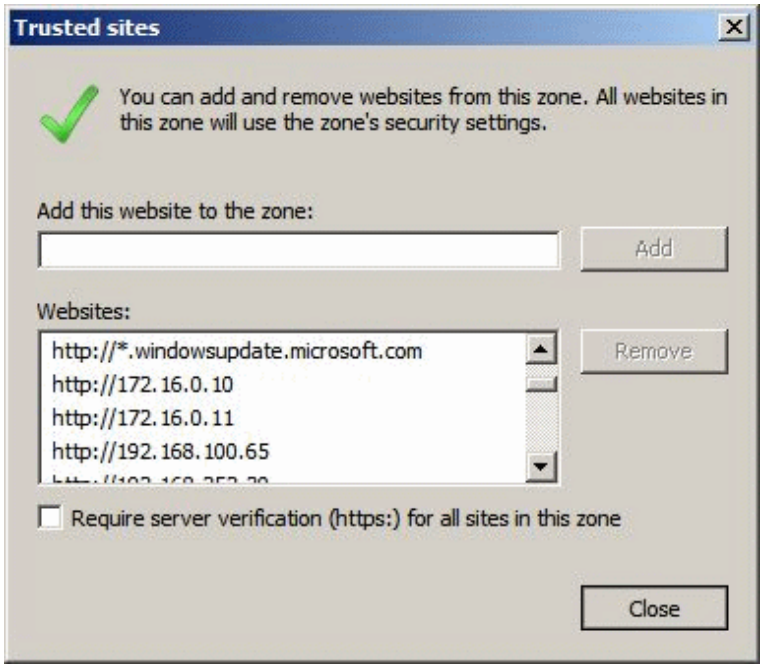
4. Enter the URL to connect to the primary management server in the format http://<IP address> and then click [Add].



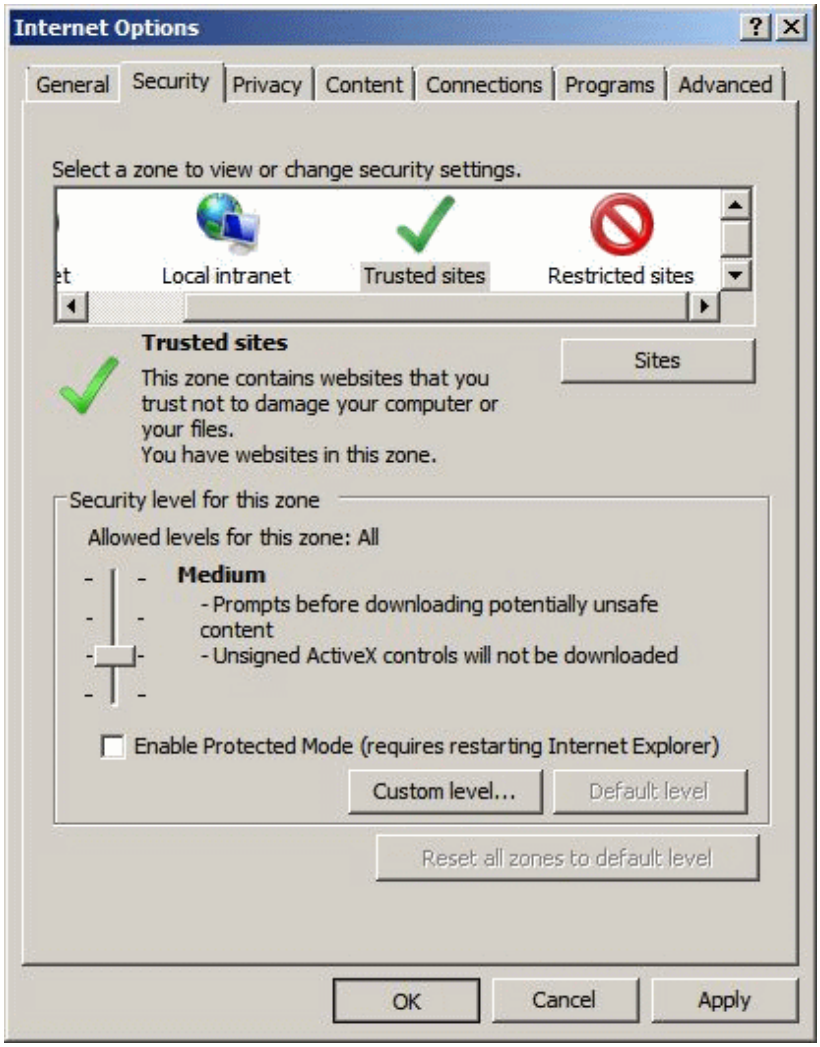
5. Enter the URL to connect to the secondary management server in the format http://<IP address> and then click [Add].



6. Make sure that connection URLs for both the primary management server and the secondary management server are set. Click [Close].

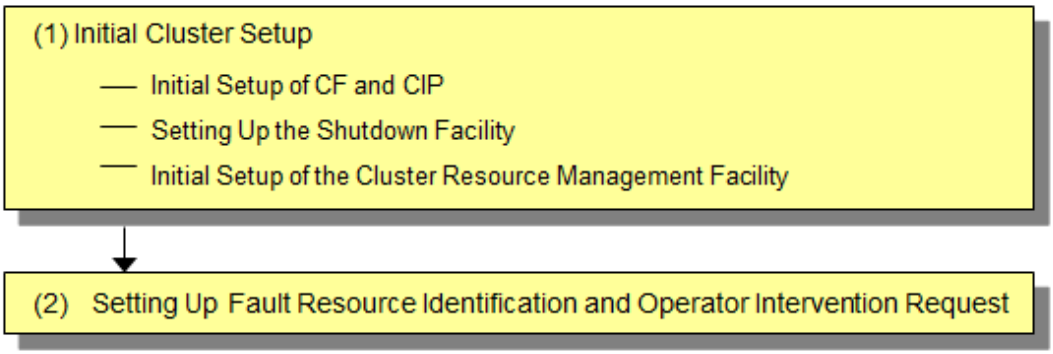


7. Click [OK] to exit from the Internet Options window.



2.7 Building a Cluster

The procedure for building a PRIMECLUSTER cluster is shown below:



2.7.1 Initial Cluster Setup

This section describes the initial cluster setup for PRIMECLUSTER.

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

	Details	Manual reference location*
1	2.7.1.1 Initial Setup of CF and CIP (setting up cluster configuration information and IP addresses)	**CF** "2.1 CF, CIP, and CIM configuration"
2	2.7.1.2 Setting Up the Shutdown Facility	**CF** "8 Shutdown Facility (SF)"
3	2.7.1.3 Initial Setup of the Cluster Resource Management Facility	**CF** "4.3 Resource Database configuration"

*****The PRIMECLUSTER manual name is abbreviated as follows:

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

2.7.1.1 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.

2.7.1.2 Setting Up the Shutdown Facility

On K5, only SA_vmk5 shutdown agent is available for setup.

This section explains the method for setting up the SA_vmk5 shutdown agent as the shutdown facility.

For details on the survival priority, see "5.1.2.2 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, see "1.4 Test" in "PRIMECLUSTER Installation and Administration Guide."
- The contents of the SA_vmk5r.cfg, rcsd.cfg files of all the nodes should be identical. If not, a malfunction will occur.
- If you changed a user password created in "2.1.1 Creating the User for the Forced Stop", perform this step again with a new password.
- Be sure to perform the following operations on all the nodes.

1. Setting up the shutdown daemon

   Create /etc/opt/SMAW/SMAWsf/rcsd.cfg with the following contents on all the 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
```

2. Encrypt the password

   Execute the sfcipher command to encrypt a password of a user for the instance control on K5. For details on how to use the sfcipher command, see 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. Setting up the shutdown agent

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

   Delimit each item with a single space.

```
CFNameX InstanceName user passwd {cycle | leave-off}
CFNameX InstanceName user passwd {cycle | leave-off}
```

```
CFNameX          : Specify the CF node name of the cluster host.
InstanceName     : Specify the instance name on K5 for which a cluster operates.
user             : Specify a user name for the instance control on K5.
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 instance names are instance1 and instance2.
   - The user name to control the instance is pcl.
   - The node will be restarted when it is forcibly stopped.

```
# cat /etc/opt/SMAW/SMAWsf/SA_vmk5r.cfg
node1 instance1 pcl O/gm+AYuWwE7ow3dgVG/Nw== cycle
node2 instance2 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 is correctly set. If the setting is incorrect, the shutdown facility cannot be performed normally.
- Make sure that the instance name (InstanceName) 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 inappropriate node may be forcibly stopped.

4. Starting up the shutdown facility

   Start or restart the shutdown facility on all the nodes in the cluster system.

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

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

     ```
     # sdtool -b
     ```

   - If the shutdown daemon (rcsd) is active:

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

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

   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. Checking the status of the shutdown facility

   Execute the following command with all the 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 "2.1.1 Creating the User for the Forced Stop" needs a periodical change of the password (every 90 days). For procedure on changing a password, see "5.1 Changing a Password Periodically."
   - If you changed the virtual server name created in "2.1.4 Creating the Virtual Server for the Cluster Node", perform the procedure from step 3 to 5 again.

   Information

   Display results of the sdtool -s command

   - If Init State displays Unknown or Init-ing, wait for about one minute and then check it again.
   - If "Unknown" is displayed as the stop or initial status, it means that the SF has still not executed node stop, path testing, or SA initialization. "Unknown" is displayed temporarily as the test status and initial status until the actual status can be confirmed.
   - If "TestFailed" is displayed as the test status, it means that a problem occurred while the agent was testing whether or not the node displayed in the cluster host field could be stopped. Some sort of problem probably occurred in the software, hardware, or network resources being used by that agent.
   - If InitFailed in Init State is displayed, communication with the endpoint of the identity service or the compute service on K5 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 cluster can communicate with the identity service from the active instance.

        ```
        # curl -k -s -X GET <URL of the endpoint of identity service>/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 "2.1.4.6 Application of the Necessary OS Patch."

- .curlrc. must be created.
- Refer to "2.1.4.7 Creating .curlrc" to make sure that .curlrc is created according to the procedure.

- The security service groups and the firewall service on K5 must be set properly.

- The virtual router of K5 must be created.

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

- URL of the endpoint of the identity service must be correct.

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

b. Execute the following command and check if the cluster can communicate with the compute service from the active instance.

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

If the following message is 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 a message other than the above message is displayed, check the following.

- The security service groups and the firewall service on K5 must be set properly.
- The virtual router of K5 must be created.
- The default router of the cluster host must be set in the virtual router.
- URL of the endpoint of the compute service must be correct.
- The DNS server used in a 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 identity service, and URL of the endpoint for the compute service for FUJITSU Cloud Service on K5 information file (/opt/SMAW/SMAWRrms/etc/k5_endpoint.cfg)

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

2.7.1.3 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 (hereinafter referred to as "CRM"). Set the iSCSI used in the mirroring among the servers of GDS and register to the resource data base in this setting.

2.7.2 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.

2.8 Building Cluster Application

For the detail on how to build the cluster application, refer to "Chapter 6 Building Cluster Application" 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 "2.6.1 Initial GLS Setup" above.

Note

Change and set the values of the following tuning parameters configured with "Setting Tuning Parameters" in setting procedures of iSCSI device for GDS ("Disk Setting for Performing Mirroring among Servers" in "PRIMECLUSTER Global Disk Services Configuration and Administration Guide 4.4").

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
```

Also, specify the IP address of 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.

Chapter 3 Operations

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

See

For detail on how to operate GDS, see "Operations and Maintenance" of "PRIMECLUSTER Global Link Services Configuration and Administration Guide" and how to operate GLS, see "Operations in the Cluster System" in "PRIMECLUSTER Global Link Services Configuration and Administration Guide: Redundant Line Control Function."

Note

On K5, when a failure occurs in the network node or the storage controller, or due to the schedule maintenance for the infrastructure, a heartbeat fails and a cluster application might have to be switched.

Chapter 4 Changing the Configurations

For details on changing the configuration information of PRIMECLUSTER system, environmental settings, the configuration of cluster application, the operation attributes of a cluster system, see "Chapter 9 Changing the Cluster System Environment", "Chapter 10 Configuration Change of Cluster Application", "Chapter 11 Changing the Operation Attributes of a Cluster System" in "PRIMECLUSTER Installation and Administration Guide." For details on changing the GDS configuration, see "Chapter 8 Changing the Configuration" in "PRIMECLUSTER Global Disk Services Guide."

Note

If you change the network environment, the single user mode cannot be set. Refer to the procedure of "4.1 Changing IP address for the Cluster Interconnect " to perform the operation for changing the IP address on K5.

4.1 Changing IP address for the Cluster Interconnect

This section describes how to change the IP address for the cluster interconnect after building the cluster system.

Operation Procedure:

1. Edit /etc/default/cluster on all the nodes in the cluster system to change the IP address and the IP address of the remote node.

```
nodename <CF node name>
clustername <cluster name>
device /dev/ip0 <IP address> <broadcast address> <IP address of the remote node>
```

2. Shut down the system on all the nodes in the cluster.
3. Change the setting of the port created in "2.1.4.2 Creating the Port for the Cluster Interconnect" to change the IP address.
4. Start the system on all the nodes in the cluster..
5. Check the CF settings.

   Check the following items.

   - Checking that all the nodes join the cluster.
   - Execute the following command on any one of nodes in the cluster system and then make sure that CF node name of all the nodes are displayed on "Node" and "UP" is displayed on "State" for each node.

```
# cftool -n
```

Example

```
# cftool -n
Node   Number   State   Os     Cpu
node1       1      UP   Linux  EM64T
node2       2      UP   Linux  EM64T
```

Make sure that CF node name of all the nodes are displayed on "Node" and "UP" is displayed on "State" for each node.

   - Checking that the setting of CF over IP is enabled.

     Execute the following command on all the nodes in the cluster system and then make sure that "Device" displays "/dev/ip0."

```
# cftool -d
```

Example

```
# cftool -d
Number Device   Type  Speed  Mtu   State   Configured   Address
4      /dev/ip0  6     n/a   1392  UP      YES          0a.00.00.c9.00.0
```

Make sure that "Device" displays "/dev/ip0" only.

If an error occurs in the above step, double-check that both the IP address of the cluster interconnect and the IP address of the remote node are correctly set in /etc/default/cluster.

See

For details on the "cftool" command, see the manual page describing "cftool."

Chapter 5 Maintenance

When you maintain the PRIMECLUSTER system on K5, note the following points:

- For the procedure for applying/deleting urgent corrections on K5, see "5.2 Software Maintenance."
- For details on other items and procedures required for maintenance of the PRIMECLUSTER system, see "Chapter 12 Maintenance of the PRIMECLUSTER System" in "PRIMECLUSTER Installation and Administration Guide." For details on how to maintain GDS, see "Chapter 7 Operations and Maintenance" in "PRIMECLUSTER Global Disk Services Guide." For details on how to maintain GLS, see "Chapter 6 Maintenance" in "PRIMECLUSTER Global Link Services Configuration and Administration Guide : Redundant Line Control Function."

5.1 Changing a Password Periodically

A user created in "2.1.1 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 on K5.

2. Set the shutdown facility again with the changed password according to step 2 to 5 in "2.7.1.2 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
```

5.2 Software Maintenance

5.2.1 Notes on Applying Corrections to the PRIMECLUSTER System

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

On K5, refer to "5.2.2 Overview of the Procedure for Applying/Deleting Corrections" to apply/delete the corrections in multi-user mode.

5.2.2 Overview of the Procedure for Applying/Deleting Corrections

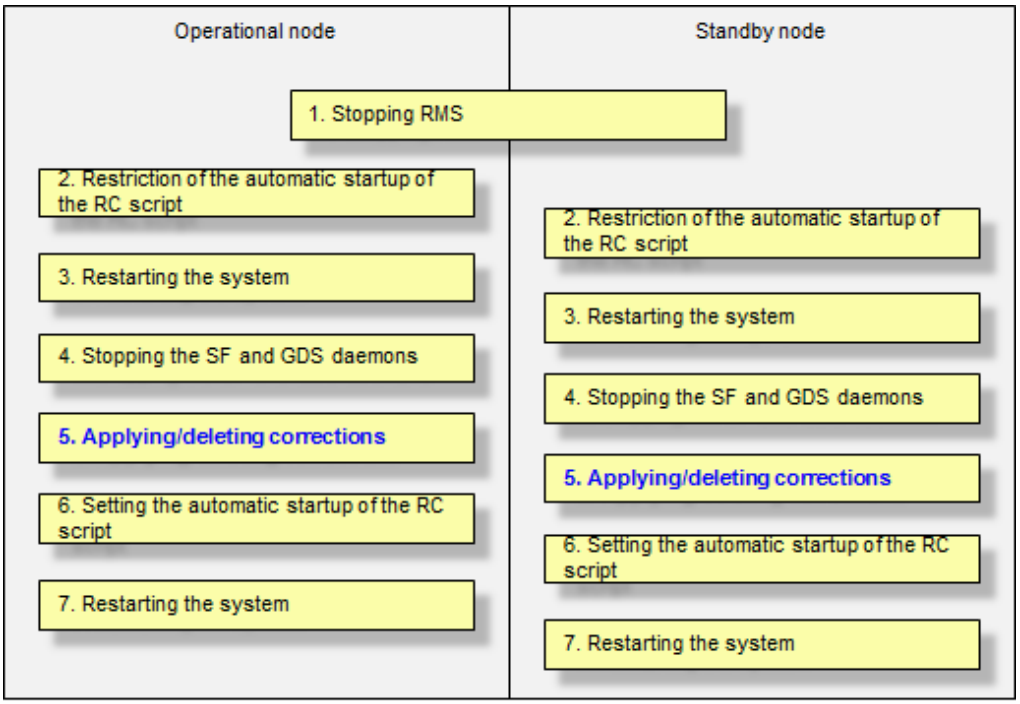
Overview of the procedure is shown for applying each correction including intensive corrections to the cluster system on K5. On the environment not using GDS, the procedure related to GDS is not necessary.

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

5.2.2.1 Procedure for Applying/Deleting Corrections by Stopping an Entire System

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

Flow of operation



GDS: Global Disk Services

Operation procedure

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

1. Stopping RMS

   If RMS is running, execute the following command on one of the cluster nodes to stop RMS.

   ```
   # hvshut -a
   ```

   Note

   If RMS is stopped on all the 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 the 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 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. Restriction of the automatic startup of the RC script

   Restrict the automatic startup of the RC script by executing the following command on all the nodes.

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

3. Restarting the system

   Restart the system on all the nodes.

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

4. Stopping the SF and GDS daemons

   Execute the following command on all the nodes to stop the SF daemon.

```
# initctl stop sf
```

   Execute the following command on all the nodes to stop the GDS daemon.

```
# initctl stop sdxm
```

5. Applying/deleting corrections

   Apply the corrections that were copied to the local file 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>
```

     At this time, 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>
```

     At this time, 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
```

6. Setting the automatic startup of the RC script

   Execute the following command on all the nodes and return the RC script setting restricted in step 2.

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

7. Restarting the system

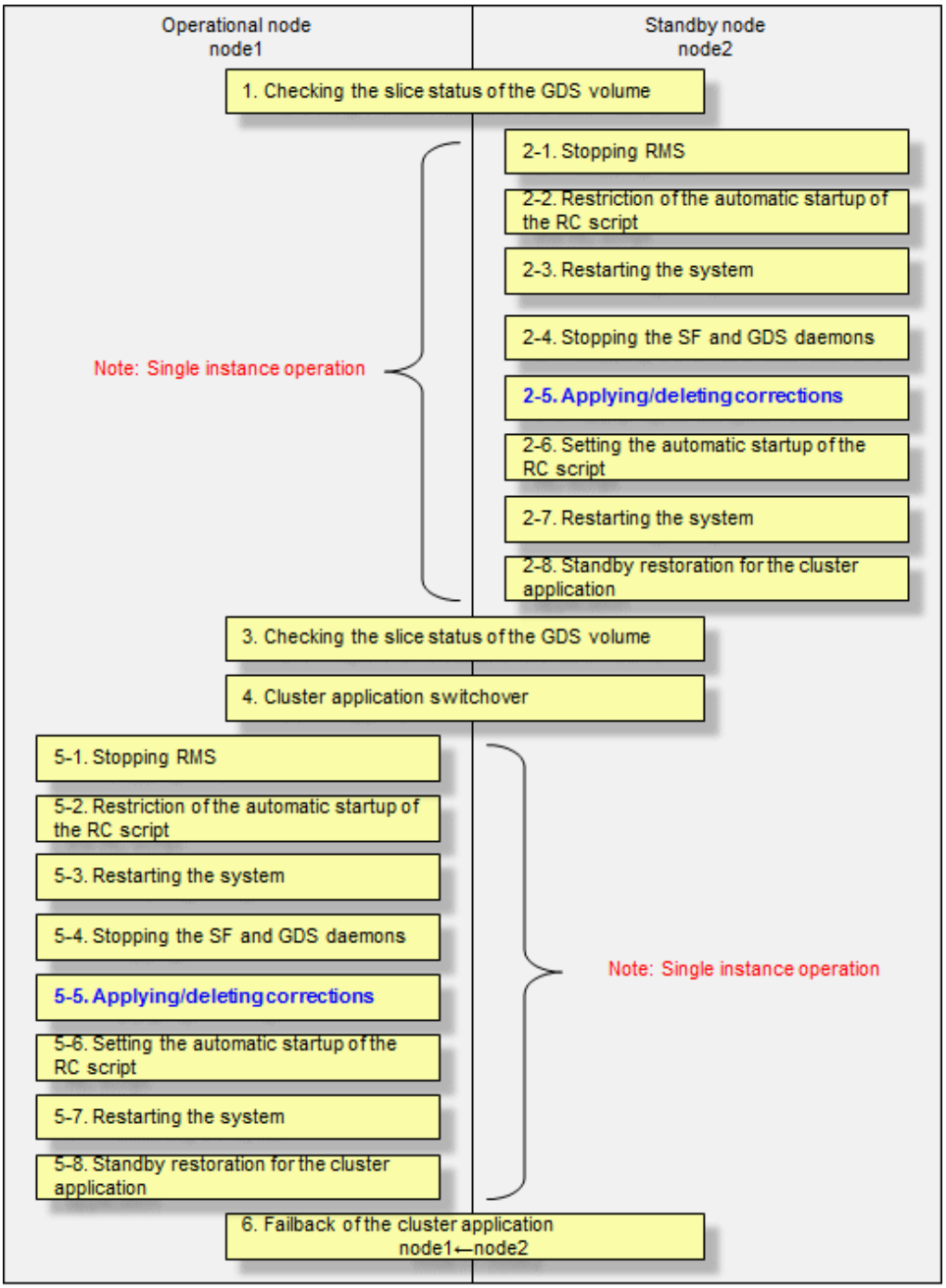
   Restart the system on all the nodes.

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

5.2.2.2 Procedure for Applying/Deleting Corrections by Rolling Update

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

Flow of operation



Operation procedure:

1. Checking 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.

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

   1. Stopping 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. Restriction of the automatic startup of the RC script

      Execute the following command to restrict the automatic startup of the RC script.

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

   3. Restarting the system

      Restart the system.

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

   4. Stopping the SF and GDS daemons

      Execute the following command on all the nodes to stop the SF daemon.

      ```
      # initctl stop sf
      ```

      Execute the following command on all the nodes to stop the GDS daemon.

      ```
      # initctl stop sdxm
      ```

   5. Applying/deleting corrections

      Apply/delete 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>
      ```

      At this time, 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>
      ```

      At this time, 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
      ```

6. Setting the automatic startup of the RC script

   Execute the following command to return the RC script setting restricted in step 2.

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

7. Restarting the system

   Restart the system.

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

8. 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, see "7.2.2.1 Starting Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

3. Checking 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 the 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. Cluster application switchover

   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, see "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. Stopping 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. Restriction of the automatic startup of the RC script

      Execute the following command to restrict the automatic startup of the RC script.

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

   3. Restarting the system

      Restart the system.

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

   4. Stopping the SF and GDS daemons

      Execute the following commands to stop the SF and GDS daemons.

      ```
      # initctl stop sf
      # initctl stop sdxm
      ```

   5. Applying/deleting corrections

      Apply/delete 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>
```

At this time, 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>
```

At this time, 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
```

6. Setting the automatic startup of the RC script

   Execute the following command and return the RC script setting restricted in 2 of step 5.

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

7. Restarting the system

   Restart the system.

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

8. 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, see "7.2.2.1 Starting Cluster Application" in "PRIMECLUSTER Installation and Administration Guide."

6. Failback of the cluster application

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

5.3 Procedure for Restoring OS with the Snapshot Function

If OS is restored with the snapshot function on K5, perform the following procedure.

5.3.1 Procedure for Restoring One Node While the Operation is Working

1. Refer to "5.3.3 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 "5.3.3 Restoring the Virtual Server from the Snapshot ", refer to "5.3.3 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 "2.7.1.2 Setting Up the Shutdown Facility" for both nodes.

5.3.2 Procedure for Restoring Nodes While the Operation does not Work

1. If any node has been already started before restoring nodes, stop the RMS. Perform this procedure with either one of both nodes that are started.

```
# hvshut -a
```

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

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

3. See "5.3.3 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 "5.3.3 Restoring the Virtual Server from the Snapshot ", refer to "5.3.3 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, see "5.3.3 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 "5.3.3 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 both 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 with either one of both nodes.

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

8. Start the RMS with the node restored in step 1.

```
# hvcm -a
```

Note

If the virtual server name is changed in step 3 or step 5, perform step 3 to 5 in "2.7.1.2 Setting Up the Shutdown Facility" for both nodes.

5.3.3 Restoring the Virtual Server from the Snapshot

Note

To use the service provided by FUJITSU Cloud Service K5 IaaS with API, it is necessary to build an environment for using API. Refer to "FUJITSU Cloud Service K5 IaaS API User Guide" to build an environment for using API.

1. Check the type of the virtual server to be restored and ID of the expanded volume.
2. Delete the virtual server to be restored.
3. Perform "2.1.4.1 Creating the Port for the Public LAN (Used also for the Administrative LAN)" and "2.1.4.2 Creating the Port for the Cluster Interconnect" to create a port.
4. Restore the virtual server from the snapshot. OS is started simultaneously with restoring the virtual server.

   Set the virtual server to be restored as follows.

Item	Value
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 (flavor)	Virtual server type checked with step 1(flavor) *Specify ID corresponding to the virtual server type(flavor) for API. For details on how to acquire ID, see "FUJITSU Cloud Service K5 IaaS API User Guide."
Connection port	Port (eth0) created in "2.1.4.1 Creating the Port for the Public LAN (Used also for the Administrative LAN)" Port (eth1) created in "2.1.4.2 Creating the Port for the Cluster Interconnect"
Security group	Not specified (Specified in the port)
Automatic failover	Disabled
Server group ID	Server Group ID created in "2.1.3 Creating the Server Group"
Minimum number of servers	1

Item	Value
Maximum number of servers	1
Snapshot ID	ID of Snapshot
Expanded volume ID	Expanded volume ID checked with step 1
Size of the expanded volume	Size of the expanded volume checked with step 1
Device path of the expanded volume	Device path of the expanded volume checked with step 1

Execution API (Example)

```
# VM_NAME=<virtual server name>
# FLAVOR_REF=<ID of the virtual server type(flavor)>
# SNAPSHOT_ID=<ID of the snapshot>
# VOL_SIZE=<size of the system volume>
# IS_DELETE=<0:Do not delete the block storages in deleting the virtual server.1: Delete the
block storages in deleting the virtual server>
# ADDITIONAL_VOL_ID=<expanded volume ID>
# KEYNAME=<Keypair name>
# PORT_ID1=<port ID of eth0>
# PORT_ID2=<port ID of eth1>
# SERVER_GROUP_ID=<server group ID>
# curl --tlsv1.2 -i $COMPUTE/v2/$PROJECT_ID/servers -X POST -H "X-Auth-Token: $OS_AUTH_TOKEN"
-H "Content-Type: application/json" -d '{"server": {"name": "'$VM_NAME'", "imageRef":
"","flavorRef": "'$FLAVOR_REF'", "block_device_mapping_v2":[ {"boot_index":
"0","uuid":"'$SNAPSHOT_ID'", "volume_size": "'$VOL_SIZE'", "device_name": "/dev/
vda","source_type": "snapshot", "destination_type": "volume",
"delete_on_termination":"'$IS_DELETE'"}, {"boot_index": "1", "uuid":"'$ADDITIONAL_VOL_ID'",
"volume_size": "","device_name": "/dev/vdb", "source_type": "volume", "destination_type":
"volume","delete_on_termination": "0"} ] , "key_name": "'$KEYNAME'", "max_count": "1",
"min_count":"1", "networks": [{"port": "'$PORT_ID1'"}, {"port": "'$PORT_ID2'"}] },
"os:scheduler_hints":{"group": "'$SERVER_GROUP_ID'"} }'
```

Index