



# Interstage Big Data Parallel Processing Server V1.0.0



## User's Guide

Linux(64)

J2UL-1576-01ENZ0(00)  
June 2012

# Preface

---

## Purpose

This manual gives an overview of the functions 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 operation of this product.

## Intended readers

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

## Structure of this document

This document is structured as follows:

Header	Content
Chapter 1 Overview	Gives an overview of this product
Chapter 2 Designing the System and Preparing for Installation	Describes the system design to be considered and the preparatory work that needs to be performed before installing this product.
Chapter 3 System Build Method	Describes how to build a system for this product
Chapter 4 Installing and Uninstalling the Product	Describes how to install and uninstall this product
Chapter 5 Operation	Describes how to operate this product
Chapter 6 Maintenance	Describes the maintenance methods for this product
Appendix A Reference	Describes the commands, definition files, and so on, of this product
Appendix B Messages	Explains the messages output by this product
Appendix C Mandatory Packages	Describes the packages required for the system software that runs this product
Appendix D DFS Environment Construction	Describes how to build the DFS environment
Appendix E DFS Operation Changes	Describes how to change DFS operation
Glossary	Explains the terminology used for this product. Refer to the glossary as required.

## Conventions

The following methods of expression are used in this document:

- Where functions 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 with " ".
- Parts that can be changed are enclosed with { }.
- Screen names, dialog names, menu names, and tab names are enclosed with [ ].
- Button names are enclosed with <>.
- The menu selection sequence is shown in the format [ ] > [ ].
- Character strings and number values requiring special emphasis are enclosed in double quotation marks ("").
- In menu names, "... " indicating the start of settings and operation screens is not shown.
- In usage examples, Prompt is expressed 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 June 2012)

## Related documents

The following manuals are bundled with this product:

- PRIMECLUSTER 4.3A10
- ServerView Resource Orchestrator Virtual Edition V3.0.0
- 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

{DVD drive}:\DISK1\products\PCL\documents\manuals\En

DISK1: ServerView Resource Orchestrator Virtual Edition manual

{DVD drive}:\DISK1\products\ROR\DISK1\Manual\en\VirtualEdition

DISK1: Primesoft Distributed File System for Hadoop manual

{DVD drive}:\DISK1\products\PDFS\documents\manuals\en

In the bundled manuals, only the functions 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) 6 (for Intel64)

Abbreviation	Product
	Red Hat(R) Enterprise Linux(R) 6.1 (for Intel64) Red Hat(R) Enterprise Linux(R) 6.2 (for Intel64)

## Export restriction

This document might contain special technology designated under Foreign Exchange and Foreign Trade Laws. If so, when this document is to be exported or provided overseas, it is necessary to check Foreign Exchange and Foreign Trade Laws, and take the necessary procedure according to these laws.

## 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.
- 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.
- 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
June 2012, First Edition	J2UL-1576-01Z0(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 2012 FUJITSU LIMITED

# Contents

---

Chapter 1 Overview.....	1
1.1 Features of the Interstage Big Data Parallel Processing Server.....	1
1.1.1 Background.....	1
1.1.2 Apache Hadoop.....	1
1.1.3 Interstage Big Data Parallel Processing Server for Enterprise Systems.....	2
1.2 Characteristics and Major Features.....	2
1.2.1 Fujitsu Distributed File System.....	3
1.2.2 Master Server Repliation.....	3
1.2.3 Smart Setup.....	4
Chapter 2 Designing the System and Preparing for Installation.....	5
2.1 System Configuration and Requirements.....	5
2.1.1 Hardware Configuration.....	5
2.1.1.1 Server Configuration.....	5
2.1.1.2 Network Configuration.....	6
2.1.1.3 Hardware Requirements.....	8
2.1.1.3.1 Static Disk Size.....	8
2.1.1.3.2 Dynamic Disk Size.....	9
2.1.1.3.3 Memory Size.....	10
2.1.1.3.4 Expansion Card Options.....	10
2.1.2 Software Configuration.....	11
2.1.2.1 Software Conditions.....	11
2.1.2.2 System Software.....	11
2.1.2.3 Mandatory Software.....	12
2.1.2.4 Exclusive Software.....	13
2.1.2.5 Related Software.....	13
2.2 Designing the System.....	13
2.2.1 Designing the System Configuration.....	13
2.2.2 Designing the bdpp.conf Configuration File.....	16
2.3 Preparing the Base System.....	16
2.3.1 Preparing the Base System.....	16
2.3.2 Performing Advance Preparations for Installing Servers.....	17
2.3.3 Installing Mandatory Software.....	19
2.3.3.1 Obtaining and Setting Up the Mandatory Software.....	20
Chapter 3 System Build Method.....	22
3.1 System Build Flow.....	22
3.1.1 Installing to Master Servers.....	23
3.1.2 Installing the First Slave Server.....	23
3.1.3 Installing a Development Server.....	23
3.1.4 Installing a Collaboration Server.....	23
3.1.5 Distributing the DFS File System Configuration Information.....	23
3.1.6 Checking the Connection to the First Slave Server.....	24
3.1.7 Adding Second and Subsequent Slave Servers.....	25
Chapter 4 Installing and Uninstalling the Product.....	27
4.1 Installing to Master Servers.....	27
4.1.1 Installing Master Server Functions.....	28
4.1.1.1 SSH Settings Between the Master Server (primary) and the Master Server (secondary).....	28
4.1.1.2 Creating the Configuration File bdpp.conf.....	28
4.1.1.3 Installing Master Server Functions.....	28
4.1.2 HA Cluster Setup.....	29
4.1.3 DFS Setup.....	30
4.1.3.1 Creating a Management Partition.....	30
4.1.3.2 Creating the file system.....	31

4.1.3.3	Setting the User ID for Executing MapReduce.....	31
4.1.3.4	Registering the Slave Server, Development Server and Collaboration Server (DFS Client) Information.....	32
4.1.3.5	Creating the Mount Point and Setting fstab Settings.....	33
4.1.3.6	Mounting.....	34
4.1.3.7	Generating and Distributing the DFS File System Configuration Information.....	34
4.1.4	Executing Network Repliation and Hadoop Setup.....	35
4.1.4.1	Editing the Slave Server Definition File 'slaves'.....	35
4.1.4.2	Executing Network Repliation and Hadoop Setup.....	35
4.2	Installing to a Slave Server.....	36
4.2.1	Installing Slave Server Functions.....	36
4.2.1.1	Creating the Configuration File bdpp.conf.....	36
4.2.1.2	Installing Slave Server Functions.....	36
4.2.2	DFS Setup.....	37
4.2.2.1	Creating the Mount Point and Setting fstab Settings.....	37
4.2.2.2	Mounting.....	38
4.2.3	Executing Network Repliation and Hadoop Setup.....	38
4.3	Installing to a Development Server.....	39
4.3.1	Installing Development Server Functions.....	39
4.3.1.1	Creating the Configuration File bdpp.conf.....	39
4.3.1.2	Installing Development Server Functions.....	39
4.3.2	DFS Setup.....	40
4.3.2.1	Mount Point Creation and fstab Settings.....	40
4.3.2.2	Mounting.....	41
4.3.3	Executing Hadoop Setup.....	41
4.4	Installing to a Collaboration Server.....	42
4.4.1	Installing Collaboration Server Functions.....	42
4.4.1.1	Creating the Configuration File bdpp.conf.....	42
4.4.1.2	Installing Collaboration Server Functions.....	42
4.4.2	DFS Setup.....	43
4.4.2.1	Creating the Mount Point and Setting fstab Settings.....	43
4.4.2.2	Registering the hadoop Group and mapred User.....	44
4.4.2.3	Mounting.....	45
4.5	Uninstalling from a Master Server.....	45
4.5.1	Network Repliation Unsetup.....	46
4.5.2	DFS Unsetup.....	47
4.5.3	HA Cluster Unsetup.....	47
4.5.4	Uninstalling Master Server Function.....	47
4.5.5	Post-Uninstallation Tasks.....	48
4.5.5.1	Directories and Files Remaining after an Uninstall.....	48
4.5.5.2	Uninstall (middleware) Uninstallation.....	49
4.5.6	Restarting the System.....	49
4.6	Uninstalling from a Slave Server.....	50
4.6.1	Network Repliation and Hadoop Unsetup.....	50
4.6.2	DFS Unsetup.....	50
4.6.3	Uninstalling Slave Server Functions.....	50
4.6.4	Post-Uninstall Tasks.....	51
4.6.4.1	Directories and Files Remaining after an Uninstall.....	52
4.6.4.2	Uninstall (middleware) Uninstallation.....	52
4.6.5	Restarting the System.....	52
4.7	Uninstalling from a Development Server.....	52
4.7.1	Unsetup DFS.....	52
4.7.2	Uninstalling Development Server Functions.....	52
4.7.3	Post-Uninstall Tasks.....	53
4.7.3.1	Directories and Files Remaining after an Uninstall.....	54
4.7.4	Restarting the System.....	54
4.8	Uninstalling from a Collaboration Server.....	54
4.8.1	DFS Unsetup.....	54

4.8.2 Uninstalling Collaboration Server Functions.....	54
4.8.3 Post-Uninstall Tasks.....	55
4.8.3.1 Directories and Files Remaining after an Uninstall.....	56
4.8.4 Restarting the System.....	56
4.9 Adding and Deleting Slave Servers.....	56
4.9.1 Adding Slave Servers.....	56
4.9.1.1 Adding a Slave Server During Initial Installation.....	56
4.9.1.1.1 Registering a Slave Server.....	57
4.9.1.1.2 Registering the Network Parameter Automatic Configuration Feature.....	58
4.9.1.1.3 Creating a Cloning Image.....	58
4.9.1.1.4 Cloning.....	60
4.9.1.1.5 Adding a Slave Server Linkage Server/Scale-out.....	60
4.9.1.2 Adding Slave Servers after Starting Operation.....	60
4.9.1.2.1 Editing and Reflecting Slave Server Definition File.....	62
4.9.2 Deleting Slave Servers.....	63
4.9.2.1 Editing and Reflecting a Slave Server Definition File.....	64
4.9.2.2 Deleting a DFS client.....	64
<b>Chapter 5 Operation.....</b>	<b>65</b>
5.1 Starting and Stopping the Interstage Big Data Parallel Processing Server.....	65
5.1.1 Starting the Interstage Big Data Parallel Processing Server.....	65
5.1.2 Stopping the Interstage Big Data Parallel Processing Server.....	66
5.1.3 Displaying the Interstage Big Data Parallel Processing Server Status.....	66
5.2 Developing and Registering Applications.....	67
5.2.1 Application Development Environment.....	67
5.2.2 Developing Applications.....	67
5.2.2.1 Application Overview.....	67
5.2.2.2 Designing Applications.....	67
5.2.2.3 Creating Applications.....	68
5.2.2.4 References for Developing MapReduce Applications.....	68
5.2.3 Registering Applications.....	68
5.3 Executing and Stopping Jobs.....	68
5.3.1 Executing Jobs.....	69
5.3.2 Stopping Jobs.....	69
5.3.3 Displaying Job Status.....	69
5.4 Operations When There are Errors.....	69
5.4.1 Master Server Error Operation.....	69
5.4.1.1 Errors while the System is Running.....	70
5.4.1.2 Errors when the System Starts up.....	71
5.4.1.3 If an Error Occurs in the Network.....	72
5.4.2 Slave Server Error Operation.....	73
5.4.3 How to Check Errors.....	73
5.5 Adding and Deleting Storage Systems.....	75
5.5.1 Adding a Storage System.....	75
5.5.2 Deleting a Storage System.....	75
<b>Chapter 6 Maintenance.....</b>	<b>76</b>
6.1 Maintenance Method.....	76
6.1.1 Applying Updates.....	76
6.1.2 Collecting Trouble shooting data.....	76
6.1.2.1 Interstage Big Data Parallel Processing Server logs.....	76
6.1.2.2 Trouble shooting data Collection Tools.....	77
6.1.2.3 Function-specific trouble shooting data.....	77
6.1.2.3.1 If HA cluster setup or switching problems occurred.....	77
6.1.2.3.2 If cloning image creation or cloning problems occur when adding or deleting slave servers.....	81
6.1.2.3.3 If DFS setup or shared disk problems occurred.....	83
6.1.2.3.4 When a problem occurs in Hadoop.....	85

Appendix A Reference.....	87
A.1 Commands.....	87
A.1.1 bdpp_addserver.....	88
A.1.2 bdpp_changeimagedir.....	88
A.1.3 bdpp_changeslaves.....	89
A.1.4 bdpp_deployimage.....	90
A.1.5 bdpp_getimage.....	91
A.1.6 bdpp_lanctl.....	92
A.1.7 bdpp_listimage.....	92
A.1.8 bdpp_listserver.....	93
A.1.9 bdpp_removeimage.....	95
A.1.10 bdpp_removeserver.....	95
A.1.11 bdpp_start.....	96
A.1.12 bdpp_stat.....	97
A.1.13 bdpp_stop.....	98
A.2 Definition Files.....	99
A.2.1 bdpp.conf.....	99
A.2.2 slaves.....	112
A.2.3 clone.conf.....	112
A.2.4 FJSVrcx.conf.....	115
A.2.5 ipaddr.conf.....	115
A.3 Port List.....	118
A.4 Tuning System Parameters.....	122
A.4.1 System Parameter Tuning Values.....	122
A.4.2 Tuning Work Procedures.....	123
Appendix B Messages.....	125
B.1 Messages During Installation.....	125
B.2 Messages During Setup.....	127
B.3 Messages Output During Operation.....	128
B.3.1 Messages Output During Command Execution.....	128
B.3.2 Other Messages.....	132
Appendix C Mandatory Packages.....	133
C.1 List of Mandatory Packages.....	133
Appendix D DFS Environment Construction.....	137
D.1 File System Design.....	137
D.1.1 Server Layout.....	137
D.1.2 Disk Layout.....	137
D.1.3 Preparing for Future Extension.....	138
D.1.4 Selecting the Data Block Size.....	138
D.1.5 Relationship between File System Size, Data Block Size and Maximum File Size.....	138
D.2 Mount Option Design.....	139
D.2.1 fstab Settings for DFS Management Server.....	139
D.2.2 fstab Settings for DFS Clients.....	140
D.3 Management Partition Creation.....	141
D.3.1 Shared Disk Settings.....	142
D.3.2 Management Partition Initialization.....	143
D.3.3 Registering DFS Management Server Information to the Management Partition.....	143
D.3.4 pdfsfrmd Daemon Startup.....	144
D.4 File System Creation.....	144
D.4.1 Shared Disk Settings.....	145
D.4.2 Creating a File System.....	146
D.4.3 Registering Slave Server, Development Server and Collaboration Server Information.....	146
D.4.4 Distributing File System Configuration Information.....	148
D.4.5 fstab Settings.....	149
D.4.6 Mount.....	149

D.4.7 Checking the File System Status.....	150
D.4.8 Backing Up Management Partition Information.....	150
D.5 Hadoop Configuration Settings.....	151
D.5.1 Configuration File and Properties Settings.....	152
D.6 User-Specific Hadoop Usage Settings.....	160
D.6.1 Settings Implemented at root.....	160
D.6.2 Settings Implemented by Hadoop Users.....	161
Appendix E DFS Operation Changes.....	163
E.1 Deleting a File System.....	163
E.2 Re-creating a File System.....	164
E.3 Adding a Slave Server/Scaling Out.....	165
E.3.1 Adding Slave Server, Development Server and Collaboration Server Information.....	166
E.3.2 Distributing File System Configuration Information.....	167
E.3.3 fstab Settings.....	167
E.3.4 Mount.....	168
E.4 Deleting a Slave Server.....	168
E.5 Adding a Shared Disk/Scaling Out.....	169
E.6 Deleting a Shared Disk.....	171
Glossary.....	172

# Chapter 1 Overview

This chapter explains the characteristics of this product and gives an overview of the features added to handle Apache Hadoop open source software.

## 1.1 Features of the Interstage Big Data Parallel Processing Server

### 1.1.1 Background

Not only are enormous amounts of data collected from smart devices, such as smart phones and tablets, and from sensors, but the formats and structures are many and varied, and these are continuously increasing.

This is known as Big Data and it is a major focus, for leading corporations in particular, as use of Big Data progresses and unprecedented business advantages are created.

#### 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

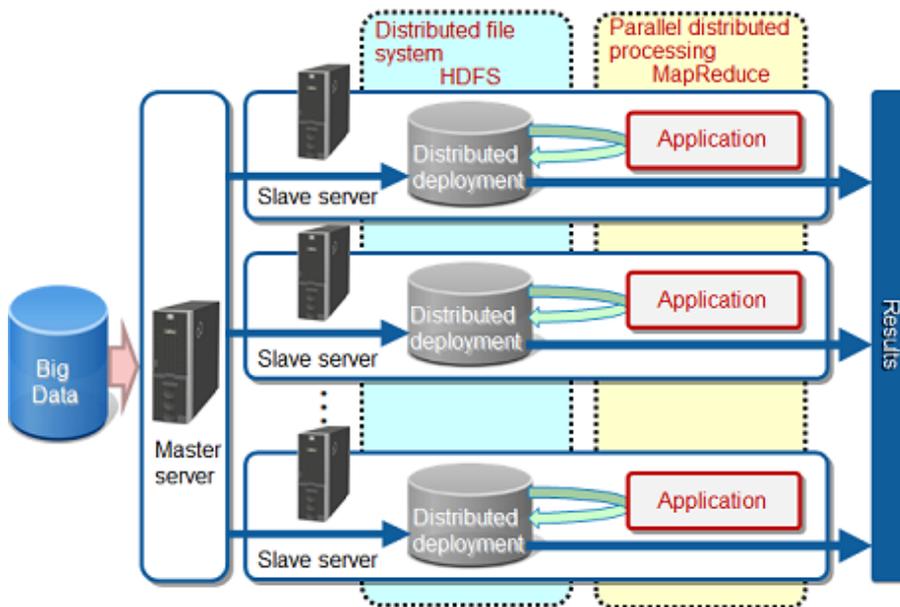
"Apache Hadoop" (\*1) is widely used and is the world standard for applications that can resolve the above Items 1 and 2 in Big Data processing (processing data of massive size, and variety of data).

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

### 1.1.2 Apache Hadoop

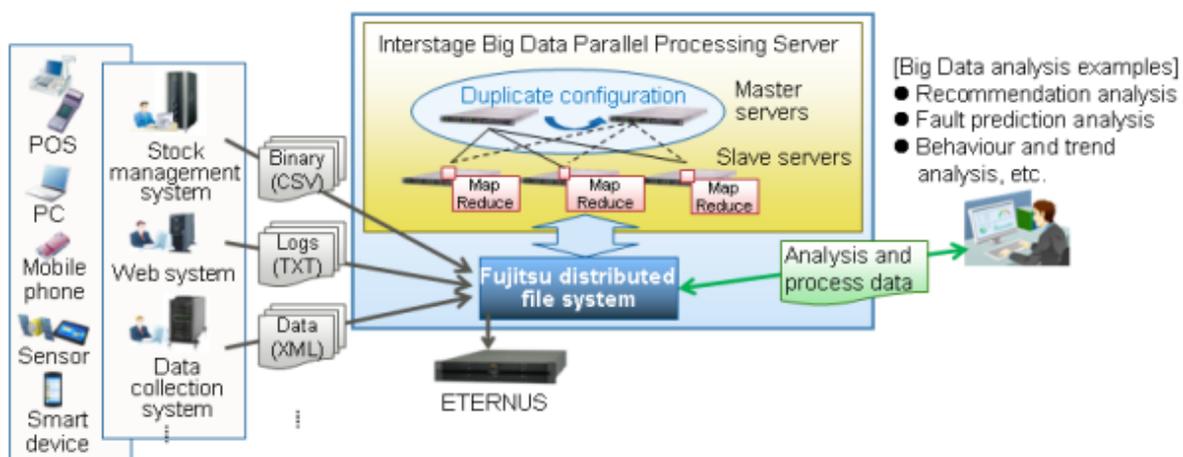
Apache Hadoop technology splits Big Data, distributes it to tens or tens of thousands of servers, and performs parallel processing, thereby performing batch processing of Big Data in a small amount of time. This technology has the following features:

- 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.
- 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.



### 1.1.3 Interstage Big Data Parallel Processing Server for Enterprise Systems

The Interstage Big Data Parallel Processing Server software incorporates Fujitsu proprietary technology with "Apache Hadoop", the world standard in Big Data processing. This assists with the use of Big Data in enterprise systems by greatly improving the reliability and processing performance and by reducing system installation time.



## 1.2 Characteristics and Major Features

This section explains the features added to Apache Hadoop by this product

- Fujitsu distributed file system
- Master server Repliation
- Smart Setup

## 1.2.1 Fujitsu Distributed File System

### Reducing time for data transfer to Hadoop processing server

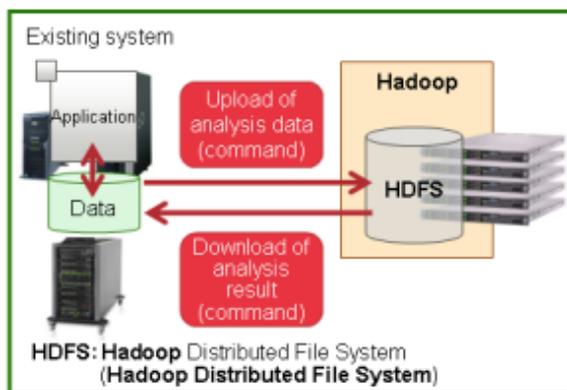
The data stored in the storage system can be accessed directly, and processed, using the Fujitsu distributed file system in addition to the "Apache Hadoop" distributed file system.

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.

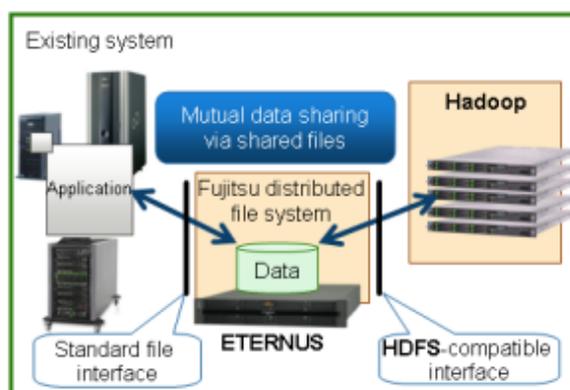
### Use of existing tools without modification

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.

[When standard **Hadoop** file system is used]

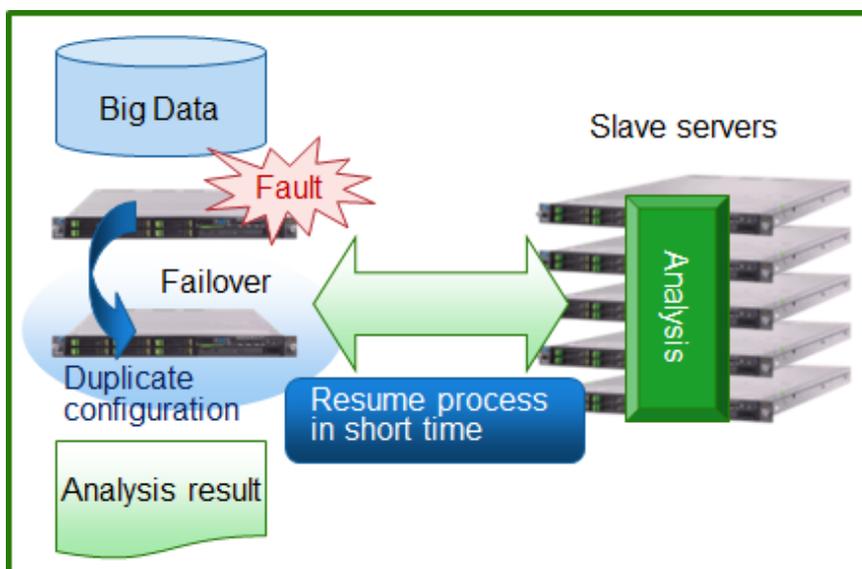


[When Fujitsu distributed file system is used]



## 1.2.2 Master Server Repliation

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.



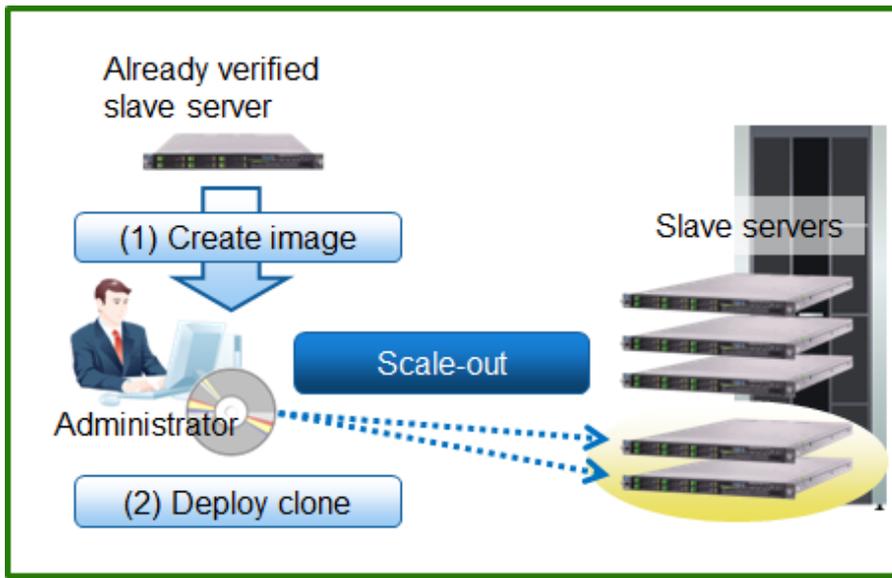
### 1.2.3 Smart Setup

---

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.

# Chapter 2 Designing the System and Preparing for Installation

This chapter explains how to design a system to use with this product and the preparations required for installation.

## 2.1 System Configuration and Requirements

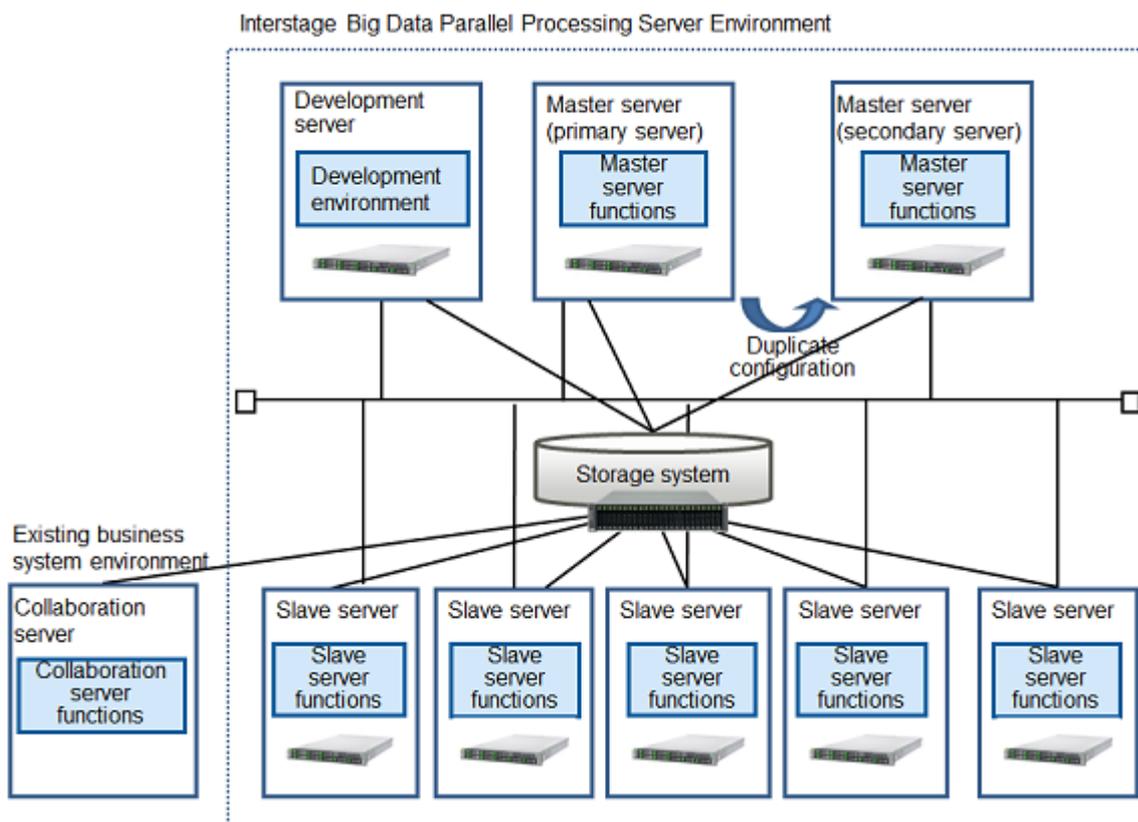
This section describes the system configuration and requirements for using this product.

### 2.1.1 Hardware Configuration

This section describes the hardware configuration for using this product.

#### 2.1.1.1 Server Configuration

##### Server configuration



##### 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.

This product requires that the master server is duplicated (a primary server and a secondary server).

This product's master server functions are installed on both the primary and secondary server of 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 functions are 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 functions are 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.

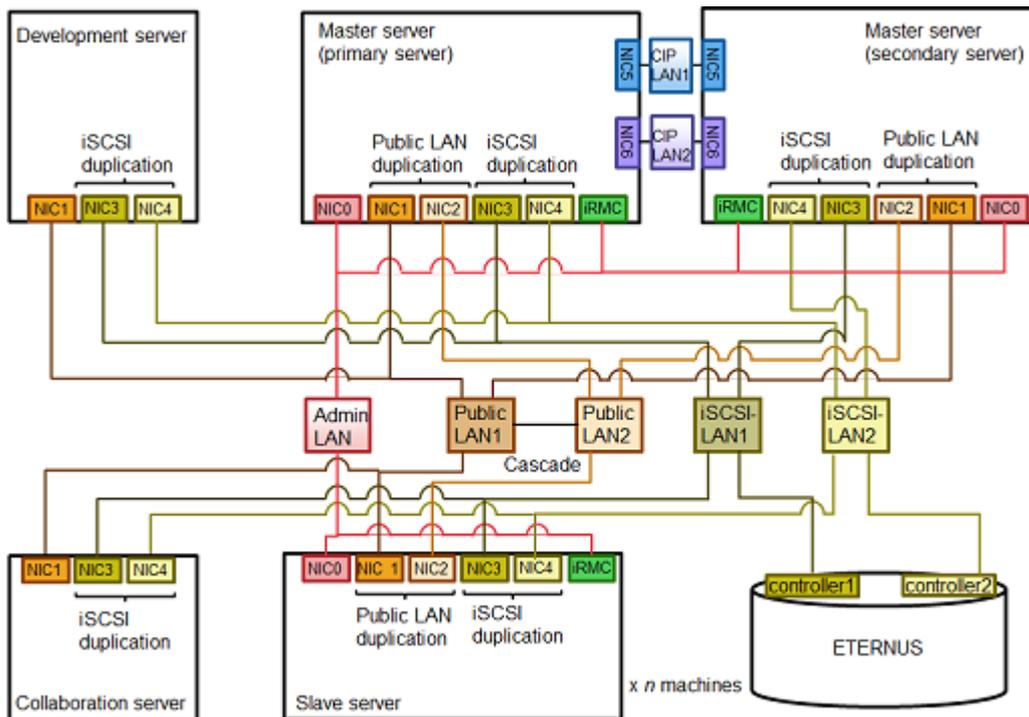
This product's collaboration server functions are installed at the collaboration server.

#### Note

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.

### 2.1.1.2 Network Configuration

## Network configuration



### 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.

### Public LAN

This is the LAN for analysis processing tasks between a master server and slave servers.

A configuration without redundancy is possible. However, by using the network redundancy software incorporated in this product, public LAN redundancy is possible even if a fault occurs in one LAN.

The configuration file (bdpp.conf) used during installation can be used to set whether or not redundancy is used. Refer to "[A.2.1 bdpp.conf](#)" for details of bdpp.conf.

### 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.

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

We recommend a redundancy configuration by means of the ETERNUS multipath driver as a precaution in case of a LAN fault.

### Cluster interconnect (CIP) LAN

This is the LAN used for a HA cluster configuration for the primary server and secondary server. A configuration without redundancy is possible, but we recommend redundancy as a precaution in case of a LAN fault.

The configuration file (bdpp.conf) used during installation can be used to set whether or not redundancy is used. Refer to "[A.2.1 bdpp.conf](#)" for details of bdpp.conf.

### 2.1.1.3 Hardware Requirements

The following hardware conditions must be met when using this product.

Table 2.1 Mandatory hardware conditions

Function	Hardware	Notes
Master server	PRIMERGY RX Series, PRIMERGY TX Series(*1)	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(*1)	The CPU must be at least a dual-core CPU.
Collaboration server	PRIMERGY RX Series, PRIMERGY TX Series(*1)	The CPU must be at least a dual-core CPU.
External storage apparatus	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.

- 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>

#### 2.1.1.3.1 Static Disk Size

The static disk sizes below are required for a new installation of this product. Disk sizes vary in accordance with differences in the environment being checked.

#### Static disk size (not including OS)

[Master server]

OS type	Directory	Disk size (unit: megabytes)
Linux	/opt	900
	/etc	16
	/var	220
	/usr	60

[Slave server]

OS type	Directory	Disk size (unit: megabytes)
Linux	/opt	120
	/etc	1
	/var	25
	/usr	60

[Development server]

OS type	Directory	Disk size (unit: megabytes)
Linux	/opt	10
	/etc	1
	/var	1
	/usr	60

[Collaboration server]

OS type	Directory	Disk size (unit: megabytes)
Linux	/opt	30
	/etc	1
	/var	1
	/usr	60

### 2.1.1.3.2 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.

Table 2.2 Dynamic disk size

Installation type	Directory	Disk size (unit: megabytes)
Master server	/etc	2
	/var/opt	2510
	Cloning image file storage directory Default: /var/opt/FJSVscw-deploysv/depot	Cloning image file storage area
Agent	/etc	1
	/var/opt	1

### Cloning image file storage area

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

Allocate area on the master server as an area to store the slave server cloning image files that are collected when cloning is used.



#### Note

- Create the cloning 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 cloning image and the servers targeted as clones must be the same model. If there are different models, a separate cloning image must be created for each model. Refer to "4.9.1.1.3 Creating a Cloning Image" for details.

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

$\text{Cloning image file storage area} = \text{Disk space used by one slave server} * \text{Compression ratio} * \text{Number of models}$
--

### 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.

## 2.1.1.3.3 Memory Size

The following memory sizes are required in order to use this product.

### Memory size (not including OS)

[Master server]

OS type	Memory size (unit: gigabytes)
Linux	8.0 or more

[Slave server]

OS type	Memory size (unit: gigabytes)
Linux	4.0 or more

[Development server]

OS type	Memory size (unit: gigabytes)
Linux	4.0 or more

[Collaboration server]

OS type	Memory size (unit: gigabytes)
Linux	4.0 or more

## 2.1.1.3.4 Expansion Card Options

Two or more network interface cards will be required when building LAN redundancy.

## 2.1.2 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	Function overview
Interstage Big Data Parallel Processing Server Standard Edition V1.0.0 Master server	The master server functions of this product
Interstage Big Data Parallel Processing Server Standard Edition V1.0.0 Slave server	The slave server functions of this product
Interstage Big Data Parallel Processing Server Standard Edition V1.0.0 <b>Development server</b>	The development server functions of this product
Interstage Big Data Parallel Processing Server Standard Edition V1.0.0 <b>Collaboration server</b>	The collaboration server functions of this product

### 2.1.2.1 Software Conditions

This section describes the software conditions that apply when installing this product.

### 2.1.2.2 System Software

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

#### System software

[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) 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)	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 function 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) 6 (for Intel64) Red Hat(R) Enterprise Linux(R) 6.1 (for Intel64) Red Hat(R) Enterprise Linux(R) 6.2 (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.

## Mandatory packages

The mandatory packages shown in "[C.1 List of 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.

### 2.1.2.3 Mandatory Software

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

[Master server]

Software name	Version	Notes
IPMI drivers	Obtain the latest edition.	IPMI drivers appended to ServerView
Microsoft(R) LAN Manager module	-	Mandatory on only the primary server. Obtain from the Microsoft FTP site. (*1)

\*1: Obtain from the following Microsoft FTP site:

Microsoft FTP site

URL: <code>ftp://ftp.microsoft.com/bussys/clients/msclient/DSK3-1.EXE (As of June 2012)</code>
---

[Slave server]

Software name	Version	Notes
ServerView Agent for Linux	V4.50.12 or later	-
IPMI drivers	Obtain the latest edition.	IPMI drivers appended to ServerView

[Development server]

Software name	Version	Notes
IPMI driver	Obtain the latest edition.	IPMI drivers appended to ServerView

[Collaboration server]

Software name	Version	Notes
IPMI driver	Obtain the latest edition.	IPMI drivers appended to ServerView

### 2.1.2.4 Exclusive Software

None.

### 2.1.2.5 Related Software

Consider installation of the following software, if required:

[Master server]

Software name	Version	Notes
ETERNUS Multipath Driver	V2.0L22	Required if providing redundant connections to the storage system

[Slave server]

Software name	Version	Notes
ETERNUS Multipath Driver	V2.0L22	Required if providing redundant connections to the storage system

[Development server]

Software name	Version	Notes
Interstage Application Server Enterprise Edition V10 or Interstage Application Server Standard-J Edition V10	V10.0.0	Required if developing analysis applications (MapReduce applications) that run under Hadoop
ETERNUS Multipath Driver	V2.0L22	Required if providing redundant connections to the storage system

[Collaboration server]

Software name	Version	Notes
ETERNUS Multipath Driver	V2.0L22	Required if providing redundant connections to the storage system

## 2.2 Designing the System

---

This section describes items that need to be designed before using this product.

1. Design the system configuration

### 2.2.1 Designing the System Configuration

---

The system configuration design items shown below are required when building this product.

1. Design the server configuration
2. System redundancy configuration design
3. DFS file system design

### Design the server configuration

This product scales out slave servers, and through the addition of slave servers, scalability is improved. It is recommended to make an estimate of how many servers are required by performing prototype tests before using the system in business, as the time required for processing depends on such factors as the number of slave servers used, 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.

### System redundancy configuration design

A system having higher reliability can be built by installing related products in addition to this product.

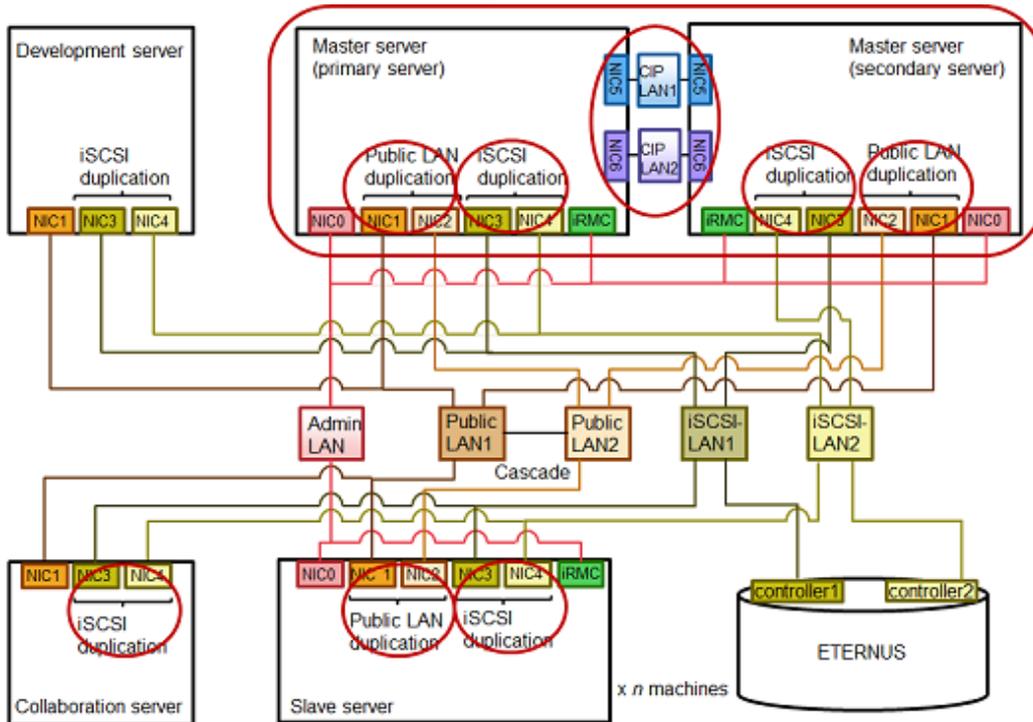
The systems that can be built using this product and related products are as follows.

Table 2.3 Supported redundancy configurations

Supported configuration	Supported server type	Means of support
Server repliation	Master server	Supported with just this product HA cluster configuration 1:1 active standby
Cluster interface repliation / non-repliation	Master server	Supported with just this product
Public LAN repliation / non-repliation	Master server Slave server	Supported with just this product Master server: NIC switching method (logical IP address inheritance) Slave server: NIC switching method (physical IP address inheritance)
Storage system connection repliation / non-repliation	Master server Slave server Development server Collaboration server	Supported if a related product, the " ETERNUS Multipath Driver ", is installed

Design the system configuration by selecting from the "Redundancy configuration targets" (parts outlined in red) the configurations that suit the customer requirements.

Figure 2.1 Redundancy configuration targets



## DFS file system design

A file system design, which should be investigated in advance, is required for using a DFS file system.

Refer to "D.1 File System Design" in "Appendix D DFS Environment Construction" for information on file system design.

This manual uses the following configuration as an example to explain file system configurations:

- 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
- Master server (primary): master1
- Master server (secondary): master2

## Point

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

Check whether each server can recognize the disk partitions having by-id names.

- Red Hat Enterprise Linux 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
```

- Red Hat Enterprise Linux 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
```

---

## 2.2.2 Designing the bdpp.conf Configuration File

To install the functions for the master server, slave server, development server and collaboration server, design and check the parameters specified in the bdpp.conf configuration file.

Refer to "[A.2.1 bdpp.conf](#)" for information on the bdpp.conf configuration file.

---

## 2.3 Preparing the Base System

This section describes points that need to be designed before using this product.

---

### 2.3.1 Preparing the Base System

The preparations below are required in order to build a system using this product.

#### Server environment preparations

Firstly, new installation of the system software and the system software mandatory packages at each of the servers shown below is a prerequisite.

Refer to "[2.1.2.2 System Software](#)" and "[Appendix C Mandatory Packages](#)" for information on the system software and mandatory packages.

- Two master servers (primary server and secondary server)
- One slave server
- At least one development server
- At least one collaboration server

Then, if ETERNUS is used as the storage system and connection redundancy is to be used, the "ETERNUS multipath driver" must also be installed in advance.



---

#### About the SELinux (Security-Enhanced Linux) function

This product cannot be installed in a Linux environment where the SELinux (Security-Enhanced Linux) function is enabled. To install this product, disable the SELinux function, and then install this product. Refer to the Linux online manual for the method for disabling the SELinux function.

## Network environment preparations

These servers and the storage system must be connected in the configuration shown at "[2.1.1.2 Network Configuration](#)".

Check the hardware connection configuration of all hardware.

## 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.

## 2.3.2 Performing Advance Preparations for Installing Servers

---

This section describes the advance preparations required before installing this product.

- Server BIOS settings  
Refer to "[Server BIOS settings](#)".
- File system when installing the OS  
Refer to "[File system when installing the OS](#)".
- System parameter settings  
Refer to "[System parameter settings](#)".
- Firewall settings  
Refer to "[Firewall settings](#)".
- Host name settings  
Refer to "[Host name settings](#)".
- Public LAN network interface settings  
Refer to "[Public LAN network interface settings](#)".
- System time settings  
Refer to "[System time settings](#)".
- Mandatory software preparations and checking  
Refer to "[Mandatory software preparations and checking](#)".

### Server BIOS settings

To introduce another slave server by cloning, it is necessary to set the BIOS settings of the slave servers as follows.

#### System BIOS

Set the boot sequence as follows.

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

## 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".

## System parameter settings

The system parameters must be tuned. Refer to "[A.4 Tuning System Parameters](#)" for the method for setting tuning settings, and tune the master server system parameters.

## 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 "[A.3 Port List](#)" for the ports used by this product, and ensure connection is authorized.

## Host name settings

For correct operation on the master server, slave server, development server, and collaboration server, it is necessary to set the host name (the official host name output by a command such as the "hostname" command) of the public LAN used by Hadoop.

The host names of the servers must have 11 alphanumeric characters with the first character being alphabetical. The names must be entered in the hosts file of the master server.

Make sure the hosts file on the primary and secondary master servers match.

Also add the host name information added to the master server to the hosts files of the slave server, development server and collaboration server.

hosts file

```
/etc/hosts
```

Refer to the help for "hosts" for information on the hosts file.

## 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.

In the following example, the host name "master1" having the IP address "10.10.10.10" is set.

```
10.10.10.10 master1
127.0.0.1 localhost.localdomain localhost
```

## 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 eth5 <Enter>
eth5      Link encap:Ethernet  HWaddr 00:19:99:8D:7E:90
          inet addr:10.10.10.11  Bcast:192.168.1.255  Mask:255.255.255.0
          ... omitted ...
          RX bytes:2440357552 (2.2 GiB)  TX bytes:1144015405 (1.0 GiB)
          Memory:cd400000-cd480000
```

## 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.

## Mandatory software preparations and checking

Before installing this product, check that the software in "[2.1.2.3 Mandatory Software](#)" is installed in the system. If not installed, it must be installed before installing this product.

### Note

#### ServerView Agent settings

In order for this product to operate correctly with the PRIMERGY Series, set the required SNMP service settings during ServerView Agent installation.

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.

## 2.3.3 Installing Mandatory Software

Install the software indicated in "[2.1.2.3 Mandatory Software](#)" on the master server, slave server, development server, and collaboration server.

The following explains the mandatory software settings requiring particular attention.

### 2.3.3.1 Obtaining and Setting Up the Mandatory Software

#### Obtaining mandatory software

- Obtaining and deploying the Microsoft LAN Manager module

This task is required on only the primary server, not the secondary server.

Obtain the Microsoft LAN Manager module from the FTP site shown below.

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

```
URL:
ftp://ftp.microsoft.com/bussys/clients/msclient/DSK3-1.EXE (as of June 2012)
```

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.

The obtained module must be expanded in advance under the Windows of the CPU architecture (x86) by means of the *Expand* command. 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 system being installed:

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

#### Mandatory software settings

- SNMP trap daemon settings

This task is required on only the primary server, not the secondary server.

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. If the file does not exist, create the file, and then add the following setting:

```
disableAuthorization yes
```

- kdump settings

##### 1. Check kdump

Check that the kdump server function can be used. If it cannot be used, make it usable. Use the runlevel command and the chkconfig command to perform the checking operation.

- 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
```

---

- 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 3:off 4:off 5:off 6:off
```

---

- 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
```

# Chapter 3 System Build Method

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

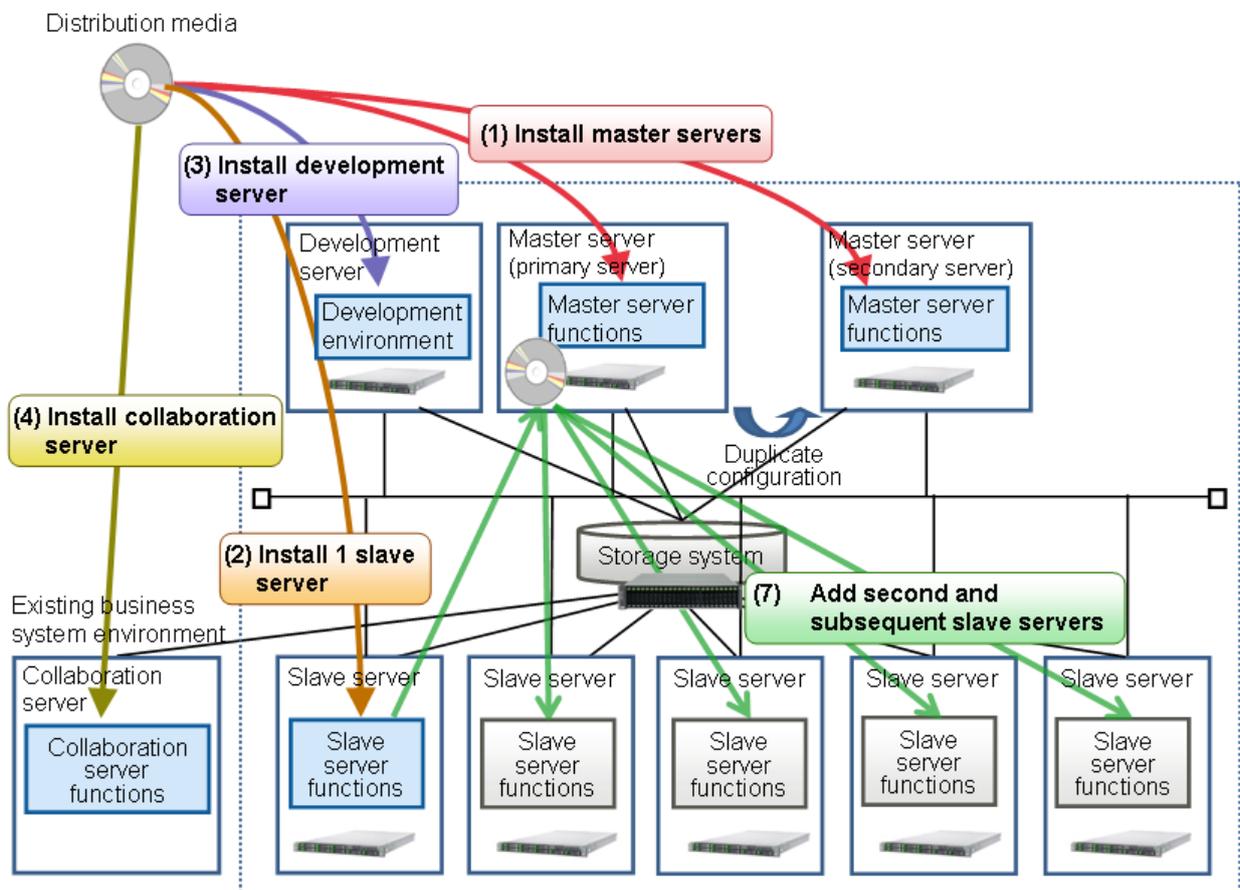
## 3.1 System Build Flow

The work flow for building a system using this product is shown below.

1. Installing to master servers.
2. Installing the first slave server.
3. Installing a development server.
4. Installing a collaboration server.
5. Distributing the DFS file system configuration information.
6. Checking the connection to the first slave server.
7. Adding second and subsequent slave servers. (\*1)

\*1: Smart Setup can be used to add second and subsequent slave servers.

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.



### 3.1.1 Installing to Master Servers

---

Check "[2.3.2 Performing Advance Preparations for Installing Servers](#)", and when pre-installation preparations are done, install the master servers.

Refer to "[4.1 Installing to Master Servers](#)" for the master server installation procedure.

After master server installation, refer to "[6.1.1 Applying Updates](#)" and apply the updates for this product and the updates for software included with this product.

### 3.1.2 Installing the First Slave Server

---

Check "[2.3.2 Performing Advance Preparations for Installing Servers](#)", and when pre-installation preparations are done, install a slave server.

Refer to "Installing to a Slave Server" for the slave server installation procedure.

After slave server installation, refer to "[6.1.1 Applying Updates](#)" and apply the updates for this product and the updates for software included with this product.

### 3.1.3 Installing a Development Server

---

Check "[2.3.2 Performing Advance Preparations for Installing Servers](#)", and when pre-installation preparations are done, install a development server.

Refer to "Installing to a Development Server" for the development server installation procedure.

Refer to "[6.1.1 Applying Updates](#)" after installing the development server environment, and apply all patches to this product and the software that came with it.

### 3.1.4 Installing a Collaboration Server

---

Check "[2.3.2 Performing Advance Preparations for Installing Servers](#)" and when pre-installation preparations are done, install a collaboration server.

Refer to "Installing to a Collaboration Server" for the collaboration server installation procedure.

Refer to "[6.1.1 Applying Updates](#)" after installing the collaboration server and apply all patches to this product and the software that came with it.

### 3.1.5 Distributing the DFS File System Configuration Information

---

Distribute the DFS file system configuration file generated on the master server to the slave server development server, and collaboration server. Refer to "[4.1.3.7 Generating and Distributing the DFS File System Configuration Information](#)" for information on generating DFS file systems.

The procedure for distributing file system configuration information is described below using the following environment as an example:

- File system ID: 1
- Logical file system name: pdfs1
- Slave servers, development server, collaboration server: slave1, slave2, slave3, slave4, slave5, develop, collaborate

1. Log in to the master server with root permissions.
2. Confirm that the user ID for executing MapReduce is set in the configuration file created in "[4.1.3.7 Generating and Distributing the DFS File System Configuration Information](#)".  
If the user ID for executing MapReduce is not set, set the user ID for executing MapReduce by checking in "[4.1.3.3 Setting the User ID for Executing MapReduce](#)".

```
# cat <directory where configuration file was created>/client.conf.pdfsl
FSID 1
MDS kawa4 29000
MDS kawa5 29000
DEV /dev/disk/by-id/scsi-1FUJITSU_300000370107 0 7341778
DEV /dev/disk/by-id/scsi-1FUJITSU_300000370108 0 6578704
MAPRED mapred <-- Check that the user ID for executing MapReduce is set
```

3. Distribute the configuration information file to the slave servers, development server and the collaboration server.

```
# cd {directory where configuration file was created} <Enter>
# scp -p ./client.conf.pdfsl root@slave1:/etc/pdfs/client.conf.pdfsl <Enter>
# scp -p ./client.conf.pdfsl root@develop:/etc/pdfs/client.conf.pdfsl <Enter>
# scp -p ./client.conf.pdfsl root@collaborate:/etc/pdfs/client.conf.pdfsl <Enter>
```



**Note**

Place the configuration information file under "/etc/pdfs" at each server.

### 3.1.6 Checking the Connection to the First 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.

1. If not logged in to the master server (primary), log in with root permissions.
2. 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
```

3. 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
slaves1,slaves2,slaves3,slaves4,slaves5
```

After editing:

```
# cat /etc/opt/FJSVbdpp/conf/slaves
slaves1
```

4. Execute `bdpp_changeslaves` command to effect the changes that were made to the slave server definition file. Refer to "[A.1.3 bdpp\\_changeslaves](#)" for information on the `bdpp_changeslaves` command.

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

5. Start the Hadoop of the Interstage Big Data Parallel Processing Server.  
Log in to the master server with root permissions and use the `bdpp_start` command to start the Hadoop of the Interstage Big Data Parallel Processing Server.  
Refer to "[A.1.11 bdpp\\_start](#)" for details of the `bdpp_start` command.

```
# /opt/FJSVbdpp/bin/bdpp_start
```

6. Use the Interstage Big Data Parallel Processing Server status display to check whether or not TaskTracker is started at the targeted slave server.  
After checking that the system for the slave server is running, execute the `bdpp_stat` command on the master server to display the status of the Hadoop of the Interstage Big Data Parallel Processing Server.  
Refer to "[A.1.11 bdpp\\_start](#)" for details of the `bdpp_stat` command.

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

7. 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.  
Refer to "[5.3 Executing and Stopping Jobs](#)" for information on executing Hadoop jobs.
8. Stop the Hadoop of the Interstage Big Data Parallel Processing Server.  
Use the `bdpp_stop` command to stop the Hadoop of the Interstage Big Data Parallel Processing Server.  
Refer to "[A.1.13 bdpp\\_stop](#)" for details of the `bdpp_stop` command.

```
# /opt/FJSVbdpp/bin/bdpp_stop
```

9. 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
# cp /etc/opt/FJSVbdpp/conf/slaves.bak /etc/opt/FJSVbdpp/conf/slaves
```

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

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



When starting or stopping the Hadoop of the Interstage Big Data Parallel Processing Server, a password input prompt will be displayed because Hadoop is started and stopped via SSH.

To reduce the workload by avoiding password input, distribute the public key of the SSH of the root user of the master server to the slave server and development server so that SSH without a password can be performed. Refer to the `ssh-keygen` command help for details.

### 3.1.7 Adding Second and Subsequent Slave Servers

A second and subsequent slave servers can be added quickly and easily using the cloning function.

1. Stop the Hadoop of the Interstage Big Data Parallel Processing Server if it is running before adding slave servers.  
Use the `bdpp_stop` command to stop the Hadoop of the Interstage Big Data Parallel Processing Server.  
Refer to "[A.1.13 bdpp\\_stop](#)" for information on the `bdpp_stop` command.

```
# /opt/FJSVbdpp/bin/bdpp_stop
```

2. Refer to "[4.9.1.1 Adding a Slave Server During Initial Installation](#)" for the procedure for adding slave servers.
3. Start the Hadoop of the Interstage Big Data Parallel Processing Server .  
Log in to the master server with root permissions and use the `bdpp_start` command to start the Hadoop of the Interstage Big Data Parallel Processing Server.  
Refer to "[A.1.11 bdpp\\_start](#)" for information on the `bdpp_start` command.

```
# /opt/FJSVbdpp/bin/bdpp_start
```

# Chapter 4 Installing and Uninstalling the Product

This chapter explains how to install and uninstall this product.

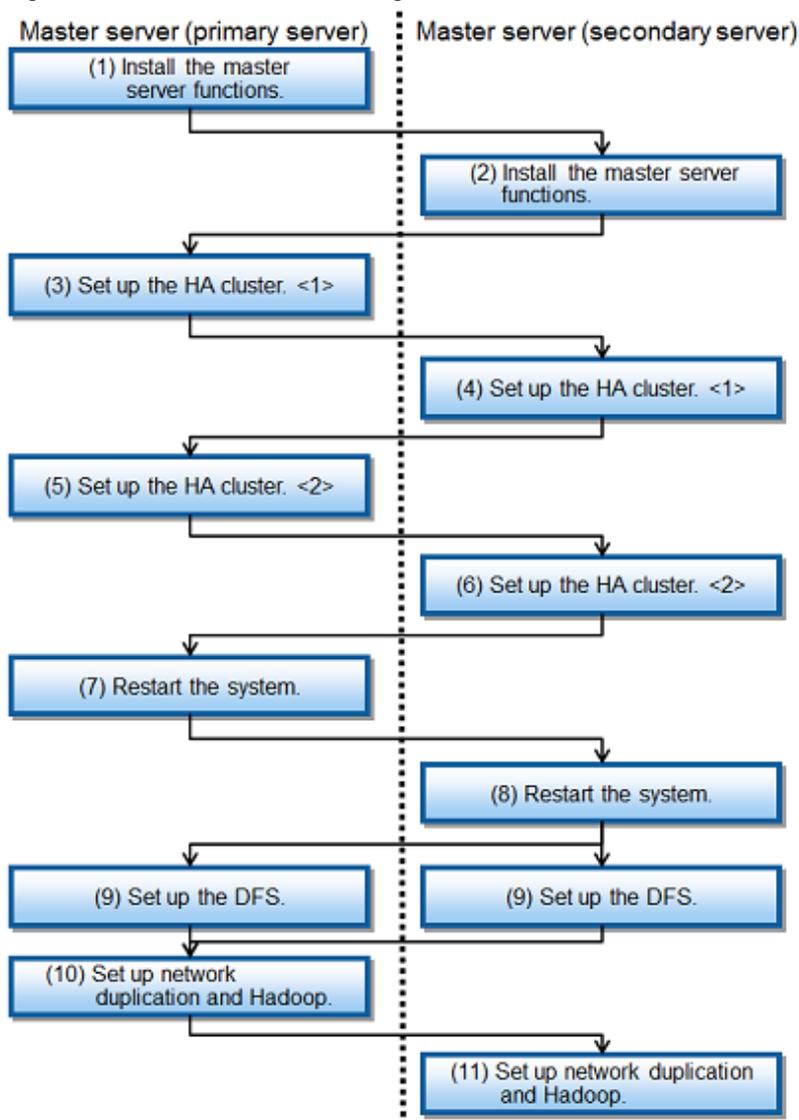
## 4.1 Installing to Master Servers

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

The installation sequence is to the master server (primary) first, then the master server (secondary), in order to build a master server duplicate configuration.

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

Figure 4.1 Procedure for installing master server functions



## 4.1.1 Installing Master Server Functions

---

Follow the procedures below to install the master server functions.

1. SSH Settings Between the Master Server (primary) and the Master Server (secondary)
2. Create the configuration file `bdpp.conf`
3. Install the master server functions

### 4.1.1.1 SSH Settings Between the Master Server (primary) and the Master Server (secondary)

Create a public key on both the master server (primary) and master server (secondary) so that the SSH can communicate between the two without a password.

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

### 4.1.1.2 Creating the Configuration File `bdpp.conf`

Before installing the master server functions, set the previously designed values in the "`bdpp.conf`" configuration file.

Refer to "Master server" in "[A.2.1 `bdpp.conf`](#)" for information on the configuration file "`bdpp.conf`".

### 4.1.1.3 Installing Master Server Functions

Install the master server functions to the master server (primary) first, and then to the master server (secondary).

Follow the procedures below to install the master server functions.

1. Log in to the master server with root permissions.
2. Set the "`BDPP_CONF_PATH`" environment variable.  
In the "`BDPP_CONF_PATH`" environment variable, set the directory where the "`bdpp.conf`" configuration file is stored.

#### Example

.....  
If the "`bdpp.conf`" configuration file was created under the "`/home/data/master`" directory:

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

.....

#### Note

.....  
The "`bdpp.conf`" configuration file must have the same settings at both the master server (primary) and the master server (secondary).

Before the master server (secondary) is installed, the "`bdpp.conf`" configuration file used at the master server (primary) must be copied to the master server (secondary).

.....

3. 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 {DVD-ROM mount point} <Enter>
```

4. Start the installer.  
Execute `bdpp_inst` to start the installer.

```
# {DVD-ROM mount point}/bdpp_inst <Enter>
```

5. Select "1. Master Server installation".

```
=====
                    Interstage Big Data Parallel Processing Server
                        V1.0.0
                    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 came up 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.

## 4.1.2 HA Cluster Setup

The sequence for setting up master server HA clusters is master server (primary) first, then the master server (secondary).

The HA cluster setup procedure is shown below.

1. Setup starting with the master server (primary).  
Log in to the master server (primary) with root permissions.
2. 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
```

### Note

- ethX is the network interface used by the cluster interconnect.  
Specify a numeric 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>
```

3. Execute HA cluster setup 1.  
Execute cluster\_setup.

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

4. Next setup with the master server (secondary).  
Log in to the master server (secondary) with root permissions and perform steps 2 and 3.

5. Execute HA cluster setup 2 first on the master server (primary) and then on the master server (secondary).  
Execute cluster\_setup2 on each server.

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

6. After performing the setup procedure, restart the system first on the master server (primary) and then on the master server (secondary).

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



### Note

If setup fails, refer to the messages came up 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.

## 4.1.3 DFS Setup

The DFS setup sequence is as follows:

1. Create a management partition
2. Create the file system
3. Set the user ID for executing MapReduce
4. Register the slave server, development server and collaboration server (DFS client) information
5. Create the mount point and set fstab settings
6. Mount
7. Generate and distribute the DFS file system configuration information

### 4.1.3.1 Creating a Management Partition

Refer to "D.3 Management Partition Creation" in "Appendix D DFS Environment Construction" for the method for creating a management partition.



### Example

1. Create a management partition (implement at the master server (primary))

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

2. Add the DFS management server information to the management partition (implement at the master server (primary) and implement the same operation at the master server (secondary server))

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

3. Check that the DFS management server information was added (implement at the master server (primary) and implement the same operation at the master server (secondary server))

```
# pdfssetup <Enter>
HOSTID          CIPNAME          MP_PATH
80000001        master1RMS       yes
80000000        master2RMS       yes
# pdfssetup -p <Enter>
/dev/disk/by-id/scsi-1FUJITSU_300000370105
```

4. Start the pdfsfrmd daemon (implement the same operation at the master server (primary) and the master server (secondary server))

```
# pdfsfrmstart <Enter>
```

### 4.1.3.2 Creating the file system

Refer to "D.4 File System Creation" in "Appendix D DFS Environment Construction" for the method for creating a file system.



#### Example

1. Create the file system (Master server (primary))

- 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
- Master server (primary):master1
- Master server (secondary):master2

```
# 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>
```

2. Check the file system information

```
# 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_300000370106 (864) 232256 DATA -----
1 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 7341778 DATA -----
1 /dev/disk/by-id/scsi-1FUJITSU_300000370108 (896) 6578704 DATA -----
```

### 4.1.3.3 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.

1. If not logged in to the master server, log in with root permissions.
2. Unmount the DFS if it is mounted.

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

3. 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>
```

4. Check that the user ID has been set.

```
# pdfsinfo -e /dev/disk/by-id/scsi-1FUJITSU_300000370106 <Enter>
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.

### 4.1.3.4 Registering the Slave Server, Development Server and Collaboration Server (DFS Client) Information

Register slave server, development server and collaboration server information in the connection authorization list. Register this connection approval list file on the master server (primary) then distribute it to the master server (secondary). Refer to "D.4.3 Registering Slave Server, Development Server and Collaboration Server Information" in "Appendix D DFS Environment Construction" for information on registering slave server, development server and collaboration server (DFS client) information.



## Example

1. Check the file system ID.

```
# 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_300000370106 (864) 232256 DATA -----
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 -p ./server.conf.sample ./server.conf.1 <Enter>
```

3. Edit the connection authorization list file and define the slave server, development server and collaboration server information. Only one slave server is installed initially, but include all the slave servers, development servers and collaboration servers that will be connection targets and define them in advance.

```
#
# 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 collaborate <-- Collaboration server to be added
CLIENT slave1 <-- Slave server to be added
CLIENT slave2 <-- Slave server to be added
CLIENT slave3 <-- Slave server to be added
CLIENT slave4 <-- Slave server to be added
CLIENT slave5 <-- Slave server to be added

```

### 4.1.3.5 Creating the Mount Point and Setting fstab Settings

#### Creating the mount point

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

Mount points should be created on both the master server (primary) and the master server (secondary).



#### Example

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

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

#### fstab settings

At "/etc/fstab", define the mount points created above and the DFS representative partitions.

fstab settings should be implemented on both the master server (primary) and the master server (secondary).



#### 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

/dev/disk/by-id/scsi-1FUJITSU_300000370106 /mnt/pdfsf       pdfsf   noauto,noatime 0 0

```

### 4.1.3.6 Mounting

Mount the DFS file system on the master server (primary). Refer to "D.4.6 Mount" in "Appendix D DFS Environment Construction" for information on mounting a DFS file system.

#### Example

Mount the DFS file system on the master server (primary).

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

### 4.1.3.7 Generating and Distributing the DFS File System Configuration Information

Generate the DFS file system configuration file on the master server (primary).

The procedure for generating the file system configuration information is explained using the following environment as an example.

- File system ID :1
- Logical file system name :pdfs1
- Slave server, development server and collaboration server :slave1, slave2, slave3, slave4, slave5, develop, collaborate

1. Log in to the master server (primary) with root permissions.
2. 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_300000370106 (864) 232256 DATA -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 7341778 DATA -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370108 (896) 6578704 DATA -----
```

3. Generate the DFS configuration file with the pdfsmkconf command.  
Execute the pdfsmkconf command on the master server (primary).

```
# pdfsmkconf <Enter>
```

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

```
# cd pdfsmkconf_out <Enter>
# mv ./client.conf.1 client.conf.pdfs1 <Enter>
```

#### Note

The configuration file is created in the directory where the pdfsmkconf command was executed.

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

5. Distribute the configuration file to the master server (secondary).

```
# scp ./client.conf.pdfs1 root@master2:/etc/pdfs/client.conf.pdfs1 <Enter>
```

### Note

Put the configuration file into the "/etc/pdfs" directories on each server.

### See

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

## 4.1.4 Executing Network Repliation and Hadoop Setup

The sequence for network repliation and Hadoop setup is the master server (primary) first, then the master server (secondary).

The setup sequence is as follows:

1. Edit the slave server definition file 'slaves'.
2. Execute the network repliation and Hadoop setup.

### 4.1.4.1 Editing the Slave Server Definition File 'slaves'

Edit the slave server definition file (/etc/opt/FJSVbdpp/conf/slaves) at the master server (primary ) to define the slave servers that will be connection targets.

Refer to "[A.2.2 slaves](#)" for details of the slave server definition file 'slaves'.

Copy the edited slave server definition file to the master server (secondary).

### 4.1.4.2 Executing Network Repliation and Hadoop Setup

The network repliationand Hadoop setup procedure is shown below.

1. If not logged in the master server, log in with root permissions.
2. Execute the network repliationand Hadoop setup.

Execute bdpp\_setup.

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

### Note

If setup fails, refer to the messages came up 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 "[D.5 Hadoop Configuration Settings](#)" for information on tuning the configuration settings.

If the configuration settings are changed, copy the configuration settings information of each master server (master server (primary) and the master server (secondary), slave server and development server and ensure that all the settings are the same.

## 4.2 Installing to a Slave Server

---

This section describes the installation of the slave server functions 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 functions on a system where only the operating system has been installed.

The installation and setup procedures are shown below.

1. Installing slave server functions
2. DFS Setup
3. Network replication and Hadoop setup



Refer to "2.3.2 Performing Advance Preparations for Installing Servers" - "File system when installing the OS" before installing a second slave server by adding a clone slave server, and ensure that the slave server has been installed with the file system type "EXT3".

### 4.2.1 Installing Slave Server Functions

---

The procedure for installing slave server functions is shown below.

1. Creating the configuration file `bdpp.conf`
2. Installing slave server functions
3. DFS setup
4. Network replication and Hadoop setup

#### 4.2.1.1 Creating the Configuration File `bdpp.conf`

Before installing the slave server functions, set the previously designed values in the "`bdpp.conf`" configuration file.

Refer to "Slave server" under "[A.2.1 `bdpp.conf`](#)" for details of the "`bdpp.conf`" configuration file.



For the values in the "`BDPP_PDFS_FILESYSTEM`" and "`BDPP_PDFS_MOUNTPOINT`" parameters in the "`bdpp.conf`" file, specify the same values as the parameter values specified in the master server "`bdpp.conf`" file.

#### 4.2.1.2 Installing Slave Server Functions

Follow the procedures below to install the slave server functions.

1. Log in to the slave server with root permissions.
2. Set the "`BDPP_CONF_PATH`" environment variable.  
In the "`BDPP_CONF_PATH`" environment variable, set the directory where the "`bdpp.conf`" configuration file is stored.

## Example

If the "bdpp.conf" configuration file was created under the "/home/data/slaves" directory:

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

### 3. 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 {DVD-ROM mount point} <Enter>
```

### 4. Start the installer.

Execute bdpp\_inst to start the installer.

```
# {DVD-ROM mount point}/bdpp_inst <Enter>
```

### 5. Select "2. Slave Server installation".

```
=====
                    Interstage Big Data Parallel Processing Server
                          V1.0.0
                    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>
```

### 6. After installation is completed, restart the system.

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

## Note

If installation fails, refer to the messages came up 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.

## 4.2.2 DFS Setup

This section explains DFS setup. The DFS setup procedure is shown below.

1. Create the mount point and set fstab settings
2. Mount

### 4.2.2.1 Creating the Mount Point and Setting fstab Settings

## 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
pdfs	/mnt/pdf	pdfs	_netdev	0 0

## 4.2.2.2 Mounting

The "DFS file configuration information" must be distributed to the slave server before the DFS file system is mounted. Refer to "[3.1.5 Distributing the DFS File System Configuration Information](#)" if the "DFS file configuration information" is not yet distributed.

Refer to "[D.4.6 Mount](#)" in "[Appendix D DFS Environment Construction](#)" for information on mounting a DFS file system.



### Example

Mount the DFS file system at the slave server.

```
# mount pdfs
```

## 4.2.3 Executing Network Repliation and Hadoop Setup

The network repliation and Hadoop setup procedure is shown below.

1. If not logged in the slave server, log in with root permissions.
2. Copy the master server 'slaves' file, edited as described in "4.1.4.1 Editing 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>
```

3. Execute the network replication and Hadoop setup.

Execute bdpp\_setup.

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



- If the "BDPP\_GLS\_SERVER\_CONNECT2" parameter value is omitted from the "bdpp.conf" file, network replication setup is not performed, and only Hadoop is set up.
- If setup fails, refer to the messages came up 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.

## 4.3 Installing to a Development Server

---

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

This section describes the procedure for installing and setting up the development server functions.

### 4.3.1 Installing Development Server Functions

---

The procedure for installing development server functions is shown below.

1. Creating the configuration file bdpp.conf
2. Installing development server functions
3. DFS setup
4. Hadoop setup

#### 4.3.1.1 Creating the Configuration File bdpp.conf

Before installing the development server functions, set the previously designed values in the "bdpp.conf" configuration file.

Refer to "Development server" under "A.2.1 bdpp.conf" for details of the configuration file "bdpp.conf".

#### 4.3.1.2 Installing Development Server Functions

Follow the procedures below to install the development server functions.

1. Log in to the development server with root permissions.
2. Set the "BDPP\_CONF\_PATH" environment variable.  
In the "BDPP\_CONF\_PATH" environment variable, set the directory where the "bdpp.conf" configuration file is stored.



## Example

If the "bdpp.conf" configuration file was created under the "/home/data/slaves" directory:

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

### 3. 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 {DVD-ROM mount point} <Enter>
```

### 4. Start the installer.

Execute bdpp\_inst to start the installer.

```
# {DVD-ROM mount point}/bdpp_inst <Enter>
```

### 5. Select "3. Development Server installation".

```

=====
                Interstage Big Data Parallel Processing Server
                        V1.0.0
                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 came up 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 installaiton again.

## 4.3.2 DFS Setup

This section describes DFS setup. The DFS setup procedure is shown below.

1. Mount point creation and fstab settings
2. Mount

### 4.3.2.1 Mount Point Creation and fstab Settings

#### Mount point creation

Create the mount point for mounting a disk partition at the storage system used as the DFS.

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



## Example

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

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

### fstab settings

In "/etc/fstab", define the mount point 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
pdfs	/mnt/pdf	pdfs	_netdev	0	0

### 4.3.2.2 Mounting

The "DFS file configuration information" must be distributed to the slave server before the DFS file system is mounted. Refer to "[3.1.5 Distributing the DFS File System Configuration Information](#)" if the "DFS file configuration information" is not yet distributed.

Refer to "[D.4.6 Mount](#)" in "[Appendix D DFS Environment Construction](#)" for information on mounting a DFS file system.



## Example

Mount the DFS file system at the slave server.

```
# mount pdfs
```

### 4.3.3 Executing Hadoop Setup

The Hadoop setup procedure is shown below.

1. If not logged in to the development server, log in with root permissions.

2. Copy the master server 'slaves' file, edited as described in "[4.1.4.1 Editing 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>
```

3. Execute Hadoop setup.

Execute bdpp\_setup.

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



If setup fails, refer to the messages came up 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.

## 4.4 Installing to a Collaboration Server

---

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

This section describes the procedure for installing and setting up the collaboration server functions.

The installation and setup procedures are shown below.

1. Installing collaboration server functions
2. DFS Setup

### 4.4.1 Installing Collaboration Server Functions

---

The procedure for installing collaboration server functions is shown below.

1. Creating the configuration file bdpp.conf
2. Installing the collaboration server functions

#### 4.4.1.1 Creating the Configuration File bdpp.conf

To install collaboration server functions, configure the values that were designed in advance, in the configuration file "bdpp.conf".

Refer to "Collaboration server" under "[A.2.1 bdpp.conf](#)" for details of the configuration file "bdpp.conf".

#### 4.4.1.2 Installing Collaboration Server Functions

Follow the procedures below to install the collaboration server functions.

1. Log in to the collaboration server with root permissions.
2. Set the "BDPP\_CONF\_PATH" environment variable.  
In the "BDPP\_CONF\_PATH" environment variable, set the directory where the "bdpp.conf" configuration file is stored.



## Example

If the "bdpp.conf" configuration file was created under the "/home/data/slaves" directory:

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

### 3. 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 {DVD-ROM mount point} <Enter>
```

### 4. Start the installer.

Execute bdpp\_inst to start the installer.

```
# {DVD-ROM mount point}/bdpp_inst <Enter>
```

### 5. Select "4. Collaboration Server installation".

```

=====
                    Interstage Big Data Parallel Processing Server
                        V1.0.0
                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 came up 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 installaiton again.

## 4.4.2 DFS Setup

This section explains DFS setup. The DFS setup procedure is shown below.

1. Create the mount point and set fstab settings
2. Register the hadoop group and mapred user
3. Mount

### 4.4.2.1 Creating the Mount Point and Setting fstab Settings

#### Creating the mount point

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

## 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 "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/pdf	pdfs	_netdev	0	0

## 4.4.2.2 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 maser server.

```
# cat /etc/group <Enter>
xxx omitted xxx
hadoop:x:123:hbase
hbase:x:503:
bdppgroup:x:1500:
```

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

```
# cat /etc/passwd <Enter>
xxx omitted xxx
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:203:503::/tmp:/bin/bash
```

3. Register the hadoop group.

```
# groupadd -g 123 hadoop
```

4. Register the mapred user.

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



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.

### 4.4.2.3 Mounting

The "DFS file configuration information" must be distributed to the collaboration server before the DFS file system is mounted. Refer to ["3.1.5 Distributing the DFS File System Configuration Information"](#) if the "DFS file configuration information" is not yet distributed.

Refer to ["D.4.6 Mount"](#) in ["Appendix D DFS Environment Construction"](#) for information on mounting a DFS file system.



Mount the DFS file system at the slave server.

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

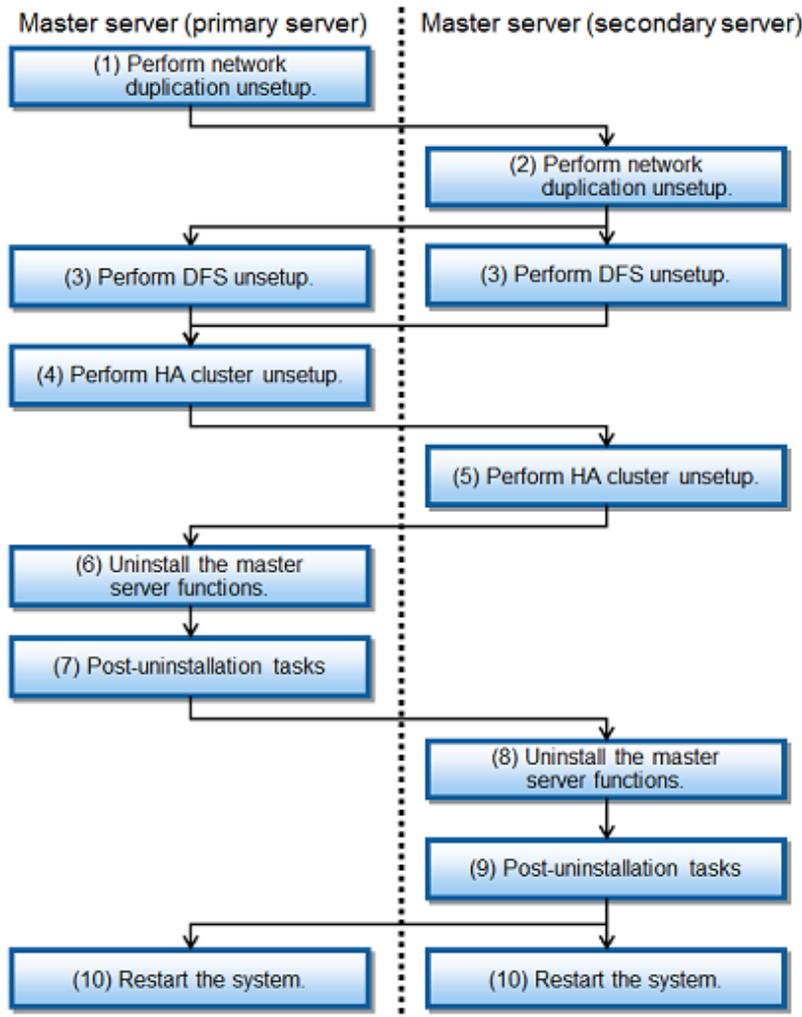
## 4.5 Uninstalling from a Master Server

---

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

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

Figure 4.2 Procedure for uninstalling master server functions



### 4.5.1 Network Repliation Unsetup

The network repliation unsetup procedure is shown below.

1. If not logged in to the master server, log in with root permissions.
2. Execute network repliation unsetup.

Execute bdpp\_unsetup.

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

 **Note**

If unsetup 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 perform the unsetup again.

## 4.5.2 DFS Unsetup

---

Refer to "E.1 Deleting a File System" in "Appendix E DFS Operation Changes" for information on DFS unsetup.

## 4.5.3 HA Cluster Unsetup

---

The HA cluster unsetup procedure is shown below.

1. If not logged in the master server, log in with root permissions.
2. Execute unsetup.

Execute `bdpp_cluster_unsetup`.

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



If unsetup fails, refer to the messages came up during the unsetup 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 unsetup again.

## 4.5.4 Uninstalling Master Server Function

---

The master server function 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 (Master server (primary) only)



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

2. If not logged in to the master server, log in with root permissions.
3. 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.
```

4. Reboot in single user mode.

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

5. Execute the uninstaller.

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

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

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

Return to the default run level that was recorded in Step 3.

```
# 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 `uninstall` again.

### Note

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)".

## 4.5.5 Post-Uninstallation Tasks

This section describes the master server post-uninstall tasks.

### 4.5.5.1 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:

- `/opt/FJSVbdpp`

- /etc/opt/FJSVbdpp
- /etc/opt/FJSVwvbs
- /etc/pdfs
- /var/opt/FJSVbdpp
- /var/opt/FJSVpdfs

### 4.5.5.2 Uninstall (middleware) Uninstallation

"Uninstall (middleware) Uninstallation" is a tool common to Fujitsu middleware products.

This section describes the uninstall method and gives notes on uninstalling the "Uninstall (middleware) Uninstallation".

Use the following procedure to "Uninstall (middleware) Uninstallation":

1. Start "Uninstall (middleware) Uninstallation" and check that no other Fujitsu middleware products remain. The start method is as follows:

```
# /opt/FJSVcir/cir/bin/cimanager.sh [-c] <Enter>
```

-c: Command interface

#### Note

Do not move to a directory that includes spaces because startup fails if the command path includes 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. If there are no Fujitsu middleware products installed, execute the following uninstall 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.

The uninstall completes within seconds.

4. After the uninstall is completed, the following directories and the files under them remain, so delete them:

- /var/opt/FJSVcir/cir/internal
- /var/opt/FJSVcir/cir/logs
- /var/opt/FJSVcir/cir/meta\_db

### 4.5.6 Restarting the System

After the uninstall is completed, if not logged in to the master server, log in with root permissions and restart the system.

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

## 4.6 Uninstalling from a Slave Server

---

This section describes how to uninstall the slave server functions from a slave server.

The install and setup procedure is shown below.

1. Unsetup network replication and Hadoop
2. Unsetup DFS
3. Uninstall the slave server functions

### 4.6.1 Network Repliation and Hadoop Unsetup

---

The network replication unsetup procedure is shown below.

1. If not logged in to the slave server, log in with root permissions.
2. Execute setup.

Execute bdpp\_unsetup.

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



If unsetup fails, refer to the messages came up during the unsetup 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.

### 4.6.2 DFS Unsetup

---

Refer to "E.1 Deleting a File System" in "Appendix E DFS Operation Changes" for information on DFS unsetup.

### 4.6.3 Uninstalling Slave Server Functions

---

The slave server function 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



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

2. If not logged in to the slave server, log in with root permissions.
3. 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.

## Note

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.
```

4. Reboot in single user mode.

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

5. Execute uninstall.

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

```
# /etc/init.d/network start <Enter>
# /opt/FJSVbdpp/agent/bdpp_agtuninstall <Enter>
```

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

Return to the default run level that was recorded in Step 3.

```
# 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 `uninstallaiton` again.

## Note

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)".

## 4.6.4 Post-Uninstall Tasks

This section describes the slave server post-uninstall tasks.

### 4.6.4.1 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:

- /opt/FJSVbdpp
- /etc/opt/FJSVbdpp
- /etc/pdfs
- /var/opt/FJSVbdpp
- /var/opt/FJSVpdfs

### 4.6.4.2 Uninstall (middleware) Uninstallation

Refer to "[4.5.5.2 Uninstall \(middleware\) Uninstallation](#)" for details.

## 4.6.5 Restarting the System

---

After the uninstall is completed, if not logged in to the master server, log in with root permissions and restart the system.

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

## 4.7 Uninstalling from a Development Server

---

This section describes how to uninstall the development server functions from a development server.

### 4.7.1 Unsetup DFS

---

Refer to "[E.1 Deleting a File System](#)" in "[Appendix E DFS Operation Changes](#)" for information on canceling the DFS setup.

### 4.7.2 Uninstalling Development Server Functions

---

The uninstall procedure is shown below.

1. If UpdateSite format updates have been applied to this product and the following product 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. If not logged in to the development server, log in with root permissions.
3. 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.

## Note

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.
```

4. Reboot in single user mode.

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

5. Execute the uninstall.

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

```
# /etc/init.d/network start <Enter>
# /opt/FJ5Vbdpp/agent/bdpp_devuninstall <Enter>
```

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

Return to the default run level that was recorded in Step 3.

```
# 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 `uninstall` again.

## Note

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)".

## 4.7.3 Post-Uninstall Tasks

This section describes the development server post-uninstall tasks.

### 4.7.3.1 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:

- /opt/FJSVbdpp
- /etc/opt/FJSVbdpp
- /etc/pdfs
- /var/opt/FJSVbdpp
- /var/opt/FJSVpdfs

### 4.7.4 Restarting the System

---

After the uninstall is completed, if not logged in to the master server, log in with root permissions and restart the system.

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

## 4.8 Uninstalling from a Collaboration Server

---

This section describes how to uninstall the collaboration server functions from a collaboration server.

The install and setup procedure is shown below.

1. Unsetup DFS
2. Uninstall

### 4.8.1 DFS Unsetup

---

Refer to "[E.1 Deleting a File System](#)" in "[Appendix E DFS Operation Changes](#)" for information on DFS unsetup.

### 4.8.2 Uninstalling Collaboration Server Functions

---

The uninstall procedure is shown below.

1. If UpdateSite format updates have been applied to this product and the following product included in this product, delete the updates.
  - Primesoft Distributed File System for Hadoop



See

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

2. If not logged in to the collaboration server, log in with root permissions.
3. 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.

## Note

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.
```

4. Reboot in single user mode.

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

5. Execute the uninstall.

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

```
# /etc/init.d/network start <Enter>
# /opt/FJ5Vbdpp/agent/bdpp_coouninstall <Enter>
```

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

Return to the default run level that was recorded in Step 3.

```
# 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 `uninstallaiton` again.

## Note

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)".

## 4.8.3 Post-Uninstall Tasks

This section describes the collaboration server post-uninstall tasks.

### 4.8.3.1 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:

- /opt/FJSVbdpp
- /etc/opt/FJSVbdpp
- /etc/pdfs
- /var/opt/FJSVbdpp
- /var/opt/FJSVpdfs

### 4.8.4 Restarting the System

---

After the uninstall is completed, if not logged in to the master server, log in with root permissions and restart the system.

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

## 4.9 Adding and Deleting Slave Servers

---

This section describes how to add and delete slave servers.

### 4.9.1 Adding Slave Servers

---

If the slave server is added through cloning, the following two patterns exist, depending on the difference in the timing of the installation of the slave server.

1. Adding a slave server during initial installation  
When performing the initial installation, collect the cloning image of the first slave server after installing it, then install all the other slave servers using cloning.
2. Adding slave servers after starting operation  
Additionally installing slave servers after starting operation.

The procedures for these patterns are described in the following pages.

#### 4.9.1.1 Adding a Slave Server During Initial Installation

The procedure for adding a slave server when installing for the first time is shown below.

1. Installing the first slave server  
Refer to "[3.1.2 Installing the First Slave Server](#)" to install the first additional slave server (the slave server whose cloning image is collected).
2. Registering a slave server
3. Registering the network parameter automatic configuration feature
4. Creating a cloning image
5. Cloning

## 6. DFS client addition/scale-out

### 4.9.1.1.1 Registering a Slave Server

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

The procedure for registering a slave server is shown below.

1. Create a cloning server definition file (clone.conf). In the cloning server definition file, specify the slave server whose cloning image is to be collected and all the slave servers to which the cloning image is to be distributed for installation. This is explained using the example shown below.

- Slave server collecting the cloning image :slave1
- Other slave servers to be installed :slave2, slave3, slave4, slave5

Refer to "A.2.3 clone.conf" for information on the cloning server definition file.

#### 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_communi
community_name,ipmi_ip_address,ipmi_user_name,ipmi_passwd,ipmi_passwd_enc,admin_lan1_nic_number,admi
n_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 cloning
image is to be collected
```

2. Log in to the master server (primary) with root permissions.
3. Execute the `bdpp_addserver` command to register the slave server.  
Refer to A.1.1 `bdpp_addserver` for information on the `bdpp_addserver` command.

```
# /opt/FJSVbdpp/bin/bdpp_addserver {clone.conf file storage directory}/clone.conf <Enter>
```

4. Use the `bdpp_listserver` command to check that the slave server has been registered successfully.  
Refer to "A.1.8 `bdpp_listserver`" for information on the `bdpp_listserver` command.

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

#### Example

Example of checking slave server registration:

```
> /opt/FJSVbdpp/bin/bdpp_listimage <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	-

#### 4.9.1.1.2 Registering the Network Parameter Automatic Configuration Feature

Create the public LAN network parameter definition beforehand so that the number of cloning images that are collected meets the whole number of slave servers. This makes it possible to automate the slave server public LAN configuration when the cloning images are distributed.

This section explains how to register the network parameter automatic configuration feature.

The procedure for registering the network parameter automatic configuration feature is shown below.

1. On the master server (primary), create the network parameter automatic configuration feature definition file (FJSVrcx.conf and ipaddr.conf).  
Refer to "[A.2.4 FJSVrcx.conf](#)" for information on the FJSVrcx.conf configuration file.  
The information about the public LAN for the slave server whose cloning image is to be collected and for all other slave servers to which the cloning image is to be distributed is specified in the definition file (ipaddr.conf). Refer to "[A.2.5 ipaddr.conf](#)" for details.
2. On the master server (primary), specify the IP address of the admin LAN of the slave server that will collect the cloning image, the definition file (FJSVrcx.conf) and the definition file (ipaddr.conf), then execute the bdpp\_lanctl command.  
Refer to "[A.1.6 bdpp\\_lanctl](#)" for information on the bdpp\_lanctl command.

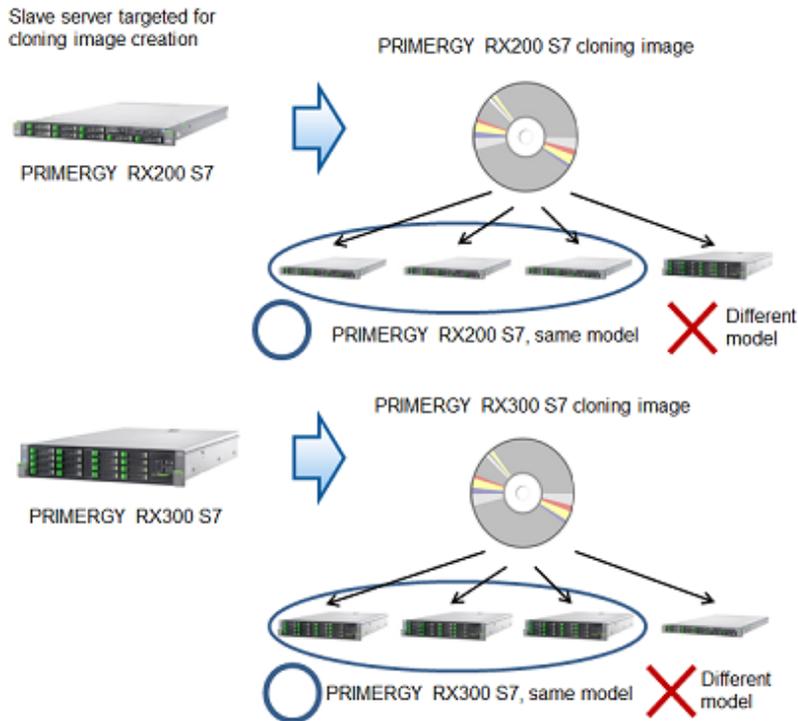
```
# /opt/FJSVbdpp/bin/bdpp_lanctl <admin LAN IP address> <FJSVrcx.conf file> <ipaddr.conf file>
<Enter>
```

#### 4.9.1.1.3 Creating a Cloning Image

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

The server for which the cloning 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 cloning 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).
- BIOS settings must be the same, according to the settings described in "[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 apparatuses. LAN switches and FC Switches that are cascade connected will be treated as one device.



The procedure for creating cloning images is shown below.

1. Check that the slave server whose cloning image is to be collected is properly connected to the master server before creating the cloning image.  
Refer to "3.1.6 Checking the Connection to the First Slave Server" for information about how to check.
2. If the login to the master server (primary) has not been performed yet, log into the server with root permissions.
3. To change the image storage directory, execute the `bdpp_changeimagedir` command.  
Refer to "A.1.2 `bdpp_changeimagedir`" for information on the `bdpp_changeimagedir` command.

```
# /opt/FJSVbdpp/bin/bdpp_changeimagedir {image directory} <Enter>
```

4. Execute the `bdpp_getimage` command that the following were specified in: the name of the slave server for collecting the cloning image and the name of the cloning image.  
Refer to "A.1.5 `bdpp_getimage`" for information on the `bdpp_getimage` command.

```
# /opt/FJSVbdpp/bin/bdpp_getimage {server name} {cloning image name} <Enter>
```

5. Use the `bdpp_listimage` command to check that the cloning image could be collected successfully.  
Refer to "A.1.7 `bdpp_listimage`" for information on the `bdpp_listimage` command.

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

## Example

Checking the cloning image collection 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 -

6. Reboot the slave server for collecting the cloning image.

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

#### 4.9.1.1.4 Cloning

The cloning procedure is shown below.

1. If the login to the master server (primary) has not been performed yet, log into the server with root permissions.
2. Execute the `bdpp_deployimage` command that the following were specified in: the image distribution destination server and the name of the cloning image.  
Refer to "[A.1.4 bdpp\\_deployimage](#)" for information on the `bdpp_deployimage` command.

```
# /opt/FJSVbdpp/bin/bdpp_deployimage {name of server} {name of cloning image} <Enter>
```

3. If the login to the image distribution destination slave server has not been performed yet, log into the server with root permissions.
4. Set the `iscsi` initiator name and IP address of the image distribution destination slave server. Refer to the "ETERNUS DX Disk Storage System User's Guide" for details on how to set these.
5. Reboot the image distribution destination slave server.

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

#### 4.9.1.1.5 Adding a Slave Server Linkage Server/Scale-out

Refer to "[E.3 Adding a Slave Server/Scaling Out](#)" in "[Appendix E DFS Operation Changes](#)" for details on adding a slave server linkage server/scale-out.

#### 4.9.1.2 Adding Slave Servers after Starting Operation

The procedure for adding a slave server after operations have started shown below.

Steps 1 to 6 shown below are the same as the steps 1 to 6 described in "[4.9.1.1 Adding a Slave Server During Initial Installation](#)".

This section explains differences to the structure described in "Adding a Slave Server During Initial Installation", such as additional definition information, based on the following additional structures:

- Slave server installed using the procedure in "Adding a Slave Server During Initial Installation" :slave1, slave2, slave3, slave4, slave5
- Newly installed slave server, with a new cloning image collected :slave6
- Other slave servers to be additionally installed :slave7, slave8, slave9, slave10

1. Slave server installation  
Install using the same procedure as for "Adding a Slave Server During Initial Installation".
2. Registering slave servers  
An example of definitions when adding is described in "[Example of the cloning server definition file for registering a slave server after operations have started](#)".  
The procedure is the same as for "Adding a Slave Server During Initial Installation".
3. Registering the network parameter automatic configuration feature  
The procedure is the same as for "Adding a Slave Server During Initial Installation".

4. Creating a cloning image  
The procedure is the same as for "Adding a Slave Server During Initial Installation".
5. Cloning  
The procedure is the same as for "Adding a Slave Server During Initial Installation".
6. Adding and scaling out the DFS client  
An example of definitions when adding is described in ["Example of the connection approval list file for registering a slave server after operations have started"](#).  
The procedure is the same as for "Adding a Slave Server During Initial Installation".
7. Editing and updating the slave server definition file  
Refer to ["4.9.1.2.1 Editing and Reflecting Slave Server Definition File"](#).

## Example of the cloning server definition file for registering a slave server after operations have started



### 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
-,,,slave1, 10.10.10.13,A0:A0:A0:A0:A0:A1,,public,192.168.1.13,admin,admin,plain,,OFF
-,,,slave2, 10.10.10.14,A0:A0:A0:A0:A0:A2,,public,192.168.1.14,admin,admin,plain,,OFF
-,,,slave3, 10.10.10.15,A0:A0:A0:A0:A0:A3,,public,192.168.1.15,admin,admin,plain,,OFF
-,,,slave4, 10.10.10.16,A0:A0:A0:A0:A0:A4,,public,192.168.1.16,admin,admin,plain,,OFF
-,,,slave5, 10.10.10.17,A0:A0:A0:A0:A0:A5,,public,192.168.1.17,admin,admin,plain,,OFF
new,,,slave6, 10.10.10.18,A0:A0:A0:A0:A0:A6,,public,192.168.1.18,admin,admin,plain,,OFF <-- Server
to be registered
new,,,slave7, 10.10.10.19,A0:A0:A0:A0:A0:A7,,public,192.168.1.19,admin,admin,plain,,OFF <-- Server
to be registered
new,,,slave8, 10.10.10.20,A0:A0:A0:A0:A0:A8,,public,192.168.1.20,admin,admin,plain,,OFF <-- Server
to be registered
new,,,slave9, 10.10.10.21,A0:A0:A0:A0:A0:A9,,public,192.168.1.21,admin,admin,plain,,OFF <-- Server
to be registered
new,,,slave10, 10.10.10.22,A0:A0:A0:A0:A0:B0,,public,192.168.1.22,admin,admin,plain,,OFF <-- Server
to be registered
[ServerAgent]
operation,server_name
new,slave6 <-- Server whose cloning image
is to be collected

```

Example of checking slave server registration:

```

> /opt/FJSVbdpp/bin/bdpp_listserver <Enter>
PHYSICAL_SERVER  SERVER                ADMIN_IP              STATUS  MAINTENANCE
-----
slave1            slave1                192.168.1.13         normal  OFF
slave2            slave2                192.168.1.14         normal  OFF
slave3            slave3                192.168.1.15         normal  OFF
slave4            slave4                192.168.1.16         normal  OFF
slave5            slave5                192.168.1.17         normal  OFF
slave6            slave6                192.168.1.18         normal  OFF
slave7            -                    192.168.1.19         stop    -
slave8            -                    192.168.1.20         stop    -
slave9            -                    192.168.1.21         stop    -
slave10           -                    192.168.1.22         stop    -

```

## Example of the connection approval list file for registering a slave server after operations have started

### Example

```
# 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 slave1
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
```

### 4.9.1.2.1 Editing and Reflecting Slave Server Definition File

The procedure for editing the slave server definition file and then effecting these changes is explained below.

1. If the login to the master server (primary) has not been performed yet, log into the server with root permissions.
2. Stop the Hadoop of the Interstage Big Data Parallel Processing Server if it is running before adding slave servers.  
Use the `bdpp_stop` command to stop the Hadoop of the Interstage Big Data Parallel Processing Server.  
Refer to "[A.1.13 bdpp\\_stop](#)" for details of the `bdpp_stop` command.

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

3. Define the slave server to be added by editing the slave server definition file (`/etc/opt/FJSVbdpp/conf/slaves`) and the TaskTracker definition file (`/etc/hadoop/mapred.include`).  
Refer to "[A.2.2 slaves](#)" for details on the slave server definition file.

### Example

Slave server definition file example:

```
slave1,slave2,slave3,slave4,slave5,slave6,slave7,slave8,slave9,slave10
```

Example TaskTracker definition file:

```
slave1
slave2
slave3
slave4
slave5
slave6
slave7
slave8
slave9
slave10
```

Note: The underlining indicates the added slave server.

- .....
4. Execute `bdpp_changeslaves` command to effect the changes that were made to the slave server definition file. Refer to "[A.1.3 bdpp\\_changeslaves](#)" for information on the `bdpp_changeslaves` command.

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

5. Copy the edited file to the master server (secondary), slave server and development 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>
Also copy the edited file to the slave server and development server.
```

6. Start the Hadoop of the Interstage Big Data Parallel Processing Server. Use the `bdpp_start` command to start the Hadoop of the Interstage Big Data Parallel Processing Server. Refer to "[A.1.11 bdpp\\_start](#)" for details of the `bdpp_start` command.

```
# /opt/FJSVbdpp/bin/bdpp_start
```

7. Use the Interstage Big Data Parallel Processing Server status display to check whether or not TaskTracker is started at the targeted slave server. After checking that the targeted slave server is started, use the `bdpp_stat` command to display the Hadoop status of the Interstage Big Data Parallel Processing Server. Refer to "[A.1.12 bdpp\\_stat](#)" for details of the `bdpp_stat` command.

```
# /opt/FJSVbdpp/bin/bdpp_stat -all
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.
```

## 4.9.2 Deleting Slave Servers

---

The procedure for deleting slave servers is shown below.

1. Edit and reflect slave server definition file
2. Delete DFS clients

## 4.9.2.1 Editing and Reflecting a Slave Server Definition File

The procedure for editing the slave server definition file and then effecting these changes is explained below.

1. If the login to the master server (primary) has not been performed yet, log in to the server with root permissions.
2. Stop the Hadoop of the Interstage Big Data Parallel Processing Server if it is running before deleting slave servers. Use the "bdpp\_stop" command to stop the Hadoop of the Interstage Big Data Parallel Processing Server. Refer to "[A.1.13 bdpp\\_stop](#)" for details of the bdpp\_stop command.

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

3. Delete the slave server to be removed from the slave server definition file (/etc/opt/FJSVbdpp/conf/slaves) and the TaskTracker definition file (/etc/hadoop/mapred.include). Refer to "[A.2.2 slaves](#)" for details on the slave server definition file.
4. Execute bdpp\_changeslaves command to effect the changes that were made to the slave server definition file. Refer to "[A.1.3 bdpp\\_changeslaves](#)" for information on the bdpp\_changeslaves command.

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

5. Copy the edited file to the master server (secondary), slave server and development 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>
  Also copy the edited file to the slave server and development server.
```

6. Start the Hadoop of the Interstage Big Data Parallel Processing Server. Use the bdpp\_start command to start the Hadoop of the Interstage Big Data Parallel Processing Server. Refer to "[A.1.11 bdpp\\_start](#)" for details of the bdpp\_start command.

```
# /opt/FJSVbdpp/bin/bdpp_start
```

## 4.9.2.2 Deleting a DFS client

Refer to "[E.4 Deleting a Slave Server](#)" in "[Appendix E DFS Operation Changes](#)" for information on deleting a DFS client.

# Chapter 5 Operation

This chapter describes operation of this product.

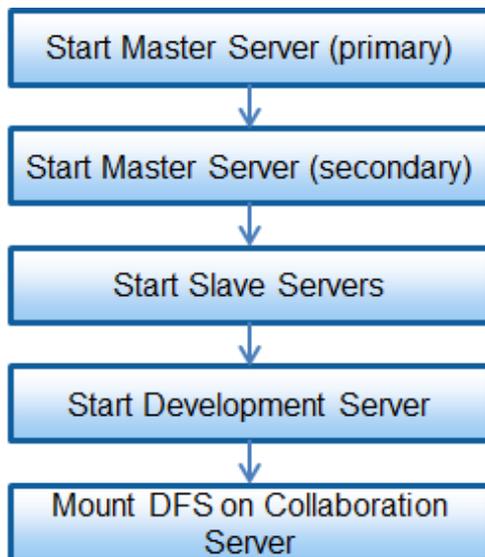
Note that, in addition to the Apache Hadoop operation specifications, the following operation features are also provided for operation of this product:

- Master server and slave server start/stop
- Master server repliation

## 5.1 Starting and Stopping the Interstage Big Data Parallel Processing Server

### 5.1.1 Starting the Interstage Big Data Parallel Processing Server

Begin starting the servers with the master server (primary) and then the master server (secondary), then start the slave servers and the development server. With the collaboration server, the DFS can be mounted after the master server has completed startup.



The Hadoop of the Interstage Big Data Parallel Processing Server can be started once the master server and all of the slave servers have been started.

Use the `bdpp_start` command to start the Hadoop of the Interstage Big Data Parallel Processing Server.

Refer to "[A.1.11 bdpp\\_start](#)" for details of the `bdpp_start` command.

#### Note

The master server uses an HA cluster configuration, so the Apache Hadoop "JobTracker" on the master server cannot be started or stopped directly.

To start or stop the Hadoop of the Interstage Big Data Parallel Processing Server, the `bdpp_start` or `bdpp_stop` commands must be used.

Refer to the Hadoop Project Top Page (<http://hadoop.apache.org/>) for details on how to use Apache Hadoop.

When Interstage Big Data Parallel Processing Server is started, the following Hadoop processes start:

- Master server
  - JobTracker
- Slave server
  - TaskTracker

## 5.1.2 Stopping the Interstage Big Data Parallel Processing Server

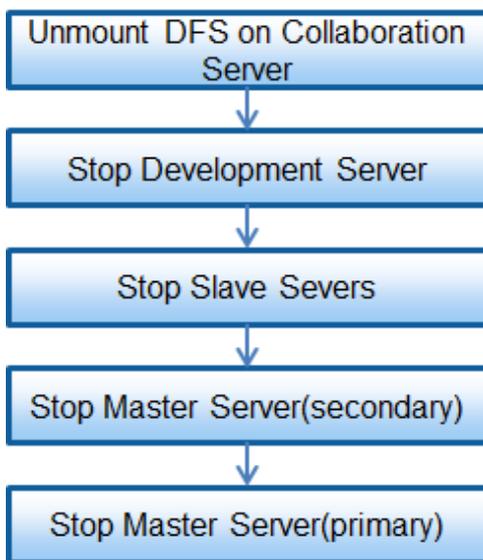
---

Use the `bdpp_stop` command to stop Interstage Big Data Parallel Processing Server.

Execution of this command stops the Hadoop processes operating at the master server and slave servers.

Refer to "A.1.13 `bdpp_stop`" for details of the `bdpp_stop` command.

To stop the servers, first unmount the DFS on the collaboration server, then stop the development server, slave servers, master server (secondary) and master server (primary) in that order.



### Note

The master server uses an HA cluster configuration, so the Apache Hadoop "JobTracker" on the master server cannot be started or stopped directly. To start or stop the Hadoop of the Interstage Big Data Parallel Processing Server, the `bdpp_start` or `bdpp_stop` commands must be used.

Refer to the Hadoop Project Top Page (<http://hadoop.apache.org/>) for details on how to use Apache Hadoop, for example.

## 5.1.3 Displaying the Interstage Big Data Parallel Processing Server Status

---

Use the `bdpp_stat` command to display the Interstage Big Data Parallel Processing Server status.

The status of Hadoop processes operating at the master server can be displayed by executing this command.

Refer to "A.1.12 `bdpp_stat`" for details of the `bdpp_stat` command.

## 5.2 Developing and Registering Applications

---

This section describes how to develop the applications executed by Hadoop.

### 5.2.1 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.

### 5.2.2 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.

#### 5.2.2.1 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, the Hadoop API is not used. These scripts are written using the Pig Latin language, which has functions equivalent to SQL language functions.

- 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.

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.

#### 5.2.2.2 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.

For MapReduce applications, 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

### 5.2.2.3 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 box.

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.

### 5.2.2.4 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/r1.0.1/mapred\\_tutorial.html](http://hadoop.apache.org/common/docs/r1.0.1/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.

## 5.2.3 Registering Applications

---

Store the created Java program (jar file) in any directory on the development server.

## 5.3 Executing and Stopping 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 of 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 the `bdpp.conf` file during installation. Refer to "[D.6 User-Specific Hadoop Usage Settings](#)" for information on user account settings.

## 5.3.1 Executing Jobs

---

Using an account that can execute Hadoop jobs, use the `hadoop` command to execute MapReduce programs (Java programs). Use either the `jar` option or the `job` option.

### Example

---

Command execution example:

```
$ hadoop jar sample.jar input output
```

---

## 5.3.2 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 <job ID>
```

---

## 5.3.3 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 the job ID, status, start time, and so on.

### Example

---

Command execution example:

```
$ hadoop job -list all <Enter>
```

---

## 5.4 Operations When There are Errors

---

### 5.4.1 Master Server Error Operation

---

This section explains the operations below when a master server error has occurred.

- Errors while the system is running
- Errors when the system starts up

### 5.4.1.1 Errors while the System is Running

When the events shown below occur on the master server (primary) while the system is operating, automatic switching to the master server (secondary) occurs.

- The system stops due to the cause of such as system panic, forced power off.
- The system has hung up
- An error has occurred in the cluster interconnection  
Error occurred on both cluster interconnect network in case redundant cluster interconnect is configured.
- An error has occurred in the public LAN network  
Error occurred on both public LAN network in case redundant public LAN is configured.
- An error has occurred in the iSCSI network  
Error occurred on both iSCSI LAN network in case redundant iSCSI LAN is configured.
- If JobTracker has ended abnormally, or been stopped using a method other than the `bdpp_stop` command (\*1)

\*1: JobTracker was stopped using the Apache Hadoop feature directly



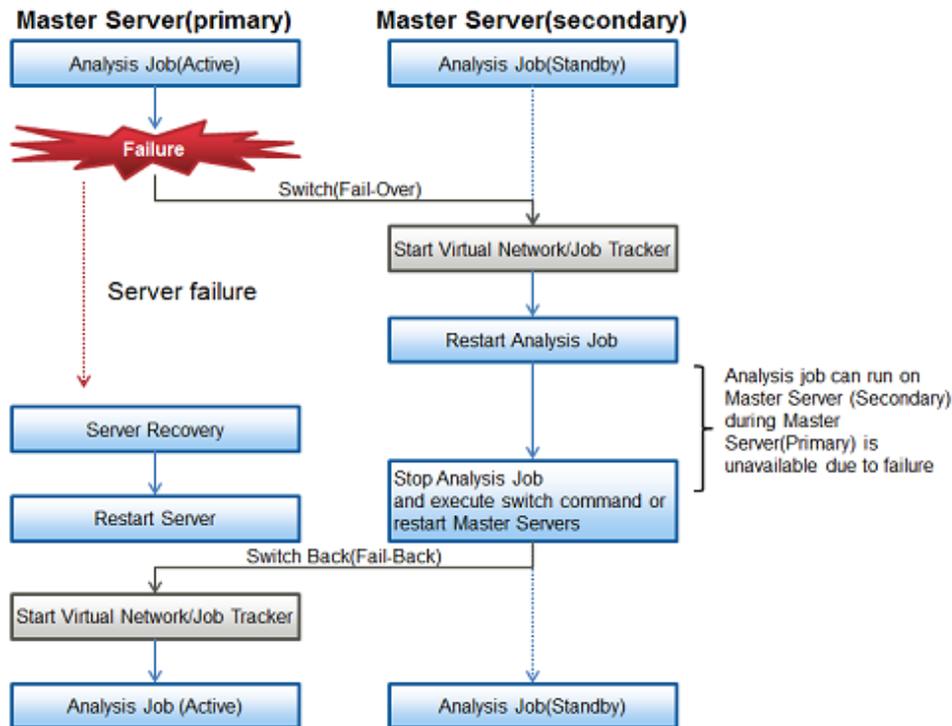
If an error occurs in the public LAN network or the iSCSI network

If an error occurs in the public LAN network or the iSCSI network, and there is a switch from the master server (primary) to the master server (secondary), refer to "[5.4.1.3 If an Error Occurs in the Network](#)" and perform the necessary operations.

If the switch from the master server occurs while a job is running, the execution of that job is suspended. Check the execution of the job status, and if necessary re-execute the job on the server that the job is switched to (master server (secondary)).

If an error occurs in the master server (primary) and there is a switch to the master server (secondary), it is possible to switch back to the master server (primary) after the cause of the error has been removed. Restart the system by executing the command to switch back on the master server (secondary) or switch back to the master server (primary).

Figure 5.1 Flow of operations between the occurrence of an error to the resumption of operations



To switch back from the master server (secondary) to the master server (primary), use the hvswitch command.

```
# hvswitch appl <Enter>
```

"appl" is a fixed string.

The switch back takes about 30 seconds, so execute the bdpp\_stat command on the master server (primary) and see that Hadoop is running properly on the master server (primary).

Refer to "hvswitch" in "7.2.2 Switching an application" of the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide" and Online help (man hvswitch) for information on the hvswitch bdpp\_stat commands.

Refer to "A.1.12 bdpp\_stat" for information on the bdpp\_stat commands.



### Note

The master server feature that is used for the execution of "4.9.1 Adding Slave Servers" is not a target of switching.

Perform "4.9.1 Adding Slave Servers" again after removing the cause of the problem from the master server (primary) if an error has occurred on it, and there was a switch to the master server (secondary).

## 5.4.1.2 Errors when the System Starts up

If the events shown below occur in the master server when the system starts up, the normal start operation cannot be performed.

This section explains the operations that are used to start the system on the other server.

- Starting the system when the master server (primary) is not running

Hadoop can be started with the bdpp\_start command if the system has been running with the master server (secondary) because an error rendered the master server (primary) inoperable.

Refer to "A.1.11 bdpp\_start" for details on starting Hadoop.

- Starting the system when the master server (secondary) is not running

Hadoop can be started with the `bdpp_start` command if the system has been running with the master server (primary) because an error rendered the master server (secondary) inoperable.

Refer to "[A.1.11 bdpp\\_start](#)" for details on starting Hadoop.

### 5.4.1.3 If an Error Occurs in the Network

The operations described below are required if there has been an error on the network causing a switch from the master server (primary) to the master server (secondary).

1. Log in to the server where the error occurred with root permissions.
2. Check DFS status.

```
# 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  <-- Master server (primary)
                                     is in run state
  1  AC      run    -         0       0       0  master1
  1  MDS(S)  wait   -         0       0       0  master2
  1  AC      run    -         0       0       0  master2
```

3. On the server where the error occurred, unmount the DFS, and stop or restart the system.

```
# pdfsumount /mnt/pdfs <Enter> or;
# shutdown -h now <Enter> or;
# shutdown -r now <Enter>
```

4. Next log in to the switched server with root permissions.
5. Unmount the DFS.

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

6. Mount the DFS.

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

7. Check that the status of the DFS of the switched server is "run" using the `pdfsrscinfo` command.

```
# pdfsrscinfo -m <Enter>
/dev/disk/by-id/scsi-36000c298b3c931387b26aaa0a9ee314f-part2:
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  <-- Master server (secondary) is
in run state
  1  AC      run    -         0       0       0  master2
```

8. Log into the servers connected to the master server (the slave servers, development server and collaboration server) using root permissions, unmount the DFS, then mount it again.

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

This completes the operation. After this it is possible to restart operations.

9. When the server where the error occurred is restored and there is a switch back to the original server, complete steps 1 to 7, replacing "server where the error occurred" with "switched server" and "switched server" with "server after restoration".
10. Switch the server back using the hvswitch command.

```
# hvswitch appl <Enter>
```

"appl" is a fixed string.



Refer to the "Appendix A Command Reference" of the "Primesoft Distributed File System for Hadoop V1 User's Guide" for information on the pdfsrscinfo, pdfsumount and pdfsmount commands.

Refer to "hvswitch" in "7.2.2 Switching an application" of the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide" for information on the hvswitch command.

## 5.4.2 Slave Server Error Operation

---

This section explains operation when a slave server error has occurred.

When the events shown below occur at a slave server, other slave servers take over the job currently being executed, enabling processing to continue.

- If the slave server stopped
  - If slave server shutdown, system panic, forced power cut, or similar occurs
- If the slave server does not respond, for example because it hangs up
- If an error occurs in the slave server network
  - If an error occurs in both networks, if the public LAN is duplicated

If a slave server error occurs, check the slave server status with information such as system log file. Take action, and then execute the bdpp\_start command on the master server to restart Hadoop on the slave server where the error occurred.

Refer to "A.1.11 bdpp\_start" for information on starting Hadoop.



If the bdpp\_start command is executed to restart Hadoop on some slave servers, the "bdpp:WARN:001" message is output, but this does not indicate a problem with restarting the Hadoop on the slave servers.

## 5.4.3 How to Check Errors

---

Check the messages output to the system log (/var/log/messages) if there is an error in the master server (primary) that causes a switch to the master server (secondary) or if an error in one side of a duplicated 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 master server (primary) and the master server (secondary).

Refer to "hvdisp" in "5.3.5 Displaying object states" of the "PRIMECLUSTER Reliant Monitor Services (RMS) with Wizard Tools Configuration and Administration Guide" and Online help (`man hvdisp`) for information on the `hvdisp` command.

The following example shows the status in the master server (secondary) after switching due to an error in the master server (primary).

Display the status of the HA cluster of the master server (primary) (example)

```
# hvdisp -a

Local System:  master1RMS
Configuration: /opt/SMAW/SMAWRrms/build/bdpp.us

Resource          Type      HostName          State      StateDetails
-----
master1RMS        SysNode                   Online
master2RMS        SysNode                   Online
appl              userApp                   Faulted    Failed Over
Machine001_appl  andOp    master1RMS
Machine000_appl  andOp    master2RMS        Offline
ManageProgram000_Cmd_APP1 gRes                   Offline
Ippaddress000_Gls_APP1 gRes                   Offline
```

Display the status of the HA cluster of the master server (secondary) (example)

```
# hvdisp -a

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    kawa4RMS          Online
Machine001_appl  andOp    kawa5RMS          Online
Machine000_appl  andOp    kawa4RMS
ManageProgram000_Cmd_APP1 gRes                   Online
Ippaddress000_Gls_APP1 gRes                   Online
```

### Checking Network Repliation 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" in "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 the network repliation of the master server (primary) (example)

```
# /opt/FJSVhanet/usr/sbin/dsphanet
[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 the network replication of the slave server (example)

```
# /opt/FJSVhanet/usr/sbin/dsphanet
[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
+-----+-----+-----+-----+
```

## 5.5 Adding and Deleting Storage Systems

---

This section explains how to add and delete storage systems.

### 5.5.1 Adding a Storage System

---

Refer to "[E.5 Adding a Shared Disk/Scaling Out](#)" in "[Appendix E DFS Operation Changes](#)" for information on adding shared disks and scaling out.

### 5.5.2 Deleting a Storage System

---

Refer to "[E.6 Deleting a Shared Disk](#)" in "[Appendix E DFS Operation Changes](#)" for information on deleting shared disks.

# Chapter 6 Maintenance

This chapter describes the maintenance method for this product.

## 6.1 Maintenance Method

This section describes the application of product updates, and investigations when problems arise.

### 6.1.1 Applying Updates

1. Contact Fujitsu Technical staff for obtaining any necessary updates(patches) for this product including patches for following components:

- ServerView Resource Orchestrator Virtual Edition V3.0.0
- PRIMECLUSTER 4.3A10
- Primesoft Distributed File System for Hadoop V1.0

Note: Other relevant products are also bundled with this product, so application of updates for those products also is recommended.

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 June 2012)

### 6.1.2 Collecting Trouble shooting data

This section describes how to collect the trouble shooting data that Fujitsu technical staff need to determine the causes of problems occurring at systems using this product.

This product has the following trouble shooting data:

- Interstage Big Data Parallel Processing Server logs  
Refer to "[6.1.2.1 Interstage Big Data Parallel Processing Server logs](#)".
- trouble shooting data collection tools  
Refer to "[6.1.2.2 Trouble shooting data Collection Tools](#)".

For trouble shooting data for problems related to the "[B.3.2 Other Messages](#)" in the messages output in Interstage Big Data Parallel Processing Server logs and the messages output in system logs ("/var/log/messages"), the trouble shooting data for specific functions incorporated in this product must also be collected.

Refer to "[6.1.2.3 Function-specific trouble shooting data](#)" for information on how to collect function-specific data.

#### 6.1.2.1 Interstage Big Data Parallel Processing Server logs

This section describes the Interstage Big Data Parallel Processing Server logs.

Interstage Big Data Parallel Processing Server 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 information on the logs specific to Apache Hadoop.

**Table 6.1 Log types and storage locations**

Log type	Log file and storage location
Install/uninstall logs	/var/tmp/bdpp_install.log
Setup/unsetup logs	/var/opt/FJSVbdpp/log/bdpp_setup.log
Interstage Big Data Parallel Processing Server 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

### 6.1.2.2 Trouble shooting data Collection Tools

This section describes the Interstage Big Data Parallel Processing Server trouble shooting data collection tools.

Execute the tools shown in the following table and collect the output results as trouble shooting data.

**Table 6.2 Trouble shooting data collection tool types and location where the data is stored**

Trouble shooting data collection tool type	Trouble shooting 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

### 6.1.2.3 Function-specific trouble shooting data

trouble shooting data must be collected for each of the functions incorporated in this product.

#### 6.1.2.3.1 If HA cluster setup or switching problems occurred

This section explains how to collect trouble shooting data when there is a problem setting up an HA cluster or when switching.

#### How to collect trouble shooting data

Collect the following information required for investigating a fault in the HA cluster system from the master server (primary) and the master server (secondary).

1. HA cluster trouble shooting data
  - Use fjsnap to collect the information necessary to investigate the fault. Refer to "[Executing the fjsnap command](#)".
  - Collect the trouble shooting data for the system.

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

Once 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. Manual for reproducing faults if fault reproduction is possible

## Information

When reporting fault information to Fujitsu technical support, it is necessary for the information necessary to investigate 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 master server (primary) and the master server (secondary). The information to be collected by fjsnap, 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 master server (primary) and the master server (secondary) with 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 as soon after the problem occurs as possible.

Collect the crash dump if the fjsnap command cannot be executed due to a system hang-up. After that, start in single-user mode and execute the fjsnap command. Refer to "[Crash dump](#)" for details about 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 multi-user mode, execute the fjsnap command.

Collect the crash dump when trouble shooting 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 master server (primary) and the master server (secondary) with 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 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 (/dev/st0, etc.) 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 as soon after the problem occurs as possible.

Collect the crash dump if the pclsnap command cannot be executed due to a system hang-up. After that, start in the single user mode and execute the pclsnap command. Refer to "[Crash dump](#)" for details about 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 multi-user mode, execute the pclsnap command.

Collect the crash dump when trouble shooting 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 (guide) (MB)
Output directory	Current directory at the time of execution	300
Temporary directory	/tmp	500



## Note

The guide values given above (300MB and 500MB) may be insufficient in some systems.

If information collection could not be completed successfully due to an insufficiency in the capacity of the directory, the pclsnap command outputs an error or warning message when it ends. In this case, take the action shown below and then rerun 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 failed 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 available space, then re-execute the command.

#### Example:

When /var/crash is made 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 available space, 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) see 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/ (SMAWsf/log/rcsd.log, etc.)

(b) Total size of the core file

- GFS core file: /var/opt/FJSVsfdfs/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 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 trouble shooting data.

### Timing for collecting the crash dump

- If an Oops occurs in the kernel
- If a panic occurs in the kernel
- If <Alt>+<SysRq>+<C> key was 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 there are crash dumps from after the time that the switch occurred, collect the crash dumps. If there are no crash dumps from after the time that 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> key 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.

## 6.1.2.3.2 If cloning image creation or cloning problems occur when adding or deleting slave servers

The following describes how to collect trouble shooting data if cloning image creation or cloning problems occur when adding or deleting slave servers.

Refer to "Chapter 15 Troubleshooting" in the "ServerView Resource Orchestrator Virtual Edition V3.0.0 Operation Guide" for details.

### Types of trouble shooting data

Collect the trouble shooting 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 trouble shooting data. Collect the data that is required for the purposes described below.

#### 1. Collecting initial trouble shooting data

Collect 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 "[Collecting initial trouble shooting data](#)" for details.

#### 2. Collecting detailed trouble shooting data

It is sometimes possible to determine the cause using just the initial trouble shooting data, but more trouble shooting data may be required for some problems.

It is then necessary to collect more detailed trouble shooting data. Detailed investigation involves collecting a 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 trouble shooting data collected to triage the problem.

If requested by Fujitsu technical support, send the detailed trouble shooting data that has been collected.

Refer to "[Collecting detailed trouble shooting data](#)" for details.



### Note

Collect the trouble shooting data quickly when a problem occurs. The information required to investigate a problem disappears as time passes.

### Collecting initial trouble shooting data

This section describes how to collect trouble shooting data needed to triage the cause of a problem.

#### How to collect trouble shooting data

Collect the initial trouble shooting data using the following procedure.

Collect the resources using the appropriate method, depending on the characteristics of the collection method and the environment and system where the problem occurred.

- Collecting resources from the master server

Collect trouble shooting data from the master server (rcxadm mgrctl snap -all)

trouble shooting data from the 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 65MB of available space required to execute the rcxadm mgrctl snap -all command, approximately 30MB of space is required for each server.

- Collecting resources from the servers

Collect trouble shooting data from the servers (rcxadm mgrctl snap, rcxadm agtctl snap)

Refer to "[Collecting resources from the servers \(rcxadm mgrctl snap, rcxadm agtctl snap\)](#)" and collect the information.

65MB of available space is required to execute the rcxadm mgrctl snap command.

30MB 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 trouble shooting data (rcxadm mgrctl snap -all) on the admin server, the trouble shooting data for the managed servers is collected in a batch.

The following describes collecting trouble shooting data with the command (rcxadm mgrctl snap -all).

#### How to collect

Use the following procedure to collect the resources on the admin server.

1. Log in to the admin server as a user with OS administrator privileges.
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 admin server, the manager must be operating on the admin server. If the manager is not operating, collect the resources on the individual servers.
- The trouble shooting data cannot be collected from the managed servers in the following cases.
  - When a communications route has not been established
  - If there is a managed server that is stopped

In either case, collection of trouble shooting data on the other managed servers continues uninterrupted.

Check the command execution results in the execution log.

Refer to "rcxadm mgrctl" in the "ServerView Resource Orchestrator Virtual Edition V3.0.0 Command Reference" for details.

When collection has failed on a managed server, either execute the rcxadm mgrctl snap -all on the admin server again or execute the rcxadm agtctl snap command on the managed server.

### Collecting resources from the servers (rcxadm mgrctl snap, rcxadm agtctl snap)

Apart from the rcxadm mgrctl snap -all command that is executed on the admin server and that can collect the trouble shooting data in a batch from the 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 trouble shooting data with the command (rcxadm mgrctl snap or rcxadm agtctl snap).

#### How to collect

Use the following procedure to collect the resources on the servers.

1. Log into the server with OS administrator privileges.
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 Virtual Edition V3.0.0 Command Reference" for details.

### Collecting detailed trouble shooting data

This section describes how to collect detailed trouble shooting data needed to determine the cause of a problem.

When the cause of the problem cannot be determined just from the initial trouble shooting data, more detailed trouble shooting data is required.

#### How to collect trouble shooting data

The trouble shooting data required to determine the cause of a problem is collected by executing the trouble shooting data collection commands (`rcxadm mgrctl snap -full` and `rcxadm agtctl snap -full`) on the servers.

80MB of available space is required to execute this feature.

#### How to collect

On the server where resources are to be collected, use the following procedure to collect the resources.

1. Log into the server with OS administrator privileges.
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 Virtual Edition V3.0.0 Command Reference" for details.

### 6.1.2.3.3 If DFS setup or shared disk problems occurred

This section explains how to collect trouble shooting 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.

#### Collecting DFS trouble shooting 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 the resources in the 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 and fjsnap)
2. Crash dump
3. Execution results of the pdfsck command
4. Collecting the core image for the daemon

When it is necessary to send the trouble shooting data quickly, collect the following as initial trouble shooting data.

- a. Output results of the resource collection tool (pdfssnap)
- b. /var/log/messages\*

After collecting the resources for initial investigation, make sure the other resources are also collected.

#### Output results of the resource collection tool (pdfssnap and fjsnap)

Use pdfssnap.sh and the fjsnap command to collect the trouble shooting data.

Collect from all servers that shared the DFS, as far as possible.

##### Executing pdfssnap.sh

```
# /etc/opt/FJSVpdfs/bin/pdfssnap.sh <Enter>
```

#### Note

With pdfssnap.sh, the trouble shooting data is output to the directory where the command is executed. For this reason, at least 100MB of free space must be available in the file system that will execute the command.

##### Executing fjsnap

```
# /opt/FJSVsnap/bin/fjsnap -a Any file name <Enter>
```

#### Collecting the crash dump

If there was a panic on the server, for example, also collect the crash dump file as part of the trouble shooting data.

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

Collect if there is a mismatch in the DFS and it needs to be restored.

```
# pdfsck -N -o nolog Block-particular file for the representative partition <Enter>
```

#### 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 the various daemons.

Collect core images on all of the DFS admin servers.

The procedure is explained below using the example of collecting the core image of the pdfsfrmd daemon.

## 1. Determining process ID

Identify the process ID 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 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.

## 6.1.2.3.4 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 master server (primary), master server (secondary), slave server and development server, separated by commas. Do not use spaces.

### Privilege Required/Execution Environment

#### Privileges

Operating system administrator

#### Execution environment

Master server



## Example

```
# /opt/FJSVbdpp/products/HADOOP/bin/HADOOP-collect.sh --servers=master1,master2,slave1,slave2,slave3,slave4,slave5,develop  
<Enter>
```



## Note

There will be a request for a password during execution if settings have not been made to allow SSH connection with the root user of the source of the information with the destination server for the information.

To reduce the workload by avoiding password input, distribute the public key of the SSH of the root user of the master server to the slave server and development server so that SSH without a password can be performed. Refer to the help for the ssh-keygen command for details.

# Appendix A Reference

This appendix explains the commands and definition files of this product.

## A.1 Commands

This section gives an overview of the commands provided by this product.

Command list

Command name	Function	Execution environment				
		Master server		Slave server	Development Server	Collaboration Server
		Primary	Secondary			
<a href="#">bdpp_addserver</a>	Registers the slave server that collects the cloning image and the slave server to which this cloning image is deployed.	Yes	No	No	No	No
<a href="#">bdpp_changeimagedir</a>	Changes the cloning image storage directory.	Yes	No	No	No	No
<a href="#">bdpp_changeslaves</a>	Effects the slave server information that was changed in the slave server definition file (slaves).	Yes	Yes(*1)	No	No	No
<a href="#">bdpp_deployimage</a>	Deploys the slave server system image that was collected using the <a href="#">bdpp_getimage</a> command to the other server.	Yes	No	No	No	No
<a href="#">bdpp_getimage</a>	Takes a system image to use for slave server cloning.	Yes	No	No	No	No
<a href="#">bdpp_lanctl</a>	Enables the automatic configuration of the network parameter.	Yes	No	No	No	No
<a href="#">bdpp_listimage</a>	Checks the registration of a cloning image	Yes	No	No	No	No
<a href="#">bdpp_listserver</a>	Displays the registration status of the slave server that was registered using the <a href="#">bdpp_addserver</a> command.	Yes	No	No	No	No
<a href="#">bdpp_removeimage</a>	Deletes a cloning images that was taken using the <a href="#">bdpp_getimage</a> command.	Yes	No	No	No	No
<a href="#">bdpp_removeserver</a>	Deletes a slave server that was added by the <a href="#">bdpp_addserver</a> command.	Yes	No	No	No	No
<a href="#">bdpp_start</a>	Starts Hadoop	Yes	Yes(*1)	No	No	No
<a href="#">bdpp_stat</a>	Checks Hadoop start and stop statuses	Yes	Yes(*1)	No	No	No
<a href="#">bdpp_stop</a>	Stops Hadoop	Yes	Yes(*1)	No	No	No

Yes: Can be used

No: Cannot be used

\*1: Command that becomes available when a failover has occurred and operation has been switched to the master server (secondary). Do not use this command when operating on the master server (primary).

## Point

---

Commands are positioned under the following directory:

- /opt/FJSVbdpp/bin

---

### A.1.1 bdpp\_addserver

---

#### Name

/opt/FJSVbdpp/bin/bdpp\_addserver - Registers the slave server that collects the cloning image and the slave server that this cloning image is deployed to

#### Format

bdpp\_addserver *clone.confFileStorageDirectory*

#### Function description

This command registers the slave server that collects the cloning image and the slave server that this cloning image is deployed to.

#### Arguments

*clone.confFileStorageDirectory*

Specify the name of the directory that the clone.conf file is stored in, or the full path name of the clone.conf file.

#### Required permissions and execution environment

##### Permissions

Operating system administrator

##### Execution environment

Master server

#### End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

### A.1.2 bdpp\_changeimagedir

---

#### Name

/opt/FJSVbdpp/bin/bdpp\_changeimagedir - Changes the cloning image storage directory

#### Format

bdpp\_changeimagedir *cloningImageStorageDirectory*

#### Function description

This command changes the cloning image storage directory.



When this command is executed, the cloning image file is copied. This may take a while.

## Arguments

### *CloningImageStorageDirectory*

Specify the name of the change destination cloning image storage directory.

The directory that is specified must satisfy the following conditions:

- The path including the cloning image file storage folder must not exceed 100 characters.
- The directory name must not contain the following symbols:  
", |, :, \*, ?, ., <>, ,, %, &, ^, =, !, ;, #, ', +, [, ], {, }, \
- The destination that is specified must not be under the following directories:
  - /opt/FJSVrcvmr
  - /etc/opt/FJSVrcvmr
  - /var/opt/FJSVrcvmr
- The directory that is specified must be empty.  
If a new file system is created as the cloning image file storage destination, a directory with the name "lost+found" is created, therefore the directory is not empty.
- Create a directory under the file system that was created, then specify that directory.
- There must not be a problem with the permissions.

For reasons of security, set one of the following as the access permissions for all new directories of the cloning image file storage directory:

- Set write permission only for the system administrator.
- Set a sticky bit so that other users cannot change file names or delete files.

If neither of these settings are applied, processing to change the cloning image file storage folder may fail.

## Required permissions and execution environment

### Permissions

Operating system administrator

### Execution environment

Master server

## End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

## A.1.3 bdpp\_changeslaves

---

### Name

/opt/FJSVbdpp/bin/bdpp\_changeslaves - Effect changes made to slave server definition file

#### Format

bdpp\_changeslaves

#### Function description

This command effects the slave server information that was changed in the slave server definition file (slaves).

#### Options

None

#### Required permissions and execution environment

##### Permissions

Operating system administrator

##### Execution environment

Master server

#### End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred



#### Note

Do not execute more than one of these commands at the same time.

## A.1.4 bdpp\_deployimage

---

#### Name

/opt/FJSVbdpp/bin/bdpp\_deployimage - Deploy cloning image

#### Format

bdpp\_deployimage *servername* [[:*hostname*],...] *imagename*

#### Function description

This command deploys a server system image to other servers.



#### Note

- When this command is executed, the cloning image file is distributed to the server. This may take a while.
- Four cloning image collection and deployment processes may be executed simultaneously. When five or more are requested, it will be on standby until the process being executed is complete.

#### Arguments

*servername* [[:*hostname*],...]

Specify the name of the physical server that the cloning image is distributed to.

Multiple physical servers can be specified by using a comma (,) as a delimiter.

The name of the server after distribution can be specified in `hostname`. Add a colon (:) after each physical server name to specify the name of the server after distribution in `hostname`.

#### *imagename*

Specify the name of the cloning image that is to be distributed.

#### Required permissions and execution environment

##### Permissions

Operating system administrator

##### Execution environment

Master server

#### End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

## A.1.5 `bdpp_getimage`

---

#### Name

`/opt/FJSVbdpp/bin/bdpp_getimage` - Get cloning image

#### Format

`bdpp_getimage servername imagename`

#### Function description

This command takes a server system image to be used for cloning slave servers.



- Executing this command may take some time as it distributes cloning image files to servers.
- Four cloning image collection and deployment processes may be executed simultaneously. When five or more are requested, only four will be executed immediately, with others executed in turn as earlier processes complete.

#### Arguments

##### *servername*

Specify the name of the server of which the cloning image is to be taken.

##### *imagename*

Specify a name to be used to identify the cloning image being taken.

For the name of the cloning image, specify a string of no more than 32 characters, starting with a letter. Alphanumeric characters and the underscore ( ) can be used.

Specify a different name from the cloning image that has been collected. The command returns an error if the same name is specified.

#### Required permissions and execution environment

##### Permissions

Operating system administrator

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

## A.1.6 **bdpp\_lanctl**

---

Name

/opt/FJSVbdpp/bin/bdpp\_lanctl - Register the automatic configuration of the network parameter

Format

*bdpp\_lanctl ipaddr file1 file2*

Function description

This command enables the automatic configuration of the network parameter for the collection and distribution of the cloning image.

Arguments

*ipaddr*

Specify the IP address of the admin LAN of the slave server that will collect the cloning image.

*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 cloning image was collected.

*file2*

Using an absolute path, specify the definition file (ipaddr.conf) used for network parameter automatic configuration.

Required permissions and execution environment

Permissions

Operating system administrator

Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

## A.1.7 **bdpp\_listimage**

---

Name

/opt/FJSVbdpp/bin/bdpp\_listimage - Cloning image registration display

## Format

bdpp\_listimage

## Function description

This command checks cloning image registration.

The following items are displayed in a list.

Item name	Content
NAME	Image name of the cloning image
VERSION	The display is fixed at "1".
CREATIONDATE	Date and time cloning image was collected
COMMENT	The display is fixed at "-".

## Required permissions and execution environment

### Permissions

Operating system administrator

### 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.1.8 bdpp\_listserver

### Name

/opt/FJSVbdpp/bin/bdpp\_listserver - Slave server registration display

### Format

bdpp\_listserver

### Function description

This command displays the registration status of the slave server that collects the cloning image and the slave server to which this cloning image is deployed.

Following items are displayed:

Item name	Content
PHYSICAL_SERVER	Physical server name
SERVER	Server (physical OS/VM host) name
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"> <li>- normal</li> <li>- warning</li> <li>- unknown</li> <li>- stop</li> <li>- error</li> <li>- fatal</li> </ul> <p>Refer to "5.2 Resource Stat" in the "ServerView Resource Orchestrator Virtual Edition V3.0.0 Operation Guide" for information on the statuses of the server.</p>
MAINTENANCE	<p>Maintenance mode status</p> <ul style="list-style-type: none"> <li>- If the maintenance mode has been set "ON" is displayed.</li> <li>- If the maintenance mode has been cancelled "OFF" is displayed.</li> </ul>

#### Required permissions and execution environment

##### Permissions

Operating system administrator

##### 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.1.9 bdpp\_removeimage

---

### Name

/opt/FJSVbdpp/bin/bdpp\_removeimage - Cloning image deletion

### Format

bdpp\_removeimage *imagename*

### Function description

This command deletes a cloning image that was taken using the bdpp\_getimage command.

### Arguments

*imagename*

Specify the cloning image name.

### Required permissions and execution environment

#### Permissions

Operating system administrator

#### Execution environment

Master server

### End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

## A.1.10 bdpp\_removeserver

---

### Name

/opt/FJSVbdpp/bin/bdpp\_removeserver - Slave server deletion

### Format

bdpp\_removeserver *servername*

### Function description

This command deletes the slave server that collects the cloning image and the slave server to which this cloning image is deployed both of which were registered using the bdpp\_addserver command.

### Arguments

*servername*

Specify the name of the server to be deleted.

### Required permissions and execution environment

#### Permissions

Operating system administrator

#### Execution environment

Master server

End status

The following status codes are returned:

0

Processed normally

Other than 0

Error occurred

## A.1.11 bdpp\_start

---

Name

/opt/FJSVbdpp/bin/bdpp\_start - Interstage Big Data Parallel Processing Server start

Format

bdpp\_start

Function description

This command starts Hadoop.

It starts the Hadoop JobTracker of the master server and the Hadoop TaskTracker of all the connection target slave servers.

Options

None

Required permissions and execution environment

Permissions

Operating system administrator

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 of the Hadoop on the master server or in the TaskTracker of the Hadoop 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 JobTrackers of the Hadoop on the master server or the TaskTrackers of the Hadoop 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

Master server JobTracker	Slave server TaskTracker	bdpp_start execution results
	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
	Running on all slave servers	- "bdpp: WARN : 001: bdpp Hadoop is already started." message only is output

## A.1.12 bdpp\_stat

### Name

/opt/FJSVbdpp/bin/bdpp\_stat - Interstage Big Data Parallel Processing Server status check

### Format

bdpp\_stat [-all]

### Function description

This command checks whether the Hadoop status is started or stopped.

### Options

-all

Displays the JobTracker on the master server and the TaskTracker processes on the slave servers.

The following items are displayed in a list.

Content
Server 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

Operating system administrator

#### Execution environment

Master server

## End status

The following status codes are returned:

0

Hadoop is started.

1

Hadoop is not started.

## Example of the display

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.
```



.....  
An error may occur if multiple bdpp\_stat commands are executed at the same time. If this occurs, execute the bdpp\_stat command again.  
.....

## A.1.13 bdpp\_stop

---

### Name

/opt/FJSVbdpp/bin/bdpp\_stop - Interstage Big Data Parallel Processing Server stop

### Format

bdpp\_stop

### Function description

This command stops Hadoop.

It stops the Hadoop JobTracker of the master server and the Hadoop TaskTracker of all the connection target slave servers.

### Options

None

### Required permissions and execution environment

#### Permissions

Operating system administrator

#### 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 in the TaskTracker on the slave server.

If this occurs, check the cautions for the `bdpp_start` command, restart the Hadoop JobTracker and TrackTracker, then stop if necessary.

.....

## A.2 Definition Files

---

This section gives an overview of the definition files provided by this product.

Definition file list

Definition file	File overview	Explanation
<a href="#">bdpp.conf</a>	Configuration file	This file defines the information required for installation and setup of this product.
<a href="#">slaves</a>	Slave server definition file	This file defines the slave servers that are the connection targets of the master server.
<a href="#">clone.conf</a>	Cloning server definition file	This file is used for Smart Setup. It defines adding slave servers using cloning and deleting slave servers.
<a href="#">FJSVrcx.conf</a>	Network parameter automatic configuration definition file	This definition file is used to configure the network parameter automatically after cloning.
<a href="#">ipaddr.conf</a>	Network parameter automatic configuration definition file	This definition file is used to configure the network parameter automatically after cloning.

### A.2.1 `bdpp.conf`

---

This section describes the configuration file (`bdpp.conf`) used by this product.

File storage location

- `/etc/opt/FJSVbdpp/conf/bdpp.conf`

### Point

.....

Separate "`bdpp.conf`" file templates for each server type are stored under `"/DISK1/config_template"` on the distribution media (DVD-ROM) 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.

.....

## File content

The file content defines the information required for installation and setup of this product.  
A separate file must be created for each of the server functions.

### Note

For parameters for which specifying a value is optional, enter up to "parameter=" but omit the value.

Example when value is omitted:

BDPP\_PCL\_PRIMARY\_CONNECT2=

[Master server]

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.
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 master server (primary).  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 master server (primary) public LAN.  (The IP address that will be used to connect to public LAN1 in the graphic)
5	BDPP_PRIMARY_ADM_IP	Mandatory	192.168.1.1	Specify the IP address that will be used to connect to the master server (primary) admin LAN.  (The IP address that will be used to connect to the admin LAN in the graphic)
6	BDPP_SECONDARY_NAME	Mandatory	master2	Specify the representative host name of the master server (secondary).  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
7	BDPP_SECONDARY_IP	Mandatory	10.10.10.12	Specify the representative IP address that will be used to connect to the master server (secondary) public LAN.  (The IP address that will be used to connect to public LAN1 in the graphic)
8	BDPP_SECONDARY_ADM_IP	Mandatory	192.168.1.2	Specify the IP address that will be used to connect to the master server (secondary) admin LAN.  (The IP address that will be used to connect to the admin LAN in the graphic)
9	BDPP_PCL_CLUSTER_GROUP	Mandatory	bdppcluster	Specify the cluster group name to be set for the cluster.
10	BDPP_PCL_PRIMARY_iRMC_IP	Mandatory	192.168.1.11	Specify the IP address of the master server (primary) remote management controller.
11	BDPP_PCL_PRIMARY_iRMC_ACCOUNT	Mandatory	root	Specify the account of the master server (primary) remote management controller.
12	BDPP_PCL_PRIMARY_iRMC_PASS	Mandatory	password	Specify the password of the master server (primary) remote management controller.
13	BDPP_PCL_SECONDARY_iRMC_IP	Mandatory	192.168.1.12	Specify the IP address of the master server (secondary) remote management controller.
14	BDPP_PCL_SECONDARY_iRMC_ACCOUNT	Mandatory	root	Specify the account of the master server (secondary) remote management controller.
15	BDPP_PCL_SECONDARY_iRMC_PASS	Mandatory	password	Specify the password of the master server (secondary) remote management controller.
16	BDPP_PCL_PRIMARY_CIP	Mandatory	192.168.2.11	Specify the IP address (*1) used for the master server (primary) CIP (cluster interconnect protocol).
17	BDPP_PCL_SECONDARY_CIP	Mandatory	192.168.2.12	Specify the IP address (*1) used for the master server (secondary) CIP (cluster interconnect protocol).

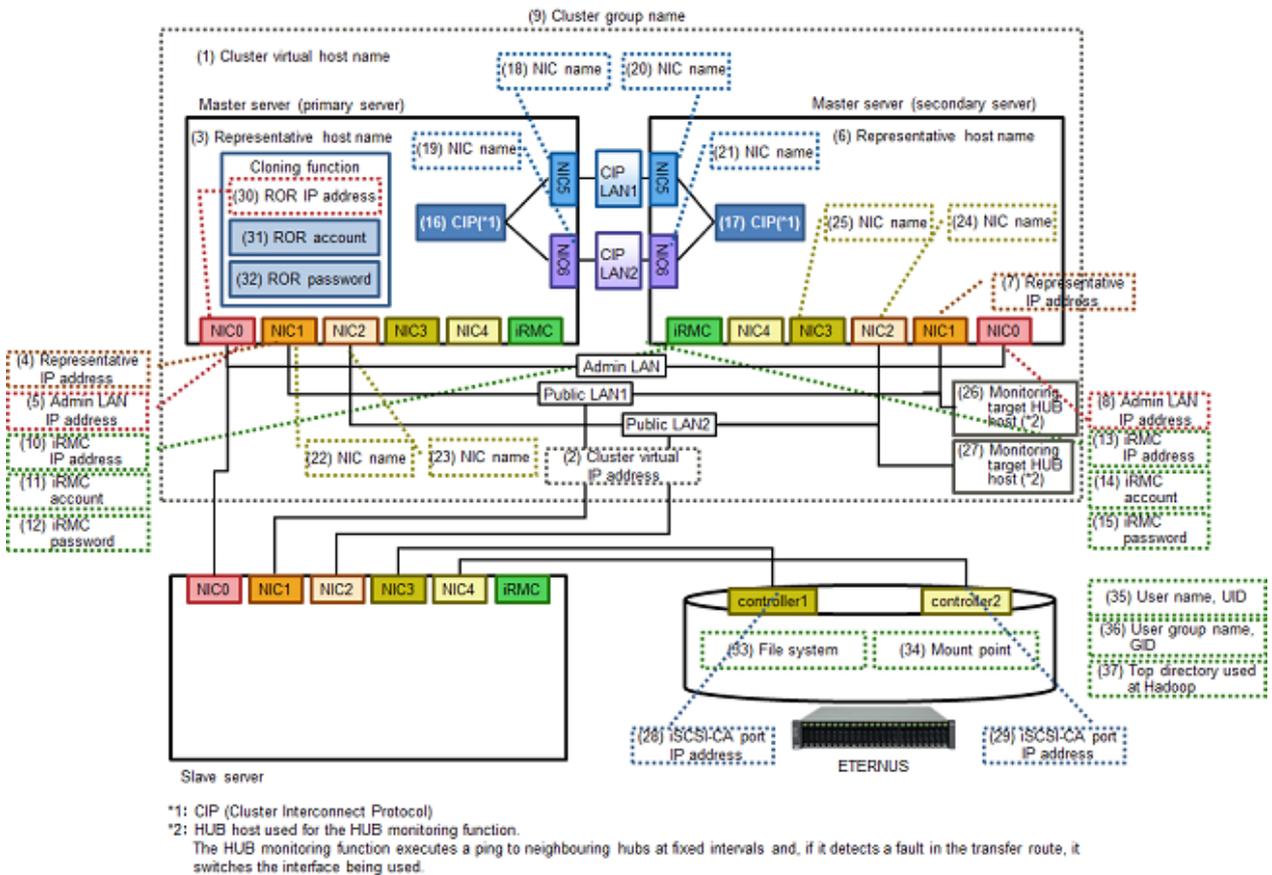
Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
18	BDPP_PCL_PRIMARY_CONNECT1	Mandatory	eth5	Specify the master server (primary) side NIC name (NIC5) used for connections between clusters.
19	BDPP_PCL_PRIMARY_CONNECT2	Optional	eth6	Specify the master server (primary) side NIC name (NIC6) used for connections between clusters.
20	BDPP_PCL_SECONDARY_CONNECT1	Mandatory	eth5	Specify the master server (secondary) side NIC name (NIC5) used for connections between clusters.
21	BDPP_PCL_SECONDARY_CONNECT2	Optional	eth6	Specify the master server (secondary) side NIC name (NIC6) used for connections between clusters.
22	BDPP_GLS_PRIMARY_CONNECT1	Mandatory	eth1	Specify the master server (primary) NIC name (NIC1) used for connections between slave servers.  (The NIC that will be used to connect to public LAN1 in the graphic)
23	BDPP_GLS_PRIMARY_CONNECT2	Optional	eth2	Specify the master server (primary) NIC name (NIC2) used for connections between slave servers.  (The NIC that will be used to connect to public LAN2 in the graphic)
24	BDPP_GLS_SECONDARY_CONNECT1	Mandatory	eth1	Specify the master server (secondary) NIC name (NIC1) used for connections between slave servers.  (The NIC that will be used to connect to public LAN1 in the graphic)
25	BDPP_GLS_SECONDARY_CONNECT2	Optional	eth2	Specify the master server (secondary) NIC name (NIC2) used for connections between slave servers.  (The NIC that will be used to connect to public LAN2 in the graphic)
26	BDPP_GLS_POLLING1_IP	Mandatory	10.10.10.100	Specify the IP address of the monitoring target HUB host.

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
				(Monitoring target HUB host on public LAN1 in the graphic)
27	BDPP_GLS_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 graphic)
28	BDPP_DISKARRAY_iSCSI1_IP	Mandatory	10.10.11.11	Specify the IP address of the storage system disk array iSCSI-CA port to which the connection will be made
29	BDPP_DISKARRAY_iSCSI2_IP	Optional	10.10.12.11	If the storage system connection has been duplicated, specify the other IP address of the storage system disk array iSCSI-CA port to which the connection will be made.
30	BDPP_ROR_ADMIN_IP	Mandatory	192.168.1.1	Specify the master server (primary) admin LAN IP address to be specified when installing ROR.
31	BDPP_ROR_ADMIN_ACCOUNT	Mandatory	root	Specify the account of the local server used when installing ROR.
32	BDPP_ROR_ADMIN_PASS	Mandatory	password	Specify the password of the local server used when installing ROR.
33	BDPP_PDFS_FILESYSTEM	Mandatory	YES	If DFS is to be used as the file system, specify YES (or yes).
34	BDPP_PDFS_MOUNTPOINT	Optional	/mnt/pdfs	Specify the DFS mount point. This parameter is mandatory if YES is specified for BDPP_PDFS_FILESYSTEM.
35	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. When registering multiple user IDs, use a space ( ) to separate each user name and UID. Example

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
				bdppuser1,1500 bdppuser2,1501 bdppuser2,1502
36	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).
37	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.

\*1: An interconnect private IP interface can be built by using an IP address provided for a private network.

Figure A.1 Network configuration of the master server and bdpp.conf parameters

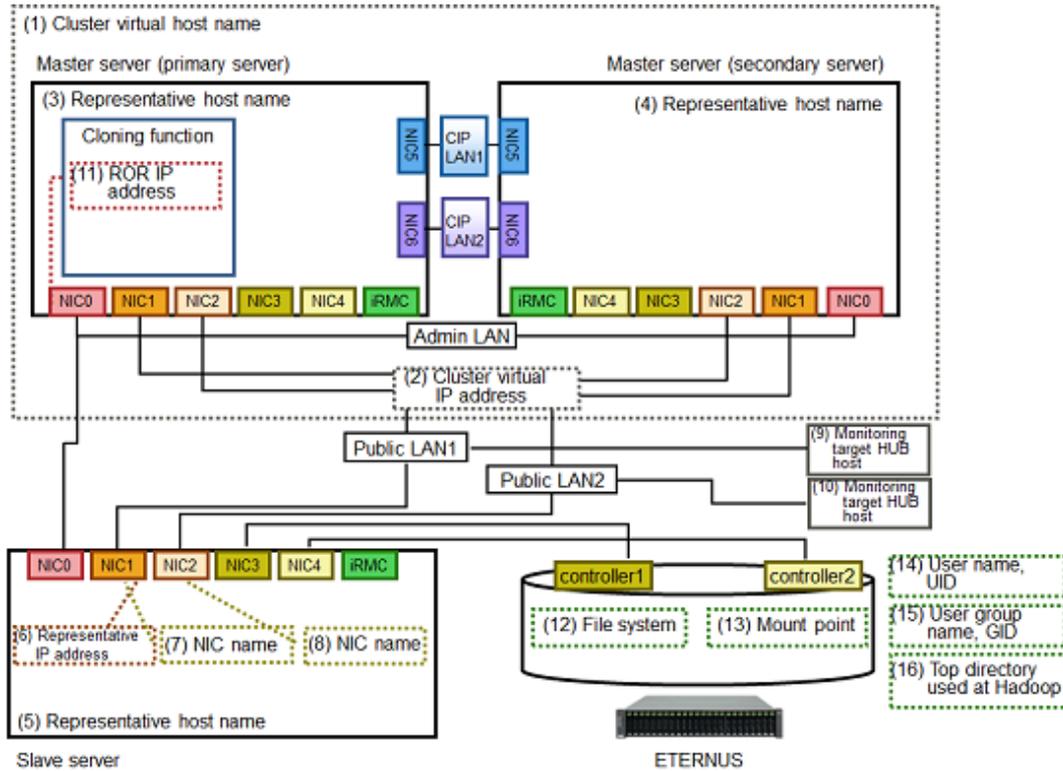


[Slave server]

Item No.	Parameter	Mandatory/Optional	Value to be specified (settings example)	Explanation
1	BDPP_CLUSTER_NAME	Mandatory	cluster	Specify the value that was specified for the master server bdpp.conf BDPP_CLUSTER_NAME parameter.
2	BDPP_CLUSTER_IP	Mandatory	10.10.10.1	Specify the value that was specified for the master server bdpp.conf BDPP_CLUSTER_IP parameter.
3	BDPP_PRIMARY_NAME	Mandatory	master1	Specify the value that was specified for the master server bdpp.conf BDPP_PRIMARY_NAME parameter.
4	BDPP_SECONDARY_NAME	Mandatory	master2	Specify the value that was specified for the master server bdpp.conf BDPP_SECONDARY_NAME parameter.
5	BDPP_SERVER_NAME	Mandatory	slave1	Specify the representative host name of the slave server.  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	10.10.10.21	Specify the representative IP address that will be used to connect to the slave server public LAN
7	BDPP_GLS_SERVER_CONNECT1	Optional	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 graphic)
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 graphic)
9	BDPP_GLS_SERVER_POLLING1_IP	Optional	10.10.10.100	Specify the IP address of the monitoring target HUB host.  (Monitoring target HUB host on public LAN1 in the graphic)

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
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 graphic)
11	BDPP_ROR_ADMIN_IP	Mandatory	192.168.1.1	Specify the value that was specified for the master server bdpp.conf BDPP_ROR_ADMIN parameter.
12	BDPP_PDFS_FILESYSTEM	Mandatory	YES	Specify the value that was specified for the master server bdpp.conf BDPP_PDFS_FILESYSTEM parameter.
13	BDPP_PDFS_MOUNTPOINT	Optional	/mnt/pdfs	Specify the value that was specified for the master server bdpp.conf BDPP_PDFS_MOUNTPOINT parameter.
14	BDPP_HADOOP_DEFAULT_USERS	Optional	bdppuser1,1500	Specify the value that was specified for the master server bdpp.conf BDPP_HADOOP_DEFAULT_USERS parameter.
15	BDPP_HADOOP_DEFAULT_GROUP	Optional	bdppgroup,1500	Specify the value that was specified for the master server bdpp.conf BDPP_HADOOP_DEFAULT_GROUP parameter.
16	BDPP_HADOOP_TOP_DIR	Optional	/hadoop	Specify the value that was specified for the master server bdpp.conf BDPP_HADOOP_TOP_DIR parameter.

Figure A.2 Slave server network configuration and bdpp.conf parameter

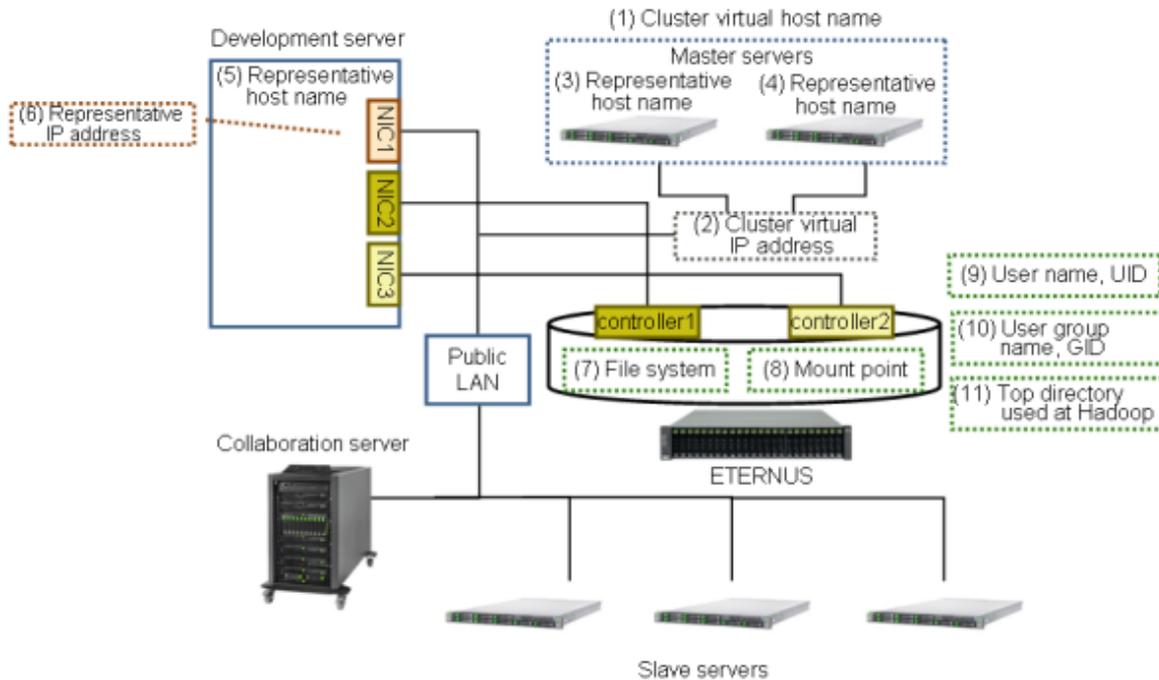


[Development server]

Item No.	Parameter	Mandatory/Optional	Value to be specified (settings example)	Explanation
1	BDPP_CLUSTER_NAME	Mandatory	cluster	Specify the value that was specified for the master server bdpp.conf BDPP_CLUSTER_NAME parameter.
2	BDPP_CLUSTER_IP	Mandatory	10.10.10.1	Specify the value that was specified for the master server bdpp.conf BDPP_CLUSTER_IP parameter.
3	BDPP_PRIMARY_NAME	Mandatory	master1	Specify the value that was specified for the master server bdpp.conf BDPP_PRIMARY_NAME parameter.
4	BDPP_SECONDARY_NAME	Mandatory	master2	Specify the value that was specified for the master server bdpp.conf BDPP_SECONDARY_NAME parameter.
5	BDPP_SERVER_NAME	Mandatory	develop	Specify the representative host name of the development server.

Item No.	Parameter	Mandatory/ Optional	Value to be specified (settings example)	Explanation
				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	If PDFS is to be used as the file system, specify YES (or yes). This parameter must match the master server settings.
8	BDPP_PDFS_MOUNTPOINT	Optional	/mnt/pdfs	Specify the PDFS mount point. This parameter is mandatory if YES is specified for BDPP_PDFS_FILESYSTEM. This parameter must match the master server settings.
9	BDPP_HADOOP_DEFAULT_USERS	Optional	bdppuser1,1500	Specify the value that was specified for the master server bdpp.conf BDPP_HADOOP_DEFAULT_USERS parameter.
10	BDPP_HADOOP_DEFAULT_GROUP	Optional	bdppgroup,1500	Specify the value that was specified for the master server bdpp.conf BDPP_HADOOP_DEFAULT_GROUP parameter.
11	BDPP_HADOOP_TOP_DIR	Optional	/hadoop	Specify the value that was specified for the master server bdpp.conf BDPP_HADOOP_TOP_DIR parameter.

Figure A.3 Development server network configuration and bdpp.conf parameter

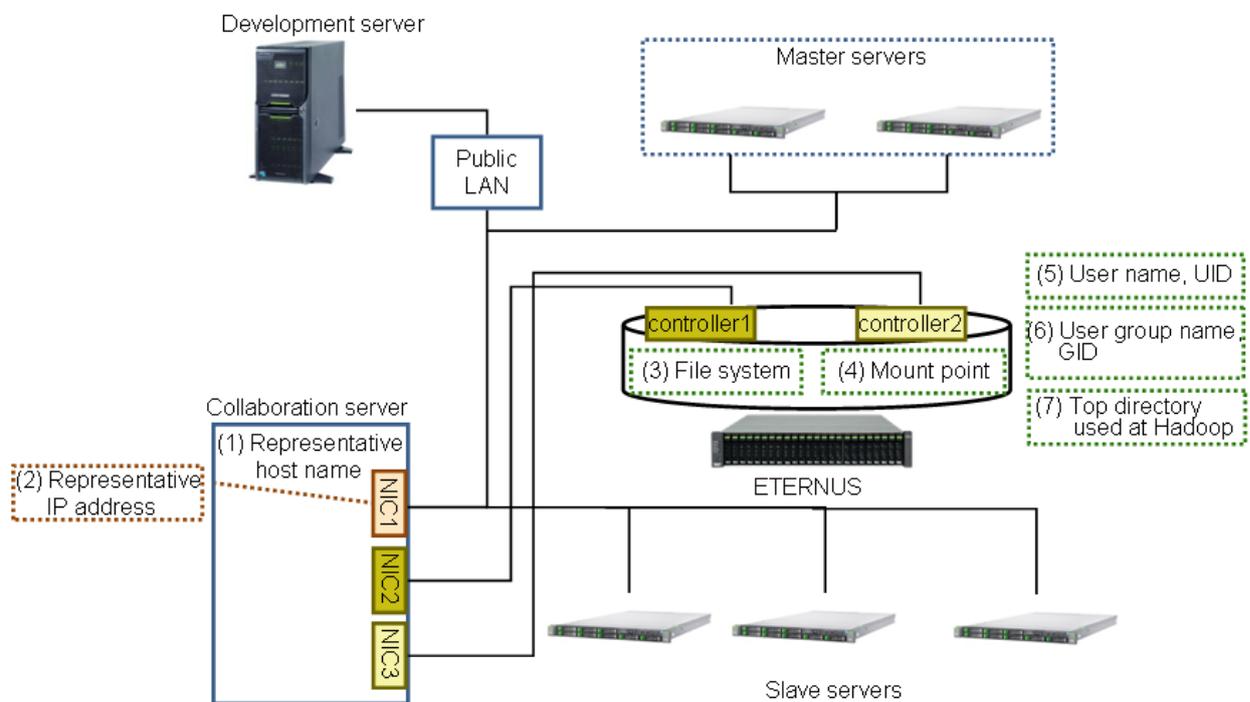


[Collaboration server]

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. 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	If PDFS is to be used as the file system, specify YES (or yes). This parameter must match the master server settings.
4	BDPP_PDFS_MOUNTPOINT	Optional	/mnt/pdfs	Specify the PDFS mount point. This parameter is mandatory if YES is specified for BDPP_PDFS_FILESYSTEM. This parameter must match the master server settings.
5	BDPP_HADOOP_DEFAULT_USERS	Optional	bdppuser1,1500	Specify the value that was specified for the master server bdpp.conf

Item No.	Parameter	Mandatory/Optional	Value to be specified (settings example)	Explanation
				BDPP_HADOOP_DEFAULT_USERS parameter.
6	BDPP_HADOOP_DEFAULT_GROUP	Optional	bdppgroup,1500	Specify the value that was specified for the master server bdpp.conf BDPP_HADOOP_DEFAULT_GROUP parameter.
7	BDPP_HADOOP_TOP_DIR	Optional	/hadoop	Specify the value that was specified for the master server bdpp.conf BDPP_HADOOP_TOP_DIR parameter.

Figure A.4 Collaboration server network configuration and bdpp.conf parameter



#### Definition examples

##### [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=192.168.1.1
BDPP_SECONDARY_NAME=master2
BDPP_SECONDARY_IP=10.10.10.12
BDPP_SECONDARY_ADM=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_IP=10.10.10.1
BDPP_CLUSTER_NAME=cluster
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_POLLING1_IP=10.10.10.100
BDPP_GLS_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_IP=10.10.10.1
BDPP_CLUSTER_NAME=cluster
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
```

## A.2.2 slaves

---

This section describes the connection target slave server definition file (slaves) used by this product.

File storage location

- /etc/opt/FJSVbdpp/conf/slaves

File content

This file defines the slave servers targeted for connection. If there are multiple connection target slave servers, define them separated by commas.

Definition example

```
slave1,slave2,slave3,slave4,slave5
```

## A.2.3 clone.conf

---

This section describes the cloning server definition file (clone.conf) used by this product.

File storage location

- Any directory

File content

This file is used to define the slave server that collects the cloning image and the slave server that is added using Smart Setup. The configuration definition file codes the items defined in each line, separated by commas.

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 (#)



### Example

```
# Development environment definition
```

- 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

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 host name for the NIC that connects to the slave server admin LAN.  For the host name, the first character must be a letter, and it must consist of 11 or fewer alphanumeric characters.   <b>Note</b> ..... If the host 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 host name for the slave server must also be 11 or fewer characters in order to preserve consistency across the system. .....
ip_address	192.168.1.13	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").
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	10.10.10.13	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).

Section header	Value to be specified (settings example)	Content
ipmi_passwd	admin	<p>Specify the password of the remote management controller that manages the server.</p> <p>Specify a string of up to 16 characters comprised of single-byte alphanumerics and ASCII character symbols (0x20 to 0x7e).</p> <p>Nothing need be specified if a password is not set.</p> <p> <b>Note</b></p> <p>.....</p> <p>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.</p> <p>.....</p>
ipmi_passwd_enc	plain	<p>Specify one of the following:</p> <ul style="list-style-type: none"> <li>- If the ipmi_passwd character string is plain text: "plain"</li> <li>- If encrypted: "encrypted"</li> </ul>
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	<p>Specify the following as the resource operation content:</p> <p>Register: new</p> <p>Change: change</p> <p>Hyphen (-): Do nothing</p>
server_name	slave1	<p>Specify the host name corresponding to the Network Interface Card connected to the admin LAN of the slave server whose cloning image is to be collected.</p> <p>For the host name, the first character must be a letter, and it must consist of 11 or fewer alphanumeric characters.</p>

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, 10.10.10.13, A0:A0:A0:A0:A0:A1, , public, 192.168.1.13, admin, admin, plain, , , OFF

```

```
new,,,slave2, 10.10.10.14,A0:A0:A0:A0:A0:A2,,,public,192.168.1.14,admin,admin,plain,,,OFF
new,,,slave3, 10.10.10.15,A0:A0:A0:A0:A0:A3,,,public,192.168.1.15,admin,admin,plain,,,OFF
new,,,slave4, 10.10.10.16,A0:A0:A0:A0:A0:A4,,,public,192.168.1.16,admin,admin,plain,,,OFF
new,,,slave5, 10.10.10.17,A0:A0:A0:A0:A0:A5,,,public,192.168.1.17,admin,admin,plain,,,OFF
[ServerAgent]
operation,server_name
new,slave1
```

## A.2.4 FJSVrcx.conf

This section describes the definition file (FJSVrcx.conf) that is used to configure the network parameter automatically after cloning.

File content

```
admin_LAN=192.168.1.11
hostname=slave1
```

- admin\_LAN

Specify the IP address of the admin LAN of the slave server collecting the cloning image.

- hostname

Specify the host name of the slave server collecting the cloning image.

For the host name, the first character must be a letter, and it must consist of 11 or fewer alphanumeric characters.

## A.2.5 ipaddr.conf

This section describes the definition file (ipaddr.conf) that is used to configure the network parameter automatically after cloning.

File content

When using Smart Setup to add slave servers, define the public LAN for the server whose clone image is to be collected and all of the servers to receive the distribution as slave servers. The definition file consists of the following entries.

- One or more node entries
- One or more interface entries under the node entry (with/without redundancy)

Node entries

Define the value that is set for the definition keyword shown in the table below.



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	Set the server name that was set when the slave sever was registered.

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 starting from 0 in ascending order must be used as the number that is added at the end of the definition keyword.

Defined content	Keyword	Value to be set	Explanation
Interface name	IF_NAME	eth $X$	$X$ is an integer of 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.

Defined content	Keyword	Value to be set	Explanation
Virtual interface name	VIF_NAME	sha $X$	$X$ is an integer between 0 and 255.
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 of 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 of 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	A maximum of two IP addresses can be specified. Use a comma (,) to separate them. To perform HUB-HUB monitoring, specify two monitoring targets.
Standby patrol virtual interface name	PAT_NAME	sha $Y$	$Y$ is an integer between 0 and 255. Specify a value that is different from the virtual interface name. If there is no standby patrol setting, do not set anything here.
Existence of HUB-HUB monitoring	POL_HUBS	ON/OFF	To perform HUB-HUB monitoring, specify ON. Otherwise, specify OFF.

Defined content	Keyword	Value to be set	Explanation
			<p>This definition is only enabled when two monitoring targets have been specified.</p> <p>When only one monitoring target has been specified, this definition is disabled and processed as OFF.</p>

#### Definition example

```

NODE_NAME="slave1"                                <-- Slave server collecting the cloning image
VIF_NAME0="sha0"
VIF_IPAD0="192.168.20.11"
VIF_MASK0="255.255.255.0"
PRI_NAME0="eth2"
SCD_NAME0="eth3"
POL_ADDR0="192.168.20.100,192.168.20.200"
PAT_NAME0="sha1"
POL_HUBS0="ON"

NODE_NAME="slave2"                                <-- Slave server to which the cloning image is to be
distributed
VIF_NAME0="sha0"
VIF_IPAD0="192.168.20.12"
VIF_MASK0="255.255.255.0"
PRI_NAME0="eth2"
SCD_NAME0="eth3"
POL_ADDR0="192.168.20.100,192.168.20.200"
PAT_NAME0="sha1"
POL_HUBS0="ON"

NODE_NAME="slave3"                                <-- Slave server to which the cloning image is to be
distributed
VIF_NAME0="sha0"
VIF_IPAD0="192.168.20.13"
VIF_MASK0="255.255.255.0"
PRI_NAME0="eth2"
SCD_NAME0="eth3"
POL_ADDR0="192.168.20.100,192.168.20.200"
PAT_NAME0="sha1"
POL_HUBS0="ON"

NODE_NAME="slave4"                                <-- Slave server to which the cloning image is to be
distributed
VIF_NAME0="sha0"
VIF_IPAD0="192.168.20.14"
VIF_MASK0="255.255.255.0"
PRI_NAME0="eth2"
SCD_NAME0="eth3"
POL_ADDR0="192.168.20.100,192.168.20.200"
PAT_NAME0="sha1"
POL_HUBS0="ON"

NODE_NAME="slave5"                                <-- Slave server to which the cloning image is to be
distributed
VIF_NAME0="sha0"
VIF_IPAD0="192.168.20.15"
VIF_MASK0="255.255.255.0"
PRI_NAME0="eth2"
SCD_NAME0="eth3"
POL_ADDR0="192.168.20.100,192.168.20.200"

```

PAT\_NAME0="sha1"  
POL\_HUBS0="ON"

## A.3 Port List

This section describes the ports used by this product.

Table A.2 Master server

Function overview	Send source				Send destination				Protocol
	Server	Service	Port	Change	Server	Service	Port	Change	
Backup/ restore and cloning	Master server	-	4972	Cannot change	Slave server	-	4973	Cannot change	udp
	Slave server	-	4973	Cannot change	Master server	-	4972	Cannot change	udp
		bootpc	68	Cannot change		bootpc	67	Cannot change	udp
		-	Value not fixed	-		pxc	4011	Cannot change	udp
		-	Value not fixed	-		tftp	69	Cannot change	udp
	Master server	-	Value not fixed	-	Master server	-	4971	Cannot change	tcp
Taking backup cloning	Master server	-	4974 to 4989	-	Master server	-	4974 to 4989	-	udp
Distributing backup cloning	Slave server	-	Value not fixed	-	Master server	-	4974 to 4989	-	tcp udp
Hadoop	Slave server	Hadoop (DataNode)	Not fixed	Cannot change	Master server	Hadoop (NameNode) (*1)	8020	Cannot change	tcp
	Development server	Hadoop (Client)	Not fixed	Cannot change					
	Development server	Hadoop (Client)	Not fixed	Cannot change	Slave server	Hadoop (DataNode) (*1)	50010	Cannot change	tcp
	Slave server	Hadoop (DataNode)	Not fixed	Cannot change	Slave server	Hadoop (DataNode) (*1)	50020	Cannot change	tcp
	Development server	Hadoop (Client)	Not fixed	Cannot change					
	Master server	Hadoop (Secondary NodeName)	Not fixed	Cannot change	Master server	Hadoop (NameNode) (*1)	50070	Cannot change	tcp
	Development server	Hadoop (Client)	Not fixed	Cannot change					
	Slave server	Hadoop (DataNode)	Not fixed	Cannot change	Slave server	Hadoop (DataNode) (*1)	50075	Cannot change	tcp

Function overview	Send source				Send destination				Protocol
	Server	Service	Port	Change	Server	Service	Port	Change	
	Development server	Hadoop (Client)	Not fixed	Cannot change					
	Master server	Hadoop (NodeName)	Not fixed	Cannot change	Master server	Hadoop (Secondary NodeName) (*1)	50090	Cannot change	tcp
	Master server	Hadoop (Secondary NodeName)	Not fixed	Cannot change	Master server	Hadoop (NameNode) (*1)	50470	Cannot change	tcp
	Development server	Hadoop (Client)	Not fixed	Cannot change					
	Slave server	Hadoop (DataNode)	Not fixed	Cannot change	Slave server	Hadoop (DataNode) (*1)	50475	Cannot change	tcp
	Development server	Hadoop (Client)	Not fixed	Cannot change					
	Slave server	Hadoop (TaskTracker)	Not fixed	Cannot change	Master server	Hadoop (JobTracker)	9000	Cannot change	tcp
	Development server	Hadoop (Client)	Not fixed	Cannot change	Master server				
	Development server	Hadoop (Client)	Not fixed	Cannot change	Master server	Hadoop (JobTracker)	50030	Cannot change	tcp
	Slave server	Hadoop (TaskTracker)	Not fixed	Cannot change	Slave server	Hadoop (TaskTracker)	50060	Cannot change	tcp
	Slave server	HBase (RegionServer)	Not fixed	Cannot change	Master server	HBase (Hmaster) (BackupHmaster)	60000	Cannot change	tcp
	Development server	HBase (Client)	Not fixed	Cannot change					
	Development server	HBase (Client)	Not fixed	Cannot change	Master server	HBase (Hmaster) (BackupHmaster)	60010	Cannot change	tcp
	Master server	HBase (Hmaster) (BackupHmaster)	Not fixed	Cannot change	Slave server	HBase (RegionServer)	60020	Cannot change	tcp
	Slave server	HBase (RegionServer)	Not fixed	Cannot change					
	Development server	HBase (Client)	Not fixed	Cannot change					
	Development server	HBase (Client)	Not fixed	Cannot change	Slave server	HBase (RegionServer)	60030	Cannot change	tcp
	Master server	HBase (Hmaster) (BackupHmaster)	Not fixed	Cannot change	Slave server	ZooKeeper	2081	Cannot change	tcp

Function overview	Send source				Send destination				Protocol
	Server	Service	Port	Change	Server	Service	Port	Change	
	Slave server	HBase (RegionServer)	Not fixed	Cannot change					
	Development server	HBase (Client)	Not fixed	Cannot change					
	Slave server	ZooKeeper	Not fixed	Cannot change	Slave server	ZooKeeper	2888	Cannot change	tcp
	Slave server	ZooKeeper	Not fixed	Cannot change	Slave server	ZooKeeper	3888	Cannot change	tcp

Table A.3 Slave server

Function overview	Send source				Send destination				Protocol
	Server	Service	Port	Change	Server	Service	Port	Change	
Cloning	Master server	-	4972	Cannot change	Slave server	-	4973	Cannot change	udp
	Slave server	-	4973	Cannot change	Master server	-	4972	Cannot change	udp
		-	Value not fixed	-		tftp	69	Cannot change	udp
Hadoop	Slave server	Hadoop (DataNode)	Not fixed	Cannot change	Master server	Hadoop (NameNode) (*1)	8020	Cannot change	tcp
	Development server	Hadoop (Client)	Not fixed	Cannot change					
	Development server	Hadoop (Client)	Not fixed	Cannot change	Slave server	Hadoop (DataNode) (*1)	50010	Cannot change	tcp
	Slave server	Hadoop (DataNode)	Not fixed	Cannot change	Slave server	Hadoop (DataNode) (*1)	50020	Cannot change	tcp
	Development server	Hadoop (Client)	Not fixed	Cannot change					
	Master server	Hadoop (Secondary NodeName)	Not fixed	Cannot change	Master server	Hadoop (NameNode) (*1)	50070	Cannot change	tcp
	Development server	Hadoop (Client)	Not fixed	Cannot change					
	Slave server	Hadoop (DataNode)	Not fixed	Cannot change	Slave server	Hadoop (DataNode) (*1)	50075	Cannot change	tcp
	Development server	Hadoop (Client)	Not fixed	Cannot change					
	Master server	Hadoop (NodeName)	Not fixed	Cannot change	Master server	Hadoop (Secondary NodeName) (*1)	50090	Cannot change	tcp
Master server	Hadoop (Secondary NodeName)	Not fixed	Cannot change	Master server	Hadoop (NameNode) (*1)	50470	Cannot change	tcp	

Function overview	Send source				Send destination				Protocol
	Server	Service	Port	Change	Server	Service	Port	Change	
	Development server	Hadoop (Client)	Not fixed	Cannot change					
	Slave server	Hadoop (DataNode)	Not fixed	Cannot change	Slave server	Hadoop (DataNode) (*1)	50475	Cannot change	tcp
	Development server	Hadoop (Client)	Not fixed	Cannot change					
	Slave server	Hadoop (TaskTracker)	Not fixed	Cannot change	Master server	Hadoop (JobTracker)	9000	Cannot change	tcp
	Development server	Hadoop (Client)	Not fixed	Cannot change	Master server				
	Development server	Hadoop (Client)	Not fixed	Cannot change	Master server	Hadoop (JobTracker)	50030	Cannot change	tcp
	Slave server	Hadoop (TaskTracker)	Not fixed	Cannot change	Slave server	Hadoop (TaskTracker)	50060	Cannot change	tcp
	Slave server	HBase (RegionServer)	Not fixed	Cannot change	Master server	HBase (Hmaster) (BackupHmaster)	60000	Cannot change	tcp
	Development server	HBase (Client)	Not fixed	Cannot change					
	Development server	HBase (Client)	Not fixed	Cannot change	Master server	HBase (Hmaster) (BackupHmaster)	60010	Cannot change	tcp
	Master server	HBase (Hmaster) (BackupHmaster)	Not fixed	Cannot change	Slave server	HBase (RegionServer)	60020	Cannot change	tcp
	Slave server	HBase (RegionServer)	Not fixed	Cannot change					
	Development server	HBase (Client)	Not fixed	Cannot change					
	Development server	HBase (Client)	Not fixed	Cannot change	Slave server	HBase (RegionServer)	60030	Cannot change	tcp
	Master server	HBase (Hmaster) (BackupHmaster)	Not fixed	Cannot change	Slave server	ZooKeeper	2081	Cannot change	tcp
	Slave server	HBase (RegionServer)	Not fixed	Cannot change					
	Development server	HBase (Client)	Not fixed	Cannot change					
	Slave server	ZooKeeper	Not fixed	Cannot change	Slave server	ZooKeeper	2888	Cannot change	tcp
	Slave server	ZooKeeper	Not fixed	Cannot change	Slave server	ZooKeeper	3888	Cannot change	tcp

## A.4 Tuning System Parameters

Refer to "A.4.1 System Parameter Tuning Values" below for the system parameters that require tuning, and their values.

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

- If the type is "Maximum":

If the value already set (initial value or previously set value) is greater than the value shown in the table, it need not be changed. If it is smaller than the value in the table, change it to the value in the table.

- If the type is "Additional":

Add the value shown in the table to the value that is already set (initial value or previously set 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.

### A.4.1 System Parameter Tuning Values

#### Master server

##### 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.shmmni	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 / collaboration server

### Semaphore

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

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

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>SEMOPMvalue</i>	Number of semaphore operators for the system as a whole	2 Minimum of 22 required	Additional

## A.4.2 Tuning Work Procedures

1. Use the following command to check the values set for the relevant parameters in the current system:

```
/sbin/sysctl -a
```

### Example

```
# /sbin/sysctl -a
...
(snip)
...
kernel.sem = 250          32000    32      128
kernel.msgmnb = 4194304
kernel.msgmni = 8192
kernel.msgmax = 65536
kernel.shmmni = 4096
kernel.shmall = 4294967296
kernel.shmmax = 68719476736
...
(snip)
...
```

2. Refer to "A.4.1 System Parameter Tuning Values" and compare with the current values that are set. Calculate the appropriate value to set for each parameter, taking into account the type (maximum or additional).
3. Edit "/etc/sysctl.conf" as shown in the example below.

### Example

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

4. Use the following command to check that the edited content is reflected in "/etc/sysctl.conf":

```
# /bin/cat /etc/sysctl.conf
```

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

- Rebooting the system to reflect the settings

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

- Using /sbin/sysctl -p to reflect the settings

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

\*1: Reboot is not required if this command is used.

6. Check the output of the following command to check that the set system parameters are reflected:

```
# /sbin/sysctl -a
```

### Example

```
# /sbin/sysctl -a
...
(snip)
...
kernel.sem = 250          32002   32      130
kernel.msgmnb = 4194304
kernel.msgmni = 8192
kernel.msgmax = 65536
kernel.shmmni = 4096
kernel.shmall = 4294967296
kernel.shmmax = 68719476736
...
(snip)
...
```

# Appendix B Messages

This appendix explains the messages output or displayed by this product.

## B.1 Messages During Installation

This section explains the messages output or displayed during installation of this product.

---

### **bdpp\_inst: ERROR: 001: Configuration file bdpp.conf does not exist.**

#### Description

The configuration file bdpp.conf does not exist.

#### Action method

Create bdpp.conf and specify the bdpp.conf file path in the BDPP\_CONF\_PATH environment variable.

---

### **bdpp\_inst: ERROR: 002: bdpp.conf parameter error. (\$NUM: \$LINE)**

#### Description

There is an error in the parameter shown at \$NUM (line number): \$LINE (parameter and specification value) in the bdpp.conf configuration file.

#### Action method

Refer to bdpp.conf, check the parameter and correct the content.

---

### **bdpp\_inst: ERROR: 004: This server's IP does not match primary IP (\$BDPP\_PRIMARY\_IP) nor secondary IP (\$BDPP\_SECONDARY\_IP).**

#### Description

The server IP address for the master server installation attempt does not match the value of the parameters (BDPP\_PRIMARY\_IP or BDPP\_SECONDARY\_IP) specified in the bdpp.conf configuration file.

#### Action method

Check the following and take the appropriate actions.

- The BDPP\_PRIMARY\_IP or BDPP\_SECONDARY\_IP specified in the configuration file bdpp.conf may not match the IP address allocated for connection to the public LAN.  
Make sure the IP address allocated for connection to the public LAN has been specified, then try installing again.
- Make sure 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

Regardless of whether there is an attempt to use the DFS file system for "BDPP\_PDFS\_FILESYSTEM", no mountpoint is specified for "BDPP\_PDFS\_MOUNTPOINT".

#### Action method

Specify the PDFS mount point at "BDPP\_PDFS\_MOUNTPOINT".

---

**bdpp\_inst: ERROR: 010: Installation failed. (\$PROCESS)****Description**

Installation failed at the installer processing shown at \$PROCESS.

**Action method**

- When \$PROCESS is "groupadd" or "useradd"  
groupadd or useradd command processing fails. Check whether the value specified for "BDPP\_HADOOP\_GROUP" or "BDPP\_HADOOP\_USER" in the bdpp.conf file is already registered in the system. If it is, delete it then re-execute the installation.
- When \$PROCESS is "PDFS"  
The directories "/opt/FJSVpdfs", "/etc/opt/FJSVpdfs", "/var/opt/FJSVpdfs", or "/etc/pdfs" may have remained. Make a backup if necessary, then delete the relevant directory and rerun the installer.
- When \$PROCESS is "ServerView Resource Orchestrator"  
The Microsoft LAN Manager module described in "Obtaining and deploying the Microsoft LAN Manager module" in "2.3.3.1 Obtaining and Setting Up the Mandatory Software" - "Obtaining mandatory software" does not exist in the system's work directory (/tmp). Put the Microsoft LAN Manager module into the system's work directory (/tmp), and rerun the installer.
- When \$PROCESS is other than the above  
Refer to the install log "/var/tmp/bdpp\_inst.log" and remove the cause of the error. Then perform the installation again.

---

**bdpp\_inst: ERROR: 011: Uninstallation failed. (\$PROCESS)****Description**

Uninstallation failed at the uninstaller processing shown at \$PROCESS.

**Action method**

Refer to the install log "/var/tmp/bdpp\_inst.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 administrator privilege.

**Action method**

Execute again with system administrator privilege.

---

**bdpp\_inst: ERROR: 102: \$PKG package is not installed.****Description**

The mandatory package shown at \$PKG is not installed.

**Action method**

Install the mandatory package shown at \$PKG.



Contact Fujitsu technical staff if an error message other than the ones shown above is output or displayed.

---

## B.2 Messages During Setup

---

This section explains the messages output or displayed during setup of this product.

---

### **bdpp\_setup: ERROR: 001: Setup of the cluster failed.**

#### Description

Cluster setup failed.

#### Action method

- When setup fails with cluster\_setup2

The NIC settings of the cluster interconnect are not "ONBOOT=YES".

Check that the setting in "/etc/sysconfig/network-scripts", "ifcfg-ethX"(\*) is "ONBOOT=YES", changing "NO" to "YES" if necessary, then restart the network interface and retry cluster\_setup2.

\*: ethX is a network interface used for cluster interconnection. Specify numerals at x.

- In all other cases

Check the message that was output prior to this message and take action. Then, perform cluster setup again.

---

### **bdpp\_setup: ERROR: 002: Removal of the cluster failed.**

#### Description

Cluster unsetup failed.

#### Action method

Check the message that was output prior to this message and take action. Then, perform cluster unsetup again.

---

### **bdpp\_setup: ERROR: 003: Setup failed. (\$PROCESS)**

#### Description

Setup failed at the setup processing shown at \$PROCESS.

#### Action method

- When \$PROCESS is "PRIMECLUSTER GLS"

The system may not have been restarted after installation. Restart the system, then redo setup.

- 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

Unsetup failed at the unsetup processing shown at \$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 unsetup 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 whether the mount point specified at fstab is specified for "BDPP\_PDFS\_MOUNTPOINT" in the bdpp.conf configuration file. Change it to the correct value.

---

**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 administrator privilege.

**Action method**

Execute again with system administrator privilege.



.....  
Contact Fujitsu technical staff if an error message other than the ones shown above is output or displayed.  
.....

---

## B.3 Messages Output During Operation

---

This section explains the messages output or displayed during operation of this product.

### B.3.1 Messages Output During Command Execution

---

This section explains the messages output or displayed by 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 "[A.3 Port List](#)", permit connections for ports used by Interstage Big Data Processing Server, then retry the command.
- The SSH daemon is not running or SSH communications may have failed. Review the SSH settings, then re-execute the command.
- Communications between the primary and secondary master servers may not have been set so that SSH is possible without a password. Create a public key for both the primary and secondary master servers so that SSH communication is possible without a password, then retry 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 retry 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 log, 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 a cloning server definition file (clone.conf) does exist. If a 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 collecting cloning 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 collecting cloning images could not be removed.

#### Action method

Refer to the log ("`/var/opt/FJSVbdpp/log/bdpp_ROR.log`") and 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 cloning image could not be collected.

#### Action method

- The required communications (port) may have been shut by a firewall. Refer to "[A.3 Port List](#)", permit connections for ports used by Interstage Big Data Processing Server, then retry 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 is first in the boot order, then retry 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 retry the command.



.....

If there is an error in `clone.conf` definitions (MAC address, etc.) or cloning image collection fails due to a timeout, the network interface of the slave server that performed the cloning image collection may have changed to "DHCP" due to the cloning process, and not changed back.

If this is the case, log into the slave server and manually restore the settings of the network interface. Next, reset by doing "[4.6.1 Network Repliation and Hadoop Unsetup](#)" and "[4.2.3 Executing Network Repliation and Hadoop Setup](#)" again.

.....

---

### **bdpp: ERROR: 016: Failed to get the list of server image.**

#### Description

The cloning 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 cloning image could not be removed.

## Action method

Refer to the log ("/var/opt/FJSVbdpp/log/bdpp\_ROR.log") and 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 cloning image could not be deployed.

## Action method

- The required communications (port) may have been shut by a firewall. Refer to "[A.3 Port List](#)", permit connections for ports used by Interstage Big Data Processing Server, then retry 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 is the first in the boot order, then retry the command.
- If the message "FJSVnrm: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 cloning image is to be collected is not in the ipaddr.conf file.  
Include the definition for the slave server whose cloning image is to be collected in the ipaddr.conf file, redo "[4.9.1.1.2 Registering the Network Parameter Automatic Configuration Feature](#)", then recollect the cloning image and distribute it.

If none of the above apply, refer to the log ("/var/opt/FJSVbdpp/log/bdpp\_ROR.log") and the syslog, take the actions for messages with the prefix "FJSVrcx:" then retry the command.

---

## **bdpp: ERROR: 019: Failed to regist 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 "[A.3 Port List](#)", permit connections for ports used by Interstage Big Data Processing Server, then retry 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 retry the command.

If none of the above apply, refer to the log ("/var/opt/FJSVbdpp/log/bdpp\_ROR.log") and the syslog, take the actions for messages with the prefix "FJSVrcx:" then retry the command.

---

## **bdpp: ERROR: 020: Failed to change the cloning image directory.**

## Description

The cloning image storage directory could not be changed.

## Action method

Refer to the log ("/var/opt/FJSVbdpp/log/bdpp\_ROR.log") and 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.

---

## **bdpp: ERROR: 101: Administrator privilege is required.**

## Description

The user attempting execution does not have administrator privilege.

## Action method

Execute again with system administrator privilege.

## B.3.2 Other Messages

---

Refer to the manuals for the relevant functions 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	Function	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"
"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 Virtual Edition V3.0.0 Messages"

# Appendix C Mandatory Packages

This appendix explains the mandatory packages for the system software operated by this product.

## C.1 List of Mandatory Packages

The mandatory packages below are required in order 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.

[Master server]

Software name	Mandatory package	Notes
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)	kernel-devel	-
	compat-libstdc++-33(i686)	-
	ruby	-
	alsa-lib(i686)	-
	apr(i686)	-
	apr-util(i686)	-
	compat-expat1(i686)	-
	compat-libtermcap(i686)	-
	compat-openldap(i686)	-
	compat-readline5(i686)	-
	cyrus-sasl-lib(i686)	-
	db4(i686)	-
	expat(i686)	-
	glibc(i686)	-
	keyutils-libs(i686)	-
	krb5-libs(i686)	-
	libcom_err(i686)	-
	libgcc(i686)	-
	libICE(i686)	-
	libselinux(i686)	-
	libSM(i686)	-
	libstdc++(i686)	-
	libtool-ltdl(i686)	-
	libuuid(i686)	-
	libX11(i686)	-
	libXau(i686)	-
	libxcb(i686)	-
	libXext(i686)	-
	libXi(i686)	-
	libXt(i686)	-

Software name	Mandatory package	Notes
	libXtst(i686)	-
	ncurses-libs(i686)	-
	net-snmp	-
	net-snmp-utils	-
	nss-softokn-freebl(i686)	-
	openssl(i686)	-
	openssl098e(i686)	-
	redhat-lsb	-
	sqlite(i686)	-
	unixODBC(i686)	-
	zlib(i686)	-
	kexec-tools(i686)	-
	system-config-kdump	-
	crash(i686)	-
	sysstat(i686)	-
	ntp(i686)	-
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)	kernel-devel	-
	compat-libstdc++-33(i686)	-
	ruby	-
	apr(i686)	-
	apr-util(i686)	-
	libxml2(i686)	-
	libxslt(i686)	-
	net-snmp	-
	net-snmp-utils	-
	postgresql-libs(i686)	-
	redhat-lsb	-
	sqlite(i686)	-
	kexec-tools(i686)	-
	system-config-kdump	-
	crash(i686)	-
	sysstat(i686)	-
ntp(i686)	-	

**[Slave server]**

Software name	Mandatory package	Notes
Red Hat(R) Enterprise Linux(R) 6 (for Intel64)	kernel-devel	-
	alsa-lib(i686)	-
Red Hat(R) Enterprise Linux(R) 6.1 (for Intel64)	glibc(i686)	-
	libgcc(i686)	-

Software name	Mandatory package	Notes
Red Hat(R) Enterprise Linux(R) 6.2 (for Intel64)	libICE(i686)	-
	libSM(i686)	-
	libstdc++(i686)	-
	libtool-ltdl(i686)	-
	libuuid(i686)	-
	libX11(i686)	-
	libXau(i686)	-
	libxcb(i686)	-
	libXext(i686)	-
	libXi(i686)	-
	libXt(i686)	-
	libXtst(i686)	-
	ncurses-libs(i686)	-
	net-snmp-utils	-
	readline(i686)	-
	sqlite(i686)	-
	sysfsutils	-
	unixODBC(i686)	-
	kexec-tools(i686)	-
	system-config-kdump	-
	crash(i686)	-
sysstat(i686)	-	
ntp(i686)	-	
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)	kernel-devel	-
	alsa-lib(i686)	-
	glibc(i686)	-
	libgcc(i686)	-
	libICE(i686)	-
	libseltlinux(i686)	-
	libsepol	-
	libSM(i686)	-
	libstdc++(i686)	-
	libX11(i686)	-
	libXau(i686)	-
	libXdmcp	-
	libXext(i686)	-
	libXi(i686)	-
libXt(i686)	-	
libXtst(i686)	-	

Software name	Mandatory package	Notes
	ncurses-libs(i686)	-
	net-snmp-utils	-
	readline(i686)	-
	sqlite(i686)	-
	kexec-tools(i686)	-
	system-config-kdump	-
	crash(i686)	-
	sysstat(i686)	-
	ntp(i686)	-

**[Collaboration server]**

Software name	Mandatory package	Notes
Red Hat(R) Enterprise Linux(R) 6 (for Intel64)	kexec-tools(i686)	-
Red Hat(R) Enterprise Linux(R) 6.1 (for Intel64)	system-config-kdump	-
Red Hat(R) Enterprise Linux(R) 6.2 (for Intel64)	crash(i686)	-
Red Hat(R) Enterprise Linux(R) 5.6 (for Intel64)	sysstat(i686)	-
Red Hat(R) Enterprise Linux(R) 5.7 (for Intel64)	ntp(i686)	-
Red Hat(R) Enterprise Linux(R) 5.8 (for Intel64)		

# Appendix D DFS Environment Construction

This appendix describes the design and creation of a DFS and the construction of a Hadoop cluster that uses the DFS.

## D.1 File System Design

This section describes the items required for creating a DFS.

### D.1.1 Server Layout

Determine the servers that will share the DFS.

In a DFS, one file system can be shared by up to 128 servers at one time, including master servers, slave servers, development servers and collaboration servers.

The MDS that manages metadata runs on a master server. DFS client functions (AC) run on slave servers, development servers and collaboration servers.

### D.1.2 Disk Layout

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

Table D.1 Three partition areas

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

Follow the procedure below to determine the disk layout.

1. Estimate the shared disk device size needed for the DFS.

Table D.2 Shared disk device size estimation

Area type	Estimation method
(a) File data area	Space required for file data
(b) Metadata area	If file data area <= 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.
<b>Total</b>	<b>(a)+(b)</b>

2. Determine whether the file data area will be separate from the metadata area partition.

Separate the metadata area and file data area and allocate them to separate partitions. This separates the input-output processes to each area and can avoid contention.

Separation of the file data area can be specified using the `dataopt` option of the `pdfsmkfs` command.

3. Determine the partitions to be used in the DFS.

Determine the shared disk device partitions to be used in the DFS, based on the size requirement, and check the device name. A maximum of 256 partitions can be used.

### Point

A DFS can be constructed with only one partition, but the input-output distribution benefits can be obtained by constructing it with multiple partitions.

- Determine the metadata area partitions.

### Information

If the DFS is constructed with multiple partitions, representative partitions are used as the partitions for the metadata area.

## D.1.3 Preparing for Future Extension

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.

The maximum size can be specified using the `maxdsz` option of the `pdfsmkfs` command.

## D.1.4 Selecting the 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.

### Note

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.

### See

Refer to "[D.1.5 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.

## D.1.5 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 D.3 Relationship between file system size, data block size and maximum file size

File system size	Data block size	Maximum file size
to 1 TB	8 KB	1 TB - 8 KB
( 1 TB + 1 Byte) to 2 TB	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)

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

## D.2 Mount Option Design

This section describes the items that must be written to `/etc/fstab` to mount a DFS.

### D.2.1 fstab Settings for DFS Management Server

DFS entries must be added to `/etc/fstab` in order to perform mount at the DFS management server.

The DFS entries to be set in `/etc/fstab` are explained below.

- Field 1 (fs\_spec)  
Specify the representative partition of the DFS to be mounted.
- Field 2 (fs\_file)  
Specify the directory to be used as the default mount point.  
Specify a directory that exists at all servers where the DFS is to be mounted.
- 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.

Table D.4 Determining mount options

Item to be checked	Option to be specified
If either of the following applies:	<code>noatime</code>

Item to be checked	Option to be specified
- The file final reference time is not updated - The DFS is shared by five or more servers	
If not performing mount at DFS management server startup	noatrc
If mounting DFS as read only	ro



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.

An example of /etc/fstab settings is shown below.

Example: DFS management server fstab settings

```
# cat /etc/fstab <Enter>
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-hda5        swap             swap    defaults    0 0

/dev/disk/by-id/scsi-1FUJITSU_300000370106 /mnt/pdfs pdfs    noauto, noatime    0 0
```

## D.2.2 fstab Settings for DFS Clients

DFS entries must be added to /etc/fstab in order to perform mount at the slave servers, development servers and collaboration servers.

Determine the DFS entries to be set in /etc/fstab.

- Field 1 (fs\_spec)  
Specify the logical file system name of the DFS to be mounted.



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

- Field 2 (fs\_file)  
This is the directory to be used as the default mount point.  
Specify the directory that is specified at the master server.
- 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.

Table D.5 Determining mount options

Item to be checked	Option to be specified
If either of the following applies: <ul style="list-style-type: none"> <li>- The file final reference time is not updated</li> <li>- The DFS is shared by five or more machines</li> </ul>	noatime
If not performing mount at slave server, development server and collaboration server startup	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.

An example of /etc/fstab settings is shown below.

Example: slave server and collaboration server fstab settings

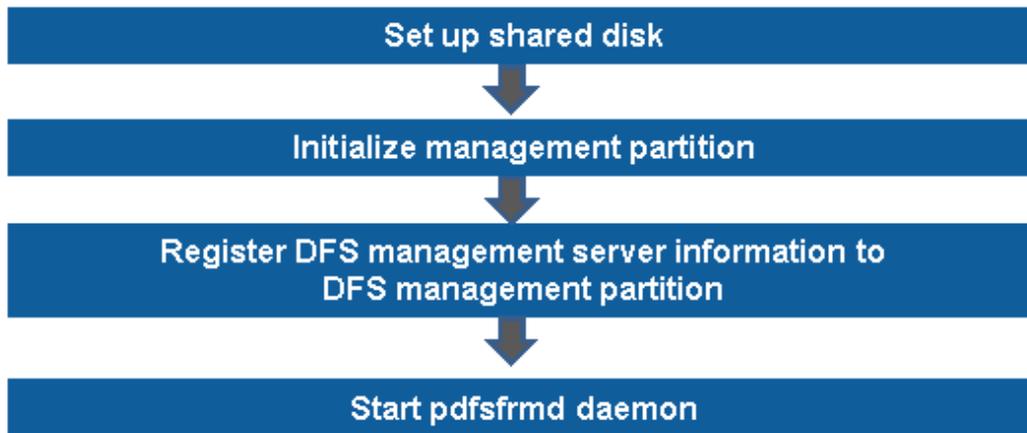
```
# cat /etc/fstab <Enter>
LABEL=/                /                ext3    defaults    1 1
LABEL=/boot            /boot           ext3    defaults    1 2
tmpfs                  /dev/shm       tmpfs   defaults    0 0
devpts                 /dev/pts       devpts  gid=5,mode=620 0 0
sysfs                  /sys           sysfs   defaults    0 0
proc                   /proc          proc    defaults    0 0
LABEL=SWAP-hda5       swap           swap    defaults    0 0

pdfs1                  /mnt/pdfsf     pdfs    _netdev     0 0
```

## D.3 Management Partition Creation

This section describes how to create a management partition.

Figure D.1 Workflow for creating a management partition



### Notes on creating a management partition

If there is master server redundancy, before creating a management partition, be sure to check that a cluster partition has not been generated. If a cluster partition has been generated, first cancel the cluster partition.



See

Refer to "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 the method for checking whether a cluster partition has been generated and the cluster partition cancellation method.

## D.3.1 Shared Disk Settings

The DFS management partition can be used from all master servers and is created on a shared disk at a shared device.

The by-id name generated by the udev function is used as the shared disk device name. The by-id name is a device name generated from the unique identification information set in the hard disk.

Use of the by-id name enables each server to always use the same device name to access a specific disk.

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

Name of compatible device where management partition is created:/dev/sdb

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

```
# udevinfo -q symlink -n /dev/sdb <Enter>
disk/by-id/scsi-1FUJITSU_300000370106 (omitted)
```

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

```
# udevadm info -q symlink -n /dev/sdb <Enter>
block/8:48 disk/by-id/ scsi-1FUJITSU_300000370106 (omitted)
```



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.



See

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



Note

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.



Point

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.

## D.3.2 Management Partition Initialization

Use the -c option of the pdfssetup command to initialize the management partition.

The example below shows how to initialize /dev/disk/by-id/scsi-1FUJITSU\_300000370105 as the management partition.

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



See

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.

## D.3.3 Registering DFS Management Server Information to the Management Partition

Use the -a option of the pdfssetup command to register master server information to the management partition.

The example below shows how to register master server information.

1. Register master server information to the management partition.

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



Note

Register the information at all master servers.

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_300000370106
```

### D.3.4 pdfsfrmd Daemon Startup

---

Start the `pdfsfrmd` daemon in order to start operations. Use the `pdfsfrmstart` command to start the `pdfsfrmd` daemon.



All master server information must be registered to the management partition before the `pdfsfrmd` daemon is started.

1. Execute the command shown below at all master servers.

```
# pdfsfrmstart <Enter>
```



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.

## D.4 File System Creation

---

This section describes the workflow for creating a DFS.

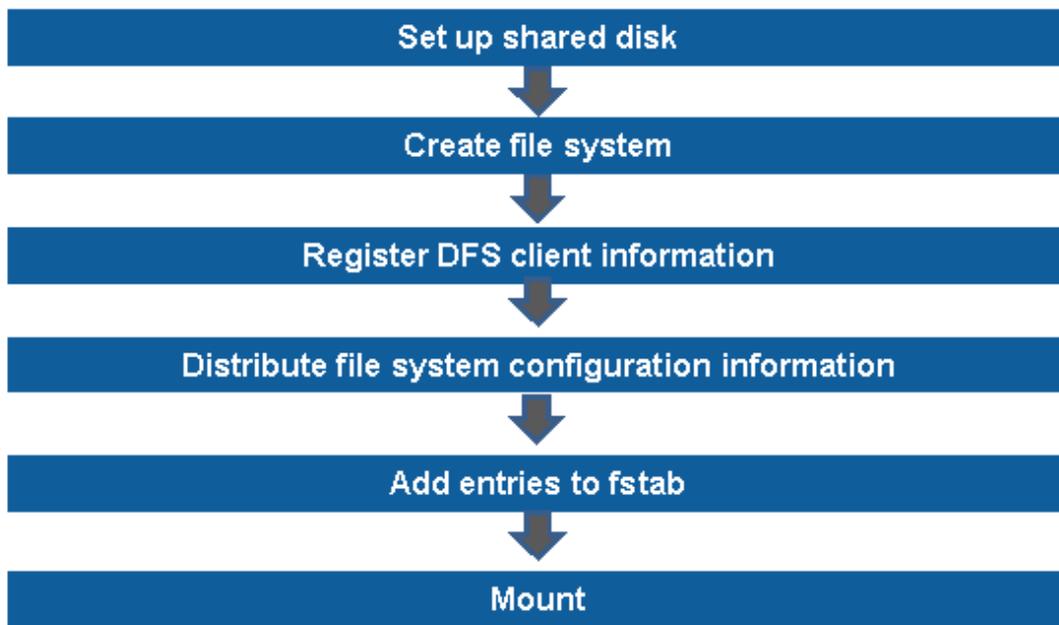
The management partition must be set and the file system must be designed before a file system is created.



Refer to "[D.3.2 Management Partition Initialization](#)" for the method for setting a management partition.

Refer to "[D.1 File System Design](#)" for the file system design method.

Figure D.2 Workflow for file system creation operations



## D.4.1 Shared Disk Settings

---

A DFS can be used from master servers, slave servers, development servers and collaboration servers and is created on the shared disk of a shared device.

The by-id name generated by the udev function is used as the shared disk device name. The by-id name is a device name generated from the unique identification information set in the hard disk.

Use of the by-id name enables each server to always use the same device name to access a specific disk.

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

---

Name of compatible device where file system is created: /dev/sdc

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

```
# udevinfo -q symlink -n /dev/sdc <Enter>
disk/by-id/scsi-1FUJITSU_300000370107 (omitted)
```

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

```
# udevadm info -q symlink -n /dev/sdb <Enter>
block/8:48 disk/by-id/ scsi-1FUJITSU_300000370107 (omitted)
```

---



### 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.

---



See

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



Note

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.



Point

The DFS file system 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.

## D.4.2 Creating a File System

Use the pdfsmkfs command to create a DFS. Execute the pdfsmkfs command at the master server.

The parameters to be specified in pdfsmkfs are the parameter values determined in advance as described at "D.1 File System Design".

The DFS creation example below is for the following environment:

- 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
- Master server (primary): master1
- Master server (secondary): master2

```
# 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>
```



Note

Specify the node option of the pdfsmkfs command in the final option.

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

## D.4.3 Registering Slave Server, Development Server and Collaboration Server Information

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

The management method involves creating a connection authorization list file at the master server and registering the host names of the slave servers, development server, and collaboration server to be connected.

The slave server, development server and collaboration server information registration procedure is described using the following environment as an example:

Representative partition: /dev/disk/by-id/scsi-1FUJITSU\_300000370106  
File system ID: 1  
Master server: master1, master2  
Slave server, development server and collaboration server: slave1, slave2, slave3, slave4, slave5, develop, collaborate

1. Check the file system ID.

Check the ID of the targeted file system from 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_300000370106 (864) 232256 DATA -----
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 connection authorization list file.

```
# 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. Add the slave server, development server and collaboration server information to the connection authorization list file.

Specify the host names of the slave servers, development servers and collaboration server that are authorized to connect to the master server.

Use the following format to enter the names:

```
CLIENT "slave server, development server, and collaboration server host names"
```

 **Example**

To authorize connection of slave1, slave2, slave3, slave4, and slave5:

```
# 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 slave1
CLIENT slave2
CLIENT slave3
CLIENT slave4
CLIENT slave5
```

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 connection authorization list file to each master server.

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

## D.4.4 Distributing File System Configuration Information

After the DFS configuration information file is generated at the master server, distribute it to each of the slave servers, development server and collaboration server.

The file system configuration information distribution procedure is described using the following environment as an example:

File system ID:	1
Logical file system name:	pdfs1
Slave server, development server, and collaboration server:	slave1, slave2, slave3, slave4, slave5, develop, collaborate

1. Check the file system ID.

Check the ID of the targeted file system from 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_300000370106 (864) 232256 DATA -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 7341778 DATA -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370108 (896) 6578704 DATA -----
```

2. Use the pdfsmkconf command to generate the DFS configuration information file.

Execute the pdfsmkconf command at the master server (primary).

```
# 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.pdfs1 <Enter>
```



In the configuration information file name, change only the file system ID part, not the other part (client.conf.).

4. Distribute the configuration information file to each of the slave servers, the development server, and the 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>
  . . . omitted . . .
# 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

Place the configuration information file under /etc/pdfs at each of the slave servers , development server and collaboration server.

### See

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.

## D.4.5 fstab Settings

---

Add the DFS description to /etc/fstab.

The /etc/fstab description content varies at the master server, slave servers, development server and collaboration server.

The parameters to be specified in each field of /etc/fstab are the parameter values determined in advance as described at "[D.2 Mount Option Design](#)".

Double-check the following points in the master server /etc/fstab:

- The device name is the by-id name of a representative partition.
- The mount options include noauto.

### Note

Make the /etc/fstab description content the same at all master servers.

Double-check the following points in the /etc/fstab of the slave servers, development server and collaboration server:

- The device name is a logical file system name.
- The mount options include \_netdev.

### Note

Make the /etc/fstab description content the same at all slave servers, development server and the collaboration server.

## D.4.6 Mount

---

Mount the DFS. Ensure that the DFS mount sequence is master server, then slave servers, development server and collaboration server. If the slave servers, development server or collaboration server mount is first, mount will fail because the master server (MDS) does not exist.

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

Representative partition:	/dev/disk/by-id/scsi-1FUJITSU_300000370106
Logical file system name:	pdfsl

## Note

- Create the mount point for mounting the DFS and perform the /etc/fstab description before performing the mount.
- Different mount commands are used for master servers than for slave servers, development server and collaboration servers.

1. Perform mount at the master server.

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

## Information

The pdfsmntgl command mounts the DFS at all DFS management servers.

2. Perform mount at the slave servers, development servers and collaboration servers.

Perform mount at all slave servers, development servers and collaboration servers that use the DFS.

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

## See

Refer to mount.pdfs under "Appendix A Command Reference" in the "Primesoft Distributed File System for Hadoop V1 User's Guide" for mount details for slave servers, development servers and collaboration servers.

## D.4.7 Checking the File System Status

Use the pdfsmntstat command to check that the DFS is mounted and in a usable state. Execute the pdfsmntstat command from the master server.

```
# pdfsmntstat <Enter>
MountPoint /mnt/fs1
Device /dev/disk/by-id/scsi-1FUJITSU_300000370106
SV MDS master1 29000 Active
SV MDS master2 29000 Standby
SV AC master1 Normal
SV AC master2 Normal
CL AC host1 Normal
CL AC host2 Normal
```

The target file system is operating normally if all the following states are confirmed:

- The status is "Active" at one of the SV MDS rows, indicating master servers.
- The status is "Normal" at all CL AC rows, indicating slave servers, development servers and collaboration servers.

## See

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

## D.4.8 Backing Up Management Partition Information

File system information that shows the configuration of each DFS is recorded in the management partition information and is used mainly during maintenance work.

Therefore, back up the management partition information after creating a DFS.

Use the pdfsgetconf command to make the backup.

## Point

If file system settings or changes are made the management partition information is changed. Therefore, back up the management partition information.

The following commands change the management partition information:

- pdfsmkfs, pdfsadd, pdfsadm, pdfsnode

1. Back up the management partition information.

```
# pdfsgetconf _backup_file_ <Enter>
```

## See

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

## Note

The pdfsfrmd daemon must be started at the DFS management server that executes the pdfsgetconf command.

2. Check the content of the backup.

The pdfsgetconf command generates a shell script having the name \_backup\_file\_.

The shell script content is shown below.

```
# cat _backup_file_ <Enter>
#!/bin/sh
# This file is made by:
#   pdfsgetconf _backup_file_
# Thu Jun 18 09:08:06 2009
#
#---- fsid : 1 ----
# MDS primary (port) : master1 (pdfs-1)
# MDS secondary (port) : master2 (pdfs-1)
# MDS other      :
# AC           : master1,master2
# options      :
# device       : /dev/disk/by-id/scsi-1FUJITSU_300000370106
pdfsadm -m master1,master2 -g master1,master2 -p pdfs-1,pdfs-1 /dev/disk/by-id/
scsi-1FUJITSU_300000370106
. . . omitted . . .
```

## Note

Store "\_backup\_file\_" very carefully because it is used for maintenance work.

## D.5 Hadoop Configuration Settings

This section describes the Hadoop configuration settings.

## D.5.1 Configuration File and Properties Settings

---

This section describes the configuration file settings that are to be set in order for the DFS to be used under Hadoop.

Configuration files include the types shown below. These files are placed in the "/etc/hadoop" directory.

### Configuration file types

- [hadoop-env.sh](#)
- [core-site.xml](#)
- [mapred-site.xml](#)
- [pdfs-site.xml](#)

This section describes the items to be set in each of these files.

### hadoop-env.sh file

Set the following environment variables in the hadoop-env.sh file:

```
export HADOOP_CLASSPATH="$HADOOP_CLASSPATH:/opt/FJSVpdfs/lib/pdfs.jar"
export HADOOP_USER_CLASSPATH_FIRST="true"
export HADOOP_SSH_OPTS="-o StrictHostKeyChecking=no -o BatchMode=yes"
```

If required, also set JAVA\_HOME.

```
export JAVA_HOME=/usr/java/default
```

### core-site.xml file

This section describes the properties to be set in the core-site.xml file.

#### fs.default.name

Specify the default file system in the "pdfs://<directory>" format (the <directory> part can be omitted).

The default file system is used to determine the URI path for path specifications that are not in URI format.

For example, if "pdfs:///" is set for fs.default.name, the "/mydir/myfile" is determined as being the "pdfs:///mydir/myfile" URI.

Default value

file:///

Settings example

```
<property>
  <name>fs.default.name</name>
  <value>pdfs:///</value>
</property>
```

#### fs.<scheme>.impl

Allocate the DFS file system class to any scheme.

Specify com.fujitsu.pdfs.fs.PdfsDirectFileSystem as the value.

For example, if set as the property named "fs.pdfs.impl", the path specification becomes the "pdfs:///mydir/myfile" URI.

Settings example

```
<property>
  <name>fs.pdfs.impl</name>
  <value>com.fujitsu.pdfs.fs.PdfsDirectFileSystem</value>
</property>
```

### io.file.buffer.size

Specify the default buffer size used during Read/Write.

Value to be specified

Multiple of 4096 (Bytes)

Default value

4096

Recommended value

128 KB

Settings example

```
<property>
  <name>io.file.buffer.size</name>
  <value>131072</value>
</property>
```

### hadoop.tmp.dir

Specify the directory used by Hadoop to store temporary files.

Default value

/tmp/hadoop-*<user name>*

Recommended value

Default

Settings example

```
<property>
  <name>hadoop.tmp.dir</name>
  <value>/var/tmp/hadoop-{user.name}</value>
</property>
```

## mapred-site.xml file

This section describes the properties to be set in the mapred-site.xml file.

### mapred.local.dir

Specify the directory that holds temporary data and Map intermediate output files while TaskTracker is executing MapReduce jobs. Specify a directory (or directories) in the local disk as the directory.

Default value

*\${hadoop.tmp.dir}/mapred/local*

Recommended value

*directory in in-built disk/mapred/local*



Do not include *{user.name}* in the specified path.

Settings example

Example: If mounted at the three in-built disks /data/1, /data/2, and /data/3:

```
<property>
  <name>mapred.local.dir</name>
  <value>/data/1/mapred/local,/data/2/mapred/local,/data/3/mapred/local</value>
</property>
```

#### mapred.system.dir

Specify the directory that stores the MapReduce processing control file.

##### Default value

`${hadoop.tmp.dir}/mapred/system`

##### Recommended value

`/mapred/system`

##### Settings example

```
<property>
  <name>mapred.system.dir</name>
  <value>/mapred/system</value>
</property>
```

#### mapreduce.jobtracker.staging.root.dir

Specify the top directory for the directories that store user-specific MapReduce job information files.

Job information files are stored in the "`${mapreduce.jobtracker.staging.root.dir}/<user name>/mapred/staging`" directory.

##### Default value

`${hadoop.tmp.dir}/mapred/staging`

##### Recommended value

Same as the `pdfs.fs.local.homedir` property (default: `/user`)(Refer to [pdfs.fs.local.homedir](#))

##### Settings example

```
<property>
  <name> mapreduce.jobtracker.staging.root.dir</name>
  <value>/user</value>
</property>
```

#### mapred.job.tracker

Specify, in *host:port* format (*port* cannot be omitted), the host name and port number of the RPC server that runs JobTracker. For the port number, specify an unused port number between 1024 and 61000.

##### Settings example

```
<property>
  <name>mapred.job.tracker</name>
  <value>host1:50001</value>
</property>
```

#### mapred.tasktracker.map.tasks.maximum

Specify the number of Map tasks executed in parallel at one node.

##### Default value

2

## Recommended value

Whichever is the larger of the following:

- Number of CPU cores - 1  
The number of CPU cores can be checked using "cat /proc/cpuinfo".
- Total number of physical disks comprising the DFS / Number of slave nodes in the Hadoop cluster (rounded up to whole number)  
If the disk device side has a RAID-1 mirror configuration, for example, the total number of physical disks is 2 \* the number of LUNs.

## Settings example

```
<property>
  <name>mapred.tasktracker.map.tasks.maximum</name>
  <value>8</value>
</property>
```

## mapred.tasktracker.reduce.tasks.maximum

Specify the number of Reduce tasks executed in parallel at one node.

### Default value

2

### Recommended value

Whichever is the larger of the following:

- Number of CPU cores - 1  
The number of CPU cores can be checked using "cat /proc/cpuinfo".
- Total number of LUNs comprising the DFS / Number of slave nodes in the Hadoop cluster (rounded up to whole number)

## Settings example

```
<property>
  <name>mapred.tasktracker.reduce.tasks.maximum</name>
  <value>4</value>
</property>
```

## mapred.map.tasks.speculative.execution

Specify whether to enable Map task speculative execution.

If "true" is specified, a task that is the same as the Map task currently being executed is processed in parallel at a node that has space for task execution, and whichever finishes first is used.

### Default value

true (enable)

If MapReduce jobs used in the Hadoop cluster include processing that write directly from Map tasks to DFS files, specify "false".  
If in doubt, specify "false".

## Settings example

```
<property>
  <name>mapred.map.tasks.speculative.execution</name>
  <value>>false</value>
</property>
```

## mapred.reduce.tasks.speculative.execution

Specify whether to enable Reduce task speculative execution.

If "true" is specified, a task that is the same as the Reduce task currently being executed is processed in parallel at a node that has space for task execution, and whichever finishes first is used.

#### Default value

true (enable)

If MapReduce jobs used in the Hadoop cluster include processing that writes directly from Reduce tasks to DFS files, specify "false". If in doubt, specify "false".

#### Settings example

```
<property>
  <name>mapred.reduce.tasks.speculative.execution</name>
  <value>>false</value>
</property>
```

#### mapred.userlog.limit.kb

Specify the maximum log size for one task in the MapReduce task log. The actual log size might be a little larger than the specified value due to the Hadoop control method.

The MapReduce task log is created under the \$HADOOP\_LOG\_DIR/userlogs directory.

#### Value to be specified

1 ( KB) or more

#### Default value

0 (unlimited)

#### Recommended value

Around 1000 (1 MB) (enough for about 10,000 rows)

This is large enough for the usual usage range, but adjust to suit the amount of logs output for MapReduce jobs.

#### Settings example

```
<property>
  <name>mapred.userlog.limit.kb</name>
  <value>1000</value>
</property>
```

#### mapred.userlog.retain.hours

Specify the retention period for MapReduce task logs. The task logs might be required for investigating the causes of MapReduce job errors.

MapReduce task logs are created under the \$HADOOP\_LOG\_DIR/userlogs directory.

#### Value to be specified

Retention time

#### Default value

24 (hours)

#### Recommended value

About 168 (enough for one week)

If the specified time is short, the required information might not be available when investigating past jobs. However, this needs to be balanced against the disk space (disk space available in the HADOOP\_LOG\_DIR directory) that can be used for retaining logs. Set a suitable time that takes the following into account:

- Space that can be permitted
- Operation work schedule

- `mapred.userlog.limit.kb` property value (refer to "[mapred.userlog.limit.kb](#)")

If, for example, the maximum log output for one minute is assumed to be 1 MB (about 10,000 rows), a space of about 10GB would be required to retain a week's worth of logs.

#### Settings example

```
<property>
  <name>mapred.userlog.retain.hours</name>
  <value>168</value>
</property>
```

### pdfs-site.xml file

If required, set the properties below.

Normally, set the `pdfs.fs.local.basedir` property (refer to "[pdfs.fs.local.basedir](#)") and the `pdfs.security.authorization` property (refer to "[pdfs.security.authorization](#)"), and the defaults can be used for the other properties.

#### pdfs.fs.local.basedir

Specify the DFS mount directory path.

If `"/mnt/pdfs/hadoop"` is set for `pdfs.fs.local.basedir`, for example, the URI `"pdfs:///user/bdppuser1"` becomes `"/mnt/pdfs/hadoop/user/bdppuser1"` path at the operating system.

#### Default value

/

#### Settings example

```
<property>
  <name>pdfs.fs.local.basedir</name>
  <value>/mnt/pdfs/hadoop</value>
</property>
```

#### pdfs.fs.local.homedir

Specify the home directory path for users in the DFS FileSystem class.

If `"/user"` is set for `pdfs.fs.local.homedir`, for example, the DFS home directory URI path for the user with the name `"bdppuser1"` becomes `"pdfs:///user/bdppuser1"`.

#### Default value

/user (same as HDFS)

#### Recommended value

Default (/user)

#### Settings example

```
<property>
  <name>pdfs.fs.local.homedir</name>
  <value>/home</value>
</property>
```

#### pdfs.security.authorization

Specify whether or not to use the DFS's own MapReduce job user authentication.

#### Value to be specified

- true: Use

- false: Do not use

Default value

false

Recommended value

true

To use Kerberos authentication under Hadoop, specify "false".

Settings example

```
<property>
  <name>pdfs.security.authorization</name>
  <value>true</value>
</property>
```

pdfs.fs.local.buffer.size

Specify the default buffer size to be used during Read/Write.

Note that either this property value or the io.file.buffer.size value, whichever is larger, is used (refer to "[io.file.buffer.size](#)").

Value to be specified

Multiple of 4096 (Bytes)

Default value

128 KB

Recommended value

128 KB to 512 KB

Settings example

```
<property>
  <name>pdfs.fs.local.buffer.size</name>
  <value>524288</value>
</property>
```

pdfs.fs.local.block.size

Specify the data size into which Map tasks are split for MapReduce jobs.

As a guide, specify <total input data size of main MapReduce jobs / Number of slave nodes> or less.

Note that this specification need not match the block size (blocksz option) specified for pdfsmkfs at the time of "[D.4.2 Creating a File System](#)".

Value to be specified

Multiple of 33554432 (32 MB)

Default value

256 MB

Maximum value

1GB (1073741824)

Recommended value

256 MB to 1GB

Settings example

```
<property>
  <name>pdfs.fs.local.block.size</name>
```

```
<value>1073741824</value>
</property>
```

#### pdfs.fs.local.posix.umask

Specify whether or not the process umask value is reflected to the access permissions set during file or directory creation.

##### Value to be specified

- true: Use umask value (POSIX compatible)
- false: Do not use umask value (HDFS compatible)

##### Default value

true

##### Settings example

```
<property>
  <name>pdfs.fs.local.posix.umask</name>
  <value>>false</value>
</property>
```

#### pdfs.fs.local.cache.location

Specify whether or not to use the cache local MapReduce function.

When enabled, this function fetches 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, thus speeding up Map phase processing.

##### Value to be specified

- true: Use the cache local MapReduce function
- false: Do not use the cache local MapReduce function

##### Default value

true

##### Settings example

```
<property>
  <name>pdfs.fs.local.cache.location</name>
  <value>>false</value>
</property>
```

#### pdfs.fs.local.cache.minsize

Specify the file size that is excluded from the cache local MapReduce function targets. Since there is a cost associated with fetching memory cache retention node information, it is possible to set for this information to not be fetched if the file is less than the specified size.

##### Value to be specified

1 (Byte) or more

##### Default value

1048576 (1 MB)

##### Settings example

```
<property>
  <name>pdfs.fs.local.cache.minsize</name>
  <value>1048576</value>
</property>
```

## pdfs.fs.local.cache.shell

Specify the remote command execution parameter used when the cache local MapReduce function fetches memory cache information.

### Default value

```
/usr/bin/ssh -o IdentityFile=%HOME/.pdfs/id_hadoop -o ConnectTimeout=10 -o BatchMode=yes -o StrictHostKeyChecking=no  
-o EscapeChar=none
```

Note that "%HOME" in the character string is replaced by "\${pdfs.fs.local.basedir}/\${pdfs.fs.local.homedir}/<user name>", and "%USER" is replaced by the user name.

### Settings example

```
<property>  
  <name>pdfs.fs.local.cache.shell</name>  
  <value> /usr/bin/ssh -o IdentityFile=/home/%USER/.ssh/id_rsa -o ConnectTimeout=10 -o  
BatchMode=yes -o StrictHostKeyChecking=no -o EscapeChar=none</value>  
</property>
```

## pdfs.fs.local.cache.procs

Specify the number of multiplex executions when the cache local MapReduce function fetches memory cache information.

### Default value

10

### Settings example

```
<property>  
  <name>pdfs.fs.local.cache.procs</name>  
  <value>40</value>  
</property>
```

## D.6 User-Specific Hadoop Usage Settings

---

This section describes the settings required for each user account that uses Hadoop.

Set the settings below for each user account that uses Hadoop.

### D.6.1 Settings Implemented at root

---

Create user accounts in advance at all nodes comprising the DFS. At that time, ensure the correspondence between user names and user IDs is the same at all nodes. This is to achieve consistency at all nodes because DFS uses the user IDs for control and Hadoop operates on the basis of user names.

Create a user home directory ( $\{\text{pdfs.fs.local.homedir}\}/\text{<user name>}$ ) under the DFS FileSystem class.



#### Example

---

If the user is called bdppuser1:

Since the settings are shared at all nodes, execute as shown below at any of the nodes where DFS is mounted.

Execute using root permission.

```
$ su  
# hadoop fs -mkdir /user/bdppuser1  
# hadoop fs -chown bdppuser1:bdppgroup /user/bdppuser1  
# hadoop fs -chmod 0700 /user/bdppuser1  
# exit
```

Set Group and Permission to suit the environment being used.

## D.6.2 Settings Implemented by Hadoop Users

### SSH authentication key settings to be used by the cache local function

If MapReduce jobs that use the cache local MapReduce function are executed, DFS fetches file cache information from all Hadoop cluster nodes.

Since 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 function 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. Since the settings are shared at all nodes, `bdppuser1` executes as shown below at any of the nodes where DFS is mounted.

```
$ cd ~
$ ssh-keygen -t rsa -N "" -f id_hadoop
$ echo -n "command=\"/opt/FJSVpdfs/sbin/pdfscachelocal.sh\",no-pty,no-port-forwarding,no-X11-
forwarding " > authorized_keys
$ cat id_hadoop.pub >> authorized_keys
$ hadoop fs -mkdir .pdfs
$ hadoop fs -chmod 700 .pdfs
$ hadoop fs -moveFromLocal id_hadoop id_hadoop.pub authorized_keys .pdfs/
$ hadoop fs -chmod 600 .pdfs/id_hadoop .pdfs/authorized_keys
$ hadoop fs -chmod 644 .pdfs/id_hadoop.pub
```

An entry like the one below is set in the `authorized_keys` file shown above.

```
$ hadoop fs -cat .pdfs/authorized_keys
"command=/opt/pdfs/sbin/pdfscachelocal.sh",no-pty,no-port-forwarding,no-X11-forwarding ssh-rsa
public key bdppuser1@local node name
```

2. Reflect the settings to `.ssh/authorized_keys` of the user `bdppuser1` home directory at all slave nodes 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
$ ssh develop "umask 0077; mkdir -p .ssh; cat /mnt/pdfs/hadoop/user/bdppuser1/.pdfs/
authorized_keys >> .ssh/authorized_keys"
```

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 remote node name /opt/FJSVpdfs/sbin/pdfscachelocal.sh
2 (*1)
```

(\*1) Normally "2" is output.

### MapReduce job user authentication key settings

The tasks here are not required if "[SSH authentication key settings to be used by the cache local function](#)" were implemented.

Since authentication key checking is performed during the MapReduce job user authentication of DFS, an authentication key file must be created in advance. Use the `pdfs.security.authorization` property in the `pdfs-site.xml` file 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 "[pdfs-site.xml file](#)" for information on the pdfs.security.authorization property.



## Example

---

If the user is bdppuser1 and the DFS mount directory is /mnt/pdfs:

Since the settings are shared at all nodes, bdppuser1 executes as shown below at any of the nodes where DFS is mounted.

```
$ cd ~
$ cat > id_hadoop
any keyword character string
ctrl-d
$ hadoop fs -mkdir .pdfs
$ hadoop fs -chmod 700 .pdfs
$ hadoop fs -moveFromLocal id_hadoop .pdfs/
$ hadoop fs -chmod 600 .pdfs/id_hadoop
$
```

---

# Appendix E DFS Operation Changes

This appendix describes changes to the operation of a file system using a DFS.

## E.1 Deleting a File System

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:

- Representative partition: /dev/disk/by-id/scsi-1FUJITSU\_300000370107
- File system ID: 2
- Logical file system name: pdfs2

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

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

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

```
# umount pdfs2 <Enter>
```

Unmount the DFS at the master server.

```
# pdfsumntgl /dev/disk/by-id/scsi-1FUJITSU_300000370107 <Enter>
```

2. 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) 14422 META -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 5116 LOG -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 95112 DATA -----
/dev/disk/by-id/scsi-1FUJITSU_300000370107:
FSID special size Type mount
 2 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 14422 META -----
 2 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 5116 LOG -----
 2 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 95112 DATA -----
```

3. Delete the DFS.

```
# pdfsadm -D /dev/disk/by-id/scsi-1FUJITSU_300000370107 <Enter>
```

4. Check that the DFS was deleted.

```
# pdfsinfo -a <Enter>
/dev/disk/by-id/scsi-1FUJITSU_300000370106:
FSID special size Type mount
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 14422 META -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 5116 LOG -----
 1 /dev/disk/by-id/scsi-1FUJITSU_300000370106 (864) 95112 DATA -----
```

5. Back up the management partition information.

Store `_backup_file_` very carefully because it is used for maintenance work.

```
# pdfsgetconf _backup_file_ <Enter>
```

6. On all slave servers, development servers and collaboration servers, delete the file system configuration information file.

```
# cd /etc/pdfs <Enter>
# rm ./client.conf.pdfs2 <Enter>
```

7. On the master server (primary) and the master server (secondary), delete the connection authorization list file.

```
# cd /etc/pdfs <Enter>
# rm ./server.conf.2 <Enter>
```

8. Delete the targeted DFS description from the `/etc/fstab` of the master server(primary), the master server (secondary), slave servers, development servers and collaboration servers.

### Information

The DFS device names entered in `/etc/fstab` are as follows:

- Master server (primary), master server (secondary): Representative partition by-id name
- Slave server, development server, collaboration server: Logical file system name

## E.2 Re-creating a 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 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 "[E.1 Deleting a File System](#)" for the DFS deletion procedure.
- Refer to "[D.4 File System Creation](#)" for the DFS creation procedure.

### Note

When a file system is re-created, all old file system data is deleted. Therefore, back up 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 procedures at the master server.

1. 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.

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

2. Fetch the row of the pdfsmkfs command that re-creates a file system.

```
# pdfsmkfs -m /dev/disk/by-id/scsi-1FUJITSU_300000370106 > _mkfs_param_ <Enter>
```

3. Add specification of the force option, used specifically for re-creation, to the pdfsmkfs command row.

```
# cat _mkfs_param_ <Enter>
pdfsmkfs -o force,[omitted] /dev/disk/by-id/scsi-1FUJITSU_300000370106
```

4. Re-create the DFS.

```
# sh _mkfs_param_ <Enter>
```

5. Delete the pdfsmkfs command row.

```
# rm _mkfs_param_ <Enter>
```

6. Mount the DFS and restart operation.

Mount the DFS, starting with the master server.

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

Mount the DFS at all slave servers, development servers and collaboration servers that use the DFS.

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

## E.3 Adding a Slave Server/Scaling Out

Under a DFS, the server processing performance can be improved by adding slave servers to scale out the system.

The slave servers, development servers and collaboration servers can be added without stopping operation.



### Note

After deletion of the slave servers, development servers and collaboration servers, servers cannot be added while operation continues.

In this case, unmount the DFS and then add the servers.

The slave server, development server and collaboration server addition procedure is described below using the following environment as an example:

- |  |  |
|--|--|
| - Logical file system name:  | pdfs1  |
| - Mount point:   | /mnt/pdfs  |
| - Representative partition:  | /dev/disk/by-id/scsi-1FUJITSU_300000370106                   |
| - File system ID:  | 1  |
| - Master server:   | master1, master2   |
| - Existing slave server, development server, collaboration server: | slave1, slave2, slave3, slave4, slave5, develop, collaborate |
| - Additional slave server:   | slave6, slave7, slave8, slave9, slave10                      |

## E.3.1 Adding Slave Server, Development Server and Collaboration Server Information

When adding slave servers, development servers and collaboration servers, add to the master server connection authorization list file the host names of the slave servers, development servers or collaboration servers being added.

The following tasks are required at the master server (primary):

1. From the FSID field in the pdfsinfo command output, check the file system ID of the file system to which the slave servers, development servers and collaboration servers are being 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. In the connection authorization list file applicable to the file system ID, add a CLIENT field for the slave servers, development servers and collaboration servers being added

Add the CLIENT field at the end of the connection authorization list file.

If multiple slave servers, development servers or collaboration servers are being added, add CLIENT fields for just that number of clients.

```
# 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
CLIENT slave4
CLIENT slave5
CLIENT develop
CLIENT collaborate
```

### 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.

1. Distribute the updated connection authorization list file to the master server (secondary).

If there are redundant master servers, distribute to the other master servers.

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

### Note

If the connection authorization list file contents differ at different master servers, MDS-down recovery might fail and file system inhibition or system panic might occur.

## E.3.2 Distributing File System Configuration Information

Generate the DFS file system configuration information, and distribute it to the slave servers, development servers and collaboration servers being added.

1. Generate the DFS configuration information file.

Execute the pdfsmkconf command at the master server (primary).

```
# pdfsmkconf <Enter>
```

### See

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

1. Convert the file name of the configuration information file applicable to the file system ID to a logical file system name.

```
# cd pdfsmkconf_out <Enter>
# mv ./client.conf.1 client.conf.pdfs1 <Enter>
```

### Note

In the file name of the configuration information file, do not change anything other than the file system ID part (client.conf).

2. Distribute the configuration information file to the slave servers, development servers and collaboration servers being added.

```
# scp -p ./client.conf.pdfs1 root@rainy3:/etc/pdfs/client.conf.pdfs1 <Enter>
```

### Note

Place the configuration information file under /etc/pdfs at the slave servers, development servers and collaboration servers.

## E.3.3 fstab Settings

Entries related to the file system to be mounted must be defined in the /etc/fstab of the slave servers, development servers and collaboration servers being added.

Also, entries related to the same file system must be made the same as for the other slave servers, development servers and collaboration servers that share the file system.



## Example

```
# cat /etc/fstab | grep pdfs <Enter>
pdfs1                /mnt/pdfs           pdfs    _netdev    0 0
```

### E.3.4 Mount

Mount the file system at the slave servers, development servers and collaboration servers being added.

Use the mount command to mount the file system at the slave servers, development servers and collaboration servers.

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

### E.4 Deleting a Slave Server

Under DFS, if there is a hardware error or similar reason, slave servers, development servers and collaboration servers can be deleted without stopping operation.

The slave server, development servers and collaboration server deletion procedure is described below using the following environment as an example:

- Logical file system name: pdfs1
- Mount point: /mnt/pdfs
- File system ID: 1
- Master server: master1, master2
- Existing Slave server, development server, collaboration server: slave1, slave2, slave3, slave4, slave5, develop, collaborate
- Deletion target slave server: slave6

1. If the DFS is mounted at the slave servers, development servers and collaboration servers targeted for deletion, unmount the DFS.

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

2. Delete the file system configuration information file at the slave server, development servers or collaboration server targeted for deletion.

```
# cd /etc/pdfs <Enter>
# rm ./client.conf.pdfs1 <Enter>
```

3. Delete the DFS description from the /etc/fstab of the slave server, development server or collaboration server targeted for deletion.
4. From the master server (primary) connection authorization list file, delete the CLIENT fields of the slave server, development server or collaboration server targeted for deletion.

```
# cat /etc/sfcfs/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 slave1
CLIENT slave2
CLIENT slave3
CLIENT slave4
CLIENT slave5
CLIENT develop
CLIENT collaborate
CLIENT slave6      <-- Delete the row here.
CLIENT slave7
CLIENT slave8
CLIENT slave9
CLIENT slave10
```

5. Distribute the updated connection authorization list file to the master servers.

Example: If executing from the master server (primary) master1:

```
# cd /etc/pdfs <Enter>
# scp -p ./server.conf.pdfs1 root@master2:/etc/pdfs/server.conf.pdfs1 <Enter>
```



### Note

If the connection authorization list file contents differ at different master servers, MDS down recovery might fail and file system inhibition or system panic might occur.

## E.5 Adding a Shared Disk/Scaling Out

Under DFS, adding partitions to scale-out the system can provide either or both improvement in processing performance, and increased file system size of a shared disk device.

When adding partitions, unmount the targeted DFS.

The procedure for adding partitions to the DFS is described below using the following environment as an example:

- Logical file system name: pdfs1
- Mount point: /mnt/pdfs
- Representative partition: /dev/disk/by-id/scsi-1FUJITSU\_300000370106
- File system ID: 1
- Master server: master1,master2
- Slave server, development server, collaboration server: slave1, slave2, slave3, slave4, slave5, develop, collaborate
- Existing partition: /dev/disk/by-id/scsi-1FUJITSU\_300000370107
- Additional partition: /dev/disk/by-id/scsi-1FUJITSU\_300000370108

1. 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>
```

2. 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_300000370106 (864) 232256 DATA -----
  1 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 7341778 DATA -----
```

3. Add the partition.

Add /dev/disk/by-id/scsi-1FUJITSU\_300000370108 as a file data area.

```
# pdfsadd -D /dev/disk/by-id/scsi-1FUJITSU_300000370108 /dev/disk/by-id/
scsi-1FUJITSU_300000370106 <Enter>
```

4. 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_300000370106 (864) 232256 DATA -----
  1 /dev/disk/by-id/scsi-1FUJITSU_300000370107 (880) 7341778 DATA -----
  1 /dev/disk/by-id/scsi-1FUJITSU_300000370108 (896) 6578704 DATA -----
```

5. Regenerate the DFS configuration information file.

Execute the pdfsmkconf command at the master server.

```
# pdfsmkconf <Enter>
```

### Information

A configuration information file is generated, as pdfsmkconf\_out/client.conf.fsid under the directory where the pdfsmkconf command is executed, for each file system ID, fsid.

6. Convert the 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>
```

7. Distribute the configuration information file to each of the slave servers, development servers and collaboration servers.

```
# scp -p ./client.conf.pdfs1 root@slave1:/etc/pdfs/client.conf.pdfs1 <Enter>
# scp -p ./client.conf.pdfs1 root@slave2:/etc/pdfs/client.conf.pdfs1 <Enter>
```

### Note

Place the configuration information file under /etc/pdfs at each of the slave servers, development servers and collaboration servers.

8. Mount the DFS and restart operation.

Mount the DFS, starting from the master server.

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

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

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

## E.6 Deleting a Shared Disk

---

Under DFS, the shared disks that comprise a file system cannot be deleted.

Use the procedure below to reconfigure a DFS when deleting shared disks.

1. Back up resources that will still be required after reconfiguration.

Perform backup in file/directory units.

### Note

Since shared disks are deleted, ensure 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".

2. Delete the targeted DFS.

### See

Refer to "E.1 Deleting a File System" for the DFS deletion procedure.

3. Re-create the DFS using the changed configuration.

### See

Refer to "D.4 File System Creation" for the DFS creation procedure.

4. If resources were backed up in Step 1, restore the resources.

# Glossary

---

## Apache Hadoop

Open source Hadoop software developed by Apache Software Foundation (ASF).

## Cloning master

A master obtained by removing server-specific information (system node names and IP addresses) from the contents of a system disk. The cloning master is copied to a virtual server system disk when a virtual server is deployed.

## DataNode

Name of the nodes that comprise a cluster in the HDFS file system used by Hadoop. The Big Data processed by Hadoop is distributed and replicated in block units and mapped on DataNodes.

## 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 MapReduce parallel distributed processing technology.

## HDFS (Hadoop Distributed File System)

The distributed file system used in Hadoop. HDFS maps the 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.

## MapReduce

The distributed parallel 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.

## 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.

## 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.

## 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 (text, voice, videos, etc.) via the Internet are called social media to differentiate these from older information media (newspapers, television, etc.).

---

## Terabyte (TB)

A unit of data size, indicating  $10^{12}$  bytes.