

FUJITSU Software Interstage Big Data Parallel Processing Server V1.0.1

User's Guide

Linux(64)

J2UL-1563-02ENZ0(00)
October 2013

Preface

Purpose of this document

This document gives an overview of the features of Interstage Big Data Parallel Processing Server (hereafter, referred to as "this product"). It also describes the operations required during installation and the settings and operations of this product.

Intended readers

This document is intended for administrators building Big Data analysis systems using this product and who have the knowledge of building infrastructure, along with the knowledge of building and operating Apache Hadoop systems, and of developing Apache Hadoop applications.

Structure of this document

This document is structured as follows:

Part 1 - Product Overview	
Chapter 1 Overview	Provides an overview of this product
Chapter 2 Functions	Describes the features provided by this product
Part 2 - Installation	
Chapter 3 System Configuration and Design	Explains the server configuration, file system configuration, network configuration, and user account design that should be considered when using this product
Chapter 4 System Requirements	Explains the hardware and software requirements for using this product
Chapter 5 Preparing to Build the System	Describes the preparatory tasks that should be performed prior to installing this product
Chapter 6 Installation	Explains how to install and set up this product
Chapter 7 Uninstallation	Explains how to uninstall this product
Part 3 - Operations	
Chapter 8 Starting and Stopping	Explains how to start and stop Hadoop on this product
Chapter 9 Developing and Registering Applications	Describes how to develop applications executed by Hadoop
Chapter 10 Executing and Stopping Jobs	Explains how to use this product to execute and stop Hadoop jobs
Chapter 11 Managing Job Execution Users	Explains the management of job execution users who perform Hadoop job operations
Chapter 12 Adding and Deleting Slave Servers	Explains how to add and delete slave servers after installing this product (after starting operations)
Chapter 13 Adding and Deleting Storage Systems	Explains how to add and delete storage systems after installing this product (after starting operations)
Chapter 14 Backup and Restore	Explains how to back up and restore the system configuration of this product
Chapter 15 Operations when There are Errors	Describes the corrective action to take when an error occurs on a system that uses this product
Chapter 16 Troubleshooting	Describes the troubleshooting data to collect when an issue occurs with a system that uses this product
Appendixes	
Appendix A Commands	Explains the commands provided by this product

Appendix B Definition Files	Describes the definition files for the system configuration information used by this product
Appendix C Hadoop Configuration Parameters	Explains the various parameters to configure for using Hadoop on this product
Appendix D Port List	Describes the ports used by this product
Appendix E Messages	Explains the meaning of messages output by this product and the corresponding action that should be taken
Appendix F Mandatory Packages	Covers the packages required for the system software that runs this product
Glossary	Explains the terminology used for this product.

Conventions

The following notation is used in this document:

- Where features differ in accordance with the system software required to use this product, information is distinguished as shown below.

[Master server]	Information intended for master servers
[Slave server]	Information intended for slave servers
[Development server]	Information intended for development servers
[Collaboration server]	Information intended for collaboration servers

- Unless indicated otherwise, "rack server" in this document refers to the PRIMERGY RX Series.
- References are enclosed in " ".
- Variable information or content that can be modified is italicized and written in mixed case (for example: *newBkpDir*).
- GUI elements, such as window, menu, and tab names, are formatted bold (for example: **File**).
- Key names are enclosed in < >.
- Strings and numeric values requiring special emphasis are enclosed in double quotation marks ("").
- In usage examples, the prompt is represented by the Linux "#".

Interstage Big Data Parallel Processing Server website

The latest manuals and technical information is published on the Interstage Big Data Parallel Processing Server website.

It is recommended to refer to that website before using this product. The URL is shown below.

URL: http://www.fujitsu.com/global/services/software/interstage/solutions/big-data/bdpp/ (as of October2013)

Related documents

The following manuals are bundled with this product:

- PRIMECLUSTER 4.3A10
- ServerView Resource Orchestrator Virtual Edition V3.1.0
- Primesoft Distributed File System V1

To refer to the contents of the manuals bundled with this product, refer to the manuals stored at the following locations in the product media:

DISK1: PRIMECLUSTER manuals

`dvdDrive:\DISK1\products\PCL\documents\manuals\En`

DISK1: ServerView Resource Orchestrator Virtual Edition manual

`dvdDrive:\DISK1\products\ROR\DISK1\Manual\en\VirtualEdition`

DISK1: Primesoft Distributed File System for Hadoop manual

`dvdDrive:\DISK1\products\PDFS\documents\manuals\en`

In the bundled manuals, only the features provided by Interstage Big Data Parallel Processing Server can be used.

Abbreviation

The following abbreviations are used in this document:

Abbreviation	Product
Linux or Red Hat Enterprise Linux	Red Hat(R) Enterprise Linux(R) 5.6 (for Intel64)
	Red Hat(R) Enterprise Linux(R) 5.7 (for Intel64)
	Red Hat(R) Enterprise Linux(R) 5.8 (for Intel64)
	Red Hat(R) Enterprise Linux(R) 5.9 (for Intel64)
	Red Hat(R) Enterprise Linux(R) 6 (for Intel64)
	Red Hat(R) Enterprise Linux(R) 6.1 (for Intel64)
	Red Hat(R) Enterprise Linux(R) 6.2 (for Intel64)
	Red Hat(R) Enterprise Linux(R) 6.3 (for Intel64)
	Red Hat(R) Enterprise Linux(R) 6.4 (for Intel64)

Export restriction

If this document is to be exported or provided overseas, confirm legal requirements for the Foreign Exchange and Foreign Trade Act as well as other laws and regulations, including U.S. Export Administration Regulations, and follow the required procedures.

Trademarks

- Apache Hadoop, Hadoop, HDFS, HBase, Hive, and Pig are trademarks of The Apache Software Foundation in the United States and/or other countries.
- Adobe, Adobe Reader, and Flash are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries.
- Linux is a registered trademark of Linus Torvalds.
- Red Hat, RPM, and all Red Hat-based trademarks and logos are trademarks or registered trademarks of Red Hat, Inc. in the United States and other countries.
- Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.
- Microsoft, Windows, MS, MS-DOS, Windows XP, Windows Server, Windows Vista, Windows 7, Excel, and Internet Explorer are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.
- VMware, the VMware logo, VMware vSphere, VMware vCenter, ESXi, vMotion, Storage DRS, Server Appliance, and DirectPath I/O are trademarks or registered trademarks of VMware, Inc. in the United States and other countries.
- Interstage, ServerView, Symfoware, and Systemwalker are registered trademarks of Fujitsu Limited.
- Other company names and product names used in this document are trademarks or registered trademarks of their respective owners.

Note that registration symbols (TM or R) are not appended to system names or product names in this manual.

Issue date and version

Edition	Manual code
October 2013: Second edition	J2UL-1563-02ENZ0(00)

Notice

No part of the content of this manual may be reproduced without the written permission of Fujitsu Limited.

The contents of this manual may be changed without notice.

Copyright 2013 FUJITSU LIMITED

Contents

Part 1 Product Overview.....	1
Chapter 1 Overview.....	2
1.1 Product Summary.....	2
1.2 Features.....	3
1.2.1 High Performance.....	4
1.2.2 High Reliability.....	4
1.2.3 High Operability.....	5
Chapter 2 Functions.....	7
2.1 Fujitsu Distributed File System (DFS).....	7
2.2 Master Server Replication.....	7
2.3 Smart Setup.....	8
Part 2 Installation.....	9
Chapter 3 System Configuration and Design.....	10
3.1 Design Overview.....	10
3.2 Server Design.....	10
3.2.1 Server Configuration.....	11
3.2.2 Design the Server Configuration.....	12
3.2.3 Designing System Resources.....	13
3.2.3.1 Static Disk Size.....	13
3.2.3.2 Dynamic Disk Size.....	15
3.2.3.3 Memory Size.....	16
3.2.3.4 System Parameters.....	16
3.3 File System Design.....	19
3.3.1 File System Configuration.....	19
3.3.2 File System Configuration Design.....	19
3.3.2.1 Relationship between File System Size, Data Block Size and Maximum File Size.....	21
3.4 Network Design.....	22
3.4.1 Network Configuration.....	22
3.4.2 Network Configuration Design.....	24
3.5 Designing User Accounts.....	25
Chapter 4 System Requirements.....	27
4.1 Hardware Requirements.....	27
4.1.1 Hardware Conditions.....	27
4.1.2 Expansion Card Options.....	27
4.2 Software Configuration.....	27
4.3 Software Requirements.....	28
4.3.1 System Software.....	28
4.3.2 Mandatory Software.....	29
4.3.3 Exclusive Software.....	29
4.3.4 Related Software.....	30
Chapter 5 Preparing to Build the System.....	31
5.1 Server BIOS Settings.....	31
5.2 File System When Installing the OS.....	32
5.3 Disabling SELinux Functionality.....	32
5.4 System Parameter Settings.....	33
5.5 Firewall Settings.....	34
5.6 Host Name Settings.....	34
5.7 Server Name Settings.....	36
5.8 Public LAN Network Interface Settings.....	36
5.9 System Time Settings.....	36

5.10 Storage System Environment Preparations.....	37
5.11 ssh Settings.....	37
5.12 SNMP Trap Daemon Settings.....	38
5.13 ServerView Agent Settings.....	38
5.14 kdump Shutdown Agent Settings.....	39
5.15 KVM Shutdown Agent Settings.....	39
5.16 Expanding the Microsoft LAN Manager Module.....	41
Chapter 6 Installation.....	43
6.1 Installing to Master Servers.....	44
6.1.1 Installing the Master Server Functionality.....	45
6.1.1.1 Creating bdpp.conf.....	45
6.1.1.2 Installing the Master Server Functionality.....	46
6.1.1.3 Applying Software Updates.....	47
6.1.2 HA Cluster Setup.....	47
6.1.2.1 HA Cluster Setup 1.....	48
6.1.2.2 HA Cluster Setup 2.....	48
6.1.2.3 Restarting the System.....	48
6.1.3 DFS Setup.....	48
6.1.3.1 Checking Shared Disk Settings.....	51
6.1.3.2 Checking Cluster Status.....	52
6.1.3.3 Management Partition Initialization.....	52
6.1.3.4 Registering DFS Management Server Information to the Management Partition.....	52
6.1.3.5 Starting the pdfsfrmd daemon.....	53
6.1.3.6 Creating the File System.....	53
6.1.3.7 Setting the User ID for Executing MapReduce.....	54
6.1.3.8 Registering the DFS Client Information.....	54
6.1.3.9 Creating the Mount Point and Configure fstab Settings.....	56
6.1.3.10 Mounting.....	57
6.1.3.11 Generating the DFS File System Configuration Information.....	58
6.1.4 Network Replication and Hadoop Setup.....	59
6.1.4.1 Creating the Slave Server Definition File 'slaves'.....	59
6.1.4.2 Network Replication and Hadoop Setup.....	60
6.2 Installing to a Slave Server.....	60
6.2.1 Installing the Slave Server Functionality.....	60
6.2.1.1 Creating bdpp.conf.....	60
6.2.1.2 Installing the Slave Server Functionality.....	61
6.2.1.3 Applying Software Updates.....	61
6.2.2 DFS Setup.....	62
6.2.2.1 Creating the Mount Point and Configuring fstab Settings.....	62
6.2.2.2 Mounting.....	63
6.2.3 Network Replication and Hadoop Setup.....	64
6.2.4 Checking the Connection to the Slave Server.....	64
6.3 Adding Second and Subsequent Slave Servers.....	65
6.3.1 Adding a Slave Server to a Physical Environment.....	66
6.3.1.1 Installing the Slave Server.....	68
6.3.1.2 Registering the Slave Server.....	68
6.3.1.3 Registering the Network Parameter and iSCSI Name Automatic Configuration Feature.....	69
6.3.1.4 Creating a Clone Image.....	70
6.3.1.5 Restarting the System.....	71
6.3.1.6 Cloning.....	71
6.3.1.7 Restarting the System.....	71
6.3.2 Adding a Slave Server to a Virtual Environment.....	71
6.3.2.1 Installing the Slave Server.....	72
6.3.2.2 Registering the Network Parameter and iSCSI Name Automatic Configuration Feature.....	72
6.3.2.3 Cloning.....	74
6.3.2.4 Configuring the Server Name.....	74

6.3.2.5 Restarting the Virtual Machine.....	74
6.4 Installing to a Development Server.....	74
6.4.1 Installing the Development Server Functionality.....	75
6.4.1.1 Creating bdpp.conf.....	75
6.4.1.2 Installing the Development Server Functionality.....	75
6.4.1.3 Applying Software Updates.....	76
6.4.2 DFS Setup.....	76
6.4.2.1 Mount Point Creation and fstab Settings.....	76
6.4.2.2 Mounting.....	77
6.4.3 Hadoop Setup.....	78
6.5 Installing to a Collaboration Server.....	78
6.5.1 Installing the Collaboration Server Functionality.....	78
6.5.1.1 Creating bdpp.conf.....	78
6.5.1.2 Installing the Collaboration Server Functionality.....	79
6.5.1.3 Applying Software Updates.....	79
6.5.2 DFS Setup.....	80
6.5.2.1 Mount Point Creation and fstab Settings.....	80
6.5.2.2 Registering the hadoop Group and mapred User.....	81
6.5.2.3 Mounting.....	81
Chapter 7 Uninstallation.....	83
7.1 Uninstalling from a Master Server.....	83
7.1.1 Removing the Network Replication Setup.....	84
7.1.2 Removing the DFS Setup.....	84
7.1.3 Removing the HA Cluster Setup.....	86
7.1.4 Uninstalling the Master Server Functionality.....	86
7.1.5 Post-Uninstallation Tasks.....	87
7.1.5.1 Directories and Files Remaining after an Uninstall.....	87
7.1.5.2 Uninstalling "Uninstall (middleware)".....	88
7.1.6 Restarting the System.....	89
7.2 Uninstalling from a Slave Server.....	89
7.2.1 Removing the Network Replication Setup.....	89
7.2.2 Removing the DFS Setup.....	89
7.2.3 Uninstalling the Slave Server Functionality.....	90
7.2.4 Post-Uninstallation Tasks.....	91
7.2.4.1 Directories and Files Remaining after an Uninstall.....	91
7.2.5 Restarting the System.....	92
7.3 Uninstalling from a Development Server.....	92
7.3.1 Removing the DFS Setup.....	92
7.3.2 Uninstalling the Development Server Functionality.....	92
7.3.3 Post-Uninstallation Tasks.....	94
7.3.3.1 Directories and Files Remaining after an Uninstall.....	94
7.3.4 Restarting the System.....	94
7.4 Uninstalling from a Collaboration Server.....	94
7.4.1 Removing the DFS Setup.....	94
7.4.2 Uninstalling the Collaboration Server Functionality.....	95
7.4.3 Post-Uninstallation Tasks.....	96
7.4.3.1 Directories and Files Remaining after an Uninstall.....	96
7.4.4 Restarting the System.....	96
Part 3 Operations.....	97
Chapter 8 Starting and Stopping.....	98
8.1 Starting.....	98
8.2 Stopping.....	99
8.3 Displaying Status.....	100
Chapter 9 Developing and Registering Applications.....	101

9.1 Application Development Environment.....	101
9.2 Developing Applications.....	101
9.2.1 Application Overview.....	101
9.2.2 Designing Applications.....	102
9.2.3 Creating Applications.....	102
9.2.4 References for Developing MapReduce Applications.....	102
9.3 Registering Applications.....	103
Chapter 10 Executing and Stopping Jobs.....	104
10.1 Executing Jobs.....	104
10.2 Stopping Jobs.....	104
10.3 Displaying Job Status.....	104
Chapter 11 Managing Job Execution Users.....	105
11.1 Adding Job Execution Users.....	105
11.1.1 Create a User Account.....	105
11.1.2 Create a Home Directory for the User on the DFS.....	106
11.1.3 Configure SSH Authentication Keys to be Used by the Cache Local Feature.....	106
11.1.4 Configure MapReduce Job User Authentication Keys.....	107
11.2 Deleting Job Execution Users.....	107
11.2.1 Deleting the Home Directory of the User Created on the DFS.....	107
11.2.2 Deleting the User Account.....	108
Chapter 12 Adding and Deleting Slave Servers.....	109
12.1 Adding Slave Servers.....	109
12.1.1 Configuring Host Name Settings.....	110
12.1.2 Registering the DFS Client Information.....	111
12.1.3 Adding by Cloning.....	112
12.1.4 Stopping Hadoop.....	113
12.1.5 Remounting.....	113
12.1.6 Editing and Reflecting the Slave Server Definition File.....	113
12.1.7 Changing Hadoop Configuration Parameters.....	114
12.1.8 Starting Hadoop.....	114
12.2 Deleting Slave Servers.....	115
12.2.1 Stopping Hadoop.....	116
12.2.2 Editing and Reflecting the Slave Server Definition File.....	116
12.2.3 Changing Hadoop Configuration Parameters.....	117
12.2.4 Starting Hadoop.....	117
12.2.5 Unmounting and Removing the fstab Configuration.....	117
12.2.6 Deleting the DFS Client Information.....	118
Chapter 13 Adding and Deleting Storage Systems.....	119
13.1 Adding a Storage System.....	119
13.1.1 Stopping Hadoop.....	120
13.1.2 Unmounting.....	121
13.1.3 Adding a Partition.....	121
13.1.4 Recreating and Distributing the DFS File System Configuration Information.....	121
13.1.5 Mounting.....	122
13.1.6 Starting Hadoop.....	122
13.2 Deleting a Storage System.....	122
13.2.1 Stopping Hadoop.....	124
13.2.2 Unmounting.....	124
13.2.3 Deleting a File System.....	124
13.2.4 Creating a File System.....	125
13.2.5 Setting the User ID for Executing MapReduce.....	125
13.2.6 Recreating and Distributing the DFS File System Configuration Information.....	125
13.2.7 Mounting.....	126
13.2.8 Starting Hadoop.....	126

Chapter 14 Backup and Restore.....	127
14.1 Backup.....	127
14.1.1 Resources Saved when the Backup Command is Executed.....	127
14.1.2 Backup Method.....	129
14.1.2.1 Backing Up a Master Server, Development Server, or Collaboration Server.....	129
14.1.2.2 Backing Up a Slave Server.....	130
14.2 Restore.....	131
14.2.1 Restore Method.....	131
14.2.1.1 Restoring a Master Server, Development Server, or Collaboration Server.....	131
14.2.1.2 Restoring a Slave Server.....	133
Chapter 15 Operations when There are Errors.....	135
15.1 Operations when Errors Occur on a Master Server.....	136
15.1.1 If the Master Servers Use Replicated Configuration.....	136
15.1.2 If the Master Server Does Not Use Replicated Configuration.....	141
15.2 Operations when Errors Occur on a Slave Server.....	143
15.3 Operations when Errors Occur on a Development Server.....	145
15.4 Operations when Errors Occur on a Collaboration Server.....	145
15.5 Operations when Errors Occur on the File System.....	146
15.6 How to Check Errors.....	147
Chapter 16 Troubleshooting.....	150
16.1 When a Problem Occurs with an HA Cluster.....	151
16.1.1 Collecting Troubleshooting Data.....	151
16.2 When a Problem Occurs during Cloning.....	154
16.2.1 Types of Troubleshooting Data.....	154
16.2.2 Collecting Initial Troubleshooting Data.....	155
16.2.3 Collecting Detailed Troubleshooting Data.....	156
16.3 If DFS Setup or Shared Disk Problems Occurred.....	157
16.3.1 Collecting DFS Troubleshooting Data.....	157
16.4 When a Problem Occurs in Hadoop.....	159
Appendix A Commands.....	160
A.1 bdpp_addserver.....	161
A.2 bdpp_backup.....	162
A.3 bdpp_changeimagedir.....	163
A.4 bdpp_changeslaves.....	164
A.5 bdpp_deployimage.....	165
A.6 bdpp_getimage.....	166
A.7 bdpp_lanctl.....	167
A.8 bdpp_listimage.....	168
A.9 bdpp_listserver.....	169
A.10 bdpp_prepareserver.....	171
A.11 bdpp_removeimage.....	172
A.12 bdpp_removeserver.....	173
A.13 bdpp_restore.....	173
A.14 bdpp_start.....	174
A.15 bdpp_stat.....	176
A.16 bdpp_stop.....	177
Appendix B Definition Files.....	179
B.1 bdpp.conf.....	179
B.2 slaves.....	195
B.3 clone.conf.....	195
B.4 FJSVrcx.conf.....	198
B.5 ipaddr.conf.....	199
B.6 initiator.conf.....	202

Appendix C Hadoop Configuration Parameters.....	203
C.1 hadoop-env.sh.....	203
C.2 core-site.xml.....	205
C.3 mapred-site.xml.....	205
C.4 pdfs-site.xml.....	207
C.5 sysctl.conf.....	209
C.6 limits.conf.....	209
C.7 HDFS Settings (Reference Information).....	209
Appendix D Port List.....	212
Appendix E Messages.....	216
E.1 Messages During Installation.....	216
E.2 Messages During Setup.....	218
E.3 Messages Output During Operation.....	220
E.3.1 Messages Output During Command Execution.....	220
E.3.2 Other Messages.....	225
Appendix F Mandatory Packages.....	227
Glossary.....	230

Part 1 Product Overview

This part describes the features and functionality of this product.

Chapter 1 Overview	2
Chapter 2 Functions	7

Chapter 1 Overview

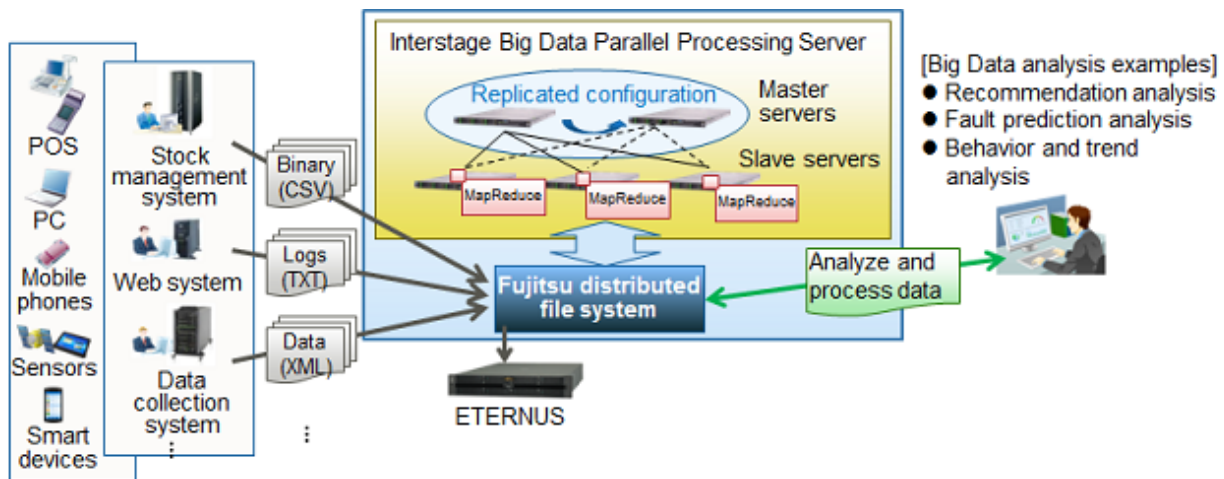
This chapter presents an overview of this product and a description of its features.

1.1 Product Summary

The Interstage Big Data Parallel Processing Server uses the industry standard for Big Data processing, "Apache Hadoop", with Fujitsu proprietary technology incorporated.

In recent years, not only are massive amounts of data collected from sensors and smart devices, such as smart phones and tablets, but also the formats and structures are many and varied, and these are continuously increasing. This is known as **Big Data**, and it is increasingly being adopted, particularly by leading corporations, and is gathering a lot of attention, because it is providing unprecedented business advantages.

This product vastly improves reliability and processing performance, decreases system installation time, reduces the burden in operation management, and supports Big Data in enterprise systems.



Features of Big Data

Big Data has the following features:

1. Massive size of data
Enormous amounts of data, with data sizes reaching the terabyte to petabyte range
2. Variety of data
Data in a variety of formats: structured data (database data), non-structured data (sensor information, text data such as access log information), semi-structured data (data having the qualities of both structured data and non-structured data)
3. Data frequently generated
Continuous generation of new data from sensors and similar
4. Need to use data in real-time
Performing analysis in a short amount of time and using the data in real-time

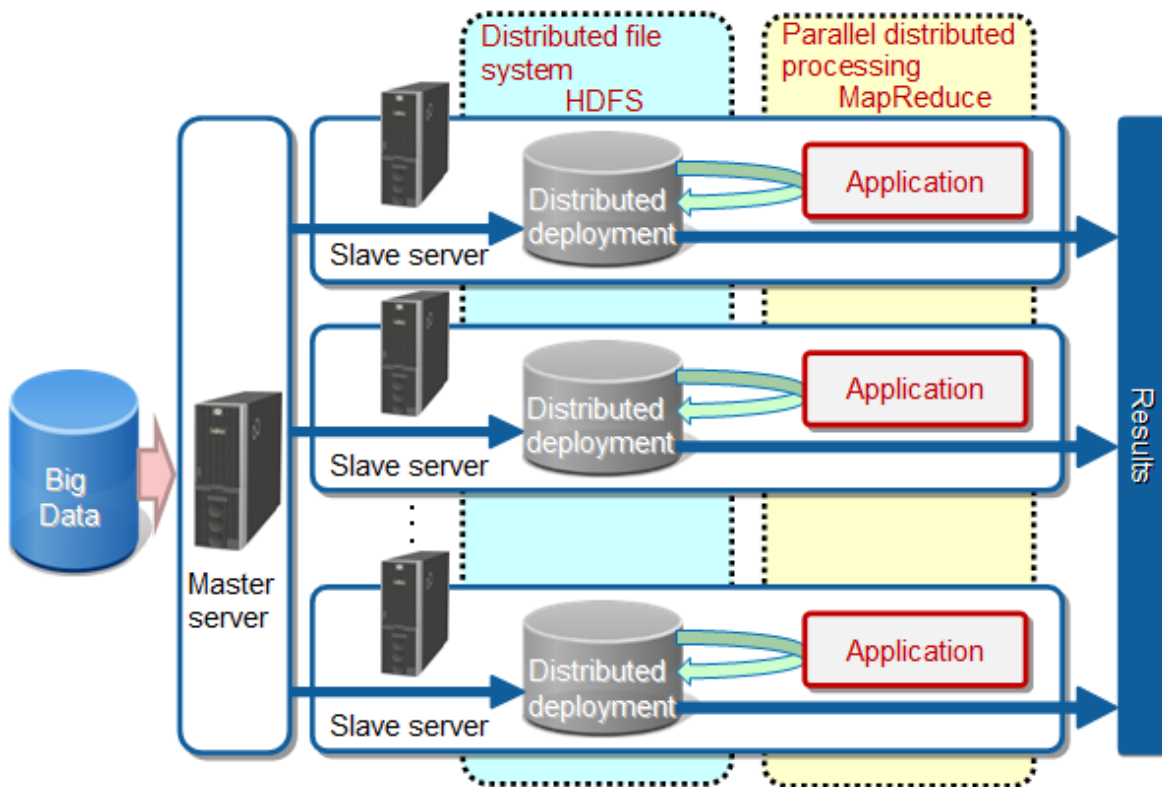
Apache Hadoop (*1) is widely used and is the world standard for applications that can resolve the above Items 1 and 2 in Big Data processing (processing data of massive size, and variety of data).

*1 Apache Hadoop: Open source software, developed by Apache Software Foundation (ASF), that efficiently performs distribution and parallel processing of Big Data

Apache Hadoop

Apache Hadoop technology splits Big Data, distributes it to tens or tens of thousands of servers, and performs parallel processing, thereby performing batch processing of Big Data in a small amount of time. This technology has the following features:

- Low cost
Economical systems can be built by using large numbers of comparatively cheap servers that perform parallel processing.
- High availability
Processing can continue even if two machines stop simultaneously because the split data is distributed to three or more of the servers (slave servers) that execute parallel processing.
- Scalability
Systems can be scaled-out easily by adding slave servers.
- Variety of data processes
Parallel analysis processing applications (MapReduce applications) can be developed for a range of uses, from simple analysis such as character string searches through to high-level analysis logic for image analysis or similar, and can process data in a variety of formats.



1.2 Features

This section describes the features of this product.

- High performance
- High reliability
- High operability

1.2.1 High Performance

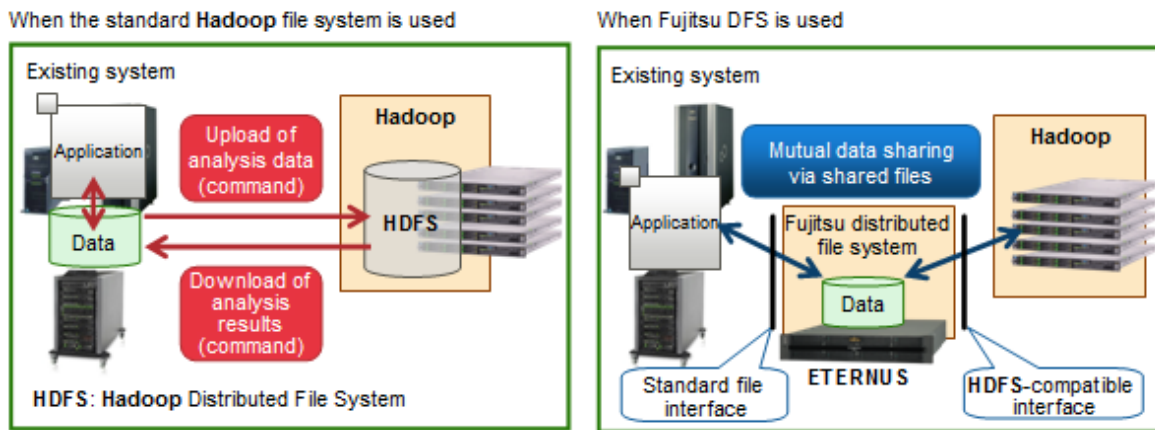
Reducing time for data transfer to Hadoop processing server

The data stored in the storage system can be accessed directly, and processed, using the Fujitsu distributed file system in addition to the Apache Hadoop distributed file system.

Under Hadoop, business application data is temporarily transferred to "HDFS" before it is processed. In contrast, when the Fujitsu distributed file system is used, data transfer is not necessary, greatly reducing the processing time.

Use of existing tools without modification

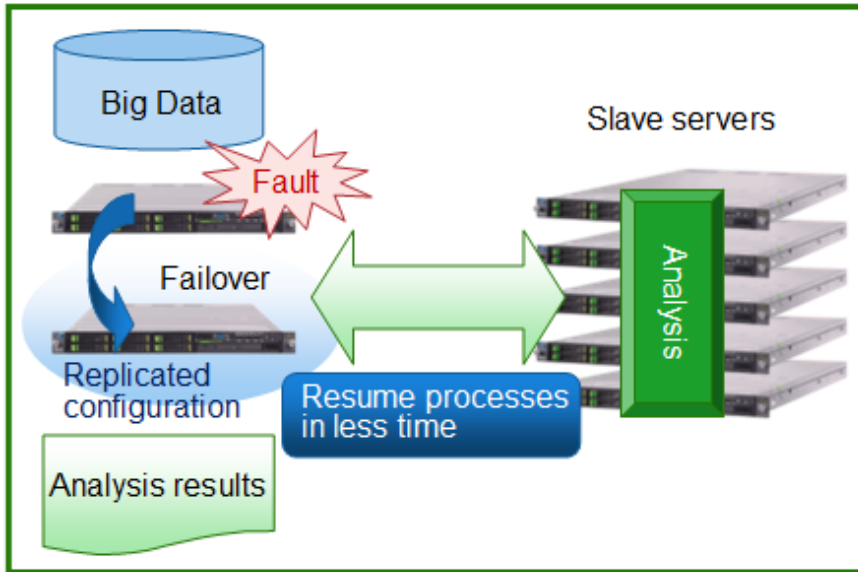
Existing tools, such as backup and print tools, can be used without modification because the interface with the storage system used to store data is the Linux standard interface.



1.2.2 High Reliability

Resolving the single point of failure issue

When a fault occurs in the master server that manages the entire system under "Apache Hadoop", "HDFS" cannot be used while the cause of the fault is being removed and the master server is being restored. This causes the stoppage time to extend over a long period (single point of failure). With this product, Fujitsu HA cluster technology provides duplicated master server operation, thus avoiding a single point of failure and achieving high reliability with restarts completed in a small timeframe.



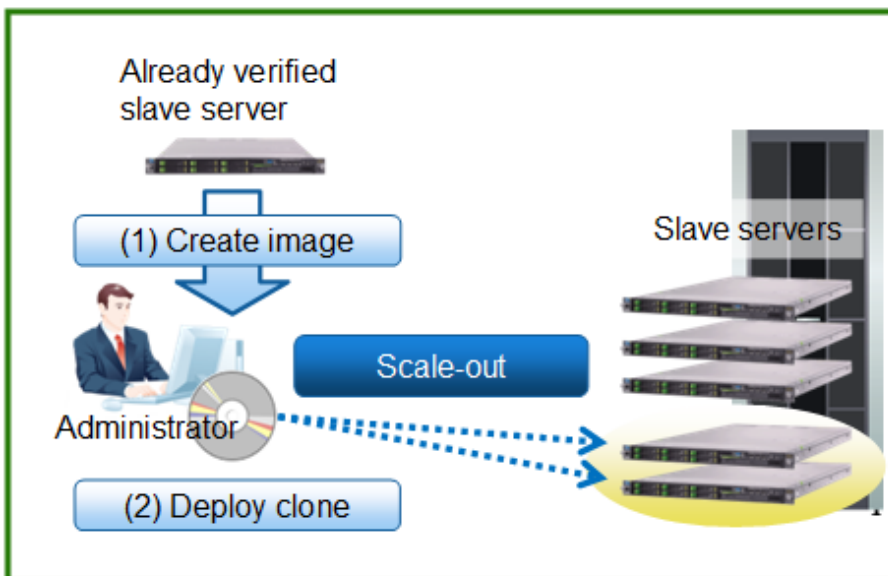
1.2.3 High Operability

Quick and simple setup

This product has in-built "Smart Setup" based on Fujitsu smart software technology (*1).

System installation time is short because a clone of an already created slave server can be deployed and set automatically at multiple servers as a batch.

Scale-out is amazingly easy because images can be deployed automatically when servers are added.

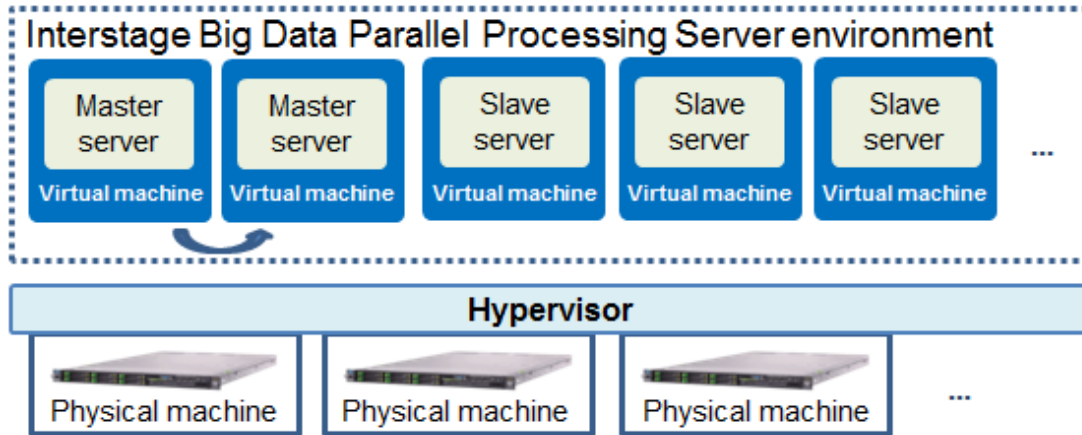


*1: Smart software technology: Fujitsu proprietary technology itself judges hardware and software conditions and is designed to perform optimization in order to improve ease of use and give peace of mind.

Cost reduction by virtualizing servers

Apache Hadoop allows you to improve the performance of distributed processing by increasing the number of slave servers, but this entails increased costs.

It is possible to use hypervisor-type virtualization software to build the system with this product. By efficiently using resources such as CPUs and memory on the minimum number of physical machines, installation and running costs can be reduced.



Chapter 2 Functions

This chapter describes the functions provided by this product.

- [Fujitsu Distributed File System \(DFS\)](#)
- [Master Server Replication](#)
- [Smart Setup](#)

2.1 Fujitsu Distributed File System (DFS)

The Fujitsu distributed file system (DFS) provides the following features:

Use of external storage devices

It can use ETERNUS, Fujitsu's external storage device that boasts high reliability and operability.

Using RAID, it is not necessary to produce data replicas with software such as HDFS, meaning much faster access.

Furthermore, you can add just the resources currently required, since it is possible to size slave servers and disks individually.

POSIX compliant interface

This product supports the standard Linux file access interface (POSIX).

Integration is easy and efficient, since data input and output are possible without having to change the application.

Speed acceleration due to memory caching

Memory caches in the slave servers can be allocated efficiently, leading to improved processing speeds.

The management data information (metadata) acquired on the various servers is maintained in caches, and this means that there is no need to constantly access the network to acquire this information.

2.2 Master Server Replication

By using replicated configuration with two master servers (a primary and a secondary), the secondary master server can take over processing even if there is a failure in the primary master server. By registering JobTracker (MapReduce) that makes up Hadoop as a cluster application, the secondary master server can take over processing if an error occurs in Hadoop.

Furthermore, you can also have redundancy in the transmission route between the master server and the slave server. Replicating by using two transmission networks means that the route where a failure occurs can be isolated so that network communications can continue.

Point

.....

A Hadoop network can also be made up of one master server or one transmission network.

However, it is recommended to use a replicated master server configuration so that operations can continue if there is a failure in the master server, which is the single point of failure for Hadoop, as explained in "[1.2.2 High Reliability](#)".

.....

2.3 Smart Setup

Cloning automates the process of building slave servers after the first one has been built.

The cloning images gathered from a slave server can be distributed to multiple slave servers so that scaling up can be achieved over a short period. Furthermore, network settings such as host name and slave server IP address are also configured automatically. This means that physical servers can be cloned, as well as slave servers built on virtual environments.

There are commands provided with this product for calling the cloning functionality, making it easy to perform all cloning operations, from gathering the cloning image, to distribution and configuration.

Part 2 Installation

This part describes how to build a system for this product to be installed.

Chapter 3 System Configuration and Design.....	10
Chapter 4 System Requirements.....	27
Chapter 5 Preparing to Build the System.....	31
Chapter 6 Installation.....	43
Chapter 7 Uninstallation.....	83

Chapter 3 System Configuration and Design

This chapter explains the configuration and design of the system for this product.

3.1 Design Overview

This section describes the items that need to be set before installing this product.

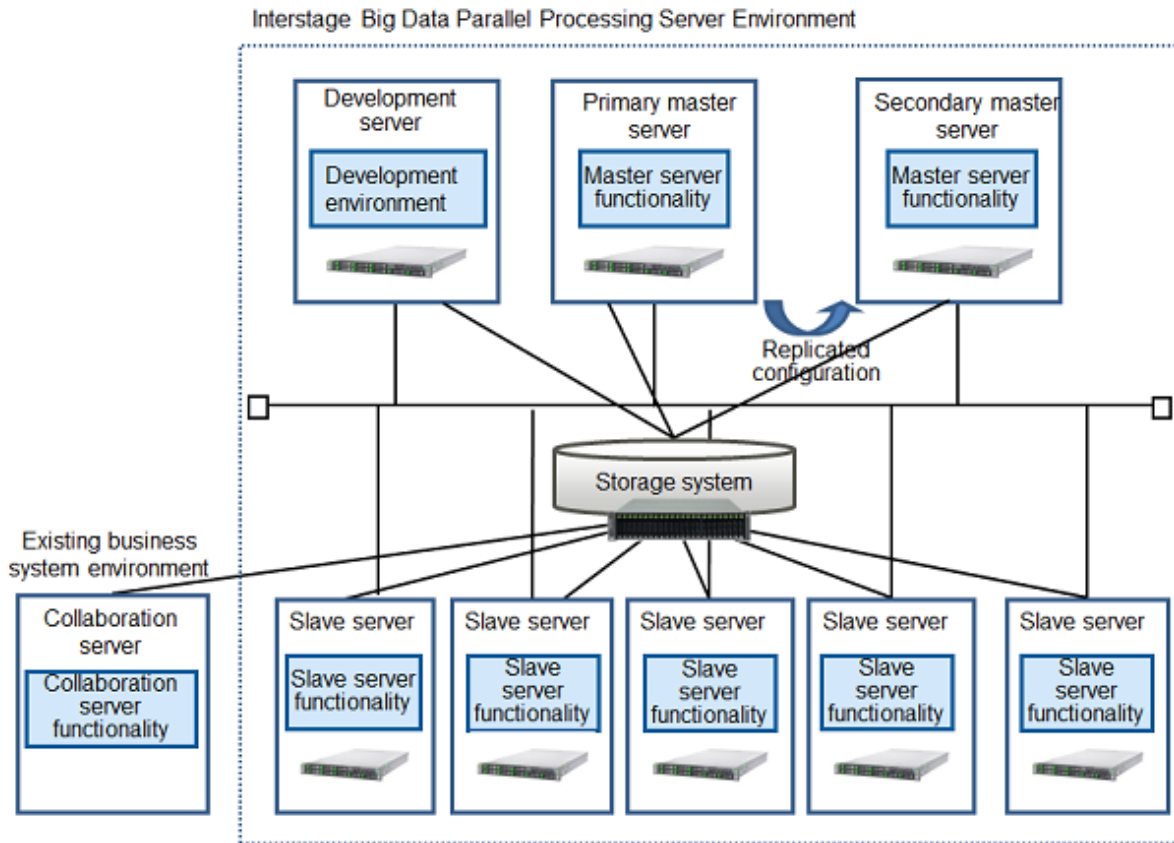
Category	Design item	Overview	Remarks
Server design	Environment to install the server	Decide whether to install on a physical environment or a virtual environment.	-
	Master server configuration	Decide whether to have replicated configuration for the master servers.	Specify in bdpp.conf used during installation.
	Slave server configuration	Decide the number of slave servers to install initially.	You can build as many slave servers as you need after the first one during installation.
	Number of machines in the virtual environment	Decide the number of machines to use as the virtual environment. (When installing on a virtual environment)	-
	Designing system resources	Confirm the disk capacity, memory capacity, and tuning parameters required to install and operate this product.	Ensure that the servers meet the requirements before installing.
File system design	Capacity to be allocated to shared disk devices	Estimate the total size of the file data and determine the disk capacity to put in the storage device.	Specify this when constructing the DFS file system during installation.
	Partitions used	Determine the partitions to be used in the DFS.	You need to complete storage system installation and configuration before installation.
	Size of the file data area	Determine the maximum size, including partitions, that may be added in the future (optional).	
	Data block size	Determine the data block size required for optimum input/output processing (optional).	
Network design	Replication of the public LAN	Determine whether to have LAN redundancy so that Hadoop parallel distribution processing is possible.	Specify in bdpp.conf used during installation.
	Replication of CIP	Determine whether to have LAN redundancy so that alive monitoring of the master server's status is possible.	The servers and storage systems must be physically connected before installation.
	Replication of iSCSI	Determine whether to have LAN redundancy for the connection of the storage systems.	
Designing user accounts	Deciding users of Hadoop	Decide the user for executing Hadoop jobs.	Specify in bdpp.conf used during installation.

3.2 Server Design

This section explains how to design the servers for this product.

3.2.1 Server Configuration

This section describes the server configurations and server types using this product.



Master server

A master server splits large data files into blocks and makes files (distributed file system), and centrally manages those file names and storage locations.

A master server can also receive requests to execute analysis processing application jobs, and cause parallel distributed processing on slave servers.

Replication of the master server is possible with this product.

Install the master server feature of this product on the master server.

Slave server

Analysis processing can be performed in a short amount of time because the data file, split into blocks by the master server, is processed using parallel distributed processing on multiple slave servers.

Furthermore, the data that is split into blocks is stored in a high-reliability system.

This product's slave server functionality is installed at each slave server.

Development server

The development server is a server where Pig or Hive is installed and executed. They enable easy development of applications that perform parallel distribution (MapReduce).

This product's development server functionality is installed at the development server.

Collaboration server

With Apache Hadoop, it was necessary to register in HDFS (the distributed file system for Hadoop) in order to analyze. Analysis can be performed by directly transferring the large amount of data on the business system to the DFS (Distributed File System), which is built on the high reliability storage system that is one of the main features of this product, from the collaboration server using the Linux standard file interface.

Installation of an existing data backup system on the collaboration server enables easy use of data backups.

This product's collaboration server functionality is installed at the collaboration server.



Make sure that the data stored in the DFS using data transfer is the data that is to be analyzed using Hadoop. Other data cannot be stored.

3.2.2 Design the Server Configuration

Make the following design decisions for each server type when designing server configuration.

Environment to install the server

Decide whether to install this product on a physical environment or a virtual environment.

When installing on a virtual environment, install the master server (both primary and secondary), slave servers, and development servers on the virtual environment.

Install the virtualization software and build the virtual environment beforehand.



Apart from the collaboration servers, it is not possible to have a server configuration with a mix of installations on physical environments and virtual environments.



Refer to the manuals of the server virtualization software you are using for information on how to build virtual environments.

Master server configuration

Decide whether to have replicated configuration for the master servers.

This product supports replicated configuration for the master server (1 to 1 active/standby type HA cluster configuration). Replicated configuration requires two master servers (a primary and a secondary one). You can configure a system using only one master server if replicated configuration is not necessary.

The replicated configuration is recommended to resolve the single point of failure issue with Hadoop.



When installing on a virtual environment (VMware)

Do not deploy the virtual machine with the master server installed on it to a VMware HA cluster. Also note that the virtual machine with the master server installed on it cannot use features such as VMware vMotion.

Refer to "Appendix H Using PRIMECLUSTER in a VMware Environment" in the "PRIMECLUSTER Installation and Administration Guide 4.3" for points to note when using VMware.

Slave server configuration

This product allows scaling out slave servers, which improves scalability. It is recommended to make an estimate of how many servers are required by performing prototype tests before using the system in a production environment, as the time required for processing depends on factors such as the number of slave servers, the Hadoop applications, and the volume and characteristics of data to be processed. On top of this, determine the maximum number of slave servers, including any future expansion.



A maximum of 128 master servers, slave servers, development servers, and collaboration servers can be set up.

Number of machines in the virtual environment

Decide the number of physical machines to use as the virtual environment.

When many virtual machines are built on one physical machine, you may not achieve the desired results as distributed processing performance reaches a ceiling, determined by things such as the Hadoop application, the amount of data being processed, and other characteristics. For this reason, perform validation using prototypes before using in actual operations, so that you can properly estimate the number of virtual machines for each physical machine.

Note that the virtual machines (guest OS's) where this product is to be installed should be created beforehand.



Refer to the manuals of the server virtualization software product for information on how to create virtual machines.



When installing on a virtual environment (KVM)

The name of the virtual machine where the master server is installed (domain name of the guest OS) must match the host name of the virtual machine. Apart from the virtual machine where the collaboration server is installed, the names of the virtual machines should be specified to match the following parameters in `bdpp.conf` specified during installation:

- `BDPP_PRIMARY_NAME` (primary master server host name)
 - `BDPP_SECONDARY_NAME` (secondary master server host name)
 - `BDPP_SERVER_NAME` (slave server host name, development server host name)
-

3.2.3 Designing System Resources

This section describes the system resources required to install this product.

3.2.3.1 Static Disk Size

The following static disk capacity (excluding the OS) is required to make a new installation for this product.

[Master server]

Installation environment	Directory	Disk size (in MB)
Physical environment Primary master server	/opt	1090
	/etc	15
	/var	470
	/usr	280
Physical environment Secondary master server	/opt	385
	/etc	2
	/var	80
	/usr	280
Virtual environment	/opt	380
	/etc	2
	/var	40
	/usr	280

[Slave server]

Installation environment	Directory	Disk size (in MB)
Physical environment	/opt	330
	/etc	1
	/var	20
	/usr	270
Virtual environment	/opt	260
	/etc	1
	/var	20
	/usr	270

[Development server]

Installation environment	Directory	Disk size (in MB)
Physical environment	/opt	230
	/etc	1
	/var	20
	/usr	400
Virtual environment	/opt	240
	/etc	1
	/var	20
	/usr	400

[Collaboration server]

Installation environment	Directory	Disk size (in MB)
Physical environment	/opt	35
	/etc	1
	/var	20

Installation environment	Directory	Disk size (in MB)
Virtual environment	/opt	40
	/etc	1
	/var	20

3.2.3.2 Dynamic Disk Size

When using this product, the disk sizes below are required in addition to the static disk size, in the master server and slave server directories.

[Master server]

Installation environment	Directory	Disk size (in MB)
Physical environment	/etc	2
	/var/opt	2520
	Clone image file storage directory Default: /var/opt/FJSVscw-deploysv/depot	Refer to " Clone image file storage area " below.
	Directory for backup and restore Refer to " Chapter 14 Backup and Restore " for details.	150
Virtual environment	/etc	1
	/var/opt	5

[Slave server]

Installation environment	Directory	Disk size (in MB)
Physical environment/ virtual environment	/etc	1
	/var/opt	1

Clone image file storage area

A clone image file storage area is required if cloning is to be performed.

Allocate area on the master server as an area to store the clone image files belonging to the cloned slave server (excluding slave servers built in virtual environments).



Note

- Create the clone image file storage area at the Master Server local disk or at SAN storage. Folders on network drives, shared Folders (NFS, SMB, etc.) on other machines on the network, or UNC format folders cannot be specified.
- The server used to create the clone image and the servers targeted as clones must be the same model. If there are different models, a separate clone image must be created for each model. Refer to the Note in "[6.3 Adding Second and Subsequent Slave Servers](#)" for details.

The method for estimating the space required as a clone image file storage area is as follows:

Clone image file storage area = Disk space used by one slave server * Compression ratio * Number of models
--

Disk size used by one slave server

If actual results are available from a system build having the same software configuration, use the same disk size as that system. If one disk is split into multiple sections, use the total size used in all sections.

Use the operating system features to check the disk size used.

If actual results are not available from a system build having the same software configuration, make an estimation on the basis of the disk space given in software installation guides or similar.

Compression ratio

This is the compression ratio when the disk area used at the slave server is stored at the master server as an image file.

The compression ratio depends on the file content, but generally a ratio of about 50% can be expected.

3.2.3.3 Memory Size

The following amount of memory (excluding the OS) is required in order to use this product.

[Master server]

Installation environment	Memory size (in GB)
Physical environment/virtual environment	8.0 or more

[Slave server]

Installation environment	Memory size (in GB)
Physical environment/virtual environment	4.0 or more

[Development server]

Installation environment	Memory size (in GB)
Physical environment/virtual environment	4.0 or more

[Collaboration server]

Installation environment	Memory size (in GB)
Physical environment/virtual environment	4.0 or more

3.2.3.4 System Parameters

This section explains the system parameters that need tuning in order to guarantee stability of this product.

[Master Server] Primary master server in a physical environment

The following are the tuning values for system parameters required for the primary master server installed on a physical environment.

Shared memory

Parameter	Explanation	Value to be set	Type
kernel.shmmax	Maximum segment size in shared memory	53116928	Maximum
kernel.shmall	Total amount of shared memory available	655360	Maximum
kernel.shmuni	Maximum number of shared memory segments	113	Additional

Semaphore

For semaphore settings, set the values for each parameter in the following format:

```
kernel.sem = SEMMSLvalue SEMMNSvalue SEMOPMvalue SEMMNIvalue
```

Parameter	Explanation	Value to be set	Type
<i>SEMMSLvalue</i>	Maximum number of semaphores for each semaphore identifier	512	Maximum
<i>SEMMNSvalue</i>	Number of semaphores for the system as a whole	14862	Additional
<i>SEMOPMvalue</i>	Maximum number of operators for each semaphore call	50	Maximum
<i>SEMMNIvalue</i>	Number of semaphore operators for the system as a whole	2208	Additional

Message queue

Parameter	Explanation	Value to be set	Type
kernel.msgmax	Maximum message size	16384	Maximum
kernel.msgmnb	Maximum value for messages that can be held in one message queue	4194304	Maximum
kernel.msgmni	Maximum value for message queues in the system as a whole	1578 Minimum of 8192 required	Additional

[Master Server] Other than the above

The following are the tuning values for system parameters required for the primary master server and the secondary master server installed on a virtual environment.

Shared memory

Parameter	Explanation	Value to be set	Type
kernel.shmmax	Maximum segment size in shared memory	1048576 + 2776 * Number of disks for the shared disk device (*1) * 3 * 2	Maximum
kernel.shmall	Total amount of shared memory available	No change	-
kernel.shmuni	Maximum number of shared memory segments	30	Maximum

*1: "Number of disks for the shared disk device" refers to the following:

- In the case of a disk array unit, this is the number of logical units (LUN).
- In all other cases, this is the number of physical disks.

Semaphore

For semaphore settings, set the values for each parameter in the following format:

```
kernel.sem = SEMMSLvalue SEMMNSvalue SEMOPMvalue SEMMNIvalue
```

Parameter	Explanation	Value to be set	Type
<i>SEMMSLvalue</i>	Maximum number of semaphores for each semaphore identifier	No change	-
<i>SEMMNSvalue</i>	Number of semaphores for the system as a whole	11 Minimum of 41 required	Additional
<i>SEMOPMvalue</i>	Maximum number of operators for each semaphore call	No change	-
<i>SEMMNIvalue</i>	Number of semaphore operators for the system as a whole	2 Minimum of 22 required	Additional

Message queue

Parameter	Explanation	Value to be set	Type
kernel.msgmax	Maximum message size	16384	Maximum
kernel.msgmnb	Maximum value for messages that can be held in one message queue	4194304	Maximum
kernel.msgmni	Maximum value for message queues in the system as a whole	8192	Maximum

[Slave server]/[Development server]/[Collaboration server]

The following are the tuning values for system parameters required for the slave server, development server, and collaboration server.

Semaphore

For semaphore settings, set the values for each parameter in the following format:

```
kernel.sem = SEMMSLvalue SEMMNSvalue SEMOPMvalue SEMMNIvalue
```

Parameter	Explanation	Value to be set	Type
<i>SEMMSLvalue</i>	Maximum number of semaphores for each semaphore identifier	No change	-
<i>SEMMNSvalue</i>	Number of semaphores for the system as a whole	11 Minimum of 41 required	Additional
<i>SEMOPMvalue</i>	Maximum number of operators for each semaphore call	No change	-

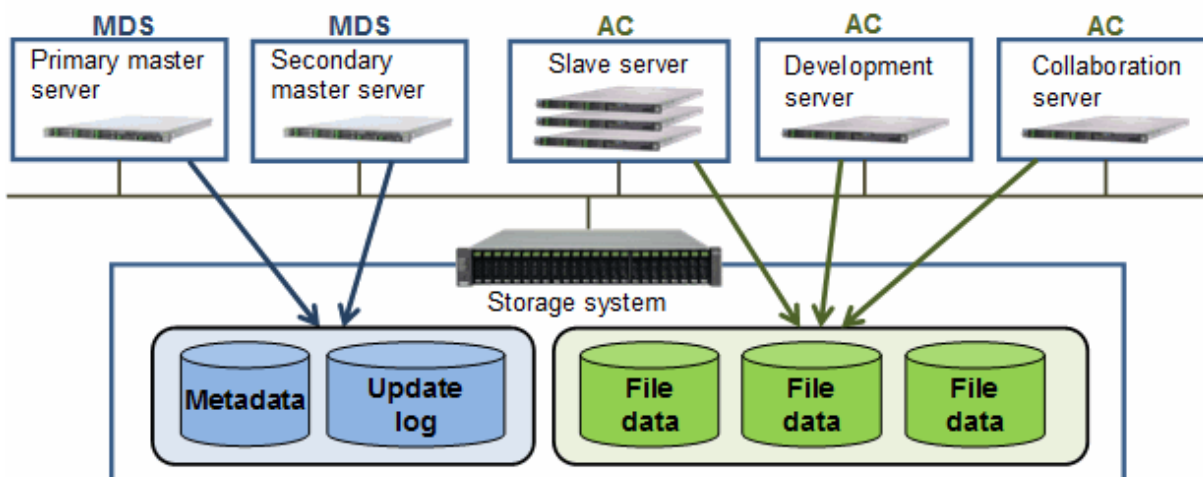
Parameter	Explanation	Value to be set	Type
<i>SEMMN</i> value	Number of semaphore operators for the system as a whole	2 Minimum of 22 required	Additional

3.3 File System Design

This section describes how to design the DFS file system of this product.

3.3.1 File System Configuration

This section describes the configuration of the file system using this product.



The DFS consists of the DFS management server and the DFS client.

The DFS management server (MDS) runs on the master server. DFS client functions (AC) run on slave servers, development servers and collaboration servers.

The partitions that comprise a file system in a DFS are split into three areas.

Area type	Component used
Metadata area	MDS
Update log area	
File data area	AC

3.3.2 File System Configuration Design

Make the following decisions when designing the configuration of the file system.

Capacity to be allocated to shared disk devices

Estimate the total size of the file data to be stored in the DFS, then based on that, estimate the capacity required by the latter on shared disk devices.

Area type	Estimation method
(a) File data area	Space required for file data
(b) Metadata area	If file data area \leq 1 TB: 100 GB If file data area $>$ 1 TB: 300 GB
(c) Update log area	Estimation is not required because it is contained within the metadata area.
(d) Management partition	Constant value (no estimation required): 1 GB
Total	(a)+(b)+(d)

Partitions used

Determine the shared disk device partitions to be used in the DFS based on the size estimated above, and check the device name.

Determine the areas to be used as the management partitions, representative partitions, and file data partitions of the multiple partitions available.



Example

Examples of the partitions for the areas that comprise the file systems:

- Management partition `:/dev/disk/by-id/scsi-1FUJITSU_300000370105`
- Representative partition `:/dev/disk/by-id/scsi-1FUJITSU_300000370106`
- File data partitions: `:/dev/disk/by-id/scsi-1FUJITSU_300000370107`
`:/dev/disk/by-id/scsi-1FUJITSU_300000370108`



Point

- Separate the metadata area and file data area and allocate them to separate partitions. The management partition also needs a separate partition.
- The metadata area and file data area of the DFS can be deployed to a single partition, but by deploying them to separate partitions, the I/O distribution can result in improved throughput.
If the DFS is constructed with multiple partitions, representative partitions are used as the partitions for the metadata area.
- A maximum of 256 partitions can be used.
- A by-id name generated by the udev function is used for shared disk device names (refer to "[6.1.3.1 Checking Shared Disk Settings](#)" for details).



Information

Deploying the metadata area and file data area to separate partitions distributes the I/O processes to each area, thus avoiding conflicts. Specify separation of the file data area using `pdfsmkfs` with the `dataopt` option when creating the file system.

Size of the file data area

If it is likely that the DFS size will be extended, allow for future extensions by estimating the maximum extended size when creating the file system.

Specify the maximum size using pdfsmkfs with the maxdsz option when creating the file system.



Refer to "pdfsmkfs" in the "Appendix A Command Reference" of the "Primesoft Distributed File System for Hadoop V1 User's Guide" for information on the size of the file data area.

Data block size

The blocksz option of the pdfsmkfs command can be used to specify the data block size during DFS creation. Specification of the data block size enables contiguous allocation of shared disk device area, which enables efficient input-output processing.

A data block size of 8 MB is recommended if the DFS is used by Hadoop.



If 8 MB is specified as the data block size, an 8 MB area is used on the shared disk even if the file size is less than 8 MB. If a large quantity of small-sized files are stored, give priority to space efficiency and do not specify a data block size.



Refer to "3.3.2.1 Relationship between File System Size, Data Block Size and Maximum File Size" for the relationship between the data block size the maximum file system size, and maximum file size.

3.3.2.1 Relationship between File System Size, Data Block Size and Maximum File Size

If a file system is created without specifying a data block size value, the data block size is calculated automatically on the basis of the file data area size or the maximum size of the partitions comprising the file system. The greater the file system size, the greater the data block size. The greater the data block size, the greater the maximum file size.

The table below shows the relationship between the file system size, the data block size and the maximum file size when a file system is created without the data block size being specified.

Table 3.1 Relationship between file system size, data block size and maximum file size

File system size	Data block size	Maximum file size
to 1TB	8 KB	1 TB - 8 KB
(1 TB + 1 Byte) to 2TB	8 KB to 16 KB	(1 TB - 8 KB) to (2 TB - 16 KB)
(2 TB + 1 Byte) to 4 TB	8 KB to 32 KB	(1 TB - 8 KB) to (4 TB - 32 KB)
(4 TB + 1 Byte) to 8 TB	8 KB to 64 KB	(1 TB - 8 KB) to (8 TB - 64 KB)
(8 TB + 1 Byte) to 16 TB	16 KB to 128 KB	(2 TB - 16 KB) to (16 TB - 128 KB)
(16 TB + 1 Byte) to 32TB	32 KB to 256 KB	(4 TB - 32 KB) to (32 TB - 256 KB)
(32 TB + 1 Byte) to 64 TB	64 KB to 512 KB	(8 TB - 64 KB) to (64 TB - 512 KB)
(64 TB + 1 Byte) to 128 TB	128 KB to 1 MB	(16 TB - 128 KB) to (128 TB - 1 MB)
(128 TB + 1 Byte) to 256 TB	256 KB to 1 MB	(32 TB - 256 KB) to (128 TB - 1 MB)
(256 TB + 1 Byte) to 512 TB	512 KB to 1 MB	(64 TB - 512 KB) to (128 TB - 1 MB)

File system size	Data block size	Maximum file size
(512 TB + 1 Byte) to 1 PB	1 MB	128 TB - 1 MB
(1 PB + 1 Byte) to 2 PB	2 MB	256 TB - 2 MB

 **Note**

The file system size is not just the file data area size. It also includes the sizes of areas such as the metadata area and the update log area. Therefore, near the boundary values for file system sizes in the above table, the data block size values might be one step smaller.

 **See**

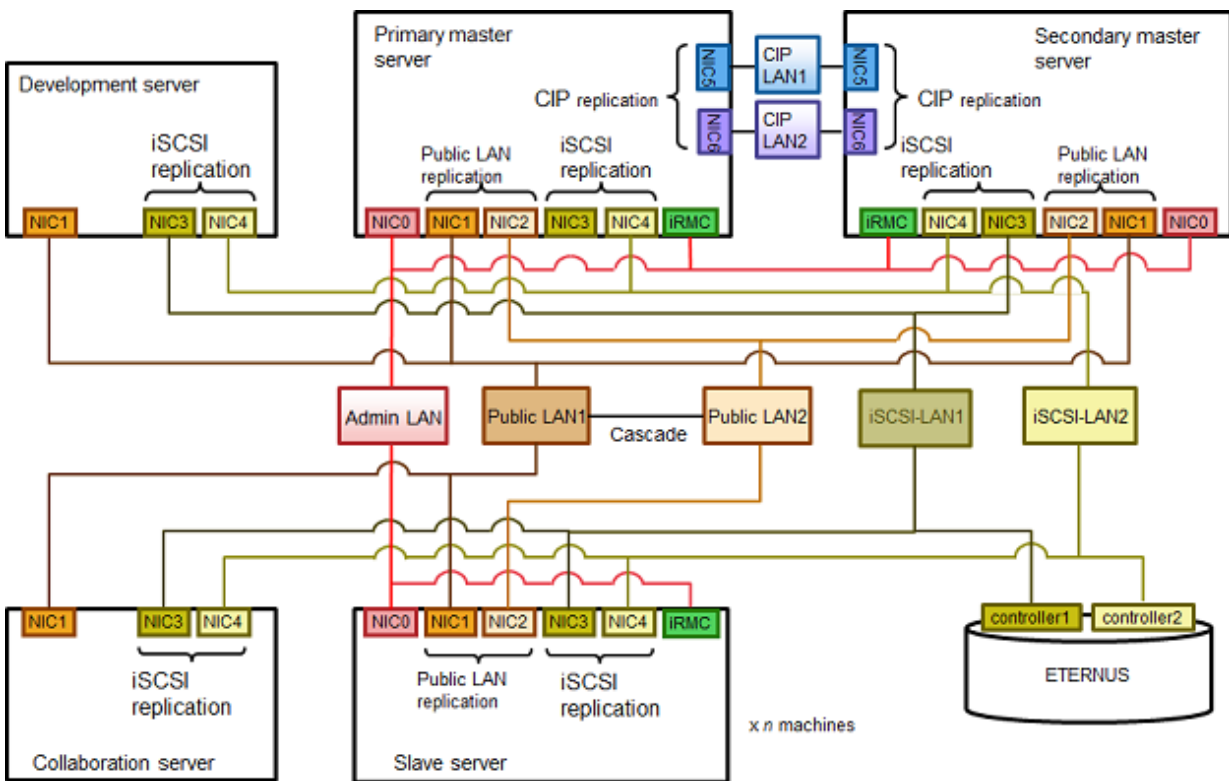
Refer to "B.3 Limit Values" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for the file system maximum values. The file system data block size can be changed during file system creation. Refer to pdfsmkfs under "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for details.

3.4 Network Design

This section explains how to design the network for this product.

3.4.1 Network Configuration

This section describes the network configurations and LAN types using this product.



Admin LAN

LAN used to perform the cloning processes in Smart Setup.

This is established separately to the public LAN used for Hadoop parallel distributed processing.

Use the first network interface as the connection for the admin LAN.

Note

When installing on a virtual environment (VMware)

When configuring a replicated master server for installation, make sure you match the names of the NIC interfaces connected to the admin LAN.

For example, if the name of the admin LAN interface of the primary master server is 'eth0', then configure the network so that the admin LAN interface of the secondary master server is also 'eth0'.

Public LAN

This is the LAN for parallel distribution processing between a master server and slave servers.

You can add redundancy to the public LAN using the network redundancy software included with this product.

When installing on a virtual environment, public LAN redundancy can be realized using the NIC teaming feature on the host machine.

Information

Public LAN redundancy using network redundancy software

Switching of transmission routes is controlled with NIC switching on the master server and slave server. IP inheritance occurs as follows with NIC switching:

- Master server: logical IP address inheritance
- Slave server: physical IP address inheritance

See

Refer to the manuals of the server virtualization software product for information on the NIC teaming feature.

Cluster interconnect (CIP) LAN

This is the LAN used for a HA cluster configuration for the primary server and the secondary server.

You can add redundancy using the high-reliability infrastructure software included with this product.

When installing on a virtual environment, cluster interconnect (CIP) LAN redundancy can be realized using the NIC teaming feature on the host machine.

Note

Cluster interconnect (CIP) settings are still required even if the master server is not to be replicated.

See

- Refer to "[3.2.2 Design the Server Configuration](#)" for information on how to design replicated configuration for the master server.
- Refer to the manuals of the server virtualization software product for information on the NIC teaming feature.

iSCSI-LAN

This product uses Internet Small Computer System Interface (iSCSI) as the interface between the servers and the storage systems. This is the LAN for this iSCSI connection.

A redundancy configuration by means of the ETERNUS multipath driver is possible.

Note that redundancy is not possible when installed on a virtual environment.

Point

It is recommended to make the transfer speed between the storage systems and the network switch 10Gbit/s or more.

Note

When installing on a virtual environment

- For the admin LAN, public LAN, and cluster interconnect (CIP) LAN to communicate with outside networks, it is necessary to have a bridge connection to the physical NIC corresponding to the virtual machine.
- If the virtual machine (virtual disk image) is to be stored in shared storage, configure a separate storage network to connect the host machine and the shared storage.
- If network communication is required by features of the virtualization software, configure so that it does not share with the public LAN, iSCSI-LAN, and cluster interconnect (CIP) LAN used by this product.

3.4.2 Network Configuration Design

Make decisions regarding the various LANs indicated in the figure in "3.4.1 Network Configuration" when determining the design of the network configuration.

Replication of the public LAN

Decide whether to have public LAN redundancy.

This is recommended so that parallel distribution processing can continue even when some parts of the LAN experiences a failure.

Replication of CIP

When the master server is replicated, decide whether to have cluster interconnect (CIP) redundancy.

Cluster interconnect (CIP) is required to perform alive monitoring between the primary master server and the secondary master server, so it is recommended to introduce redundancy to prepare for any failures in the LAN.

Replication of iSCSI

When using the ETERNUS external storage device, decide whether to have redundancy for the iSCSI-LAN that connects the storage systems.

When installed on a physical environment, it is recommended to have redundancy in preparation for LAN failures.

Point

If redundancy with the LANs is to be achieved, include redundancies as follows, according to the configuration to be installed.

LAN type	Physical environment	Virtual environment
Public LAN	Redundancy is achieved by setting the following parameters in bdpp.conf. [Master server] BDPP_GLS_PRIMARY_CONNECT2 BDPP_GLS_SECONDARY_CONNECT2 BDPP_GLS_POLLING2_IP [Slave server] BDPP_GLS_SERVER_CONNECT2 BDPP_GLS_SERVER_POLLING2_IP	Redundancy on the host machine (required). In this case, the parameters listed in the previous cell do not need to be set.
CIP LAN	Redundancy is achieved by setting the following parameters in bdpp.conf . [Master server] BDPP_PCL_PRIMARY_CONNECT2 BDPP_PCL_SECONDARY_CONNECT2	Redundancy on the host machine (recommended). In this case, the parameters listed in the previous cell do not need to be set. If redundancy is required on the virtual machine, configure in the same way as the physical environment.
iSCSI-LAN	Redundancy using the ETERNUS multipath driver is performed.	Redundancy unavailable.



See

Refer to "B.1 bdpp.conf" for information on bdpp.conf.

3.5 Designing User Accounts

This section explains the user accounts for installing, running, and maintaining this product.

User type	User name	Description
System administrator	root	System administrator with root permissions (superuser) Installs and maintains this product, and starts and stops Hadoop.
MapReduce executing user	mapred (fixed)	Executes JobTracker and TaskTracker processes.
Hadoop user	Optional	Executes, stops, and displays status of Hadoop jobs (job executing user)

Make the following design decisions in order to operate this product.

Deciding users of Hadoop

You can register more than one user account for the Hadoop user.

You can register Hadoop users by specifying the user accounts and groups that users belong to when installing and setting up this product.

Check the anticipated Hadoop users (user names and user IDs) beforehand.



See

.....
Refer to "[11.1 Adding Job Execution Users](#)" for information on how to manage Hadoop users (job execution users).
.....

Chapter 4 System Requirements

This chapter explains the system requirements for this product.

4.1 Hardware Requirements

This section describes the hardware requirements for using this product.

4.1.1 Hardware Conditions

The following hardware conditions must be met for each of the features.

Functionality	Hardware	Notes
Master server	PRIMERGY RX Series, PRIMERGY TX Series	The CPU must be at least a dual-core CPU.
Slave server	PRIMERGY RX Series, PRIMERGY TX Series (*1)	The CPU must be at least a dual-core CPU.
Development server	PRIMERGY RX Series, PRIMERGY TX Series	The CPU must be at least a dual-core CPU.
Collaboration server	PRIMERGY RX Series, PRIMERGY TX Series	The CPU must be at least a dual-core CPU.
External storage device	ETERNUS DX series	

*1: Refer to the supported model information at the following site for the PRIMERGY RX and TX Series models supported by this product. Note that there are no restrictions on supported models when building virtual environments.

- Supported model information

Refer to following URL for detail information on supported PRIMERGY RX/TX:

<http://globalsp.ts.fujitsu.com/dmsp/Publications/public/ds-ror-ve-v3-0-ww-en.pdf>

4.1.2 Expansion Card Options

Additional network interface cards are required when building LAN redundancy.

Add network interface cards according to the design prepared in "[3.4.1 Network Configuration](#)".

4.2 Software Configuration

This product is comprised of the following DVD-ROMs:

- Interstage Big Data Parallel Processing Server

This product is comprised of the following software:

Software name	Functionality overview
Interstage Big Data Parallel Processing Server Standard Edition V1.0.1	The master server functionality of this product

Software name	Functionality overview
Master server	
Interstage Big Data Parallel Processing Server Standard Edition V1.0.1 Slave server	The slave server functionality of this product
Interstage Big Data Parallel Processing Server Standard Edition V1.0.1 Development server	The development server functionality of this product
Interstage Big Data Parallel Processing Server Standard Edition V1.0.1 Collaboration server	The collaboration server functionality of this product

4.3 Software Requirements

This section describes the software requirements for using this product.

4.3.1 System Software

The following system software is required in order to use this product:

[Master server], [Slave server]

OS type	System software	Notes
Linux	Red Hat(R) Enterprise Linux(R) 5.6 (for Intel64) (*1) Red Hat(R) Enterprise Linux(R) 5.7 (for Intel64) (*1) Red Hat(R) Enterprise Linux(R) 5.8 (for Intel64) (*1) Red Hat(R) Enterprise Linux(R) 5.9 (for Intel64) (*1) Red Hat(R) Enterprise Linux(R) 6 (for Intel64) (*1) Red Hat(R) Enterprise Linux(R) 6.1 (for Intel64) (*1) Red Hat(R) Enterprise Linux(R) 6.2 (for Intel64) (*1) Red Hat(R) Enterprise Linux(R) 6.3 (for Intel64) (*1) Red Hat(R) Enterprise Linux(R) 6.4 (for Intel64) (*1)	If software such as driver or update kits is mandatory, prepare the software. Refer to the server manual or the Linux installation guide for information on mandatory software.

*1: Only the cloning functionality runs as a 32-bit application on the WOW64 (Windows 32-bit On Windows 64-bit) subsystem.

[Development server], [Collaboration server]

OS type	System software	Notes
Linux	Red Hat(R) Enterprise Linux(R) 5.6 (for Intel64) Red Hat(R) Enterprise Linux(R) 5.7 (for Intel64) Red Hat(R) Enterprise Linux(R) 5.8 (for Intel64) Red Hat(R) Enterprise Linux(R) 5.9 (for Intel64) Red Hat(R) Enterprise Linux(R) 6 (for Intel64) Red Hat(R) Enterprise Linux(R) 6.1 (for Intel64) Red Hat(R) Enterprise Linux(R) 6.2 (for Intel64) Red Hat(R) Enterprise Linux(R) 6.3 (for Intel64) Red Hat(R) Enterprise Linux(R) 6.4 (for Intel64)	If software such as driver or update kits is mandatory, prepare the software. Refer to the server manual or the Linux installation guide for information on mandatory software.

When installing on a virtual environment

When installing this product on a virtual environment, the following virtualization system software is required on the host machine.

Build virtual machines on these host machines (host OS's), and install the system software (guest OS's) on the servers above.

Virtual environment type	Operating system (host OS)	Remarks
KVM	Red Hat(R) Enterprise Linux(R) 6 (for Intel64) Red Hat(R) Enterprise Linux(R) 6.1 (for Intel64) Red Hat(R) Enterprise Linux(R) 6.2 (for Intel64) Red Hat(R) Enterprise Linux(R) 6.3 (for Intel64) Red Hat(R) Enterprise Linux(R) 6.4 (for Intel64)	If software such as driver or update kits is mandatory, prepare the software.
VMware (vSphere ESXi)	VMware vSphere 5.0 Standard VMware vSphere 5.0 Enterprise VMware vSphere 5.0 Enterprise Plus VMware vSphere 5.1 Standard VMware vSphere 5.1 Enterprise VMware vSphere 5.1 Enterprise Plus	Refer to the server manual or the Linux installation guide for information on mandatory software.

Mandatory Packages

The mandatory packages shown in "[Appendix F Mandatory Packages](#)" are required in order to use this product.

If any packages are not installed during installation of this product, an error occurs. Check the message and install the required packages.

4.3.2 Mandatory Software

The following software is required in order to use this product:

[Master server]

Software name	Version	Notes
Microsoft(R) LAN Manager module	-	Only required for the primary master server. Obtain from the Microsoft FTP site. (*1) This is not required when installing on a virtual environment.

*1: Obtain from the following Microsoft FTP site:

URL:

`ftp://ftp.microsoft.com/bussys/clients/msclient/DSK3-1.EXE (as of October 2013)`

Note that the Microsoft LAN Manager module can be used regardless of the CPU architecture (x86, x64).

If this product is being installed in an environment where ServerView Deployment Manager is installed, it is not necessary to obtain the Microsoft LAN Manager module.

[Slave server]

Software name	Version	Notes
ServerView Agent for Linux	V4.50.12 or later	This is not required when installing on a virtual environment.

4.3.3 Exclusive Software

None.

4.3.4 Related Software

Consider installation of the following software, if required:

[Master server]

Software name	Version	Notes
ETERNUS Multipath Driver	V2.0L22 or later	Required if providing redundant connections to the storage system This is not required when installing on a virtual environment.

[Slave server]

Software name	Version	Notes
ETERNUS Multipath Driver	V2.0L22 or later	Required if providing redundant connections to the storage system This is not required when installing on a virtual environment.

[Development server]

Software name	Version	Notes
Interstage Application Server Enterprise Edition or Interstage Application Server Standard-J Edition	V10.0.0 or later	Required if developing analysis applications (MapReduce applications) that run under Hadoop
ETERNUS Multipath Driver	V2.0L22 or later	Required if providing redundant connections to the storage system This is not required when installing on a virtual environment.

[Collaboration server]

Software name	Version	Notes
ETERNUS Multipath Driver	V2.0L22 or later	Required if providing redundant connections to the storage system This is not required when installing on a virtual environment.

Others

Software name	Version	Notes
VMware vCenter Server	5.0 5.1	When using a virtual environment that uses VMware vSphere, this software is required for managing VM guests and VM hosts.
VMware vSphere Client	5.0 5.1	When using a virtual environment that uses VMware vSphere, this software is required to use functionality that links to the VM management product or the VMware for Managed Servers on the management client.

Chapter 5 Preparing to Build the System

This chapter describes tasks that must be performed before installing this product.

Setting sequence/Setting items		Installation environment				
		Master server		Slave server	Development server	Collaboration server
		Primary	Secondary			
1	Server BIOS Settings	Y	Y	Y	Y	N
2	File System When Installing the OS	N	N	Y (*1)	N	N
3	Disabling SELinux Functionality	Y	Y	Y	Y	Y
4	System Parameter Settings	Y	Y	Y	Y	Y
5	Firewall Settings	Y	Y	Y	Y	N
6	Host Name Settings	Y	Y	Y	Y	Y
7	Server Name Settings	Y	Y	Y	Y	N
8	Public LAN Network Interface Settings	Y	Y	Y	Y	Y
9	System Time Settings	Y	Y	Y	Y	Y
10	Storage System Environment Preparations	Y	Y	Y	Y	Y
11	ssh Settings	Y	Y	N	N	N
12	SNMP Trap Daemon Settings	Y (*1)	N	N	N	N
13	ServerView Agent Settings	N	N	Y (*1)	N	N
14	kdump Shutdown Agent Settings	Y (*1)	Y (*1)	N	N	N
15	KVM Shutdown Agent Settings	Y (*2)	Y (*2)	N	N	N
16	Expanding the Microsoft LAN Manager Module	Y (*1)	N	N	N	N

Y: Required

N: Not required

*1: Required only when installing in a physical environment

*2: Required only when installing in a virtual environment (KVM)

5.1 Server BIOS Settings

This section describes the system BIOS configuration required.

Hyper-threading

Disable hyper-threading in the master servers, slave servers, and development servers.

Point

- There is no guarantee that hyper-threading will improve performance. For this reason, this product optimizes parameters for Hadoop under the assumption that hyper-threading is disabled.
Refer to "[Appendix C Hadoop Configuration Parameters](#)" for information on Hadoop parameters.
- Configure the system BIOS of the host machine when installing on a virtual environment.

Boot sequence

Set the boot sequence for the slave server as follows:

1. Boot from the network interface used in the admin LAN (the first network interface)
2. Boot from DVD-ROM or CD-ROM (if a DVD-ROM or CD-ROM is connected)
3. Boot from storage

Note

- Do not change the boot sequence even if operation is started from the managed Server and then booted from the disk.
- If "UEFI" and "Legacy" are displayed when setting the boot from the network interface, select "Legacy".

Point

- Settings are necessary to install other slave servers by cloning.
- The boot sequence setting is not required when installing the slave server on a virtual environment.

5.2 File System When Installing the OS

To introduce another slave server by cloning, it is necessary to make the file system type "EXT3" when installing the OS on the slave server.

If the file system type used is "LVM (Logical Volume Manager)", it is not possible to introduce a slave server using cloning. Install the OS with the file system type "EXT3".

Point

This setting is not required when installing the slave server on a virtual environment.

5.3 Disabling SELinux Functionality

This product cannot be installed in a Linux environment where the SELinux (Security-Enhanced Linux) functionality is enabled.

To install this product, disable the SELinux functionality, and then install this product.



See

Refer to the Linux online manual for the method for disabling the SELinux functionality.

5.4 System Parameter Settings

The system parameters must be tuned.

Use the following procedure to tune system parameters:

1. Use the following command to check the current settings for the system parameters:

```
# /sbin/sysctl -a <Enter>
```



Example

```
# /sbin/sysctl -a <Enter>
... omitted ...
kernel.sem = 250      32000   32      128
kernel.msgmnb = 4194304
kernel.msgmni = 8192
kernel.msgmax = 65536
kernel.shmni = 4096
kernel.shmall = 4294967296
kernel.shmmax = 68719476736
... omitted ...
```

2. Refer to "[3.2.3.4 System Parameters](#)", and compare the current settings. Calculate an appropriate value for each parameter, taking into account the parameter type ("Maximum" or "Addition").



Point

Set values as shown below in accordance with the parameter "Type".

- If the type is "Maximum":

If the value already set (initial value or previous value) is greater than "Value to set (tuning value)", it need not be changed. If it is smaller than the tuning value, change it to the tuning value.

- If the type is "Additional":

Add the "Value to set (tuning value)" to the value already set (initial value or previous value). Check the system maximum values before adding this value and, if adding that value would exceed the system maximum value, set the system maximum value.

Refer to a Linux manual or similar for details.

3. Edit /etc/sysctl.conf as shown in the following example.



Example

```
kernel.sem = 250 32002 32 130
kernel.shmni = 4096
kernel.msgmnb = 4194304
kernel.msgmni = 8192
```

4. Use the following command to check whether the changes have been applied to the /etc/sysctl.conf file:

```
# /bin/cat /etc/sysctl.conf <Enter>
```

5. Use one of the following methods to enable the settings made in Step 3.

- Restart the system to apply the settings

```
# /sbin/shutdown -r now <Enter>
```

- Apply the settings

```
# /sbin/sysctl -p /etc/sysctl.conf <Enter> (*1)
```

*1: You do not need to restart the system if you use this command.

6. Use the following command to check whether the specified system parameters have been updated:

```
# /sbin/sysctl -a <Enter>
```



Example

```
# /sbin/sysctl -a <Enter>
... omitted ...
kernel.sem = 250          32002   32       130
kernel.msgmnb = 4194304
kernel.msgmni = 8192
kernel.msgmax = 65536
kernel.shmmni = 4096
kernel.shmall = 4294967296
kernel.shmmax = 68719476736
... omitted ...
```

5.5 Firewall Settings

If this product is installed in an environment where firewall functions are used, the required communications (ports) must be authorized in the firewall function.

Refer to the operating system manual for the method for authorizing the required communications (ports) in the firewall function.

Refer to "[Appendix D Port List](#)" for the ports used by this product, and ensure connection is authorized.

5.6 Host Name Settings

The host names used by Hadoop need to be set for the master servers, slave servers, development servers, and collaboration servers.

In the hosts file, enter the host names for the NICs to be connected to the public LANs of the servers.

hosts file

```
/etc/hosts
```

Refer to the online manual for information on the hosts file.

Point

- Specify strings meeting the following conditions as the host names:
 - Must contain only lowercase alphanumeric characters
 - The first character must be a letter
 - Must be up to 11 characters
- The content of the hosts file on the primary master server must match the secondary master server in order to build a master server replicated configuration.
- Add the host name information added to the master server to the hosts files of the slave servers, development servers and collaboration servers.
- Match the server names with the host names of the master servers, slave servers, and development servers (refer to "5.7 Server Name Settings" for details).

Note

Note the following if registering a local host in the hosts file.

- If setting a local host name for "127.0.0.1", it is essential to first enter an IP address that can be referenced remotely. Alternatively, do not set a local host name for "127.0.0.1".
- Do not use an alias for the Host name.

Example

Example of IP addresses and host names allocated to the NICs connected to the public LANs of the different servers:

- | | | |
|---------------------------|-----------------------------|-------------------|
| - Primary master server | :10.10.10.11 | :master1 |
| - Secondary master server | :10.10.10.12 | :master2 |
| - Slave server | :10.10.10.21 to 10.10.10.25 | :slave1 to slave5 |
| - Development server | :10.10.10.30 | :develop |
| - Collaboration server | :10.10.10.31 | :collaborate |

Set the hosts file on the master servers as follows:

```
10.10.10.11 master1
10.10.10.12 master2
127.0.0.1 localhost.localdomain localhost

10.10.10.21 slave1
10.10.10.22 slave2
10.10.10.23 slave3
10.10.10.24 slave4
10.10.10.25 slave5
10.10.10.30 develop
10.10.10.31 collaborate
```

5.7 Server Name Settings

The server name (the host name of the system as output by a command such as `hostname`) must be the same as the host name corresponding to the NIC connected to the public LAN.

If it is different, set the host name that was set in the network file in "5.6 Host Name Settings".

network file

```
/etc/sysconfig/network
```

Set the setting name in the `HOSTNAME` parameter.



Example

Confirm the server name of the primary master server.

```
# hostname <Enter>
master1
```

Edit the network file to change the server name.

```
NETWORKING=yes
HOSTNAME=master1
```



Information

If an existing server is to be used as a collaboration server, the server name does not need to match the host name corresponding to the NIC connected to the public LAN.

5.8 Public LAN Network Interface Settings

Before this product is installed the IP address must be set for the public LAN network interface of each server, and the interfaces must be in the active state.

Use the `ifconfig` command to check the status of the targeted network interfaces. If settings are not yet set or if the status is inactive, set the IP address and make the interface active.



Example

```
# ifconfig eth1 <Enter>
eth1      Link encap:Ethernet  HWaddr 00:19:99:8D:7E:90
          inet addr:10.10.10.11  Bcast:10.10.10.255  Mask:255.255.255.0
          ... omitted ...
          RX bytes:2440357552 (2.2 GiB)  TX bytes:1144015405 (1.0 GiB)
          Memory:cd400000-cd480000
```

5.9 System Time Settings

Use NTP (Network Time Protocol) to set the system times of the master servers, slave servers, development servers and collaboration servers to the same time.

Note

When installing on a virtual environment (KVM)

With the host machine as the NTP server, synchronize times with the virtual machines as NTP clients.

If virtual machines have been built on multiple physical machines, also synchronize times on the host machine using NTP.

5.10 Storage System Environment Preparations

If DFS is used for the distributed file system, a shared LUN must be created for the storage system, and iSCSI settings and multipath settings (if iSCSI is duplicated) must be completed in advance.

Refer to the respective manuals for the storage system and iSCSI settings and for information on installing and setting the ETERNUS multipath driver.

Note

ETERNUS multipath drivers do not need to be installed or configured in virtual environments.

5.11 ssh Settings

The master server needs to be configured so that it can perform ssh communication with itself without a password.

Furthermore, in a master server replicated configuration, also configure so that the primary master server and the secondary master server can use ssh communications with each other without a password.

Create a public key on both the primary master server and the secondary master server so that ssh connection between the two occurs with root permissions and does not require a password.

See

Refer to the help for the ssh-keygen command for information on how to create a public key.

Note

When starting or stopping the Hadoop of this product, Hadoop communicates via ssh.

Furthermore, the content of the definitions files on the master server need to be matched with the various servers during Hadoop setup and DFS file system configuration. In this document, the procedure is to copy the definition files on the master server using scp.

Be aware that prompts for the password are displayed when using ssh/scp.

The following shows whether ssh connections between the servers require root permission.

		Connected server				
		Master server		Slave server	Development server	Collaboration server
		Primary	Secondary			
Initiating server	Primary master server	Y (*1)	Y (*1)	Y (*2)	Y	Y

		Connected server				
		Master server		Slave server	Development server	Collaboration server
		Primary	Secondary			
	Secondary master server	Y (*1)	Y (*1)	Y (*2)	N	N
	Slave server	Y	N	-	N	N
	Development server	Y	N	N	-	N
	Collaboration server	Y	N	N	N	-

Y: Required

N: Not required

*1: It is necessary to be able to perform ssh connection without a password

*2: The master server communicates via ssh with the slave servers when starting and stopping Hadoop, therefore a prompt will be displayed for each slave server if you do not make the settings for ssh connection without a password

5.12 SNMP Trap Daemon Settings

In order for this product to operate correctly, the net-snmp package must be installed and the setting below must be added to the `/etc/snmp/snmptrapd.conf` file on the primary master server. If the file does not exist, create the file, and then add the following setting:

```
disableAuthorization yes
```

Point

This setting is not required when installing the master server on a virtual environment.

5.13 ServerView Agent Settings

Install ServerView Agent on the slave servers and set the required SNMP service.

Refer to the ServerView Agent manual for the method for setting the SNMP service.

- For the SNMP community name, set the same value as the SNMP community name set in the iRMC.
- Set Read (reference permission) or Write (reference and update permission) for the SNMP community name.
- Set the Master Server admin LAN IP address as the Host that receives SNMP packets.
- Set the Master Server IP address as the SNMP trap send destination.
- If the Master Server that is the SNMP trap send destination has multiple NICs, set the IP address of the admin LAN of the side connected to the slave servers.

Point

This setting is not required when installing the slave server on a virtual environment.

5.14 kdump Shutdown Agent Settings

Set so that the kdump can be used to gather crash dumps if a system panic occurs in a master server.

This must be done on both the primary master server and the secondary master server in order to build a master server replicated configuration.

1. Use the runlevel command to check the current run level.

Example

In the following example, the current run level is 3.

```
# /sbin/runlevel <Enter>
N 3
```

2. Use the chkconfig command to check the kdump usability status.

Example

In the following example, the current run level 3 kdump is off.

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

3. If kdump is off at the current run level, use the chkconfig command to switch it on, then use the service command to start kdump.

```
# /sbin/chkconfig kdump on <Enter>
# /sbin/service kdump start <Enter>
```

Point

This setting is not required when installing the master server on a virtual environment.

5.15 KVM Shutdown Agent Settings

The shutdown agent needs to be configured if the master server is installed on a virtual environment (KVM).

libvirt-guests settings

1. Log in to the host machine with root permissions.
2. Define the following values in the /etc/sysconfig/libvirt-guests file:
 - ON_SHUTDOWN=shutdown
 - SHUTDOWN_TIMEOUT=*timeToWaitForVirtualMachineToShutdown*

Example

If the time taken for the virtual machine to shutdown is less than 300 seconds (5 minutes):

```
# cat /etc/sysconfig/libvirt-guests <Enter>
... omitted ...
```

```
ON_SHUTDOWN=shutdown
SHUTDOWN_TIMEOUT=300
```

Point

This setting is made so that if the host machine is mistakenly shutdown while virtual machines are running, the virtual machines that contain master servers can be stopped properly. In SHUTDOWN_TIMEOUT, set the longest time it takes for shutdown when multiple virtual machines have been created on the host machine.

Note

When making a replicated configuration for the master server

When the primary master server and the secondary master server are to be installed on separate host machines, do this on both host machines.

Creating a user for the shutdown agent

Create user accounts on the host machine.

1. Log in to the host machine with root permissions.
2. Set any password for the user.

Example

Creating the user pcluser for the shutdown agent:

```
# useradd pcluser <Enter>
# passwd pcluser <Enter>
```

3. Allow users to execute commands as the root user.

Execute visudo, and in the editor started by the command, add a setting line for the user.

```
# visudo <Enter>
```

Example

Configuring to allow the user pcluser to execute commands as root without entering a password:

```
pcluser          ALL=(ALL)          NOPASSWD: ALL
```

4. Log in to the virtual machine where the master server is to be installed.
5. Go through the process of the initial user query (RSA key generation) that happens with ssh connection.

This must be done on both the primary master server and the secondary master server in order to build a master server replicated configuration.

Example

If the IP address of the KVM host machine is 192.168.1.100 and the user account for the shutdown agent is pcluser:

```
# ssh -l pcluser 192.168.1.100 <Enter>
The authenticity of host '192.168.1.100 (192.168.1.100)' can't be established.
```

```
RSA key fingerprint is xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx.
Are you sure you want to continue connecting (yes/no)? yes <Enter>
Warning: Permanently added '192.168.1.100' (RSA) to the list of known hosts.
pcluser@192.168.1.100's password: passwordForShutdownAgentUser <Enter>
$ exit
```

Point

The IP address for the KVM host machine, the created user accounts, and passwords need to be set in `BDPP_KVMHOST_IP`, `BDPP_KVMHOST_ACCOUNT`, and `BDPP_KVMHOST_PASS` parameters (respectively) in `bdpp.conf` specified during installation of the master server.

Refer to "Master server" in "B.1 [bdpp.conf](#)" for information on `bdpp.conf`.

Note

When making a replicated configuration for the master server

When the primary master server and the secondary master server are to be installed on separate host machines, do this on both host machines.

In this case, set the same user account and password for the shutdown agent. Furthermore, also perform the initial access required for ssh connection on the host machines for the primary master server and the secondary master server to the other host machines.

5.16 Expanding the Microsoft LAN Manager Module

Use the `Expand` command to expand the module obtained at "[4.3.2 Mandatory Software](#)" in the CPU architecture (x86) Windows in advance. Refer to the example below for the expansion method.

Example

If `dsk3-1.exe` is deployed to `C:\temp`:

```
> cd /d c:\temp
> dsk3-1.exe
> Expand c:\temp\protman.do_ /r
> Expand c:\temp\protman.ex_ /r
```

Use Windows 8.3 format (*1) for folder and file names.

The expanded Microsoft LAN Manager module is no longer required after the Manager is installed.

*1: The rules are a maximum of eight characters for the file name part and a maximum of three characters for the extension part.

Place the following expanded modules in the work directory (`/tmp`) of the master server:

- PROTMAN.DOS
- PROTMAN.EXE
- NETBIND.COM

Information

The modules deployed in the working directory are loaded when the master server feature of this product is installed.

 Point

.....
This setting is not required when installing the master server on a virtual environment.
.....

Chapter 6 Installation

This chapter explains how to build a system using this product.

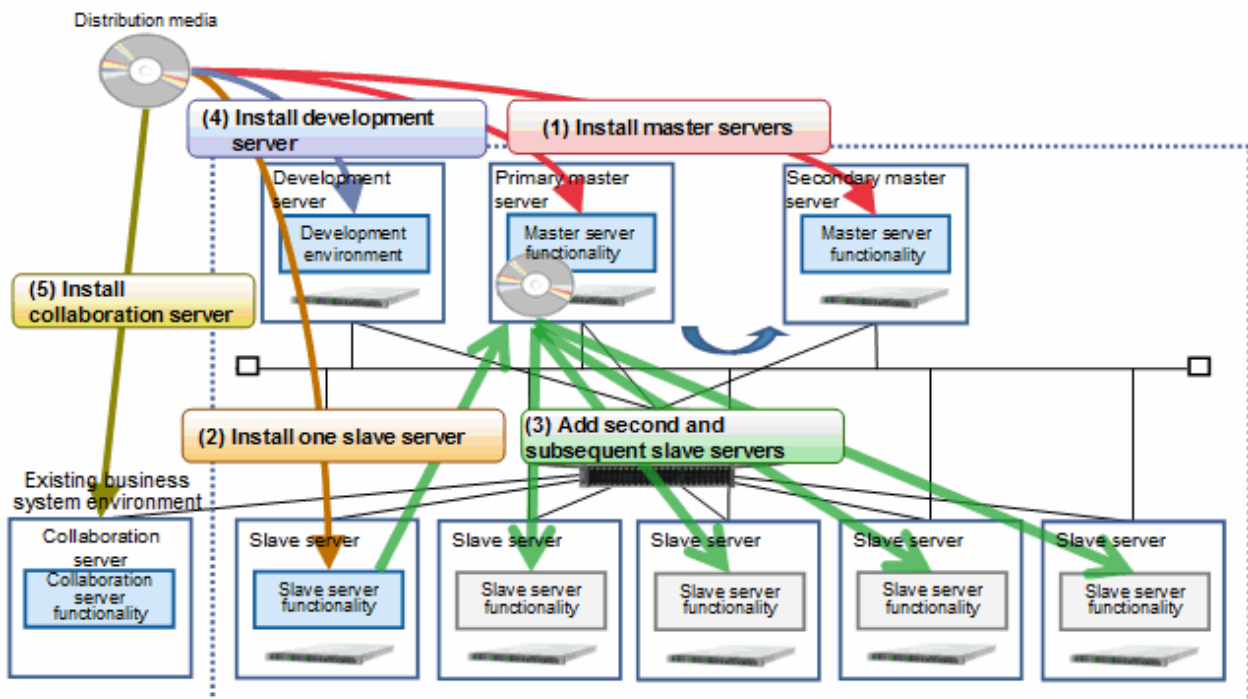
The sequence for building the system is shown below:

1. Installing to master servers
Refer to "6.1 Installing to Master Servers" for details.
2. Installing the first slave server
Refer to "6.2 Installing to a Slave Server" for details.
3. Adding second and subsequent slave servers
Smart Setup can be used to add second and subsequent slave servers (refer to "6.3 Adding Second and Subsequent Slave Servers" for details).

Information

As an example, this manual explains an initial installation of five slave servers, but the same operations can be used to add a greater number of slave servers.

4. Installing a development server
Refer to "6.4 Installing to a Development Server" for details.
5. Installing a collaboration server
Refer to "6.5 Installing to a Collaboration Server" for details.



Note

After installing to the various servers, apply the updates for this product and the updates for software included with this product.

1. Contact a technological staff of Fujitsu and obtain the product shown in the table below included in this product and the patch of the component.
2. Latest manuals (including any update from ones bundled in this product DVD) are available from the following site:

URL:
<http://www.fujitsu.com/global/support/software/manual/> (as of October 2013)

Products to which updates should be applied are as follows:

Product name	Product version	Installation environment				
		Master server		Slave server	Development server	Collaboration server
		Primary	Secondary			
Primesoft Distributed File System for Hadoop	1.0.0	Y	Y	Y	Y	Y
PRIMECLUSTER Clustering Base	4.3A10	Y	Y	-	-	-
PRIMECLUSTER GLS	4.3A10	Y	Y	Y(*1)	-	-
ServerView Resource Orchestrator Virtual Edition	V3.1.0A	Y(*1)	-	Y(*1)	-	-

Y: Required

*1: Not required if installing in a virtual environment.

 **Point**

It is recommended to back up the configuration information for the master servers, development servers, and collaboration servers after completing building of the system with this product and before starting operations. Refer to "14.1 Backup" for information on backup.

6.1 Installing to Master Servers

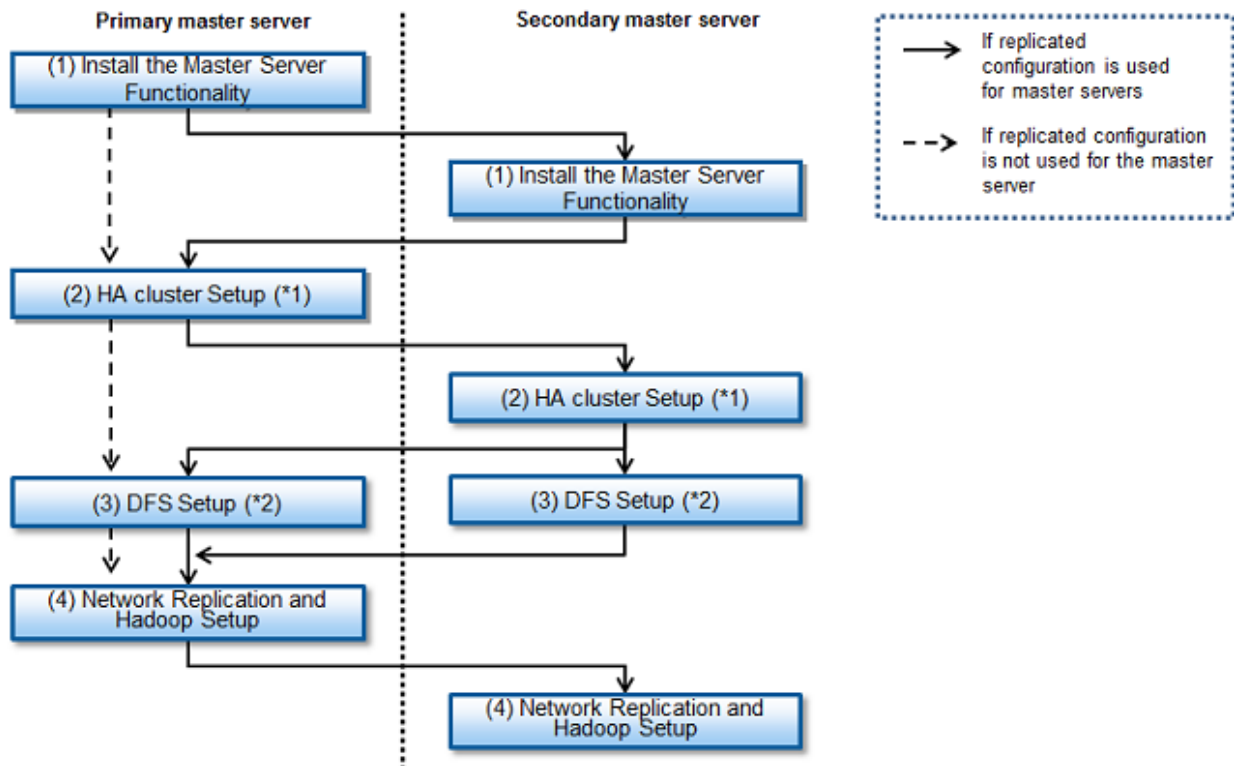
This section describes the installation of the master server functionality on master servers.

1. [Installing the Master Server Functionality](#)
2. [HA Cluster Setup](#)
3. [DFS Setup](#)
4. [Network Replication and Hadoop Setup](#)

The figure below shows the work flow for master server installation.

Perform the setup by first installing the primary master server first, and then the secondary master server, in order to build a master server replicated configuration.

Install and set up using root permissions for all tasks.



*1 Refer to "6.1.2 HA Cluster Setup" for information on the sequence for setting up HA clusters.

*2 Refer to "6.1.3 DFS Setup" for information on the sequence for setting up the DFS.

6.1.1 Installing the Master Server Functionality

Follow the procedures below to install the master server functionality.

1. [Creating bdpp.conf](#)
2. [Installing the Master Server Functionality](#)
3. [Applying Software Updates](#)

6.1.1.1 Creating bdpp.conf

Before installing the master server functionality, set the previously designed values in bdpp.conf.

The settings in bdpp.conf at the primary master server must match the settings at the secondary master server. When creating a master server replicated configuration, before installing the secondary master server, copy bdpp.conf used on the primary master server to the secondary master server.



Example

If bdpp.conf is stored in /home/data/master:

```
# cd /home/data/master <Enter>
# scp -p ./bdpp.conf root@master2:/home/data/master/bdpp.conf <Enter>
```


Note

If a master server replicated configuration is installed in a virtual environment (KVM) and the virtual machines to which the primary master server and the secondary master server are to be installed are on different host machines, you need to set the IP addresses for the two host machines in the `BDPP_KVMHOST_IP` parameter of the configuration file.

See

Refer to "Master server" in "B.1 `bdpp.conf`" for information on `bdpp.conf`.

6.1.1.2 Installing the Master Server Functionality

Perform the setup by first installing the primary master server first, and then the secondary master server, in order to build a master server replicated configuration.

1. Set the "BDPP_CONF_PATH" environment variable.

In the "BDPP_CONF_PATH" environment variable, set the directory where `bdpp.conf` is stored.

Example

If `bdpp.conf` is stored in `/home/data/master`:

```
# export BDPP_CONF_PATH=/home/data/master <Enter>
```

2. Mount the DVD-ROM.

Place the DVD-ROM in the DVD drive, then execute the command below to mount the DVD-ROM. If the DVD-ROM is automatically mounted by the automatic mount daemon (autofs), installer startup will fail because "noexec" is set in the mount options.

```
# mount /dev/hdc dvdRomMountPoint <Enter>
```

3. Start the installer.

Execute `bdpp_inst` to start the installer.

```
# dvdRomMountPoint/bdpp_inst <Enter>
```

4. Select "1. Master Server installation".

```
=====
                Interstage Big Data Parallel Processing Server
                        V1.0.1
                All Rights Reserved, Copyright(c) FUJITSU LIMITED 2012
=====

<< Menu >>

1. Master Server installation
2. Slave Server installation
3. Development Server installation
4. Collaboration Server installation

=====
Please select number to start specified Server installation [?,q]
=> 1 <Enter>
```

Note

- Before starting the installer, the IP address must be set in the Public LAN network interface and the interface must be in the active state.
- If installation fails, refer to the messages output during the installation operation and/or installation log file (/var/tmp/bdpp_install.log) to diagnose the failure. Then, remove the cause of the failure and perform installation again.

6.1.1.3 Applying Software Updates

After installing the master server functionality, apply the updates for this product and the updates for software included with this product. Refer to the note in "[Chapter 6 Installation](#)" for details.

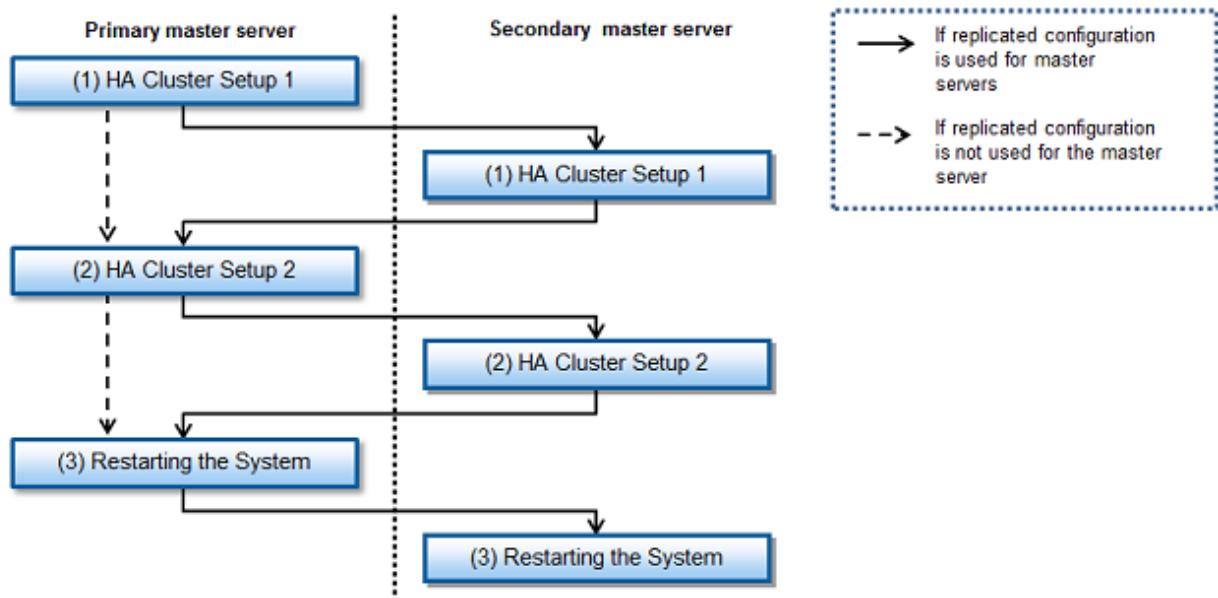
6.1.2 HA Cluster Setup

The HA cluster setup procedure is shown below.

1. [HA Cluster Setup 1](#)
2. [HA Cluster Setup 2](#)
3. [Restarting the System](#)

The sequence of the HA cluster setup procedure is shown below.

Perform the setup by first installing the primary master server and then the secondary master server, in order to build a master server replicated configuration.



Note

If setup fails, refer to the messages output during the setup operation and/or setup log file (/var/opt/FJSVbdpp/log/bdpp_setup.log) to diagnose the failure. Then, remove the cause of the failure and perform setup again.

6.1.2.1 HA Cluster Setup 1

Execute HA cluster setup 1.

This must be done on the primary master server first and then the secondary master server in order to build a master server replicated configuration.

1. Set the cluster interconnect and activate it.
 - a. Edit the `/etc/sysconfig/network-scripts/ifcfg-ethX` file

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

```
ONBOOT=yes
```



ethX is the network interface used for the cluster interconnect.

Specify a numeral at X.

- b. Check the cluster interconnect

Use the command below to check the status of the interface for the interconnect:

```
# ifconfig <relevant interface> <Enter>
```

- c. If the output results of the above command indicate that the relevant interface status is not "UP", execute the following command, then check that the interface is "UP".

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

2. Execute `cluster_setup`.

```
# /opt/FJSVbdpp/setup/cluster_setup <Enter>
```

6.1.2.2 HA Cluster Setup 2

Execute HA cluster setup 2.

This must be done on the primary master server first and then the secondary master server in order to build a master server replicated configuration.

Execute `cluster_setup2`.

```
# /opt/FJSVbdpp/setup/cluster_setup2 <Enter>
```

6.1.2.3 Restarting the System

Restart the system.

This must be done on the primary master server first and then the secondary master server in order to build a master server replicated configuration.

```
# shutdown -r now <Enter>
```

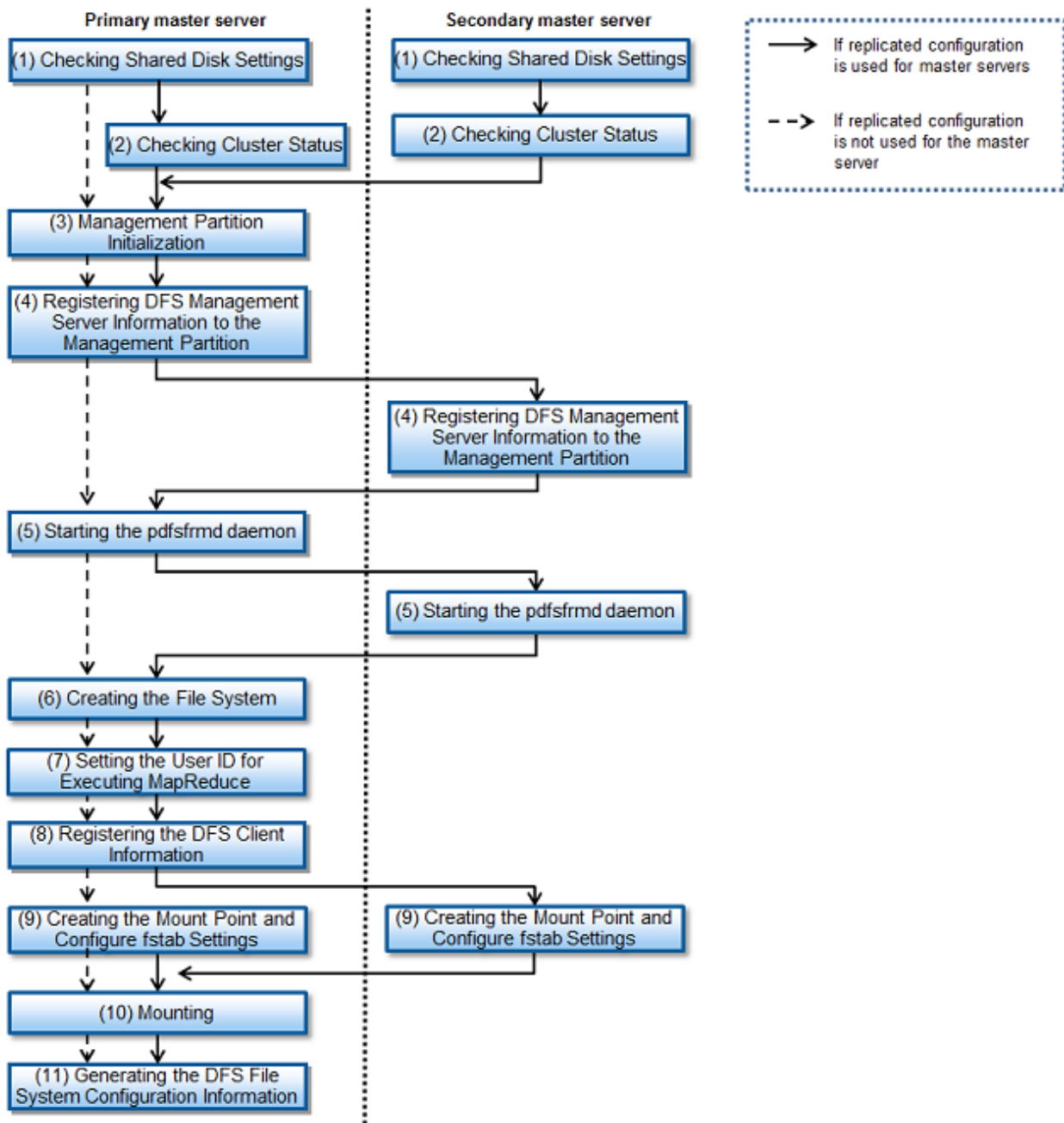
6.1.3 DFS Setup

The DFS setup sequence is as follows:

1. [Checking Shared Disk Settings](#)
2. [Checking Cluster Status](#)
3. [Management Partition Initialization](#)
4. [Registering DFS Management Server Information to the Management Partition](#)
5. [Starting the pdfsfrmd daemon](#)
6. [Creating the File System](#)
7. [Setting the User ID for Executing MapReduce](#)
8. [Registering the DFS Client Information](#)
9. [Creating the Mount Point and Configure fstab Settings](#)
10. [Mounting](#)
11. [Generating the DFS File System Configuration Information](#)

The DFS setup sequence is as follows:

Perform the setup by first installing the primary master server and then the secondary master server, in order to build a master server replicated configuration.



Point

The table below provides examples of information such as the device name specified in setup with the following file system configuration. Refer to the actual information for your system when you are setting up.

- Management partition :/dev/disk/by-id/scsi-1FUJITSU_300000370105
- Representative partition :/dev/disk/by-id/scsi-1FUJITSU_300000370106
- File data partition :/dev/disk/by-id/scsi-1FUJITSU_300000370107
/dev/disk/by-id/scsi-1FUJITSU_300000370108
- File system ID :1
- Logical file system name :pdfs1
- Master server : master1 (primary), master2 (secondary)
- Slave server :slave1, slave2, slave3, slave4, slave5

- Development server :develop
- Collaboration server :collaborate

6.1.3.1 Checking Shared Disk Settings

A by-id name generated by the udev function is used for shared disk device names.

Use either the udevinfo or udevadm command to ascertain the by-id name from the conventional compatible device name.

An example of checking the by-id name is shown below.

Example

Determining device names from compatible device names using by-id name

- Under Red Hat(R) Enterprise Linux(R) 5:

```
# udevinfo -q symlink -n /dev/sdb <Enter>
disk/by-id/scsi-1FUJITSU_300000370105
# udevinfo -q symlink -n /dev/sdc <Enter>
disk/by-id/scsi-1FUJITSU_300000370106
# udevinfo -q symlink -n /dev/sdd <Enter>
disk/by-id/scsi-1FUJITSU_300000370107
# udevinfo -q symlink -n /dev/sde <Enter>
disk/by-id/scsi-1FUJITSU_300000370108
```

- Under Red Hat(R) Enterprise Linux(R) 6:

```
# udevadm info -q symlink -n /dev/sdb <Enter>
block/8:48 disk/by-id/scsi-1FUJITSU_300000370105
# udevadm info -q symlink -n /dev/sdc <Enter>
block/8:48 disk/by-id/scsi-1FUJITSU_300000370106
# udevadm info -q symlink -n /dev/sdd <Enter>
block/8:48 disk/by-id/scsi-1FUJITSU_300000370107
# udevadm info -q symlink -n /dev/sde <Enter>
block/8:48 disk/by-id/scsi-1FUJITSU_300000370108
```

Note

- In order to use the by-id name checked using the udevinfo or udevadm command, "/dev/" must be added at the start of the name.
- If shared disk partition information is changed using the fdisk, parted, or similar command, refer to "4.2.4 Partition Information of Shared Disk Device Modified with fdisk(8) is not Reflected" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" and refresh the partition information at all servers.

See

Refer to the online manual pages for details of the udevinfo and udevadm commands.

Point

- The by-id name is a device name generated from the unique identification information set in the hard disk. Use of the by-id names enables each server to always use the same device name to access a specific disk.

- The DFS management partition can operate in either Logical Unit (physical) units or disk partition (logical) units. If volume copy using ETERNUS SF AdvancedCopy Manager is performed, take into account the device units supported by ETERNUS SF AdvancedCopy Manager.
Refer to the "ETERNUS SF AdvancedCopy Manager Operation Guide" for ETERNUS SF AdvancedCopy Manager details.

6.1.3.2 Checking Cluster Status

When creating a master server replicated configuration, before creating a management partition, be sure to check that a cluster partition has not occurred.

Perform this action on both the primary master server and the secondary master server.

1. Execute the `cftool(1M)` command.

Confirm that the displayed state (in the State column) is the same for the two master servers.

```
# cftool -n <Enter>
Node      Number State      Os      Cpu
master1  1      UP        Linux   EM64T
master2  2      UP        Linux   EM64T
```



Refer to the online help of `cftool` for details of the `cftool(1M)` command.

2. If the display result is not identical on the two master servers, cluster partitioning has occurred.
Cancel the cluster partitioning if this is the case.



Refer to the "4.2.1 Corrective Action when the `pdfsfrmd` Daemon Does Not Start" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for information on canceling cluster partitions.

6.1.3.3 Management Partition Initialization

Initialize the management partition.

Perform this action on the primary master server.

Specify the `-c` option and the path name of the management partition in the `pdfssetup` command and execute.

```
# pdfssetup -c /dev/disk/by-id/scsi-1FUJITSU_300000370105 <Enter>
```



Refer to `pdfssetup` under "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for details of the `pdfssetup` command.

6.1.3.4 Registering DFS Management Server Information to the Management Partition

Register master server information to the management partition.

This must be done on the primary master server first and then the secondary master server in order to build a master server replicated configuration.

1. Register master server information to the management partition.

Specify the -a option in the pdfssetup command and execute.

```
# pdfssetup -a /dev/disk/by-id/scsi-1FUJITSU_300000370105 <Enter>
```

2. Check the registered master server information.

The registered information can be checked by executing the pdfssetup command without any options specified.

```
# pdfssetup <Enter>
HOSTID      CIPNAME      MP_PATH
80380000    master1RMS   yes
80380001    master2RMS   yes
```

The management partition path name that has been set can be checked by executing the pdfssetup command with the -p option specified.

```
# pdfssetup -p <Enter>
/dev/disk/by-id/scsi-1FUJITSU_300000370105
```

6.1.3.5 Starting the pdfsfrmd daemon

Start the pdfsfrmd daemon in order to start operations.

This must be done on the primary master server first and then the secondary master server in order to build a master server replicated configuration.

```
# pdfsfrmstart <Enter>
```



See

Refer to pdfsfrmstart under "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for details of the pdfsfrmstart command.

6.1.3.6 Creating the File System

Create the DFS in the partitions to be used.

Perform this action on the primary master server.

1. Create the file system.

Specify the following options and the representative partition in the pdfsmkfs command and execute.

- dataopt option

Specify y to separate the file data area from the representative partition.

- blocksz option

Specify the data block size. 8388608 (8MB) is recommended.

- data option

Specify the path names of the file data partitions separated with commas.

- node option

Specify the host name of the master server (the host name corresponding to the NIC that connects public LAN).

Separate the primary master server and the secondary master server with a comma when building a master server replicated configuration. This option can be omitted if a master server replicated configuration is not required.

Note

Specify the node option in the final option.

When making replicated configuration for the master server:

```
# pdfsmkfs -o dataopt=y,blocksz=8388608,data=/dev/disk/by-id/  
scsi-1FUJITSU_300000370107,data=/dev/disk/by-id/  
scsi-1FUJITSU_300000370108,node=master1,master2 /dev/disk/by-id/scsi-1FUJITSU_300000370106  
<Enter>
```

See

- Refer to "3.3.2 File System Configuration Design" for information on data block size.
- Refer to "pdfsmkfs" in the "Appendix A Command Reference" of the "Primesoft Distributed File System for Hadoop V1 User's Guide" for information on the pdfsmkfs command.

2. Confirm the file system information created.

```
# pdfsinfo /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>  
/dev/disk/by-id/scsi-1FUJITSU_300000370106:  
FSID special size Type mount  
1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 25418 META -----  
1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 5120 LOG -----  
1 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 7341778 DATA -----  
1 /dev/disk/by-id/scsi-1FUJITSU_300000370108 (896) 6578704 DATA -----
```

6.1.3.7 Setting the User ID for Executing MapReduce

Users must be set to the DFS in order for mapred users to execute Hadoop JobTracker and TaskTracker.

This section describes the procedure for setting mapred users to the DFS.

Perform this action on the primary master server.

1. Set the user ID.

Use the pdfsadm command to set the user ID for executing MapReduce in the MAPRED variable.

```
# pdfsadm -o MAPRED=mapred /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
```

2. Check that the user ID has been set.

```
# pdfsinfo -e /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>  
/dev/disk/by-id/scsi-1FUJITSU_300000370106:  
MAPRED=mapred
```

See

Refer to "pdfsadm" under "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for the method for deleting a set MAPRED variable and other pdfsadm command details.

6.1.3.8 Registering the DFS Client Information

Register the information for the slave servers, development servers and collaboration servers (DFS client information) in the connection authorization list.

Create and register the connection authorization list on the primary master server and then distribute to the secondary master server in order to build a master server replicated configuration.

Information

The DFS manages the DFS clients that can connect to the master server (MDS).

Create a connection authorization list on the master server and register the host names of the server that will connect with the DFS.

1. Check the file system ID.

Check the target file system ID in the file system information recorded in the management partition.

```
# pdfsinfo /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
/dev/disk/by-id/scsi-1FUJITSU_300000370106:
FSID special size Type mount
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 25418 META -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 5120 LOG -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 7341778 DATA -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370108 (896) 6578704 DATA -----
```

2. Create a file listing approved connections.

```
# cd /etc/pdfs <Enter>
# cp ./server.conf.sample server.conf.1 <Enter>
```

Note

Place the connection authorization list file under /etc/pdfs at the master server.

In the connection authorization list file name, change only the file system ID part, not the other part (server.conf.).

3. Register the host names (the host names corresponding to the NICs connected to the public LAN) of servers (slave servers, development servers, and collaboration servers) permitted to connect in the connection authorization list file.

Use the following format to enter the names:

```
CLIENT hostNameToBePermittedToConnect
```

Example

When permitting connection of slave1, slave2, slave3, slave4, slave5, develop, and collaborate:

```
# cat /etc/pdfs/server.conf.1 <Enter>
#
# Copyright (c) 2012 FUJITSU LIMITED. All rights reserved.
#
# /etc/pdfs/server.conf.<FSID>
#
# List of client hostnames of a file system.
#
# Notes:
# Do not describe hostnames of management servers.
#
# example:
#CLIENT nodeac1
#CLIENT nodeac2
#CLIENT nodeac3
#CLIENT nodeac4
#CLIENT nodeac5
CLIENT develop <-- development environment server you are adding
CLIENT collaborate <-- collaboration server you are adding
```

```
CLIENT slave1 <-- slave server you are adding
CLIENT slave2 <-- slave server you are adding
CLIENT slave3 <-- slave server you are adding
CLIENT slave4 <-- slave server you are adding
CLIENT slave5 <-- slave server you are adding
```

4. Check the content of the connection authorization list file.

Mount is not possible at the master server if there is an error in the connection authorization list. Therefore, check the following:

- The total number of Master servers, slave servers, development servers and collaboration servers does not exceed the maximum number of shared servers.
- The specified slave Server, development Server and collaboration Server hosts can be referenced correctly via the network.
- The slave server, development server and collaboration server specifications are not duplicated.

5. Distribute the updated connection authorization list file to the master servers. (only when creating a master server replicated configuration)

```
# cd /etc/pdfs <Enter>
# scp -p ./server.conf.1 root@master2:/etc/pdfs/server.conf.1 <Enter>
```

6.1.3.9 Creating the Mount Point and Configure fstab Settings

Create the mount point and add the DFS entry to /etc/fstab.

This must be done on both the primary master server and the secondary master server in order to build a master server replicated configuration.

Creating the mount point

Create the mount point for mounting the disk partitions on the storage system used as the DFS.

The mount point created must be the same as the value specified in the BDPP_PDFS_MOUNTPOINT parameter in bdpp.conf.



Example

Create the mount point "pdfs" under "/mnt".

```
# mkdir /mnt/pdfs <Enter>
```

fstab settings

Add the DFS entry to /etc/fstab.

The parameters specified in the fields for the added entry are as follows:

- Field 1 (fs_spec)
Specify the representative partition of the DFS to be mounted.
- Field 2 (fs_file)
Specify the mount point you created above.
- Field 3 (fs_vfstype)
Specify the pdfs.
- Field 4 (fs_mntops)
Specify the mount options to be used when mount is performed.
Ensure that the noauto option is specified.

Determine other option specifications as shown below.

Item to be checked	Option to be specified
If either of the following applies: <ul style="list-style-type: none"> - The file final reference time is not updated - The DFS is shared by five or more servers 	noatime
If not performing mount at DFS management server startup	noatrc
If mounting DFS as read only	ro

See

Refer to pdfsmount under "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for mount option details.

- Field 5 (fs_freq)
Specify 0.
- Field 6 (fs_passno)
Specify 0.

Example

This example shows the mount point "/mnt/pdfsf" and DFS representative partitions defined at "/etc/fstab".

LABEL=/	/	ext3	defaults	1	1
LABEL=/boot	/boot	ext3	defaults	1	2
tmpfs	/dev/shm	tmpfs	defaults	0	0
devpts	/dev/pts	devpts	gid=5,mode=620	0	0
sysfs	/sys	sysfs	defaults	0	0
proc	/proc	proc	defaults	0	0
LABEL=SWAP-sda3	swap	swap	defaults	0	0
<hr/>					
/dev/disk/by-id/scsi-1FUJITSU_300000370106	/mnt/pdfsf	pdfsf	noauto,noatime	0	0

Note

The entries in /etc/fstab must be the same on both the primary master server and the secondary master server in order to build a master server replicated configuration.

6.1.3.10 Mounting

Mount the DFS file system.

Perform this action only on the primary master server.

```
# pdfsmntgl /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
```

Note

- The `pdfsmntgl` command mounts the DFS on all DFS management servers. For this reason, it is not necessary to mount on the secondary master server.
- Ensure that the DFS mount sequence is: master servers, slave servers, development servers and collaboration servers. If the slave servers, development servers or collaboration servers are mounted first, then mount will fail, because the master server (MDS) does not exist.

6.1.3.11 Generating the DFS File System Configuration Information

Generate the DFS configuration information file on the master server.

Perform this action on the primary master server.

1. Check the file system ID.

Check the target file system ID in the file system information recorded in the management partition.

```
# pdfsinfo /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
/dev/disk/by-id/scsi-1FUJITSU_300000370106:
FSID special size Type mount
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 25418 META -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 5120 LOG -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 7341778 DATA -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370108 (896) 6578704 DATA -----
```

2. Generate the DFS configuration file with the `pdfsmkconf` command.

```
# pdfsmkconf <Enter>
```

3. Convert the generated configuration information file name to a logical file system name from the file system ID.

```
# cd pdfsmkconf_out <Enter>
# mv ./client.conf.1 client.conf.pdfsl <Enter>
```

Note

The DFS configuration information file is created as `pdfsmkconf_out/client.conf.fsid` in the directory where the `pdfsmkconf` command is executed.

Other than the file system ID part (`client.conf`), do not change the name of the DFS configuration file.

4. Confirm that the user ID for executing MapReduce that was set in "6.1.3.7 Setting the User ID for Executing MapReduce" has been set properly in the DFS configuration information file.

```
# cat ./client.conf.pdfsl <Enter>
FSID 1
MDS master1 29000
MDS master2 29000
DEV /dev/disk/by-id/scsi-1FUJITSU_300000370107 0 7341778
DEV /dev/disk/by-id/scsi-1FUJITSU_300000370108 0 6578704
MAPRED mapred
```

Point

The generated DFS configuration information files must be placed in `/etc/pdfs` on the DFS clients (slave servers, development servers, and collaboration servers).

When setting up the slave servers, development servers, and collaboration servers explained later, distribute the DFS configuration information file on the master server to each of the servers.



Refer to the pdfsmkconf command under "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for details of pdfsmkconf.

6.1.4 Network Replication and Hadoop Setup

The procedure for duplicating the network and setting up Hadoop is shown below.

1. [Creating the Slave Server Definition File 'slaves'](#)
2. [Network Replication and Hadoop Setup](#)

6.1.4.1 Creating the Slave Server Definition File 'slaves'

Create a slave server definition file (/etc/opt/FJSVbdpp/conf/slaves) on the primary master server to define the slave servers to be connected.

Distribute the slave server definition files defined on the primary master server to the secondary master server in order to build a master server replicated configuration.

1. Create a slave server configuration file.



Defining five slave servers; slave1, slave2, slave3, slave4, and slave5:

```
# cat /etc/opt/FJSVbdpp/conf/slaves <Enter>
slave1,slave2,slave3,slave4,slave5
```

2. Distribute the slave server configuration file to the secondary master server. (**Only when creating a master server replicated configuration**)

```
# scp /etc/opt/FJSVbdpp/conf/slaves root@master2:/etc/opt/FJSVbdpp/conf/slaves <Enter>
```



The settings in the slave server configuration file must be the same on the primary master server, the secondary master server, the slave servers, and the development servers.

When setting up the slave servers and development servers explained later, distribute the slave server configuration file on the master server to each of the servers.



Refer to "B.2 slaves" for details of the slave server definition file 'slaves'.

6.1.4.2 Network Replication and Hadoop Setup

Replicate the network and setup Hadoop.

This must be done on the primary master server first and then the secondary master server in order to build a master server replicated configuration.

Execute `bdpp_setup`.

```
# /opt/FJSVbdpp/setup/bdpp_setup <Enter>
```

Note

If setup fails, refer to the messages output during the setup operation and/or setup log file (`/var/opt/FJSVbdpp/log/bdpp_setup.log`) to diagnose the failure. Then, remove the cause of the failure and perform setup again.

Information

Hadoop setup sets Hadoop and DFS configurations as default settings of this product. Refer to "[Appendix C Hadoop Configuration Parameters](#)" for information on how to tune the configuration settings.

6.2 Installing to a Slave Server

This section describes the installation of the slave server functionality to a slave server.

The procedures here build a new slave server. This section describes the procedure for installing and setting up the slave server functionality on a system where only the operating system has been installed.

Install and set up using root permissions for all tasks.

1. [Installing the Slave Server Functionality](#)
2. [DFS Setup](#)
3. [Network Replication and Hadoop Setup](#)
4. [Checking the Connection to the Slave Server](#)

6.2.1 Installing the Slave Server Functionality

The procedure for installing the slave server functionality is shown below.

1. [Creating `bdpp.conf`](#)
2. [Installing the Slave Server Functionality](#)
3. [Applying Software Updates](#)

6.2.1.1 Creating `bdpp.conf`

Before installing the slave server functionality, set the previously designed values in `bdpp.conf`.

See

Refer to "Slave server" under "[B.1 `bdpp.conf`](#)" for information on `bdpp.conf`.

6.2.1.2 Installing the Slave Server Functionality

Follow the procedures below to install the slave server functionality.

1. Set the "BDPP_CONF_PATH" environment variable.
In the "BDPP_CONF_PATH" environment variable, set the directory where bdpp.conf is stored.

Example

If bdpp.conf is stored in /home/data/slaves:

```
# export BDPP_CONF_PATH=/home/data/slave <Enter>
```

2. Mount the DVD-ROM.
Place the DVD-ROM in the DVD drive, then execute the command below to mount the DVD-ROM. If the DVD-ROM is automatically mounted by the automatic mount daemon (autofs), installer startup will fail because "noexec" is set in the mount options.

```
# mount /dev/hdc dvdRomMountPoint <Enter>
```

3. Start the installer.
Execute bdpp_inst to start the installer.

```
# dvdRomMountPoint/bdpp_inst <Enter>
```

4. Select "2. Slave Server installation".

```
=====
                Interstage Big Data Parallel Processing Server
                        V1.0.1
                All Rights Reserved, Copyright(c) FUJITSU LIMITED 2012
=====

<< Menu >>

1. Master Server installation
2. Slave Server installation
3. Development Server installation
4. Collaboration Server installation

=====
Please select number to start specified Server installation [?,q]
=> 2 <Enter>
```

5. After installation is completed, restart the system.

```
# shutdown -r now <Enter>
```

Note

If installation fails, refer to the messages output during the installation operation and/or installation log file (/var/tmp/bdpp_install.log) to diagnose the failure. Then, remove the cause of the failure and perform installation again.

6.2.1.3 Applying Software Updates

After installing the slave server functionality, apply the updates for this product and the updates for software included with this product.

Refer to the note in "[Chapter 6 Installation](#)" for details.

6.2.2 DFS Setup

The DFS setup sequence is as follows:

1. [Create the Mount Point and Configure fstab Settings](#)
2. [Mounting](#)

6.2.2.1 Creating the Mount Point and Configuring fstab Settings

Create the mount point and add the DFS entry to `/etc/fstab`.

Creating the mount point

Create the mount point for mounting the disk partitions on the storage system used as the DFS.

Make the mount point the same as that specified at the master server.



Example

.....

Create the mount point "pdfs" under "/mnt".

```
# mkdir /mnt/pdf <Enter>
```

fstab settings

Add the DFS entry to `/etc/fstab`.

The parameters specified in the fields for the added entry are as follows:

- Field 1 (`fs_spec`)
Specify the logical file system name of the DFS to be mounted.



Information

.....

The logical file system name is used to identify the DFS file system. Use the name defined when generating the configuration information of the file system on the master server.

The configuration information file is used to manage the file system structure at slave servers, development servers and collaboration servers. Use a logical file system name to identify the DFS configuration information file.

.....

- Field 2 (`fs_file`)
Specify the mount point you created above.
- Field 3 (`fs_vfstype`)
Specify the pdfs.
- Field 4 (`fs_mntops`)
Specify the mount options to be used when mount is performed.

Ensure that the `_netdev` option is specified.

Determine other option specifications as shown below.

Item to be checked	Option to be specified
If not updating the time when a file was last referenced or when a DFS is shared by five or more servers	noatime
If not mounting when starting the DFS client	noauto
If mounting DFS as read only	ro

See

Refer to the mount command page in the online manual for details of the noauto option.

Refer to "mount.pdfs" under "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for details of other mount options.

- Field 5 (fs_freq)
Specify 0.
- Field 6 (fs_passno)
Specify 0.

Example

This example shows the mount point "/mnt/pdfsf" and the logical file system name "pdfsf1" defined at "/etc/fstab".

LABEL=/	/	ext3	defaults	1 1
LABEL=/home	/home	ext3	defaults	1 2
LABEL=/boot	/boot	ext3	defaults	1 2
tmpfs	/dev/shm	tmpfs	defaults	0 0
devpts	/dev/pts	devpts	gid=5,mode=620	0 0
sysfs	/sys	sysfs	defaults	0 0
proc	/proc	proc	defaults	0 0
LABEL=SWAP-sda3	swap	swap	defaults	0 0
pdfsf1	/mnt/pdfsf	pdfsf	_netdev	0 0

6.2.2.2 Mounting

Distribute the DFS configuration information file generated on the master server to the slave server to mount the DFS file system.

1. Copy the DFS configuration information file generated in "6.1.3.11 Generating the DFS File System Configuration Information" for the master server to /etc/pdfsf.

```
# scp -p root@master1:dfsConfigurationFileDirectory/client.conf.pdfsf1 /etc/pdfsf <Enter>
```

2. Mount the DFS file system.

Example

Mount the DFS file system at the slave server.

```
# mount pdfsf1 <Enter>
```

Note

- The DFS configuration information file must be distributed to the slave server before the DFS file system is mounted.
- Ensure that the DFS mount sequence is: master servers, slave servers, development servers and collaboration servers. If the slave servers, development servers or collaboration servers are mounted first, then mount will fail, because the master server (MDS) does not exist.

See

Refer to the "mount.pdfs" under "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for details of mounting.

6.2.3 Network Replication and Hadoop Setup

The network replication and Hadoop setup procedure is shown below.

1. Copy the master server 'slaves' file, edited as described in "[6.1.4.1 Creating the Slave Server Definition File 'slaves'](#)", and place the copy under "/etc/opt/FJSVbdpp/conf".

```
# scp -p root@master1:/etc/opt/FJSVbdpp/conf/slaves /etc/opt/FJSVbdpp/conf <Enter>
```

2. Execute the network replication and Hadoop setup.

Execute bdpp_setup.

```
# /opt/FJSVbdpp/setup/bdpp_setup <Enter>
```

Note

- If the BDPP_GLS_SERVER_CONNECT2 parameter value is omitted from bdpp.conf, the setup for duplicating the network is not performed, and only Hadoop is set up.
- If setup fails, refer to the messages output during the setup operation and/or setup log file (/var/opt/FJSVbdpp/log/bdpp_setup.log) to diagnose the failure. Then, remove the cause of the failure and perform setup again.

Information

Hadoop setup sets Hadoop and DFS configurations as default settings of this product. Refer to "[Appendix C Hadoop Configuration Parameters](#)" for information on how to tune the configuration settings.

6.2.4 Checking the Connection to the Slave Server

Before building the second and subsequent slave servers, check that the first already installed slave server can connect correctly to the master server.

Check this on the primary master server. Also make sure to execute with root permissions.

1. Backup the slave server definition file (/etc/opt/FJSVbdpp/conf/slaves) so that it can be edited to check the connection of the first slave server.

```
# cp /etc/opt/FJSVbdpp/conf/slaves /etc/opt/FJSVbdpp/conf/slaves.bak <Enter>
```

2. Edit the slave server definition file, and leaving only the slave server that is to be checked for connection, delete the other slave servers.

Before editing:

```
# cat /etc/opt/FJSVbdpp/conf/slaves <Enter>
slave1,slave2,slave3,slave4,slave5
```

After editing:

```
# cat /etc/opt/FJSVbdpp/conf/slaves <Enter>
slave1
```

3. Execute `bdpp_changeslaves` command to effect the changes that were made to the slave server definition file.

```
# /opt/FJSVbdpp/bin/bdpp_changeslaves <Enter>
```

4. Execute the `bdpp_start` command to start the Hadoop of this product.

```
# /opt/FJSVbdpp/bin/bdpp_start <Enter>
```

5. Execute the `bdpp_stat` command to display the status of the Hadoop of this product and check whether or not TaskTracker is running on the slave server.

```
# /opt/FJSVbdpp/bin/bdpp_stat -all <Enter>
cluster mapred 30595 jobtracker
slave1 mapred 11381 tasktracker <-- Check that TaskTracker is started at slave1.
bdpp: INFO : 003: bdpp Hadoop JobTracker is alive.
```

6. Run the Hadoop job to check operation.

Use the sample that came with Hadoop (teragen, terasort, etc.) to check the operation of the Hadoop job.

7. Execute the `bdpp_stop` command to stop the Hadoop of this product.

```
# /opt/FJSVbdpp/bin/bdpp_stop <Enter>
```

8. After checking the connection between the master server and the first slave server, restore the backup of the slave server definition file.

```
# rm -fr /etc/opt/FJSVbdpp/conf/slaves <Enter>
# cp /etc/opt/FJSVbdpp/conf/slaves.bak /etc/opt/FJSVbdpp/conf/slaves <Enter>
```

9. Execute `bdpp_changeslaves` command to effect the changes that were made to the slave server definition file again.

```
# /opt/FJSVbdpp/bin/bdpp_changeslaves <Enter>
```



See

- Refer to "[A.4 bdpp_changeslaves](#)" for information on the `bdpp_changeslaves` command.
- Refer to "[A.14 bdpp_start](#)" for information on the `bdpp_start` command.
- Refer to "[A.15 bdpp_stat](#)" for information on the `bdpp_stat` command.
- Refer to "[Chapter 10 Executing and Stopping Jobs](#)" for information on executing Hadoop jobs.
- Refer to "[A.16 bdpp_stop](#)" for information on the `bdpp_stop` command.

6.3 Adding Second and Subsequent Slave Servers

This section describes how to add slave servers using the cloning functionality.

- [Adding a Slave Server to a Physical Environment](#)
- [Adding a Slave Server to a Virtual Environment](#)

This is explained using the example shown below.

- Slave server already installed :slave1
- Slave server being added :slave2, slave3, slave4, slave5

Note

If Hadoop is working, execute the `bdpp_stop` command on the master server to stop the Hadoop in advance.

```
# /opt/FJSVbdpp/bin/bdpp_stop <Enter>
```

See

This section describes the procedure for adding a slave server during initial installation.

Refer to "[12.1 Adding Slave Servers](#)" when further adding slave servers (scaling out) after starting operations.

6.3.1 Adding a Slave Server to a Physical Environment

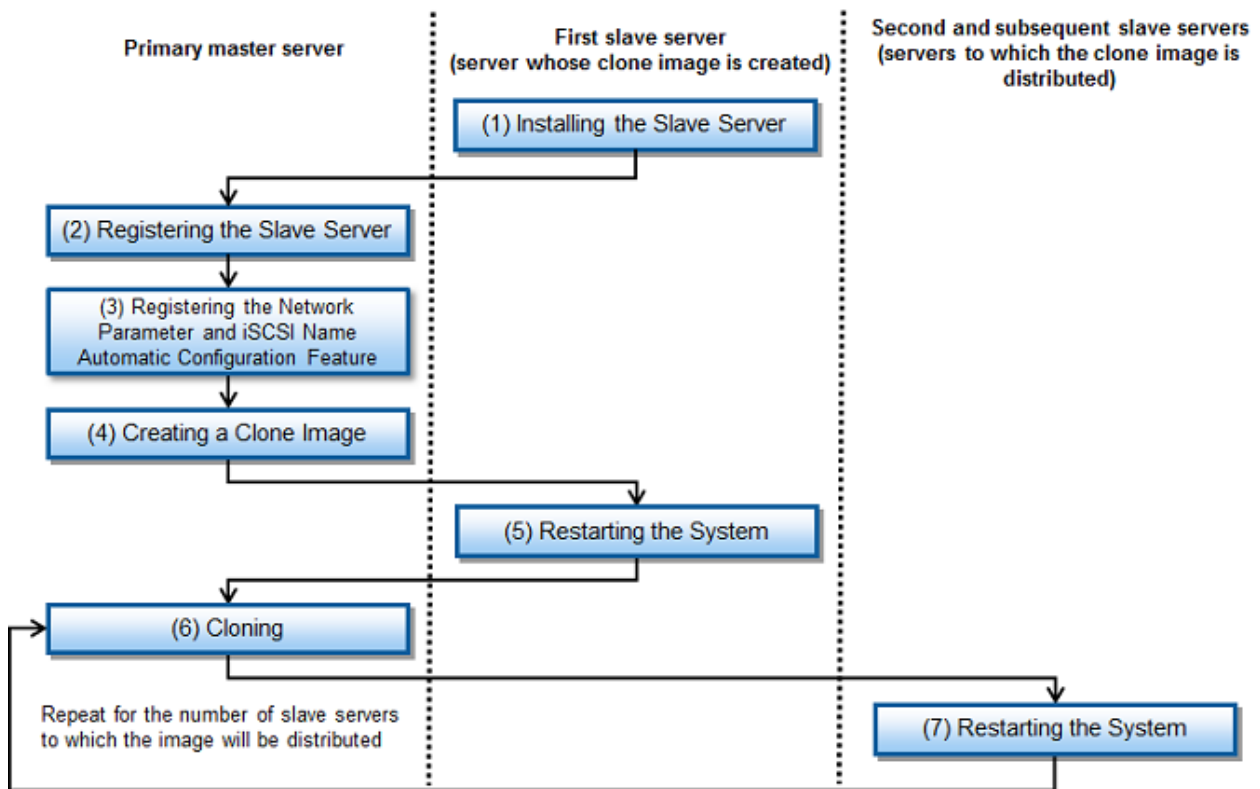
The procedure for adding a slave server to a physical environment is shown below.

1. [Installing the Slave Server](#)
2. [Registering the Slave Server](#)
3. [Registering the Network Parameter and iSCSI Name Automatic Configuration Feature](#)
4. [Creating a Clone Image](#)
5. [Restarting the System](#)
6. [Cloning](#)
7. [Restarting the System](#)

The procedure for installing a slave server is shown below.

Create the clone image of the first slave server (one that is already installed) and clone as many times as the number of slave servers to be installed.

All installation tasks must be performed with root permissions.



Note

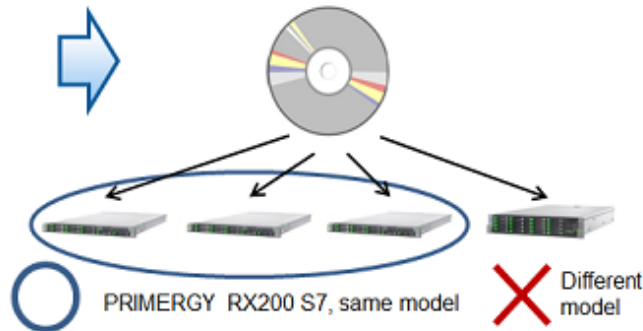
The server for which the clone image is to be created and the server to be cloned must meet the following conditions. If the conditions differ, it is necessary to create groups for the clone images where the conditions are the same.

- The model name must be identical.
- The hardware conditions must be the same (the same option cards and expansion boards and their locations).
- The BIOS settings must be the same, according to the settings described in "[5.1 Server BIOS Settings](#)".
- The LAN and SAN connections use the same redundancy method, have the same number of redundant paths, and have access to the same network devices and storage devices.
LAN switches and FC Switches that are cascade connected will be treated as one device.

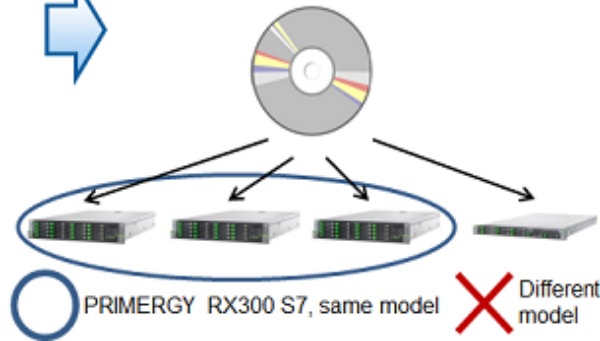
Slave server targeted for clone image creation



PRIMERGY RX200 S7 clone image



PRIMERGY RX300 S7 clone image



6.3.1.1 Installing the Slave Server

Refer to "6.2 Installing to a Slave Server" for information on how to install the first additional slave server (the slave server whose clone image is created).

6.3.1.2 Registering the Slave Server

To perform the cloning tasks, register all the slave servers that are expected to be installed in the master server.

1. Create a cloning server definition file (clone.conf).

In the cloning server definition file, specify the slave server whose clone image is to be created and all the slave servers to which the clone image is to be distributed for installation.

Example

Example cloning server definition file (clone.conf):

```
RCXCSV, V3.4
[Server]
operation, chassis_name, slot_no, server_name, ip_address, mac_address, second_mac_address, snmp_community_name, ipmi_ip_address, ipmi_user_name, ipmi_passwd, ipmi_passwd_enc, admin_lan1_nic_number, admin_lan2_nic_number, admin_lan_nic_redundancy
new, , slave1, 10.10.10.13, A0:A0:A0:A0:A0:A1, , public, 192.168.1.13, admin, admin, plain, , OFF <---
Server to be registered
new, , slave2, 10.10.10.14, A0:A0:A0:A0:A0:A2, , public, 192.168.1.14, admin, admin, plain, , OFF <---
Server to be registered
new, , slave3, 10.10.10.15, A0:A0:A0:A0:A0:A3, , public, 192.168.1.15, admin, admin, plain, , OFF <---
```

```

Server to be registered
new,,,slave4, 10.10.10.16,A0:A0:A0:A0:A0:A4,,public,192.168.1.16,admin,admin,plain,,,OFF <--
Server to be registered
new,,,slave5, 10.10.10.17,A0:A0:A0:A0:A0:A5,,public,192.168.1.17,admin,admin,plain,,,OFF <--
Server to be registered
[ServerAgent]
operation,server_name
new,slave1 <-- Server whose clone image is to be created

```

2. Execute the `bdpp_addserver` command to register the slave server.

```
# /opt/FJSVbdpp/bin/bdpp_addserver clone.confFileDirectory/clone.conf <Enter>
```

3. Use the `bdpp_listserver` command to check that the slave server has been registered successfully.

```
# /opt/FJSVbdpp/bin/bdpp_listserver <Enter>
```



Example

Example of checking slave server registration:

```

# /opt/FJSVbdpp/bin/bdpp_listserver <Enter>
PHYSICAL_SERVER  SERVER                ADMIN_IP             STATUS  MAINTENANCE
-----
slave1           slave1                192.168.1.21        normal  OFF
slave2           -                     192.168.1.22        stop    -
slave3           -                     192.168.1.23        stop    -
slave4           -                     192.168.1.24        stop    -
slave5           -                     192.168.1.25        stop    -

```



See

- Refer to "[B.3 clone.conf](#)" for information on the cloning server definition file.
- Refer to "[A.1 bdpp_addserver](#)" for information on the `bdpp_addserver` command.
- Refer to "[A.9 bdpp_listserver](#)" for information on the `bdpp_listserver` command.

6.3.1.3 Registering the Network Parameter and iSCSI Name Automatic Configuration Feature

This section explains how to register the network parameter and iSCSI initiator name automatic configuration feature.

By setting the network parameters for the public LAN and iSCSI-LAN and the iSCSI initiator for the slave server beforehand and then creating the clone image, it is possible to automatically set the network and iSCSI initiator names for all slave servers when distributing the clone images.

1. Create the network parameter automatic configuration feature definition file (`FJSVrcx.conf` and `ipaddr.conf`).
 - Specify the admin LAN information for the slave server from which the clone image is to be created in the `FJSVrcx.conf` definition file.
 - Specify the public LAN and iSCSI-LAN information for all slave servers in the `ipaddr.conf` definition file.
2. Create the `initiator.conf` iSCSI initiator configuration file.
 - Specify the iSCSI initiator names corresponding to all slave servers that will receive images.

- Specify the IP address of the admin LAN of the slave server from which the clone image is to be created, the FJSVrcx.conf, ipaddr.conf, and initiator.conf definition files in the bdpp_prepareserver command.

```
# /opt/FJSVbdpp/bin/bdpp_prepareserver -m adminLanIpAddress -o FJSVrcx.confFile -i
ipaddr.confFile -f initiator.confFile <Enter>
```

See

- Refer to "B.4 FJSVrcx.conf" and "B.5 ipaddr.conf" for information on the definition file used for network parameter automatic configuration.
- Refer to "B.6 initiator.conf" for information on the iSCSI initiator configuration file.
- Refer to "A.10 bdpp_prepareserver" for information on the bdpp_prepareserver command.

6.3.1.4 Creating a Clone Image

Create a clone image from a slave server that is already installed.

- To change the image storage directory, execute the bdpp_changeimagedir command.

```
# /opt/FJSVbdpp/bin/bdpp_changeimagedir imageDirectory <Enter>
```

- Execute the df command and confirm the amount of disk space available (Avail) in the image storage directory.

Refer to "Clone image file storage area" for information on how to confirm whether there is enough space to store the clone image.

Example

When confirming the disk usage of the file system in /data/imagedir (image storage directory):

```
# df -h /data/imagedir <Enter>
Filesystem      Size  Used Avail Use% mount point
/dev/sda3       1.9T   62G  1.8T   4% /
```

The free space on disk is 1.8TB in this example.

- Specify the name of the slave server for creating the clone image and the name of the clone image in the bdpp_getimage command.

```
# /opt/FJSVbdpp/bin/bdpp_getimage serverName cloneImageName <Enter>
```

- Use the bdpp_listimage command to check that the clone image could be created successfully.

```
# /opt/FJSVbdpp/bin/bdpp_listimage <Enter>
```

Example

Checking the clone image creation status:

```
# /opt/FJSVbdpp/bin/bdpp_listimage <Enter>
IMAGEDIRECTORY
-----
imagedir:  /data/imagedir

NAME                VERSION  CREATIONDATE          COMMENT
-----
RX200img            1        2012/04/20-22:39:59  -
RX300img            1        2012/04/20-22:41:13  -
```



See

- Refer to "[A.3 bdpp_changeimagedir](#)" for information on the `bdpp_changeimagedir` command.
- Refer to "[A.6 bdpp_getimage](#)" for information on the `bdpp_getimage` command.
- Refer to "[A.8 bdpp_listimage](#)" for information on the `bdpp_listimage` command.

6.3.1.5 Restarting the System

Restart the slave server from which the image was gathered.

```
# shutdown -r now <Enter>
```

6.3.1.6 Cloning

Clone as many times as the number of slave servers where the image is to be distributed.

Specify the image distribution destination server and the name of the clone image and execute the `bdpp_deployimage` command.

```
# /opt/FJSVbdpp/bin/bdpp_deployimage serverName cloneImageName <Enter>
```



See

Refer to "[A.5 bdpp_deployimage](#)" for information on the `bdpp_deployimage` command.

6.3.1.7 Restarting the System

Restart slave server to which the image was distributed.

```
# shutdown -r now <Enter>
```

6.3.2 Adding a Slave Server to a Virtual Environment

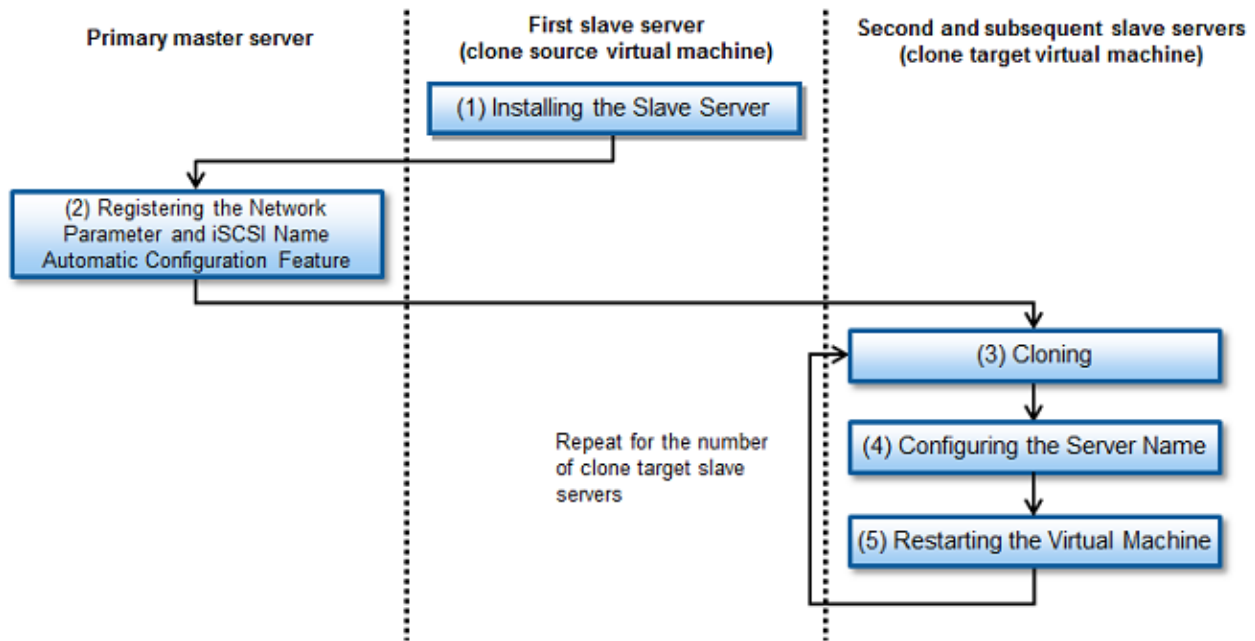
The procedure for adding a slave server to a virtual environment is shown below.

1. [Installing the Slave Server](#)
2. [Registering the Network Parameter and iSCSI Name Automatic Configuration Feature](#)
3. [Cloning](#)
4. [Configuring the Server Name](#)
5. [Restarting the Virtual Machine](#)

The procedure for installing a slave server is shown below.

Create the clone image of the first slave server (one that is already installed) from the virtual machine where it was built, and clone as many times as the number of slave servers to be installed.

Perform this action with root permissions unless otherwise indicated.



Note

The area where tasks are performed on the clone target virtual machines depends on the virtualization software.

6.3.2.1 Installing the Slave Server

Refer to "6.2 Installing to a Slave Server" for information on how to install the first additional slave server (the clone source slave server).

6.3.2.2 Registering the Network Parameter and iSCSI Name Automatic Configuration Feature

This section explains how to register the network parameter and iSCSI initiator name automatic configuration feature.

By setting the network parameters for the admin LAN, public LAN and iSCSI-LAN and the iSCSI initiator for the clone source virtual machine beforehand, it is possible to automatically set the slave server network and iSCSI initiator names when cloning.

1. Create the network parameter automatic configuration feature definition file (FJSVrcx.conf and ipaddr.conf).
Specify the admin LAN information for the slave server of the clone source in the FJSVrcx.conf definition file.
Specify the admin LAN, public LAN, and iSCSI-LAN information for all slave servers in the ipaddr.conf definition file.
2. Create the initiator.conf iSCSI initiator configuration file.
Specify the iSCSI initiator names corresponding to all slave servers that are clone targets.
3. Specify the IP address of the admin LAN of the slave servers which are clone sources, the FJSVrcx.conf, ipaddr.conf, and initiator.conf definition files and the -v option in the bdpp_prepareserver command.

```
# /opt/FJSVbdpp/bin/bdpp_prepareserver -m adminLanIpAddress -o FJSVrcx.confFile -i ipaddr.confFile -f initiator.confFile -v <Enter>
```

Example

The example that follows shows an ipaddr.conf network parameter automatic configuration file for the following configuration:

- NIC to connect to the admin LAN :eth0
- NIC to connect to the public LAN :eth1
- NIC to connect to iSCSI-LAN :eth2, eth3 (redundancy added)

```

NODE_NAME="slave1" <-- Slave server of the clone source
IF_NAME0="eth0"
IF_IPAD0="192.168.1.21"
IF_MASK0="255.255.255.0"
IF_NAME1="eth1"
IF_IPAD1="10.10.10.21"
IF_MASK1="255.255.255.0"
IF_NAME2="eth2"
IF_IPAD2="10.10.11.21"
IF_MASK2="255.255.255.0"
IF_NAME3="eth3"
IF_IPAD3="10.10.12.21"
IF_MASK3="255.255.255.0"

NODE_NAME="slave2" <-- Slave server of the clone target
IF_NAME0="eth0"
IF_IPAD0="192.168.1.22"
IF_MASK0="255.255.255.0"
IF_NAME1="eth1"
IF_IPAD1="10.10.10.22"
IF_MASK1="255.255.255.0"
IF_NAME2="eth2"
IF_IPAD2="10.10.11.22"
IF_MASK2="255.255.255.0"
IF_NAME3="eth2"
IF_IPAD3="10.10.12.22"
IF_MASK3="255.255.255.0"

NODE_NAME="slave3" <-- Slave server of the clone target
IF_NAME0="eth0"
IF_IPAD0="192.168.1.23"
IF_MASK0="255.255.255.0"
IF_NAME1="eth1"
IF_IPAD1="10.10.10.23"
IF_MASK1="255.255.255.0"
IF_NAME2="eth2"
IF_IPAD2="10.10.11.23"
IF_MASK2="255.255.255.0"
IF_NAME3="eth3"
IF_IPAD3="10.10.12.23"
IF_MASK3="255.255.255.0"

:

```



See

- Refer to "[B.4 FJSVrcx.conf](#)" and "[B.5 ipaddr.conf](#)" for information on the definition file used for network parameter automatic configuration.
- Refer to "[B.6 initiator.conf](#)" for information on the iSCSI initiator configuration file.
- Refer to "[A.10 bdpp_prepareserver](#)" for information on the bdpp_prepareserver command.

6.3.2.3 Cloning

Clone virtual machines from the clone source slave server virtual machine as many times as the number of clone target slave servers.



See

Refer to the manuals for the virtualization software for information on how to clone virtual machines.

6.3.2.4 Configuring the Server Name

Start the slave server of the clone target and set a server name.

Change the HOSTNAME parameter in the /etc/sysconfig/network file to the server name of the slave server on the clone target.



Example

Cloning slave2 from the slave server slave1

Before setting:

```
HOSTNAME=slave1
```

After setting:

```
HOSTNAME=slave2
```



Point

In the HOSTNAME parameter, specify server name defined in the ipaddr.conf network parameter automatic configuration file specified in "[6.3.2.2 Registering the Network Parameter and iSCSI Name Automatic Configuration Feature](#)".



See

Refer to "[5.7 Server Name Settings](#)" for information on how to set server names.

6.3.2.5 Restarting the Virtual Machine

Restart the slave server of the clone target.



See

Refer to the manuals for the virtualization software for information on how to start and restart virtual machines.

6.4 Installing to a Development Server

This section describes the installation of the development server functionality to a development server.

Install and set up using root permissions for all tasks.

1. [Installing the Development Server Functionality](#)
2. [DFS Setup](#)
3. [Hadoop Setup](#)

6.4.1 Installing the Development Server Functionality

Follow the procedures below to install the development server functionality.

1. [Creating bdpp.conf](#)
2. [Installing the Development Server Functionality](#)
3. [Applying Software Updates](#)

6.4.1.1 Creating bdpp.conf

Before installing the development server functionality, set the previously designed values in bdpp.conf.



See

.....
 Refer to "Development server" under "B.1 bdpp.conf" for information on bdpp.conf.

6.4.1.2 Installing the Development Server Functionality

Follow the procedures below to install the development server functionality.

1. Set the "BDPP_CONF_PATH" environment variable.
 In the "BDPP_CONF_PATH" environment variable, set the directory where bdpp.conf is stored.



Example

.....
 If bdpp.conf is stored in /home/data/develop:


```
# export BDPP_CONF_PATH=/home/data/develop <Enter>
```

2. Mount the DVD-ROM.
 Place the DVD-ROM in the DVD drive, then execute the command below to mount the DVD-ROM. If the DVD-ROM is automatically mounted by the automatic mount daemon (autofs), installer startup will fail because "noexec" is set in the mount options.

```
# mount /dev/hdc dvdRomMountPoint <Enter>
```

3. Start the installer.
 Execute bdpp_inst to start the installer.

```
# dvdRomMountPoint/bdpp_inst <Enter>
```

4. Select "3. Development Server installation".

```
=====
                          Interstage Big Data Parallel Processing Server
                              V1.0.1
                    All Rights Reserved, Copyright(c) FUJITSU LIMITED 2012
=====
```

```
<< Menu >>

1. Master Server installation
2. Slave Server installation
3. Development Server installation
4. Collaboration Server installation

=====
Please select number to start specified Server installation [?,q]
=> 3 <Enter>
```

Note

If installation fails, refer to the messages output during the installation operation and/or installation log file (/var/tmp/bdpp_install.log) to diagnose the failure. Then, remove the cause of the failure and perform installation again.

6.4.1.3 Applying Software Updates

After installing the development server functionality, apply the updates for this product and the updates for software included with this product.

Refer to the note in "[Chapter 6 Installation](#)" for details.

6.4.2 DFS Setup

The DFS setup sequence is as follows:

1. [Mount Point Creation and fstab Settings](#)
2. [Mounting](#)

6.4.2.1 Mount Point Creation and fstab Settings

Create the mount point and add the DFS entry to /etc/fstab.

Creating the mount point

Create the mount point for mounting the disk partitions on the storage system used as the DFS.

Make the mount point the same as that specified at the master server.

Example

Create the mount point "pdfs" under "/mnt".

```
# mkdir /mnt/pdf <Enter>
```

fstab settings

At "/etc/fstab", define the mount points created above and the logical file system name.

The logical file system name is used to identify the DFS file system. Use the name defined when generating the configuration information of the file system on the master server.

Example

This example shows the mount point "/mnt/pdfs" and the logical file system name "pdfs1" defined at "/etc/fstab".

LABEL=/	/	ext3	defaults	1 1
LABEL=/home	/home	ext3	defaults	1 2
LABEL=/boot	/boot	ext3	defaults	1 2
tmpfs	/dev/shm	tmpfs	defaults	0 0
devpts	/dev/pts	devpts	gid=5,mode=620	0 0
sysfs	/sys	sysfs	defaults	0 0
proc	/proc	proc	defaults	0 0
LABEL=SWAP-sda3	swap	swap	defaults	0 0
pdfs1	/mnt/pdfs	pdfs	_netdev	0 0

See

Refer to "[6.2.2.1 Creating the Mount Point and Configuring fstab Settings](#)" used when installing slave servers for information about the parameters specified in the fields of the entries added to /etc/fstab.

6.4.2.2 Mounting

Distribute the DFS configuration information file generated on the master server to the development server to mount the DFS file system.

1. Copy the DFS configuration information file generated in "[6.1.3.11 Generating the DFS File System Configuration Information](#)" for the master server to /etc/pdfs.

```
# scp -p root@master1:dfsConfigurationFileDirectory/client.conf.pdfs1 /etc/pdfs <Enter>
```

2. Mount the DFS file system.

Example

Mount the DFS file system at the development server.

```
# mount pdfs1 <Enter>
```

Note

- The DFS configuration information file must be distributed to the development server before the DFS file system is mounted.
- Ensure that the DFS mount sequence is: master servers, slave servers, development servers and collaboration servers. If the slave servers, development servers or collaboration servers are mounted first, then mount will fail, because the master server (MDS) does not exist.

See

Refer to the "mount.pdfs" under "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for details of mounting.

6.4.3 Hadoop Setup

The Hadoop setup procedure is shown below.

1. Copy the master server 'slaves' file, edited as described in "[6.1.4.1 Creating the Slave Server Definition File 'slaves'](#)", and place the copy under "/etc/opt/FJSVbdpp/conf".

```
# scp -p root@master1:/etc/opt/FJSVbdpp/conf/slaves /etc/opt/FJSVbdpp/conf <Enter>
```

2. Execute Hadoop setup.

Execute bdpp_setup.

```
# /opt/FJSVbdpp/setup/bdpp_setup <Enter>
```



.....

If setup fails, refer to the messages output during the setup operation and/or setup log file (/var/opt/FJSVbdpp/log/bdpp_setup.log) to diagnose the failure. Then, remove the cause of the failure and perform setup again.

.....



.....

Hadoop setup sets Hadoop and DFS configurations as default settings of this product. Refer to "[Appendix C Hadoop Configuration Parameters](#)" for information on how to tune the configuration settings.

.....

6.5 Installing to a Collaboration Server

This section describes the installation of the collaboration server functionality to a collaboration server.

Install and set up using root permissions for all tasks.

1. [Installing the Collaboration Server Functionality](#)
2. [DFS Setup](#)

6.5.1 Installing the Collaboration Server Functionality

Follow the procedures below to install the collaboration server functionality.

1. [Creating bdpp.conf](#)
2. [Installing the Collaboration Server Functionality](#)
3. [Applying Software Updates](#)

6.5.1.1 Creating bdpp.conf

To install the collaboration server functionality, configure in bdpp.conf the values that were designed in advance.



.....

Refer to "Collaboration server" under "[B.1 bdpp.conf](#)" for information on bdpp.conf.

.....

6.5.1.2 Installing the Collaboration Server Functionality

Follow the procedures below to install the collaboration server functionality.

1. Set the "BDPP_CONF_PATH" environment variable.

In the "BDPP_CONF_PATH" environment variable, set the directory where bdpp.conf is stored.

Example

If bdpp.conf is stored in /home/data/collaborate:

```
# export BDPP_CONF_PATH=/home/data/collaborate <Enter>
```

2. Mount the DVD-ROM.

Place the DVD-ROM in the DVD drive, then execute the command below to mount the DVD-ROM. If the DVD-ROM is automatically mounted by the automatic mount daemon (autofs), installer startup will fail because "noexec" is set in the mount options.

```
# mount /dev/hdc dvdRomMountPoint <Enter>
```

3. Start the installer.

Execute bdpp_inst to start the installer.

```
# dvdRomMountPoint/bdpp_inst <Enter>
```

4. Select "4. Collaboration Server installation".

```
=====
                Interstage Big Data Parallel Processing Server
                        V1.0.1
                All Rights Reserved, Copyright(c) FUJITSU LIMITED 2012
=====

<< Menu >>

1. Master Server installation
2. Slave Server installation
3. Development Server installation
4. Collaboration Server installation

=====
Please select number to start specified Server installation [?,q]
=> 4 <Enter>
```

Note

If installation fails, refer to the messages output during the installation operation and/or installation log file (/var/tmp/bdpp_install.log) to diagnose the failure. Then, remove the cause of the failure and perform installation again.

6.5.1.3 Applying Software Updates

Apply the updates for this product and the updates for software included with this product.

Refer to the note in "[Chapter 6 Installation](#)" for details.

6.5.2 DFS Setup

The DFS setup sequence is as follows:

1. [Mount Point Creation and fstab Settings](#)
2. [Register the hadoop Group and mapred User](#)
3. [Mounting](#)

6.5.2.1 Mount Point Creation and fstab Settings

Create the mount point and add the DFS entry to `/etc/fstab`.

Creating the mount point

Create the mount point for mounting the disk partitions on the storage system used as the DFS.

Make the mount point the same as that specified at the master server.



Example

Create the mount point "pdfs" under "/mnt".

```
# mkdir /mnt/pdf <Enter>
```

fstab settings

At `/etc/fstab`, define the mount points created above and the logical file system name.

The logical file system name is used to identify the DFS file system. Use the name defined when generating the configuration information of the file system on the master server.



Example

This example shows the mount point `/mnt/pdf` and the logical file system name `pdfs` defined at `/etc/fstab`.

LABEL=/	/	ext3	defaults	1	1
LABEL=/home	/home	ext3	defaults	1	2
LABEL=/boot	/boot	ext3	defaults	1	2
tmpfs	/dev/shm	tmpfs	defaults	0	0
devpts	/dev/pts	devpts	gid=5,mode=620	0	0
sysfs	/sys	sysfs	defaults	0	0
proc	/proc	proc	defaults	0	0
LABEL=SWAP-sda3	swap	swap	defaults	0	0
pdfs1	/mnt/pdf	pdfs	_netdev	0	0



See

Refer to "[6.2.2.1 Creating the Mount Point and Configuring fstab Settings](#)" used when installing slave servers for information about the parameters specified in the fields of the entries added to `/etc/fstab`.

6.5.2.2 Registering the hadoop Group and mapred User

Register the hadoop group and mapred user.

The hadoop group GID and the mapred user UID must be the same values as the master server hadoop group GID and mapred user UID.

Check the master server "/etc/passwd" file and register the same values.



Example

1. Check the hadoop group amongst the groups registered at the master server.

```
# cat /etc/group <Enter>
... omitted ...
hadoop:x:123:hbase
hbase:x:1503:
bdppgroup:x:1500:
```

2. Check the mapred user amongst the users registered at the master server.

```
# cat /etc/passwd <Enter>
... omitted ...
bdppuser1:x:1500:1500::/home/bdppuser1:/bin/bash
bdppuser2:x:1501:1500::/home/bdppuser2:/bin/bash
mapred:x:202:123:Hadoop MapReduce:/tmp:/bin/bash
hdfs:x:201:123:Hadoop HDFS:/tmp:/bin/bash
hbase:x:203:1503::/tmp:/bin/bash
```

3. Register the hadoop group.

```
# groupadd -g 123 hadoop <Enter>
```

4. Register the mapred user.

```
# useradd -g hadoop -u 202 -c "Hadoop MapReduce" -d /tmp -s /bin/bash mapred <Enter>
```



Note

If the GID and UID need to be changed to different values because the GID and UID are already registered at the collaboration server, the hadoop group GID and the mapred user UID registered at the master server, slave servers and development server must all be changed to the same values.

6.5.2.3 Mounting

Distribute the DFS configuration information file generated on the master server to the collaboration server to mount the DFS file system.

1. Copy the DFS configuration information file generated in "[6.1.3.11 Generating the DFS File System Configuration Information](#)" for the master server to /etc/pdfs.

```
# scp -p root@master1:dfsConfigurationFileDirectory/client.conf.pdfs1 /etc/pdfs <Enter>
```

2. Mount the DFS file system.



Example

Mount the DFS file system on the collaboration server.

```
# mount pdfs1 <Enter>
```

Note

- The DFS configuration information file must be distributed to the collaboration server before the DFS file system is mounted.
- Ensure that the DFS mount sequence is: master servers, slave servers, development servers and collaboration servers. If the slave servers, development servers or collaboration servers are mounted first, then mount will fail, because the master server (MDS) does not exist.

See

Refer to the "mount.pdfs" under "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for details of mounting.

Chapter 7 Uninstallation

This chapter explains how to uninstall the server features used to build the system for this product.

- [Uninstalling from a Master Server](#)
- [Uninstalling from a Slave Server](#)
- [Uninstalling from a Development Server](#)
- [Uninstalling from a Collaboration Server](#)

7.1 Uninstalling from a Master Server

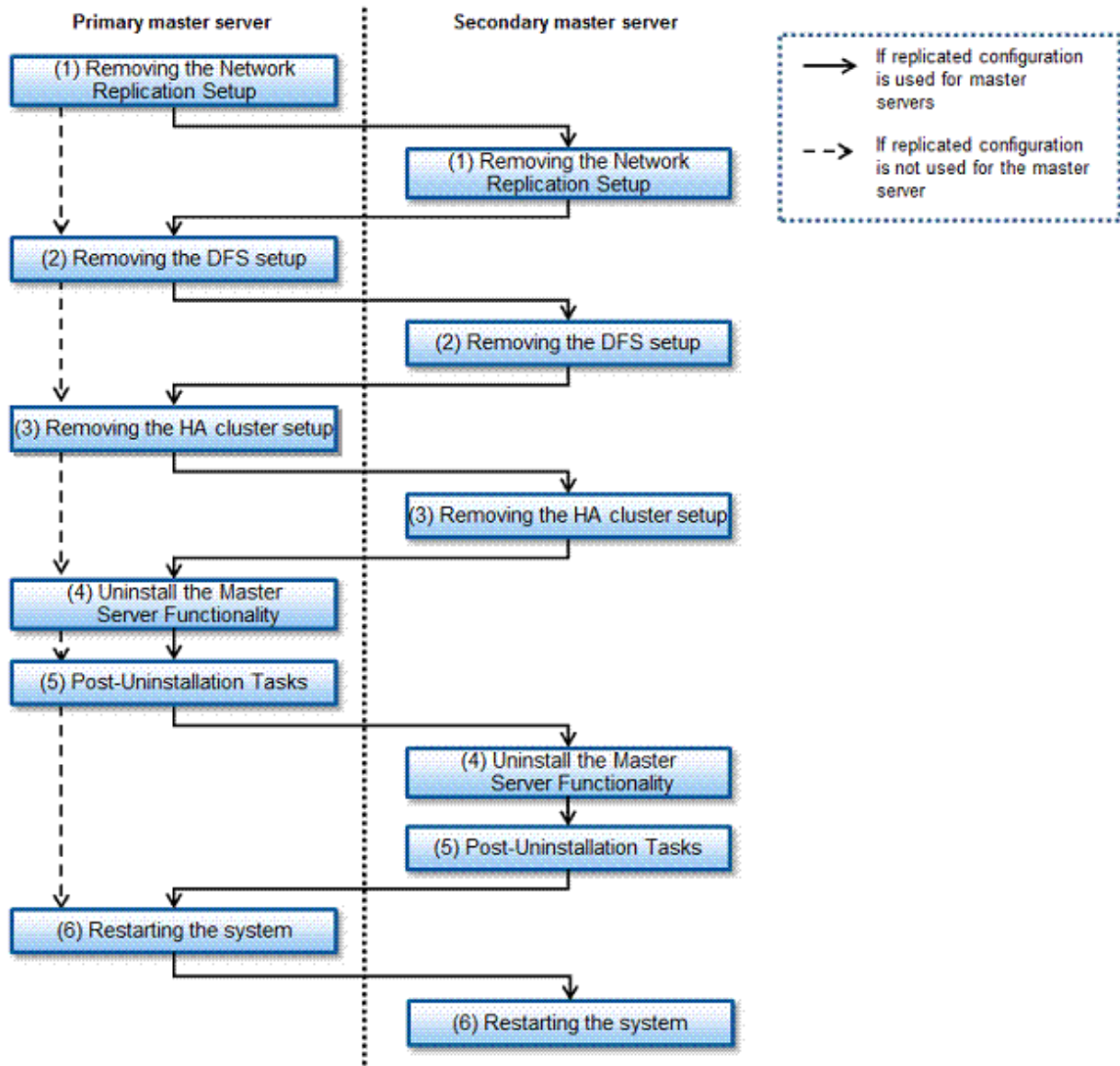
This section describes how to uninstall the master server functionality from a master server.

1. [Removing the Network Replication Setup](#)
2. [Removing the DFS Setup](#)
3. [Removing the HA Cluster Setup](#)
4. [Uninstalling the Master Server Functionality](#)
5. [Post-Uninstallation Tasks](#)
6. [Restarting the System](#)

The work flow for uninstalling a master server is shown below.

Perform setup cancellation alternating between the primary master server first, then the secondary master server, in order to uninstall a master server replicated configuration.

Uninstall and remove the setup using root permissions for all tasks.



7.1.1 Removing the Network Replication Setup

Remove the network replication setup.

Execute `bdpp_unsetup`.

```
# /opt/FJSVbdpp/setup/bdpp_unsetup <Enter>
```

Note

If setup removal fails, refer to the explanation and action in the message that is output. Or, if required, refer to the setup log (`/var/opt/FJSVbdpp/log/bdpp_setup.log`). Remove the cause of the error and remove the setup again.

7.1.2 Removing the DFS Setup

The DFS setup removal sequence is as follows:

In a DFS, information on the partitions comprising a file system is recorded in the management partition.

Therefore, when the DFS currently being used is no longer required, it must be deleted.

Note

When a file system is deleted, all file system data is deleted. Therefore, back up required data before deletion.

The DFS deletion procedure is described below using the following environment as an example:

- Management partition :/dev/disk/by-id/scsi-1FUJITSU_300000370105
- Representative partition :/dev/disk/by-id/scsi-1FUJITSU_300000370106
- File data partition :/dev/disk/by-id/scsi-1FUJITSU_300000370107
:/dev/disk/by-id/scsi-1FUJITSU_300000370108
- File system ID :1
- Logical file system name :pdfs1

Unless specifically indicated otherwise, execute at the primary master server.

1. Check the current file system information.

```
# pdfsinfo -a <Enter>
/dev/disk/by-id/scsi-1FUJITSU_300000370106:
FSID special size Type mount
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 25418 META -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 5120 LOG -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 7341778 DATA -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370108 (896) 6578704 DATA -----
```

2. If the targeted DFS is mounted, unmount it.

Unmount the DFS at all slave servers, development servers and collaboration servers.

```
# umount pdfs1 <Enter>
```

Unmount the DFS at the master server.

```
# pdfsumntgl /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
```

3. Delete the DFS.

```
# pdfsadm -D /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
```

4. Check that the DFS was deleted.

Check that the file system information recorded in the management partition does not exist.

```
# pdfsinfo -a <Enter>
```

5. On the primary master server and the secondary master server, delete the connection authorization list file.

```
# cd /etc/pdfsf <Enter>
# rm ./server.conf.1 <Enter>
```

6. Delete the targeted DFS description from the /etc/fstab of the primary master server and the secondary master server.

Relevant entries are those with a representative partition (device name using by-id name) as the first field.

7. Stop the pdfsfrmd daemon.

Stop the pdfsfrmd daemons running on the primary master server and the secondary master server.

```
# pdfsfrmstop <Enter>
```


8. Delete the DFS management server information on the primary master server and the secondary master server.

```
# pdfssetup -d <Enter>
```



Refer to the "Appendix A Command Reference" of the "Primesoft Distributed File System for Hadoop V1 User's Guide" for information on pdfsinfo, pdfsumntgl, pdfsadm, pdfsfrmstop, and pdfssetup.

7.1.3 Removing the HA Cluster Setup

Execute HA cluster setup cancellation.

Execute cluster_unsetup.

```
# /opt/FJSVbdpp/setup/cluster_unsetup <Enter>
```



If setup removal fails, refer to the explanation and action in the message that is output. Or, if required, refer to the setup log (/var/opt/FJSVbdpp/log/bdpp_setup.log). Remove the cause of the error and remove the setup again.

7.1.4 Uninstalling the Master Server Functionality

The master server functionality uninstall procedure is shown below.

1. If UpdateSite format updates have been applied to this product and the following products included in this product, delete the updates.
 - PRIMECLUSTER Clustering Base
 - PRIMECLUSTER GLS
 - Primesoft Distributed File System for Hadoop
 - ServerView Resource Orchestrator Virtual Edition (*1)

*1: Only the primary master server installed on a physical environment



Refer to the "UpdateAdvisor (middleware)" help and the update information files of update patches for details.

2. Change the default run level so that the system starts in single user mode.

Change the ID entry content to 1 in the "/etc/inittab" file.



In order to return to the original run-level after uninstallation, record and keep the pre-change default run level.

```
# Default runlevel. The runlevels used are:  
# 0 - halt (Do NOT set initdefault to this)
```

```
# 1 - Single user mode
# 2 - Multiuser, without NFS (The same as 3, if you do not have networking)
# 3 - Full multiuser mode
# 4 - unused
# 5 - X11
# 6 - reboot (Do NOT set initdefault to this)
#
id:1:initdefault:          <-- Change this row.
```

3. Reboot in single user mode.

```
# shutdown -r now <Enter>
```

4. Execute the uninstaller.

Start the network interface, and then execute `bdpp_mgruninstall`.

```
# /etc/init.d/network start <Enter>
# /opt/FJSVbdpp/manager/bdpp_mgruninstall <Enter>
```

5. Change the default run level at system startup back to the original setting.

Return to the default run level that was recorded in Step 2.

```
# Default runlevel. The runlevels used are:
# 0 - halt (Do NOT set initdefault to this)
# 1 - Single user mode
# 2 - Multiuser, without NFS (The same as 3, if you do not have networking)
# 3 - Full multiuser mode
# 4 - unused
# 5 - X11
# 6 - reboot (Do NOT set initdefault to this)
#
id:5:initdefault:          <-- Change this row back to the original setting.
```

Note

- If uninstallation fails, refer to the messages came up during the uninstallation operation and/or installation log file (`/var/tmp/bdpp_install.log`) to diagnose the failure. Then, remove the cause of the failure and perform uninstallation again.
- This product also installs "UpdateAdvisor (middleware)" in preparation for applying patches to this product and related products, but the uninstaller of this product does not uninstall "UpdateAdvisor (middleware)". Refer to the "UpdateAdvisor (middleware)" help for information on uninstalling "UpdateAdvisor (middleware)".

7.1.5 Post-Uninstallation Tasks

This section describes the master server post-uninstall tasks.

7.1.5.1 Directories and Files Remaining after an Uninstall

Directories and files might remain after an uninstall.

Delete the following directories (includes directories and files under that directory) and files:

Directories:

- `/opt/FJSVbdpp`
- `/opt/SMAW`
- `/etc/opt/FJSVbdpp`

- /etc/opt/FJSVhanet
- /etc/opt/FJSVwvbs
- /etc/opt/SMAW
- /etc/pdfs
- /var/opt/FJSVbdpp
- /var/opt/FJSVpdfs

Files:

- /etc/cip.cf
- /etc/default/cluster
- /etc/default/cluster.config
- /var/adm/cfreg.data

7.1.5.2 Uninstalling "Uninstall (middleware)"

"Uninstall (middleware)" is a tool common to Fujitsu middleware products.

This section describes how to uninstall "Uninstall (middleware)" and provides additional notes.

To uninstall "Uninstall (middleware)", follow the procedure below:

1. Start "Uninstall (middleware)" and check that other Fujitsu middleware products do not remain.

```
# /opt/FJSVcir/cir/bin/cimanager.sh [-c] <Enter>
```

-c: command interface

Note

If the command path contains a blank space, it will fail to start. Do not specify a directory with a name containing blank spaces.

Information

To start in command mode, specify -c. If -c is not specified, startup is in GUI mode if there is a GUI environment, and in command mode if there is no GUI environment.

2. After checking that no Fujitsu middleware product has been installed, execute the following uninstallation command:

```
# /opt/FJSVcir/bin/cirremove.sh <Enter>
```

3. If "This software is a common tool of Fujitsu products. Are you sure you want to remove it? [y/n]:" is displayed, enter "y" and continue.

Uninstallation will be completed in seconds.

4. After uninstallation is complete, delete the following directory and contained files.

- /var/opt/FJSVcir/cir/internal
- /var/opt/FJSVcir/cir/logs
- /var/opt/FJSVcir/cir/meta_db

Point

This task is not required when the master server is installed on a virtual environment.

7.1.6 Restarting the System

Restart the system.

```
# shutdown -r now <Enter>
```

7.2 Uninstalling from a Slave Server

This section describes how to uninstall the slave server functionality from a slave server.

Uninstall and remove the setup using root permissions for all tasks.

1. [Removing the Network Replication Setup](#)
2. [Removing the DFS Setup](#)
3. [Uninstalling the Slave Server Functionality](#)
4. [Post-Uninstallation Tasks](#)
5. [Restarting the System](#)

7.2.1 Removing the Network Replication Setup

Remove the network replication setup.

Execute `bdpp_unsetup`.

```
# /opt/FJSVbdpp/setup/bdpp_unsetup <Enter>
```

Note

If setup removal fails, refer to the explanation and action in the message that is output. Or, if required, refer to the setup log (`/var/opt/FJSVbdpp/log/bdpp_setup.log`). Remove the cause of the error and remove the setup again.

7.2.2 Removing the DFS Setup

The DFS setup removal sequence is as follows:

The setup removal procedure is described below using the following environment as an example:

- Representative partition `:/dev/disk/by-id/scsi-1FUJITSU_300000370106`
- File system ID `:1`
- Logical file system name `:pdfs1`

1. If the targeted DFS is mounted, unmount it.

```
# umount pdfs1 <Enter>
```

2. Delete the DFS configuration information file.

```
# cd /etc/pdfs <Enter>
# rm ./client.conf.pdfs1 <Enter>
```

3. Remove the target DFS entries from /etc/fstab.

Relevant entries are those with a logical file system name as the first field.

7.2.3 Uninstalling the Slave Server Functionality

The slave server functionality uninstall procedure is shown below.

1. If UpdateSite format updates have been applied to this product and the following products included in this product, delete the updates.
 - PRIMECLUSTER GLS
 - Primesoft Distributed File System for Hadoop
 - ServerView Resource Orchestrator Virtual Edition (*1)

*1: Only slave servers installed on a physical environment



Refer to the "UpdateAdvisor (middleware)" help and the update information files of update patches for details.

2. Change the default run level so that the system starts in single user mode.

Change the ID entry content to 1 in the "/etc/inittab" file.



In order to return to the original run-level after uninstallation, record and keep the pre-change default run level.

```
# Default runlevel. The runlevels used are:
# 0 - halt (Do NOT set initdefault to this)
# 1 - Single user mode
# 2 - Multiuser, without NFS (The same as 3, if you do not have networking)
# 3 - Full multiuser mode
# 4 - unused
# 5 - X11
# 6 - reboot (Do NOT set initdefault to this)
#
id:1:initdefault:          <-- Change this row.
```

3. Restart in single-user mode.

```
# shutdown -r now <Enter>
```

4. Execute the uninstall.

Start the network interface, and then execute `bdpp_agtuninstall`.

```
# /etc/init.d/network start <Enter>
# /opt/FJSVbdpp/agent/bdpp_agtuninstall <Enter>
```

5. Change the default run level at system startup back to the original setting.

Return to the default run level that was recorded in Step 2.

```
# Default runlevel. The runlevels used are:
# 0 - halt (Do NOT set initdefault to this)
# 1 - Single user mode
# 2 - Multiuser, without NFS (The same as 3, if you do not have networking)
# 3 - Full multiuser mode
# 4 - unused
# 5 - X11
# 6 - reboot (Do NOT set initdefault to this)
#
id:5:initdefault:          <-- Change this row back to the original setting.
```

Note

- If uninstallation fails, refer to the messages came up during the uninstallation operation and/or installation log file (`/var/tmp/bdpp_install.log`) to diagnose the failure. Then, remove the cause of the failure and perform uninstallation again.
- This product also installs "UpdateAdvisor (middleware)" in preparation for applying patches to this product and related products, but the uninstaller of this product does not uninstall "UpdateAdvisor (middleware)". Refer to the "UpdateAdvisor (middleware)" help for information on uninstalling "UpdateAdvisor (middleware)".

7.2.4 Post-Uninstallation Tasks

This section describes the slave server post-uninstall tasks.

7.2.4.1 Directories and Files Remaining after an Uninstall

Directories and files might remain after an uninstall.

Delete the following directories (includes directories and files under that directory) and files:

Directories:

- `/opt/FJSVbdpp`
- `/opt/FJSVnrmp (*1)`
- `/opt/FJSVrcxat (*1)`
- `/opt/FJSVssagt (*1)`
- `/opt/systemcastwizard (*1)`
- `/etc/opt/FJSVbdpp`
- `/etc/opt/FJSVhanet`
- `/etc/opt/FJSVnrmp (*1)`
- `/etc/opt/FJSVrcxat (*1)`
- `/etc/opt/FJSVssagt (*1)`
- `/etc/pdfs`
- `/var/opt/FJSVbdpp`
- `/var/opt/FJSVpdfs`
- `/var/opt/FJSVnrmp (*1)`
- `/var/opt/FJSVrcxat (*1)`
- `/var/opt/FJSVssagt (*1)`
- `/var/opt/systemcastwizard (*1)`

Files:

- /boot/clcomp2.dat (*1)
- /etc/init.d/scwagent (*1)
- /etc/scwagent.conf (*1)

*1: Only in slave servers installed on a physical environment

7.2.5 Restarting the System

Restart the system.

```
# shutdown -r now <Enter>
```

7.3 Uninstalling from a Development Server

This section describes how to uninstall the development server functionality from a development server.

Uninstall and remove the setup using root permissions for all tasks.

1. [Removing the DFS Setup](#)
2. [Uninstalling the Development Server Functionality](#)
3. [Post-Uninstallation Tasks](#)
4. [Restarting the System](#)

7.3.1 Removing the DFS Setup

The DFS setup removal sequence is as follows:

The setup removal procedure is described below using the following environment as an example:

- Representative partition :/dev/disk/by-id/scsi-1FUJITSU_300000370106
- File system ID :l
- Logical file system name :pdfs1

1. If the targeted DFS is mounted, unmount it.

```
# umount pdfs1 <Enter>
```

2. Delete the DFS configuration information file.

```
# cd /etc/pdfs <Enter>
# rm ./client.conf.pdfs1 <Enter>
```

3. Remove the target DFS entries from /etc/fstab.

Relevant entries are those with a logical file system name as the first field.

7.3.2 Uninstalling the Development Server Functionality

Follow the procedures below to uninstall the development server functionality.

1. If UpdateSite format updates have been applied to this product and the following products included in this product, delete the updates.

- Primesoft Distributed File System for Hadoop



Refer to the "UpdateAdvisor (middleware)" help and the update information files of update patches for details.

2. Change the default run level so that the system starts in single user mode.

Change the ID entry content to 1 in the "/etc/inittab" file.



In order to return to the original run-level after uninstallation, record and keep the pre-change default run level.

```
# Default runlevel. The runlevels used are:
# 0 - halt (Do NOT set initdefault to this)
# 1 - Single user mode
# 2 - Multiuser, without NFS (The same as 3, if you do not have networking)
# 3 - Full multiuser mode
# 4 - unused
# 5 - X11
# 6 - reboot (Do NOT set initdefault to this)
#
id:1:initdefault:          <-- Change this row.
```

3. Restart in single-user mode.

```
# shutdown -r now <Enter>
```

4. Execute the uninstall.

Start the network interface, and then execute bdpp_devuninstall.

```
# /etc/init.d/network start <Enter>
# /opt/FJSVbdpp/agent/bdpp_devuninstall <Enter>
```

5. Change the default run level at system startup back to the original setting.

Return to the default run level that was recorded in Step 2.

```
# Default runlevel. The runlevels used are:
# 0 - halt (Do NOT set initdefault to this)
# 1 - Single user mode
# 2 - Multiuser, without NFS (The same as 3, if you do not have networking)
# 3 - Full multiuser mode
# 4 - unused
# 5 - X11
# 6 - reboot (Do NOT set initdefault to this)
#
id:5:initdefault:          <-- Change this row back to the original setting.
```



- If uninstallation fails, refer to the messages came up during the uninstallation operation and/or installation log file (/var/tmp/bdpp_install.log) to diagnose the failure. Then, remove the cause of the failure and perform uninstallation again.

- This product also installs "UpdateAdvisor (middleware)" in preparation for applying patches to this product and related products, but the uninstaller of this product does not uninstall "UpdateAdvisor (middleware)". Refer to the "UpdateAdvisor (middleware)" help for information on uninstalling "UpdateAdvisor (middleware)".
-

7.3.3 Post-Uninstallation Tasks

This section describes the development server post-uninstall tasks.

7.3.3.1 Directories and Files Remaining after an Uninstall

Directories and files might remain after an uninstall.

Delete the following directories (includes directories and files under that directory) and files:

Directories:

- /opt/FJSVbdpp
- /etc/opt/FJSVbdpp
- /etc/pdfs
- /var/opt/FJSVbdpp
- /var/opt/FJSVpdfs

7.3.4 Restarting the System

Restart the system.

```
# shutdown -r now <Enter>
```

7.4 Uninstalling from a Collaboration Server

This section describes how to uninstall the collaboration server functionality from a collaboration server.

Uninstall and remove the setup using root permissions for all tasks.

1. [Removing the DFS Setup](#)
2. [Uninstalling the Collaboration Server Functionality](#)
3. [Post-Uninstallation Tasks](#)
4. [Restarting the System](#)

7.4.1 Removing the DFS Setup

The DFS setup removal sequence is as follows:

The setup removal procedure is described below using the following environment as an example:

- Representative partition :/dev/disk/by-id/scsi-1FUJITSU_300000370106
- File system ID :1
- Logical file system name :pdfs1

1. If the targeted DFS is mounted, unmount it.

```
# umount pdfs1 <Enter>
```

2. Delete the DFS configuration information file.

```
# cd /etc/pdfs <Enter>
# rm ./client.conf.pdfs1 <Enter>
```

3. Remove the target DFS entries from /etc/fstab.

Relevant entries are those with a logical file system name as the first field.

7.4.2 Uninstalling the Collaboration Server Functionality

Follow the procedure below to uninstall the collaboration server functionality.

1. If UpdateSite format updates have been applied to this product and the following products included in this product, delete the updates.

- Primesoft Distributed File System for Hadoop



Refer to the "UpdateAdvisor (middleware)" help and the update information files of update patches for details.

2. Change the default run level so that the system starts in single user mode.

Change the ID entry content to 1 in the "/etc/inittab" file.



In order to return to the original run-level after uninstallation, record and keep the pre-change default run level.

```
# Default runlevel. The runlevels used are:
# 0 - halt (Do NOT set initdefault to this)
# 1 - Single user mode
# 2 - Multiuser, without NFS (The same as 3, if you do not have networking)
# 3 - Full multiuser mode
# 4 - unused
# 5 - X11
# 6 - reboot (Do NOT set initdefault to this)
#
id:1:initdefault:          <-- Change this row.
```

3. Restart in single-user mode.

```
# shutdown -r now <Enter>
```

4. Execute the uninstall.

Start the network interface, and then execute bdpp_coouninstall.

```
# /etc/init.d/network start <Enter>
# /opt/FJSVbdpp/agent/bdpp_coouninstall <Enter>
```

5. Change the default run level at system startup back to the original setting.

Return to the default run level that was recorded in Step 2.

```
# Default runlevel. The runlevels used are:
# 0 - halt (Do NOT set initdefault to this)
# 1 - Single user mode
# 2 - Multiuser, without NFS (The same as 3, if you do not have networking)
# 3 - Full multiuser mode
# 4 - unused
# 5 - X11
# 6 - reboot (Do NOT set initdefault to this)
#
id:5:initdefault:          <-- Change this row back to the original setting.
```

Note

- If uninstallation fails, refer to the messages came up during the uninstallation operation and/or installation log file (/var/tmp/bdpp_install.log) to diagnose the failure. Then, remove the cause of the failure and perform uninstallation again.
- This product also installs "UpdateAdvisor (middleware)" in preparation for applying patches to this product and related products, but the uninstaller of this product does not uninstall "UpdateAdvisor (middleware)". Refer to the "UpdateAdvisor (middleware)" help for information on uninstalling "UpdateAdvisor (middleware)".

7.4.3 Post-Uninstallation Tasks

This section describes the collaboration server post-uninstall tasks.

7.4.3.1 Directories and Files Remaining after an Uninstall

Directories and files might remain after an uninstall.

Delete the following directories (includes directories and files under that directory) and files:

Directories:

- /opt/FJSVbdpp
- /etc/opt/FJSVbdpp
- /etc/pdfs
- /var/opt/FJSVbdpp
- /var/opt/FJSVpdfs

7.4.4 Restarting the System

Restart the system.

```
# shutdown -r now <Enter>
```

Part 3 Operations

This part describes the registration and execution of applications using this product, along with other general operating procedures such as system management, modifications to the system configuration, and maintenance.

Chapter 8 Starting and Stopping.....	98
Chapter 9 Developing and Registering Applications.....	101
Chapter 10 Executing and Stopping Jobs.....	104
Chapter 11 Managing Job Execution Users.....	105
Chapter 12 Adding and Deleting Slave Servers.....	109
Chapter 13 Adding and Deleting Storage Systems.....	119
Chapter 14 Backup and Restore.....	127
Chapter 15 Operations when There are Errors.....	135
Chapter 16 Troubleshooting.....	150

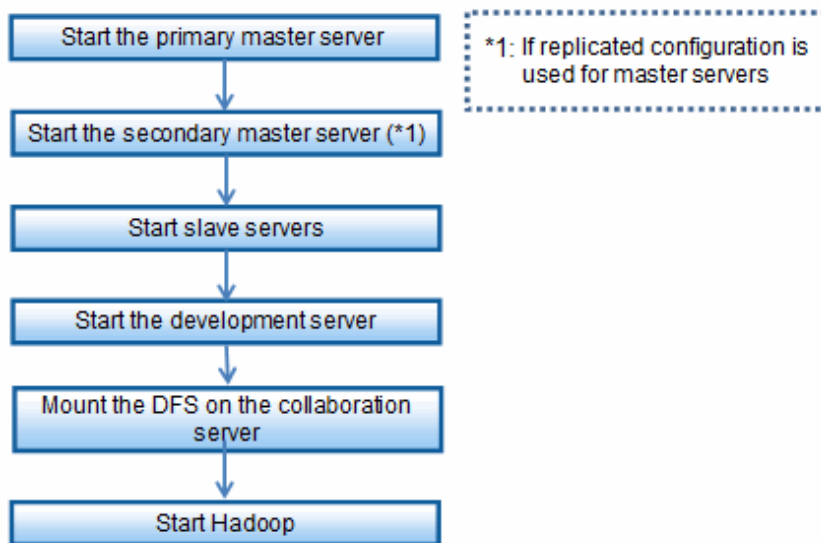
Chapter 8 Starting and Stopping

This chapter explains how to start, stop, and display the status of systems where this product is installed.

8.1 Starting

Perform the tasks in the following order when starting a system where this product is installed.

Figure 8.1 Sequence for starting the system where this product is installed



Mounting the DFS

When mounting the DFS, mount the DFS management server first, then mount the DFS clients.

For this reason, first start the master servers, which are the DFS management server. Start the primary master server and then the secondary master server when operating using a master server replicated configuration.

After starting the master servers, start the slave servers, development servers, and collaboration servers, which are the DFS clients.

When operating an existing system where the collaboration servers are already running, only mounting of the DFS on the collaboration server needs to be performed.

Note

- When using replicated configuration for the master server, the system is not running until the DFS management server has started on both the primary master server and the secondary master server. For this reason, start the secondary master server system without waiting for the primary master server system to finish starting.
- If DFS mounting is not set to automatic when starting the servers, mount the DFS manually after the servers have started.

Starting Hadoop

After DFS mounting is finished on all servers, use `bdpp_start` on the master servers to start the Hadoop of this product (refer to "A.14 `bdpp_start`" for details).

The following processes start when Hadoop is started:

- Master server
 - JobTracker (Hadoop)
- Slave server
 - TaskTracker (Hadoop)

Note

The "JobTracker" of Apache Hadoop on the master server cannot be started directly - you must start the Hadoop of this product using `bdpp_start`.

Refer to the Hadoop Project Top Page (<http://hadoop.apache.org/>) for details on how to use Apache Hadoop.

Information

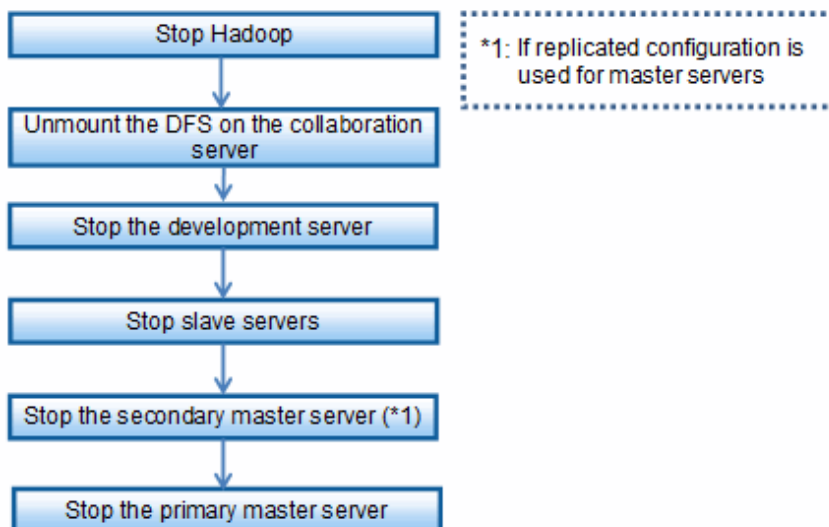
Set whether to mount the DFS automatically in `/etc/fstab`. Mounting is automatic when the DFS mount options are as follows:

- On a master server: The `noatrc` option has not been specified
- On any other server: The `noauto` option has not been specified

8.2 Stopping

Perform the tasks in the following order when stopping a system where this product is installed.

Figure 8.2 Sequence for stopping the system where this product is installed



Stopping Hadoop

Stop the servers after using `bdpp_stop` to stop the Hadoop of this product (refer to "A.16 `bdpp_stop`" for details).

Execution of this command stops the Hadoop processes operating at the master server and slave servers.

Unmounting the DFS

When unmounting the DFS, unmount the DFS clients first, then unmount the DFS management server.

For this reason, first stop the collaboration servers, development servers, and slave servers, which are the DFS clients.

When operating an existing system where the collaboration servers cannot be stopped, only unmounting of the DFS on the collaboration server needs to be performed.

Stop the master servers after DFS clients have been stopped.

First stop the primary master server, and then the secondary master server when operating using a master server replicated configuration.



Note

The "JobTracker" of Apache Hadoop on the master servers cannot be stopped directly - you must stop the Hadoop of this product using `bdpp_stop`.

Refer to the Hadoop Project Top Page (<http://hadoop.apache.org/>) for details on how to use Apache Hadoop.

8.3 Displaying Status

This section describes how to display the status of this product's Hadoop.

Use `bdpp_stat` to display the status of Hadoop processes running on the master server and slave servers (refer to "[A.15 bdpp_stat](#)" for details).

Chapter 9 Developing and Registering Applications

This chapter describes how to develop applications that can be executed under Hadoop.

9.1 Application Development Environment

Applications are developed using the APIs and language interfaces provided by the Hadoop, Hive, Pig and HBase functions installed on the development server.

Development of Java programs using the Hadoop API is made easier by using Eclipse and the Eclipse Hadoop-specific plug-in.

9.2 Developing Applications

Conventionally, in order to achieve parallel distributed processing of Big Data, complicated programs need to be created for synchronization processing and so on.

Under Hadoop, there is no need to consider parallel distributed processing when creating programs. The user just creates programs as two applications: applications that perform Map processing and Reduce processing in accordance with MapReduce algorithms. The distributed storage and extraction of data and the parallel execution of created processing is all left up to Hadoop.

9.2.1 Application Overview

The applications for performing processing under Hadoop include the following types:

- MapReduce application

Java programs that operate in the Hadoop MapReduce framework are developed using the Hadoop API.

- Hive query

These are queries written in an SQL-equivalent language (HiveQL) using Apache Hive, developed by The Apache Software Foundation, rather than using the Hadoop API.

- Pig script

Like Hive, these scripts are written using the Pig Latin language without using the Hadoop API.

- HBase application

The HBase API is used to develop Java programs that perform HBase data input-output and perform operations on the data in HBase.



Information

Installation directory

The applications are installed in the following directories on the master servers, slave servers, and development servers.

Application	Installation directory	Master server	Slave server	Development server
MapReduce	/usr/bin (command) /usr/share/hadoop (library) /etc/hadoop (setup file)	Y	Y	Y
Hive	/usr/lib/hive- <i>version</i>	N	N	Y

Application	Installation directory	Master server	Slave server	Development server
	/etc/hive (setup file) (*1)			
Pig	/usr/bin (command) /usr/share/pig (library) /etc/pig (setup file)	N	N	Y
Hbase	/usr/lib/hbase- <i>version</i> /etc/hbase (setup file) (*2)	Y	Y	Y

Y: Installed.

N: Not installed.

*1:/etc/hive is a symbolic link to /usr/lib/hive-*version*/conf

*2:/etc/hbase is a symbolic link to /usr/lib/hbase-*version*/conf.

The development of MapReduce applications is described below. Refer to the website and similar of the Apache Hadoop project for information on developing other applications.

9.2.2 Designing Applications

Design the application processing logic. The processing such as input file splitting, merge, and so on, which needs to be designed under conventional parallel distributed processing, does not need to be designed because that is executed by the Hadoop framework. Therefore, developers can concentrate on designing the logic required for jobs.

The application developer must understand the Hadoop API and design applications in accordance with the MapReduce framework. The main design tasks required are:

- Determining items corresponding to Key and Value
- Content of Map processing
- Content of Reduce processing

9.2.3 Creating Applications

Create applications based on the application design result.

MapReduce applications

As with the creation of ordinary Java applications, create a Java project and perform coding.

The Hadoop API can be used by adding the Hadoop jar file to the Eclipse build path.

Note that specification of `hadoop-core-xxx.jar` is mandatory (enter the Hadoop version at "xxx"). Specify other suitable Hadoop libraries in accordance with the Hadoop API being used.

9.2.4 References for Developing MapReduce Applications

Refer to the following information provided by Apache Hadoop for MapReduce application references:

If Hadoop API is used

- Hadoop project top page

<http://hadoop.apache.org/mapreduce/>

- Getting started
- Release page
- Documentation

- MapReduce project page

<http://hadoop.apache.org/mapreduce/>

- MapReduce tutorial

http://hadoop.apache.org/common/docs/r1.0.3/mapred_tutorial.html

Note: The samples and explanations in the above tutorial are based on the "org.apache.hadoop.mapred" packages, but these packages are not recommended for the current stable Hadoop version. Use the tutorial content for reference only since some of the content does not apply.

9.3 Registering Applications

Store the created Java program (jar file) in any directory on the development server.

Chapter 10 Executing and Stopping Jobs

This chapter explains how to use this product to execute, stop, and display the status of Hadoop jobs.

The commands provided by Hadoop can be used for operations such as executing and stopping Hadoop jobs and displaying their status.

Refer to the web page for the Apache Hadoop project for information about Hadoop, the system required to run Hadoop, tuning Java, and the command specifications for Hadoop.

Note that user account settings need to be changed if Hadoop jobs are to be executed using a user ID other than the one specified in the `BDPP_HADOOP_DEFAULT_USERS` parameter in `bdpp.conf` during installation. Refer to "[Chapter 11 Managing Job Execution Users](#)" for information on user account settings.

10.1 Executing Jobs

Using an account that can execute Hadoop jobs, use the `hadoop` command to execute MapReduce programs (Java programs). Use either the `jar` or `job` option.



Example

Command execution example:

```
$ hadoop jar sample.jar input output <Enter>
```

The above example uses the `jar` option to execute `sample.jar` containing a MapReduce program.

10.2 Stopping Jobs

Using an account that can execute Hadoop jobs, the `hadoop` command can be used to stop Hadoop jobs that are being executed. Use the `job` option.



Example

Command execution example:

```
$ hadoop job -kill jobId <Enter>
```

10.3 Displaying Job Status

Using an account that can execute Hadoop jobs, the `hadoop` command can be used to display and check the status of a Hadoop job during execution. Use the `job` option.

The example below shows the command to display a list of all Hadoop jobs (executing, completed, failed, and pending). This command checks, for example, the job ID, status, and start time.



Example

Command execution example:

```
$ hadoop job -list all <Enter>
```

Chapter 11 Managing Job Execution Users

This chapter explains the management of job execution users that perform Hadoop job operations.

11.1 Adding Job Execution Users

Hadoop job execution users are configured by specifying user names and user IDs in the `BDPP_HADOOP_DEFAULT_USERS` parameter of `bdpp.conf` when this product is installed.

This section describes how to add user accounts, during systems operations when using this product, for users who execute jobs but who are not specified in `bdpp.conf`.

The following table shows the tasks required for adding a job execution user.

Process	Permissions and conditions that can be implemented
Create a User Account	Use root permissions.
Create a Home Directory for the User on the DFS	Use root permissions.
Configure SSH Authentication Keys to be Used by the Cache Local Feature	Use the newly created user account. Perform only if using the cache local feature.
Configure MapReduce Job User Authentication Keys	Use the newly-created user account. Perform only if not using the cache local feature, but using job user authentication.

11.1.1 Create a User Account

Begin by creating a user account for the job execution user.

The user account must be created for all nodes comprising the DFS. The user name and the user ID must, therefore, be the same on all nodes. This is to achieve consistency at all nodes, because the DFS uses user IDs for control and Hadoop operates on the basis of user names.

Specify the group entered in the `BDPP_HADOOP_DEFAULT_GROUP` parameter of `bdpp.conf` as the group to which the newly created user account belongs.



Example

Create a user account for a user named `bdppuser1`.

The following conditions apply:

- User name: `bdppuser1`
- User ID: 1500
- Group name: `bdppgroup` (the group specified in the `BDPP_HADOOP_DEFAULT_GROUP` parameter)

```
# useradd -u 1500 -g bdppgroup bdppuser1 <Enter>
```

11.1.2 Create a Home Directory for the User on the DFS

After the user account has been created on all nodes, also create the home directory (`${pdfs.fs.local.homedir}/userName`) for the newly created user in the FileSystem class for the DFS.

Refer to "C.4 pdfs-site.xml" for information on the `pdfs.fs.local.homedir` property.

Example

Create a home directory for the user named `bdppuser1` on the DFS.

As the settings are shared at all nodes, execute as shown below at any of the nodes where the DFS is mounted.

```
# hadoop fs -mkdir /user/bdppuser1 <Enter>
# hadoop fs -chown bdppuser1:bdppgroup /user/bdppuser1 <Enter>
# hadoop fs -chmod 700 /user/bdppuser1 <Enter>
```

Set the group and permissions to suit the environment being used.

11.1.3 Configure SSH Authentication Keys to be Used by the Cache Local Feature

If MapReduce jobs that use the cache local MapReduce feature are executed, the DFS fetches the file cache information from all Hadoop cluster nodes.

As the MapReduce job execution user executes remote commands (SSH is used by default) to remote nodes from the job start node to fetch information at that time, settings that enable remote command execution must be set in advance.

Use the `pdfs.fs.local.cache.location` property to set whether or not the cache local feature is used. The default is "true" (enabled).

Note that the tasks below are not required if "false" (disabled) is set for the `pdfs.fs.local.cache.location` property.

Example

If the user is `bdppuser1` and the DFS mount directory is `/mnt/pdfs`:

1. As the settings are shared at all nodes, `bdppuser1` executes as shown below at any of the nodes where the DFS is mounted.

```
$ cd ~ <Enter>
$ ssh-keygen -t rsa -N "" -f id_hadoop <Enter>
$ echo -n "command=\"/opt/FJSVpdfs/sbin/pdfscachelocal.sh\",no-pty,no-port-forwarding,no-X11-forwarding" > authorized_keys <Enter>
$ cat id_hadoop.pub >> authorized_keys <Enter>
$ hadoop fs -mkdir .pdfs <Enter>
$ hadoop fs -chmod 700 .pdfs <Enter>
$ hadoop fs -moveFromLocal id_hadoop id_hadoop.pub authorized_keys .pdfs/ <Enter>
$ hadoop fs -chmod 600 .pdfs/id_hadoop .pdfs/authorized_keys <Enter>
$ hadoop fs -chmod 644 .pdfs/id_hadoop.pub <Enter>
```

An entry like the one below is set in the `authorized_keys` file shown above.

```
$ hadoop fs -cat .pdfs/authorized_keys <Enter>
"command=/opt/pdfs/sbin/pdfscachelocal.sh",no-pty,no-port-forwarding,no-X11-forwarding ssh-rsa
publicKey bdppuser1@localNodeName
```

2. Reflect the settings to `.ssh/authorized_keys` of the user `bdppuser1` home directory at all slave servers and development servers.

`bdppuser1` executes the following:

```
$ xargs -ti ssh {} "umask 0077; mkdir -p .ssh; cat /mnt/pdfs/hadoop/user/bdppuser1/.pdfs/authorized_keys >> .ssh/authorized_keys" < /etc/hadoop/slaves <Enter>
```

```
$ ssh develop "umask 0077; mkdir -p .ssh; cat /mnt/pdfs/hadoop/user/bdppuser1/.pdfs/authorized_keys >> .ssh/authorized_keys" <Enter>
```

3. The settings can be checked by performing remote execution as follows:

```
$ echo /dummy | ssh -o IdentityFile=/mnt/pdfs/hadoop/user/bdppuser1/.pdfs/id_hadoop -o StrictHostKeyChecking=no -o BatchMode=no remoteNodeName /opt/FJSVpdfs/sbin/pdfscachelocal.sh <Enter>
2 (*1)
```

*1: Normally "2" is output.

.....

11.1.4 Configure MapReduce Job User Authentication Keys

The tasks here are not required if "[11.1.3 Configure SSH Authentication Keys to be Used by the Cache Local Feature](#)" was implemented.

As authentication key checking is performed during the MapReduce job user authentication of the DFS, an authentication key file must be created in advance. Use the `pdfs.security.authorization` property in `pdfs-site.xml` to set whether or not job user authentication is used. The default is "false" (disabled).

Note that the tasks below are not required if "false" (disabled) is set for `pdfs.security.authorization`. Refer to "[C.4 pdfs-site.xml](#)" for information on the `pdfs.security.authorization` property.



Example

.....

If the user is `bdppuser1` and the DFS mount directory is `/mnt/pdfs`:

As the settings are shared at all nodes, `bdppuser1` executes as shown below at any of the nodes where the DFS is mounted.

```
$ cd ~ <Enter>
$ cat > id_hadoop <Enter>
anyKeywordString
ctrl-d
$ hadoop fs -mkdir .pdfs <Enter>
$ hadoop fs -chmod 700 .pdfs <Enter>
$ hadoop fs -moveFromLocal id_hadoop .pdfs/ <Enter>
$ hadoop fs -chmod 600 .pdfs/id_hadoop <Enter>
```

.....

11.2 Deleting Job Execution Users

This section describes how to delete a Hadoop job execution user that was added using the procedure described in "[11.1 Adding Job Execution Users](#)".

The procedure below is performed using root permissions.

11.2.1 Deleting the Home Directory of the User Created on the DFS

Delete the home directory of the user (`${pdfs.fs.local.homedir}/userName`) that was created in the `FileSystem` class for the DFS, as described in "[11.1.2 Create a Home Directory for the User on the DFS](#)". Refer to "[C.4 pdfs-site.xml](#)" for information on the `pdfs.fs.local.homedir` property.



Example

.....

If the user is `bdppuser1`:

As the settings are shared at all nodes, execute as shown below at any of the nodes where the DFS is mounted.
Execute using root permissions.

```
# hadoop fs -rmr /user/bdppuser1 <Enter>
```

* Specify the -rmr option to delete the specified directory and file.

11.2.2 Deleting the User Account

Delete the execution user account that is no longer required from all nodes comprising the DFS.



Example

If the user is bdppuser1:

```
# userdel -r bdppuser1 <Enter>
```

Use the above procedure to delete user bdppuser1 from all nodes comprising the DFS. In the above example, the -r option has been specified in the userdel command to simultaneously delete the home directory of the relevant user.

Chapter 12 Adding and Deleting Slave Servers

This chapter explains how to add and delete slave servers.

- [Adding Slave Servers](#)
- [Deleting Slave Servers](#)

Hadoop splits the data to be processed among multiple slave servers and processes it in parallel, enabling large-scale data to be processed in a short period of time. Therefore, the performance of systems that use this product can be improved by increasing the number of slave servers for distributed processing.

The processing load per server can be balanced by increasing the number of slave servers to achieve even better performance.

If the scale of data processed by an application developed using Hadoop increases, and improved performance is required, add a slave server.



Point

After changing the configuration by adding or deleting a slave server, but before restarting operations, ensure that you back up the configuration information for master servers, development servers, and collaboration servers.

Refer to "[14.1 Backup](#)" for details.

12.1 Adding Slave Servers

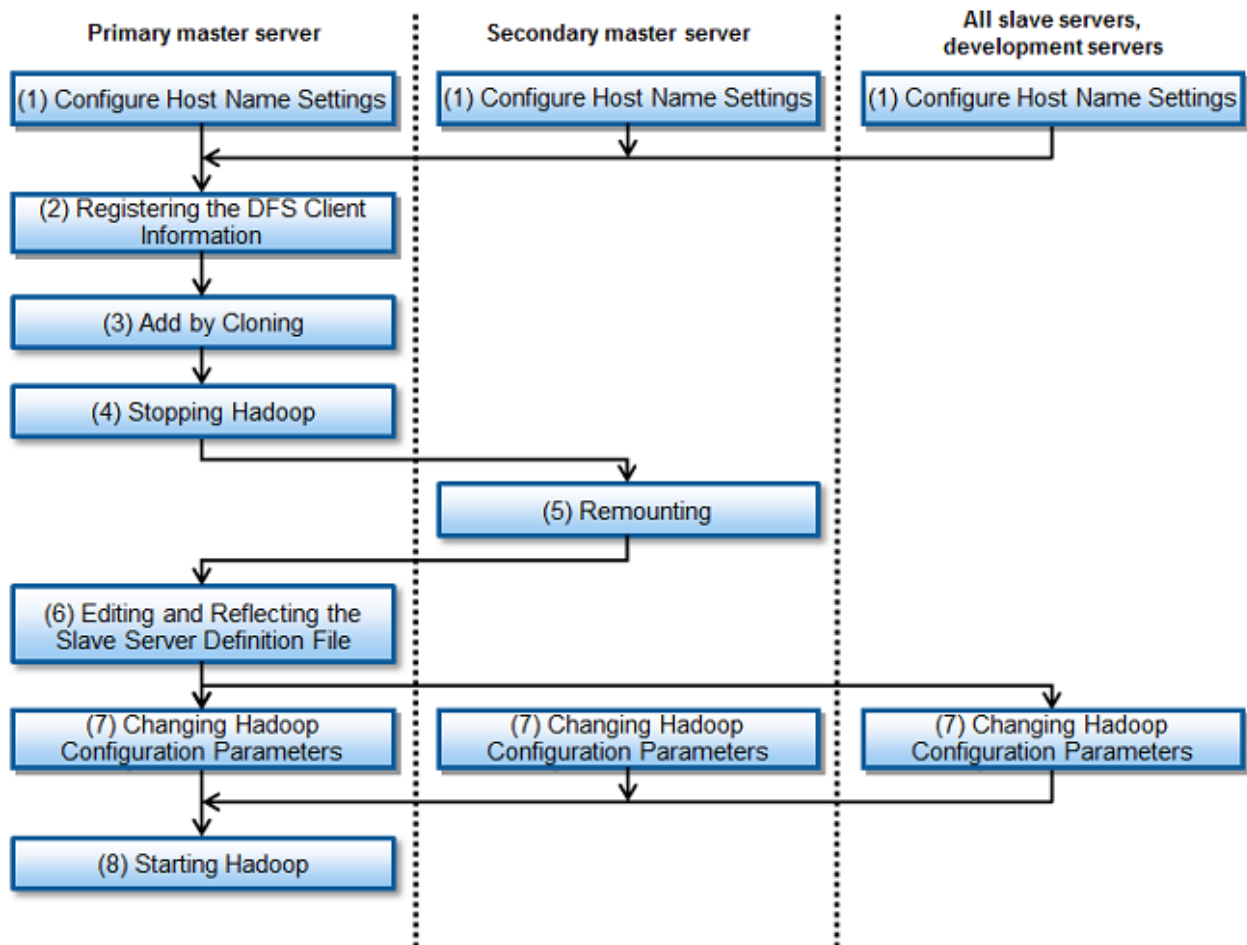
This section explains how to add slave servers after operations have started.

1. [Configuring Host Name Settings](#)
2. [Registering the DFS Client Information](#)
3. [Adding by Cloning](#)
4. [Stopping Hadoop](#)
5. [Remounting](#) (Only if replicated configuration is used for the master server)
6. [Editing and Reflecting the Slave Server Definition File](#)
7. [Changing Hadoop Configuration Parameters](#)
8. [Starting Hadoop](#)

The procedure for adding a slave server is shown below.

Use root permissions for all tasks in this procedure.

Note that the location for implementing the procedure outlined in "[12.1.3 Adding by Cloning](#)" will vary depending on the environment in which the additional slave server will be installed.



This section describes the procedure to install even more slave servers, in addition to those comprising the configuration built in "6.3 Adding Second and Subsequent Slave Servers" during initial installation, using the following environment as an example.

- Representative partition : /dev/disk/by-id/scsi-1FUJITSU_300000370106
- File system ID : 1
- Mount point : /mnt/pdfs
- Master servers : master1 (primary), master2 (secondary)
- Slave servers already built during initial installation : slave1, slave2, slave3, slave4, slave5
- Slave server to be newly installed : slave6
- Other slave servers to be additionally installed : slave7, slave8, slave9, slave10

12.1.1 Configuring Host Name Settings

Configure the host names of the slave servers being added in the master servers, slave servers, and development servers.

These settings must be configured for both the primary master server and the secondary master server if replicated configuration is used for master servers.



See

Refer to "5.6 Host Name Settings" for details.

12.1.2 Registering the DFS Client Information

Register the host names of the slave servers to be added in the connection authorization list file on the master servers.

If replicated configuration is used for master servers, update and register the connection authorization list on the primary master server and then distribute to the secondary master server.

1. Check the FSID column in the output generated by executing the pdfsinfo command, and confirm the file system ID of the file system to which the slave servers will be added.

```
# pdfsinfo -n /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
/dev/disk/by-id/scsi-1FUJITSU_300000370106:
FSID hostID status hostname
  1 80a4f75b RUN master1
  1 80960096 RUN master2
```

2. Add the CLIENT field of the slave server to be added at the end of the connection authorization list file that corresponds to the file system ID.

If adding multiple slave servers, add CLIENT fields for the number of clients.



Example

Authorizing the connection for slave servers to be added

```
# cat /etc/pdfs/server.conf.1 <Enter>
#
# Copyright (c) 2012 FUJITSU LIMITED. All rights reserved.
#
# /etc/pdfs/server.conf.<FSID>
#
# List of client hostnames of a file system.
#
# Notes:
# Do not describe hostnames of management servers.
#
# example:
#CLIENT nodeac1
#CLIENT nodeac2
#CLIENT nodeac3
#CLIENT nodeac4
#CLIENT nodeac5
CLIENT slavel1
CLIENT slave2
CLIENT slave3
CLIENT slave4
CLIENT slave5
CLIENT develop
CLIENT collaborate
CLIENT slave6 <- Slave server to be added
CLIENT slave7 <- Slave server to be added
CLIENT slave8 <- Slave server to be added
CLIENT slave9 <- Slave server to be added
CLIENT slave10 <- Slave server to be added
```



Note

- Ensure that the CLIENT fields for additional DFS clients are added at the end of the connection authorization list file. If existing CLIENT fields are changed, the mount command fails for the slave servers, development servers, and collaboration servers being added.

- Existing CLIENT fields cannot be changed or deleted.
If you want to make changes or deletions at the same time as adding slave servers, development servers, and collaboration servers, do so while the DFS is in the unmounted state.
- Ensure that the total number of slave servers, development servers, collaboration servers, and master servers does not exceed the maximum number, 128 servers, for the number of shared servers.

3. Distribute the updated connection authorization list file to the master server. (**Only if replicated configuration is used for the master server**)

```
# cd /etc/pdfs <Enter>
# scp -p ./server.conf.1 root@master2:/etc/pdfs/server.conf.1 <Enter>
```

Note

If the contents of the connection authorization list file differ at different master servers, MDS-down recovery may fail and file system inhibition or system panic may occur.

12.1.3 Adding by Cloning

Add slave servers by using the same procedure as used during initial installation, outlined in "[6.3 Adding Second and Subsequent Slave Servers](#)".

This section explains the difference between this and initial installation, such as the definition information used and other points.

Example

Adding slave servers to a physical environment

The example below shows the definitions for additional installations.

- Example of the cloning server definition file (clone.conf):

```
RCXCSV,V3.4
[Server]
operation,chassis_name,slot_no,server_name,ip_address,mac_address,second_mac_address,snmp_community_name,ipmi_ip_address,ipmi_user_name,ipmi_passwd,ipmi_passwd_enc,admin_lan1_nic_number,admin_lan2_nic_number,admin_lan_nic_redundancy
-,,slave1,192.168.1.21,A0:A0:A0:A0:A0:A1,,public,192.168.1.121,admin,admin,plain,,,OFF
-,,slave2,192.168.1.22,A0:A0:A0:A0:A0:A2,,public,192.168.1.122,admin,admin,plain,,,OFF
-,,slave3,192.168.1.23,A0:A0:A0:A0:A0:A3,,public,192.168.1.123,admin,admin,plain,,,OFF
-,,slave4,192.168.1.24,A0:A0:A0:A0:A0:A4,,public,192.168.1.124,admin,admin,plain,,,OFF
-,,slave5,192.168.1.25,A0:A0:A0:A0:A0:A5,,public,192.168.1.125,admin,admin,plain,,,OFF
new,,slave6,192.168.1.26,A0:A0:A0:A0:A0:A6,,public,192.168.1.126,admin,admin,plain,,,OFF <-( *1)
new,,slave7,192.168.1.27,A0:A0:A0:A0:A0:A7,,public,192.168.1.127,admin,admin,plain,,,OFF <-( *1)
new,,slave8,192.168.1.28,A0:A0:A0:A0:A0:A8,,public,192.168.1.128,admin,admin,plain,,,OFF <-( *1)
new,,slave9,192.168.1.29,A0:A0:A0:A0:A0:A9,,public,192.168.1.129,admin,admin,plain,,,OFF <-( *1)
new,,slave10,192.168.1.30,A0:A0:A0:A0:A0:B0,,public,192.168.1.130,admin,admin,plain,,,OFF <-( *1)
[ServerAgent]
operation,server_name
new,slave6 <-( *2)
```

*1: Server to be registered

*2: Server whose clone image is to be created

- Example of checking slave server registration

```
# /opt/FJSVbdpp/bin/bdpp_listserver <Enter>
PHYSICAL_SERVER SERVER ADMIN_IP STATUS MAINTENANCE
```

slave1	slave1	192.168.1.21	normal	OFF
slave2	slave2	192.168.1.22	normal	OFF
slave3	slave3	192.168.1.23	normal	OFF
slave4	slave4	192.168.1.24	normal	OFF
slave5	slave5	192.168.1.25	normal	OFF
slave6	slave6	192.168.1.26	normal	OFF
slave7	-	192.168.1.27	stop	-
slave8	-	192.168.1.28	stop	-
slave9	-	192.168.1.29	stop	-
slave10	-	192.168.1.30	stop	-

Note

Existing clone images cannot be used to build slave servers that are scheduled to be added. Create a new clone image when installing an additional slave server.

12.1.4 Stopping Hadoop

If Hadoop is running on this product prior to adding slave servers, execute the `bdpp_stop` command to stop it.

```
# /opt/FJSVbdpp/bin/bdpp_stop <Enter>
```

See

Refer to "[A.16 bdpp_stop](#)" for information on this command.

12.1.5 Remounting

Remount the DFS on the secondary master server.

Perform this step only if replicated configuration is used for the master server.

1. Unmount the DFS.

```
# pdfsumount /mnt/pdfs <Enter>
```

2. Mount the DFS.

```
# pdfsmount /mnt/pdfs <Enter>
```

See

Refer to the relevant commands in "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for information on the `pdfsumount` and `pdfsmount` commands.

12.1.6 Editing and Reflecting the Slave Server Definition File

The procedure for editing the slave server definition file, and then reflecting the changes is explained below.

Change the various definition files on the primary master server and distribute to all slave servers and development servers. If replicated configuration is used for master servers, also distribute to the secondary master server.

1. Edit the slave server definition file (/etc/opt/FJSVbdpp/conf/slaves), and define the slave servers to be added.

Example

Slave server definition file example:

```
slave1,slave2,slave3,slave4,slave5,slave6,slave7,slave8,slave9,slave10
```

Note: The underlining indicates the slave servers to be added.

2. Execute the bdpp_changeslaves command to reflect the changes made to the slave server definition file.

```
# /opt/FJSVbdpp/bin/bdpp_changeslaves <Enter>
```

3. Copy the edited file to the secondary master server. (Only if replicated configuration is used for the master server)

```
# scp -p /etc/opt/FJSVbdpp/conf/slaves root@master2:/etc/opt/FJSVbdpp/conf <Enter>
# scp -pr /etc/opt/FJSVbdpp/data root@master2:/etc/opt/FJSVbdpp/ <Enter>
# scp -p /etc/hadoop/mapred.include root@master2:/etc/hadoop <Enter>
```

4. Also, copy the edited file to slave servers and development servers.

See

- Refer to "[B.2 slaves](#)" for information on the slave server definition file.
- Refer to "[A.4 bdpp_changeslaves](#)" for information on this command.

12.1.7 Changing Hadoop Configuration Parameters

Change the Hadoop configuration in accordance with the changes made to the number of slave servers.

Make these changes on the primary master server, the secondary master server, slave servers, and development servers.

Edit /etc/hadoop/mapred-site.xml, and reconfigure the value in the mapred.reduce.tasks property.

See

Refer to "[C.3 mapred-site.xml](#)" for information on properties configured in mapred-site.xml.

12.1.8 Starting Hadoop

Start Hadoop on this product, and check its status.

1. Execute the bdpp_start command to start Hadoop on this product.

```
# /opt/FJSVbdpp/bin/bdpp_start <Enter>
```

2. Check the Hadoop status display to confirm if the TaskTracker has started on the corresponding slave server.

After confirming that the corresponding slave server system is running, execute the bdpp_stat command to display the status of Hadoop on this product.

```
# /opt/FJSVbdpp/bin/bdpp_stat -all <Enter>
cluster mapred 2420 jobtracker
slave1 mapred 29615 tasktracker
slave2 mapred 24417 tasktracker
slave3 mapred 16332 tasktracker
slave4 mapred 9042 tasktracker
slave5 mapred 5880 tasktracker
slave6 mapred 25126 tasktracker
slave7 mapred 18747 tasktracker
slave8 mapred 24467 tasktracker
slave9 mapred 6108 tasktracker
slave10 mapred 10425 tasktracker
bdpp: INFO : 003: bdpp Hadoop JobTracker is alive.
```



See

- Refer to "[A.14 bdpp_start](#)" for information on this command.
- Refer to "[A.15 bdpp_stat](#)" for information on this command.

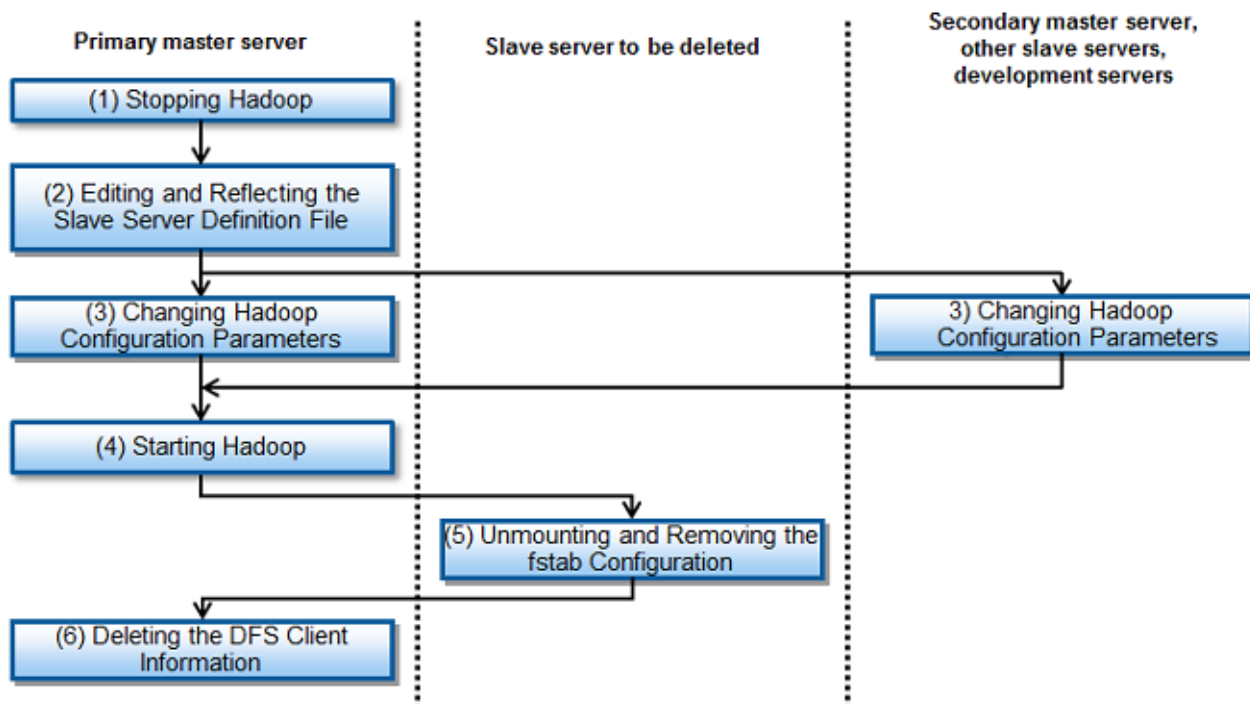
12.2 Deleting Slave Servers

This section explains how to delete slave servers.

1. [Stopping Hadoop](#)
2. [Editing and Reflecting the Slave Server Definition File](#)
3. [Changing Hadoop Configuration Parameters](#)
4. [Starting Hadoop](#)
5. [Unmounting and Removing the fstab Configuration](#)
6. [Deleting the DFS Client Information](#)

The procedure for deleting a slave server is shown below.

Use root permissions for all tasks in this procedure.



This section describes the procedure to delete a slave server from the configuration built in "6.3 Adding Second and Subsequent Slave Servers", using the following environment as an example.

- File system ID : 1
- Logical file system name : pdfs1
- Master servers : master1 (primary), master2 (secondary)
- Slave servers already built during initial installation : slave1, slave2, slave3, slave4, slave5
- Slave server to be deleted : slave3

12.2.1 Stopping Hadoop

If Hadoop is running on this product prior to deleting the slave server, execute the `bdpp_stop` command to stop it.

```
# /opt/FJSVbdpp/bin/bdpp_stop <Enter>
```



See

Refer to "A.16 `bdpp_stop`" for information on this command.

12.2.2 Editing and Reflecting the Slave Server Definition File

The procedure for editing the slave server definition file, and then reflecting the changes is explained below.

Change the various definition files on the primary master server and distribute to all slave servers and development servers. If replicated configuration is used for master servers, also distribute to the secondary master server.

1. Delete the slave server to be removed from the slave server definition file (`/etc/opt/FJSVbdpp/conf/slaves`).

2. Execute the `bdpp_changeslaves` command to reflect the changes made to the slave server definition file.

```
# /opt/FJSVbdpp/bin/bdpp_changeslaves <Enter>
```

3. Copy the edited file to the secondary master server. (Only if replicated configuration is used for the master server)

```
# scp -p /etc/opt/FJSVbdpp/conf/slaves root@master2:/etc/opt/FJSVbdpp/conf <Enter>
# scp -pr /etc/opt/FJSVbdpp/data root@master2:/etc/opt/FJSVbdpp/ <Enter>
# scp -p /etc/hadoop/mapred.include root@master2:/etc/hadoop <Enter>
```

4. Also, copy the edited file to slave servers and development servers.



See

- Refer to "[B.2 slaves](#)" for information on the slave server definition file.
- Refer to "[A.4 bdpp_changeslaves](#)" for information on this command.

12.2.3 Changing Hadoop Configuration Parameters

Change the Hadoop configuration in accordance with the changes made to the number of slave servers.

Perform the same procedure as outlined in "[12.1.7 Changing Hadoop Configuration Parameters](#)".

12.2.4 Starting Hadoop

Execute the `bdpp_start` command to start Hadoop on this product.

```
# /opt/FJSVbdpp/bin/bdpp_start <Enter>
```



See

Refer to "[A.14 bdpp_start](#)" for information on this command.

12.2.5 Unmounting and Removing the fstab Configuration

Perform the tasks listed below for the slave server to be deleted.

1. Unmount the DFS if it is mounted.

```
# umount pdfs1 <Enter>
```

2. Delete the DFS configuration information file.

```
# cd /etc/pdfs <Enter>
# rm ./client.conf.pdfs1 <Enter>
```

3. Remove the target DFS entries from `/etc/fstab`.

Relevant entries are those with a logical file system name as the first field.

12.2.6 Deleting the DFS Client Information

Delete the host name of the slave server to be deleted from the connection authorization list file on the master server.

If replicated configuration is used for master servers, update and register the connection authorization list on the primary master server and then distribute to the secondary master server.

1. From the primary master server connection authorization list file, delete the CLIENT field of the slave server targeted for deletion.



Example

Removing the connection authorization for the slave server to be deleted

```
# cat /etc/pdfs/server.conf.1 <Enter>
#
# Copyright (c) 2012 FUJITSU LIMITED. All rights reserved.
#
# /etc/pdfs/server.conf.<FSID>
#
# List of client hostnames of a file system.
#
# Notes:
# Do not describe hostnames of management servers.
#
# example:
#CLIENT nodeac1
#CLIENT nodeac2
#CLIENT nodeac3
#CLIENT nodeac4
#CLIENT nodeac5
CLIENT slavel
CLIENT slave2
CLIENT slave3      <- Delete the row here.
CLIENT slave4
CLIENT slave5
CLIENT develop
CLIENT collaborate
```

2. Distribute the updated connection authorization list file to the secondary master server. (Only if replicated configuration is used for the master server)

```
# cd /etc/pdfs <Enter>
# scp -p ./server.conf.1 root@master2:/etc/pdfs/server.conf.1 <Enter>
```



Note

If the contents of the connection authorization list file differ at different master servers, MDS-down recovery may fail and file system inhibition or system panic may occur.

Chapter 13 Adding and Deleting Storage Systems

This chapter explains how to add and delete storage systems.

- [Adding a Storage System](#)
- [Deleting a Storage System](#)

In a DFS, the storage area for data is built into shared disk devices.

To enhance the processing performance of a shared disk device for file data I/O, or to increase the file system capacity, add a new file data area to the shared disk device (an area to store the file data on the shared disk device).

Adding a new partition will increase the size of the file data area and enable larger volumes of data than previously to be stored on a DFS. When a DFS comprises multiple partitions, the file data I/O processing is distributed, which improves the performance of the shared disk device.



Point

After changing the configuration by adding or deleting a storage system, but before restarting operations, ensure that you back up the configuration information for master servers, development servers, and collaboration servers.

Refer to "[14.1 Backup](#)" for details.

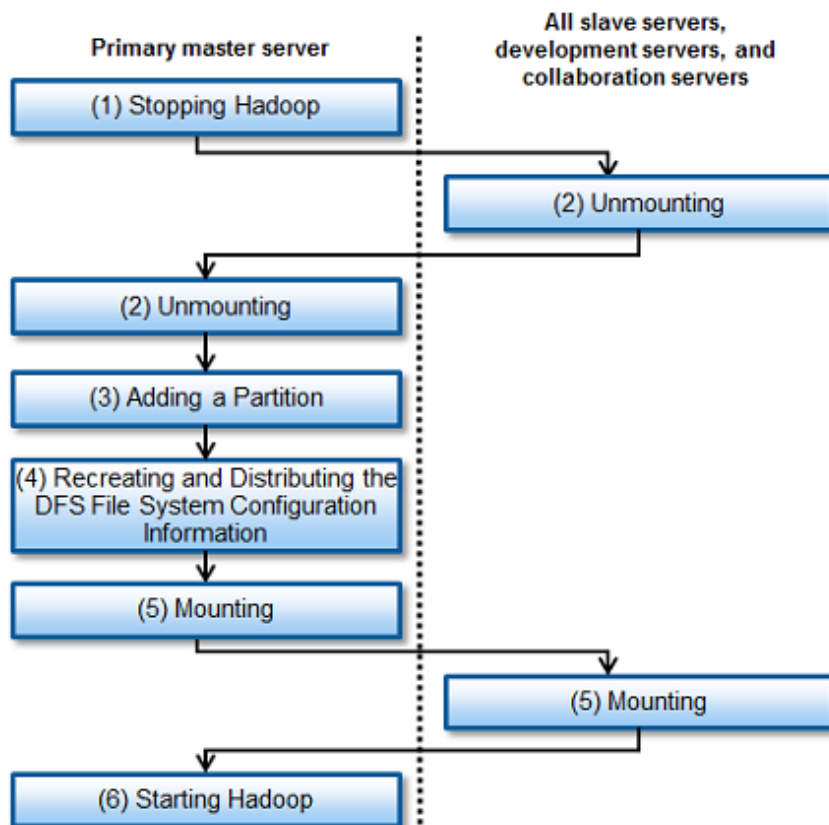
13.1 Adding a Storage System

This section explains how to add a shared disk.

1. [Stopping Hadoop](#)
2. [Unmounting](#)
3. [Adding a Partition](#)
4. [Recreating and Distributing the DFS File System Configuration Information](#)
5. [Mounting](#)
6. [Starting Hadoop](#)

The procedure for adding a shared disk is shown below.

Use root permissions for all tasks in this procedure.



This section describes the procedure for adding partitions to a DFS, using the following environment as an example.

- File system ID : 1
- Logical file system name : pdfs1
- Representative partition : /dev/disk/by-id/scsi-1FUJITSU_300000370106
- Existing partition : /dev/disk/by-id/scsi-1FUJITSU_300000370107
- Additional partition : /dev/disk/by-id/scsi-1FUJITSU_300000370108
- Master server : master1 (primary), master2 (secondary)
- Slave server : slave1, slave2, slave3, slave4, slave5
- Development server : develop
- Collaboration server : collaborate

13.1.1 Stopping Hadoop

If Hadoop is running on this product prior to adding the shared disk, execute the `bdpp_stop` command to stop it.

```
# /opt/FJSVbdpp/bin/bdpp_stop <Enter>
```



See

Refer to "A.16 `bdpp_stop`" for information on the `bdpp_stop` command.

13.1.2 Unmounting

If the targeted DFS is mounted, unmount it.

1. Unmount the DFS at all slave servers, development servers, and collaboration servers.

```
# umount pdfs1 <Enter>
```

2. Unmount the DFS at the primary master server.

```
# pdfsumntgl /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
```

13.1.3 Adding a Partition

Add a partition as the file data area.

1. Check the file system information.

From the file system information recorded in the management partition, check the configuration of the targeted file system.

```
# pdfsinfo /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
/dev/disk/by-id/scsi-1FUJITSU_300000370106:
FSID special                                size Type mount
  1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864)  25418 META -----
  1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864)   5120 LOG -----
  1 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 7341778 DATA -----
```

2. Add the partition.

Add /dev/disk/by-id/scsi-1FUJITSU_300000370108 as the file data area.

```
# pdfsadd -D /dev/disk/by-id/scsi-1FUJITSU_300000370108/dev/disk/by-id/
scsi-1FUJITSU_300000370106 <Enter>
```

3. Check that the partition was added.

From the file system information, check that the partition was added.

```
# pdfsinfo /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
/dev/disk/by-id/scsi-1FUJITSU_300000370106:
FSID special                                size Type mount
  1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864)  25418 META -----
  1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864)   5120 LOG -----
  1 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 7341778 DATA -----
  1 /dev/disk/by-id/scsi-1FUJITSU_300000370108 (896) 6578704 DATA -----
```

13.1.4 Recreating and Distributing the DFS File System Configuration Information

Recreate the DFS configuration information file on the master server.

Create the DFS configuration information file on the primary master server, and distribute to all slave servers, development servers, and collaboration servers.

1. Regenerate the DFS configuration information file.

Execute the pdfsmkconf command.

```
# pdfsmkconf <Enter>
```

Information

For each file system ID *fsid*, a DFS configuration information file is generated as `pdfsmkconf_out/client.conf.fsid` in the directory where the `pdfsmkconf` command was executed.

2. Convert the DFS configuration information file name to a logical file system name.

Convert the file system ID *fsid* to a logical file system name.

```
# cd pdfsmkconf_out <Enter>
# mv ./client.conf.1 client.conf.pdfs1 <Enter>
```

3. Distribute the DFS configuration information file to each slave server, development server, and collaboration server.

```
# scp ./client.conf.pdfs1 root@slave1:/etc/pdfs/client.conf.pdfs1 <Enter>
# scp ./client.conf.pdfs1 root@slave2:/etc/pdfs/client.conf.pdfs1 <Enter>
# scp ./client.conf.pdfs1 root@slave3:/etc/pdfs/client.conf.pdfs1 <Enter>
# scp ./client.conf.pdfs1 root@slave4:/etc/pdfs/client.conf.pdfs1 <Enter>
# scp ./client.conf.pdfs1 root@slave5:/etc/pdfs/client.conf.pdfs1 <Enter>
# scp ./client.conf.pdfs1 root@develop:/etc/pdfs/client.conf.pdfs1 <Enter>
# scp ./client.conf.pdfs1 root@collaborate:/etc/pdfs/client.conf.pdfs1 <Enter>
```

Note

Deploy the DFS configuration information file to each slave server, development server, and collaboration server under `/etc/pdfs`.

13.1.5 Mounting

Mount the DFS file system.

1. Mount the DFS, starting with the primary master server.

```
# pdfsmntgl /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
```

2. Mount the DFS at all slave servers, development servers, and collaboration servers.

```
# mount pdfs1 <Enter>
```

13.1.6 Starting Hadoop

Execute the `bdpp_start` command to start Hadoop on this product, and then restart operations.

```
# /opt/FJSVbdpp/bin/bdpp_start <Enter>
```

See

Refer to "[A.14 bdpp_start](#)" for information on the `bdpp_start` command.

13.2 Deleting a Storage System

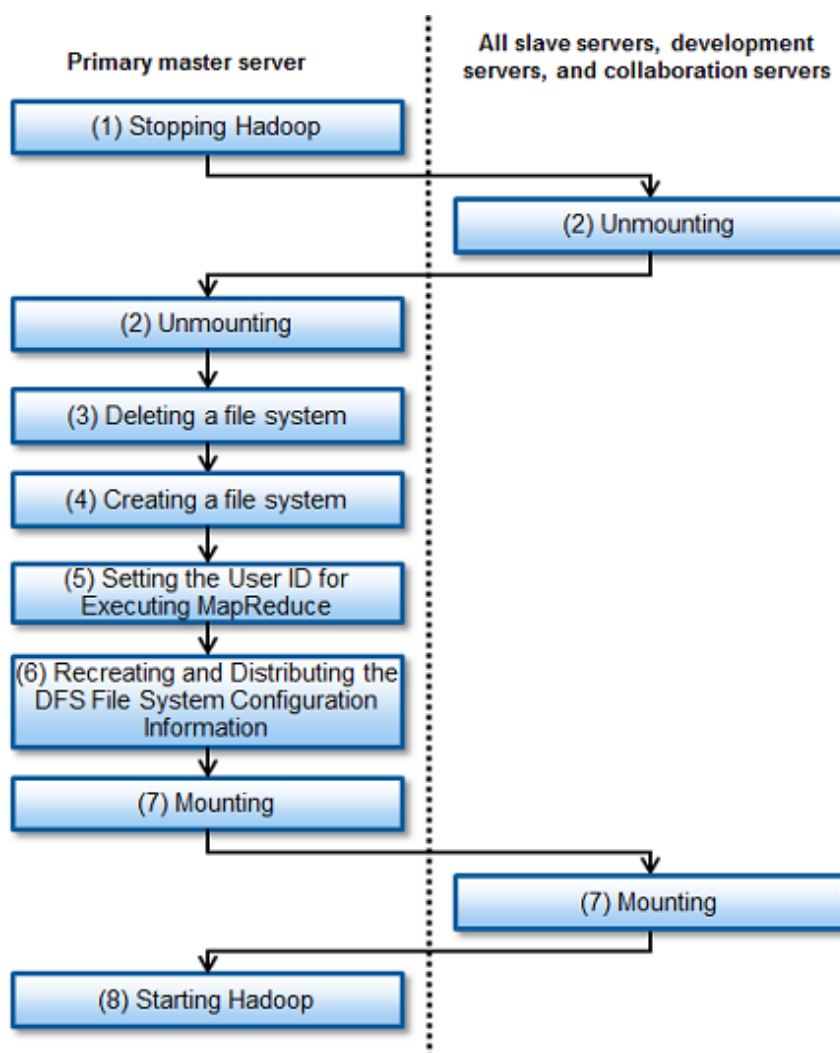
Under a DFS, the shared disks that comprise a file system cannot be deleted.

This section explains how to change the configuration of a DFS when a shared disk is deleted.

1. [Stopping Hadoop](#)
2. [Unmounting](#)
3. [Deleting a File System](#)
4. [Creating a File System](#)
5. [Setting the User ID for Executing MapReduce](#)
6. [Recreating and Distributing the DFS File System Configuration Information](#)
7. [Mounting](#)
8. [Starting Hadoop](#)

The procedure for changing a DFS configuration is shown below.

All configuration changes must be performed using root permissions.



This section describes the procedure for changing a configuration when deleting partitions from a DFS, using the following environment as an example.

- File system ID : 1
- Logical file system name : pdfs1
- Representative partition : /dev/disk/by-id/scsi-1FUJITSU_300000370106

- Existing partitions : /dev/disk/by-id/scsi-1FUJITSU_300000370107
/dev/disk/by-id/scsi-1FUJITSU_300000370108
- Partition to be deleted : /dev/disk/by-id/scsi-1FUJITSU_300000370108
- Master server : master1 (primary), master2 (secondary)
- Slave server : slave1, slave2, slave3, slave4, slave5
- Development server : develop
- Collaboration server : collaborate

Note that if any resources are required after changing the configuration, those resources must be backed up beforehand and then restored when the configuration change completes.

Perform backup in file/directory units.

Note

As shared disks are deleted, ensure that there is sufficient DFS space when backed up resources are restored.

See

Refer to "3.2.1 Backing-up and Restoring Using Standard Linux Commands (Files/Directories)" in the "Primesoft Distributed File System for Hadoop V1 User's Guide".

13.2.1 Stopping Hadoop

If Hadoop is running on this product prior to changing the DFS configuration, execute the `bdpp_stop` command to stop it.

```
# /opt/FJSVbdpp/bin/bdpp_stop <Enter>
```

See

Refer to "A.16 `bdpp_stop`" for information on the `bdpp_stop` command.

13.2.2 Unmounting

If the targeted DFS is mounted, unmount it.

1. Unmount the DFS at all slave servers, development servers, and collaboration servers.

```
# umount pdfs1 <Enter>
```

2. Unmount the DFS at the primary master server.

```
# pdfsumntgl /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
```

13.2.3 Deleting a File System

Delete the relevant DFS file system.

1. Check the current file system information.

```
# pdfsinfo -a <Enter>
/dev/disk/by-id/scsi-1FUJITSU_300000370106:
FSID special size Type mount
1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 25418 META -----
1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 5120 LOG -----
1 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 7341778 DATA -----
1 /dev/disk/by-id/scsi-1FUJITSU_300000370108 (896) 6578704 DATA -----
```

2. Delete the DFS.

```
# pdfsadm -D /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
```

3. Ensure that the DFS has been deleted.

```
# pdfsinfo -a <Enter>
```

13.2.4 Creating a File System

Refer to "[6.1.3.6 Creating the File System](#)" in the chapter that describes initial installation, and create the file system accordingly.



As shown in the example below, a user can change the configuration of the file system from which the relevant partition was deleted by recreating the file system using the pdfsmkfs command, omitting /dev/disk/by-id/scsi-1FUJITSU_300000370108 from the data option.

If replicated configuration is used for master servers:

```
# pdfsmkfs -o dataopt=y,blocksz=8388608,data=/dev/disk/by-id/
scsi-1FUJITSU_300000370107,node=master1,master2 /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
```

13.2.5 Setting the User ID for Executing MapReduce

Refer to "[6.1.3.7 Setting the User ID for Executing MapReduce](#)" in the chapter that describes initial installation, and set the user ID accordingly.

13.2.6 Recreating and Distributing the DFS File System Configuration Information

Recreate the DFS configuration information file on the master server.

Create the DFS configuration information file on the primary master server, and distribute to all slave servers, development servers, and collaboration servers.

1. Regenerate the DFS configuration information file.

Execute the pdfsmkconf command.

```
# pdfsmkconf <Enter>
```


Information

For each file system ID *fsid*, a DFS configuration information file is generated as `pdfsmkconf_out/client.conf.fsid` in the directory where the `pdfsmkconf` command was executed.

2. Convert the DFS configuration information file name to a logical file system name.

Convert the file system ID *fsid* to a logical file system name.

```
# cd pdfsmkconf_out <Enter>
# mv ./client.conf.1 client.conf.pdfsl <Enter>
```

3. Distribute the DFS configuration information file to each slave server, development server, and collaboration server.

```
# scp ./client.conf.pdfsl root@slave1:/etc/pdfs/client.conf.pdfsl <Enter>
# scp ./client.conf.pdfsl root@slave2:/etc/pdfs/client.conf.pdfsl <Enter>
# scp ./client.conf.pdfsl root@slave3:/etc/pdfs/client.conf.pdfsl <Enter>
# scp ./client.conf.pdfsl root@slave4:/etc/pdfs/client.conf.pdfsl <Enter>
# scp ./client.conf.pdfsl root@slave5:/etc/pdfs/client.conf.pdfsl <Enter>
# scp ./client.conf.pdfsl root@develop:/etc/pdfs/client.conf.pdfsl <Enter>
# scp ./client.conf.pdfsl root@collaborate:/etc/pdfs/client.conf.pdfsl <Enter>
```

Note

Deploy the DFS configuration information file to each slave server, development server, and collaboration server under `/etc/pdfs`.

13.2.7 Mounting

Mount the DFS file system.

1. Mount the DFS, starting with the primary master server.

```
# pdfsmntgl /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
```

2. Mount the DFS at all slave servers, development servers, and collaboration servers.

```
# mount pdfsl <Enter>
```

13.2.8 Starting Hadoop

Execute the `bdpp_start` command to start Hadoop on this product, and then restart operations.

```
# /opt/FJSVbdpp/bin/bdpp_start <Enter>
```

See

Refer to "[A.14 bdpp_start](#)" for information on the `bdpp_start` command.

Chapter 14 Backup and Restore

This chapter explains how to back up and restore the system configuration.

If a severe fault such as hardware failure occurs on a server, use the backup and restore functionality to recover.

Note

- The backup and restore functionality described in this chapter enables the recovery of the system configuration or the definition information. It cannot, however, recover Hadoop applications or job data.
- Backing up and restoring master servers that use replicated configuration
If a fault occurs on either the primary master server or the secondary master server, the master server system configuration can be recovered.
If a fault occurs on both the primary master server and the secondary master server, the restore functionality cannot be used to rebuild the master servers, so create a backup of the system if required.

14.1 Backup

Back up the system configuration and the definition information for a server when making changes to the configuration, in case hardware failure such as a corrupt disk causes problems.

Point

- For master servers, development servers, and collaboration servers, use the backup command `bdpp_backup`. Refer to "A.2 `bdpp_backup`" for information on this command.
- For slave servers, use the clone image of an installed slave server for backup. Note that slave servers in a virtual environment do not need to be backed up.

14.1.1 Resources Saved when the Backup Command is Executed

This section explains the definition information and the configuration information that is saved when the backup command is executed.

Resources saved when the backup command is executed are shown below.

Table 14.1 Overview of types of resources backed up on each server

Resource type	Resource overview	Master server		Development server	Collaboration server
		Primary	Secondary		
OS configuration-related information	Information configured separately for each server OS, such as the hosts file and <code>bdpp.conf</code> , which is the configuration file of this product.	Y	Y	Y	Y
Image file-related information	Image files such as slave server clone images and the related management information.	Y (*1)	N	N	N
DFS configuration information	Information relating to the configuration and definitions of the DFS, such as the DFS file	Y	Y	Y	Y

Resource type	Resource overview	Master server		Development server	Collaboration server
		Primary	Secondary		
	system configuration information.				
Configuration information for this product	Information relating to the configuration and definitions for systems that use this product, such as the slave server definitions file.	Y	Y	Y	N
Hadoop configuration information	Information relating to the configuration and definitions for Hadoop on this product, such as the definitions file for configuring Hadoop parameters.	Y	Y	Y	N

Y: Backed up

N: Not backed up

*1: Clone image backup on the primary master server can be skipped by executing the `bdpp_backup` command with the `-q` option specified. Note that the image file-related information is not backed up in a virtual environment.

Backup storage directory

When the `bdpp_backup` command is executed, a parent directory named `FJSVbdpp-backup` is created in the specified backup destination directory. The saved resources are then classified by type and stored in this parent directory.

The table below shows the types of resources (backup) stored in the directories that are created.

Table 14.2 Directories created under the `FJSVbdpp-backup` directory and the resources stored

Directory name	Stored resource type
SYS	OS-related information
ROR	Image file-related information
PDFS	DFS configuration information
BDPP	Configuration information for this product
Hadoop	Hadoop configuration information



Note

The structure of directories created under `FJSVbdpp-backup` varies for each server, because the resources to be backed up differ for the primary master server, the secondary master server, development servers, and collaboration servers.

Backup size

The table below provides an estimate of the size of resources saved at backup (the capacity of the `FJSVbdpp-backup` directory).

Check the available disk space on the relevant server before performing backup to prevent backup failure due to insufficient disk space.

Table 14.3 Disk space required on each server

Execution environment	Disk space (MB)	Remarks
Primary master server	10 + <i>cloneImageFileSize</i>	The approximate disk space required when creating a clone image file.

Execution environment	Disk space (MB)	Remarks
		If not creating a clone image file, use 10 MB as an estimate.
Secondary master server	10	-
Development server	10	-
Collaboration server	10	-

Note

When the `bdpp_backup` command is executed on the primary master server, separate disk space is required temporarily for tasks. Refer to "[Master Server]" in "3.2.3.2 Dynamic Disk Size" for information on the temporarily required disk space.

Information

Obtaining the clone image file size

This section explains how to obtain the clone image file size.

1. Log in to the master server.
2. Execute the `bdpp_listimage` command, and confirm the current storage location of the clone image file.
3. Specify the directory confirmed at Step 2, and execute the `du` command.

The following example shows `/var/opt/FJSVscw-deploysv/depot` as the storage area of the clone image file.

```
# du -hs /var/opt/FJSVscw-deploysv/depot <Enter>
```

14.1.2 Backup Method

This section explains how to back up a server.

- [Backing Up a Master Server, Development Server, or Collaboration Server](#)
- [Backing Up a Slave Server](#)

14.1.2.1 Backing Up a Master Server, Development Server, or Collaboration Server

This section explains how to back up a primary or secondary master server, development server, or collaboration server.

To back up a server other than a slave server, execute the `bdpp_backup` command on the relevant server.

Note

Execution of the `bdpp_backup` command will not back up MapReduce applications developed on the development server, nor existing business systems or business data on the collaboration server. These must be backed up and saved separately by the user.

Backup procedure

Stopping Hadoop

1. Log in to the primary master server using root permissions.

2. Execute the `bdpp_stop` command to stop Hadoop.

Perform backup

1. Log in to the server to be backed up using root permissions.
2. Execute the `bdpp_backup` command.

Example

If storing the backup under `/var/backup` on the primary master server and not backing up the clone image:

```
# /opt/FJSVbdpp/bin/bdpp_backup -d /var/backup -q <Enter>
```

Note

- If execution of the `bdpp_backup` command fails, refer to the output message and check its meaning and corrective action, or if required, refer to the backup log (`/var/opt/FJSVbdpp/log/bdpp_backup.log`) and remove the cause of the failure.
- If a backup storage directory and subdirectories were created when the previous backup failed, delete them all and then re-execute the backup command.
- If an option is not specified when the `bdpp_backup` command is executed on the primary master server, the backup will include clone images. Note that in this case, the time required for backup may increase in proportion to the file size of the clone image multiplied by the number of images, because copying the files (disk I/O) is time-consuming.
Use the `bdpp_removeimage` command to remove any unnecessary clone images prior to performing backup. If required, specify the `-q` option to skip the backup of clone images altogether.
Refer to "[A.11 bdpp_removeimage](#)" for details.

Archive the backup created by using the backup command, and save it on other media such as another server or on tape.

Information

Archiving backup

The following example shows how the `tar` command is used to archive the backup storage directory created under `/var/backup`, as well as its subdirectories.

```
# cd /var/backup <Enter>
# tar cvf FJSVbdpp-backup-master1-20121010.tar FJSVbdpp-backup <Enter>
```

14.1.2.2 Backing Up a Slave Server

This section explains how to back up a slave server.

A slave server is backed up by creating a clone image of the slave server.

Create the clone image from a slave server that is already installed.

Refer to "[6.3.1.4 Creating a Clone Image](#)" for information on how to create a clone image.

Point

Backup is not required for slave servers installed in a virtual environment.

14.2 Restore

If a fault such as hardware failure occurs on a system using this product, as a result requiring the server to be rebuilt, use the restore functionality to recover.

Point

- For master servers, development servers, or collaboration servers, use the `bdpp_restore` command to restore from backup. Refer to "A.13 `bdpp_restore`" for information on this command.
 - For slave servers, restore using the cloning functionality.
-

14.2.1 Restore Method

This section explains how to restore a server.

- [Restoring a Master Server, Development Server, or Collaboration Server](#)
- [Restoring a Slave Server](#)

14.2.1.1 Restoring a Master Server, Development Server, or Collaboration Server

This section explains how to restore a primary or secondary master server, development server, or collaboration server.

To restore a server other than a slave server, execute the `bdpp_restore` command on the relevant server.

Note

- Development server
Execution of the `bdpp_restore` command will not restore other items such as MapReduce applications developed on the development server.
 - Collaboration server
Execution of the `bdpp_restore` command will not back up business systems or business data on the collaboration server.
-

Preparing to perform restore

The preparatory tasks described below must be performed on the relevant server prior to performing restore.

Deploy backup

Deploy the backup that will be used to perform restore to any directory on the server to be restored.

If the backup has been archived, decompress it.

Example

If backup is archived in tar format

The following example shows how the `tar` command is used to decompress the backup that was deployed to `/var/backup` (`FJSVbdpp-backup-master1-20121010.tar`).

```
# cd /var/backup <Enter>
# tar xvf FJSVbdpp-backup-master1-20121010.tar <Enter>
```

Prepare to build the system

Refer to "[Chapter 5 Preparing to Build the System](#)", and configure the required settings on the relevant server.

Note that there is no need to perform the tasks specified in "[5.6 Host Name Settings](#)", because the hosts file included in the backup will be used. Replace the hosts file located on the relevant server with the hosts file in the backup storage directory (FJSVbdpp-backup/SYS).



Example

Replacing the hosts file

The following example shows how the cp command is used to replace the hosts file located on the relevant server with the hosts file in the backup storage directory that was deployed to /var/backup.

```
# cp -p /var/backup/FJSVbdpp-backup/SYS/hosts /etc <Enter>
```

Register the hadoop group and the mapred user (only if a collaboration server is restored)

Refer to "[6.5.2.2 Registering the hadoop Group and mapred User](#)", and register the hadoop group and the mapred user on the collaboration server. These tasks are only performed when a collaboration server is restored.

Restore procedure

This section explains the restore procedure. Ensure that the tasks outlined in "[Preparing to perform restore](#)" have been performed in advance.

Stop Hadoop (only if a master server that uses replicated configuration will be restored)

1. Log in to the master server that will not be restored using root permissions.
2. Execute the bdpp_stop command to stop Hadoop.



Note

If the master server uses replicated configuration, Hadoop may be running. If so, stop Hadoop.

Perform restore

1. Log in to the server to be restored using root permissions.
2. Deploy bdpp.conf located in the backup storage directory (FJSVbdpp-backup/SYS) to any directory.
3. Install the functionality for the relevant server.
 - On a master server
Refer to "[6.1.1.2 Installing the Master Server Functionality](#)", and install this functionality. Then, perform the HA cluster setup only on the master server to be restored. Refer to "[6.1.2 HA Cluster Setup](#)" for the setup procedure.
 - On a development server
Refer to "[6.4.1.2 Installing the Development Server Functionality](#)", and install this functionality.
 - On a collaboration server
Refer to "[6.5.1.2 Installing the Collaboration Server Functionality](#)", and install this functionality.
4. Execute the bdpp_restore command on the relevant server.

Example

The following example shows the backup storage directory deployed to /var/backup.

```
# /opt/FJSVbdpp/bin/bdpp_restore -d /var/backup <Enter>
```

Note

- Ensure that the configuration of the backup storage directory and its subdirectories is the same as when the backup was created.
- If the backup includes clone images, these will be decompressed (copied) when restore is executed on the primary master server. Note that in this case, the time required for restore may increase in proportion to the file size of the clone image multiplied by the number of images, because copying the files (disk I/O) is time-consuming.
- Ensure that there is sufficient disk space on the server to be restored prior to performing the restore. Refer to "[Table 14.3 Disk space required on each server](#)" for details.
- If job execution users were added using the method outlined in "[11.1 Adding Job Execution Users](#)", the settings for these users will need to be configured again on the server that will be restored. Refer to "[11.1 Adding Job Execution Users](#)" for information on how to configure these users.
- If execution of the bdpp_restore command fails, refer to the output message and check its meaning and corrective action, or if required, refer to the backup log (/var/opt/FJSVbdpp/log/bdpp_restore.log), remove the cause of the failure, and re-execute the command.

14.2.1.2 Restoring a Slave Server

This section explains how to restore a slave server.

Slave servers are restored using the cloning functionality.

Point

- Restoring a slave server in a physical environment
Clone using the clone image that was created beforehand.
- Restoring a slave server in a virtual environment
Clone from a slave server (virtual machine) that is already installed.

Preparing to perform restore

The preparatory tasks described below must be performed prior to cloning.

Preparing clone images (in a physical environment)

Confirm that the clone image to be used for restore is located in the clone image storage directory on the primary master server.

If the clone image has not yet been prepared, refer to "[6.3.1.4 Creating a Clone Image](#)" and create a clone image from a slave server that has been installed.

Checking the clone source slave server (in a virtual environment)

The network parameter and iSCSI name automatic configuration must be registered in the clone source slave server. Accordingly, check the clone source slave server that was used when the slave server to be restored was added.

If there is no clone source slave server, or if the clone source slave server itself will be restored, perform the steps outlined in "[6.3.2.2 Registering the Network Parameter and iSCSI Name Automatic Configuration Feature](#)" to ensure that an existing slave server can be used as the clone source.

Restore procedure

Rebuild the slave server to be restored using the cloning functionality.

- For a physical environment, perform the steps outlined in "[6.3.1.6 Cloning](#)".
- For a virtual environment, perform the steps outlined in "[6.3.2.3 Cloning](#)".



Restoring a slave server to a physical environment

If the clone image used for the restore was created from a configuration older than the slave server currently operating, the files below must be obtained from the primary master server after cloning, and then deployed to the slave server that was restored.

- DFS file system configuration information, `client.conf.fsid`
 - Slave server definitions file, `/etc/opt/FJSVbdpp/conf/slaves`
 - Hadoop configuration parameters, `/etc/hadoop/mapred-site.xml`
-

Chapter 15 Operations when There are Errors

This chapter describes the corrective action to take when an error occurs on a system that uses this product.

Possible errors that may occur with this product are shown below.

Error details	Possible issue
System error (physical machine) (*1)	<ul style="list-style-type: none"> - System panic occurred - System has stopped due to forced power off or other cause - System has stopped responding
System error (virtual machine)	<ul style="list-style-type: none"> - Virtual machine panic occurred - Virtual machine has stopped responding
Public LAN network error (*2)	<ul style="list-style-type: none"> - Hardware error such as a faulty NIC or cable - An error has occurred in the public LAN transmission route (*3)
Cluster interconnect (CIP) error (*2)	<ul style="list-style-type: none"> - Hardware error such as a faulty NIC or cable - Heartbeat monitoring detected an error between the primary master server and the secondary master server
iSCSI network error (*2)	<ul style="list-style-type: none"> - Hardware error such as a faulty NIC or cable - Error has occurred in the iSCSI-LAN transmission route
JobTracker error	<ul style="list-style-type: none"> - JobTracker process has ended in an error - JobTracker process was stopped by a means other than the bdpp_stop command (*4)

*1: Indicates an error on the physical environment server or on the virtual environment host machine.

*2: If a LAN uses redundancy, indicates that an error has occurred on both LANs.

*3: Errors are detected differently depending on the feature used for redundancy.

*4: Indicates the direct use of the Apache Hadoop feature to stop the JobTracker.

The following sections explain how these errors affect each server, and the corresponding action that should be taken.

- [Operations when Errors Occur on a Master Server](#)
- [Operations when Errors Occur on a Slave Server](#)
- [Operations when Errors Occur on a Development Server](#)
- [Operations when Errors Occur on a Collaboration Server](#)

This chapter also contains a section that describes how to perform operations when errors occur on a file system.

Additionally, reference information on checking error occurrences is provided.

- [Operations when Errors Occur on the File System](#)
- [How to Check Errors](#)



See

- Refer to "2.4 Monitoring function of NIC switching mode" in the "PRIMECLUSTER Global Link Services Configuration and Administration Guide 4.3 Redundant Line Control Function" for information on detecting public LAN errors in network redundancy

software (redundancy in a physical environment).
 If required, also refer to "[Checking network replication status](#)".

- Refer to the manual for the virtualization software product you are using for information on the NIC teaming feature (redundancy in a virtual environment).

15.1 Operations when Errors Occur on a Master Server

This section explains operations when an error occurs on a master server using replicated and non-replicated configurations.

15.1.1 If the Master Servers Use Replicated Configuration

If an error occurs on the primary master server, operations will switch to the secondary master server as shown below in accordance with the environment where the master server is installed.

Error details	Installation environment				
	Physical environment	Virtual environment (same host machine)		Virtual environment (different host machine)	
		KVM	VMware	KVM	VMware
System error (physical machine)	Y1	N	N	N (*1)	Y1
System error (virtual machine)	N	Y1	Y1	Y1	Y1
Public LAN network error	Y1	Y1	Y1	Y1	Y1
Cluster interconnect (CIP) error	Y2	Y2	Y2 (*1)	Y2	Y2 (*1)
iSCSI network error	Y1	Y1	Y1	Y1	Y1
JobTracker error	Y1	Y1	Y1	Y1	Y1

Y1: Switches to the secondary master server.

Y2: Continues tasks on the primary master server (switch not required).

N: Does not switch to the secondary master server.

*1: The secondary master server must also be forcibly stopped separately.

Note

- If an error occurs on the secondary master server, there will be no switch to the primary master server. Tasks can continue without the need to re-execute tasks on the primary master server.
 In this situation, investigate and remove the cause of the error on the secondary master server by referring to the system log. After the secondary master server has fully recovered, restart the secondary master server.
- If the primary master server and the secondary master server are both installed on each virtual machine on the same host machine, tasks will stop if an error occurs on the host machine.
 In this situation, refer to the system log of the host machine and remove the cause of the error. After the host machine has fully recovered, refer to "[8.1 Starting](#)" and restart the system that this product is installed in.

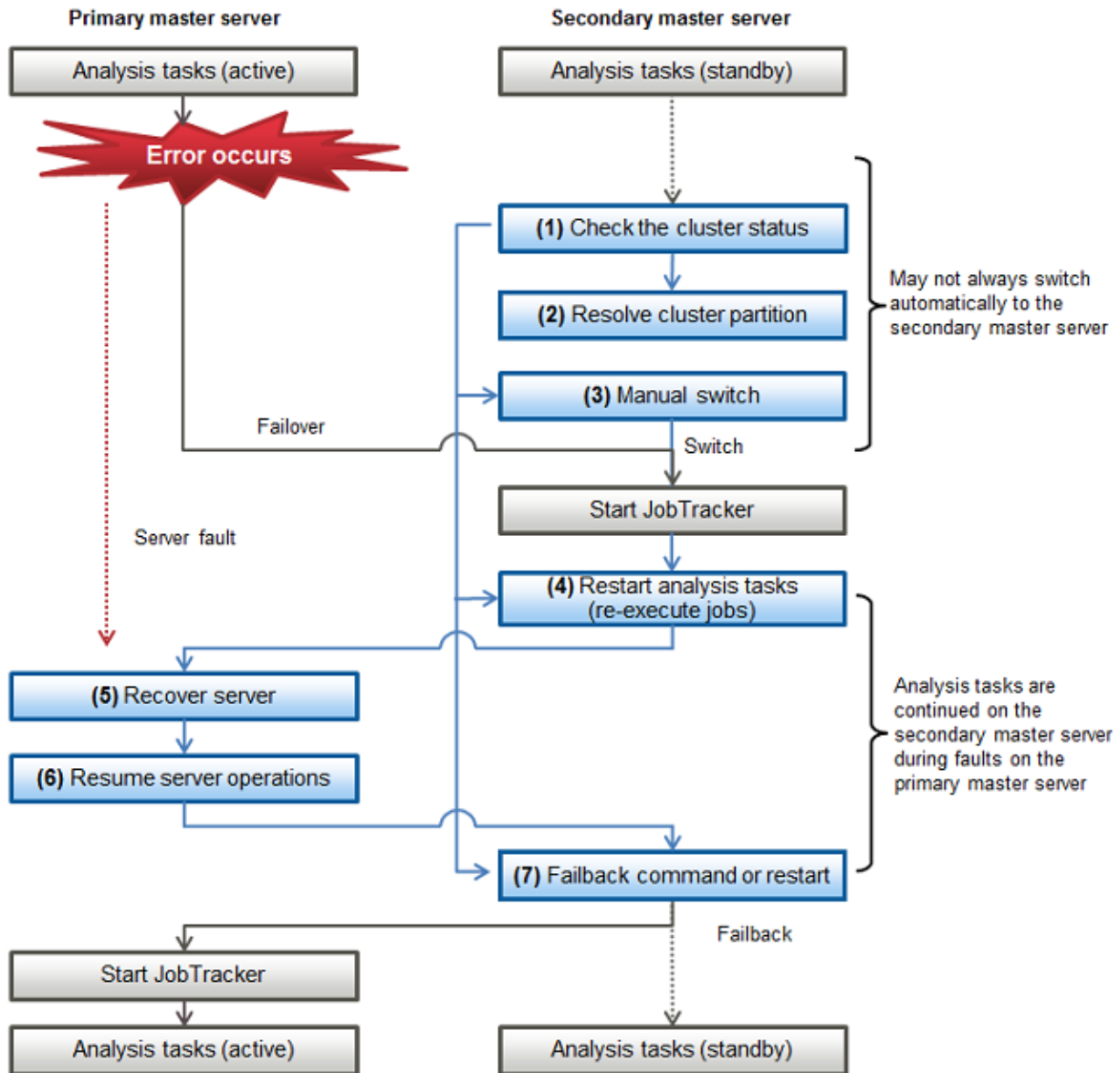
Information

If a system error occurs on the host machine in a virtual environment, there will be no switch to the secondary master server (some exceptions apply). Refer to "2.2.1 Virtual Machine Function" in "PRIMECLUSTER Installation and Administration Guide 4.3" for information on these operations, and refer to the articles on the configurations listed below:

- If a cluster system is built between the guest operating systems on a single host operating system
- If a cluster system is built between the guest operating systems on multiple host operating systems, without using the feature for switching when an error occurs on a host operating system

The following sections describe the corrective actions to take after an error occurs on the primary master server.

Figure 15.1 Procedure to resume tasks after an error occurs in a replicated configuration



(1) Check cluster status

If an error occurs on the primary master server, check that operations have switched to the secondary master server.

Note that switching will cause the status to transition, so wait 5 minutes after the error occurs before checking the status.

1. Refer to "[Checking the status of the master server HA cluster](#)", and check if operations have switched to the secondary master server.

If operations have switched, check the status of the primary master server.

- If the status of the primary master server is "Faulted" (or if the system is not running)
Continue operations on the secondary master server until the primary master server has recovered. Proceed to Step (4) "Restart analysis tasks (re-execute jobs)".
 - If the status of the primary master server is "Offline"
Failback to the primary master server is possible in this state. Proceed to Step (7) "Failback command or restart".
2. If switching has not occurred, refer to "[Checking the status of communication between master servers](#)" for information on how to check the communication status between the primary and secondary master servers.
 - If the status of the remote side of both master servers is "LEFTCLUSTER"
A cluster partition has occurred. Proceed to Step (2) "Resolve cluster partition".
 - If the status of the remote side of the secondary master server is "LEFTCLUSTER"
A switch to the secondary master server must be performed manually. Proceed to Step (3) "Manual switch".

(2) Resolve cluster partition

Tasks can be continued on the primary master server, but the cluster partition must be resolved.

1. Remove the fault on the cluster interconnect (CIP).
2. Restart the secondary master server.

If a master server has been installed in a physical environment or a virtual environment (KVM):

```
# shutdown -r now <Enter>
```

If a master server has been installed in a virtual environment (VMware) (requires the system to be forcibly stopped):

```
# reboot -f <Enter>
```

3. Refer to "[Checking the status of communication between master servers](#)", and check the communication status between the primary and secondary master servers. Confirm that the communication status has recovered (the status of both master servers is "UP (starting)").

This action will return the servers to their normal operating status (no subsequent steps required).

(3) Manual switch

Operations must be switched to the secondary master server, because a system error has occurred on the primary master server.

1. Unmount the DFS on all slave servers, development servers, and collaboration servers.

Example

If the logical file system name of the DFS is pdfs1:

```
# umount pdfs1 <Enter>
```

2. Forcibly stop the system on the secondary master server, and restart.

```
# reboot -f <Enter>
```

3. Forcibly start the pdfsfrmd daemon on the secondary master server.

```
# pdfsfrmstart -f <Enter>
```

4. Confirm that the status of the DFS on the secondary master server is "run".

Example

If the DFS management server (MDS) has been switched to the secondary master server:

```
# pdfsrscinfo -m <Enter>
/dev/disk/by-id/scsi-1FUJITSU_300000370106:
FSID  MDS/AC  STATE  S-STATE  RID-1  RID-2  RID-N  hostname
  1  MDS(P)  stop   -         0       0       0  master1
  1  AC      stop   -         0       0       0  master1
  1  MDS(S)  run    -         0       0       0  master2 <- (*1)
  1  AC      run    -         0       0       0  master2
```

*1: Status of the secondary master server is "run"

Note

If the master server is configured so that the DFS is not mounted automatically when the master server is started, mount the DFS manually.

5. Mount the DFS on all slave servers, development servers, and collaboration servers.

Example

If the logical file system name of the DFS is pdfs1:

```
# mount pdfs1 <Enter>
```

6. Forcibly start Hadoop by executing the bdpp_start command on the secondary master server.

```
# /opt/FJSVbdpp/bin/bdpp_start -f <Enter>
```

7. Refer to "[Checking the status of the master server HA cluster](#)", and check if operations have switched to the secondary master server.

See

- Refer to the relevant command in "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for information on the pdfsfrmstart and pdfsrscinfo commands.
- Refer to "[A.14 bdpp_start](#)" for information on this command.

(4) Restart analysis tasks (re-execute jobs)

If an error occurs on the primary master server during execution of jobs and operations switch to the secondary master server, these jobs may be interrupted. Check the job execution status and if required, re-execute the jobs on the secondary master server that was switched to.

Note

Starting and stopping Hadoop

Hadoop can be started or stopped solely on the secondary master server only if an error has occurred on the primary master server, or if it is not running. After the primary master server has recovered, promptly perform failback to the primary master server.

Do not run Hadoop on the secondary master server when the primary master server is in a normal state.

See

- Refer to "[Checking the status of the master server HA cluster](#)" for information on the status of the master server.

- Refer to "A.14 bdpp_start" and "A.16 bdpp_stop" for information on starting and stopping Hadoop.

(5) Recover server

Refer to the system log of the primary master server. Investigate and remove the cause of the error.

If a serious error has occurred on the master server, requiring a server to be rebuilt, use the restore feature of this product to recover. The restore feature can be used to recover the system configuration and the master server definition information.

Refer to "14.2.1.1 Restoring a Master Server, Development Server, or Collaboration Server" for information on the procedure to restore the master server.

Point

Prior to performing a restore, a backup of the master server must be created when it is running normally.

Refer to "14.1.2.1 Backing Up a Master Server, Development Server, or Collaboration Server" for information on the procedure to back up the master server.

(6) Resume server operations

After the primary master server has fully recovered, restart it.

(7) Failback command or restart

After the primary master server has fully recovered and operations have resumed, perform failback from the secondary master server to the primary master server. Failback can be performed either by executing the failover command on the secondary master server or by restarting the system.

Note that if analysis tasks were being performed on the secondary master server, operations must be stopped temporarily (after stopping jobs) prior to performing failback.

The procedure to perform failback using commands is given below.

1. Unmount the DFS on the secondary master server.

This operation will cause the DFS management server (MDS) to failback.

Example

If the DFS mount point is /mnt/pdfs:

```
# pdfsumount /mnt/pdfs <Enter>
```

2. Mount the DFS on the secondary master server.

Example

If the DFS mount point is /mnt/pdfs:

```
# pdfsmount /mnt/pdfs <Enter>
```

3. Confirm that the status of the DFS on the primary master server is "run".

Example

If the DFS management server (MDS) has failed back to the primary master server:

```
# pdfsrscinfo -m <Enter>
/dev/disk/by-id/scsi-1FUJITSU_300000370106:
FSID  MDS/AC  STATE  S-STATE  RID-1  RID-2  RID-N  hostname
  1  MDS(P)  run    -         0      0      0      master1 <- (*1)
```

1	AC	run	-	0	0	0	master1
1	MDS(S)	wait	-	0	0	0	master2
1	AC	run	-	0	0	0	master2

*1: Status of the primary master server is "run"

- Execute the hvswitch command on the secondary master server, and failback from the secondary master server to the primary master server.

```
# hvswitch appl <Enter>
```

appl is a fixed string.

- Refer to "[Checking the status of the master server HA cluster](#)", and confirm that failback to the primary master server was successful.



See

- Refer to the relevant commands in "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for information on the pdfsumount, pdfsmount, and pdfsrscinfo commands.
- Refer to the online Help of hvswitch for details of the hvswitch(1M) command.



Note

The master server functionality that is used for the execution of "[12.1 Adding Slave Servers](#)" is not a target for switching. If an error occurs on the primary master server and it switches to the secondary master server, remove the cause of the error from the primary master server and failback to the primary master server. Then, perform the steps in "[12.1 Adding Slave Servers](#)" again.



Information

If the master servers use replicated configuration and either the primary or secondary master server is not running, start operations as shown below.

- Starting the system when the primary master server is not running

Hadoop can be started by using the bdpp_start command if the system has been running solely on the secondary master server, because an error rendered the primary master server inoperable.

Refer to "[A.14 bdpp_start](#)" for information on starting Hadoop.

- Starting the system when the secondary master server is not running

Hadoop can be started by using the bdpp_start command if the system has been running solely on the primary master server, because an error rendered the secondary master server inoperable.

Refer to "[A.14 bdpp_start](#)" for information on starting Hadoop.

15.1.2 If the Master Server Does Not Use Replicated Configuration

If any of the errors listed below occur on the master server during execution of jobs, those jobs will be interrupted.

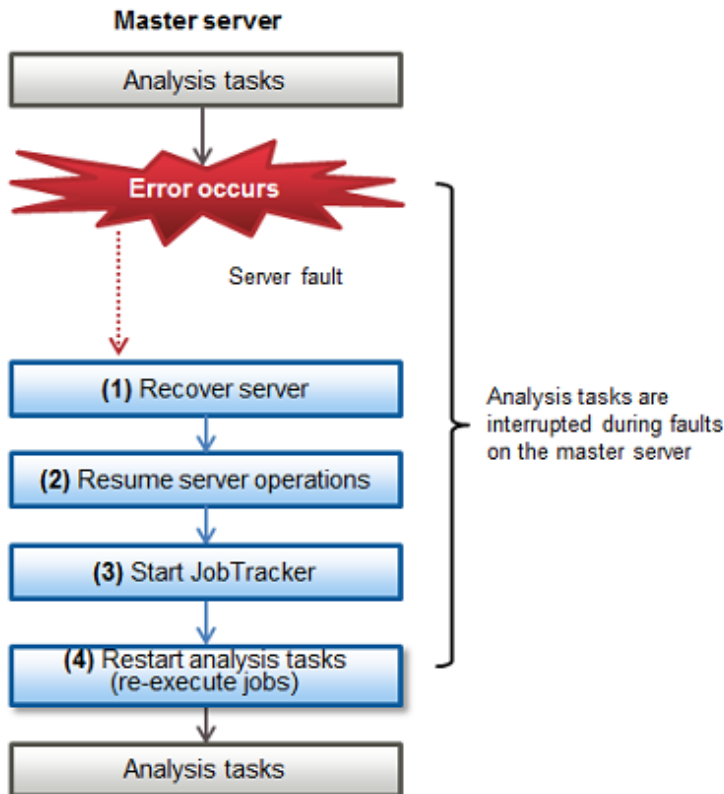
Tasks may be stopped for a long time until the master server has recovered.

- System error (physical machine)
- System error (virtual machine)
- Public LAN network error

- Cluster interconnect (CIP) error
- iSCSI network error
- JobTracker error

The following section explains the corrective action to take after an error occurs on the master server.

Figure 15.2 Procedure to resume tasks after an error occurs in a non-replicated configuration



(1) Recover server

Refer to the system log of the master server, and remove the cause of the error.

If a serious error has occurred on the master server requiring a server to be rebuilt, recover the master server.

The restore feature of this product can be used to rebuild and reconfigure the system configuration and the master server definition information to the normal running status.

Refer to "[14.2.1.1 Restoring a Master Server, Development Server, or Collaboration Server](#)" for information on the procedure to restore the master server.

Point

Prior to performing a restore, a backup of the master server must be created when it is running normally.

Refer to "[14.1.2.1 Backing Up a Master Server, Development Server, or Collaboration Server](#)" for information on the procedure to back up the master server.

(2) Resume server operations

Restart the master server that was recovered. When restarting the master server, the DFS must be unmounted and remounted on the DFS client (slave servers, development servers, and collaboration servers).

Use the following procedure to restart the master server:

1. Unmount the DFS on all slave servers, development servers, and collaboration servers.

Example

If the logical file system name of the DFS is pdfs1:

```
# umount pdfs1 <Enter>
```

2. Restart the master server.

If the master server is configured so that the DFS is not mounted automatically when the master server is started, restart it and then mount the DFS manually.

3. Mount the DFS on all slave servers, development servers, and collaboration servers.

Example

If the logical file system name of the DFS is pdfs1:

```
# mount pdfs1 <Enter>
```

(3) Start JobTracker

Start Hadoop on the master server that was recovered.

Always use the `bdpp_start` command to start Hadoop.

(4) Restart analysis tasks (re-execute jobs)

After the master server has fully recovered, execute jobs as required and resume tasks.

15.2 Operations when Errors Occur on a Slave Server

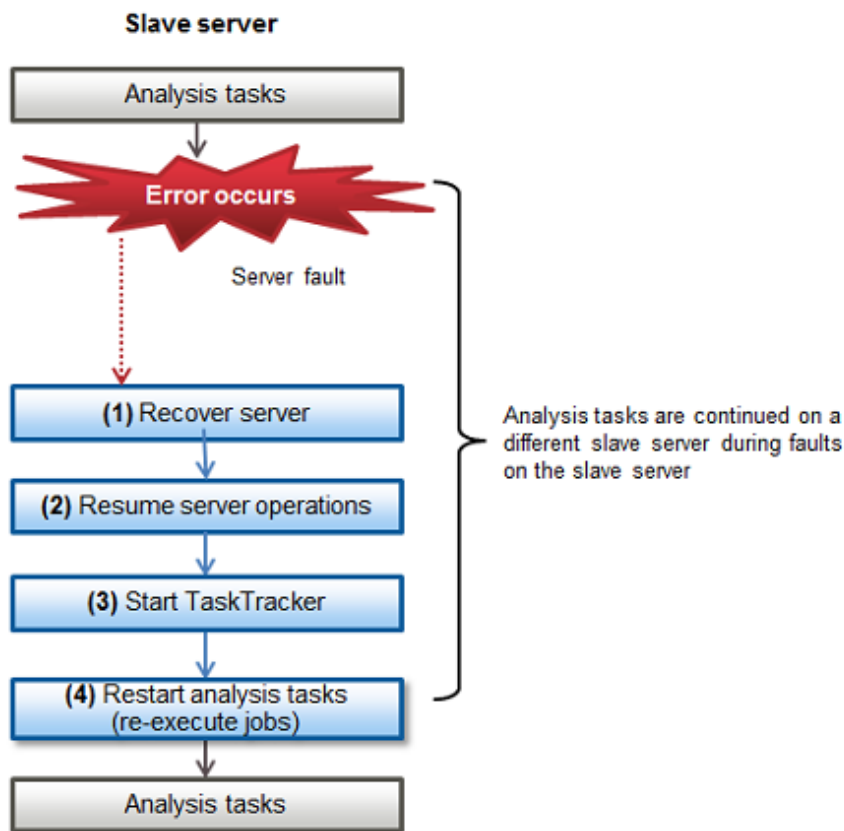
This section explains operations when an error occurs on a slave server.

When the events shown below occur on a slave server, other slave servers take over the job currently being executed, enabling processing to continue.

- System error
- Public LAN network error

- iSCSI network error

Figure 15.3 Procedure to resume tasks after an error occurs on a slave server



(1) Recover server

Refer to the system log of the relevant slave server. Investigate and remove the cause of the error.

If a serious fault occurs that cannot be resolved solely by referring to the system log and restarting the slave server, recover the slave server.

Refer to "[14.2.1.2 Restoring a Slave Server](#)" for information on the procedure to restore a slave server.

Point

Prior to performing a restore, a backup of the slave server must be created when it is running normally.

Refer to "[14.1.2.2 Backing Up a Slave Server](#)" for information on the procedure to back up a slave server.

(2) Resume server operations

After the slave server has fully recovered, restart the slave server that was recovered.

(3) Start TaskTracker

After the slave server has been recovered and operations have resumed, execute the `bdpp_start` command from the master server and restart Hadoop on the slave server where the error occurred.

Refer to "[A.14 bdpp_start](#)" for information on starting Hadoop.

Note

If the `bdpp_start` command is executed to restart Hadoop on some slave servers, the message "bdpp:WARN:001" will be output, but note that this does not indicate a problem with restarting Hadoop on the slave servers.

(4) Restart analysis tasks (re-execute jobs)

After the slave server has fully recovered, execute jobs as required and resume tasks.

15.3 Operations when Errors Occur on a Development Server

This section explains operations when an error occurs on a development server.

If any of the errors listed below occur on a development server, the MapReduce application will stop and jobs being executed may be interrupted.

- System error
- Public LAN network error
- iSCSI network error

If an error occurs on a development server, check the system log of the development server and remove the cause of the error.

If a serious error has occurred on the development server, requiring a server to be rebuilt, perform a restore on the development server and recover the server.

Refer to "[14.2.1.1 Restoring a Master Server, Development Server, or Collaboration Server](#)" for information on the procedure to restore the development server.

Point

Prior to performing a restore, a backup of the development server must be created when it is running normally.

Refer to "[14.1.2.1 Backing Up a Master Server, Development Server, or Collaboration Server](#)" for information on the procedure to back up a development server.

15.4 Operations when Errors Occur on a Collaboration Server

This section explains operations when an error occurs on a collaboration server.

If any of the following errors occur on a collaboration server, it may no longer be possible to access the DFS.

- System error
- iSCSI network error

If an error occurs on a collaboration server, check the system log of the collaboration server and remove the cause of the error.

If a serious error has occurred on the collaboration server, requiring a server to be rebuilt, perform a restore on the collaboration server and recover the server.

Refer to "[14.2.1.1 Restoring a Master Server, Development Server, or Collaboration Server](#)" for information on the procedure to restore the collaboration server.

Point

Prior to performing a restore, a backup of the collaboration server must be created when it is running normally.

Refer to "14.1.2.1 Backing Up a Master Server, Development Server, or Collaboration Server" for information on the procedure to back up a collaboration server.

15.5 Operations when Errors Occur on the File System

If an error occurs on the file system and it becomes corrupted, re-create the file system.

This section describes the procedure for re-creating a currently used DFS without changing the partition configuration.

The purpose of this is to restore all files from the backup to restore a file system.



See

If the partition configuration is changed, delete the DFS being used, and then create a DFS in the new partition configuration.

- Refer to "13.2 Deleting a Storage System" for information on the procedure to delete and create a DFS.
-



Note

When a file system is re-created, all old file system data is deleted. Therefore, back up the required data before re-creation.

The DFS re-creation procedure is described below, using the following environment as an example,

- Representative partition : /dev/disk/by-id/scsi-1FUJITSU_300000370106
- Logical file system name : pdfs1

Unless specifically indicated otherwise, execute at the primary master server.

1. If the targeted DFS is mounted, unmount it.

Unmount the DFS on all slave servers, development servers, and collaboration servers.

```
# umount pdfs1 <Enter>
```

Unmount the DFS at the master server.

```
# pdfsumntgl /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
```

2. Obtain the pdfsmkfs command line that re-creates a file system.

```
# pdfsmkfs -m /dev/disk/by-id/scsi-1FUJITSU_300000370106 > _mkfs_param_ <Enter>
```

3. Also, specify the force option in the pdfsmkfs command line. This option is used specifically for re-creation.

Before editing:

```
# cat _mkfs_param_ <Enter>
pdfsmkfs -o free=10,...[snip]...,node=master1,master2 /dev/disk/by-id/scsi-1FUJITSU_300000370106
```

After editing:

```
# cat _mkfs_param_ <Enter>
pdfsmkfs -o force,free=10,...[snip]...,node=master1,master2 /dev/disk/by-id/
scsi-1FUJITSU_300000370106
```

4. Re-create the DFS.

```
# sh _mkfs_param_ <Enter>
```

5. Delete the pdfsmkfs command line.

```
# rm _mkfs_param_ <Enter>
```

6. Mount the DFS, and restart operations.

Mount the DFS, starting with the master server.

```
# pdfsmntgl /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
```

Mount the DFS on all slave servers, development servers, and collaboration servers that use the DFS.

```
# mount pdfs1 <Enter>
```

15.6 How to Check Errors

Check the messages output to the system log (/var/log/messages) if an error occurs on the primary master server that causes a switch to the secondary master server (when using replicated configuration), or if an error on one side of a replicated network causes a switch to the other side.

The status of the HA cluster and network replication can be checked with the following commands.

Checking the status of the master server HA cluster

The status can be checked by running the hvdisp command on both the primary master server and the secondary master server.

Refer to the online Help of hvdisp for details of the hvdisp(1M) command.

The following example shows the status of the HA cluster after it switches to the secondary master server due to an error on the primary master server.

The error must be removed from the primary master server, because the status of app1 on this master server is "Faulted".

Display the status of the HA cluster of the primary master server (example)

```
# hvdisp -a <Enter>

Local System:      master1RMS
Configuration:    /opt/SMAW/SMAWRrms/build/bdpp.us

Resource          Type      HostName          State      StateDetails
-----
master1RMS        SysNode                Online
master2RMS        SysNode                Online
app1              userApp                Faulted    Failed Over
Machine001_app1  andOp    master1RMS
Machine000_app1  andOp    master2RMS        Offline
ManageProgram000_Cmd_APP1  gRes                Offline
Ippaddress000_Gls_APP1  gRes                Offline
```

Display the status of the HA cluster of the secondary master server (example)

```
# hvdisp -a <Enter>

Local System:      master2RMS
Configuration:    /opt/SMAW/SMAWRrms/build/bdpp.us

Resource          Type      HostName          State      StateDetails
-----
master1RMS        SysNode                Online
```

```

master2RMS      SysNode      Online
appl           userApp      Online
appl           userApp master1RMS  Online
Machine001_appl andOp      master2RMS  Online
Machine000_appl andOp      master1RMS
ManageProgram000_Cmd_APP1 gRes      Online
Ipaddress000_Gls_APP1 gRes      Online

```

In the example below, the cause of the error has been removed from the primary master server and the recovered state is shown. Failback to the primary master server is possible, because the status of appl on this master server is "Offline".

Display the status of the HA cluster of the primary master server (example)

```

# hvdisp -a <Enter>

Local System:      master1RMS
Configuration:    /opt/SMAW/SMAWRrms/build/bdpp.us

Resource          Type      HostName          State      StateDetails
-----
master1RMS        SysNode      Online
master2RMS        SysNode      Online
appl             userApp      Offline
appl             userApp master2RMS        Online
Machine001_appl  andOp      master2RMS
Machine000_appl  andOp      master1RMS        Offline
ManageProgram000_Cmd_APP1 gRes      Offline
Ipaddress000_Gls_APP1 gRes      Offline

```

Checking the status of communication between master servers

The status of the primary and secondary master servers can be checked by executing the cftool command on the respective server.

Refer to the online Help of cftool for details of the cftool(1M) command.

In the example below, an error has occurred in the cluster interconnect (CIP) and the status of the remote node cannot be determined. In this situation, a cluster partition is assumed to have occurred.

Display the communication status of the primary master server (example)

```

# cftool -n <Enter>
Node   Number State      Os      Cpu
master1 1      UP        Linux   EM64T
master2 2      LEFTCLUSTER Linux   EM64T

```

Display the communication status of the secondary master server (example)

```

# cftool -n <Enter>
Node   Number State      Os      Cpu
master1 1      LEFTCLUSTER Linux   EM64T
master2 2      UP        Linux   EM64T

```

Checking network replication status

Check the status using the dsphanet command. The status can be checked on the master server and slave servers.

Refer to "7.4 dsphanet Command" under "Chapter 7 Command references" in the "PRIMECLUSTER Global Link Services Configuration and Administration Guide 4.3 Redundant Line Control Function" for information on the dsphanet command.

Display the status of network replication of the primary master server (example)

```
# /opt/FJSVhanet/usr/sbin/dsphanet <Enter>
[IPv4,Patrol / Virtual NIC]
Name          Status  Mode CL  Device
+-----+-----+-----+-----+
sha0          Active  d   ON   eth5(ON),eth9(OFF)
sha1          Active  p   OFF  sha0(ON)
[IPv6]
Name          Status  Mode CL  Device
+-----+-----+-----+-----+
```

Display the status of network replication of the slave server (example)

```
# /opt/FJSVhanet/usr/sbin/dsphanet <Enter>
[IPv4,Patrol / Virtual NIC]
Name          Status  Mode CL  Device
+-----+-----+-----+-----+
sha0          Active  e   OFF  eth5(ON),eth9(OFF)
sha1          Active  p   OFF  sha0(ON)
[IPv6]
Name          Status  Mode CL  Device
+-----+-----+-----+-----+
```


Chapter 16 Troubleshooting

This chapter explains how to investigate problems that may occur with this product.

If a problem occurs on a system that uses this product, collect the data required by Fujitsu technical staff to determine the cause of the problem.

The troubleshooting data for this product is described below.

Logs

This section describes the logs generated by this product.

Logs are output for each process.

The process-specific logs are shown in the table below.

Refer to the information provided by Apache Hadoop for details on the logs specific to Apache Hadoop.

Table 16.1 Log types and storage locations

Log type	Log file and storage location
Install/uninstall logs	- /var/tmp/bdpp_install.log
Logs for setup/removing setup	- /var/opt/FJSVbdpp/log/bdpp_setup.log
Start and stop processing logs	- /var/opt/FJSVbdpp/log/bdpp_hadoop.log - /var/opt/FJSVbdpp/log/bdpp_pcl_hadoop.log
Slave server addition and deletion processing logs	- /var/opt/FJSVbdpp/log/bdpp_ROR.log
Command logs	- /var/opt/FJSVbdpp/log/ <i>commandName</i> .log

Troubleshooting data collection tools

This section describes the Interstage Big Data Parallel Processing Server troubleshooting data collection tools.

Execute the tools shown in the following table, and collect the output results as the troubleshooting data.

Table 16.2 Troubleshooting data collection tool types and location where the data is stored

Troubleshooting data collection tool type	Troubleshooting data collection tool and location where the data is stored
Installation configuration display tool	- /opt/FJSVbdpp/bin/sys/bdpp_show_config.sh
Installation status display tool	- /opt/FJSVbdpp/bin/sys/bdpp_show_status.sh

Functionality-specific troubleshooting data

The troubleshooting data for specific functionality incorporated in this product must also be collected if a message described in "[E.3.2 Other Messages](#)" is output to this product's logs or to the system log (/var/log/messages).

- [When a Problem Occurs with an HA Cluster](#)
- [When a Problem Occurs during Cloning](#)
- [If DFS Setup or Shared Disk Problems Occurred](#)
- [When a Problem Occurs in Hadoop](#)

16.1 When a Problem Occurs with an HA Cluster

This section explains how to collect the troubleshooting data when there is a problem setting up an HA cluster or when switching.

16.1.1 Collecting Troubleshooting Data

Collect the following information required for investigating a fault in the HA cluster system from the primary master server and the secondary master server.

1. HA cluster troubleshooting data
 - Use the `fjsnap` command to collect the information necessary to investigate the fault.
Refer to "[Executing the fjsnap command](#)".
 - Collect the troubleshooting data for the system.
Refer to "[Executing the pclsnap command](#)".

2. Crash dump

If it is possible to collect the crash dump from the server where the fault occurred, collect the crash dump manually before restarting the server.

The crash dump becomes useful when the problem is with the OS.

Example: A switchover occurs because of an unexpected resource failure

When switching of the cluster application has finished, collect the crash dump on the node where the resource failure occurred.

Refer to "[Crash dump](#)" for information on the crash dump.

3. If a fault can be reproduced, collect the data (in any format) summarizing the procedure to reproduce the fault



Information

When reporting the fault information to Fujitsu technical support, it is necessary for the information required for investigating the fault to be collected correctly. The information collected is used to check the problems and to reproduce the faults. This means that if the information is not accurate, reproducing and diagnosing the problem can take longer than necessary, or render it impossible.

Collect the information for the investigation quickly from the primary master server and the secondary master server. The information to be collected by the `fjsnap` command, in particular, disappears as time passes from the point when the fault occurred, so special care is needed.

Executing the `fjsnap` command

1. Log in to the primary master server and the secondary master server using root permissions.
2. Execute the `fjsnap` command on each server.

```
# /usr/sbin/fjsnap -a output <Enter>
```

Specify the output destination file name for the error information collected with the `fjsnap` command in the `output` option.



See

Refer to the README file included with the FJSVsnap package for information on the `fjsnap` command.

Information

Execution timing for the fjsnap command

For problems occurring during normal operation, such as when an error message is output, execute the fjsnap command immediately after the problem occurs.

Collect the crash dump if the fjsnap command cannot be executed because the system has stopped responding. After that, start in single-user mode and execute the fjsnap command. Refer to "[Crash dump](#)" for information on collecting the crash dump.

If the node automatically restarts after a problem occurs (unable to start in single-user mode) or if incorrectly restarted in multiuser mode, execute the fjsnap command.

Collect the crash dump when the troubleshooting data cannot be collected, because the fjsnap command ends in an error or does not produce a return.

Executing the pclsnap command

1. Log in to the primary master server and the secondary master server using root permissions.
2. Execute the pclsnap command on each server.

```
# /opt/FJSVpclsnap/bin/pclsnap {-a output or -h output} <Enter>
```

When -a is specified, all the detailed information is collected, so the data size will be large. Only the cluster control information is collected if -h is specified.

Specify the output destination and either the file name particular to the output media or the output file name (such as /dev/st0) for the error information collected with the pclsnap command in the *output* option.

Specify the path beginning with "/" if specifying a path relative to the current directory where the output file name includes the directory.

See

Refer to the README file included with the FJSVpclsnap package for information on the pclsnap command.

Information

Execution timing for the pclsnap command

For problems occurring during normal operation, such as when an error message is output, execute the pclsnap command immediately after the problem occurs.

Collect the crash dump if the pclsnap command cannot be executed, because the system has stopped responding. After that, start in single-user mode and execute the pclsnap command. Refer to "[Crash dump](#)" for information on collecting the crash dump.

If the node automatically restarts after a problem occurs (unable to start in single-user mode) or if mistakenly restarted in multiuser mode, execute the pclsnap command.

Collect the crash dump when the troubleshooting data cannot be collected, because the pclsnap command ends in an error or does not produce a return.

Information

Available directory space required to execute the pclsnap command

The following table is a guide to the available directory space required to execute the pclsnap command.

Directory type	Default directory	Free space (estimate) (MB)
Output directory	Current directory at the time of execution	300
Temporary directory	/tmp	500

Note

The estimates given above (300 MB and 500 MB) may be insufficient in some systems.

If information collection could not be completed successfully due to insufficient directory capacity, the `pclsnap` command outputs an error or warning message when it ends. In this case, take the action shown below and then re-execute the command.

Action to take when there is insufficient capacity in the output directory

The following error message is output when the `pclsnap` command is executed but fails to generate an output file.

```
ERROR: failed to generate the output file "xxx".
DIAG: ...
```

Action:

Change the output directory to a location with a large amount of space available, then re-execute the command.

Example:

When `/var/crash` is the output directory:

```
# /opt/FJSVpclsnap/bin/pclsnap -a /var/crash/output <Enter>
```

Action to take when there is insufficient capacity in the temporary directory

The following warning message may be output when the `pclsnap` command is executed.

```
WARNING: The output file "xxx" may not contain some data files.
DIAG: ...
```

If this warning message is output, the output file for the `pclsnap` command is generated, but the output file may not contain all the information that was meant to be collected.

Action:

Change the temporary directory to a location with a large amount of space available, then re-execute the command.

Example:

When the temporary directory is changed to `/var/crash`:

```
# /opt/FJSVpclsnap/bin/pclsnap -a -T /var/crash output <Enter>
```

If the same warning is output even after changing the temporary directory, investigate the following possible causes:

- (1) The state of the system is causing the information collection command to timeout
- (2) The files being collected are larger than the free space in the temporary directory

If the problem is (1), the log for the timeout is recorded to `pclsnap.elog`, part of the files output by the `pclsnap` command. Collect the crash dump if possible along with the output file of the `pclsnap` command.

If the problem is (2), check if the sizes of (a) or (b) exceed the capacity of the temporary directory.

- (a) Log file size

- /var/log/messages
- Log files in /var/opt/SMAW*/log/ (such as SMAWsf/log/rcsd.log)
- (b) Total size of the core file

- GFS core file: /var/opt/FJSVsfcs/cores/*
- GDS core file: /var/opt/FJSVsdX/*core/*

If these are larger than the capacity of the temporary directory, move the relevant files to a partition separate to the output directory and the temporary directory and re-execute the pclsnap command. Save the files moved. Do not delete them.

Crash dump

In environments where Linux Kernel Crash Dump (LKCD), Netdump, or diskdump is installed, it is possible to collect the crash dump as the troubleshooting data.

Timing for collecting the crash dump

- If an Oops occurs in the kernel
- If a panic occurs in the kernel
- If the <Alt> + <SysRq> + <C> keys were pressed at the system administrator console
- When the NMI button is pressed on the main unit

The following describes how to collect the crash dump.

1. How to collect the crash dump after a system panic
First, check if the crash dumps, for times after the switchover occurred, exist in the directory where the crash dumps are stored. If crash dumps exist from after the time when the switch occurred, collect the crash dumps. If there are no crash dumps from after the time when the switch occurred, collect the crash dumps manually as far as possible.
2. How to collect crash dumps manually
Use one of the following methods to collect the crash dumps in the directories where crash dumps are stored:
 - Press the NMI button on the main unit
 - Press the <Alt> + <SysRq> + <C> keys at the console

Directory where the crash dump is saved

The crash dump is saved as a file either on the node where the fault occurred (LKCD or diskdump) or on the Netdump server (Netdump).

The directory where it is saved is /var/crash.

16.2 When a Problem Occurs during Cloning

The following describes how to collect the troubleshooting data if clone image creation or cloning problems occur when adding or deleting slave servers.

Refer to "Troubleshooting" in the "ServerView Resource Orchestrator Virtual Edition V3.1.0 Operation Guide" for details.

16.2.1 Types of Troubleshooting Data

Collect the troubleshooting data when a problem occurs on the system where this product is being used so that Fujitsu technical support can investigate the problem.

There are two types of troubleshooting data. Collect the data that is required for the purposes described below.

1. Collecting the initial troubleshooting data

Collect the data required for initial triage of the cause of the problem that occurred and contact Fujitsu technical support. The amount of information collection is small and so can be easily sent by means such as email. Refer to "[16.2.2 Collecting Initial Troubleshooting Data](#)" for details.

2. Collecting detailed troubleshooting data

It is sometimes possible to determine the cause by using just the initial troubleshooting data, but more troubleshooting data may be required for some problems.

It is then necessary to collect more detailed troubleshooting data. Detailed investigation involves collecting large number of resources that are needed to determine the cause of a problem that has occurred.

Consequently, the information collected will be larger than the initial troubleshooting data collected to triage the problem.

If requested by Fujitsu technical support, send the detailed troubleshooting data that has been collected.

Refer to "[16.2.3 Collecting Detailed Troubleshooting Data](#)" for details.



Note

Collect the troubleshooting data quickly when a problem occurs. The information required to investigate a problem disappears as time passes.

16.2.2 Collecting Initial Troubleshooting Data

This section describes how to collect the troubleshooting data needed to triage the cause of a problem.

Collect resources by using the appropriate method, depending on the characteristics of the collection method and the environment and the system where the problem occurred.

1. Collecting resources from the master server

Collect the troubleshooting data from the master server (`rcxadm mgrctl snap -all`)

The troubleshooting data from managed servers can be collected in a batch using the network, so this method is much simpler than executing the command on each individual managed server.

Refer to "[Collecting diagnostics data from the admin sever \(rcxadm mgrctl snap -all\)](#)", and collect the information.

Along with the 65 MB of available space required to execute the `rcxadm mgrctl snap -all` command, approximately 30 MB of space is required for each server.

2. Collecting resources from servers

Collect the troubleshooting data from servers (`rcxadm mgrctl snap`, `rcxadm agtctl snap`)

Refer to "[Collecting resources from servers \(rcxadm mgrctl snap, rcxadm agtctl snap\)](#)", and collect the information.

65 MB of available space is required to execute the `rcxadm mgrctl snap` command.

30 MB of available space is required to execute the `rcxadm agtctl snap` command.

Collecting diagnostics data from the admin sever (`rcxadm mgrctl snap -all`)

By executing the command for collecting the troubleshooting data (`rcxadm mgrctl snap -all`) on the admin server, the troubleshooting data for managed servers is collected in a batch.

The following describes collecting the troubleshooting data with the command (`rcxadm mgrctl snap -all`).

1. Log in to the master server using root permissions.
2. Execute the `rcxadm mgrctl snap -all` command.

```
# /opt/FJSVrcvnr/bin/rcxadm mgrctl snap [-dir directory] -all <Enter>
```

3. Send the information collected to Fujitsu technical support.

Note

- When collecting resources from the master server, the manager must be operating on the admin server. If the manager is not operating, collect resources on individual servers.
- The troubleshooting data cannot be collected from managed servers in the following cases:
 - When a communications route has not been established
 - If a managed server has stopped

In either case, collection of the troubleshooting data on other managed servers continues uninterrupted.

Check the command execution results in the execution log.

Refer to "rcxadm mgrctl" in the "ServerView Resource Orchestrator Express/Virtual Edition V3.1.0 Command Reference" for details.

When collection has failed on a managed server, either execute the `rcxadm mgrctl snap -all` command on the master server again or execute the `rcxadm agtctl snap` command on the managed server.

Collecting resources from servers (`rcxadm mgrctl snap`, `rcxadm agtctl snap`)

Apart from the `rcxadm mgrctl snap -all` command that is executed on the master server and that can collect the troubleshooting data in a batch from managed servers, there are the `rcxadm mgrctl snap` and `rcxadm agtctl snap` commands that collect the information only on the server where they are executed.

The following describes collecting the troubleshooting data with the commands (`rcxadm mgrctl snap` or `rcxadm agtctl snap`).

1. Log in to the server where the data will be collected using root permissions.
2. Execute the `rcxadm mgrctl snap` or `rcxadm agtctl snap` command.
Note that the command executed depends on the server where the resources are collected.

When collecting on a master server

```
# /opt/FJSVrcvnr/bin/rcxadm mgrctl snap [-dir directory] <Enter>
```

When collecting on a slave server

```
# /opt/FJSVrcxat/bin/rcxadm agtctl snap [-dir directory] <Enter>
```

3. Send the information collected to Fujitsu technical support.

Refer to "rcxadm agtctl" or "rcxadm mgrctl" in the "ServerView Resource Orchestrator Express/Virtual Edition V3.1.0 Command Reference" for details.

16.2.3 Collecting Detailed Troubleshooting Data

This section describes how to collect detailed troubleshooting data needed to determine the cause of a problem.

When the cause of a problem cannot be determined just from the initial troubleshooting data, more detailed troubleshooting data is required.

The troubleshooting data required to determine the cause of a problem is collected by executing the troubleshooting data collection commands (`rcxadm mgrctl snap -full` and `rcxadm agtctl snap -full`) on servers.

80 MB of available space is required to execute this feature.

On the server where resources are to be collected, use the following procedure to collect the resources.

1. Log in to the server where the data will be collected using root permissions.

2. Execute the `rcxadm mgrctl snap -full` or `rcxadm agtctl snap -full` command.
Note that the command executed depends on the server where the resources are collected.

When collecting on a master server

```
# /opt/FJSVrcvnr/bin/rcxadm mgrctl snap -full [-dir directory] <Enter>
```

When collecting on a slave server

```
# /opt/FJSVrcxat/bin/rcxadm agtctl snap -full [-dir directory] <Enter>
```

3. Send the information collected to Fujitsu technical support.

Refer to "rcxadm agtctl" or "rcxadm mgrctl" in the "ServerView Resource Orchestrator Express/Virtual Edition V3.1.0 Command Reference" for details.

16.3 If DFS Setup or Shared Disk Problems Occurred

This section explains how to collect the troubleshooting data when there is a problem setting up a DFS or in the shared disc.

Refer to "4.6.2 Collecting DFS Troubleshooting Information" in the "Primesoft Distributed File System for Hadoop V1.0 User's Guide" for details.

16.3.1 Collecting DFS Troubleshooting Data

When requesting investigation by Fujitsu technical support as part of the action taken in response to an output message, login using root permissions and collect the following resources.

Collect resources in a state that is as close as possible to the state when the phenomena occurred.

In the information collected after the phenomena has ended or the system has been restarted, the state of the system has changed, and this may make investigation impossible.

1. Output results of the resource collection tool (`pdfssnap.sh` and the `fjsnap` command)

Use `pdfssnap.sh` and the `fjsnap` command to collect the troubleshooting data. Collect from all servers that shared the DFS, as far as possible.

- Refer to "[Executing pdfssnap.sh](#)".
- Refer to "[Executing fjsnap](#)".

2. Crash dump

If there was a panic on the server, for example, also collect the crash dump file as part of the troubleshooting data. Refer to "[Collecting the crash dump](#)" for details.

3. Execution results of the `pdfsck` command

Collect if there is a mismatch in the DFS and it needs to be restored. Refer to "[Execution results of the pdfsck command](#)".

4. Collecting the core image for the daemon

As part of the actions in response to DFS error messages, it may be necessary to collect core images as they relate to various daemons. Refer to "[Collecting the core image for a daemon](#)" for details.

When it is necessary to send the troubleshooting data quickly, collect the following as the initial troubleshooting data:

- a. Output results of the resource collection tool (`pdfssnap.sh`)

b. /var/log/messages*

After collecting the resources for initial investigation, ensure that other resources are also collected.

Executing pdfssnap.sh

1. Log in to the server where the data will be collected using root permissions.
2. Execute pdfssnap.sh.

```
# /etc/opt/FJSVpdfs/bin/pdfssnap.sh <Enter>
```

Note

With pdfssnap.sh, the troubleshooting data is output to the directory where the command is executed. For this reason, at least 100 MB of free space must be available in the file system that will execute the command.

Executing fjsnap

1. Log in to the server where the data will be collected using root permissions.
2. Execute the fjsnap command.

```
# /opt/FJSVsnap/bin/fjsnap -a anyFileName <Enter>
```

Collecting the crash dump

This is normally saved in a folder named "/var/crash/time of panic" when the server is started after a panic. Collect on all servers where a system panic has occurred.

Execution results of the pdfsck command

```
# pdfsck -N -o nolog blockParticularFileForTheRepresentativePartition <Enter>
```

Collecting the core image for a daemon

Collect core images on all DFS admin servers.

The procedure is explained below using the example of collecting the core image of the pdfsfrmd daemon.

1. Determining the process ID

Identify the process ID by using the ps command. Change the argument of the grep command if the target is other than the pdfsfrmd daemon.

```
# /bin/ps -e | /bin/grep pdfsfrmd <Enter>
5639 ? 00:00:25 pdfsfrmd
```

The beginning of the output is the process ID of the pdfsfrmd daemon. This is not output if the pdfsfrmd daemon is not running. Collect on another server if it is not operating.

Information

When collecting the MDS core image, specify pdfsmg in the argument of the grep command.



See

Refer to the online Help for information on the ps and grep commands.

2. Getting the core image

Collect the core image of pdfsfrmd to the /var/tmp/pdfsfrmd_node1.5639 file by using the gcore command. After that, compress the file with the tar command.

```
# /usr/bin/gcore -o /var/tmp/pdfsfrmd_node1 5639 <Enter>
gcore: /var/tmp/pdfsfrmd_node1.5639 dumped
# /bin/tar czvf /var/tmp/pdfsfrmd_node1.5639.tar.gz /var/tmp/pdfsfrmd_node1.5639 <Enter>
# /bin/ls -l /var/tmp/pdfsfrmd_node1.5639.tar.gz <Enter>
-rw-rw-r-- 1 root other 1075577 June 12 16:30 /var/tmp/
pdfsfrmd_node1.5639.tar.gz
```



See

Refer to the online Help for information on the tar command.

16.4 When a Problem Occurs in Hadoop

Execute /opt/FJSVbdpp/products/HADOOP/bin/HADOOP-collect.sh, and collect collectinfo.tar.gz output to the current directory.

Format

```
HADOOP-collect.sh --servers serverName[,serverName]
```

Options

```
--servers serverName[,serverName]
```

Specify the host names of the primary master server, the secondary master server, the slave server, and the development server separated by commas. Do not use spaces.

Required permissions and execution environment

Privileges

root permissions

Execution environment

Master server



Example

```
# /opt/FJSVbdpp/products/HADOOP/bin/HADOOP-collect.sh --
servers=master1,master2,slave1,slave2,slave3,slave4,slave5,develop <Enter>
```

Appendix A Commands

This appendix explains the commands provided by this product.

Command list

Command name	Description	Execution environment				
		Master server		Slave server	Development server	Collaboration server
		Primary	Secondary			
bdpp_addserver (*1)	Registers the slave server that creates a clone image and the slave server to which the clone image is deployed	Y (*3)	N	N	N	N
bdpp_backup (*1)	Backs up the configuration information	Y	Y	N	Y	Y
bdpp_changeimagedir (*1)	Changes the storage directory for the clone image	Y (*3)	N	N	N	N
bdpp_changeslaves (*1)	Reflects the changes made to the slave server definition file	Y	Y(*4)	N	N	N
bdpp_deployimage (*1)	Deploys a clone image	Y (*3)	N	N	N	N
bdpp_getimage (*1)	Creates a clone image	Y (*3)	N	N	N	N
bdpp_lanctl (*2)	Registers the automatic configuration of the network parameter (compatible with V1.0.0)	Y (*3)	N	N	N	N
bdpp_listimage	Displays the registration status of a clone image	Y (*3)	N	N	N	N
bdpp_listserver	Displays the registration status of a slave server	Y (*3)	N	N	N	N
bdpp_prepareserver (*2)	Registers the automatic configuration of the network parameter and the iSCSI name	Y	N	N	N	N

Command name	Description	Execution environment				
		Master server		Slave server	Development server	Collaboration server
		Primary	Secondary			
bdpp_removeimage (*1)	Deletes a clone image	Y (*3)	N	N	N	N
bdpp_removeserver (*1)	Deletes a slave server	Y (*3)	N	N	N	N
bdpp_restore (*1)	Restores the configuration information	Y	Y	N	Y	Y
bdpp_start (*1)	Starts Hadoop	Y	Y (*4)	N	N	N
bdpp_stat	Displays the status of Hadoop	Y	Y (*4)	N	N	N
bdpp_stop	Stops Hadoop	Y	Y (*4)	N	N	N

Y: Can be used

N: Cannot be used

*1: These commands cannot be executed simultaneously. Also, multiple occurrences of the same command cannot be executed concurrently.

*2: Multiple occurrences of the same command cannot be executed concurrently.

*3: These commands cannot be used on a master server installed in a virtual environment.

*4: These commands become available when the master server has replicated configuration and a failover causes operations to switch to the secondary master server. Do not use these commands when operating on the primary master server.



Point

.....
 Commands are positioned under the following directory:

- /opt/FJSVbdpp/bin

.....

A.1 bdpp_addserver

Name

/opt/FJSVbdpp/bin/bdpp_addserver - Registers the slave server that creates a clone image and the slave server to which the clone image is deployed

Format

bdpp_addserver *storageDirectoryOf_clone.conf*

Description

This command registers the slave server that creates a clone image and the slave server to which the clone image is deployed.

Arguments

storageDirectoryOf_clone.conf

Specify the name of the directory that clone.conf is stored in or the absolute path of clone.conf.

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

A.2 bdpp_backup

Name

/opt/FJSVbdpp/bin/bdpp_backup - Backs up the configuration information

Format

bdpp_backup -d *directory* [-q]

Description

This command backs up the configuration information associated with this product.

Options

-d *directory*

Using an absolute path, specify the storage directory for backup files.

-q

Clone images are not backed up when this option is specified.

This option can only be specified on the primary master server installed in a physical environment.

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server, development server, or collaboration server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

Note

- Backup cannot be performed if Hadoop is running. Use the `bdpp_stop` command to stop Hadoop so that backup can be performed.
- A directory that already stores backup files cannot be specified.
- If an error occurs during backup, the storage of backup files in the specified directory may not have completed. When this happens, delete all the contents of the `FJSVbdpp-backup` directory created within the storage directory.

A.3 `bdpp_changeimagedir`

Name

`/opt/FJSVbdpp/bin/bdpp_changeimagedir` - Changes the storage directory for the clone image

Format

`bdpp_changeimagedir cloneImageStorageDirectory`

Description

This command changes the storage directory for the clone image.

Note

When this command is executed, the clone image file is copied. This may take a while.

Arguments

cloneImageStorageDirectory

Specify the name of the new storage directory (destination to be changed to) for the clone image.

Note

The directory that is specified must satisfy the following conditions:

- The path including the storage directory for the clone image file must not exceed 100 characters.

Description

This command reflects the slave server information that was changed in the slave server definition file (slaves).

Options

None

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

A.5 bdpp_deployimage

Name

/opt/FJSVbdpp/bin/bdpp_deployimage - Deploys a clone image

Format

`bdpp_deployimage serverName [...] imageName`

Description

This command deploys a server system image to other servers.



Note

- When this command is executed, the clone image file is distributed to the server. This may take a while.
- Four clone image deployment processes may be executed simultaneously. When five or more are requested, it will be on standby until the process being executed completes.

Arguments

`serverName [...]`

Specify the name of the server that deploys the clone image (the host name specified in the HOSTNAME parameter of the `/etc/sysconfig/network` file).

Multiple servers can be specified by using a comma (,) as a delimiter.

Point

Specify the server name that is displayed when the `bdpp_listserver` command is executed. Refer to "A.9 `bdpp_listserver`" for information on this command.

imageName

Specify the name of the clone image to be distributed.

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

A.6 `bdpp_getimage`

Name

`/opt/FJSVbdpp/bin/bdpp_getimage` - Creates a clone image

Format

`bdpp_getimage serverName imageName`

Description

This command creates a server system image to be used for clone slave servers.



Note

Executing this command may take a while, as it creates the clone image file.

Arguments

serverName

Specify the name of the server to be cloned (the host name specified in the `HOSTNAME` parameter of the `/etc/sysconfig/network` file).

Point

Specify the server name that is displayed when the `bdpp_listserver` command is executed. Refer to "A.9 `bdpp_listserver`" for information on this command.

imageName

Specify a name to identify the clone image being created.

For the name of the clone image, specify a string of up to 32 characters, starting with a letter. Alphanumeric characters and the underscore (`_`) can be used.

Specify a different name from the clone image that has been created. The command returns an error if the same name is specified.

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

A.7 `bdpp_lanctl`

Name

`/opt/FJSVbdpp/bin/bdpp_lanctl` - Registers the automatic configuration of the network parameter

Format

`bdpp_lanctl ipAddr file1 file2`

Description

This command registers the automatic configuration of the network parameter for the creation and distribution of a clone image.

Arguments

ipAddr

Specify the IP address of the admin LAN for the slave server whose clone image was created.

file1

Using an absolute path, specify the definition file (FJSVrcx.conf) for the admin LAN information that was created along with the slave server whose clone image was created.

file2

Using an absolute path, specify the definition file (ipaddr.conf) used for the automatic configuration of the network parameter.

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred



Information

This command is provided for backward compatibility with earlier versions (V1.0.0). It is not normally used.

A.8 bdpp_listimage

Name

/opt/FJSVbdpp/bin/bdpp_listimage - Displays the registration status of a clone image

Format

bdpp_listimage

Description

This command checks clone image registration.

The following items are displayed in a list.

Item name	Content
NAME	Name of the clone image
VERSION	Display is fixed at "1"
CREATIONDATE	Date and time when the clone image was created
COMMENT	Display is fixed at "-"

Options

None

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

Example

```
# /opt/FJSVbdpp/bin/bdpp_listimage <Enter>
IMAGEDIRECTORY
-----
imagedir: /data/imagedir

NAME                VERSION  CREATIONDATE      COMMENT
-----
RX200img            1        2012/04/20-22:39:59 -
RX300img            1        2012/04/20-22:41:13 -
```

A.9 bdpp_listserver

Name

/opt/FJSVbdpp/bin/bdpp_listserver - Displays the registration status of a slave server

Format

bdpp_listserver

Description

This command displays the registration status of the slave server that creates a clone image and the slave server to which the clone image is deployed.

The following items are displayed in a list.

Item name	Content
PHYSICAL_SERVER	Physical server name
SERVER	Server (physical OS/VM host) name

Item name	Content
ADMIN_IP	IP address of the admin LAN
STATUS	<p>Server status</p> <p>One of the following is displayed:</p> <ul style="list-style-type: none"> - normal - warning - unknown - stop - error - fatal <p>Refer to "11.2 Resource State" in the "ServerView Resource Orchestrator Virtual Edition V3.1.0 Operation Guide" for information on the statuses of the server.</p>
MAINTENANCE	<p>Maintenance mode status</p> <ul style="list-style-type: none"> - If the maintenance mode has been set "ON" is displayed. - If the maintenance mode has been cancelled "OFF" is displayed.

Options

None

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

Example

```
# /opt/FJSVbdpp/bin/bdpp_listserver <Enter>
PHYSICAL_SERVER  SERVER                ADMIN_IP             STATUS  MAINTENANCE
-----
slave1           slave1                192.168.1.21        normal OFF
slave2           -                    192.168.1.22        stop   -
slave3           -                    192.168.1.23        stop   -
```

slave4	-	192.168.1.24	stop	-
slave5	-	192.168.1.25	stop	-

A.10 bdpp_prepareserver

Name

/opt/FJSVbdpp/bin/bdpp_prepareserver - Registers the automatic configuration of the network parameter and the iSCSI name

Format

bdpp_prepareserver -m *ipAddr* -o *file* -i *file* -f *file* [-v]

Description

This command enables the automatic configuration of the network parameter and the iSCSI name for cloning.

Options

-m *ipAddr*

Specify the IP address of the admin LAN for the slave server whose clone image was created.

If adding a slave server to a virtual environment, specify the IP address of the admin LAN for the slave server that is the clone source.

-o *file*

Specify FJSVrcx.conf for the automatic configuration of the network parameter.

Specify the admin LAN information of the slave server whose clone image was created.

If adding a slave server to a virtual environment, define the admin LAN information for the slave server that is the clone source.

-i *file*

Specify ipaddr.conf for the automatic configuration of the network parameter.

Define the public LAN and iSCSI-LAN information for all slave servers to be cloned. If adding a slave server to a virtual environment, define the admin LAN information.

-f *file*

Specify initiator.conf for the automatic configuration of the iSCSI initiator name.

Define the iSCSI name for each server defined in the ETERNUS DX disk array.

-v

Specify this option when adding a slave server to a virtual environment.

Omit when adding a slave server to a physical environment.

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

A.11 bdpp_removeimage

Name

/opt/FJSVbdpp/bin/bdpp_removeimage - Deletes a clone image

Format

bdpp_removeimage *imageName*

Description

This command deletes the clone image that was created using the bdpp_getimage command.

Arguments

imageName

Specify the clone image name.

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

A.12 bdpp_removeserver

Name

/opt/FJSVbdpp/bin/bdpp_removeserver - Deletes a slave server

Format

bdpp_removeserver *serverName*

Description

This command deletes the slave server that creates a clone image, and the slave server to which the clone image is deployed; both of which were registered using the bdpp_addserver command.

Arguments

serverName

Specify the name of the server to be deleted (the host name specified in the HOSTNAME parameter of the /etc/sysconfig/network file).

Point

Specify the server name that is displayed when the bdpp_listserver command is executed. Refer to "A.9 bdpp_listserver" for information on this command.

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

A.13 bdpp_restore

Name

/opt/FJSVbdpp/bin/bdpp_restore - Restores the configuration information

Format

bdpp_restore -d *directory*

Description

This command restores the configuration information associated with this product.

Options

-d *directory*

Using an absolute path, specify the storage directory for backup files.

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server, development server, or collaboration server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred



The hardware configuration and the software configuration for each server must not have been changed since the `bdpp_backup` command was used for backup.

A.14 `bdpp_start`

Name

`/opt/FJSVbdpp/bin/bdpp_start` - Starts Hadoop

Format

`bdpp_start [-f]`

Description

This command starts Hadoop for this product.

It starts the Hadoop JobTracker on the master server and the Hadoop TaskTracker on all the connection target slave servers.

Options

-f

Forces Hadoop to start.

Specify this option only to force Hadoop to start, such as when an error occurs on the master servers that use replicated configuration and the JobTracker cannot start automatically.



Refer to "[15.1.1 If the Master Servers Use Replicated Configuration](#)" for information on situations that may require forced start.

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred



If the `bdpp_start` command and the `bdpp_stop` command are executed simultaneously, a mismatch may arise in the run state of the Hadoop JobTracker on the master server or the Hadoop TaskTracker on the slave server.

If this occurs, use the `bdpp_stat -all` command to check the status, then re-execute the `bdpp_start` command so that the Hadoop JobTrackers on the master server or the Hadoop TaskTrackers on the slave servers are all running, as shown in the following table.

Table A.1 Status of the `bdpp_stat -all` command and the execution results of the `bdpp_start` command

Master server JobTracker	Slave server TaskTracker	bdpp_start execution results
Not running	Not running	- Start the master server JobTracker and all slave server TaskTrackers
	Running on some slave servers	- Start the master server JobTracker and all slave server TaskTrackers that are not running
	Running on all slave servers	- Start only the master server JobTracker
Running	Not running	- "bdpp: WARN : 001: bdpp Hadoop is already started." message is output - Start all slave server TaskTrackers
	Running on some slave servers	- "bdpp: WARN : 001: bdpp Hadoop is already started." message is output - Start all slave server TaskTrackers that are not running

Master server JobTracker	Slave server TaskTracker	bdpp_start execution results
	Running on all slave servers	- "bdpp: WARN : 001: bdpp Hadoop is already started." message is output

A.15 bdpp_stat

Name

/opt/FJSVbdpp/bin/bdpp_stat - Displays the status of Hadoop

Format

bdpp_stat [-all]

Description

This command checks the Hadoop start and stop statuses for this product.

Options

-all

Displays the JobTracker on the master server and the TaskTracker processes on slave servers.

The following items are displayed in a list.

Content
Host name where the process is running
Name of the user that started the process (fixed as "mapred")
Process ID on the server where the process is running
Process name (jobtracker or tasktracker)

Default

Displays the status of the JobTracker on the master server.

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server

End status

The following status codes are returned:

0

Hadoop is started.

Other than 0

Hadoop is not started.

Example

When -all is specified

```
# /opt/FJSVbdpp/bin/bdpp_stat -all <Enter>
cluster mapred 4303 jobtracker
slave1 mapred 5526 tasktracker
slave2 mapred 1091 tasktracker
slave3 mapred 7360 tasktracker
slave4 mapred 28256 tasktracker
slave5 mapred 26884 tasktracker
bdpp: INFO : 003: bdpp Hadoop JobTracker is alive.
```

Default

```
# /opt/FJSVbdpp/bin/bdpp_stat <Enter>
bdpp: INFO : 003: bdpp Hadoop JobTracker is alive.
```

A.16 bdpp_stop

Name

/opt/FJSVbdpp/bin/bdpp_stop - Stops Hadoop

Format

bdpp_stop

Description

This command stops Hadoop for this product.

It stops the Hadoop JobTracker on the master server and the Hadoop TaskTracker on all the connection target slave servers.

Options

None

Required permissions and execution environment

Permissions

root permissions

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

 Note

.....

If the `bdpp_start` command and the `bdpp_stop` command are executed simultaneously, a mismatch may arise in the run state of the JobTracker on the master server or the TaskTracker on the slave server.

If this occurs, check the cautions for the `bdpp_start` command, restart the Hadoop JobTracker and TaskTracker, then stop if necessary.

.....

Appendix B Definition Files

This appendix describes the definition files for the system configuration information used when building systems or cloning.

Definition file list

Definition files	File overview	Explanation
bdpp.conf	Configuration file	This file defines the information required for the installation and setup of this product.
slaves	Slave server definition file	This file defines the slave servers that are the connection targets of the master server.
clone.conf	Cloning server definition file	This file is used for Smart Setup. It defines adding slave servers using cloning and deleting slave servers.
FJSVrcx.conf	Network parameter automatic configuration definition file	This definition file is used to configure the network parameter automatically after cloning.
ipaddr.conf	Network parameter automatic configuration definition file	This definition file is used to configure the network parameter automatically after cloning.
initiator.conf	iSCSI initiator configuration definition file	This definition file is used to automatically configure the iSCSI name for each server defined in the ETERNUS DX disk array.

The definition files are used in environments shown below.

Definition file	Environment				
	Master server		Slave server	Development server	Collaboration server
	Primary	Secondary			
bdpp.conf (*1)	Y	Y	Y	Y	Y
slaves (*2)	Y	Y	Y	Y	N
clone.conf	Y (*3)	N	N	N	N
FJSVrcx.conf	Y	N	N	N	N
ipaddr.conf	Y	N	N	N	N
initiator.conf	Y	N	N	N	N

Y: Used

N: Not used

*1: A [bdpp.conf](#) configuration file is created separately for each server.

*2: The definition of the "slaves" slave server definition file must match across servers. To change this definition, the definition files on the primary master server must be updated, after which the files on the primary master server must be copied to other servers.

*3: This file cannot be used on a master server installed in a virtual environment.

B.1 [bdpp.conf](#)

This section explains the settings in [bdpp.conf](#).

File storage location

`/etc/opt/FJSVbdpp/conf/bdpp.conf`

File content

This file defines the information required for the installation and setup of this product.

A separate file must be created for each server functionality.

Format

Configure using the format shown below.

```
parameter=specifiedValue
```

Refer to the following for information on the parameters defined in each server:

- [Definition for master servers](#)
- [Definition for slave servers](#)
- [Definition for development servers](#)
- [Definition for collaboration servers](#)

Note

Parameters with "select" as the specified value

Enter up to "parameter=" but omit the value.

```
BDPP_PCL_PRIMARY_CONNECT2=
```

Parameters with "optional" as the specified value

The parameter line itself can be omitted.

Point

Separate "bdpp.conf" templates for each server type are stored under "/DISK1/config_template" on the distribution media (DVDROM) of this product. Copy the relevant template to the system work directory, and then edit the parameters to create the configuration file.

Note

When installing this product, create the configuration file in any directory. The storage directory of the created configuration file must be specified as the "BDPP_CONF_PATH" environment variable before installation.

[Master server]

Parameters set in the master server configuration file are shown below.

Item No.	Parameter	Mandatory/Optional	Value to be specified (settings example)	Explanation
1	BDPP_CLUSTER_NAME	Mandatory	cluster	Specify the virtual host name to be set for the cluster. For the host name, the first character must be a letter, and it must consist of 11 or fewer alphanumeric characters.

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
2	BDPP_CLUSTER_IP	Mandatory	10.10.10.1	Specify the virtual IP address to be set for the cluster
3	BDPP_PRIMARY_NAME	Mandatory	master1	Specify the representative host name of the primary master server (the host name corresponding to the NIC that connects to the public LAN). For the host name, the first character must be a letter, and it must consist of 11 or fewer alphanumeric characters.
4	BDPP_PRIMARY_IP	Mandatory	10.10.10.11	Specify the representative IP address that will be used to connect to the primary master server public LAN. (The IP address that will be used to connect to public LAN1 in the figure)
5	BDPP_PRIMARY_ADM_IP	Mandatory	192.168.1.1	Specify the IP address that will be used to connect to the primary master server admin LAN. (The IP address that will be used to connect to the admin LAN in the figure)
6	BDPP_SECONDARY_NAME	Optional	master2	Specify the representative host name of the secondary master server (the host name corresponding to the NIC that connects to the public LAN). For the host name, the first character must be a letter, and it must consist of 11 or fewer alphanumeric characters. If not replicating the master server, this value should be omitted.
7	BDPP_SECONDARY_IP	Optional	10.10.10.12	Specify the representative IP address that will be used to connect to the secondary master server public LAN. (The IP address that will be used to connect to public LAN1 in the figure) If not replicating the master server, this value should be omitted.

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
8	BDPP_SECONDARY_ADM_IP	Optional	192.168.1.2	Specify the IP address that will be used to connect to the secondary master server admin LAN. (The IP address that will be used to connect to the admin LAN in the figure) If not replicating the master server, this value should be omitted.
9	BDPP_PCL_CLUSTER_GROUP	Mandatory	bdppcluster	Specify the cluster group name to be set for the cluster. For the cluster group name, the first character must be a letter, and it must consist of 31 or fewer alphanumeric characters.
10	BDPP_PCL_PRIMARY_iRMC_IP	Optional	192.168.1.11	Specify the IP address of the primary master server remote management controller. If building a master server in a virtual environment, this value should be omitted.
11	BDPP_PCL_PRIMARY_iRMC_ACCOUNT	Optional	root	Specify the account of the primary master server remote management controller. If building a master server in a virtual environment, this value should be omitted.
12	BDPP_PCL_PRIMARY_iRMC_PASS	Optional	password	Specify the password of the primary master server remote management controller. If building a master server in a virtual environment, this value should be omitted.
13	BDPP_PCL_SECONDARY_iRMC_IP	Optional	192.168.1.12	Specify the IP address of the secondary master server remote management controller. If not replicating the master server, this value should be omitted. If building a master server in a virtual environment, this value should be omitted.
14	BDPP_PCL_SECONDARY_iRMC_ACCOUNT	Optional	root	Specify the account of the secondary master server remote management controller.

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
				<p>If not replicating the master server, this value should be omitted.</p> <p>If building a master server in a virtual environment, this value should be omitted.</p>
15	BDPP_PCL_SECONDARY_iRMC_PASS	Optional	password	<p>Specify the password of the secondary master server remote management controller.</p> <p>If not replicating the master server, this value should be omitted.</p> <p>If building a master server in a virtual environment, this value should be omitted.</p>
16	BDPP_PCL_PRIMARY_CIP	Mandatory	192.168.2.11	Specify the IP address (*1) used for the primary master server CIP (Cluster Interconnect Protocol).
17	BDPP_PCL_SECONDARY_CIP	Optional	192.168.2.12	<p>Specify the IP address (*1) used for the secondary master server CIP (Cluster Interconnect Protocol).</p> <p>If not replicating the master server, this value should be omitted.</p>
18	BDPP_PCL_PRIMARY_CONNECT1	Mandatory	eth5	<p>Specify the primary master server side NIC name (NIC5) used for connections between clusters.</p> <p>Configure the network interface used in the cluster interconnect so that it is activated when the system is started. Note, however, that no IP address must be assigned.</p>
19	BDPP_PCL_PRIMARY_CONNECT2	Optional	eth6	<p>Specify the primary master server side NIC name (NIC6) used for connections between clusters.</p> <p>Note, however, that no IP address must be assigned.</p> <p>When the cluster interconnect is a single configuration, this value should be omitted.</p>
20	BDPP_PCL_SECONDARY_CONNECT1	Optional	eth5	Specify the secondary master server side NIC name (NIC5) used for connections between clusters.

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
				If not replicating the master server, this value should be omitted.
21	BDPP_PCL_SECONDARY_CONNECT2	Optional	eth6	Specify the secondary master server side NIC name (NIC6) used for connections between clusters. When the cluster interconnect is a single configuration, this value should be omitted. If not replicating the master server, this value should be omitted.
22	BDPP_GLS_PRIMARY_CONNECT1	Mandatory	eth1	Specify the primary master server NIC name (NIC1) used for connections between slave servers. (The NIC that will be used to connect to public LAN1 in the figure)
23	BDPP_GLS_PRIMARY_CONNECT2	Optional	eth2	Specify the primary master server NIC name (NIC2) used for connections between slave servers. (The NIC that will be used to connect to public LAN2 in the figure) When the public LAN is a single configuration, this value should be omitted. If building a master server in a virtual environment, this value should be omitted.
24	BDPP_GLS_SECONDARY_CONNECT1	Optional	eth1	Specify the secondary master server NIC name (NIC1) used for connections between slave servers. (The NIC that will be used to connect to public LAN1 in the figure) If not replicating the master server, this value should be omitted.
25	BDPP_GLS_SECONDARY_CONNECT2	Optional	eth2	Specify the secondary master server NIC name (NIC2) used for connections between slave servers.

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
				<p>(The NIC that will be used to connect to public LAN2 in the figure)</p> <p>When the public LAN is a single configuration, this value should be omitted.</p> <p>If not replicating the master server, this value should be omitted.</p> <p>If building a master server in a virtual environment, this value should be omitted.</p>
26	BDPP_GLS_POLLING1_IP	Mandatory	10.10.10.100	<p>Specify the IP address of the monitoring target HUB host.</p> <p>(Monitoring target HUB host on public LAN1 in the figure)</p>
27	BDPP_GLS_POLLING2_IP	Optional	10.10.10.101	<p>Specify the IP address of the monitoring target HUB host.</p> <p>(Monitoring target HUB host on public LAN2 in the figure)</p> <p>When the public LAN is a single configuration, this value should be omitted.</p> <p>If building a master server in a virtual environment, this value should be omitted.</p>
28	BDPP_GLS_NETMASK	Optional	255.255.0.0	<p>Specify the net mask for the public LAN IP address.</p> <p>The default net mask is "255.255.255.0".</p>
29	BDPP_DISKARRAY_iSCSI1_IP	Optional	10.10.11.11	<p>Specify the IP address of the ETERNUS DX disk array iSCSI-CA port that is part of the DFS management partition.</p> <p>This parameter is mandatory if YES is specified for BDPP_PDFS_FILESYSTEM.</p>
30	BDPP_DISKARRAY_iSCSI2_IP	Optional	10.10.12.11	<p>If the connection to the ETERNUS DX disk array is redundant, specify the other IP address of the ETERNUS DX disk array iSCSI-CA port to which the connection will be made.</p> <p>When the iSCSI LAN is a single configuration, this value should be omitted.</p>

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
				If building a master server in a virtual environment, this value should be omitted.
31	BDPP_ROR_ADMIN_IP	Optional	192.168.1.1	Specify the primary master server admin LAN IP address to be specified when installing ROR. If building a master server in a virtual environment, this value should be omitted.
32	BDPP_ROR_ADMIN_ACCOUNT	Optional	root	Specify the account of the local server used when installing ROR. If building a master server in a virtual environment, this value should be omitted.
33	BDPP_ROR_ADMIN_PASS	Optional	password	Specify the password of the local server used when installing ROR. If building a master server in a virtual environment, this value should be omitted.
34	BDPP_PDFS_FILESYSTEM	Mandatory	YES	If DFS is to be used as the file system, specify YES (or yes).
35	BDPP_PDFS_MOUNTPOINT	Optional	/mnt/pdfs	Specify the PDFS mount point. This parameter is mandatory if YES is specified for BDPP_PDFS_FILESYSTEM.
36	BDPP_HADOOP_DEFAULT_USERS	Optional	bdppuser1,1500	To register a user ID with permissions for the execution of MapReduce and access to the file system (DFS), specify the user name and UID. Multiple user IDs can be registered (*2).
37	BDPP_HADOOP_DEFAULT_GROUP	Optional	bdppgroup,1500	Specify this to register a user group name and GID for the execution of MapReduce and to access the file system (DFS).
38	BDPP_HADOOP_TOP_DIR	Optional	/hadoop	Specify the name of the top directory that stores the data used in Hadoop on the file system (DFS). The data used in Hadoop and the results data processed in Hadoop will be stored under the directory that is specified here.
39	BDPP_KVMHOST_IP		192.168.1.100	Specify the IP address of the admin LAN for the KVM host

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
				machine used in the shutdown agent settings (*3). Specify this only when building a master server in a virtual environment (KVM).
40	BDPP_KVMHOST_ACCOUNT		pcluser	Specify the user account for the KVM host machine used in the shutdown agent settings. Specify this only when building a master server in a virtual environment (KVM).
41	BDPP_KVMHOST_PASS		password	Specify the user password for the KVM host machine used in the shutdown agent settings. Specify this only when building a master server in a virtual environment (KVM).

*1: An interconnect private IP interface can be built by using an IP address provided for a private network.

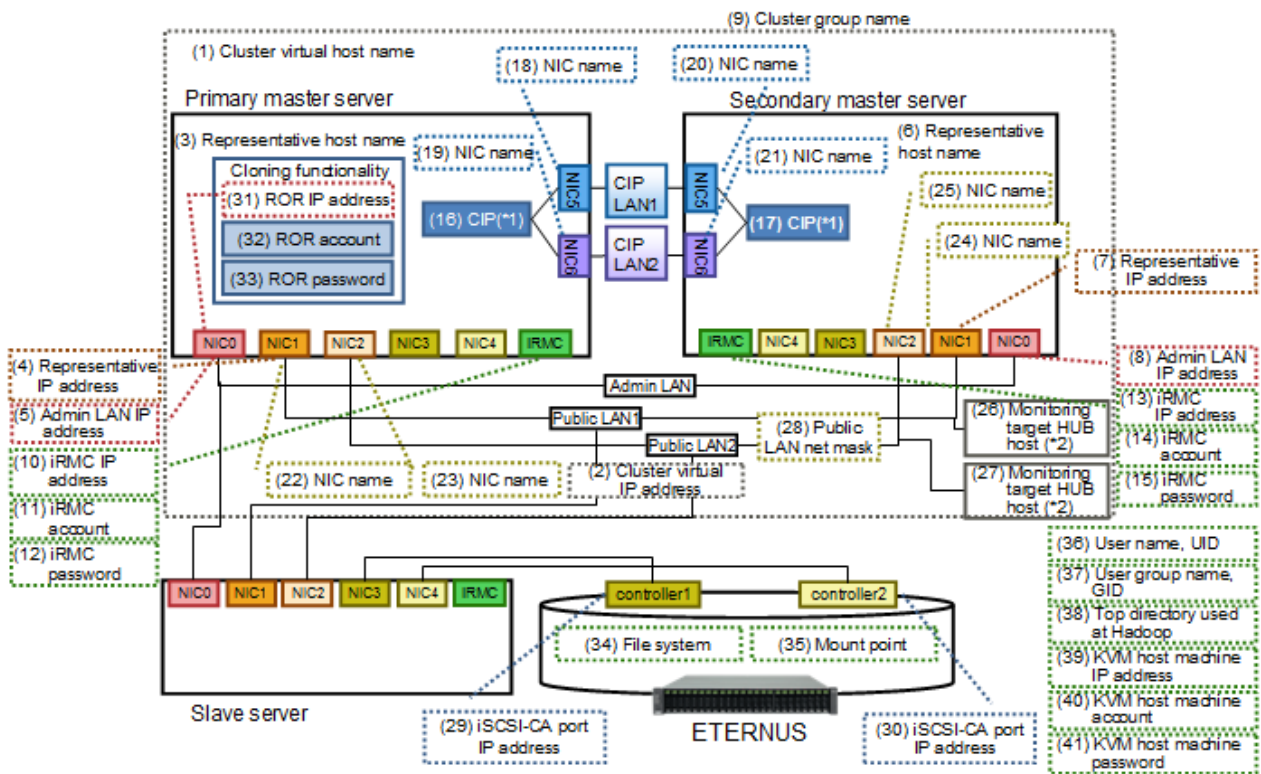
*2: When registering multiple user IDs, specify multiple user names and user IDs by delimiting with a space (" ") as shown below.

```
bdppuser1,1500 bdppuser2,1501 bdppuser3,1502
```

*3: If replicated configuration is used for master servers and the primary master server and the secondary master server are installed on virtual machines that are on different host machines, specify the IP addresses for the two host machines by delimiting with a comma (",").

```
192.168.1.100,192.168.1.101
```

Figure B.1 Network configuration of the master server and bdpp.conf parameters



*1: CIP (Cluster Interconnect Protocol)

*2: HUB host used for the HUB monitoring feature.

The HUB monitoring feature executes a ping to neighboring hubs at fixed intervals, and if it detects a fault in the transfer route, it switches the interface being used.

[Slave server]

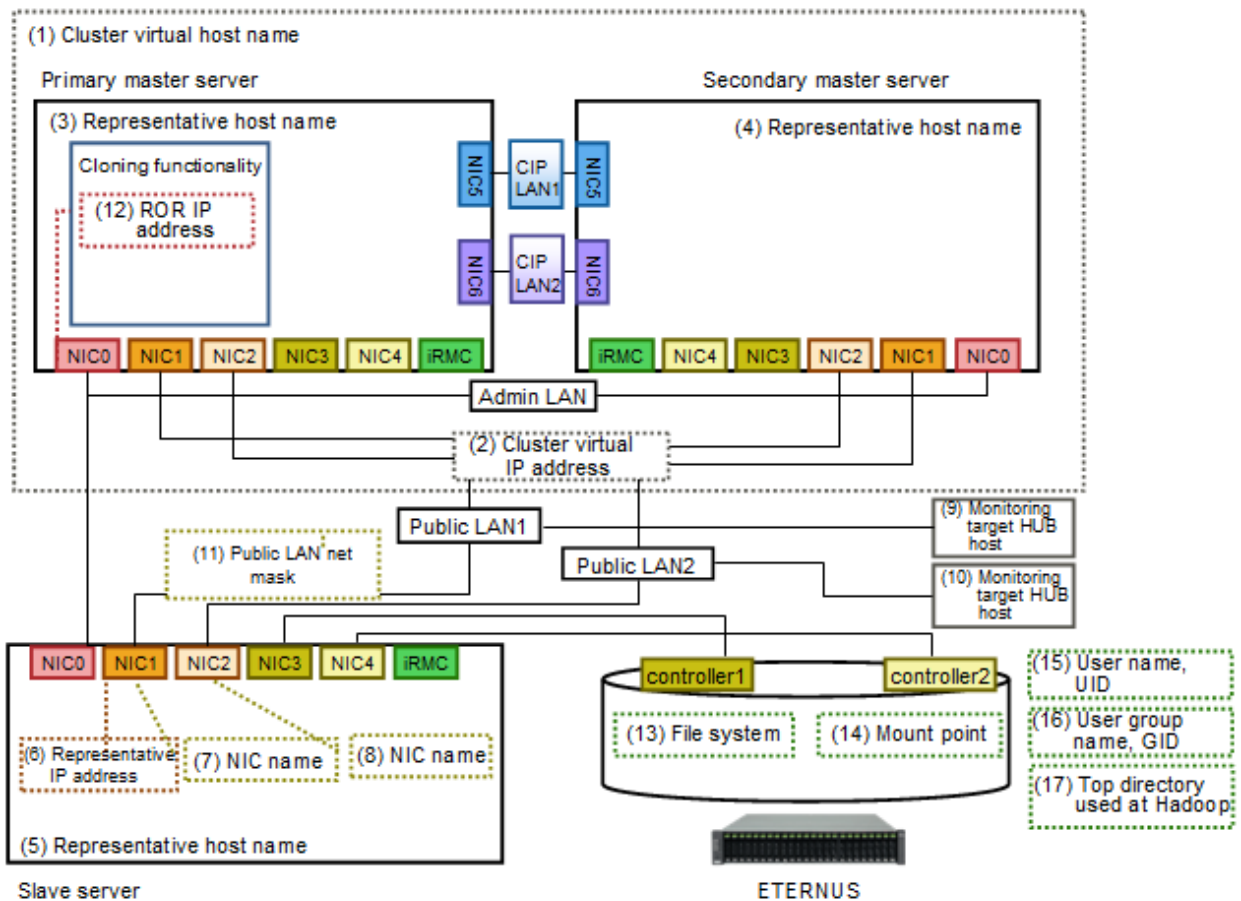
Parameters set in the slave server configuration file are shown below.

Item No.	Parameter	Mandatory/Optional	Value to be specified (settings example)	Explanation
1	BDPP_CLUSTER_NAME	Mandatory	cluster	Set the same value as the master server.
2	BDPP_CLUSTER_IP	Mandatory	10.10.10.1	Set the same value as the master server.
3	BDPP_PRIMARY_NAME	Mandatory	master1	Set the same value as the master server.
4	BDPP_SECONDARY_NAME	Optional	master2	Set the same value as the master server.
5	BDPP_SERVER_NAME	Mandatory	slave1	Specify the representative host name of the slave server (the host name corresponding to the NIC that connects to the public LAN). For the host name, the first character must be a letter, and it

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
				must consist of 11 or fewer alphanumeric characters.
6	BDPP_SERVER_IP	Mandatory	10.10.10.21	Specify the representative IP address that will be used to connect to the slave server public LAN.
7	BDPP_GLS_SERVER_CONNECT1	Mandatory	eth1	Specify the slave server NIC name (NIC1) used for connections between slave servers. (The NIC that will be used to connect to public LAN1 in the figure)
8	BDPP_GLS_SERVER_CONNECT2	Optional	eth2	Specify the slave server NIC name (NIC2) used for connections between slave servers. (The NIC that will be used to connect to public LAN2 in the figure) When the public LAN is a single configuration, this value should be omitted. If building a slave server in a virtual environment, this value should be omitted.
9	BDPP_GLS_SERVER_POLLING1_IP	Mandatory	10.10.10.100	Specify the IP address of the monitoring target HUB host. (Monitoring target HUB host on public LAN1 in the figure)
10	BDPP_GLS_SERVER_POLLING2_IP	Optional	10.10.10.101	Specify the IP address of the monitoring target HUB host. (Monitoring target HUB host on public LAN2 in the figure) When the public LAN is a single configuration, this value should be omitted. If building a slave server in a virtual environment, this value should be omitted.
11	BDPP_GLS_NETMASK	Optional	255.255.0.0	Specify the net mask for the public LAN IP address. The default net mask is "255.255.255.0".
12	BDPP_ROR_ADMIN_IP	Optional	192.168.1.1	Set the same value as the master server.

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
13	BDPP_PDFS_FILESYSTEM	Mandatory	YES	Set the same value as the master server.
14	BDPP_PDFS_MOUNTPOINT	Optional	/mnt/pdfs	Set the same value as the master server.
15	BDPP_HADOOP_DEFAULT_USERS	Optional	bdppuser1,1500	Set the same value as the master server.
16	BDPP_HADOOP_DEFAULT_GROUP	Optional	bdppgroup,1500	Set the same value as the master server.
17	BDPP_HADOOP_TOP_DIR	Optional	/hadoop	Set the same value as the master server.

Figure B.2 Slave server network configuration and bdpp.conf parameter

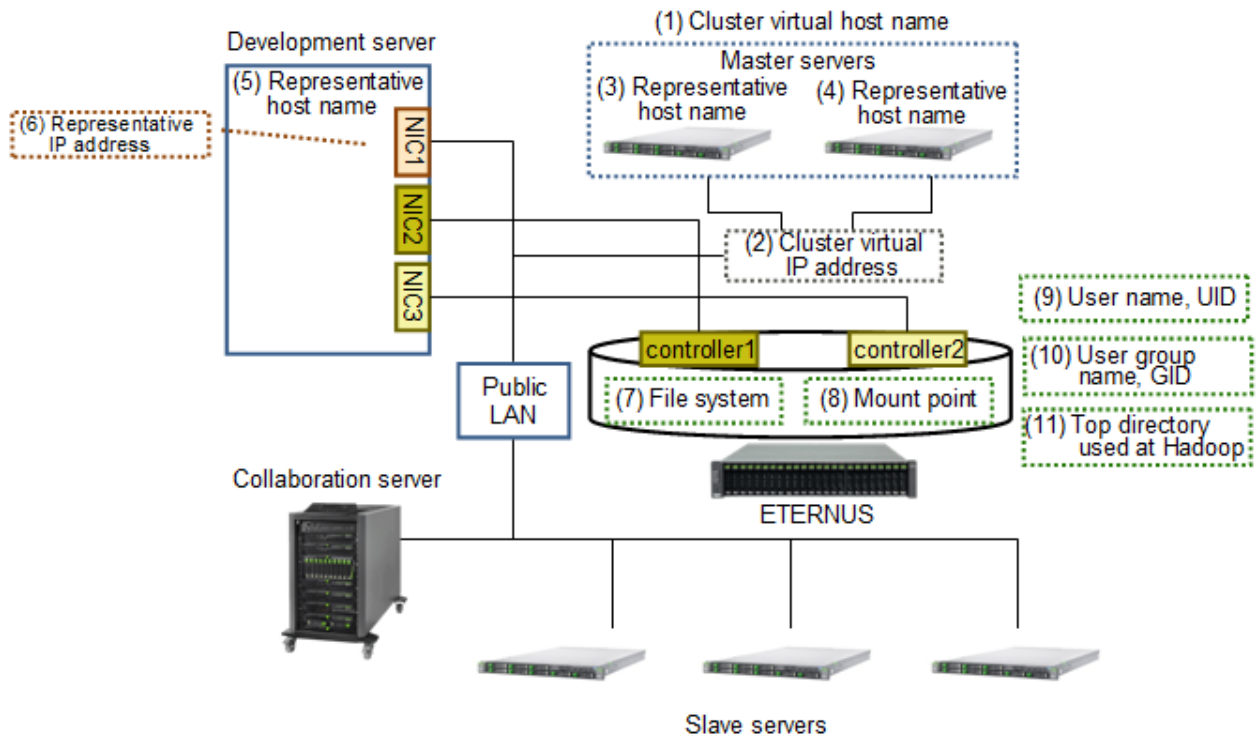


[Development server]

Parameters set in the development server configuration file are shown below.

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
1	BDPP_CLUSTER_NAME	Mandatory	cluster	Set the same value as the master server.
2	BDPP_CLUSTER_IP	Mandatory	10.10.10.1	Set the same value as the master server.
3	BDPP_PRIMARY_NAME	Mandatory	master1	Set the same value as the master server.
4	BDPP_SECONDARY_NAME	Optional	master2	Set the same value as the master server.
5	BDPP_SERVER_NAME	Mandatory	develop	Specify the representative host name of the development server (the host name corresponding to the NIC that connects to the public LAN). For the host name, the first character must be a letter, and it must consist of 11 or fewer alphanumeric characters.
6	BDPP_SERVER_IP	Mandatory	10.10.10.30	Specify the representative IP address of the development server.
7	BDPP_PDFS_FILESYSTEM	Mandatory	YES	Set the same value as the master server.
8	BDPP_PDFS_MOUNTPOINT	Optional	/mnt/pdfs	Set the same value as the master server.
9	BDPP_HADOOP_DEFAULT_USERS	Optional	bdppuser1,1500	Set the same value as the master server.
10	BDPP_HADOOP_DEFAULT_GROUP	Optional	bdppgroup,1500	Set the same value as the master server.
11	BDPP_HADOOP_TOP_DIR	Optional	/hadoop	Set the same value as the master server.

Figure B.3 Development server network configuration and bdpp.conf parameter



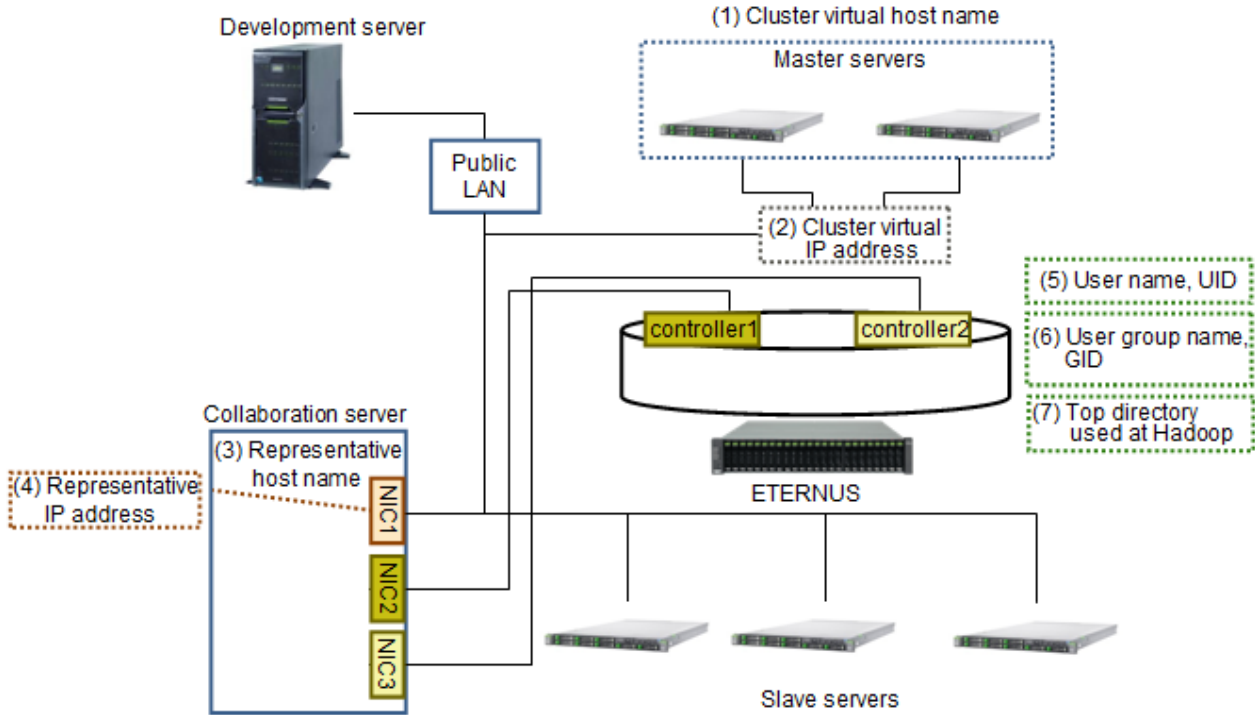
[Collaboration server]

Parameters set in the collaboration server configuration file are shown below.

Item No.	Parameter	Mandatory/Optional	Value to be specified (settings example)	Explanation
1	BDPP_SERVER_NAME	Mandatory	collaborate	Specify the representative host name of the collaboration server (the host name corresponding to the NIC that connects to the public LAN). For the host name, the first character must be a letter, and it must consist of 11 or fewer alphanumeric characters.
2	BDPP_SERVER_IP	Mandatory	10.10.10.31	Specify the representative IP address of the collaboration server.
3	BDPP_PDFS_FILESYSTEM	Mandatory	YES	Set the same value as the master server.
4	BDPP_PDFS_MOUNTPOINT	Optional	/mnt/pdfs	Set the same value as the master server.
5	BDPP_HADOOP_DEFAULT_USERS	Optional	bdppuser1,1500	Set the same value as the master server.
6	BDPP_HADOOP_DEFAULT_GROUP	Optional	bdppgroup,1500	Set the same value as the master server.

Item No.	Parameter	Mandatory/Optional	Value to be specified (settings example)	Explanation
7	BDPP_HADOOP_TOP_DIR	Optional	/hadoop	Set the same value as the master server.

Figure B.4 Collaboration server network configuration and bdpp.conf parameter



Definition examples

Definition examples for bdpp.conf are shown below.

These definition examples are provided for various servers for the following system configurations:

- Replicating the master server configuration
- Replicating the public LAN
- Replicating the cluster interconnect (CIP)
- Replicating the iSCSI

[Master server]

```
BDPP_CLUSTER_NAME=cluster
BDPP_CLUSTER_IP=10.10.10.1
BDPP_PRIMARY_NAME=master1
BDPP_PRIMARY_IP=10.10.10.11
BDPP_PRIMARY_ADM_IP=192.168.1.1
BDPP_SECONDARY_NAME=master2
BDPP_SECONDARY_IP=10.10.10.12
BDPP_SECONDARY_ADM_IP=192.168.1.2
BDPP_PCL_CLUSTER_GROUP=bdppcluster
BDPP_PCL_PRIMARY_IRMC_IP=192.168.1.11
```

```
BDPP_PCL_PRIMARY_irmc_ACCOUNT=irmc
BDPP_PCL_PRIMARY_irmc_PASS=password
BDPP_PCL_SECONDARY_irmc_IP=192.168.1.12
BDPP_PCL_SECONDARY_irmc_ACCOUNT=irmc
BDPP_PCL_SECONDARY_irmc_PASS=password
BDPP_PCL_PRIMARY_CIP=192.168.2.11
BDPP_PCL_SECONDARY_CIP=192.168.2.12
BDPP_PCL_PRIMARY_CONNECT1=eth5
BDPP_PCL_PRIMARY_CONNECT2=eth6
BDPP_PCL_SECONDARY_CONNECT1=eth5
BDPP_PCL_SECONDARY_CONNECT2=eth6
BDPP_GLS_PRIMARY_CONNECT1=eth1
BDPP_GLS_PRIMARY_CONNECT2=eth2
BDPP_GLS_SECONDARY_CONNECT1=eth1
BDPP_GLS_SECONDARY_CONNECT2=eth2
BDPP_GLS_POLLING1_IP=10.10.10.100
BDPP_GLS_POLLING2_IP=10.10.10.101
BDPP_DISKARRAY_iscsi1_IP=10.10.11.11
BDPP_DISKARRAY_iscsi2_IP=10.10.12.11
BDPP_ROR_ADMIN_IP=192.168.1.1
BDPP_ROR_ADMIN_ACCOUNT=admin
BDPP_ROR_ADMIN_PASS=password
BDPP_PDFS_FILESYSTEM=YES
BDPP_PDFS_MOUNTPOINT=/mnt/pdfs
BDPP_HADOOP_DEFAULT_USERS=bdppuser1,1500
BDPP_HADOOP_DEFAULT_GROUP=bdppgroup,1500
BDPP_HADOOP_TOP_DIR=/hadoop
```

[Slave server]

```
BDPP_CLUSTER_NAME=cluster
BDPP_CLUSTER_IP=10.10.10.1
BDPP_PRIMARY_NAME=master1
BDPP_SECONDARY_NAME=master2
BDPP_SERVER_NAME=slave1
BDPP_SERVER_IP=10.10.10.21
BDPP_GLS_SERVER_CONNECT1=eth1
BDPP_GLS_SERVER_CONNECT2=eth2
BDPP_GLS_SERVER_POLLING1_IP=10.10.10.100
BDPP_GLS_SERVER_POLLING2_IP=10.10.10.101
BDPP_ROR_ADMIN_IP=192.168.1.1
BDPP_PDFS_FILESYSTEM=YES
BDPP_PDFS_MOUNTPOINT=/mnt/pdfs
BDPP_HADOOP_DEFAULT_USERS=bdppuser1,1500
BDPP_HADOOP_DEFAULT_GROUP=bdppgroup,1500
BDPP_HADOOP_TOP_DIR=/hadoop
```

[Development server]

```
BDPP_CLUSTER_NAME=cluster
BDPP_CLUSTER_IP=10.10.10.1
BDPP_PRIMARY_NAME=master1
BDPP_SECONDARY_NAME=master2
BDPP_SERVER_NAME=develop
BDPP_SERVER_IP=10.10.10.30
BDPP_PDFS_FILESYSTEM=YES
BDPP_PDFS_MOUNTPOINT=/mnt/pdfs
BDPP_HADOOP_DEFAULT_USERS=bdppuser1,1500
BDPP_HADOOP_DEFAULT_GROUP=bdppgroup,1500
BDPP_HADOOP_TOP_DIR=/hadoop
```

[Collaboration server]

```
BDPP_SERVER_NAME=collaborate
BDPP_SERVER_IP=10.10.10.31
BDPP_PDFS_FILESYSTEM=YES
BDPP_PDFS_MOUNTPOINT=/mnt/pdfs
BDPP_HADOOP_DEFAULT_USERS=bdppuser1,1500
BDPP_HADOOP_DEFAULT_GROUP=bdppgroup,1500
BDPP_HADOOP_TOP_DIR=/hadoop
```

B.2 slaves

This section explains settings in the "slaves" slave server definition file.

File storage location

```
/etc/opt/FJSVbdpp/conf/slaves
```

File content

This file defines the slave servers targeted for connection.

Format

Specify the host name of the slave server (the host name corresponding to the NIC that connects to the public LAN).

For multiple host names, delimit with a comma (",").

```
slaveServer1[,slaveServer2][, ...]
```

Definition example

```
slave1,slave2,slave3,slave4,slave5
```

B.3 clone.conf

This section explains the settings in the clone.conf cloning server definition file.

File storage location

Any directory

File content

This file is used to define the slave server that creates the clone image and the slave server that is added using Smart Setup.

Format

Specify items to be defined on separate lines by delimiting with a comma (",").

Use the format below to code each line.

File format definition

It is essential that the first line of the file starts with the following content:

```
RCXCSV,V3.4
```

Comments

Content defined on the following lines is treated as a comment and skipped.

- Lines in which the leading character of the string is a comment symbol (#)
- Lines in which there is only a space (), tab symbol, or line feed

- Lines in which there is only a comma (,)
- Resource definitions that cannot be recognized

Example

Comment line

```
# Development environment definition
```

Resource definition


Define the Server section and the ServerAgent section.


[Section name]

Specify [Server].

[Section header]

Specify the section headers and values shown below.

Section header	Value to be specified (settings example)	Content
operation	new	Specify the following as the resource operation content: Register: new Change: change Hyphen (-): Do nothing
chassis_name	-	Omit this information.
slot_no	-	Omit this information.
server_name	slave1	Specify the name of the slave server (the host name specified in the HOSTNAME parameter of the /etc/sysconfig/network file). For the server name, the first character must be a letter, and it must consist of 11 or fewer alphanumeric characters.  Note If the server name of the slave server is within 15 characters, then input is possible, but as the master server must be 11 or fewer characters, then the server name for the slave server must also be 11 or fewer characters to preserve consistency across the system.
ip_address	192.168.1.21	Specify the IP address that is assigned to the NIC that connects to the admin LAN set in the slave server.
mac_address	A0-A0-A0-A0-A0-A0	Specify the NIC MAC address that connects to the admin LAN. Enter the address using either the hyphen ("-") or colon (:) delimited format ("xx-xx-xx-xx-xx-xx" or xx:xx:xx:xx:xx:xx).

Section header	Value to be specified (settings example)	Content
second_mac_address	-	Omit this information.
snmp_community_name	public	Specify the SNMP community name (reference permission) to be set for the server. Specify a string of up to 32 characters comprised of single-byte alphanumerics, underscores (_), and hyphens (-).
ipmi_ip_address	192.168.1.121	Specify the IP address of the remote management controller that manages the server.
ipmi_user_name	admin	Specify a user name having administrator permissions or higher for the remote management controller that manages the server. Specify a string of up to 16 characters comprised of single-byte alphanumerics and ASCII character symbols (0x20 to 0x7e).
ipmi_passwd	admin	Specify the password of the remote management controller that manages the server. Specify a string of up to 16 characters comprised of single-byte alphanumerics and ASCII character symbols (0x20 to 0x7e). Nothing needs to be specified if a password is not set.  Note If a password that is 17 characters or more is already set for the remote management controller, either add a new user or reset the password using 16 characters or less.
ipmi_passwd_enc	plain	Specify one of the following: - If the ipmi_passwd character string is plain text: "plain" - If encrypted: "encrypted"
admin_lan1_nic_number	-	Omit this information.
admin_lan2_nic_number	-	Omit this information.
admin_lan_nic_redundancy	OFF	Specify OFF for this information.

[Section name]

Specify [ServerAgent].

[Section header]

Specify the section headers and values shown below.

Section header	Value to be specified (settings example)	Content
operation	new	Specify the following as the resource operation content: Register: new

Section header	Value to be specified (settings example)	Content
		Change: change Hyphen (-): Do nothing
server_name	slave1	Specify the name of the slave server that creates the clone image (the host name specified in the HOSTNAME parameter of the /etc/sysconfig/network file). For the server name, the first character must be a letter, and it must consist of 11 or fewer alphanumeric characters.

Definition example

```

RCXCSV, V3.4
[Server]
operation, chassis_name, slot_no, server_name, ip_address, mac_address, second_mac_address, snmp_community_name, ipmi_ip_address, ipmi_user_name, ipmi_passwd, ipmi_passwd_enc, admin_lan1_nic_number, admin_lan2_nic_number, admin_lan_nic_redundancy
new, , slave1, 192.168.1.21, A0:A0:A0:A0:A0:A1, , public, 192.168.1.121, admin, admin, plain, , , OFF
new, , slave2, 192.168.1.22, A0:A0:A0:A0:A0:A2, , public, 192.168.1.122, admin, admin, plain, , , OFF
new, , slave3, 192.168.1.23, A0:A0:A0:A0:A0:A3, , public, 192.168.1.123, admin, admin, plain, , , OFF
new, , slave4, 192.168.1.24, A0:A0:A0:A0:A0:A4, , public, 192.168.1.124, admin, admin, plain, , , OFF
new, , slave5, 192.168.1.25, A0:A0:A0:A0:A0:A5, , public, 192.168.1.125, admin, admin, plain, , , OFF
[ServerAgent]
operation, server_name
new, slave1

```

B.4 FJSVrcx.conf

This section explains the settings in the FJSVrcx.conf network parameter automatic configuration file.

File storage location

Any directory

File content

Define the network parameter to be configured automatically after cloning.

Format

Configure using the format shown below.

```
parameter=specifiedValue
```

Specify the following as the parameters with their respective values.

Parameter	Value to be specified (settings example)	Content
admin_LAN	192.168.1.21	Specify the IP address of the admin LAN of the slave server creating the clone image. When cloning in a virtual environment, specify the IP address of the admin LAN for the clone source slave server.
hostname	slave1	Specify the name of the slave server that creates the clone image (the host name specified in the HOSTNAME parameter of the /etc/sysconfig/network file).

Parameter	Value to be specified (settings example)	Content
		When cloning in a virtual environment, specify the name of the clone source slave server. For the server name, the first character must be a letter, and it must consist of 11 or fewer alphanumeric characters.

Definition example

```
admin_LAN=192.168.1.21
hostname=slave1
```

B.5 ipaddr.conf

This section explains the settings in the ipaddr.conf network parameter automatic configuration file.

File storage location

Any directory

File content

When using Smart Setup to add slave servers, define the public LAN and the iSCSI-LAN for slave servers to be cloned.

If adding a slave server to a virtual environment, define the admin LAN information.

Format

Configured from the following entries:

- One or more node entries
- One or more interface entries under the node entry (with/without redundancy)

Each entry is configured using the format shown below.

```
keyword="specifiedValue"
```

Node entries

Define the value that is set for the definition keyword shown in the table below.



Note

Enter one entry per server, including all server names that may be added in the future.

Defined content	Keyword	Value to be set	Explanation
Slave server name	NODE_NAME	Slave server name	Specify the name of the slave server (the host name specified in the HOSTNAME parameter of the /etc/sysconfig/network file).

Interface entry (without redundancy)

Define the value that is set for the definition keyword shown in the table below.

Define the interface entry by specifying a number between 0 and 99 at the end of the definition keyword for each interface name.

Note

- A value in ascending order starting from 0 must be used as the number that is added at the end of the definition keyword.
- When building a redundant iSCSI-LAN, define two interfaces without redundancy, instead of defining an interface with redundancy.

Defined content	Keyword	Value to be set	Explanation
Interface name	IF_NAME	eth X	X is an integer; 0 or greater.
IP address	IF_IPAD	xxx.xxx.xxx.xxx format IP address	-
Subnet mask	IF_MASK	xxx.xxx.xxx.xxx format subnet mask	-

Interface entry (with redundancy)

Define the value that is set for the definition keyword shown in the table below.

Define the interface entry by specifying a number between 0 and 99 at the end of the definition keyword for each interface name.

Interfaces with and without redundancy can be used together within a single node entry, as long as the interface name is not duplicated.

Note

- A value in ascending order starting from 0 must be used as the number that is added at the end of the definition keyword.
- If an interface with redundancy is used together with an interface without redundancy, the value that is set for the interface name must always be in ascending order.
- Only define the interface with redundancy when adding slave servers to a physical environment and making the public LAN redundant.
The interface with redundancy cannot be defined when adding slave servers to a virtual environment.

Defined content	Keyword	Value to be set	Explanation
Virtual interface name	VIF_NAME	sha0	-
IP address that is set for the virtual interface	VIF_IPAD	xxx.xxx.xxx.xxx format IP address	-
Subnet mask	VIF_MASK	xxx.xxx.xxx.xxx format subnet mask	-
Primary interface name	PRI_NAME	Interface name (eth X)	X is an integer; 0 or greater. Where there is a pair of interface names, set the primary interface name.
Secondary interface name	SCD_NAME	Interface name (eth Y)	Y is an integer; 0 or greater. Where there is a pair of interface names, set the secondary interface name.
Monitoring target IP address	POL_ADDR	xxx.xxx.xxx.xxx format IP address	Specify the IP address of the monitoring target HUB hosts by delimiting with a comma (","). Specify BDPP_GLS_SERVER_POLLING1_IP

Defined content	Keyword	Value to be set	Explanation
			and BDPP_GLS_SERVER_POLLING2_IP for bdpp.conf of slave servers whose images will be cloned.
Standby patrol virtual interface name	PAT_NAME	sha1	-
Existence of HUB- HUB monitoring	POL_HUBS	OFF	Specify OFF as HUB-HUB monitoring will not be performed.



Refer to "2.4 Monitoring function of NIC switching mode" in the "PRIMECLUSTER Global Link Services Configuration and Administration Guide 4.3 (Redundant Line Control Function)" for information on standby patrol and HUB-HUB monitoring.

Definition examples

Definition examples for the following configurations (when cloning in a physical environment) are shown below.

- NIC that connects to the public LAN : eth1, eth2 (redundancy added)
- NIC that connects to the iSCSI-LAN : eth3, eth4 (redundancy added)

```

NODE_NAME="slave1" <-- Slave server that will create the clone image
VIF_NAME0="sha0"
VIF_IPAD0="10.10.10.21"
VIF_MASK0="255.255.255.0"
PRI_NAME0="eth1"
SCD_NAME0="eth2"
POL_ADDR0="10.10.10.100,10.10.10.101"
PAT_NAME0="sha1"
POL_HUBS0="OFF"
IF_NAME1="eth3"
IF_IPAD1="10.10.11.21"
IF_MASK1="255.255.255.0"
IF_NAME2="eth4"
IF_IPAD2="10.10.12.21"
IF_MASK2="255.255.255.0"

NODE_NAME="slave2" <-- Slave server to which the clone image will be deployed
VIF_NAME0="sha0"
VIF_IPAD0="10.10.10.22"
VIF_MASK0="255.255.255.0"
PRI_NAME0="eth1"
SCD_NAME0="eth2"
POL_ADDR0="10.10.10.100,10.10.10.101"
PAT_NAME0="sha1"
POL_HUBS0="OFF"
IF_NAME1="eth3"
IF_IPAD1="10.10.11.22"
IF_MASK1="255.255.255.0"
IF_NAME2="eth4"
IF_IPAD2="10.10.12.22"
IF_MASK2="255.255.255.0"

NODE_NAME="slave3" <-- Slave server to which the clone image will be deployed
VIF_NAME0="sha0"
VIF_IPAD0="10.10.10.23"

```

```
VIF_MASK0="255.255.255.0"  
PRI_NAME0="eth1"  
SCD_NAME0="eth2"  
POL_ADDR0="10.10.10.100,10.10.10.101"  
PAT_NAME0="sha1"  
POL_HUBS0="OFF"  
IF_NAME1="eth3"  
IF_IPAD1="10.10.11.23"  
IF_MASK1="255.255.255.0"  
IF_NAME2="eth4"  
IF_IPAD2="10.10.12.23"  
IF_MASK2="255.255.255.0"  
  
:
```

B.6 initiator.conf

This section explains the settings in the iSCSI initiator configuration definition file `initiator.conf`.

File storage location

Any directory

File content

Define the iSCSI name for each server defined in the ETERNUS DX disk array.

Format

Configure using the format shown below.

```
serverName, initiatorName
```

serverName

Specify the name of the slave server (the host name specified in the `HOSTNAME` parameter of the `/etc/sysconfig/network` file).

initiatorName

Specify the iSCSI name defined in the ETERNUS DX disk array for the relevant server.

Definition examples

```
slave1, iqn.1994-05.com.redhat:875816c7c4ba  
slave2, iqn.1994-05.com.redhat:875816c7c4bb  
slave3, iqn.1994-05.com.redhat:875816c7c4bc  
slave4, iqn.1994-05.com.redhat:875816c7c4bd  
slave5, iqn.1994-05.com.redhat:875816c7c4be
```

Appendix C Hadoop Configuration Parameters

This appendix explains the configuration parameters for Hadoop provided by this product.

Some parameters are configured automatically when setting up this product. This includes parameters for Hadoop to use a DFS and parameters for optimizing parallel distributed processing of Hadoop jobs.

This appendix shows parameters that are changed from default values and that are added to OS setup files within the setup files (excluding pdfs-site.xml) provided by the open source software Apache Hadoop.

List of setup files

Setup file	Description	Directory
hadoop-env.sh	Setup file for defining environment variables used in Hadoop	/etc/hadoop
core-site.xml	Common setup file for Hadoop	/etc/hadoop
mapred-site.xml	MapReduce-related setup file	/etc/hadoop
pdfs-site.xml	DFS-related setup file	/etc/hadoop
sysctl.conf	OS setup file specifying kernel parameters	/etc
limits.conf	OS setup file for limiting system resources	/etc/security

Note

- This appendix does not provide detailed explanation of all the parameters in setup files. If necessary, refer to a resource such as the Apache Hadoop project website.
- If tuning the configuration settings, make changes immediately following the setup. This excludes when changes are made to the number of slave servers (see "[12.1.7 Changing Hadoop Configuration Parameters](#)").
Note that if any changes need to be made to the Hadoop configuration parameters after operations begin, settings will need to be reviewed for each master server, slave server, and development server.

C.1 hadoop-env.sh

This section explains the environment variables set in hadoop-env.sh.

Environment variable	Default value/Value to be set
HADOOP_CLASSPATH CLASSPATH of the JVM that operates Hadoop	Default value None Value to be set \$HADOOP_CLASSPATH:/opt/FJSVpdfs/lib/pdfs.jar
HADOOP_JOBTRACKER_OPTS Launch option specifically for the JVM that operates the JobTracker	Default value -Dcom.sun.management.jmxremote \$HADOOP_JOBTRACKER_OPTS Value to be set -Xmx\$\$m -Dcom.sun.management.jmxremote - Dcom.sun.management.jmxremote.port=60001 - Dcom.sun.management.jmxremote.ssl=false - Dcom.sun.management.jmxremote.authenticate=false - Djava.rmi.server.hostname=localhost - Dhadoop.security.logger=INFO,DRFAS - Dmapred.audit.logger=INFO,MRAUDIT -

Environment variable	Default value/Value to be set
	Dhadoop.mapreduce.jobsummary.logger=INFO,JSA \$HADOOP_JOBTRACKER_OPTS \$\$ = <i>physicalMemorySizeInMb</i> / 3
HADOOP_OPTS Launch option for the JVM that operates Hadoop	Default value None Value to be set -Djava.net.preferIPv4Stack=true \$HADOOP_OPTS
HADOOP_SSH_OPTS Option specified in ssh when Hadoop uses ssh	Default value None Value to be set -o StrictHostKeyChecking=no -o BatchMode=yes
HADOOP_TASKTRACKER_OPTS Launch option specifically for the JVM that operates the TaskTracker	Default value None Value to be set -Xmx\$\$m -Dcom.sun.management.jmxremote -Dcom.sun.management.jmxremote.port=60002 - Dcom.sun.management.jmxremote.ssl=false - Dcom.sun.management.jmxremote.authenticate=false - Djava.rmi.server.hostname=localhost - Dhadoop.security.logger=ERROR,console - Dmapred.audit.logger=ERROR,console \$HADOOP_TASKTRACKER_OPTS \$\$ = The lesser of the following values: - 512 - <i>physicalMemorySizeInMb</i> / 6
HADOOP_USER_CLASSPATH_FIRST Add HADOOP_CLASSPATH specification to the beginning of the existing CLASSPATH	Default value None Value to be set true
JAVA_HOME Path of the JVM used by Hadoop	Default value None Value to be set JVM installation directory



Note

In some cases, values to be set for some parameters (JVM-specific launch options) may be displayed across several lines, even though the specified value does not actually include a newline.

C.2 core-site.xml

This section describes the properties to be set in core-site.xml.

Property	Default value/Value to be set
fs.default.name URI that provides access to file systems used as default	Default value file:///
	Value to be set pdfs:///
fs.pdfs.impl Allocate the file system class to any Scheme	Default value None
	Value to be set com.fujitsu.pdfs.fs.PdfsDirectFileSystem
io.file.buffer.size Size of the buffer (bytes) used when reading or writing data from the disk or network	Default value 4096 (4 KB)
	Value to be set 131072 (128 KB)
ipc.server.listen.queue.size Number of simultaneous connection requests from other servers that can be processed	Default value 128
	Value to be set 1024 (*1)

*1: Specify the same value as net.core.somaxconn in "C.5 sysctl.conf".

C.3 mapred-site.xml

This section describes the properties to be set in mapred-site.xml.

Property	Default value/Value to be set
io.sort.factor Number of process results (segments) that will be merged for each map written to the disk	Default value 10
	Value to be set 50
io.sort.mb Size of memory (MB) for retaining output results of Map tasks	Default value 100
	Value to be set The lesser of the following values: - 2047 - <i>valueOf\$\$In_mapred.child.java.opts</i> * 0.5
mapred.child.java.opts Launch option specified in the JVM that executes Map/Reduce tasks	Default value -Xmx200m
	Value to be set

Property	Default value/Value to be set
	-server -Xmx\$\$m -Djava.net.preferIPv4Stack=TRUE $$$ = (\text{physicalMemorySizeInMb} - 2048) / (\text{mapred.tasktracker.map.tasks.maximum} + \text{mapred.tasktracker.reduce.tasks.maximum})$
mapred.child.ulimit Maximum size (KB) of process (address) space for Map/Reduce tasks	Default value None Value to be set 0 (unlimited)
mapred.compress.map.output Whether to compress Map task output results	Default value false Value to be set true
mapred.local.dir Intermediate file storage directory for MapReduce jobs	Default value \${hadoop.tmp.dir}/mapred/local Value to be set /var/lib/hadoop/mapred/local
mapred.max.tracker.failures Maximum number of retries within the same TaskTracker when Map/Reduce tasks fail	Default value 4 Value to be set 40
mapred.reduce.parallel.copies Multiplicity of Map results of other TaskTrackers obtained by the TaskTracker that executes Reduce tasks	Default value 5 Value to be set 20
mapred.reduce.tasks Maximum number of Reduce tasks operated within a MapReduce job	Default value 1 Value to be set mapred.tasktracker.reduce.tasks.maximum * <i>numberOfSlaveServers</i>
mapred.task.tracker.http.address Port number of the TaskTracker HTTP server	Default value 0.0.0.0:50060 Value to be set 0.0.0.0:50060 (reset)
mapred.tasktracker.map.tasks.maximum Number of Map tasks executed simultaneously by a single TaskTracker	Default value 2 Value to be set The greater of the following values: - <i>numberOfCpuCores</i> -1


Property	Default value/Value to be set
	- <i>totalNumberOfPhysicalDisksComprisingThePdfs</i> / <i>numberOfSlaveServers</i> (rounded up to a whole number)
mapred.tasktracker.reduce.tasks.maximum Number of Reduce tasks executed simultaneously by a single TaskTracker	Default value 2 Value to be set The greater of the following values: - <i>numberOfCpuCores</i> -1 - <i>totalNumberOfLunsComprisingThePdfs</i> / <i>numberOfSlaveServers</i> (rounded up to a whole number)
mapred.userlog.limit.kb Maximum value of a userlog output by a task	Default value 0 Value to be set 1024
mapred.userlog.retain.hours Time (hours) a userlog is retained after a job is completed	Default value 24 (1 day) Value to be set 168 (1 week) (*1)
mapreduce.history.server.embedded Whether to launch the JVM dedicated to job histories or operate using the JobTracker JVM	Default value None Value to be set true
mapreduce.tasktracker.group Group to which the TaskTracker process belongs	Default value None Value to be set hadoop
mapreduce.tasktracker.outofband.heartbeat Whether to push forward a survival notification to the JobTracker when a task is completed	Default value false Value to be set false (reset)

*1: Specify within a range acceptable for HADOOP_LOG_DIR.

C.4 pdfs-site.xml

This section describes the properties to be set in pdfs-site.xml.

Property	Value to be set
pdfs.fs.local.basedir DFS mount directory path	Path joining BDPP_PDFS_MOUNTPOINT and BDPP_HADOOP_TOP_DIR in bdpp.conf

Property	Value to be set
	 Example If the bdpp.conf configuration file has the following settings, this path is /mnt/pdfs/hadoop. <pre>BDPP_PDFS_MOUNTPOINT=/mnt/pdfs BDPP_HADOOP_TOP_DIR=/hadoop</pre>
pdfs.fs.local.homedir Home directory path for users in the DFS FileSystem class	/user
pdfs.security.authorization Whether to use the MapReduce job user authentication of the DFS	true
pdfs.fs.local.buffer.size Default buffer size (bytes) used during Read/Write	524288 (512 KB) (*1)
pdfs.fs.local.block.size Data size (bytes) into which Map tasks are split for MapReduce jobs	268435456 (256 MB)
pdfs.fs.local.posix.umask Whether to reflect the process umask value in access permissions set when creating a file or directory	true (*2)
pdfs.fs.local.cache.location Whether to use the cache local MapReduce feature	true
pdfs.fs.local.cache.minsize Size (bytes) of files excluded from the cache local MapReduce feature	1048576 (1 MB)
pdfs.fs.local.cache.procs Number of multiplex executions when the cache local MapReduce feature obtains the memory cache information	40

*1: Use the greater of the value to be set or the io.file.buffer.size value in "C.2 core-site.xml".

*2: true: Uses the umask value (POSIX-compatible)/false: Does not use the umask value (HDFS-compatible)

Information

The cache local MapReduce feature

When pdfs.fs.local.cache.location is enabled (true), this feature obtains the memory cache retention node information of the target file when a MapReduce job starts. It also prioritizes assignment of the Map task to the node that has the cache, as a result, speeding up Map phase processing.

Additionally, as it is costly to obtain the memory cache retention node information, settings can be configured to not obtain information from files that are smaller than the size specified in pdfs.fs.local.cache.minsize.

C.5 sysctl.conf

This section explains the parameters configured in sysctl.conf.

Parameter name	Value to be set
net.ipv4.ip_local_reserved_ports Ports not automatically assigned	8020,9000,10101,10102,50010,50020,50030,50060,50070,50075,50090,50470,50475,60001,60002 (*1)
net.core.somaxconn Maximum length of the queue storing connection requests received by the socket	1024 (*2)

*1: Specify all ports used with this product.

*2: Specify the same value as ipc.server.listen.queue.size in "C.2 core-site.xml".

C.6 limits.conf

This section explains the parameters configured in limits.conf.

Parameter name	Value to be set
nfile Number of files that can be opened	32768
nproc Number of processes (threads) that can be started	32000



Example

Specify using the following group units:

- hadoop groups
- Groups specified in BDPP_HADOOP_DEFAULT_GROUP in the bdpp.conf configuration file

```
@hadoop - nfile 32768
@bdppgroup - nfile 32768
```

C.7 HDFS Settings (Reference Information)

This section describes configuration changes for other parameters.

The properties listed below cannot be used with this product. This information is provided for reference only.

hadoop-env.sh

Environment variable	Default value/Value to be set
HADOOP_DATANODE_OPTS	Default value -Dcom.sun.management.jmxremote \$HADOOP_DATANODE_OPTS

Environment variable	Default value/Value to be set
Launch option specifically for the JVM that operates DataNode	<p>Value to be set</p> <p>-Xmx\$\$m -Dhadoop.security.logger=ERROR,DRFAS \$HADOOP_DATANODE_OPTS</p> <p>\$\$ = The lesser of the following values:</p> <ul style="list-style-type: none"> - 512 - <i>physicalMemorySizeInMb</i> / 12
<p>HADOOP_NAMENODE_OPTS</p> <p>Launch option specifically for the JVM that operates NameNode</p>	<p>Default value</p> <p>-Dcom.sun.management.jmxremote \$HADOOP_NAMENODE_OPTS</p> <p>Value to be set</p> <p>-Xmx\$\$m -Dhadoop.security.logger=INFO,DRFAS -Dhdfs.audit.logger=INFO,DRFAAUDIT \$HADOOP_NAMENODE_OPTS</p> <p>\$\$ = <i>physicalMemorySizeInMb</i> / 4</p>
<p>HADOOP_SECONDARYNAMENODE_OPTS</p> <p>Launch option specifically for the JVM that operates SecondaryNameNode</p>	<p>Default value</p> <p>-Dcom.sun.management.jmxremote \$HADOOP_SECONDARYNAMENODE_OPTS</p> <p>Value to be set</p> <p>-Xmx\$\$m -Dhadoop.security.logger=INFO,DRFAS -Dhdfs.audit.logger=INFO,DRFAAUDIT \$HADOOP_SECONDARYNAMENODE_OPTS</p> <p>\$\$ = <i>physicalMemorySizeInMb</i> / 4</p>

core-site.xml

Property	Default value/Value to be set
<p>fs.checkpoint.dir</p> <p>Directory storing the HDFS management information for SecondaryNameNode</p>	<p>Default value</p> <p>\${hadoop.tmp.dir}/dfs/secondary</p> <p>Value to be set</p> <p>/var/lib/hadoop/hdfs/secondary</p>
<p>fs.checkpoint.period</p> <p>Interval (seconds) at which the NameNode change history is reflected (redundancy is performed) to SecondaryNameNode</p>	<p>Default value</p> <p>3600 (60 minutes)</p> <p>Value to be set</p> <p>1800 (30 minutes)</p>

hdfs-site.xml

This setup file cannot be used with this product.

Property	Default value/Value to be set
dfs.block.size Size of the local file (bytes) when storing the HDFS file blocks in the local file	Default value 67108864 (64 MB) Value to be set 134217728 (128 MB)
dfs.datanode.du.reserved Size of the disk (bytes) to which dfs.data.dir is deployed, where dfs.data.dir is reserved for use with data other than HDFS data	Default value 0 Value to be set Set at 40% (bytes) of the available space on the disk where dfs.data.dir is deployed
dfs.datanode.handler.count Number of threads performing processes within DataNode	Default value 3 Value to be set The greater of the following values: - 3 - Number of CPU cores
dfs.datanode.ipc.address Address for sending and receiving data between DataNodes	Default value 0.0.0.0:50020 Value to be set 0.0.0.0:50020 (reset)
dfs.datanode.max.xcievers Maximum number of connections to DataNode	Default value 256 Value to be set 4096
dfs.permissions.supergroup HDFS supergroup name	Default value supergroup Value to be set hadoop
dfs.support.append Whether the HDFS append feature is used	Default value false Value to be set true

Appendix D Port List

This appendix describes the ports used by this product.

Note

Port numbers listed in the Port column are fixed and cannot be changed.

Ports used with Hadoop (MapReduce)

The following ports are used for MapReduce processes.

LAN type	Send source			Send destination			Protocol
	Server	Service	Port	Server	Service	Port	
Public LAN	Slave server	Hadoop (TaskTracker)	Not fixed	Master server	Hadoop (JobTracker)	9000	tcp
	Development server	Hadoop (Client)	Not fixed	Master server			
	Development server	Hadoop (Client)	Not fixed	Master server	Hadoop (JobTracker)	50030	tcp
	Slave server	Hadoop (TaskTracker)	Not fixed	Slave server	Hadoop (TaskTracker)	50060	tcp

Ports used with Fujitsu distributed file system (DFS)

The following ports are used for iSCSI connections to the ETERNUS DX disk array.

LAN type	Send source			Send destination			Protocol
	Server	Service	Port	Server	Service	Port	
iSCSI LAN	Master server	-	Not fixed	ETERNUS DX disk array	iscsi-target	3260	tcp
	Slave server	-	Not fixed				
	Development server	-	Not fixed				
	Collaboration server	-	Not fixed				

Ports used for master server replication

The following ports are used for the master server replication feature.

LAN type	Send source			Send destination			Protocol
	Server	Service	Port	Server	Service	Port	
Admin LAN	Master server	-	Not fixed	Master server (*1)	-	2316	udp
		-	Not fixed		-	623	udp

LAN type	Send source			Send destination			Protocol
	Server	Service	Port	Server	Service	Port	
CIP LAN	Master server	-	Not fixed	Master server (*2)	-	9331	tcp
	Master server	-	Not fixed	Master server (*1) (*2)	-	9379	tcp
	Master server	-	Not fixed	Master server (*2)	-	9378	tcp
	Master server	-	Not fixed	Master server (*1) (*2)	-	9377	tcp
	Master server	-	Not fixed	Master server (*2)	-	9376	tcp
		-	Not fixed		-	9375	tcp
	Master server	-	Not fixed	Master server (*1)	-	9383	tcp
		-	Not fixed		-	9384	tcp
		-	9786		-	9786	tcp
		-	8000		-	8000	udp
		-	11111		-	11111	udp

*1: Used for communication between the primary master server and the secondary master server if replicated configuration is used for master servers.

*2: Used for communication between processes within master servers. Can also be used if replicated configuration is not used for master servers.

Ports used for Smart Setup

The following ports are used when adding (cloning) slave servers to a physical environment.

They are not used for adding slave servers to virtual environments.

LAN type	Send source			Send destination			Protocol
	Server	Service	Port	Server	Service	Port	
Admin LAN	Primary master server	-	4972	Slave server	-	4973	udp
	Slave server	-	4973	Primary master server	-	4972	udp
		bootpc	68		bootpc	67	udp
		-	Not fixed		pxc	4011	udp
		-	Not fixed		tftp	69	udp
	Primary master server	-	Not fixed	Primary master server	-	4971	tcp
	Slave server	-	4974~4989	Primary master server	-	4974~4989	udp

LAN type	Send source			Send destination			Protocol
	Server	Service	Port	Server	Service	Port	
		-	Not fixed		-	4974~4989	tcp udp

 **Point**

The method for adding slave servers differs depending on the server configuration. Refer to "[6.3 Adding Second and Subsequent Slave Servers](#)" for details.

Ports used for Hadoop (HBase, ZooKeeper) (reference information)

The following ports are used for HBase and ZooKeeper processes.

LAN type	Send source			Send destination			Protocol
	Server	Service	Port	Server	Service	Port	
Public LAN	Slave server	HBase (RegionServer)	Not fixed	Master server	HBase (Hmaster) (BackupHmaster)	60000	tcp
	Development server	HBase (Client)	Not fixed				
	Development server	HBase (Client)	Not fixed	Master server	HBase (Hmaster) (BackupHmaster)	60010	tcp
	Master server	HBase (Hmaster) (BackupHmaster)	Not fixed	Slave server	HBase (RegionServer)	60020	tcp
	Slave server	HBase (RegionServer)	Not fixed				
	Development server	HBase (Client)	Not fixed				
	Development server	HBase (Client)	Not fixed	Slave server	HBase (RegionServer)	60030	tcp
	Master server	HBase (Hmaster) (BackupHmaster)	Not fixed	Slave server	ZooKeeper	2081	tcp
	Slave server	HBase (RegionServer)	Not fixed				
	Development server	HBase (Client)	Not fixed				
	Slave server	ZooKeeper	Not fixed	Slave server	ZooKeeper	2888	tcp
	Slave server	ZooKeeper	Not fixed	Slave server	ZooKeeper	3888	tcp

Ports used for HDFS (reference information)

The following ports are used with the HDFS file system.

They cannot be used when using DFS file systems (when YES is specified in the BDPP_PDFS_FILESYSTEM parameter of bdpp.conf).

LAN type	Send source			Send destination			Protocol
	Server	Service	Port	Server	Service	Port	
Public LAN	Slave server	Hadoop (DataNode)	Not fixed	Master server	Hadoop (NameNode)	8020	tcp
	Development server	Hadoop (Client)	Not fixed				
	Development server	Hadoop (Client)	Not fixed	Slave server	Hadoop (DataNode)	50010	tcp
	Slave server	Hadoop (DataNode)	Not fixed	Slave server	Hadoop (DataNode)	50020	tcp
	Development server	Hadoop (Client)	Not fixed				
	Master server	Hadoop (Secondary NameNode)	Not fixed	Master server	Hadoop (NameNode)	50070	tcp
	Development server	Hadoop (Client)	Not fixed				
	Slave server	Hadoop (DataNode)	Not fixed	Slave server	Hadoop (DataNode)	50075	tcp
	Development server	Hadoop (Client)	Not fixed				
	Master server	Hadoop (NameNode)	Not fixed	Master server	Hadoop (Secondary NameNode)	50090	tcp
	Master server	Hadoop (Secondary NameNode)	Not fixed	Master server	Hadoop (NameNode)	50470	tcp
	Development server	Hadoop (Client)	Not fixed				
	Slave server	Hadoop (DataNode)	Not fixed	Slave server	Hadoop (DataNode)	50475	tcp
	Development server	Hadoop (Client)	Not fixed				



See

Refer to "B.1 bdpp.conf" for information on this configuration file.

Appendix E Messages

This appendix explains the messages output or displayed by this product.

E.1 Messages During Installation

This section explains the messages output or displayed during the installation of this product.

bdpp_inst: ERROR: 001: Configuration file bdpp.conf does not exist.

Description

The bdpp.conf configuration file does not exist.

Action method

Create bdpp.conf, and specify the bdpp.conf path in the BDPP_CONF_PATH environment variable.

bdpp_inst: ERROR: 002: bdpp.conf parameter error. (\$INFO1: \$INFO2)

Description

There is an error in the parameter shown in \$INFO1 (line number or description code):\$INFO2 (parameter and specified value) in bdpp.conf.

The description codes displayed in \$INFO1 include the following:

- NOT
Displayed when the mandatory parameter in \$INFO2 has not been specified.
- IMPROPER
Displayed when a combination of the parameter in \$INFO2 and other parameters is inappropriate.
For example, IMPROPER will be displayed if the BDPP_SECONDARY_ADM_IP parameter was specified, even though the BDPP_SECONDARY_NAME parameter is not specified.

Action method

Refer to "[B.1 bdpp.conf](#)" for the parameters that can be defined in bdpp.conf, and ensure that their specified values are correct. Then, re-execute the bdpp_inst command.

bdpp_inst: ERROR: 003: The LAN Manager module is not exist at /tmp.

Description

The "Microsoft LAN Manager module" cannot be found on the system work directory (/tmp).

Action method

Deploy the "Microsoft LAN Manager module", as described in "[5.16 Expanding the Microsoft LAN Manager Module](#)", to the system work directory (/tmp), and then re-execute the installer.

Note that the "Microsoft LAN Manager module" should be placed directly under /tmp without creating a subdirectory.

bdpp_inst: ERROR: 004: This server's IP does not match primary IP (\$BDPP_PRIMARY_IP) nor secondary IP (\$BDPP_SECONDARY_IP).

Description

The IP address of the server attempting to install the master server functionality does not match the value of the parameters (BDPP_PRIMARY_IP or BDPP_SECONDARY_IP) specified in bdpp.conf.

Action method

Check the following, and take appropriate actions:

- The BDPP_PRIMARY_IP or BDPP_SECONDARY_IP parameters specified in bdpp.conf may not match the IP address allocated for connection to the public LAN.
Ensure that the IP address allocated for connection to the public LAN has been specified, then try installing again.
- Ensure that the network interface of the Public LAN is active, activate it if it is not, then try installing again.

Example

```
# /etc/init.d/network restart <Enter>
```

bdpp_inst: ERROR: 005: The value is not specified for BDPP_PDFS_MOUNTPOINT.

Description

An attempt is being made to use the DFS file system with BDPP_PDFS_FILESYSTEM specified in bdpp.conf, but a mount point has not been specified for BDPP_PDFS_MOUNTPOINT.

Action method

Specify the PDFS mount point at BDPP_PDFS_MOUNTPOINT.

bdpp_inst: WARN: 006: Since file systems other than ext3 are contained, cloning cannot be performed.

Description

Cloning cannot be performed, because a file system other than ext3 is included in the file system of the system attempting to install the slave server functionality.

Action method

Enter "Y" or "N" in the message that follows ("Is installation continued? [y/n]") to continue or cancel the installation.

Note that although the slave server functionality can be installed if "Y" is selected, it will no longer be possible to scale-out the slave server by cloning. To use cloning, use ext3 to rebuild the file system of the system on which the functionality will be installed and perform the installation again.

bdpp_inst: ERROR: 010: Installation failed. (\$PROCESS)

Description

The installation failed at the installer processing shown in \$PROCESS.

Action method

- When \$PROCESS is "groupadd" or "useradd"

The groupadd or useradd command processing failed. Check if the value specified for BDPP_HADOOP_DEFAULT_GROUP or BDPP_HADOOP_DEFAULT_USERS in bdpp.conf is already registered in the system. If it is, delete it and then perform the installation again.

- When \$PROCESS is "PDFS"

The directories /opt/FJSVpdfs, /etc/opt/FJSVpdfs, /var/opt/FJSVpdfs, or /etc/pdfs may still exist. Perform backup if necessary, then delete the relevant directory and rerun the installer.

- When \$PROCESS is "ServerView Resource Orchestrator"

The IP address of the admin LAN may contain an incorrect value. Check the IP address settings, and revise if necessary. Then, perform the installation again.

- When \$PROCESS is other than the above

Refer to the install log /var/tmp/bdpp_install.log, and remove the cause of the error. Then, perform the installation again.

bdpp_inst: ERROR: 011: Uninstallation failed. (\$PROCESS)

Description

The uninstallation failed at the uninstaller processing shown in \$PROCESS.

Action method

Refer to the install log /var/tmp/bdpp_install.log, and remove the cause of the error. Then, perform the uninstallation again.

bdpp_inst: ERROR: 101: Administrator privilege is required.

Description

The user attempting execution does not have root permissions.

Action method

Re-execute using root permissions.

bdpp_inst: ERROR: 102: \$PKG package is not installed.

Description

The mandatory package shown in \$PKG is not installed.

Action method

Install the mandatory package shown in \$PKG.



Note

Contact Fujitsu technical staff if an error message other than the ones shown above is output or displayed.

E.2 Messages During Setup

This section explains the messages output or displayed during the setup of this product.

bdpp_setup: ERROR: 001: Setup of the cluster failed.

Description

The cluster setup failed.

Action method

- When the setup fails on using the cluster_setup2 command

The NIC settings of the cluster interconnect are not "ONBOOT=YES".

Check that the setting in "/etc/sysconfig/network-scripts", "ifcfg-ethX" (*1) is "ONBOOT=YES", changing "NO" to "YES" if necessary. Then, restart the network interface and re-execute the cluster_setup2 command.

*1: ethX is the network interface used for the cluster interconnect. Specify a numeral at X.

- In all other cases

Check the message that was output prior to this message, and take action. Then, perform the cluster setup removal again.

bdpp_setup: ERROR: 002: Removal of the cluster failed.

Description

Removal of the cluster setup failed.

Action method

Check the message that was output prior to this message, and take action. Then, perform the cluster setup removal again.

bdpp_setup: ERROR: 003: Setup failed. (\$PROCESS)

Description

The setup failed at the setup processing shown in \$PROCESS.

Action method

- When \$PROCESS is "PRIMECLUSTER GLS"

The system may not have been restarted after installation. Restart the system, then perform the setup again.

- When \$PROCESS is other than the above

Refer to the setup log /var/opt/FJSVbdpp/log/bdpp_setup.log, and remove the cause of the error. Then, perform the setup again.

bdpp_setup: ERROR: 004: Removal failed. (\$PROCESS)

Description

The setup removal failed at the setup removal processing shown in \$PROCESS.

Action method

Refer to the setup log /var/opt/FJSVbdpp/log/bdpp_setup.log, and remove the cause of the error. Then, perform the setup removal again.

bdpp_setup: ERROR: 005: The slaves file is not in "/etc/opt/FJSVbdpp/conf".

Description

The slave server definition file "slaves" is not under /etc/opt/FJSVbdpp/conf.

Action method

Create the slave server definition file "slaves", and store it under /etc/opt/FJSVbdpp/conf.

bdpp_setup: ERROR: 006: DFS is set up but it is not mounted.

Description

The DFS is not mounted.

Action method

Mount the DFS.

bdpp_setup: ERROR: 007: The mount point specified for fstab does not correspond with the value of BDPP_PDFS_MOUNTPOINT.

Description

The mount point specified at fstab does not match the BDPP_PDFS_MOUNTPOINT value.

Action method

Check if the mount point specified at fstab is specified for BDPP_PDFS_MOUNTPOINT in bdpp.conf. Change it to the correct value.

bdpp_setup: ERROR: 008: It is not connectable with Cluster interconnect.

Description

Unable to establish a connection between the primary and secondary master server HA clusters.

Action method

Check the following, and take appropriate actions:

- Check the connection status of the LAN used for the cluster interconnect, and take action to ensure that it can communicate correctly. Next, execute the cluster_unsetup command and then re-execute the cluster_setup and cluster_setup2 commands.
- Check "6.1.2 HA Cluster Setup", and configure and activate the cluster interconnect. Next, execute the cluster_unsetup command and then re-execute the cluster_setup and cluster_setup2 commands.

bdpp_setup: ERROR: 009: It is not connectable with an ETERNUS DX disk array.**Description**

Unable to connect to an ETERNUS DX disk array.

Action method

Check the following, and take appropriate actions:

- Check if there is an error in the IP address specified in the BDPP_DISKARRAY_iSCSI1_IP or BDPP_DISKARRAY_iSCSI2_IP parameters of bdpp.conf in the /etc/opt/FJSVbdpp/conf directory. If an error is found, revise using the correct IP address. Next, execute the bdpp_unsetup command and then re-execute the bdpp_setup command.
- Check the connection status of the LAN used to connect to the ETERNUS DX disk array, and take action to ensure that it can communicate correctly. Then, re-execute the bdpp_setup command.
- Connect and log in to the target ETERNUS DX disk array. Next, execute the bdpp_unsetup command and then re-execute the bdpp_setup command.

bdpp_setup: WARN : 010: Setup or Removal has already been executed.**Description**

The setup or removal has already been executed.

bdpp_setup: ERROR: 101: Administrator privilege is required.**Description**

The user attempting execution does not have root permissions.

Action method

Re-execute using root permissions.



Contact Fujitsu technical staff if an error message other than the ones shown above is output or displayed.

E.3 Messages Output During Operation

This section explains the messages output or displayed during operation of this product.

E.3.1 Messages Output During Command Execution

This section explains the messages output or displayed during command execution.

bdpp: WARN : 001: bdpp Hadoop is already started.**Description**

Hadoop is already started.

bdpp: WARN : 002: bdpp Hadoop is already stopped.**Description**

Hadoop is already stopped.

bdpp: INFO : 003: bdpp Hadoop JobTracker is alive.**Description**

The Hadoop process (JobTracker) is starting.

bdpp: INFO : 004: bdpp Hadoop JobTracker is not alive.**Description**

The Hadoop process (JobTracker) is not starting.

bdpp: ERROR: 005: bdpp Hadoop start failed (\$PROCESS).**Description**

The Hadoop process shown in \$PROCESS failed to start.

Action method

- The required communications (port) may have been shut by a firewall. Refer to "[Appendix D Port List](#)", permit connections for ports used by Interstage Big Data Processing Server, then re-execute the command.
- The SSH daemon is not running or SSH communications may have failed. Review the SSH settings, then re-execute the command.
- The settings may not permit SSH communications on master servers without a password. Create a public key for both the primary and secondary master servers, and configure settings to enable the servers to communicate without a password both separately and mutually. Then, re-execute the command.

If none of the above applies, refer to the log (/var/opt/FJSVbdpp/log/bdpp_hadoop.log), take the actions corresponding to the Apache Hadoop error content and Apache Hadoop-related logs, then re-execute the command.

bdpp: ERROR: 006: bdpp Hadoop stop failed (\$PROCESS).**Description**

The Hadoop process shown in \$PROCESS failed to stop.

Action method

Refer to the log (/var/opt/FJSVbdpp/log/bdpp_hadoop.log), check the Apache Hadoop error content and Apache Hadoop related logs, take the appropriate action, then re-execute the command.

bdpp: ERROR: 010: clone.conf does not exist.**Description**

The cloning server definition file clone.conf does not exist.

Action method

Specify a directory where the cloning server definition file clone.conf does exist. If the cloning server definition file clone.conf does not yet exist, create one.

bdpp: ERROR: 011: Failed to get the information from slave server.**Description**

The slave server for creating clone images could not be registered.

Action method

Check that there are no errors in the content of the cloning server definition file clone.conf.

bdpp: ERROR: 012: Failed to get the list of slave servers.**Description**

The server information could not be obtained.

Action method

Check the bdpp_addserver command execution result, then re-execute the command.

bdpp: ERROR: 013: server name and image name are needed to get the cloning image.**Description**

Required parameters are not specified.

Action method

Check the usage method of the command.

bdpp: ERROR: 014: Failed to remove the slave server.**Description**

The slave server for creating clone images could not be removed.

Action method

Refer to the log (/var/opt/FJSVbdpp/log/bdpp_ROR.log) and the syslog, take the action advised after the message prefix "FJSVrcx:", then re-execute the command.

bdpp: ERROR: 015: Failed to create the slave server image.**Description**

The clone image could not be created.

Action method

- The required communications (port) may have been shut by a firewall. Refer to "[Appendix D Port List](#)", permit connections for ports used by Interstage Big Data Processing Server, then re-execute the command.
- In the system BIOS settings of the slave server, it is possible that booting from the network interface used by the admin LAN has not been set as first in the boot order. Check the system BIOS settings of the slave server, set booting from the network interface used by the admin LAN as first in the boot order, then re-execute the command.

If none of the above applies, refer to the log (/var/opt/FJSVbdpp/log/bdpp_ROR.log) and the syslog, take the actions for messages with the prefix "FJSVrcx:", then re-execute the command.

 **Note**

.....
If there is an error in clone.conf definitions (such as MAC address) or clone image creation fails due to a timeout, the network interface of the slave server that performed the clone image creation may have changed to "DHCP" due to the cloning process, and not changed back.

If this is the case, log in to the slave server and manually restore the settings of the network interface. Next, reset by performing "[7.2.1 Removing the Network Replication Setup](#)" and "[6.2.3 Network Replication and Hadoop Setup](#)" again.
.....

bdpp: ERROR: 016: Failed to get the list of server image.**Description**

The clone image information could not be obtained.

Action method

Check the `bdpp_getimage` command execution result, then re-execute the command.

bdpp: ERROR: 017: Failed to remove the cloning image.

Description

The clone image could not be removed.

Action method

Refer to the log (`/var/opt/FJSVbdpp/log/bdpp_ROR.log`) and the syslog, take the action advised after the message prefix "FJSVrcx:", then re-execute the command.

bdpp: ERROR: 018: Failed to deploy the slave server image.

Description

The clone image could not be deployed.

Action method

- The required communications (port) may have been shut by a firewall. Refer to "[Appendix D Port List](#)", permit connections for ports used by Interstage Big Data Processing Server, then re-execute the command.
- In the system BIOS settings of the slave server, it is possible that booting from the network interface used by the admin LAN has not been set as first in the boot order. Check the system BIOS settings of the slave server, set booting from the network interface used by the admin LAN as first in the boot order, then re-execute the command.

- If the message "FJSVnrmpr:ERROR:64780:invalid file format(file=ipaddr.conf, detail=line X)" is output to the log (`/var/opt/FJSVbdpp/log/bdpp_ROR.log`) (X is the line number in `ipaddr.conf`), the definition for the slave server whose clone image is to be created is not in `ipaddr.conf`.

Include the definition for the slave server whose clone image is to be created in `ipaddr.conf`, again perform "[6.3.1.3 Registering the Network Parameter and iSCSI Name Automatic Configuration Feature](#)", then recreate the clone image and distribute it.

If none of the above applies, refer to the log (`/var/opt/FJSVbdpp/log/bdpp_ROR.log`) and the syslog, take the actions for messages with the prefix "FJSVrcx:", then re-execute the command.

bdpp: ERROR: 019: Failed to register FJSVrcx.conf or ipaddr.conf in the slave server.

Description

Failed to register `FJSVrcx.conf` or `ipaddr.conf` to the slave server.

Action method

- The required communications (port) may have been shut by a firewall. Refer to "[Appendix D Port List](#)", permit connections for ports used by Interstage Big Data Processing Server, then re-execute the command.
- The `FJSVrcx.conf` or `ipaddr.conf` file may not exist, or the specified path may be incorrect.
Create the file if `FJSVrcx.conf` or `ipaddr.conf` does not exist, or correct the path, then re-execute the command.

If none of the above applies, refer to the log (`/var/opt/FJSVbdpp/log/bdpp_ROR.log`) and the syslog, take the actions for messages with the prefix "FJSVrcx:", then re-execute the command.

bdpp: ERROR: 020: Failed to change the cloning image directory.

Description

The clone image storage directory could not be changed.

Action method

Refer to the log (`/var/opt/FJSVbdpp/log/bdpp_ROR.log`) and the syslog, take the action advised after the message prefix "FJSVrcx:", then re-execute the command.

bdpp: INFO : 021: Processing was completed normally.

Description

The command processing completed normally.

Supplementary note:

When cloning, this is also output to the system log on slave servers where the clone image is distributed (or slave servers at the clone destination in a virtual environment).

bdpp: ERROR: 022: Failed to set the initiator name in the slave server.

Description

Failed to set the iSCSI initiator name.

Action method

Check if the settings can be found for the server name specified in the setup files clone.conf, ipaddr.conf, and initiator.conf when cloning.

If the settings cannot be found, review the setup files and then re-execute the bdpp_prepareserver command for the slave server whose clone image will be created (or the slave server at the clone source in a virtual environment).

Supplementary note:

This message is output to the system log on slave servers where the clone image is distributed (or slave servers at the clone destination in a virtual environment).

bdpp: ERROR: 031 : Parameters or options are invalid.

Description

The command cannot be executed, because there is an error in the argument or the option specified in the command.

Action method

Check the reference for each command, and then re-execute the command after confirming how it is used.

bdpp: ERROR: 032: Command cannot be executed while (\$PROCESS) is running.

Description

The command cannot be executed, because Hadoop or an associated command is running.

Action method

- If \$PROCESS is "Hadoop"

Use the bdpp_stop command to stop Hadoop, and then re-execute the command.

- If \$PROCESS is any of the following commands:

- bdpp_addserver
- bdpp_backup
- bdpp_changeimagedir
- bdpp_changeslaves
- bdpp_deployimage
- bdpp_getimage
- bdpp_lanctl
- bdpp_prepareserver
- bdpp_removeimage
- bdpp_removeserver

- bdpp_restore
- bdpp_start

Refer to the output information. When the associated commands have finished executing, re-execute the command.

bdpp: ERROR: 033: Command abnormally finished.

Description

Processing failed, because a problem occurred during command execution. A message about the cause of the error will be output to the log (to the file under /var/opt/FJSVbdpp/log).

Action method

- Typical messages and actions:
 - If the message "No enough space." is output, the command has failed because of insufficient disk space at the specified backup destination or on the server to be restored.
Refer to "[Chapter 14 Backup and Restore](#)", secure the required space, and then re-execute the command.
 - If the message "\$BACKUP_PATH: Already exists." (where \$BACKUP_PATH is the backup destination directory specified in the command) is output to the log, the backup has failed because the backup storage directory in the specified directory (FJSVbdpp-backup) or a file with the same name already exists. Delete the directory or specify a different directory as required, and re-execute the command.
 - If the message "Backup files are not found." is output to the log, the restore could not be executed because the backup files were not found in the specified directory. Specify the correct directory that stores the backup files, and then re-execute the command.
- Other messages:
Refer to the log, remove the cause of the error, and then re-execute the command.

bdpp: ERROR: 101: Administrator privilege is required.

Description

The user attempting execution does not have root permissions.

Action method

Re-execute using root permissions.

E.3.2 Other Messages

Refer to the manuals of the relevant features for explanations of the messages output to the system log (/var/log/messages) or to the log of this product with the prefixes below.

Prefix	Feature	Reference manual
FJSVcluster	Master server switching	"PRIMECLUSTER Cluster Foundation(CF) Configuration and Administration Guide"
CF		"PRIMECLUSTER Cluster Foundation(CF) Configuration and Administration Guide"
SMAWsf		"PRIMECLUSTER Cluster Foundation(CF) Configuration and Administration Guide"
RMS		"PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide"
"SDX: " or "sfdisk: "		"PRIMECLUSTER Global Disk Services Configuration and Administration Guide" - "Appendix E GDS Messages"

Prefix	Feature	Reference manual
"sfdfs" or "sfdfs"		"PRIMECLUSTER Global File Services 4.3 Configuration and Administration" - "Appendix A List of Messages"
hanet		"PRIMECLUSTER Global File Services Configuration and Administration Guide 4.3 Redundant Line Control Function" - "Appendix A Messages and corrective actions"
FJSVrcx:	Smart Setup	"ServerView Resource Orchestrator V3.1.0 Messages"

Appendix F Mandatory Packages

This appendix covers the mandatory packages for the system software operated by this product.

The mandatory packages below are required to use this product. Install the required packages in advance.

The architecture of the mandatory packages to be installed is shown in parentheses "()".

If the architecture to be installed is not specified, install the package having the same architecture as the operating system.

Red Hat(R) Enterprise Linux(R) 6 (for Intel64)
Red Hat(R) Enterprise Linux(R) 6.1 (for Intel64)
Red Hat(R) Enterprise Linux(R) 6.2 (for Intel64)
Red Hat(R) Enterprise Linux(R) 6.3 (for Intel64)
Red Hat(R) Enterprise Linux(R) 6.4 (for Intel64)

Mandatory package	Installation environment				
	Master server		Slave server	Development server	Collaboration server
	Primary	Secondary			
alsa-lib(i686)	Y	N	Y	N	N
apr(i686)	Y	N	N	N	N
apr-util(i686)	Y	N	N	N	N
compat-expat1(i686)	Y	N	N	N	N
compat-libstdc++-33(i686)	Y	Y	N	N	N
compat-libtermcap(i686)	Y	N	N	N	N
compat-openldap(i686)	Y	N	N	N	N
compat-readline5(i686)	Y	N	N	N	N
crash	Y	Y	Y	Y	Y
cyrus-sasl-lib(i686)	Y	N	N	N	N
db4(i686)	Y	N	N	N	N
expat(i686)	Y	N	N	N	N
gcc	Y	Y	N	N	N
glibc(i686)	Y	Y	Y	Y	Y
kernel-devel	Y	Y	Y	Y	Y
kexec-tools	Y	Y	Y	Y	Y
keyutils-libs(i686)	Y	Y	Y	Y	Y
krb5-libs(i686)	Y	Y	Y	Y	Y
libcom_err(i686)	Y	Y	Y	Y	Y
libgcc(i686)	Y	Y	Y	Y	Y
libICE(i686)	Y	N	Y	N	N
libselinux(i686)	Y	Y	Y	Y	Y
libSM(i686)	Y	N	Y	N	N
libstdc++(i686)	Y	Y	Y	Y	Y
libtool-ltdl(i686)	Y	N	Y	N	N
libuuid(i686)	Y	N	Y	N	N
libX11(i686)	Y	N	Y	N	N
libXau(i686)	Y	N	Y	N	N

Mandatory package	Installation environment				
	Master server		Slave server	Development server	Collaboration server
	Primary	Secondary			
libxcb(i686)	Y	N	Y	N	N
libXext(i686)	Y	N	Y	N	N
libXi(i686)	Y	N	Y	N	N
libXt(i686)	Y	N	Y	N	N
libXtst(i686)	Y	N	Y	N	N
ncurses-libs(i686)	Y	N	Y	N	N
net-snmp	Y	N	N	N	N
net-snmp-utils	Y	N	Y	N	N
nss-softokn-freebl(i686)	Y	Y	Y	Y	Y
ntp	Y	Y	Y	Y	Y
openssl(i686)	Y	Y	Y	Y	Y
openssl098e(i686)	Y	N	N	N	N
readline(i686)	N	N	Y	N	N
redhat-lsb	Y	N	N	N	N
ruby	Y	N	N	N	N
sqlite(i686)	Y	N	Y	N	N
sysfsutils	N	N	Y	N	N
sysstat	Y	Y	Y	Y	Y
system-config-kdump	Y	Y	Y	Y	Y
unixODBC(i686)	Y	N	Y	N	N
zlib(i686)	Y	Y	Y	Y	Y

Y: Installation is required.

N: Installation is not required.

Red Hat(R) Enterprise Linux(R) 5.6 (for Intel64)

Red Hat(R) Enterprise Linux(R) 5.7 (for Intel64)

Red Hat(R) Enterprise Linux(R) 5.8 (for Intel64)

Red Hat(R) Enterprise Linux(R) 5.9 (for Intel64)

Mandatory package	Installation environment				
	Master server		Slave server	Development server	Collaboration server
	Primary	Secondary			
alsa-lib(i686)	N	N	Y	N	N
apr(i686)	Y	N	N	N	N
apr-util(i686)	Y	N	N	N	N
compat-libstdc++-33(i686)	Y	Y	N	N	N
crash	Y	Y	Y	Y	Y
gcc	Y	Y	N	N	N

Mandatory package	Installation environment				
	Master server		Slave server	Development server	Collaboration server
	Primary	Secondary			
glibc(i686)	N	N	Y	N	N
e2fsprogs(i686)	Y	Y	Y	Y	Y
kernel-devel	Y	Y	Y	Y	Y
kexec-tools	Y	Y	Y	Y	Y
krb5-libs(i686)	Y	Y	Y	Y	Y
libgcc(i686)	Y	Y	Y	Y	Y
libICE(i686)	N	N	Y	N	N
libselinux(i686)	N	N	Y	N	N
libsepol	N	N	Y	N	N
libSM(i686)	N	N	Y	N	N
libstdc++(i686)	Y	Y	Y	Y	Y
libX11(i686)	N	N	Y	N	N
libXau(i686)	N	N	Y	N	N
libXdmcp	N	N	Y	N	N
libXext(i686)	N	N	Y	N	N
libXi(i686)	N	N	Y	N	N
libxml2(i686)	Y	N	N	N	N
libxslt(i686)	Y	N	N	N	N
libXt(i686)	N	N	Y	N	N
libXtst(i686)	N	N	Y	N	N
ncurses-libs(i686)	N	N	Y	N	N
net-snmp	Y	N	N	N	N
net-snmp-utils	Y	N	Y	N	N
ntp	Y	Y	Y	Y	Y
openssl(i686)	Y	Y	Y	Y	Y
postgresql-libs(i686)	Y	N	N	N	N
readline(i686)	N	N	Y	N	N
redhat-lsb	Y	N	N	N	N
ruby	Y	Y	N	N	N
sqlite(i686)	N	N	Y	N	N
sysstat	Y	Y	Y	Y	Y
system-config-kdump	Y	Y	Y	Y	Y
zlib(i686)	Y	Y	Y	Y	Y

Y: Installation is required.

N: Installation is not required.

Glossary

admin LAN

The network used mainly to perform cloning processes in Smart Setup.

Apache Hadoop

Open source Hadoop software developed by Apache Software Foundation (ASF).

clone master

A master obtained by removing server-specific information (system node names and IP addresses) from the contents of a system disk. The clone master is copied to a virtual server system disk when a virtual server is deployed.

Cluster Interconnect Protocol (CIP) LAN

The network that performs heartbeat monitoring between primary and secondary master servers in an HA cluster configuration.

collaboration server

An existing server that accesses data files by using a standard Linux file access interface.

DataNode

Name of the nodes that comprise a cluster in the HDFS file system used by Hadoop. Big data processed by Hadoop is distributed and replicated in block units and mapped on DataNodes.

development server

The server that develops and executes applications (MapReduce) that perform parallel distribution.

Hadoop

The name of technology that performs efficient distribution and parallel processing for big data. Its broad components are the HDFS distributed file system and the MapReduce parallel distributed processing technology.

HDFS (Hadoop Distributed File System)

The distributed file system used in Hadoop. HDFS maps big data that is distributed and replicated in block units on multiple nodes, called DataNodes, and this mapping is managed by a node called a NameNode.

image information

Information expressing the structure of a virtual image.

Image information is required when constructing a virtual image. Separate image information is required for creating each virtual image.

iSCSI-LAN

The network that uses iSCSI connections between servers and storage systems.

MapReduce

The parallel distributed processing technology at the core of Hadoop. MapReduce performs parallel processing of distributed data and compiles/consolidates the processing results. Its broad components are the TaskTracker that is in charge of the processing at each cluster, and the JobTracker that manages the overall processing and allocates processing to the TaskTracker.

master server

The server that generates blocks from data files and centrally manages them. It receives requests to execute jobs for analysis processing, and performs parallel distributed processing on slave servers.

NameNode

Name of a node that manages the HDFS file system used in Hadoop.

petabyte (PB)

A unit of data size, indicating 10^{15} bytes.

public LAN

The network for parallel distributed processing tasks between master servers and slave servers.

sensing data

Refers to data sent from various types of sensors.

single point of failure

A single part which, if it fails, will be fatal to the entire system, is called a single point of failure. With HDFS, if the NameNode that manages all DataNodes fails, HDFS will be unusable. Therefore, the NameNode is a single point of failure for HDFS.

slave server

The server that accesses blocks of data files. Analysis processing can be performed in a short period of time through parallel distributed processing by using multiple slave servers.

social media

The services and applications that underpin the communication-based society that results from all sorts of people exchanging and sharing a variety of content (such as text, voice, and videos) via the Internet are called social media to differentiate these from older information media (such as newspaper and television).

terabyte (TB)

A unit of data size, indicating 10^{12} bytes.
